2024 18th International Conference on
Control, Automation, Robotics and Vision (ICARCV)
December 12-15, 2024. Dubai

QLSTM-Based Microgrid Daily Operation with Renewable Uncertainty

Zeina Bahij, Najmus Sahar, Bing Yan

Abstract— **Microgrid is a small-scale grid where generation is close to the demand allowing more penetration of renewables, like photovoltaic (PV). However, the intermittent nature of PV power generation poses a significant challenge in microgrid operation, especially on days with highly variable weather conditions. In this paper, a deep reinforcement Q-learning long short-term memory (QLSTM) model is developed to predict the operation strategy of a microgrid for the next day at a 15-minute time interval. To address the uncertainty of PV power and demand, the previous three days' PV and load data are added as inputs to the model since weather conditions on consecutive days may depend on similar atmospheric conditions. Also, to address the effect of propagation of error in the long forecasting horizon with multiple steps, a moving window training method is implemented. The moving window will be shifted by 15 minutes at each step once the actual PV and load data are available till the end of the day. The model is tested in a microgrid consisting of combined cooling, heating and power, heat pump, PV, battery, and heating and cooling energy storage systems. Results show that our model outperforms gated recurrent unit, LSTM, and Q-learning for testing data from different months. Also, it shows better performance than MATLAB 2023 Optimization Toolbox (the branch-and-bound method) which uses forecasted data, especially on a day with highly variable weather conditions.**

Keywords: Microgrid operation, PV, QLSTM, uncertainty.

I. Introduction

Microgrid supports a flexible and efficient grid by enabling the integration of distributed energy resources. It allows more penetration of renewables, like photovoltaic (PV) along with other conventional generators and storage units. With multiple units generating power at the same time, there is a need to optimize the daily microgrid operation to specify the hourly power generation amount needed from each unit to satisfy the variable demand. The main difficulty in the operation optimization of microgrids lies in the intermittency and uncertainty of renewables and demand. Also, predicting PV and demand at once for the next day at multiple time intervals results in a high forecasting error.

The daily operation of microgrids has been studied by many researchers as reviewed in Section II. A commonly used approach is deterministic where the PV power and demand forecasted data are used without considering the uncertainties. This approach shows a high deviation from optimal generation levels due to the forecasting error in the input data (PV and demand). In the stochastic approach, renewables are modeled using scenario-based approaches without considering their changes with time. The most recent approach for solving the problem is machine learning where a model is trained to predict the operation strategy using PV and load forecasted data as input. It shows better results than previous approaches. However, during days with highly variable weather conditions, the high forecasting error results in improper scheduling of microgrids. Also, the effect of propagation of forecasting error is high in a long forecasting horizon with multiple steps.

In this paper, a deep reinforcement Q learning long short-term memory (QLSTM) model is developed to predict the operation strategy of a microgrid. The microgrid structure and problem formulation are presented in Section III. The QLSTM model is developed in Section IV. To address the problem of uncertainty of the input data (PV and demand), the last three days' data of actual PV and demand and the LSTM forecasted data for the next day time steps are added as inputs to the QLSTM model since consecutive days' weather conditions may depend on similar atmospheric conditions. Adding previous days data method is not possible in previous mathematical built approaches. Also, to address the problem of propagation of error in a long forecasting horizon with multiple steps, a moving window training method is implemented. As time passes, the actual input data will be available and added to the model inputs. Then the moving window is shifted by one time step till the end of the day. Flexibility of QLSTM model in updating input information, which previous models lack, will provide more accurate results.

The testing results are presented in section V. Datasets consist of the PV power data that was collected from a 2-MW PV farm at Rochester Institute of Technology (RIT) in Rochester, New York, and power, cooling, and heating demand data that were collected from a large hotel in New York State. Using the actual PV and demand data, the operation problem of the microgrid under study is solved by MATLAB Optimization Toolbox (branch and bound) to generate the training and testing data. The prediction performance of the developed QLSTM model was compared to Gated Recurrent Unit (GRU), LSTM, and Q-Network performance prediction using testing data from different months where QLSTM outperformed all other models. Also, the QLSTM model is tested on three different days with low, medium, and high variance in PV power and shows better performance than solving using the mathematical model with forecasted PV and demand data, especially on a day with high PV power variance.

The main contributions of this paper are as follows:

1. A deep reinforcement learning QLSTM model is developed to address the uncertainty of PV and demand by adding the last three days of actual data of PV and demand along with the next day's forecasted data.
2. A moving window training method is implemented to manage the issue of the effect of propagation of error in a long forecasting horizon with multiple steps.

Zeina Bahij, Najmus Sahar, and Bing Yan are with the Department of Electrical and Microelectronic Engineering, Rochester Institute of Technology, Rochester, NY 14623, USA (e-mail: zb6781@rit.edu, nl1991@rit.edu,, bxyeee@rit.edu).

This work is based upon work supported by the National Science Foundation (NSF) under Award ESSC-2340095. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

2024 18th International Conference on Control, Automation, Robotics and Vision (ICARCV) | 979-8-3315-1849-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICARCV63323.2024.10821662

Authorized licensed use limited to: Rochester Institute of Technology. Downloaded on June 26,2025 at 20:04:54 UTC from IEEE Xplore. Restrictions apply.

II. Literature Review

This section gives a summary of the existing work on microgrid operation under the uncertainty of renewables. Subsection A summarizes existing work done using math optimization, and subsection B summarizes some research papers that used machine learning models.

A. Math Optimization

Modeling renewable uncertainties to solve the operation problem of microgrids has been studied by many researchers. In [1], Bing et al. developed a stochastic mixed-integer linear programming model with uncertain renewable generation modeled by a Markovian process for a microgrid that includes different energy conversion devices and thermal storage systems. The problem is aimed at reducing energy costs and environmental impacts. Markovian process was also used in [2] by M. Di Somma et al to optimize energy costs and exergy efficiency. A multi-objective linear programming problem was formulated in a microgrid consisting of CCHP, biomass, heat pump, thermal solar plant, and thermal store units. In [3], An economic model predictive control scheme is developed to achieve optimal economic performance in the operational costs of the microgrid. The control scheme was tested in a simulated microgrid composed of a wind turbine, a set of PV panels, battery in a grid-connected mode.

However, due to the intermittent nature of weather conditions, mathematical models are inaccurate. An optimization problem should consider the uncertainty in the PV generation. In [4], Duan et al. developed a two-stage expected-scenario robust optimization approach to address the uncertainty effect of renewables output and load in a microgrid. The first stage was the day ahead scheduling stage where all possible scenarios, including the worst case, were considered; the second was the real-time rescheduling stage. This approach showed effectiveness in reducing uncertainty impact.

B. Machine learning methods

With the recent advancement in artificial intelligence (AI) methods to solve optimization problems, some researchers used AI-based approaches to solve the operation problem in microgrids. A. Chaouachi et al [5] developed an intelligent energy management system that combines the artificial neural network technique and the linear programming-based multi-objective optimization technique. The developed model aimed to minimize generation cost and environmental impacts in a microgrid consisting of PV, Wind Turbine (WT), fuel cell, batteries, and a microturbine. Inputs of the model were the predicted PV power generation, wind power, and load demand for the coming 24 hours at one hour step. Then, the power generation levels from each generation unit were determined. In [6], M. Gamez et al. developed a new recurrent neural network model to optimize the operation of a microgrid made of PV, WT, batteries, and electric vehicles (EV). The optimization problem was solved using the multiagent system, where each agent was defined to be a unit in the microgrid and sent information about its power generation levels and constraints. The developed model determines the optimal power over one week for wind, solar, and battery to minimize the power delivered from the grid and maximize the power supplied from renewable energy sources. Results showed that the developed model achieved optimal generation levels from each source.

Due to the effect of deep reinforcement learning in optimization problems, it has been used in the operation optimization of microgrids. In [7], Kuznetsova developed a 2-steps ahead reinforcement learning model to optimize the operation of a microgrid made of WT and batteries. Optimization aimed at increasing the utilization rate of batteries and wind turbines during high-load power. Battery scheduling actions were defined by the reinforcement learning model that was fed by predicted WT power generation. In [8], a deep reinforcement learning model was used to predict the operation strategy of a microgrid, inputs of the model were the forecasted PV generation and the states of each energy device, which is the power of each device generation. During days with highly variable weather conditions, the forecasting error is high leading to improper scheduling of generation. Also, the effect of propagation of errors in forecasting PV and load data in one shot for the next 24 hours will result in improper scheduling.

III. Problem Formulation

This section gives a summary of the microgrid structure under study and the problem formulation. Subsection A presents the structure of the microgrid; modeling of energy devices is discussed in subsection B; power and thermal balance are presented in subsection C; and the objective function is presented in subsection D.

A. Microgrid structure

The microgrid under study consists of a CCHP, heat pump, PV, battery, and heating and cooling energy storage systems. The microgrid structure is shown in Fig. 1. The CCHP includes a gas turbine that generates power from burning natural gas. The recovered exhaust gas can be used by the heat recovery boiler and the absorption chiller along with the direct flow of natural gas for space cooling and space heating. The heat pump supplies space heating and space cooling for charging thermal storage and satisfying the load. PV provides electricity to satisfy the demand, charge the battery, and supply the heat pumps. Power demand and heat pump required power are satisfied by the CCHP generated power from the gas turbine, PV power, the grid power, and the power discharged from the battery. Space heating demand is satisfied by the thermal energy provided by the CCHP through the heat recovery boiler, heat pump, and the thermal energy discharged from the space heating thermal storage. Space cooling demand is satisfied by the thermal energy provided by the absorption chiller, heat pump, and the thermal cooling energy discharged.

B. Modelling of energy devices

For all the devices except for PV, battery, and thermal storage, there is a set of binary variables x to present the device's on/off status. If the device is on ($x_{dev}(t)$ is 1), its generation level $P_{dev}(t)$ has to be within minimum and maximum levels as shown in equation (1),

$$x_{dev}(t)P_{dev}^{min} \leq P_{dev}(t) \leq x_{dev}(t)P_{dev}^{max}. \quad (1)$$

1) CCHP

CCHP consists of a gas turbine, an absorption chiller, and a heat recovery boiler. The volumetric flow rate of natural gas in the gas turbine is presented in equation (2),

$$G_{ICE}(t) = E_{CCHP}/\eta_e LHV_{gas}, \quad (2)$$

where E_{CCHP} is the electricity provided by the engine, η_e is the engine gas-to-electric efficiency, and LHV_{gas} is the lower heat value of natural gas. In addition, between two consecutive

Authorized licensed use limited to: Rochester Institute of Technology. Downloaded on June 26,2025 at 20:04:54 UTC from IEEE Xplore. Restrictions apply.

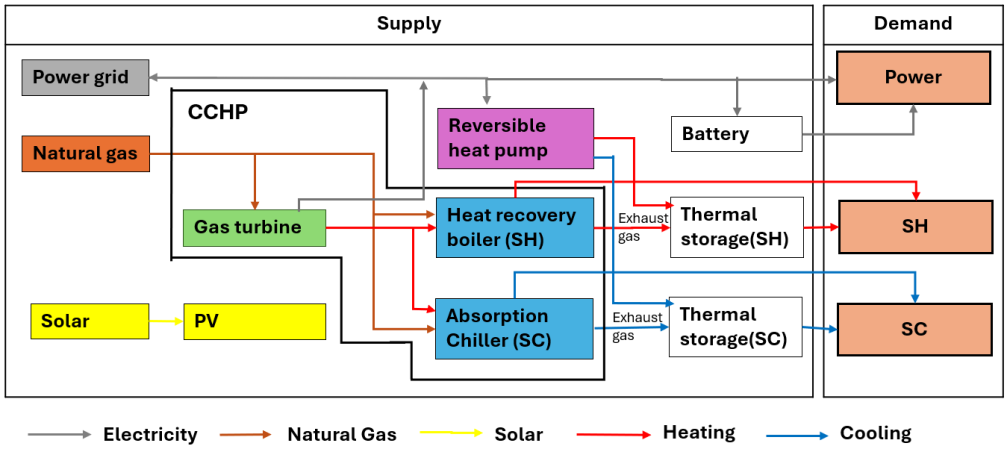

Fig. 1: Microgrid Structure

time steps, the change in electricity level should not exceed ramp-up or go below ramp-down rates.

The heat rate available from the exhaust gas recovered from the gas turbine is shown in equation (3),

$$Q_{GT,ex}(t) = P_{GT}(t)(1-\eta_e-\mu_{GT})/\eta_e, \quad (3)$$

where μ_{GT} is the present heat loss of the gas turbine. This exhaust gas is recovered by the heat recovery boiler and the gas turbine to generate heating and cooling as shown in equations (4) and (5), respectively,

$$H_{HRB,ex}(t) = Q_{GT,ex}(t).\ \xi_{SH}(t).\eta_{HR,HRB}\ .\ COP_{HRB}, \quad (4)$$

$$C_{abs,ex}(t) = Q_{GT,ex}(t).\ \Xi_{SC}(t).\eta_{HR,abs}\ .\ COP_{abs}, \quad (5)$$

where $\xi_{SH}(t)$and $\Xi_{SC}(t)$ are the fractions of exhaust gas for space heating and cooling respectively; their summation should be one, $\eta_{HR,HRB}$ and $\eta_{HR,abs}$ are the waste heat recovery efficiency, and COP_{HRB} and COP_{abs} are the coefficients of performance of the heat boiler and the absorption chiller, respectively,

The volumetric flow rate of natural gas by the heat recovery boiler and the absorption chiller are presented in equations (6) and (7), respectively,

$$G_{HRB,di}(t) = H_{HRB,di}(t)/(LHV_{gas}.\eta_{HP}\ .\ COP_{HRB}\ ), \quad (6)$$

$$G_{abs,di}(t) = C_{abs,di}(t)/(LHV_{gas}.\eta_{abs}\ .\ COP_{abs}), \quad (7)$$

where $H_{HRB,di}$ and $C_{abs,di}$ are the heating and cooling rates generated by natural gas, and η_{HRB} and η_{abs} are the efficiencies of the heat recovery boiler and absorption chiller combustors, respectively.

2) Heat Pump

The power consumption of the heat pump for heating is shown in (8),

$$P_{HP,H}(t) = (H_{HP}(t) \times \Delta t)/COP_{HP,H}, \quad (8)$$

where $COP_{HP,H}$ is the coefficient of performance of the heat pump in the heating mode and $H_{HP}(t)$ is the heating rate in the heat pump.

The cooling in the heat pump is modeled the same way as in heating.

3) Modeling of the battery

The state-of-charge (SOC) dynamics is presented in equation (9),

$$SOC_{BESS}(t+1) = SOC_{BESS}(t).(1-\delta_{BESS}) - \frac{P_{discharge}(t)}{\eta_{discharge}}.\Delta t + \eta_{charge}.P_{charge}(t)\ .\ \Delta t, \quad (9)$$

where $SOC_{BESS}(t)$ and $SOC_{BESS}(t+1)$ are the energy of BESS at the time t and t+1, respectively, $\eta_{discharge}$ and η_{charge} are the charge and discharge efficiencies respectively, and δ_{BESS} is the energy loss ratio of the storage.

The battery cannot be charged and discharged simultaneously, and the SOC level cannot go below or above the SOC maximum and minimum limits.

4) Modeling of the thermal storage

The energy stored for space heating is shown in (10),

$$H_{sto}(t) = H_{sto}(t-\Delta t).\ \left(1-\varphi_{sto}(\Delta t)\right) + \left(H^{in}_{sto}(t) - H^{out}_{sto}(t)\right)\Delta t, \quad (10)$$

where φ_{sto} is the heat loss through the tank wall, Δt is the length of the time interval, $H^{in}_{sto}(t)$, $H^{out}_{sto}(t)$ are the heat rates brought in and taken out by the flow in and out of water respectively, and $H_{sto}(t)$, $H_{sto}(t-\Delta t)$ are the energy stored and the current and previous time steps, respectively.

The space heating stored energy cannot go below and above the space heating storage minimum or maximum limits. The space cooling in thermal storage is modeled the same way as the heating storage.

C. Power and thermal balance

A balance between power demand, $P_{demand}(t)$ heating demand, $H_{demand}(t)$, and cooling demand, $C_{demand}(t)$, and the generation should be established by the energy devices generation and the grid power as shown in equations (11)-(13), respectively,

$$P_{demand}(t) + P_{charge}(t) + P_{sell}(t) + P_{HP}(t) + P_{CP}(t) = P_{discharge}(t) + P_{buy}(t) + P_{PV}(t) + P_{CCHP}(t), \quad (11)$$

$$H_{demand}(t) = H_{CCHP}(t) + H_{HP}(t) + H^{out}_{sto}(t) - H^{in}_{sto}(t), \quad (12)$$

$$C_{demand}(t) = C_{CCHP}(t) + C_{CP}(t) + C^{out}_{sto}(t) - C^{in}_{sto}(t), \quad (13)$$

where $P_{sell}(t)$ and $P_{buy}(t)$ are the sold and bought power levels from and to the grid, respectively.

D. Objective Function

This work aims at minimizing the operation cost which is the grid power cost, the difference between buying power from the grid and selling power to the grid, as shown in equation

Authorized licensed use limited to: Rochester Institute of Technology. Downloaded on June 26,2025 at 20:04:54 UTC from IEEE Xplore. Restrictions apply.

(14), and the natural gas cost as shown in equation (15). The total operation cost is the summation of grid power cost and natural gas cost as shown in equation (16),

$$C_{grid} = \sum_{t=1}^{T}\left(c_{buy}(t)P_{buy}(t).\Delta t - c_{sell}(t)P_{sell}(t).\Delta t\right), \quad (14)$$

$$C_{gas\,turbine} = \sum_{t=1}^{T}\left(c_{gas}G_{CCHP}(t)\right), \quad (15)$$

$$C_T = C_{gas\,turbine} + C_{grid}, \quad (16)$$

where c_{buy} and c_{sell} are the cost/revenue of buying/selling power from/to the grid, and c_{gas} is the cost of gas.

IV. Methodology

In this section, the QLSTM model for predicting the operation strategy in the microgrid under study is presented. The main challenge in the microgrid optimization problem lies in the intermittency and uncertainty of renewables and demand data. A hybrid prediction model that has the ability of both addressing the uncertainty in input variables and predicting for the optimal reward value can address the above issue. Q-learning is a type of reinforcement learning that shows effectiveness in predicting the optimal state-action reward value, that matches with the operation problem of a microgrid. However, it is not a time series prediction model, it will take the input data at once to predict the operation strategy for the next day. To address the above issue, an LSTM layer is added to predict the operation strategy in a time series manner. The PV data mainly depend on weather conditions, weather conditions on consecutive days depend on the same atmospheric conditions, so a history of the last three days of PV and demand data achieves the goal of using experience.

The environment, agents, actions, and reward of the QLSTM model are taken as below:

1. *Environment*

The environmental observations space, $E(t)$, are taken to be the input data, the PV power data at time t, $P_{PV}(t)$, the power demand data at time t, $P_{demand}(t)$,the space cooling load at time t, $C_{demand}(t)$, and the space heating load at time t, $H_{demand}(t)$, as shown in equation (17),

$$E(t):\{P_{PV}(t), P_{demand}(t), C_{demand}(t), H_{demand}(t)\}. \quad (17)$$

2. *Agents*

Agents, G, are the energy devices to be scheduled, gas turbine, absorption chiller, heat recovery boiler, battery, heat pump, thermal storage, and grid power as shown in equation (18),

$$G:\left\{\begin{matrix} gas\ turbine. absorption\ chiller, heat\ recovery \\ boiler, battery, heat\ pump, thermal\ storage, \\ grid\ power \end{matrix}\right\}. \quad (18)$$

3. *Actions*

Actions, $A(t)$, are presented in equation (19) which are the gas turbine natural gas flow rate at time t, $G_{GT}(t)$, the ratio of supplied exhaust gas to the heat recovery boiler for heating at time t, $\xi_{SH}(t)$, the flow rate of natural gas in the heat recovery boiler at time t, $G_{HRB,di}(t)$, the ratio of supplied exhaust gas to the absorption chiller for cooling at time t, $\xi_{SC}(t)$, cooling directly from the absorption chiller at time t, $G_{abs,di}(t)$, the heat pump heating rate at time t, $H_{HP}(t)$, the heat pump cooling rate at time t, $C_{CP}(t)$, battery charging power at time t, $P_{charge}(t)$, battery discharging power at time t, $P_{discharge}(t)$, heating rate input to the storage, $H_{sto}^{in}(t)$, heating rate output to the storage, $H_{sto}^{out}(t)$, cooling rate input to the storage, $C_{sto}^{in}(t)$, cooling rate output to the storage, $C_{sto}^{out}(t)$,

$$A(t):\left\{\begin{matrix} G_{GT}(t), \xi_{SH}(t), G_{HRB,di}(t), \xi_{SC}(t), \\ G_{abs,di}(t), H_{HP}(t), C_{CP}(t), P_{charge}(t), \\ P_{discharge}(t), H_{sto}^{in}(t), H_{sto}^{out}(t), \\ C_{sto}^{in}(t), C_{sto}^{out}(t) \end{matrix}\right\}. \quad (19)$$

4. *States*

States are presented in equation (20) which are the gas turbine power at time t, $P_{GT}(t)$, heating rate generated by the exhaust gas $H_{HRB,ex}(t)$, heating rates generated by natural gas at time t, $H_{di,HRB}(t)$, cooling rate generated by exhaust gas at time t $C_{abs,ex}(t)$, cooling rates generated by natural gas $C_{di,abs}(t)$, power consumption of the heat pump for heating at time t, $P_{HP,H}(t)$, the power consumption of the heat pump for cooling at time t, $P_{HP,C}(t)$, energy stored for space heating at time t, $H_{sto}(t)$, and the energy stored for space heating at time t, $C_{sto}(t)$,

$$S(t):\left\{\begin{matrix} P_{GT}(t), H_{HRB,ex}(t), H_{di,HRB}(t), C_{abs,ex}(t), \\ C_{di,abs}(t), P_{HP,H}(t), P_{HP,C}(t), H_{sto}(t), C_{sto}(t) \end{matrix}\right\}. \quad (20)$$

5. *Reward*

Since in the Q learning model, the model tends to maximize the reward and the microgrid operation aims at minimizing the total operation cost, the reward R_T is taken to be the negative of the total operation cost of one day $-C_T$, as expressed in equation (21),

$$R_T = -(C_{gas\,turbine} + C_{grid}), \quad (21)$$

where T represents the testing day.

The QLSTM model developed to predict the operation strategy of the microgrid under study is shown in Fig. 2. The LSTM model will take the data for the previous three days starting from t-3×96 till t+96 where t=0 represents the start of the testing day for both the environmental observations $O(t$-3x96),…, $O(t$+96)} actions $A(t$-3x96),…, $A(t$+96)}, and the output will be the predicted states, $S(t)$,…, $S(t$+96) the Q-Network will take the series of predicted states as inputs and will output the Q-value reward R_T which is the total operation cost of the microgrid on the testing day.

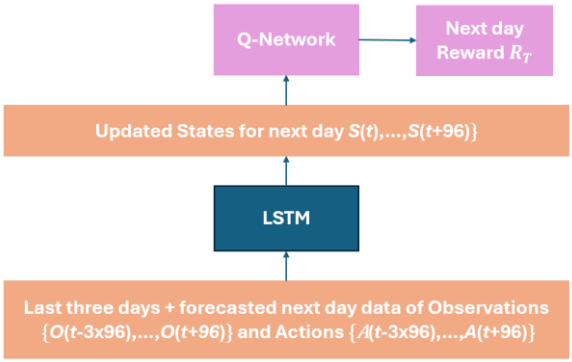

Fig. 2: QLSTM model

V. Results

The QLSTM model is tested in this section. The microgrid's configuration under study is presented in

Authorized licensed use limited to: Rochester Institute of Technology. Downloaded on June 26,2025 at 20:04:54 UTC from IEEE Xplore. Restrictions apply.

subsection A; the dataset description is presented in subsection B; testing criteria in subsection C; QLSTM configuration and testing results in subsection D; comparison to other models results in subsection E; and testing during variable weather conditions in subsection F.

A. Microgrid Configuration

Microgrid configuration and the size per device are shown in Table I. The microgrid configuration was determined based on the demand. Coefficients shown in equations (1) – (11) for modeling energy devices are taken from [9].

Table I: Microgrid Size

Energy Device	Size Per Unit
Gas Turbine	40-300 MW
Battery	5-50 MW
Absorption Chiller	10-100 MW
Heat Recovery Boiler	15-150 MW
Heat Pump	192 MW – 128 MW (Heating, Cooling)

B. Dataset Description

The dataset used to test the model consists of PV, demand price data, and the training and testing datasets.

1) *Demand*

Power, heating, and cooling demand data are taken from a large hotel in New York State for each 15- minutes for the year 2020 [10] with a total of 35,040 data points.

2) *PV*

The PV power data was taken from a 2-MW solar farm from RIT for the year 2020 [11]. The power generation was scaled up 10 times to adjust to the power demand of the large hotel. The time step is 5 minutes, but to get the same time interval data as the available demand data, the data points at every 15 minutes are considered in this study.

3) *Price*

The price of natural gas is fixed at 180 /MWh for the whole day [12] as shown in equation (22),

$$\text{Cost}(P_{GT}(t)) = 180 \text{ \$/MWh} \quad \text{for } 0 \le t \le 23 \tag{22}$$

The cost of buying power from the grid is shown in equation (23). The price is higher during peak hours 9:00 am to 6:00 pm, [12]

$$\text{Cost}(P_{grid}(t)) = \begin{cases} 140 \text{ \$/MWh} & \text{for } 0 \le t \le 9 \\ 210 \text{ \$/MWh} & \text{for } 9 \le t \le 18 \\ 140 \text{ \$/MWh} & \text{for } 18 \le t \le 23 \end{cases} \tag{23}$$

The price of selling power to the grid is 80% of the buying cost as shown below in equation (24),

$$\text{Revenue}(P_{sell}(t)) = 0.8 \times \text{Cost}(P_{buy}(t)) \quad \text{for } 0 \le t \le 23 \tag{24}$$

4) *Training and testing dataset*

A dataset is generated by solving the operation optimization problem of microgrids using MATLAB 2023 Optimization Toolbox (the branch-and-bound method) for the year 2020. The average solving time for one day is 10.54 seconds. A total of 35,040 data points are created. The dataset includes input variables, (PV power, power heating and cooling demand), the actions and states described in IV.B.3 and IV.B.4, respectively.

From the above data, 80% of data from each month was taken for training, 10% for validation, and 10% for testing to consider all possible changes in the weather conditions.

C. Testing criteria

The prediction error is calculated based on the total operation cost. After predicting the operation strategy for the whole day, and specifying the generation levels from each device the operation cost is calculated using equation (19). First, correlation (r), mean absolute error (MAE), and root mean square error (RMSE) as shown in equations (25)- (27) are calculated between the predicted cost and the optimal cost in case the actual data is used to solve the problem by MATLAB 2023 Optimization Toolbox (branch and bound) for the testing data.

$$r = \frac{\sum_{i=1}^{n}((x_i-\bar{x})(y_i-\bar{y}))}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2 \sum_{i=1}^{n}(y_i-\bar{y})^2}}, \tag{25}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - y_i|, \tag{26}$$

$$RMSE = \sqrt{\frac{1}{n}(\sum_{i=1}^{n}(x_i - y_i))^2}. \tag{27}$$

where xi and yi are the real and forecasted values, respectively, and $\bar{x}$and $\bar{y}$ are the average values of the real and the forecasted data, respectively.

D. QLSTM configuration and testing results

After performing a hyper parameter tuning, the number of hidden layers in LSTM is 4, each with 96 units which is the number of time steps in one day. The epoch size is set to 100, and the batch size is 64 using the "Adam" optimizer. In the Q-learning layer, same number of hidden layers (4) is and units (96) are used, the learning rate is set to 0.01 to prevent very fast or slow convergence to optimal value and the discount rate to 0.95 to present the importance of the future rewards since the model aims at minimizing the total cost of microgrid operation of one day. The QLSTM model has been implemented in MATLAB 2023 and simulation is performed on a DESKTOP-MAMPLNJ with a 2.5 GHz Dual-Core Intel i5 processor and system memory of 8 GB 1.5 GHz.

Training and testing results of the QLSTM model are shown in Table II. The training time is 344 seconds while the testing time is 197 seconds for the whole training and testing datasets, respectively.

Table II: Training and testing accuracy and time

Dataset	R(%)	MAE($)	RMSE($)	Time (sec)
Training	96.28	104.30	182.80	344
Testing	95.50	103.18	193.72	197

The prediction time for one day is 5.41 seconds using QLSTM while solving time for one day using MATLAB 2023 Optimization Toolbox is 10.54 seconds.

E. Testing results: comparison to other models

For the whole testing dataset, the correlation, MAE, and RMSE between the predicted total operation cost and the optimal cost using LSTM, Q-learning, and QLSTM are shown in Table III. For a fair comparison, the same moving window method was implemented in all models.

As shown in Table III, GRU and LSTM show the poorest performance, where R are only 84.6% and 85.79%, MAE are 168.2$ and 155.3$ and RMSE are 283.8$ and 263.6$, respectively. The Q-Network model shows better performance than LSTM since LSTM will predict the operation strategy of the microgrid (generation levels of devices) by learning from the previous sequence without considering the objective function or the reward. By adding LSTM and Q-learning

Authorized licensed use limited to: Rochester Institute of Technology. Downloaded on June 26,2025 at 20:04:54 UTC from IEEE Xplore. Restrictions apply.

models together, the hybrid model will take advantage of learning from previous sequences to predict the operation strategy to achieve the reward. QLSTM model outperformed all other models with R of 95.5%, MAE of 103.1$, and RMSE of 193.7$.

Table III: Total cost prediction error using different models

	R (%)	MAE ($)	RMSE ($)
GRU	84.60	168.23	283.80
LSTM	85.79	155.33	263.69
Q-learning	87.71	144.64	240.84
QLSTM	95.50	103.18	193.72

F. Testing during variable weather conditions

Since the PV power represents the weather conditions, the testing days are chosen based on the variance of the PV power. Three days with low, medium, and high variance in PV power days are chosen for further testing. Table shows the variance and forecasting correlation between actual and forecasted values of the PV power.

Table IV: Testing days

Day	Variance	PV Forecasting Error
January, 27	2.78	75.38%
April, 26	4591.01	62.52%
September, 8	1650.58	90.48%

In the microgrid operation problem, the goal is to predict the operation strategy for the next day using the forecasted PV and demand. The error using math optimization or prediction models in this problem is a result of the forecasting error in the input data. In this test, the operation cost using the branch-and-bound method in MATLAB 2023 Optimization Toolbox with the forecasted data and using the QLSTM model are compared as shown in Table V to test the effectiveness of the developed QLSTM model over math optimization to address the uncertainty in the input data.

Table V: Operation cost using math optimization and QLSTM

	OPERATION COST ($)		
Testing Day	Optimal	Math Optimization	QLSTM
January, 27	5892.44	6084.51	6019.10
April, 26	4918.45	5341.28	5101.59
September, 8	1816.80	1919.20	1877.89

As shown in Table V, the QLSTM's total operation cost was closer to the optimal scheduling operation cost on all days. April 26 shows the highest variance in PV power and prediction error and hence highly variable weather conditions. On this day, the branch-and-bound optimization method with forecasted data shows a high difference of 423$ from the optimal cost because of the high forecasting error. QLSTM addressed this issue because of the LSTM layer that works in a time series manner with a difference of 127$.

VI. CONCLUSION

Solving the microgrid operation optimization problem is challenging due to the intermittent and uncertain nature of PV and load, especially during days with highly variable weather conditions. To address this issue, a novel QLSTM model is developed to predict the operation strategy for the next 24 hours with 15- minutes as a time step. The key idea is adding the last three days of actual input data to the model since consecutive days' weather conditions are close to each other. Also, to address the effect of the propagation of errors in multiple-step forecasting, the model has been trained by a moving window method. The moving window is shifted by one time step to get the actual data and ends by the end of the day. The model has been tested using actual PV power data from RIT and demand actual data from a large hotel located in New York State. Testing and training datasets have been created by solving the operation problem using the branch and bound method in MATLAB 2023 Optimization Toolbox. The model performance is compared to three other models, GRU, LSTM, and Q-learning and it shows better performance for testing data from different months. Also, compared to the branch-and-bound method with forecasted data, it shows lower cost, especially on a day with highly variable weather conditions. The developed model needs longer time to coverage in a large action-state space, but still gives high accuracy in predicting the operation strategy. Also, in a transition from sunny days to an intermittent day, the LSTM layer in the developed model can mismatch the relation between the previous days and coming days data, but still predicting the operation strategy of the lower cost compared to other prediction models or solving methods.

References:

[1] Bing Yan, Marialaura Di Somma, Giorgio Graditi, Peter B. Luh, Markovian-based stochastic operation optimization of multiple distributed energy systems with renewables in a local energy community, Electric Power Systems Research, Volume 186,2020, 106364, ISSN 0378-7796.

[2] M. Di Somma, B. Yan, N. Bianco, P. B. Luh, G. Graditi, L. Mongibello, V. Naso, "Multi-objective operation optimization of a Distributed Energy System for a large-scale utility customer", Applied Thermal Engineering, Volume 101, 2016, Pages 752-761, ISSN 1359-4311.

[3] M. Avirup, D. Debapriy, "Optimal operation of microgrid using four different optimization techniques." Sustainable Energy Technologies and Assessments. Volume 21, 2017, 100-120. 10.1016/j.seta.2017.04.005.

[4] Qing Duan, Wanxing Sheng, Haoqing Wang, Caihong Zhao, and Chunyan Ma Power Distribution Technology Center, China Electric Power Research Institute, Beijing 100192, China.

[5] A. Chaouachi, R.M. Kamel, R. Andoulsi, K. Nagasaka, "Multi objective intelligent energy management for a microgrid, "IEEE Transactions on Industrial Electronics Volume 60, Issue 4, 2013, Article number 6157610, Pages 1688-1699.

[6] M. Gamez Urias; E. Sanchez; L. Ricalde, "Electrical Microgrid Optimization via a New Recurrent Neural Network", IEEE SYSTEMS JOURNAL, VOL. 9, NO. 3, SEPTEMBER 2015 945.

[7] E. Kuznetsova, Y. Li, Carlos Ruiz, E. Zio, G. Ault, K. Bell, "Reinforcement learning for microgrid energy management," Energy, Volume 59, 2013, Pages 133-146, ISSN 0360-5442.

[8] H. Alabdullah, Mohammad A. Abido, Microgrid energy management using deep Q-network reinforcement learning, Alexandria Engineering Journal, Volume 61, Issue 11, 2022, Pages 9069-9078, ISSN 1110-0168.

[9] M. Di Somma, B. Yan, N. Bianco, G. Graditi, P.B. Luh, L. Mongibello, V. Naso, "Multi-objective design optimization of distributed energy systems through cost and exergy assessments, Applied Energy, Volume 204, 2017, Pages 1299-1316, ISSN 0306-2619.

[10] Department of Energy, National Renewable Energy Laboratory, "Updated U.S. Low-Temperature Heating and Cooling Demand by County and Sector," July 2023.

[11] RIT's New Solar Energy Farm, "2-megawatt array among largest for any NY college" Aug. 14, 2015.

[12] Nyserda, energy prices "New York 2020 Residential process", Jan.2020.

Authorized licensed use limited to: Rochester Institute of Technology. Downloaded on June 26,2025 at 20:04:54 UTC from IEEE Xplore. Restrictions apply.