2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON) | 978-1-6654-0382-5/20/$31.00 ©2021 IEEE | DOI: 10.1109/ECTI-CON51831.2021.9454878

Forecasting the Quantity and Concentration of Flocculant in Clarification Process for Sugarcane Industry

Singhadid Chantaruk
Computer Engineering
Chulalongkorn University
Bangkok, Thailand
6170296721@student.chula.ac.th

Diew Koolpiruck
Control System and Instrumentation Engineering
King Mongkut's University of Technology Thonburi
Bangkok, Thailand
diew.koo@kmutt.ac.th

Prabhas Chongstitvatana
Computer Engineering
Chulalongkorn University
Bangkok, Thailand
prabhas.c@chula.ac.th

***Abstract*— The clarification process is an important part of sugarcane production. This process is used for separating sediment and sugarcane juice by adding flocculant. The addition of quantity and concentration of flocculant directly affects the settling rate and turbidity of sugarcane juice. This paper proposes a model for forecasting quantity and concentration of flocculant by using Long Short-Term Memory (LSTM) Neural Network. Input data consists of green cane, burn cane, turbidity, and rainfall. Output data includes quantity and concentration of flocculant. Raw data was collected from top sugarcane factory and meteorological department in Thailand. The results are the forecast of the quantity and concentration of flocculant for one day in advance. The performance of LSTM is compared to the autoregressive integrated moving average (ARIMA), recurrent neural network (RNN), and gated recurrent unit (GRU) using root mean square error and mean absolute percent error. The result indicates that LSTM has the best performance. The forecast helps the operator in clarification process to prepare the flocculant.**

Keywords*— *Forecasting, Flocculant, LSTM neural network, Clarification process, Sugarcane industry

I. Introduction

The clarification process is important part to extract dirty particles in sugarcane juice by using flocculant to interact with particle and fall as sediment [1] thus, layers between sediment and sugarcane juice is occurred. The adding of flocculant part is controlled by settling rate. The settling rate measure from sampling sugarcane juice in clarified tank and adjust flocculant amount. The operator will not know the optimization of flocculant amount, thus factory will prepare extra amount in every day. Generally, factory estimates quantity of flocculant from amount of sugarcane compare with last usage flocculant. This research shows the forecasting of quantity and concentration of flocculant to improve the efficiency of clarification process in sugarcane extraction.

To forecast time series data, a common traditional statistical method is widely used, called the Autoregressive Integrated Moving Average (ARIMA) [2]. Moreover, for the problem which has complex data Artificial Neural Network (ANN) is better than ARIMA in prediction accuracy [3]. Recurrent Neural Network (RNN) model collects activations of each time step in internal state (temporary memory) to process sequence of input. However, RNN model has problem of vanishing gradient [4]. To solve this problem, Long-Short Term Memory (LSTM) model was introduced for processing long interval time series datasets [5].

There are forecasting models which apply for clarification process such as Neural network and Heuristic Dynamic Programming (HDP) [6]. This research combined the concepts of dynamic programming and reinforcement learning to optimize and control the neutralized pH value for sugar clarifying section. Karthik C. [7] studied the control neutralizing pH value by using neural combined with reinforcement learning to optimize and control the neutralized pH value for sugar clarifying. Shaojing Song [8] studied the forecasting of purified juice color value and alkalinity using genetic dynamic fuzzy neural network (GDFNN). They compared it with back propagation (BP) network and found that GDFNN is more accurate and is suitable for production.

This research studies LSTM which is a popular forecasting method for time series data. There are previous works that compared LSTM model with ARIMA, FFNN, RNN, etc. The result shows that LSTM has higher accuracy than other models [9]. Therefore, this research selects LSTM model for forecasting the quantity and concentration of flocculant to help factory preparing suitable quantity and concentration of flocculant in the clarification process.

II. Theory

A. Clarification Process

The main function of clarification process is to separate dirtiness from sugarcane juice. There are three of main procedure; temperature control, PH control and separating dirtiness, showing in Fig. 1. In this paper focus on the separating dirtiness from sugarcane juice which is composed of soil, sand and the aggregate from burning sugarcane. The isolation starts with adding flocculant into the clarified tank in order to make the chemical reaction called sedimentation. Flocculant is function of forming the particles to become mud at the bottom causing separate between mud and sugarcane juice. Turbidity of

978-0-7381-1127-8/21/$31.00 ©2021 IEEE

juice depends on settling rate, The best estimate of settling rate was that it varied between about 15 and 40 cm/mins [10].

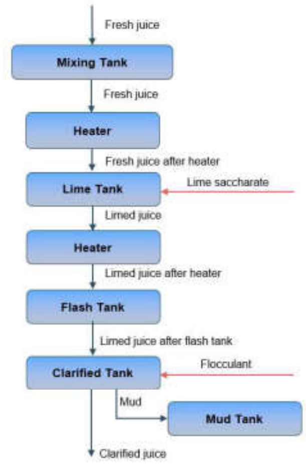

Fig. 1. Clarification Process

B. Inverse Distance Weighting Method

Inverse Distance Weighting (IDW) is one of the popular ways to interpolate value. The method of IDW can be calculated with a weighted average of the values available at the known points, which can be given as follows:

$$w_i = \frac{1}{d_i} \tag{1}$$

Where, w_i is weight value at the i location

d_i is distance value at the i location

When we find the weight in each position, it will be interpolated value for the right position, which can be given as follows:

$$P_x = \frac{\sum_{i=1}^{n} w_i P_i}{\sum_{i=1}^{n} w_i} \tag{2}$$

Where, w_i is weight value at the i location

P_i is measured value at the i location

P_x is interpolation value at the i location

C. Long Short-Term Memory (LSTM) Neural Network

The LSTM is developed from Recurrent Neural Network (RNN) by Hochreiter and Schmidhuber [5]. It composes of memory cell, input gate, output gate and forget gate, shown in formular as below.

Forget gate

$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f) \tag{3}$$

Current state

$$i_t = \sigma(w_f[h_{t-1}, x_t] + b_i) \tag{4}$$

$$c_t = tanh(w_c[h_{t-1}, x_t] + b_c) \tag{5}$$

Remember cell

$$c_t = (f_t * c_{t-1}) + (i_t * c_t) \tag{6}$$

Output gate

$$o_t = \sigma(w_o[h_{t-1}, x_t] + b_0) \tag{7}$$

$$h_t = o_t * \tanh(c_t) \tag{8}$$

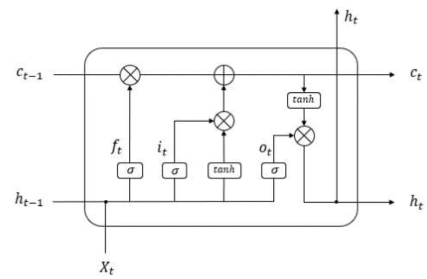

Fig. 2. Long Short-Term Memory (LSTM) structure

D. Reference model

Forecasting models for time series data have various models such as traditional statistical method (ARIMA), recurrent neural network (RNN), and gated recurrent unit (GRU). The forecasting models that employ in the paper are of:

1. The ARIMA Model

The Autoregression Integrated Moving Average (ARIMA) algorithm represents a statistical based. ARIMA algorithm have to adjust (p, d, q) parameters by using grid search. For each combination of these parameters for the new ARIMA model, we use Akaike Information Criterion (AIC) value to choose the best combination of these parameters [11].

2. The RNN Model

The Recurrent Neural Network (RNN) model is one of the recursive neural networks approaches that can be applied for modeling sequential data. However, its capability to process short-term sequential data, the weakness of RNN is carried out when learning long-range dependencies is demanded in time series forecasting applications [12].

3. The GRU Model

The Gated Recurrent Unit (GRU) model is one of the recurrent neural networks. Its model is similar to the LSTM model with the exception that GRU includes only two gates rather than three [13]

III. Method

A. Data Preprocessing & Preparation

Data are collected from the production line of a sugarcane factory in Thailand. The first part is inside factory information including green cane, burn cane and turbidity. The second part is outside factory information including the rainfall from Meteorological Department in Thailand. However, it has to be preprocessed before use it to train the model. The processes are as follow.

Step1: Handle missing value imputation

Handle missing value imputation in form of time series data by using linear interpolation for managing the missing value imputation. Data that has missing value is turbidity by human error when collecting data.

Step2: Rainfall interpolation

Since amount of rainfall from rainfall station, Meteorological Department about 68 station in 38 provinces found that the rainfall stations are not covered for farming of sugarcane area directly in Fig. 3. There for the data of rainfall is not available, it has to be interpolation for finding amount rainfall in covered area.

Interpolation method applied for finding amount of rainfall in the areas by using IDW method section 2. It starts with drawing the table 10 kilometers distance in polygon in Fig. 3. Then, bring the rainfall interpolation value in each covered area and finding the average of rainfall.

Fig. 3 Polygon 10 km

Step3: weight average rainfall

Both amount of sugarcane received from customer and amount of rainfall in each area are different. Therefore, we cannot use amount of rainfall from rainfall interpolation value directly. It has to find weight average value of rainfall, which can be given as follows:

$$R_x = \frac{\sum_{i=1}^{n} w_i r_i}{\sum_{i=1}^{n} w_i} \tag{9}$$

Where, R_x is weight average rainfall at the i location

w_i is weight of cane at the i location

r_i is average rainfall at the i location

Step4: Data Normalization

The raw data before training the models have to be normalized data in order to reduce bias data by using min-max normalization. It is the popular way for normalizing data and period of data should be 0 to 1.

$$z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \tag{10}$$

Where, z_i is data normalization

x_i is data at i

B. Forecating model

The input data of proposed model has 4 data and output also has to 2 data that showing distribution of input variables as shown in table I and II. All the input data will be normalized and sampling rate of collecting data in every day. The number of datasets is 396 samples which is historical data about December to April in the past 4 years of each sugarcane production. The datasets were divided for training 70% of data set (276 samples) and testing 30% of data set (120 samples).

TABLE I. DISTRIBUTION OF INPUT PARAMETERS

Inputs	**Input Parameters**		
	Parameter	***Mean***	***σ***
1	Green cane	7292.527	2516.887
2	Burn cane	8190.039	2615.319
3	Turbidity	5.870	1.053
4	Rainfall	0.570	1.995

TABLE II. DISTRIBUTION OF OUTPUT PARAMETERS

Inputs	**Output Parameters**		
	Parameter	***Mean***	***σ***
1	Quantity of Flocculant	60.881	4.0104
2	Concentration of Flocculant	16.231	0.943

The main structure is LSTM cell and hidden layer shown in Fig. 5. TensorFlow library was used to create LSTM model. The model has adjusted hyper-parameters including the number of hidden, epochs and lag time. The hyper-parameters compose of the lag time in period of 1 to 2 samples referred from amount of rainfall data that affect to amount of mud, dropout of training is 0.2. Adam optimizer is applied for optimization because it is the best optimizer for deep learning [14]. The epoch of training is 300 epochs depend on loss of training in Fig. 4.

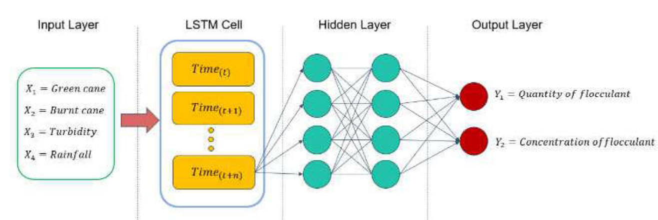

Fig. 4 LSTM Model structure

IV. RESULT & DISCUSSION

In this section shows the results of forecasting quantity and concentration of flocculant in one day advanced. The paper compares performance of forecasting model including traditional statistical method (ARIMA), RNN, GRU and proposed model (LSTM). The results are shown in Table III, IV, V and VI.

For ARIMA, (p, d, q) parameters of the best performance is (2, 1, 1). For Recurrent Neural Network (RNN) the best results for forecasting quantity and concentration are 8 and 32 of hidden units. Gated Recurrent Unit (GRU) uses 8 hidden units. The proposed LSTM uses (16, 32) and 8 of hidden units.

TABLE III. BEST REULTS OF ARIMA

(p, d, q)	Quantity of Flocculant		Concentration of Flocculant	
	RMSE	*MAPE*	*RMSE*	*MAPE*
(1, 1, 1)	13.616	24.771	1.061	13.384
(2, 1, 0)	13.413	24.724	1.014	13.030
(2, 1, 1)	12.791	24.652	0.922	11.948

TABLE IV. BEST REULTS OF RNN

No. of layer	No. of hidden units	Lag time	Quantity of Flocculant		Concentration of Flocculant	
			RMSE	*MAPE*	*RMSE*	*MAPE*
1	[8]	1	7.448	14.391	0.630	9.984
1	[16]	1	6.577	10.943	0.632	10.626
1	[32]	1	6.474	11.842	0.635	11.465
2	[8,16]	2	7.414	13.819	0.665	10.571
2	[16,32]	1	6.965	13.029	0.606	10.340

TABLE V. BEST REULTS OF GRU

No. of layer	No. of hidden units	Lag time	Quantity of Flocculant		Concentration of Flocculant	
			RMSE	*MAPE*	*RMSE*	*MAPE*
1	[8]	1	5.874	10.567	0.643	10.524
1	[32]	1	6.102	11.039	0.632	10.626
1	[64]	2	6.293	11.142	0.729	11.620
2	[16,32]	1	6.101	10.586	0.755	11.331
2	[32,16]	1	6.070	10.562	0.940	12.240

TABLE VI. BEST REULTS OF LSTM

No. of layer	No. of hidden units	Lag time	Quantity of Flocculant		Concentration of Flocculant	
			RMSE	*MAPE*	*RMSE*	*MAPE*
1	[8]	1	6.059	10.775	0.617	9.648
1	[16]	2	6.975	11.854	0.635	10.863
1	[32]	1	5.961	10.935	0.668	11.091
2	[8,16]	1	6.667	11.431	0.670	11.213
2	[32,16]	1	5.730	9.912	0.841	12.007

TABLE VII. OVERALL COMPARISON MODEL (BEST RESULT)

Model	Quantity of Flocculant		Concentration of Flocculant	
	RMSE	*MAPE*	*RMSE*	*MAPE*
ARIMA	12.791	24.652	0.922	11.948
RNN	6.474	11.842	0.630	9.984
GRU	5.874	10.567	0.643	10.524
LSTM	5.730	9.912	0.628	9.718

The results show that the proposed model (LSTM) has the best performance in term of forecasting quantity and concentration of flocculant. LSTM has 5.7% RMSE of forecasting quantity and 0.6% RMSE for concentration.

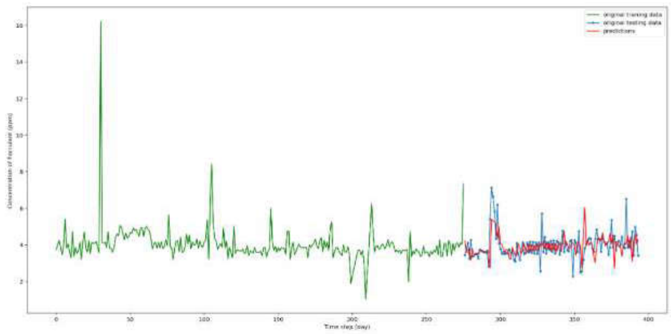

Fig. 6 One day ahead quantity of flocculant by LSTM model

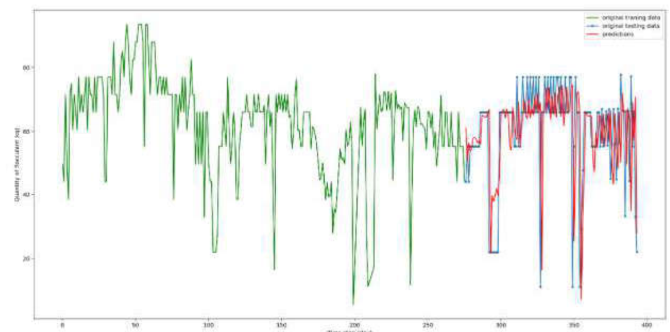

Fig. 7 One day ahead concentration of flocculant by LSTM model

V. CONCLUSION

This paper proposed LSTM model for forecasting the quantity and concentration of flocculant to solve flocculant preparation problem in clarification process. This benefit is to operator for preparation of flocculant efficiently. The results show that the proposed model (LSTM) is the best performer in term of forecasting quantity and concentration of flocculant compared to other models. Fig. 6 and 7 show the actual forecasting graph. It can be observed that the error (in red) is very small. This forecasting is suitable to be used in a real production.

REFERENCES

[1] W. O. S. Doherty, C. M., S. Gorjian, E. Senogles, W. H. Cheung, "Flocculation and sedimentation of cane sugar juice particles with catiomic homo- and copolymers," Journal of Applied Polymer Science 90 (1), 2003, pp. 316-325.

[2] S.L. Ho and M. Xie , The use of ARIMA models for reliability forecasting and analy- sis, Comput. Ind. Eng. 35, 1998, pp. 213–216.

[3] G. Zhang , B.E. Patuwo and M.Y. Hu , Forecasting with artificial neural networks: the state of the art, Int. J. Forecast. 14, 1998, pp. 35–62.

[4] Pascanu , Razvan , T. Mikolov , Y. Bengio , On the difficulty of training recurrent neural networks, in: Proceedings of the 30th International Conference on Ma- chine Learning (3), volume 28, 2013, pp. 1310–1318.

[5] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation 9 (8), 1997, pp. 1735-1780.

[6] X. Lin, S. Leu and H. Liu, "Neural Netowrk Modeling and HDP for Neutralized pH Value Control in the Clarifying Process of Sugar Cane Juice," World Congress on Engineering and Computer Science, 2008.

[7] K. C*, V. K. and P. R., "Modelling and Control of Chemical Process in Sugar Industry," International Journal of Cumputer Applications, 2011, pp. 24-28.

[8] S. Song, J. Wu, X. Lin and H. Liu, "Predictive Model of Production Index for Sugar Clarification Process by GDFNN," Advances in Neural Networks – ISNN, 2012, pp. 585-593.

[9] A. Sagheer and M. Kotb, "Time series forecasting of petroleum production using deep LSTM recurrent networks," Neurocomputing, 2019, pp. 203-213.

[10] R.J. Steindl, "Calrification of Cane Juice for Fermentation," International Society of Sugar Cane Technologists, 2010.

[11] E.J. Bedrick and C.L. Tsai, "Model selection for multivariate regression in small samples," Biometrics 50, 1994, pp. 226-231.

[12] J. Bayer, "Learning Sequence Representations," Technische Universitt Mnchen, 2015.

[13] J. Chung, C. Gulcehre, K. Cho and Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," Deep Learning and Representation Learning Workshop at NIPS, 2014.

[14] D. P. Kingma and J. L. Ba, "Adam : A Method for Stochastic Optimization," 3rd International Conference for Learning Representations, 2015.