Analysing and Optimizing the Refrigeration System Using Machine Learning Algorithm

Avesahemad S. N. Husainy[1], Bhushan S. Kumbhar[2], Prajwal S. Chavan[3], Ajeem A. Attar[4] and Hemant A. Patil[5]
[1] Assistant Professor, [2,3,4&5]UG Research Scholar,
Department of Mechanical Engineering, Sharad Institute of Technology,
College of Engineering, Yadrav, Ichalkarnji, Maharashtra, India
E-mail: avesahemad@gmail.com

Abstract - The Cold chain/Refrigeration plays an important role for preserving vaccines as per the vaccine storage and guidelines prescribed by World Health Organizations and it is better than most of other preservation methods. The primary goal of adopting phase change material is to increase performance, cooling duration, capacity for storage, and to sustain the steady cooling effect for a longer length of time during power outages. In experimental set up it is decoded to use Inorganic and Eutectic phase change materials which is encapsulated in plastic containers inside the vaccine storage. The project will involve collecting data from sensors installed in cabinet and using machine learning algorithms to analyse the data and generate insights. The results of the project can be used to inform best practices for vaccine storage and contribute to the development of more efficient and effective cabinet management systems. This project explores the use of machine learning techniques to optimize the storage of vaccines in cabinet. Proper storage is critical to maintain vaccine efficacy, and therefore it is important to ensure that the temperature and humidity are kept within specific ranges. Machine learning models can be trained to predict the temperature and humidity levels in the cold rooms and to detect anomalies that may indicate a potential problem. This can help improve the monitoring and maintenance of the cabinet and reduce the risk of vaccine spoilage. Machine Learning algorithms like Regression and Classification are used on web-app because they provide us with continuous as well as discrete value as an output. Libraries like pandas, Numpy, sklearn, matplotlib are used to predict along with visualization because of which it will be possible to forecast the actual aspects like temperature and humidity. User Interface has also been developed which acquires input from any user and displays the data according to user's inputs. The project will involve collecting data from sensors installed in cabinet and using machine learning algorithms to analyse the data and generate insights. The results of the project can be used to inform best practices for vaccine storage and contribute to the development of more efficient and effective cabinet management systems.

Keywords: Refrigeration, Thermal Storage, IoT, Machine Learning, HVAC

I. INTRODUCTION

We can predict the future with the help of machine learning. We gather information from routines and utilise it to anticipate the future. Although predictions cannot always be correct, we can acquire a feel of what will happen. The forecast is, nevertheless, more accurate the more data there are. Regression is used to anticipate the rise in temperature and humidity in the box in this instance. [1]. For this, Python libraries are employed. Because machine learning predicts the future and informs us if an anomaly will occur or not, it is a highly effective and potent approach that enables us to take the necessary precautions to avoid any harm.

II. PREDICTIVE MODEL FOR REFRIGERATION SYSTEM

A. Ascertainment of Input Values

Regression is employed in this instance to anticipate the rise in temperature and humidity within the box. Some of the numerous readings that we produced for the model-building process were sorted out to be fed into the linear regression model. It makes predictions about one variable's value based on another's value. The dependent variable is the one you want to be able to forecast. The independent variable is the variable used to forecast the value of the other variable.[2]. he coefficients of a linear equation including one or more independent variables that accurately predict the value of the dependent variable are calculated in this sort of analysis. To directly match a hypothesis between expected and actual output values, linear regression is utilised. It is possible to obtain the best-fit line for a set of paired data using straightforward linear regression calculators that employ the "least squares" approach. Then, using Y (the independent variable), the value of X (the dependent variable) is approximated. When there is power available, the VCRS system used in the setup keeps the cold storage box at the ideal temperature of -5 to 8 degrees Celsius. The Phase Change Materials are maintained at their particular freezing point so that they can be implemented during the power outage condition [3]. The data including the temperature, humidity values, its fluctuation if any, the location of the cold box in terms of latitude and longitude will be collected to the NODEMCU which is nothing but

ESP8266 that acts as Wi-Fi module conveys all the information to the cloud storage for further process. For the best output, constant input is selected as shown in Table I.

All of the null values and missing values are filled in data cleaning and pre-processing.

TABLE I RECORDED INPUT VALUES

	A	B	C	D	E	F	G
1	created_at	entry_id	field1	field2	field3	field4	field5
982	2021-12-21T10:51:04+00:00	981	27.2	54.6	16.68483	74.47332	12.83333
983	2021-12-21T10:51:25+00:00	982	27.3	54.6	16.68483	74.47331	12.83333
984	2021-12-21T10:55:17+00:00	983	27.3	54.5	16.68477	74.47321	12.83333
985	2021-12-21T10:55:38+00:00	984	27.3	54.5	16.68482	74.47331	12.83333
986	2021-12-21T10:55:59+00:00	985	27.3	54.5	16.68483	74.47332	12.83333
987	2021-12-21T10:56:20+00:00	986	27.3	54.4	16.68482	74.47331	12.83333
988	2021-12-21T10:56:41+00:00	987	27.3	54.4	16.68485	74.47327	12.83333
989	2021-12-21T10:57:01+00:00	988	27.3	54.4	16.68485	74.47327	12.83333
990	2021-12-21T10:57:22+00:00	989	27.3	54.4	16.68489	74.47331	12.83333

B. Development of Predictive Model of Vaccine Storage System using Linear Regression

A linear regression model gives us continuous values so that prediction and analysis is comparatively easier. PCM error detection can be determined using fluctuations in temperature and humidity. Initially, two main parameters are set for predictive graph plotting. The dependent variables are set to temperature and humidity and independent variable as Timestamp. In MATLABfor predicting data in real time using linear regression algorithm. By computing model coefficients and presenting the model on top of the data, the MATLAB Basic Fitting UI helps you fit your data. For further details, see Example: Using Basic Fitting UI. In order to fit your data to a linear model with coefficients, you may also utilise the MATLAB polyfit and polyval functions.

Steps:
1. Load data into MATLAB framework. The "load" function is used to get dataset for a file. The data which is pre-processed and cleaned for better model prediction and accuracy.
2. Assigning the data attributes in forms of variables of the model, the data is split into X and Y for training and testing. For training 70% data is given to model and 30% for testing. X is determined as prediction variable and Y as response variable. Here "regress" function is used to build the model.
3. To visualization the graph and hypothesis line, "plot" function is used and to analyse "scatter" function is used.

C. Linear Regression Algorithm

Based on the value of another variable, linear regression analysis forecasts the value of the first variable. The dependent variable is the one you want to be able to forecast. The independent variable is the one you're using to make a prediction about the value of the other variable. The

coefficients of a linear equation including one or more independent variables that accurately predict the value of the dependent variable are calculated in this sort of analysis. The disparities between the expected and actual output values are minimised by using linear regression to match a straight line or surface. There are straightforward linear regression calculators that employ the "least squares" technique to choose the optimal line given a set of paired data. [3]. The value of X (dependent variable) is then estimated from Y (independent variable).

The linear regression method can be used in a variety of programs and environments, including:
1. R linear regression
2. MATLAB linear regression
3. Sklearn linear regression
4. Linear regression Python
5. Excel linear regression

Leaders in businesses and organisations may improve decision-making by using linear regression techniques. Organisations gather enormous quantities of data, and linear regression helps them to use that data, rather than depending on experience and intuition, to better manage reality. It is possible to turn a lot of raw data into useful knowledge. Additionally, linear regression may be utilised to highlight patterns and correlations that your coworkers may have previously seen and assumed they understood, giving you superior insights. Analysing sales and purchase data, for instance, might show you specific buying habits on particular days or at particular times.

Business executives may estimate when there will be a strong demand for their company's products using insights from regression analysis. We have utilised linear regression to forecast data in real-time using MATLAB. By computing model coefficients and presenting the model on top of the data, the MATLAB Basic Fitting UI helps you fit your data [4]. Using the Basic Fitting UI to obtain more details. In order to fit your data to a linear model with coefficients, you

may also utilise the MATLAB polyfit and polyval functions. For a sample, see Programmatic Fitting. The variables should be transformed to make the connection linear if a nonlinear model is needed to match the data. The Optimisation Toolbox TMlsqcurvefit function, the Statistics and Machine Learning Toolbox TMnlinfit function, or functions from the Curve Fitting Toolbox TM can also be used to directly fit a nonlinear function.

This topic describes how to
1. Perform simple linear regression using the operator.
2. Use correlation analysis to determine whether two quantities are related to justify fitting the data.
3. Fit a linear model to the data.
4. Evaluate the goodness of fit by plotting residuals and looking for patterns.
5. Calculate measures of goodness of fit R^2 and adjusted R^2
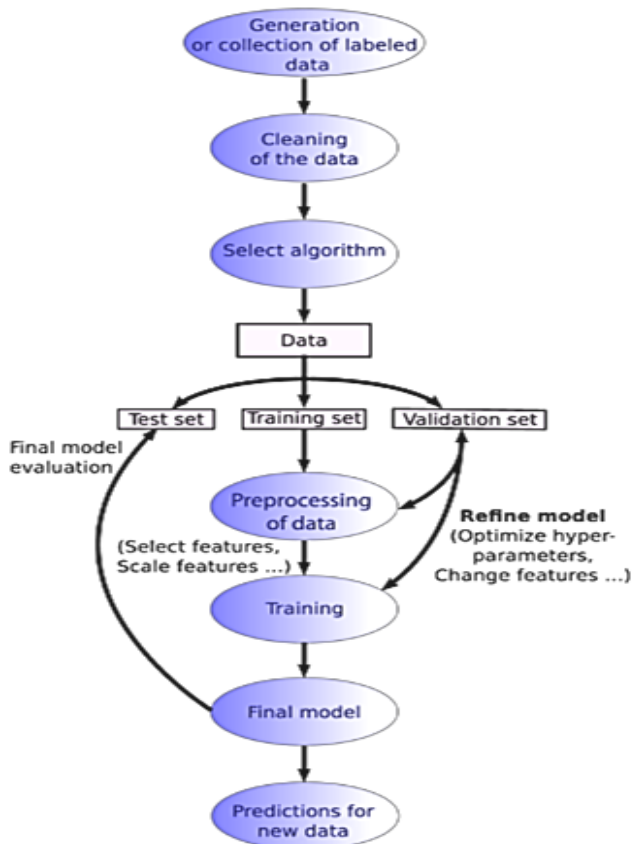
D. Steps Involved in Machine Learning

Fig. 1 Flow Chart for Machine Learning Workflow

1. Data Collection: When there is electricity available, the VCRS system used in the setup keeps the cold box at the ideal temperature of -5 to 8 degrees Celsius. The Phase Change Materials are kept at their specific freezing point so they may be used in the event of a power loss. [5]The data including the temperature, humidity values, its fluctuation if any, the location of the cold box in terms of latitude and longitude will be collected to the NODEMCU which is nothing but ESP8266 that acts as Wi-Fi module conveys all the information to the cloud storage for further process.

2. Cleaning and Pre-Processing of Data: Missing data, Null values, duplicate values were updated/removed for better precision and accuracy of prediction.

3. Choosing an Algorithm: A statistical method called linear regression may be used to forecast the value of a variable based on the value of another variable. Thus, the optimal fit for the linear regression procedure is to provide continuous values of projected data. You may use regression analysis to gauge how strongly two variables are related. Regression analysis can tell you how much of the overall variability in the data is explained by your model using statistical metrics like R-squared / adjusted R-squared. It makes the modelling process easier and yields more comprehensible results.

Because of its strong statistical base and mathematical evidence of performance guarantees, linear discriminant analysis is a preferable choice when you need to make predictions on fresh data points (i.e., the test set).In consequence, linear regression algorithm is most suitable and relevant for analysing and predicting two main parameters that are Temperature and Humidity for vaccine storage.

4. Training the Model: With the help of linear regression algorithm, we have trained our model. This is where we've used all the data. Two splits as Testing data and Training data were created, The first dataset you use for teaching a machine learning programme to recognise patterns or perform to your criteria is called training data, whereas the testing or validation dataset is used to assess the correctness of your model [6].

5. Check if the Model Works: Furthermore, after model was trained we have checked accuracy and precision of the model, which are further discussed in the report.

6. Deploying is the Last Stage: Over the succeeding results model is deployed on web-app accompanied by other aspects [7].

III. CODE FOR PREDICTION OF PARAMETERS USING MATLAB AND LINEAR REGRESSION ALGORITHM

A. Data Analysis

In this code, load is used to load your data from a text file / CSV file into MATLAB. X represents your predictor variables, and y represents your response variable. The stepwise function is used to perform stepwise linear regression, and the resulting model is stored in mdl [8-9]. The predict function is used to predict new data using the model, with the new predictor variables stored in newX. Finally, yPred represents the predicted response variable.

Avesahemad S. N. Husainy, Bhushan S. Kumbhar, Prajwal S. Chavan, Ajeem A. Attar and Hemant A. Patil

1. Prediction of Temperature

Name

predicted temperature

MATLAB Code

```matlab
1  % Load your data into MATLAB
2  data = load('feeds.csv');
3
4  % Separate the predictor variables from the response variable
5  X = data("field1");
6  y = data("created_at");
7
8  % Perform stepwise linear regression
9  mdl = stepwiselm(X, y);
10
11 % Display the results
12 disp(mdl);
13
14 % Predict new data using the model
15 newX = [1 2 3 4]; % example new predictor variables
16 yPred = predict(mdl, newX);
17
18 % Display the predicted response variable
19 disp(yPred);
```

Save and Run   Save

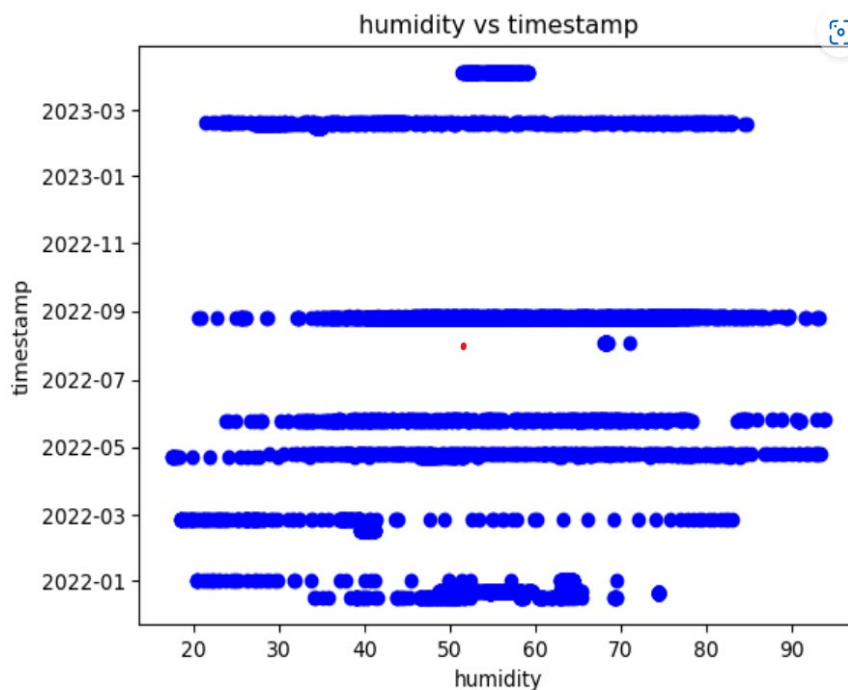Fig. 2 Prediction of Temperature Input

Output:

Fig. 3 Graph of humidity vs timestamp

2. Prediction of Humidity

Name

predicted humidity

MATLAB Code

```
1  % Load your data into MATLAB
2  data = load('feeds.csv');
3
4  % Separate the predictor variables from the response variable
5  X = data("field2");
6  y = data("created_at");
7
8  % Perform stepwise linear regression
9  mdl = stepwiselm(X, y);
10
11 % Display the results
12 disp(mdl);
13
14 % Predict new data using the model
15 newX = [1 2 3 4]; % example new predictor variables
16 yPred = predict(mdl, newX);
17
18 % Display the predicted response variable
19 disp(yPred);
```

Save and Run   Save

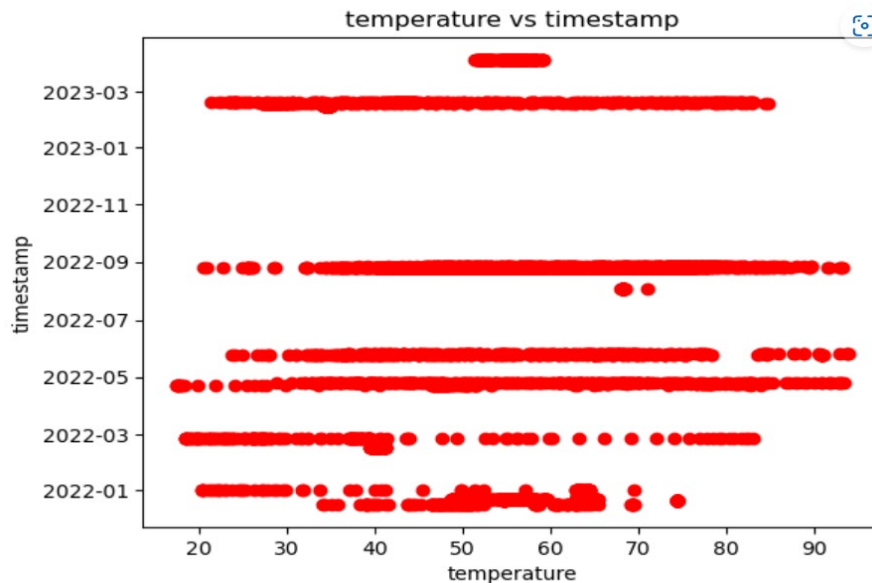Fig. 4 Prediction of Humidity

Output

Fig. 5 Graph of temperature vs timestamp

B. Data Visualization

In this code, we first load the data and perform stepwise linear regression as in the previous example. We then create a scatter plot of the data using the first predictor variable (X(:,1)) and the response variable (y).[7] We add a regression line to the plot by generating a range of x-values (xRange) and computing the predicted y-values (yRange) using the predict function with the mdl model object. We then plot the regression line on top of the scatter plot using a solid red line with a width of 2 ('r-', 'LineWidth', 2). Finally, we add axis labels and a title to the plot, turn on the legend (which automatically labels the scatter points and the regression line), and turn on the grid.

1. Visualization of Temperature Parameter

```
In [3]: import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        plt.rcParams['figure.figsize'] = (7.0, 5.0)
```

```
In [4]: df = pd.read_csv('vaccdataset.csv')
```

```
In [20]: df.describe()
```
Out[20]:

	42.19	35.06	-3.13	33.06	28.9	1
count	253.000000	253.000000	253.000000	253.000000	253.000000	253.000000
mean	39.591344	34.972688	2.299170	24.794625	14.225296	31.726877
std	1.677435	0.889779	7.802006	4.897018	4.849242	13.285847
min	36.000000	33.560000	-18.190000	12.060000	7.200000	1.000000
25%	38.310000	34.060000	-3.690000	22.250000	8.800000	29.400000

```
In [6]: X = np.array(df.iloc[:100,1]).reshape(-1, 1)
        y = np.array(df.iloc[:100,5]).reshape(-1,1)
```

```
In [7]: from sklearn.model_selection import train_test_split
```

```
In [8]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=34)
```

```
In [9]: from sklearn.linear_model import LinearRegression
```

```
In [10]: lm = LinearRegression()
```

```
In [11]: lm.fit(X_train,y_train)
```
Out[11]: LinearRegression()

```
In [12]: predictions = lm.predict(X_test)
```

```
In [13]: print(predictions)
        [[19.20693987]
         [19.15059865]
         [19.06139172]
         [19.32431741]
         [17.361765  ]
         [18.56371097]
         [18.68108851]
         [19.17876926]
         [19.20693987]
         [19.00505051]
         [19.47456065]
         [18.94401419]
         [18.9721848 ]
         [16.94859607]
         [19.23980558]
         [18.26791958]
         [18.03316451]
         [18.06603022]
         [17.85944575]
         [19.2961468 ]]
```

```
In [15]: plt.scatter(X, y)
        plt.plot([min(X), max(X)], [min(predictions), max(predictions)], color='green')
        plt.show()
```

```
In [16]: print(lm.intercept_)

         [-0.18852425]
```

```
In [17]: print(lm.coef_)

         [[0.46951015]]
```

```
In [18]: from sklearn import metrics
```

```
In [19]: print('MAE:', metrics.mean_absolute_error(y_test, predictions))
         print('MSE:', metrics.mean_squared_error(y_test, predictions))
         print('RMSE:', np.sqrt(metrics.mean_squared_error(y_test, predictions)))

         MAE: 2.2571671059916207
         MSE: 6.922541087557178
         RMSE: 2.6310722315354966
```
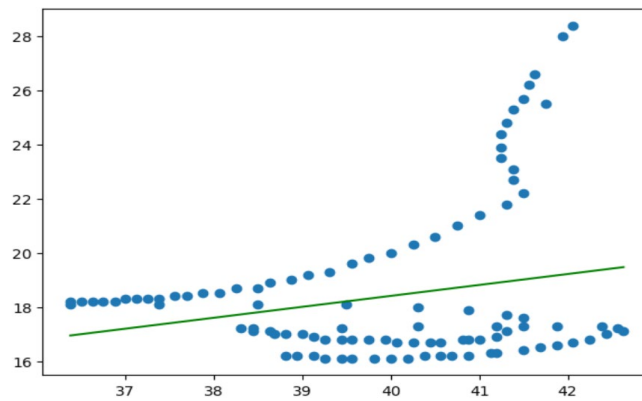
Fig. 6 Visualization of Temperature Parameter

Output

Fig. 7 Graph of timestamp vs aspect

C. Experiment Setup

Fig. 8 Experimental Set Up

Avesahemad S. N. Husainy, Bhushan S. Kumbhar, Prajwal S. Chavan, Ajeem A. Attar and Hemant A. Patil

IV. RESULTS AND DISCUSSION

The accuracy of results demonstrated that the linear regression model successfully captured the relationship between temperature/humidity and Timestamp, allowing for accurate predictions of PCM error detection. Real-time monitoring and analysis of temperature and humidity fluctuations using this model can provide valuable insights into potential errors or deviations in the PCM system. The predictive graph plotted using the linear regression model can serve as a visual representation of the predicted data. This graph can aid in identifying patterns, trends, and potential anomalies, allowing for proactive measures to be taken to address PCM errors promptly.

V. CONCLUSION

The system used for keeping and distributing vaccines in good condition at the recommended temperature range (-5 to +8°C). Also, a technology that can transport larger volume of vaccines with is also needed as existing cool boxes can transport a limited number of vaccines at a time the temperature, humidity and other parameters should be continuously monitored during its transportation as it may affect the vaccine life drastically. Proposed experimental set up will easily transport because it is portable and will be available in low cost. Due to thermal storage, it will help to maintain constant temperature (2 to 8 degree) for longer duration of time even though power outage exists. It will helpful for doctors, pharmacist, medical stores, hospitals to store vaccine and medicines. We use machine learning to track and forecast variables like temperature and humidity while displaying data on a specified smartphone, tablet, laptop, or PC. The performance is tested with machine learning algorithms while condition of power cut off and power on. The following are the parameters on basis of which the performance of the project is tested Prediction of future temperature after certain time period. Prediction and analysis of humidity. Their data can go through ML

algorithms which will predict the data so user can be aware of the probability of an outcome thus, we can create an opportunity to override any circumstances before any damage happens. So that the system can safely store and transport the product to the recipient.

REFERENCES

[1] H. Hu, J. Xu, M. Liu and M. K. Lim, "Vaccine supply chain management: An intelligent system utilizing blockchain, IoT and machine learning," *Journal of Business Research,* Vol. 156, pp. 113480, 2023.

[2] R. Jain and S. K. Gawre, "Monitoring and Control of COVID Vaccine Storage Temperature Using IoT and Machine Learning," In *2022 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), IEEE,* pp. 1-6, February 2022.

[3] X. Tang, Q. Guo, M. Li, C. Wei, Z. Pan, and Y. Wang, "Performance analysis on liquid-cooled battery thermal management for electric vehicles based on machine learning," *Journal of Power Sources,* Vol. 494, pp. 229727, 2021.

[4] Y. Zhou, S. Zheng, Z. Liu, T. Wen, Z. Ding, J. Yan and G. Zhang, "Passive and active phase change materials integrated building energy systems with advanced machine-learning based climate-adaptive designs, intelligent operations, uncertainty-based analysis and optimisations: A state-of-the-art review," *Renewable and Sustainable Energy Reviews,* Vol. 130, pp. 109889, 2020.

[5] Y. Li, Y. Wen, D. Tao and K. Guan, "Transforming cooling optimization for green data center via deep reinforcement learning," *IEEE transactions on cybernetics,* Vol. 50, No. 5, pp. 2002-2013, 2019.

[6] C. Johansson, M. Bergkvist, D. Geysen, O. De Somer, N. Lavesson and D. Vanhoudt, "Operational demand forecasting in district heating systems using ensembles of online machine learning algorithms," *Energy Procedia,* Vol. 116, pp. 208-216, 2017.

[7] W. Zhang, Y. Wen, Y. W. Wong, K. C. Toh and C. H. Chen, "Towards joint optimization over ICT and cooling systems in data centre: A survey," *IEEE Communications Surveys & Tutorials,* Vol. 18, No. 3, pp.1596-1616, 2016.

[8] Y. Guo, Z. Tan, H. Chen, G. Li, J. Wang, R. Huang, J. Liu and T. Ahmad, "Deep learning-based fault diagnosis of variable refrigerant flow air-conditioning system for building energy saving," *Applied Energy,* Vol. 225, pp. 732-745, 2018.

[9] S. Liu and G. P. Henze, "Evaluation of reinforcement learning for optimal control of building active and passive thermal storage inventory," 2007.