Stochastic optimization approaches in solving energy scheduling problems under uncertainty

P S Pravin * **Xiaonan Wang** **

* *Department of Chemical & Biomolecular Engineering,
National University of Singapore, Singapore 117585, Singapore.
(e-mail: chepps@nus.edu.sg).
** Department of Chemical Engineering,
Tsinghua University, Beijing 100084, China.
(e-mail: wangxiaonan@mail.tsinghua.edu.cn)*

Abstract: This paper focuses on evaluating the performances of various approaches to solve a day-ahead energy scheduling Mixed Integer Linear Programming (MILP) cost minimization problem, with hybrid energy generation system consisting of solar photo voltaic (PV), waste to energy (WTE) and main grid. The system under consideration is a highly energy-intensive industrial facility that can be generalized to any industrial process regardless of the domain. Out of the three case studies considered in this paper, the first case study deals with adopting unsupervised machine learning approach to generate a reduced number of probabilistic scenarios of the uncertain parameters and implements the Here & Now and Wait & See algorithms to solve the resulting stochastic optimization problem. While the second case study directly collects and assumes a finite set of historical data of the uncertain parameters as scenarios, the third case study uses an evolutionary algorithm called Limited Evaluation Evolutionary Algorithm (LEEA) to solve the energy scheduling problem with associated constraints. Various performance metrics such as mean absolute error (MAE) and root mean square error (RMSE) are utilized to compare the performances of the various approaches presented through case studies. The values of these evaluation metrics showcased the enhanced performance of evolutionary algorithm based approach compared to the other two approaches.

Keywords: Hybrid Energy Sources, Energy scheduling, Data-driven stochastic optimization, Scenario reduction, Unsupervised machine learning, Evolutionary algorithm.

1. INTRODUCTION

The traditional strategy followed in any energy-intensive industry is to prepare a schedule of operation based on the energy demand and the energy-supply cost. However, the main technological challenge for energy-intensive industrial processes solely depending on the electricity grid lies in the (1) high uncertainty in future electricity prices and (2) non-eco-friendly use of conventional fossil fuel based power generation mechanisms. These drawbacks can be averted by the penetration of renewable energy sources (RES) into the energy market. Although the inclusion of RES leads to economic cost savings and reduced environmental effects, the main challenge with these non-conventional energy sources (especially solar PV and wind energy) lies in the intermittency in availability. For instance, the energy generation from these sources can vary based on the time of the day and seasonal variations. Scheduling decisions made without adequate consideration of these uncertainties can often lead to uneconomical plant operation and demand unfulfillment. The aforementioned drawbacks of these energy resources can be mitigated by incorporating several complementary solutions such as flexible generation, use of energy storage technologies (e.g. batteries), enforcing demand response, and the use of a hybrid combination of several energy resources.

The research area of optimization in production scheduling has been a subject of great interest both in the operations research community and the process systems engineering community (Zhang and Grossmann (2016), Pravin et al. (2020)). Several studies examined production scheduling of industrial operations particularly concentrating on energy intensive chemical processes. While Zhang et al. (2015) focused on the optimal scheduling of air separation unit taking electric energy and reserve market into account, Cao et al. (2016) considered varying market conditions to study the optimal dynamic operation of air separation plants. Several other studies have focused on the design of industrial operations with time-varying electricity prices. While Miller et al. (2008) analyzed the intermittent operation of an industrial facility using the process model and three electricity price tiers, the optimized design of an industrial facility subject to constant product demand was studied by Pattison and Baldea (2014) with two electricity price tiers. Mitra et al. (2014) developed optimization problems for energy-intensive industrial processes and applied it for an air separation unit. They formulated a set of two optimization problems including a classical problem

with a deterministic product demand and a stochastic optimization problem with uncertain product demands.

Generally, the traditional fashion in production scheduling assumes a constant future electricity price profile and at often times neglect the effect of uncertainty (Misra et al. (2021)). However, the magnitude and shape of the future electricity prices generally fluctuates based on the energy market conditions. On a similar note, while considering the highly intermittent renewable energy sources, the uncertainty associated with the energy generation based on the seasonal variations need to be considered in the scheduling phase. For handling the aforementioned uncertainties, stochastic optimization is traditionally adopted which is designed based on the past historical trends of these uncertain parameters, generally referred to as scenarios. The scenario data sets are generated for each uncertain parameters and can easily make the stochastic optimization problem computationally complex. Scenario reduction techniques are generally adopted that reduces the big dataset to a reduced number of data, eventually reducing the computational burden and memory usage.

In this work, as a case study, we consider the optimal energy scheduling of a highly energy-intensive industrial facility that can be generalized to any industrial processes irrespective of the domain. We consider three energy resources viz. solar PV, waste to energy (WTE) that utilizes the waste generated within the industry to generate energy, and main grid that generally uses a combination of several energy generation mechanisms, in a hybrid fashion that optimally supplies energy to the industrial facility based on the electricity prices and availability of each sources. While the future solar PV energy generation depends on the highly intermittent solar irradiance, the main grid electricity price in the future depends on several factors like the fluctuations in the electricity market as well as pull and push of the demand and supply. Stochastic optimization is formulated for the resulting energy scheduling problem to generate an optimal day-ahead schedule of the energy mix among the available energy sources.

The aims of this work are (1) to utilize the historical data of the uncertain parameters to generate scenarios and implement unsupervised learning algorithm (K-means clustering in this study) to reduce the scenarios to a smaller dataset with associated probabilities to perform the stochastic optimization, (2) to implement the classical Here & Now and Wait & See approaches to solve the resultant stochastic MILP based energy scheduling problem, (3) to solve the same problem using the immediate past finite dataset of the considered uncertain parameters with equal probabilities, (4) to implement and analyze the capability of evolutionary algorithms (Limited Evaluation Evolutionary Algorithm (LEEA) in this study) that completely avoids the scenario reduction method and work on a limited number of data samples for each generation to solve the scheduling problem, and finally (5) to compare the performances of the above discussed approaches using various performance metric analysis such as MAE and RMSE.

The remainder of this paper is organized as follows: A brief description of the system under consideration is presented in Section 2 followed by a discussion on the
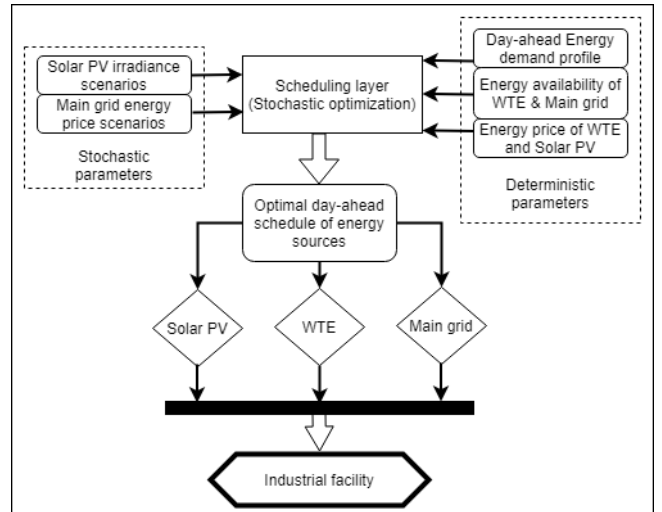

Fig. 1. Block diagram showing the hybrid power system with associated power sources

various classical stochastic optimization approaches and the problem formulation in Section 3. Section 4 elaborates on the the various case studies considered in this work with the results and discussion in Section 5. This paper ends with concluding remarks in Section 6.

2. SYSTEM DESCRIPTION

The system under consideration is a generalized energy-intensive industrial facility designed with a hybrid combination of various energy resources. A combination of three different energy sources namely the solar PV, waste to energy (WTE) and main grid are equipped in order to ensure undisrupted delivery of energy to the industrial facility. As the energy system parameters such as the main grid energy price and the solar PV irradiance/energy are stochastic in nature, the historical data for these parameters (referred to as scenarios) are collected from various sources. For demonstration purpose, all associated data and parameter values considered in this study are collected from authorized weather and electricity websites of Singapore. For instance, past solar irradiance data and main grid electricity price data were collected from the meteoblue website (Meteoblue (2020)) and the Energy Market Authority (EMA) website (EMA (2020)) of the Singapore government respectively. As illustrated in Figure 1, the stochastic optimizer in the scheduling layer uses the reduced scenario sets of these uncertain parameters together with the deterministic parameter values to generate an optimal day-ahead schedule of the individual energy contribution of the available energy sources.

3. STOCHASTIC OPTIMIZATION

Stochastic optimization or stochastic programming is an approach of solving optimization problems that involve uncertainty. In the case of uncertainty in the input data to the optimization problem, a set of discrete scenarios $s \in S$ with associated probabilities p_s for the uncertain parameters gets generated using the scenario reduction method, resulting in a probability distribution of future realizations of the uncertain parameters. We implement

two classical solution approaches: the Here & Now and Wait & See approaches to the formulated stochastic optimization problem. A brief discussion on these two approaches together with their mathematical formulation is presented in the upcoming subsections followed by the energy scheduling optimization problem formulation for the proposed industrial facility.

3.1 Here & Now approach

In the Here & Now approach, with several possible outcomes $s \in S$ of the uncertain parameters, a decision has to be made now before knowing which of the possible outcomes s realizes. The general formulation of the optimization problem for the Here & Now approach is as follows:

$$z_{HN} = min_x \mathbb{E}_{s \in S} f(x, s) \tag{1}$$
$$s.t. \quad x \in C = \cap_{s \in S} C_s$$

where, $f(x, s)$ denotes the objective function with decision variable x which has to be feasible for all possible scenarios s, and C_s is the feasible region given by the constraints in scenario s. The expected value operand in Equation 1 can be formulated as the sum of probabilities of the scenarios as shown in Equation 2.

$$\mathbb{E}_{s \in S} f(x, s) = \sum_{s \in S} p_s f(x, s) \tag{2}$$

Hence, in this approach, the optimal decision variables $x^\star$ minimizes the expected value of the objective function z_{HN} assuming that all scenarios in set S can occur.

3.2 Wait & See approach

The Wait & See approach generally assumes that we can wait to design our system until uncertainty occurs. In essence, this approach assumes perfect knowledge of the future uncertainties which is generally not implementable in practice, but provide appropriate bounds on the solution. This approach optimizes each scenario s in the reduced scenario set separately as follows:

$$z_s = min_x f(x, s) \tag{3}$$
$$s.t. \quad x \in C_s$$

For each scenario s, an optimal solution z_s with corresponding decision variables x_s is obtained, resulting in a set of S optimal solutions. Finally, the expected value of the optimal cost is calculated as follows:

$$z_{WS} = \mathbb{E}_{s \in S} z_s = \sum_{s \in S} p_s z_s \tag{4}$$

3.3 Problem formulation

A stochastic optimization problem is formulated to minimize the overall cost of energy contributed by the hybrid energy sources with associated constraints. Altogether, two uncertain parameters are considered in this case viz. solar irradiance/energy and the main grid energy price. These uncertain parameters are represented by the superscript s in the objective function J of the optimization problem as follows.

$$J = min \sum_{t=1}^{NT} (P_S^t E_S^{t,s} + P_{WTE}^t E_{WTE}^t + P_M^{t,s} E_M^t) \tag{5}$$

It need to be ensured that the energy delivered by the hybrid combination of energy sources always meets the demanded load energy by the industrial facility at all times. This constraint is formulated as follows:

$$ED^t \leq \sum_{e}^{E} E_e^t \tag{6}$$

The problem also makes sure that the power delivered by each energy sources should be always between the lower and upper bounds of the designed capacity of each energy sources. The upper and lower bounds are decided by taking the stochastic nature of the solar PV generation into account in order to ensure that the inequality holds. This constraint is formulated as follows:

$$E_{S,min}^t \leq E_S^{t,s} \leq E_{S,max}^t \tag{7}$$
$$E_{WTE,min}^t \leq E_{WTE}^t \leq E_{WTE,max}^t \tag{8}$$
$$E_{M,min}^t \leq E_M^t \leq E_{M,max}^t \tag{9}$$

where, NT is the total number of time slots in the horizon, t is the time period, s denotes stochastic parameters, $e \in E$ denotes energy sources, P_e^t is the price of energy from e^{th} energy source at time period t, E_e^t is the energy available from e^{th} energy source at time period t, ED^t is the energy demanded by the load at time period t, $E_{e,min}^t$ is the lower bound/capacity of the eth energy source at time period t and $E_{e,max}^t$ is the upper bound/capacity of the eth energy source at time period t. The average energy prices of solar PV, WTE and main grid used for solving the energy scheduling problem are 0.065, 0.088 and 0.097 SGD/kWh respectively.

4. CASE STUDIES

Three case studies are considered for demonstrating the stochastic energy scheduling problem for the energy-intensive industrial facility. The first case study adopts scenario reduction using machine learning technique (K-means clustering) to reduce the big dataset (1 year data of both solar irradiance and main grid energy price) to a reduced number of scenarios to perform stochastic optimization using the Here & Now and Wait & See approaches. The second case study uses the immediate past 10 days data of the uncertain parameters with equal probabilities to perform the stochastic optimization. In the third case study, the concept of scenario reduction is avoided and the stochastic optimization problem is solved using the Limited Evaluation Evolutionary Algorithm that works on a limited number of data samples from the big dataset for each generation. Finally, a thorough performance evaluation of the three approaches is carried out by comparing the day-ahead schedules of these approaches with the actual day profile with the help of evaluation metrics like Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). For demonstration purpose, the day of interest (for which the day-ahead schedule is generated) is taken as 31st of Dec 2020 and the associated data and the parameter values used for the simulation are that of Singapore. Also, it is assumed that the industrial facility under consideration is operational only for a duration of 10 hours starting from 8 am and ending at 5 pm Singapore time. The simulations for all the case studies are performed on a Windows IntelCore i7 (1.80 GHz) processor.

4.1 Case study 1

In this case study, as part of the scenario generation for stochastic optimization, past 1 year data (1st Jan 2020 to 30th Dec 2020) of the solar irradiance and main grid electricity price is collected from appropriate sources in Singapore. As discussed, K-means clustering algorithm is applied to this data to reduce the big dataset to a reduced number of scenarios. K-means clustering is a type of unsupervised machine learning algorithm that generally finds groups (known as clusters) within the big data that have not been explicitly labeled. In K-means clustering, the number of centroids (i.e. value of K) we wish to divide our data points into, has to be pre-determined by the user. A centroid is basically an imaginary or real location representing the center of the cluster. For this study, we have chosen the value of K as 10 meaning that the past 1 year dataset for the uncertain parameters are reduced to 10 scenarios with associated probabilities. Figure 2 and Figure 3 shows the reduced scenarios for the solar irradiance and main grid electricity price using K-means clustering with their corresponding probabilities respectively.
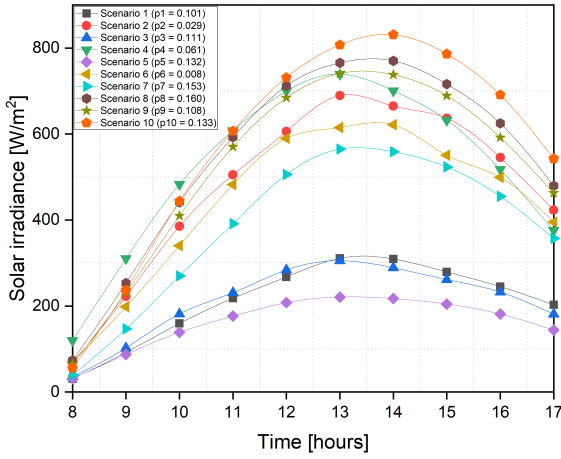

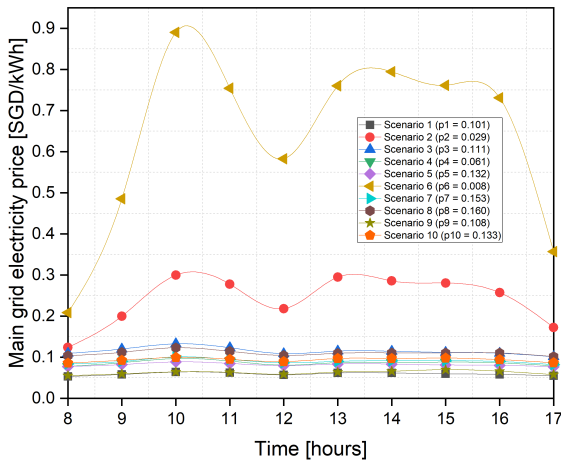Fig. 2. Reduced scenarios of solar irradiance

Fig. 3. Reduced scenarios of main grid electricity price

Both Here & Now and Wait & See approaches are performed to solve the stochastic optimization problem to generate the optimal energy mix among the available sources. The Here & Now approach involves a single optimization run over all the scenarios, where expected value is a part of the objective function of the optimization problem. The Here & Now approach for the cost minimization problem is observed to give an objective function value of 8.8290 Singapore Dollars (SGD) which indicates the total cost incurred to meet the energy demand for the day. It is important to note that the Wait & See approach is not implementable in practice, as it assumes perfect knowledge of the uncertainties in the future. However, the Wait & See approach rather provides an upper bound on the solution of the optimization problem. The Wait & See approach for the cost minimization problem is observed to give an objective function value of 10.9517 SGD which indicates the upper bound for the solution. The profile showing the energy demanded by the industrial facility and the individual contribution of energy among the available energy sources to meet the requested demand is shown in Figure 4. For demonstration purpose, we restrict our focus to show the results only for the Here & Now approach as this approach shows better affinity to reality as compared to the Wait & See approach, which is not practically implementable. However, from the objective function values of these two approaches, it is to be noted that there isn't much difference in the solutions. The cost minimization objective function with the associated constraints is solved using the CPLEX solver available in General Algebraic Modeling Language (GAMS), with a computational time of 15.75 seconds.
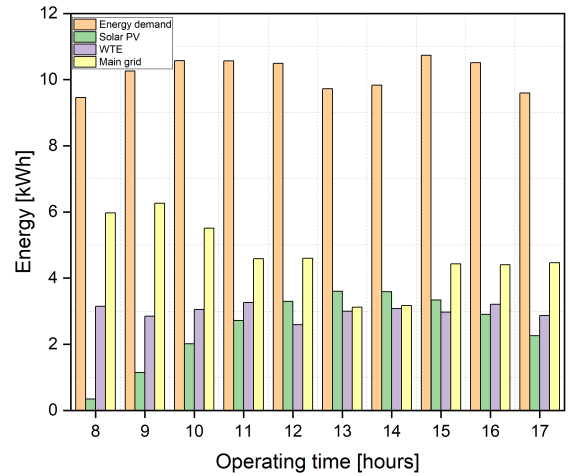

Fig. 4. Individual contribution of energy sources

4.2 Case study 2

In this case study, in order to ensure an effective comparison with the previous case study, immediate past 10 days data (21st Dec 2020 to 30th Dec 2020) of the stochastic parameters are collected from the appropriate sources in Singapore, which eventually acts as the 10 scenarios with equal probabilities (0.1 in this case) for performing the stochastic optimization energy scheduling problem. Figure 5 and Figure 6 shows the scenarios from the past 10 days for the solar irradiance and main grid electricity price. By adopting this data as the scenarios with equal probabilities, stochastic optimization is performed using the

Here & Now approach to generate the day-ahead energy schedule for the industrial facility. The individual energy contribution by the various energy sources together with the energy demanded by the load is shown in Figure 7.
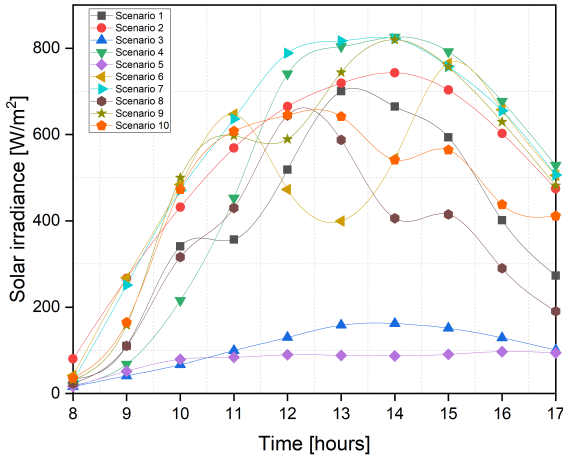

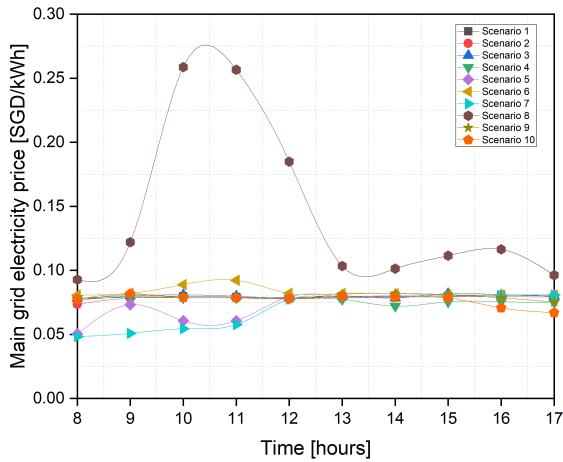Fig. 5. Past 10 days data of solar irradiance

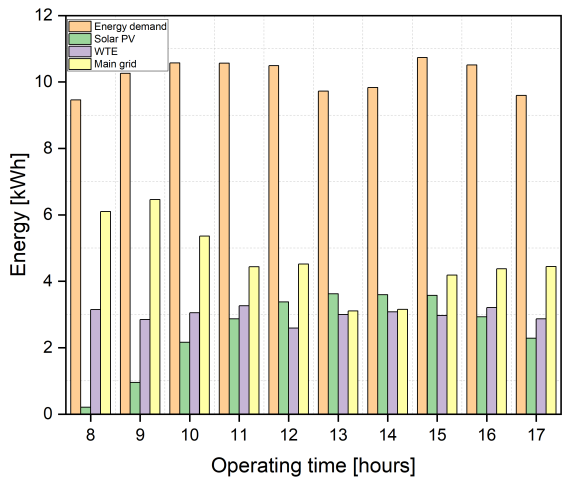Fig. 6. Past 10 days data of main grid electricity price

Fig. 7. Individual contribution of energy sources

The objective function value resulted from solving the scheduling problem in this case study is seen to be 8.8117

SGD which shows a slight variation from the value obtained in case study 1. This is expected due to the presence of stochasticity in the parameters in both the case studies. The clear differences will be evident from the final comparison of the day-ahead schedule with the actual day profile (i.e. 31st Dec 2020) which will be dealt with in the upcoming discussions. This algorithm is also solved using the CPLEX solver available in General Algebraic Modeling Language (GAMS), with a computational time of 13.08 seconds.

4.3 Case study 3

Limited Evaluation Evolutionary Algorithms (LEEA) are a special class of evolutionary algorithms specifically designed for problems with exogenous stochasticity. This algorithm is capable of handling mixed integer programs comprising of both continuous and binary variables. One of the key advantages of stochastic gradient descent is that the gradient does not have to be calculated for the complete training dataset (Prellberg and Kramer (2018)). In essence, the gradient gets calculated for a single or a small number of training examples at a time that significantly reduces the computational effort compared to working with the entire training set. This in effect, also helps in reducing the chance of solution getting stuck in local optima. LEEA is one such algorithm that evaluates the population against a limited number of datasets for each generation (Yaman et al. (2018)). This lean approach
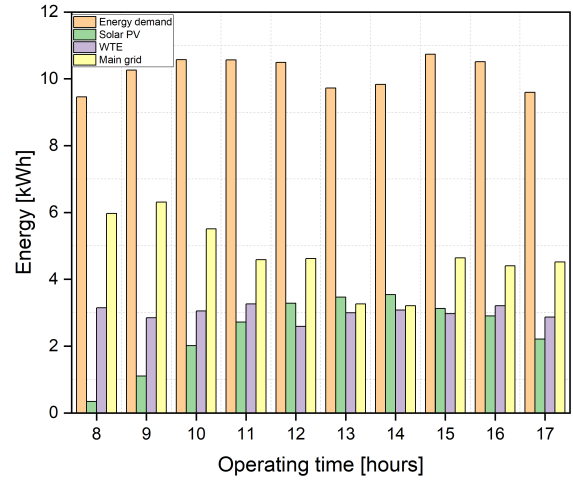

Fig. 8. Individual contribution of energy sources

of evaluation greatly relieves the computational load, particularly on large training sets. Hence, this algorithm does not confide in scenario reduction strategies compared to other traditional stochastic optimization algorithms and works on a small number of dataset within the big dataset (1st Jan 2020 to 30th Dec 2020). The main design parameters for the LEEA algorithm are the population size and the generation size of the problem under consideration which is generally user specific. For this case study, we used a population size of 40 and a generation size of 100 and the optimization problem is solved using the Python IDE: PyCharm resulting in an objective function value of 10.7231 SGD with a computational time of 23.83 seconds. The day-ahead energy schedule obtained from the LEEA algorithm is shown in Figure 8.

5. RESULTS & DISCUSSION

In order to compare the performance of algorithms in the above discussed case studies, the actual day profiles for the solar irradiance (or solar energy) and the main grid electricity price for the 31st of Dec 2020 is collected from appropriate sources in Singapore. It is interesting to note that due to the deterministic nature of the WTE source in terms of both availability and energy cost, the energy contribution from WTE source is seen to be exactly similar for all the three case studies irrespective of the algorithm adopted, as can be observed from Figures 4, 7 and 8. Figure 9 shows the comparison of the actual profiles of the hourly energy delivered from solar PV and main grid energy sources with the forecasted profiles in all the three case studies. For effective comparison of the profiles in the case studies with the actual day profile, evaluation metrics such as the mean absolute error (MAE) and root mean square error (RMSE) are calculated from the data points in Figure 9. Table 1 shows the MAE and RMSE values of

Table 1. Evaluation metric values for Solar PV
and Main grid

Metric	CS1	CS2	CS3
MAE (kWh)	0.423	0.460	0.385
RMSE (kWh)	0.545	0.563	0.523

the two uncertain energy sources for the three case studies. It can be observed that the MAE and RMSE values are smaller for the case study 3 with LEEA algorithm implying that the energy schedule generated in case study 3 better matches with the real day profile as compared to the other two case studies. The increasing order of magnitude of the MAE and RMSE values for the three discussed case studies are observed to be CS3 < CS1 < CS2.
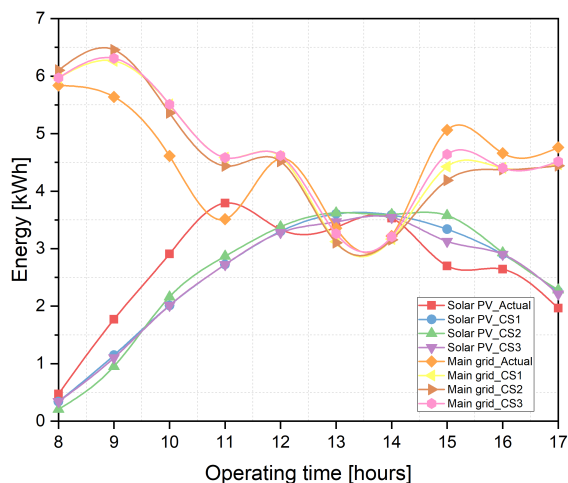

Fig. 9. Comparison of day-ahead schedules of three case studies with the actual day profile

6. CONCLUSION

A day-ahead energy scheduling problem with hybrid energy generation system combining solar PV, WTE and main grid to meet the demanded energy for a generic industrial facility has been considered in this paper. Three case studies are mainly dealt with to solve the energy

scheduling MILP optimization problem using various approaches. The first case study uses K-means clustering for scenario reduction of the training dataset to reduce the big data to a reduced set of scenarios with associated probabilities. A finite number of immediate past days data is utilized in the second case study to represent the scenarios for the stochastic optimization problem. The third case study avoids the scenario reduction of the training data and uses an evolutionary algorithm (LEEA) to solve the energy scheduling problem. It can be observed from the results that although the LEEA in case study 3 possess slightly larger computation time compared to the other two cases, it is seen to outperform the other 2 case studies in terms of lesser MAE and RMSE values.

REFERENCES

Cao, Y., Swartz, C., and Flores-Cerrillo, J. (2016). Optimal dynamic operation of a high purity air separation plant under varying market conditions. *Industrial & Engineering Chemistry Research*, 55.

EMA (2020). "https://www.ema.gov.sg".

Meteoblue (2020). "https://www.meteoblue.com".

Miller, J., Luyben, W.L., and Blouin, S. (2008). Economic incentive for intermittent operation of air separation plants with variable power costs. *Industrial & Engineering Chemistry Research*, 47(4), 1132–1139.

Misra, S., Pravin, P.S., Gudi, R.D., and Bhartiya, S. (2021). Integration of supply and demand side management using renewable power sources: Application on an air separation plant. *Industrial & Engineering Chemistry Research*, 60(9), 3670–3686.

Mitra, S., Pinto, J.M., and Grossmann, I.E. (2014). Optimal multi-scale capacity planning for power-intensive continuous processes under time-sensitive electricity prices and demand uncertainty. part i: Modeling. *Computers & Chemical Engineering*, 65, 89–101.

Pattison, R.C. and Baldea, M. (2014). Optimal design of air separation plants with variable electricity pricing. In M.R. Eden, J.D. Siirola, and G.P. Towler (eds.), *The 8th International Conference on Foundations of Computer-Aided Process Design*, volume 34 of *Computer Aided Chemical Engineering*, 393–398. Elsevier.

Pravin, P.S., Misra, S., Bhartiya, S., and Gudi, R.D. (2020). A reactive scheduling and control framework for integration of renewable energy sources with a reformer-based fuel cell system and an energy storage device. *Journal of Process Control*, 87, 147–165.

Prellberg, J. and Kramer, O. (2018). Limited evaluation evolutionary optimization of large neural networks. *CoRR*, abs/1806.09819.

Yaman, A., Mocanu, D.C., Iacca, G., Fletcher, G., and Pechenizkiy, M. (2018). Limited evaluation cooperative co-evolutionary differential evolution for large-scale neuroevolution. In *Proceedings of the Genetic and Evolutionary Computation Conference*, GECCO '18, 569–576.

Zhang, Q. and Grossmann, I.E. (2016). *Planning and Scheduling for Industrial Demand Side Management: Advances and Challenges*, 383–414. Springer.

Zhang, Q., Grossmann, I.E., Heuberger, C.F., Sundaramoorthy, A., and Pinto, J.M. (2015). Air separation with cryogenic energy storage: Optimal scheduling considering electric energy and reserve markets. *AIChE Journal*, 61(5), 1547–1558.