Optimising of vacuum distillation units using surrogate models

S X H’ng, L Y Ng, D K S Ng and V Andiappan
School of Engineering and Physical Sciences, Heriot-Watt University Malaysia, 62200, Putrajaya, Wilayah Persekutuan Putrajaya, Malaysia

Email: v.murugappan@hw.ac.uk

Abstract. Crude oil blending is an important step for the operation of crude distillation systems in the refinery to improve the yield and profitability of the products. The product’s yield and quality are strongly dependent on the properties of the crude oil. However, the products of crude distillation units, especially the vacuum distillation unit (VDU) need to satisfy the yield and quality requirements of the downstream process units in the refinery. Otherwise, the performance of downstream processes will be affected, and loss of profitability in the refinery. Hence, it is important to optimise the performance of the VDU to ensure the optimum operation of VDU. This work covers the process simulation of VDU, surrogate modelling and mathematical optimisation model. The objective of the developed optimisation model is to determine an optimal for maximum high vacuum gas oil (HVGO) yield and minimum total annualised cost (TAC) respectively. To do this, crude oil blending ratio, column temperature, column pressure, stripping steam flowrate, pump-around flowrate in the VDU are optimised. Based on the optimised result, the heavy-light crude blend achieves higher HVGO yield and lower TAC as compared to the heavy-medium crude blend and heavy-medium-light crude blend. The optimised results can provide insight into the optimal process conditions of VDU for the refiners. With this insight, effective operating strategies can be developed to overcome the limitations present in real VDU operations.

1. Introduction
Vacuum distillation unit (VDU) is part of a typical refinery process that helps to further recover the higher boiling gas oil from the atmospheric residue of an atmospheric distillation unit. The major products of the VDU include light vacuum gas oil and heavy vacuum gas oil (HVGO). In the VDU, the separation of atmospheric residue into products involves many complicated relationships between operating conditions. The operating conditions of VDU need to be adjusted from time to time due to this variation of crude oil properties. These adjustments may consequently impact the product yield and quality. Nevertheless, it is critical for the products of the VDU to meet the targeted yield and quality specifications. This is because the products of VDU are the feedstocks to the downstream process units such as hydrocracker, fluid catalytic cracker or lube oil facility to produce commercially valuable products [1]. If the product yield or product quality specifications cannot be met, the downstream process units of the refinery will suffer. It may also cause an unplanned shutdown of the units and eventually incur a financial loss. Hence, all these sequences raise the need to develop a systematic optimisation method for determining the blending ratio of different kinds of crude oil and the optimal
operating conditions of the VDU. By doing so, VDU can achieve high production throughput, satisfactory product quality and excellent economic potential.

Mathematical programming has been widely used in the synthesis, design and optimisation of crude distillation systems. The implementation of the mathematical optimisation technique helps to determine the optimal solutions of the system based on the problems of the mathematical descriptions [2]. For instance, Seo et al. [3] implemented mixed integer nonlinear programming to determine the optimal feed location and operating conditions of a crude distillation unit. With that, the proposed framework reduced the operating and capital costs of the existing crude distillation unit by 86 %. Next, Basak et al. [4] developed a non-linear, steady-state crude distillation unit to maximise the net profit while satisfying the product properties. A noticeable increase in profitability was observed with this optimisation model. Furthermore, Gu et al. [5] combined exergy analysis and mathematical programming for the energy optimisation of the multistage crude oil distillation units. The results showed a considerable reduction of 2.79 % in energy consumption which proved that this approach was implemented effectively. Besides, Inamdar et al. [6] introduced elitist non-dominated sorting genetic algorithms to address the multi-objective optimisation of the crude distillation systems. From this work, better optimal conditions and higher profit of the crude distillation system were reported. Similarly, Xiaoqiao et al. [7] also employed a multi-objective optimisation model to investigate the trade-off between economic benefit, furnace energy consumptions and CO₂ emissions of the crude distillation unit. Meanwhile, the effect of binary feed composition impact on the crude distillation unit had investigated as well. The proposed model provided a set of Pareto-optimal solutions that can solve the limitations of the refinery. Based on the literature cited above, a lot of researchers have considered the mathematical programming techniques in the optimisation work of crude distillation unit. Nevertheless, the complexity and nonlinearities associated with the distillation units may restrict the use of the mathematical optimisation model.

This has brought the attention of the researchers to employ the surrogate modelling technique into the distillation optimisation process, particularly for the crude distillation units in the refinery. A surrogate model is a sophisticated analytic model that mimics the complex behaviour of the simulation environment based on the regression of statistical input variables and output responses [8,9]. For example, López et al. [10] implemented surrogate models in the non-linear programming model to investigate the effect of crude composition and operating conditions of the crude distillation unit. An increase of 13 % in profit was observed from the proposed optimisation model. Besides, Yao and Chu [11] developed a surrogate model with the support vector regression and improved design of experiment method to optimise the operating conditions of the atmospheric distillation unit. A remarkable increase in profitability was also reported in the proposed approach. Apart from that, Ochoa-estopier et al. [12] introduced an artificial neural network model to represent the distillation columns. The proposed model associated with the heat exchanger networks was incorporated in the optimisation framework to determine the optimal operating conditions that give maximum process economy. Moreover, a cut-point temperature surrogate modelling technique was proposed by Gut et al. [13] to determine the product yield and properties of the crude distillation units. The surrogate models were proven to have accurate estimations and the models were able to be implemented into the planning, control and scheduling applications of the crude oil distillation units. Based on the literature surveys above, most of the optimisation works have been devoted to the overall crude distillation systems and atmospheric distillation units, rather less attention has been paid to the VDUs.

The performance of the VDU is significant as it needs to satisfy the feedstocks quality and rate target of the downstream process units. Therefore, this raised the focus of some researchers on the optimisation work of the VDU. With the aid of Aspen HYSYS simulation, Mittal et al. [14] developed an optimisation framework for crude oil blending and processing with the simultaneous considerations of the furnace energy consumptions, CO₂ emissions and economic potential of the VDU. Next, Gu et al. [15] proposed a methodology to analyse and evaluate three crude oil vacuum distillation processes. Recovery energy and exergy efficiencies, economical potential and product yield were also taken into considerations in the proposed methodology. Consequently, the results provided insight for engineers to determine a suitable process and outlet temperature of the VDU. Also, it is worth noting that Li et al. [16] introduced
a dividing wall column in a lubricant type VDU to enhance the product yield and quality. The proposed configuration showed that the boiling point range of the lube cut was reduced significantly which improved the product yield and quality.

A critical observation of the literature review presented above infers that the relationship between the product yield and total annualised cost (TAC) of VDU has not been studied for crude oil blending. The refiners have been looking for the opportunity in processing lower-cost heavy crude oil due to the operational and feed availability constraints. By blending the crude oil, the diversity of crude oil processing as feedstock has increased which maximise the profitability of refiners. Therefore, crude oil blending is taken into consideration in the optimisation work of the VDU. Meanwhile, another observation that can be drawn from the literature cited above is that no work has combined the process simulation, surrogate modelling and mathematical programming in the distillation system optimisation works. Process simulation is used to simulate and provide basic data of the distillation process. Nevertheless, a lot of challenges occur in the optimisation process using process simulation due to the high degree of freedom and nonlinearities associated with the distillation units. The process simulation tool does not reflect the optimal solution to run a system. It can only give the best possible solution based on a set of predefined scenarios determined by the decision-makers. To address the issues of process simulation, surrogate modelling and mathematical programming are introduced. Surrogate modelling technique is applied to mimic the behaviour of the distillation system due to their simplicity and satisfactory accuracy. Surrogate modelling techniques are employed to capture the relationship between the independent X variables and dependent Y variables of the distillation system. The independent X variables are the variables that can be controlled (i.e., crude oil blending ratio and operational variables), whereas dependent Y variables are the variables that are measured and dependent on the independent variables (i.e, HVGO yield, pump-around cooler flowrate, column diameter, etc.). The surrogate model is constructed by regressing against a set of statistical data. The data is later converted to a fairly representative equation that can describe the operation of the distillation system. Also, mathematical programming can be used to analyse and provide the optimal solution of a given distillation system. Therefore, the combination of the individual strengths of these three methods can reveal the results that are not possible to be determined when these three methods are used separately. The combination of these methods is essential for decreasing the overall process costs, predicting the process behaviour, improving the utilisation of resources and providing detailed information on how the system operates. Hence, this work considers the combination of both process simulation, surrogate modelling and mathematical programming in the optimisation work of the VDU.

In order to fulfil the research gaps discussed above, this paper develops an optimisation model with the combination of process simulation, surrogate modelling and mathematical programming. The proposed model is used to investigate the impact of crude blending composition and the operating conditions of the VDU on the HVGO yield. Besides, upon considering the economic profit of refineries, the total annualised cost (TAC) of VDU is investigated as well. This study can provide economical and operational benefits to the refiners as it can overcome the tighter supply-demand constraints in the refinery. The objectives of this research are:

- To determine the optimal VDU operating performance and conditions that give maximum HVGO.
- To determine the optimal VDU operating performance and conditions that give minimum TAC.

The remainder of this paper is organised as follows. Section 2 presents the methodology for the optimisation of VDU, based on the process simulation, surrogate modelling and mathematical programming. Section 3 illustrates the application of the proposed methodology in a case study. In the case study, different crude blending ratios are compared and analysed to identify the optimal crude blending ratio with the corresponding operating conditions that give maximum HVGO yield and minimum TAC respectively. Moreover, the development of surrogate models is illustrated in this section as well. Section 4 presents the optimised results of the maximum HVGO yield and the minimum TAC of VDU respectively. Finally, the conclusion and future work are given in Section 5.
2. Methodology
In this research study, the methodology is conducted in three stages: process simulation, surrogate modelling and mathematical optimisation.

2.1. Process Simulation
Two types of models are mainly used to design the crude distillation unit, namely shortcut model and rigorous model. A rigorous model can simulate mass balance, energy balance and equilibrium relations for every stage of the distillation unit. Therefore, the rigorous model is chosen in this research study as it provides more accurate predictions. Commercial rigorous simulation software (i.e., Aspen HYSYS, Aspen Plus, Pro/II of SimSci-Esscor, etc.) are commonly used for modelling crude distillation units in the refinery. In this work, Aspen HYSYS version 10.0 is selected to develop the crude oil blending model and rigorous distillation model of the vacuum distillation unit (VDU). This is because Aspen HYSYS contains existing routines specific to solving the distillation columns, unlike other software.

2.1.1. Crude Oil Characterisation and Blending. For the simulation, three crude assays are considered as different choices of crude blends. Heavy crude is considered as one of the crude assays. Meanwhile, higher API gravity crude assays, namely light and medium assays are considered for the remaining crudes. As mentioned in the research gap, crude oil blending plays an important role to maximise the profitability of the refinery. Consequently, a crude oil blending model is considered here to mix the heavy crude with other crudes to be processed as feedstock in the distillation operation. Thus, the assay data of the crude oil is first added and calculated by Aspen HYSYS. By doing so, the working curves such as internal true boiling point, molecular weight, density and viscosity curves are generated. Then, a set of hypothetical pseudo-components representing each crude oil is created from the working curves. By using the oil product cut option, different types of crude oil are blended in a designated ratio in the oil environment of Aspen HYSYS. Once the blending is completed, the crude oil is installed in the simulation environment and the rigorous distillation model of the crude distillation unit is built and discussed in Section 2.1.2.

2.1.2. Simulation Development. The crude distillation unit is simulated based on the operating parameters and column specifications listed in Aspen HYSYS version 10 [17], which has been validated. Peng-robinson equation-of-state is selected to simulate the simulation model as it is the most suitable thermodynamic fluid package to predict the volumetric properties and phase behaviour of the selected crude oil. The simulation model of crude distillation unit included preheating trains, atmospheric distillation unit and VDU, as shown in figure 1. The atmospheric distillation unit consisted of 50 trays, 1 total condenser, 3 side strippers and 3 pump-arounds. The products of the atmospheric distillation unit are naphtha, kerosene, diesel, atmospheric gas oil and atmospheric residue. The atmospheric residue of the atmospheric distillation unit is fed into VDU which consists of 14 trays and 2 pump-arounds. After that, VDU allowed fractionation of the atmospheric residue into off-gas, light vacuum gas oil, heavy vacuum gas oil, slop wax and vacuum residue under vacuum pressure.
It is worth noting that the number of stages for the distillation columns was determined using a series of procedures. The procedures consist of various levels of simulation rigour. Firstly, a component splitter was used to determine the removal ratios and operating conditions required to achieve the required yield. With the removal ratios and operating conditions, a shortcut distillation simulation was performed to determine the number of stages required to achieve the said yield. The number of stages here is then used as input data for the rigorous simulation model to establish the baseline performance of the distillation column. This baseline simulation model is later used to develop sampling points for the regression analysis. Thus, it is assumed that the number of stages in the column is fixed.

2.2. Surrogate Modelling

In the second stage, the surrogate modelling technique is employed to represent the simulation environment of the vacuum distillation unit (VDU). The surrogate model is built around the data generated from the simulation model with a random selection of input variables and their corresponding bound limits. The data is then regressed and converted into equations that can be used to represent the operation of the distillation system. One advantage of the surrogate model is that it allows a rapid calculation of the model output responses by applying equations to relate the input variables to the output responses. There are two types of surrogate models that have been reported in many publications such as response surface method and artificial neural network. As compared to the response surface method, an artificial neural network in the distillation column is more complicated, consists of more intensive steps and requires more investigation for the data regression [8]. On the other hand, the major advantage of the response surface method is that it requires very short computational time to study the relationship between the factors. The behaviour between the output responses and the important influencing variables can be captured easily using regression analysis [18]. Besides, the response surface method can produce a mathematical expression based on the data supplied and this mathematical expression essentially helps to represent the performance of the system mathematically. The mathematical expression from RSM is allowed to be integrated into a mathematical optimisation model to enhance and optimise a process or a system. It can be used in a mathematical model to obtain optimal responses with minimum variance and a small number of controlled parameters. These combinations allow the decision-makers to make more informed decisions about the operation of the system. Hence, the response surface method is selected to be used in this work. A general first-order polynomial function of a response surface equation is expressed as the following form:

\[
Y = c + \sum_{\alpha=1}^{N} \beta_{\alpha} X_{\alpha}
\]

where \( Y \) represents the dependent variables, \( X \) represents the independent variables, \( c \) is the intercept regression coefficient, \( \beta \) is the linear term regression coefficient and \( N \) is the number of factors.

Due to the complexity and non-linearity of VDU operations, the surrogate modelling technique is applied in this work to represent the relationship between the input and output responses. The surrogate
models of VDU are built based on the aforementioned response surface method. Before the construction of VDU surrogate models, the dependent variables and independent variables are determined. In order to achieve the objectives of this research study, 4 dependent variables are considered such as HVGO yield, HVGO pump-around cooler utility flowrate, LVGO pump-around cooler utility flowrate and column diameter. Meanwhile, crude oil blending ratio and operating conditions of VDU are selected as the independent variables for all the dependent variables. In order to determine the coefficient values that make up the response surface equations above, a large number of sample points are generated from the simulation model developed in the previous stage. The number of sample points is determined based on the amount of data available for the equipment considered, the amount of data available after data pre-processing and the quality of the regressions generated using the available data. In this work, the sample points are generated by varying the crude oil blending ratio and operating conditions of VDU. Statistical analysis software (i.e., JMP, Python, etc.) is typically used to regress the generated sample points in response surface equations [8]. In this work, JMP version 15.0 is selected to regress the response surface equations because it can easily assess the data from various sources which allows user to build their model rapidly. Besides, JMP can link the statistical data to interactive graphics which helps the user to explore and visualise their data better. After the regression analysis, validation analysis is carried out to evaluate the accuracy of surrogate models. An evaluation of R-square, P-value and leverage plots are considered. R-square of the equations must be greater than 0.90 to indicate a good fit of data, while the P-value of each variable must be smaller than 0.05 to show a significant impact on the dependent variable. The impacts of each independent variable on the dependent variables are further identified using the leverage plots. By performing the validation analysis, those independent variables that have a weak impact on the dependent variables are excluded from the development of the surrogate models. Finally, the surrogate models are built and implemented in the optimisation model, as discussed in Section 2.3.

2.3. Mathematical Optimisation
Large numbers of mathematical programming tools are available for the development of the distillation optimisation model. In this work, LINGO version 18.0 is used to develop an optimisation model to maximise the HVGO yield and minimise the TAC of the VDU respectively. The optimisation model is developed based on the response surface equations of the surrogate models, as discussed in Section 2.2.

3. Case Study
In this section, a case study is presented to illustrate the methodology that is discussed in Section 2. The case study aims to determine the optimal crude blending composition with their respective operating conditions that give maximum heavy vacuum gas oil (HVGO) yield and minimum total annualised cost (TAC) respectively.

The surrogate models are built to represent the relationship between the independent variables and dependent variables of the vacuum distillation unit (VDU). As discussed in the methodology section, HVGO yield, HVGO pump-around cooler utility flowrate, light vacuum gas oil (LVGO) pump-around cooler utility flowrate and column diameter are chosen as the dependent variables of the surrogate models. Besides, the development of the VDU surrogate model considered the crude oil blending ratio and operating conditions as the independent variables. Three different types of crude oil blending are conducted based on different crude blending ratios of light crude, medium crude and heavy crude. The bound limits of each crude oil ratio are given in table 1.

| Crude Blending Ratio   | Lower | Upper |
|------------------------|-------|-------|
| Light crude ratio      | 0.2   | 0.8   |
| Medium crude ratio     | 0.2   | 0.8   |
| Heavy crude ratio      | 0.2   | 0.8   |
The operating conditions of the case study are varied based on the crude blending ratio. 7 input variables, related to the operating conditions are selected. Those variables are furnace outlet temperature, flash zone temperature, column top pressure, column bottom pressure, stripping steam flowrate, HVGO pump-around flowrate and LVGO pump-around flowrate. Each variable with the respective lower and upper bound limits is summarised in table 2.

| Operating Conditions                  | Lower  | Upper  | References |
|---------------------------------------|--------|--------|------------|
| Furnace outlet temperature (°F)       | 720    | 780    | [19]       |
| Flash zone temperature (°F)           | 650    | 750    | [19]       |
| Top pressure (mmHg)                   | 40     | 60     | [20]       |
| Bottom pressure (mmHg)                | 65     | 85     | [19]       |
| Stripping steam flowrate (lb/hr)      | 15,000 | 30,000 | [21]       |
| HVGO pump-around flowrate (bbl/day)   | 10,000 | 45,000 | -          |
| LVGO pump-around flowrate (bbl/day)   | 10,000 | 35,000 | -          |

The lower and upper limits of each operating condition are determined by performing a literature review. Note that the lower and upper limits of HVGO and LVGO pump-around flowrate are not reported in the literature review. Hence, case studies are conducted in the simulation model to fine-tune the limits of HVGO and LVGO pump-around flowrate. Besides, the process constraints are also determined, as given in table 3.

| Constraints                          | Specifications |
|--------------------------------------|----------------|
| Feedstock (bbl/day)                  | 99,000         |
| Column Pressure Drop (mmHg)          | 25             |
| LVGO ASTM D86 95 % recovery (°F)     | 915            |
| HVGO ASTM D86 95 % recovery (°F)     | 1,050          |

The total crude blending flowrate that fed into the distillation model is held constant at the value specified in the Aspen HYSYS listed conditions. On the other hand, the pressure drop of VDU is fixed at the maximum value of 25 mmHg, suggested by Jones [19]. In order to meet the composition specification of the products, ASTM D86 95 % cut points for both HVGO and LVGO are set as constant. Consequently, the product flowrate of VDU is allowed to be varied.

After determining the independent variables and process constraints, the surrogate models for each independent variable are constructed. 654 sets of sampling points are generated from the Aspen HYSYS simulation model for the development of VDU surrogate models. For each crude oil blending, 4 response surface equations are generated based on the above-mentioned dependent variables. In the case of heavy-light crude, the sampling points are generated based on the different ratios of the heavy-light crude blends with their respective operating conditions. Then, JMP is used to regress the corresponding data in the response surface equations for the heavy-light crude blend. The following discussions use one of the dependent variables, HVGO yield as an example. The regression results of the HVGO yield in the heavy-light crude blend surrogate model are demonstrated in figure 2, figure 3 and table 4.
The regression coefficients of the crude blending ratio and 7 operating variables that are calculated by JMP version 15 are given in figure 2. As observed, the HVGO pump-around flowrate has influenced the HVGO yield the most as HVGO pump-around flowrate has the highest log worth value. Log worth value is defined as $10^{-\log(p\text{-value})}$. The lower the P-value, the higher the log worth value. This is then followed by the light crude ratio, LVGO pump-around flowrate and so on. On the other hand, it can be observed that heavy crude ratio and flash zone temperature have the least but still significant impacts on the HVGO yield. Apart from that, the leverage plots can help to further investigate the effect of the independent variable on the HVGO yield. The leverage plots of the three independent variables that have the most significant impacts on HVGO yield are illustrated in figure 3. The effect is significant at the 5% level when the confidence curve (red coloured line) crosses the hypothesis line (blue coloured horizontal line). In contrast, if the confidence curve does not cross the hypothesis line, the effect is considered not significant. With that, the leverage plots suggested that these three independent variables have a significant impact on the HVGO yield. To further validate the model, the R-square of the model must be higher than 0.9. As observed from table 4, the R-square of the HVGO yield surrogate model is 0.957 and this has confirmed that the model has high accuracy. Hence, all the selected independent variables are included in the development of the surrogate model. The constructed surrogate models are then integrated into the optimisation model.

An optimisation model of VDU is formulated using LINGO version 18.0 to determine the maximum HVGO yield and minimum TAC respectively. The optimisation work of the VDU has led to mixed-
integer non-linear programming containing linear, non-linear and integer variables as well as constraints. The model had a total of 169 variables, 12 non-linear variables, 3 integer variables, 196 constraints and 12 non-linear constraints. Based on the previous surrogate model development, 4 response surface equations of the surrogate model are generated for heavy-light crude blend, heavy medium crude blend and heavy-medium-light crude blend, respectively. The HVGO yield surrogate model is integrated into the optimisation model to address the objective function of maximising HVGO yield. Meanwhile, the other 3 surrogate models are incorporated in the optimisation model to address the objective function of minimising TAC. In order to achieve the objective functions, all the surrogate model equations in three crude blendings are incorporated in the optimisation model. In this model, two stages of optimisation procedures are being considered, covering (a) maximise HVGO yield under varying the crude oil blending ratio and operating conditions, (b) minimise TAC under varying the crude oil blending ratio and operating conditions.

4. Results and Discussions
In this section, the optimal solution of maximising heavy vacuum gas oil (HVGO) yield and minimising total annualised cost (TAC) are presented, as given in table 5.

| Objective Function | Maximise HVGO yield | Minimise TAC |
|--------------------|---------------------|--------------|
| Selected Crude Blending | Heavy - Light Crude | Heavy - Light Crude |
| HVGO Yield (bbl/day) | 23,788 | 14,028 |
| TAC ($ MM) | 1.75 | 1.70 |
| Light crude fraction | 0.8 | 0.8 |
| Medium crude fraction | 0 | 0 |
| Heavy crude fraction | 0.2 | 0.2 |
| Furnace outlet temperature (°F) | 720 | 720 |
| Flash zone temperature (°F) | 750 | 750 |
| Top pressure (mmHg) | -13.54 | -13.54 |
| Bottom pressure (mmHg) | -13.05 | -13.05 |
| Stripping steam flowrate (lb/hr) | 15,000 | 15,000 |
| HVGO pump-around flowrate (bbl/day) | 45,000 | 10,000 |
| LVGO pump-around flowrate (bbl/day) | 35,000 | 10,000 |

The purpose of these two objectives is to provide a comparative analysis of the resulting configurations. Based on the results presented above, the crude blending has considered a light crude ratio of 0.8 and a heavy crude ratio of 0.2 in both objectives. The light crude contains more distillate and it is easier to refine as compared to medium and heavy crude. It can be proven that the heavy-light crude achieves maximum HVGO yield and minimum TAC as compared to the heavy-medium crude blend and heavy-medium-light crude blend. Moreover, other decision variables of VDU such as the operating conditions have significant impacts on the HVGO yield and TAC as well. Maximum HVGO yield is achieved by operating the VDU at the recommended operating conditions, as shown in table 5. On the other hand, it is important to note that only the HVGO pump-around flowrate and LVGO pump-around flowrate are optimised and stayed at the lowest bound of the variable in order to minimise the TAC of the VDU. Hence, it can be concluded that the optimisation results offer the possibilities of the different configurations based on the objective functions. The results can be used to provide insights into the optimal operating conditions of the VDU. By doing so, it can help the refiners to develop an effective strategic planning application for the VDU operation [22]. The developed insight can also help to overcome the current limitations such as tighter crude demand present in the refineries. Apart from that, the optimisation model can be used to perform computational experiments which can help to avoid costly mistakes in the building pilot-scale plant. In terms of flexibility, the proposed surrogate-assisted
mathematical optimisation model can be extended to other equipment or industries. This can be done by incorporating the proposed methodology to develop surrogate models based on the statistical data of the equipment or industry. Then, the next thing is to develop a new optimisation model based on the surrogate models of the equipment or industry.

5. Conclusions and Future Works
In this work, a surrogate-assisted methodology is developed to capture the relationship between input and output responses of the vacuum distillation unit (VDU), and thus predicted the complex behaviour of the VDU. The constructed surrogate models are then integrated into the optimisation model to obtain an optimal solution for maximising heavy vacuum gas oil (HVGO) yield and minimising total annualised cost (TAC) respectively. The proposed methodology is implemented in a case study where different crude oil blending ratios are compared and analysed. As a result, the heavy-light crude blend provided a higher HVGO yield and lower TAC as compared to other types of crude oil blending. As concluded, the proposed model provided the optimal operating condition of VDU which can be used to provide insight for the decision-makers upon deciding a preferred solution. With this insight, the refiners can easily develop strategic planning procedures to tackle the issues present in the real VDU operation in the refinery.

Future work will focus on the development of a multi-objective optimisation model to optimise the HVGO yield and TAC of VDU simultaneously. This is because these two objectives are contradicting in nature, which means that improvement in HVGO yield may bring a negative impact on TAC or vice versa. Therefore, a multi-objective optimisation model can be used to determine the trade-off between these two objectives. Additionally, further studies will be carried out to analyse the impact of the number of stages for the distillation columns on the HVGO yield and TAC. Lastly, future work will also focus on using the real-time data from the refinery to improve the accuracy of the proposed model.

Acknowledgments
This work has supported by the School of Engineering and Physical Sciences at Heriot-Watt University Malaysia.

References
[1] Martin G R and Nigg J M 2001 Vacuum pressure control: Impact on profitability Pet. Technol. Q. 6 73–81
[2] Pintarič Z N and Kravanja Z 2006 Selection of the economic objective function for the optimisation of process flow sheets Ind. Eng. Chem. Res. 45 4222–32
[3] Seo J W, Oh M and Lee T H 2000 Design optimisation of crude oil distillation Chem. Eng. Technol. 23 157–64
[4] Basak K, Abhilash K S, Ganguly S and Saraf D N 2002 On-line optimisation of a crude distillation unit with constraints on product properties Ind. Eng. Chem. Res. 41 1557–68
[5] Gu W, Wang K, Huang Y, Zhang B, Chen Q and Hui C W 2015 Energy optimisation for a multistage crude oil distillation process Chem. Eng. Technol. 38 1243–53
[6] Inamdar S V, Gupta S K and Saraf D N 2004 Multi-objective optimisation of an industrial crude distillation unit using the elitish non-dominated sorting genetic algorithm Chem. Eng. Res. Des. 82 611–23
[7] Huang X, Zhao T, Li N, Ma Z, Song L and Li J 2017 Emission-reductive and multi-objective coordinative optimisation of binary feed for atmospheric and vacuum distillation unit China Pet. Process Petrochemical Technol. 19 101–12
[8] Loper M L 2015 Modeling and Simulation in the Systems Engineering Life Cycle: Core Concepts and Accompanying Lectures (Springer)
[9] Denimal E, Nechak L, Sinou J and Nacivet S 2016 Kriging surrogate models for predicting the complex eigenvalues of mechanical systems subjected to friction-induced vibration Shock Vib. 2016 1–22
[10] López D C, Hoyos L J, Mahecha C A, Arellano-Garcia H and Wozny G 2013 Optimisation model of crude oil distillation units for optimal crude oil blending and operating conditions Ind. Eng. Chem. Res. 52 12993–3005

[11] Yao H and Chu J 2012 Operational optimisation of a simulated atmospheric distillation column using support vector regression models and information analysis Chem. Eng. Res. Des. 90 2247–61

[12] Ochoa-estopier L M, Jobson M and Smith R 2013 Operational optimisation of crude oil distillation systems using artificial neural networks Comput. Chem. Eng. 59 178–85

[13] Gut J A W, Franzoi R E, Menezes B C, Kelly J D and Grossmann I E 2020 Cutpoint temperature surrogate modeling for distillation yields and properties Ind. Eng. Chem. Res. 59 18616–28

[14] Mittal V, Zhang J, Yang X and Xu Q 2011 E3 Analysis for crude and vacuum distillation system Chem. Eng. Technol. 34 1854–63

[15] Gu W, Huang Y, Wang K, Zhang B, Chen Q and Hui C W 2014 Comparative analysis and evaluation of three crude oil vacuum distillation processes for process selection Energy 76 559–71

[16] Li X, Cui C and Sun J 2018 Enhanced product quality in lubricant type vacuum distillation unit by implementing dividing wall column Chem. Eng. Process Process Intensif. 123 1–11

[17] Aspen HYSYS 2017 Crude Tower Simulation - HYSYS v10 MA Aspen Technology Inc

[18] Aydar A Y 2018 Utilisation of response surface methodology in optimisation of extraction plant materials Stat. Approaches With Emphas. Des. Exp. Appl. to Chem. Process. 157–69

[19] Jones D S J 2015 Handbook of Petroleum Processing vol. 1 (Springer Science & Business Media)

[20] Lousdad S 2020 What is vacuum distillation 1–5. https://www.arab-oil-naturalgas.com/what-is-vacuum-distillation/ (accessed December 20, 2020)

[21] More R K, Bulasara V K, Uppaluri R and Banjara V R 2010 Optimisation of crude distillation system using Aspen Plus: Effect of binary feed selection on grass-root design Chem. Eng. Res. Des. 88 121–34

[22] Geoffrion A M 1976 The purpose of mathematical programming is insight, not numbers Interfaces (Providence) 7 81–92