Optimization Process of Glycerol Esterification Using Real Time Adaptive Control

Iwan Aang Soenandi
Post Graduate Program of Agro-industrial Technology
Bogor Agricultural University Dramaga, Indonesia
Lecturer, Faculty of Engineering and Computer Science
Krida Wacana Christian University
iwansoenandi@gmail.com

Ani Suryani
Post Graduate Program of Agro-industrial Technology
Bogor Agricultural University. Dramaga, Indonesia
anisuryani.sbrcipb@gmail.com

Taufik Djatna
Post Graduate Program of Agro-industrial Technology and Graduate Program of Computer Science,
Bogor Agricultural University. Dramaga, Indonesia
taufikdjatna@ipb.ac.id

Irzaman
Department of Physics
Faculty Mathematics and Natural Sciences
Bogor Agricultural University. Dramaga, Indonesia
irzaman@ipb.ac.id

Abstract— **The synthesis reaction used in esterification needs high energy consumption and a precise processing time to get the best yield of target. In this study, a model was formulated to optimize glycerol esterification process by minimizing the time needed for the process and maximizing the yield of Mono-glycerides. This optimization has gained importance for boosting the esterification industry and improving the production efficiency. Optimization through adaptive monitoring and control has provided significant advances in the process efficiency, a lower energy consumption and a better product quality. This paper presents the optimization with a computational algorithm in real time and adaptive control (RTAC), as compared to the conventional (traditional) methods to monitor and control of glycerol esterification processes. The identification of esterification status based on temperature and time are evaluated to strengthen the optimization. An adaptive method as feature selection to select wavelength IR sensors at specified intervals was carried out with Relief algorithm and Adaptive Pillar K-means clustering method to set the parameter control was proposed in this paper. Many combinations were evaluated from real time condition process, to achieve the best optimization results. The experimental results demonstrate that real time adaptive control can be achieved by using three clusters, which are heating up, stabilizing and finishing. In RTAC, each cluster has its own parameter to set the control point by the servo motor that was attached to magnetic stirrer-heater. By using optimization parameter for each cluster, esterification process time can be shortened by 15-20 minutes with a higher yield (7% or more), lower range stirrer rotation (300rpm-450 rpm) and a lower final temperature of 200^{0}C - 210^{0}C.**

Keywords— ***Optimization; Real Time; Adaptive Control; Glycerol Esterification***

I. INTRODUCTION

Currently mono-glycerides and di-glycerides are important in processes involving emulsification. One of the key challenges in the process industry is to find the best operation method for the plant under different conditions such as feed compositions, production rates, energy availability, feed and product compositions that change dynamically like in esterification process [1]. In industrial chemical process synthesis by direct esterification is widely used to get mono-glycerides and di-glycerides because this process is simple and feasible in batch production system [2]. The task for optimizing process is usually tackled using a supervisory control technique, monitored manually to check the parameter of the process. One such technique that has received considerable attention in the process industry is the Real-Time Optimization (RTO) [3]. Real-Time Optimization (RTO), which refers to the online economic optimization of a process plant, is a widely employed technology to meet this challenge. The limitation of RTO is that it is not adaptive to the type of the raw material (feed composition) that affected reaction process of esterification. In this paper we address this issue by combining RTO with an Adaptive Controller (AC) for glycerol esterification reaction consisting a real time identifier and a minimum variance regulator for the identification, done by augmenting the state with the unknown parameters of the process was discussed. This method was previously proposed by Wieslander and Wittenmark [4] and state the main problems in adaptive control are to know the present system parameters and to choose the appropriate control strategy. Early application of adaptive control was presented by Astrom 1989 [5] and used in the control of a permanent synchronous motor for digital adaptive velocity to maintain invariant velocity control over the motor in the presence of varying mechanical parameters.

There are several chemical process and reaction to produced mono glycerides. We are focus on the chemical reaction process in selective synthesis of mono-glycerides by esterification of glycerol with fatty acids [6], which has complex reaction because of the immiscibility of reagents and the formation of di- and tri-glyceride by-products. Because of the process esterification characteristics that requires time and energy for heating the raw material, this process will be optimized by reducing the process time (minimization) and

---

This work was supported by the financial support of Kementerian Pendidikan dan Kebudayaan Republik Indonesia (DIKTI) with contract number 095/K3/KM/2015

maximizing the yield of product by monitoring and controlling[7]. Process monitoring is the manipulation of sensor measurements (e.g. force, vision, temperature, rate of transmittance) needed for determining the state or condition of the processes. Automatic monitoring algorithms utilize selected sensor measurements that, along with inputs, determine the process state. The states of complex processes are monitored by a sophisticated signal processing of sensor measurements. Process control is the manipulation of process variables to regulate the processes [8]. In traditional methods, the operators perform on-line and off-line process control by adjusting the temperature button and always giving attention to the rotational speeds of stirrer reactor to suppress over-temperature, and watching the time needed for the process with a traditional stopwatch.

The objectives of the study presented in this paper are: to select the sensors that are related to esterification status for the wavelength, to find the optimal number of data cluster for parameter control and to evaluate the system performance using a new proposed method. We describe Real Time Adaptive Control (RTAC) as a method of optimizing the number of cluster from real time data streaming using sensors, during the process of esterification in bioreactor and clustered it to determine the set point parameter of temperature and rotational speed of stirrer. To achieve these tasks simultaneously, RTAC system is supported by real time optimization methods using Relief (Recursive Elimination of Features) from Kira & Rendell (1992) and Pillar K-means algorithm introduced by Barabakh (2010), combined with real time data acquisition from optical mid IR sensors to set the best parameter control for the esterification process.

II. DESCRIPTION OF PROBLEM

From motivation, explained at Section I. To optimize esterification process, the main challenge is how to determine the time needed for the process to get the best yield for the product. By using sensors to identify the composition of product mixed in mono-glycerides and di-glycerides. We state our optimization problems using parameters of temperature in and rotational stirrer speed to optimize process time and yield of product. To simplify the problem, we assumed the purity of reactant, catalyst and product concentrations as the states variables. In this research, the optimization objectives are to minimize process time and maximize the yield of mono-glycerides as target of product. In mathematical formulations, they are expressed in (1-5):

$$\text{Minimize } t\,(T, R) \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \ (1)$$

$$\text{Maximize } y(T,C,R) \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$\text{Subject to :}$$

$$120\ {}^{0}\text{C} \leq T \leq 230\ {}^{0}\text{C} \ \ldots\ldots\ldots\ldots\ldots\ldots \ (3)$$

$$200 \leq R \leq 500 \text{ RPM} \ \ldots\ldots\ldots\ldots\ldots\ldots \ (4)$$

$$S \equiv \text{Status } \{\text{Sensor 1, Sensor n, ...}\} \ldots\ldots\ldots \ (5)$$

Where:
t= Esterification process time (minutes);
y=Yield of Mono-glycerides (%);
T= Temperature bioreactor (^{0}C) ;
R= Rotational speed of stirrer (RPM);
S= Status of Esterification (Clustering);

for status of esterification we get from combined the output signal from the sensor and detector and used adaptive clustering algorithm.

In RTAC there are many tasks to do, from preparing the sensors and data acquisition system to processing real time data, to finding the best parameter for the process that is being affected from the condition of the external disturbance like: reactor’s surrounding temperature, impurity of raw material and instability of the electric current. On the other side, noise problem from the input sensor is also crucial. Calibration is also important because of the variability condition from the bioreactor and the sensitivity of the sensors used for setting parameter of the controller. The scope of high performance computing is rapidly expanding from single parallel systems to clusters of heterogeneous sequential and parallel systems. Moreover, as applications become more complex, they grow more irregular, with a data-dependent execution behavior, and more dynamic, with time-varying resource demands.

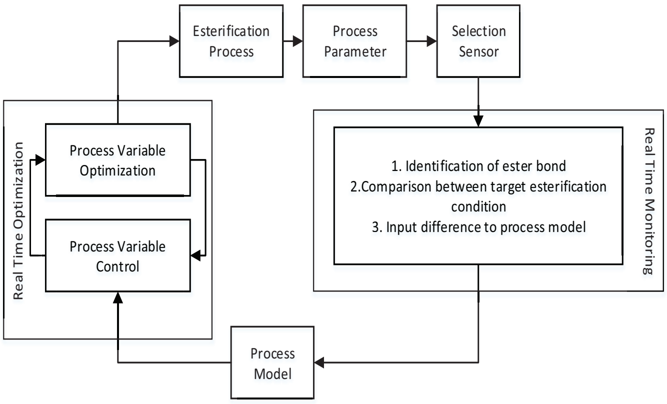

Fig. 1. Flow diagram adaptive optimization process design

In this research, we want to solve those optimization problems with an algorithm to adapt to concept changes by clustering. Real-time optimization is the process of extracting interesting and previously unknown or unexpected knowledge from large amount of real time data (Fig 1). Its correctness depends on not only to logical correctness but also time constraints. Various methods in the area of data mining have been developed in order to effectively and correctly discover knowledge.

III. METHODS

This section describes the procedure and step from monitoring the reaction in reactor to get the solution for optimization in real time esterification process by controlled the parameter using cluster. The components of this process detailed in this section are: material and equipment, mid IR sensors, dataset, feature selection, and clustering algorithm.

A. Material and Equipment

For the raw material we used pure glycerol with purity level of 85%-90% mixed and synthesized it with Oleic acids pro analysis grade specification in 1:1 ratio charged in reactor with heater and stirring blade. For the catalyst, we used Methyl Ester Sulfonic Acid (MESA) of 0.5%. Esterification reactions were carried out in a laboratory-built apparatus. Apparatus consists of 1000 ml laboratory four neck-round bottom flask with 400 ml working volume. All the reactants (oleic acids, glycerol and catalyst) were weighted and charged into the reactor. Then, the temperature was increased by adjusting the thermostat. The magnetic stirrer was allowed to operate after 5-10 min (to heat up the mixture). After passing the desired reaction time, the reactor was removed from the hot plate. Samples were withdrawn from the reaction mixture for analysis. The reaction mixture was cooled down to the ambient temperature by immersing it into water bath. Several reaction and variation condition were tested in Surfactan and Bioenergy Research Center (SBRC) IPB Bogor laboratorium using traditional and RTAC method. In order to get various conditions of the temperature and reaction time, we setup the apparatus as in Fig. 2. The temperature was controlled between 180°C-230°C and the process time was varied between 100-130 minutes.

Fig. 2. Apparatus experiment in laboratory

B. Mid LED IR Sensors

Mid IR sensors were chosen to design the real time data acquisition system to identify the esterification condition process. Infrared energy is emitted or absorbed by molecules when they change their rotational-vibrational movements. Infrared energy excites vibrational modes in a molecule through a change in the dipole moment, making it a useful frequency/wavelength range for the study of these energy states for molecules of the proper symmetry. Infrared spectroscopy examines absorption and transmission of photons in the infrared energy range [9]. The LED IR sensors with these center wavelengths: 3.4 µm, 5.5 µm and 7.0 µm were used in this study. Each comes with optically immersed specification. LEDs fabricated from III-V hetero structures grown onto InAs substrates type from Scitec Instruments were used in this study for the identification of esterification condition in the process. These sensors were connected to analog inputs using a microcontroller from Arduino for streaming data in real time using a USB 2.0 serial port with BAUD rate 9600 bps.

C. Dataset

The samples were from a sensor (Fig. 3) with a detector that was connected using the Arduino Uno version 2.0 and the real time data were collected with spreadsheet Microsoft Excel 2013 (Fig. 4) and interfaced with pyserial for computational data mining with Python Integrated Development Environment (IDLE). As long as the esterification process, real time data streaming with duration of 2.25 hours, step in minutes incremental recorded to the database. The input data were the percentage transmittance level of IR rays from wavelengths of 3.4, 5.5 and 7 µm, temperature in the reactor, rotation speed of stirrer in reactor and time of esterification process, while the output data were the condition of esterification.

Fig. 3. Optically immersed LED MID IR sensor

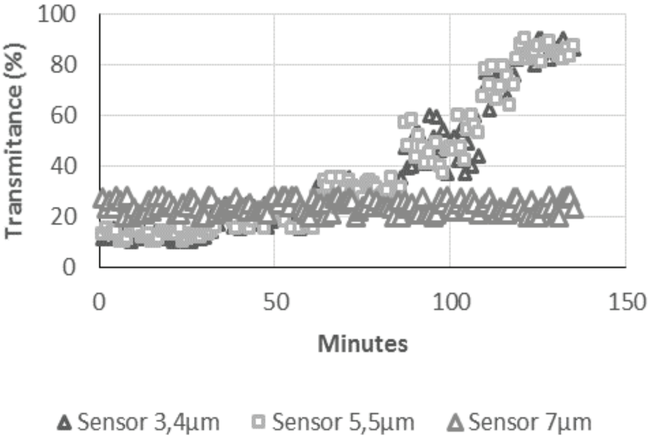

Fig. 4. Real time data from MID IR sensor

D. Feature Selection using Relief (Recursive Elimination of Features)

In this work, to select the sensors that are related to esterification status, we used the Relief (Recursive Elimination of Features) algorithm, which are generalizable to polynomial classification by decomposition into a number of binary problems developed by Kononenko [10]. We used Relief to reduce unimportant features that resulted in a more efficient computation and sensor task. This algorithm is able to improve the reliability of the probability approximation and is robust to incomplete data. We used the algorithm in three-class that changed esterification status as in Table 1 from an assay in laboratory: FTIR (Fourier Transform Infra-Red) and GCMS (Gas Chromatography–Mass Spectrometry) test. Before using Relief feature selection method, we have collected 8 dataset from the processing glycerol esterification with a range of temperature from 120^0C to 230^0C and tested in a certified laboratory as in Table I. From the output of the Relief algorithm in each time range, we got the results as the weighted average number graph (Fig. 5). For example, by comparing the value between them, we chose sensor 1 (3.4 µm), sensor 2 (5.5 µm) and eliminated sensor 3 (7.0 µm) because it has the lowest weighted value. This means that sensor 3 was not suitable for identification of esterification status.

TABLE I.EXAMPLE RELIEF DATASET

Transmittance Detection (%)			Esterification
Sensor 1	Sensor 2	Sensor 3	Status
11	13	20	A
12	15	24	A
11	13	21	A
20	14	24	B
25	20	25	B
20	20	27	B
25	25	25	C
25	30	25	C

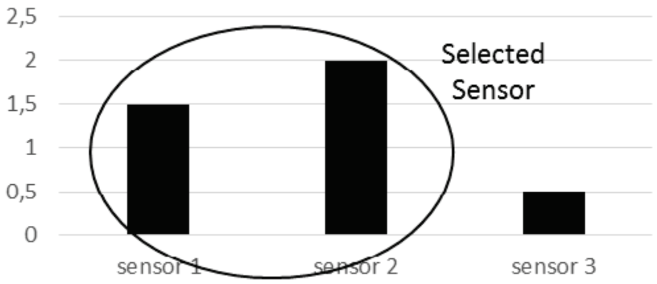

Fig. 5. Example weighted average number score from 90-120 minutes period

E. Pillar K-means Algorithm

To get the right number of data cluster from two sensors input, this study applied the K-means clustering for grouping by pillar algorithm as flowchart in Fig. 6. The pillar algorithm considers that pillars placement should be located as far as possible from each other [11]. The Pillar algorithm is very robust and superior for initial centroids optimization for K-means by positioning all centroids far separately among them in the data distribution [12]. The thought process of determining a set of pillars' locations in order to make a stable house or building inspires this algorithm. By distributing the pillars as far as possible from each other within the pressure distribution of a roof, the pillars can withstand the roof's pressure and stabilize a house or building. It considers the pillars, which should be located as far as possible from each other to withstand the pressure distribution of a roof, a number of centroids among the gravity weight of data distribution in the vector space. Therefore, this algorithm steps conducted for data sensor segmentation using pillar k-mean clustering are:

Step 1. Read the dataset as a matrix.
    Read each data input sensor in the form of matrixes.
Step 2.Calculate the data size.
    Calculate the means of each data and put them in the means arrayed.
Step 3.Find number of cluster.
    -Sort the data from the mid IR sensors.
    -Compare between current elements.
    -Count how many elements remain in the means array.
Step 4.Decide initial furthest centroid of cluster using the Pillar K-means algorithm.
Step 5. Create a group of data.
    -Apply value of centroid to group of data.
    -Re-generate the data in form of data segmentation
    -Compare Silhouette Score.
Step 6. Get control output in the form of data segmentation.
    -The cluster is used as a group set point for PID control.

In this research, to determine the optimal number of cluster from sensors, we used the silhouette number. Silhouette refers to a method of interpretation and validation of consistency within clusters of data. The technique provides a succinct graphical representation of how well each object lies within its cluster [13]

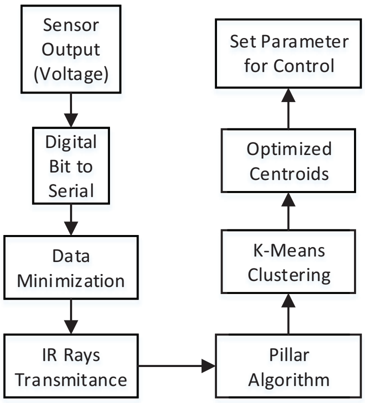

Fig. 6. Initial centroid optimization of K-means clustering

IV. INTEGRATION OF SENSOR DEVICE AND ALGORITHMS

Integration of sensor devices and the algorithms (Relief and Pillar K-means) was intended to measure the transmittance level (variable input) which is related to the esterification condition precisely. The data can be read and collected by the computer using USB 2.0 with a 1-minute interval time sampling to build the database. The esterification process needs a long time to complete (approximately 120- 180 minutes) to produce high yield and good quality of product.

By using a computational algorithm we still need time to process the data before we can get the information needed especially in dynamic reactions. In order to solve this problem, we used the data sampling size of 1 minute, and then we applied the Pillar algorithm (Fig. 6), these architecture computational models were designed on a computer with CPU processor Intel Celeron 1007U and a memory of 2GB with Software Python version 2.7.10 (www.python.org). Based on the data transmittance read by the sensor, we then used Relief method in spreadsheet Microsoft Excel 2013 with Macro programming to select the sensor that are suitable for the esterification, and sent the signal that was detected from the sensor selected to be clustered with the Pillar K-means in Python 2.7.10. After clusterring the data, the parameters of the process, which are temperature and rotational speed of the stirrer, were used to set the position levels of the PID control in Arduino Uno Microcontroller. We used a proportional-integral-derivative controller (PID controller) as shown in Fig.7 because it has a control loop feedback mechanism (controller) widely used in industrial control systems. PID control is often combined with logic, sequential functions, selectors, and simple function blocks to build complicated automation systems used for energy production, transportation, and manufacturing. In this research, a PID controller calculates an error value as the difference between a measured process variable and a desired set point, and used the values to adjust the level or position of rotary resistor switch at the magnetic stirrer and heater, using a servo motor. This control mechanism was very useful in adaptive real time process monitoring and control in esterification.

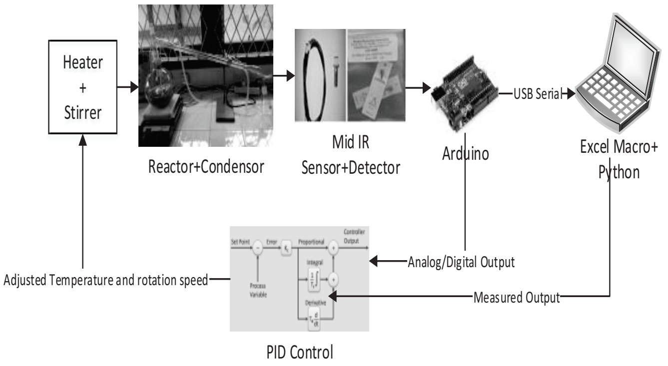

Fig.7. Block diagram of the integration real time control system

V. EXPERIMENTAL RESULTS

As setup for experimental tools, laboratory equipment and integration between sensor and algortihm that explained in Section III and IV, we conducted a performance comparison between the traditional control and the proposed method of adaptive optimization control with one sequence of batch learning for the computational algorithms to set the optimization parameter. In traditional control, we used monitoring by a human to control the set temperature at 230^0 C and process time as 120 minutes. For the adaptive proposed method, as Relief algorithm to select the sensor wavelength that closely related to esterification process as a result from comparison between Weighted Average score in batch learning process for 0,30,60,90,120 minutes (Table II.) and real time adaptive control method with the Pillar K-means algorithm used in this paper, we set α=0.25 and β=0.58 for the detection the outliers as used in Barakbah (2010) and the data points with silhouette score as in Table III-VI.

TABLE II. SELECTED MID IR SENSOR

Minutes	Sensor 3,4µm	Sensor 5,5µm	Sensor 7µm
0	ON	OFF	ON
30	OFF	ON	ON
60	OFF	ON	ON
90	ON	ON	OFF
120	ON	ON	OFF

TABLE III. DATA POINTS FOR K=5: (SILHOUETTE SCORE: 0.536)

Cluster	Data Points
0	87,89,90,91,92,94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109,110,111,112,113,114,115,116,117,118,119,120,121,122,123,124,125,126,127,128,129,130,131,132,133,134,135
1	6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,
2	1,2,3,4,5,
3	21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,46,
4	45,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,88,93,

TABLE IV. DATA POINTS FOR K=4: (SILHOUETTE SCORE: 0.489)

Cluster	Data Points
0	107,109,110,111,112,113,114,115,116,117,118,119,120,121,122,123,124,125,126,127,128,129,130,131,132,133,134,135
1	33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,58,59,60,61
2	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18 19,20,21,22,23,24,25,26,27,28,29,30,31,32,57
3	62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100,101,102,103,104,105,106,108

TABLE V. DATA POINTS FOR K=3: RAW MATERIAL TYPE 1 (SILHOUETTE SCORE: 0.604)

Cluster	Data Points
0	107,109,110,111,112,113,114,115,116,117,118,119,120,121,122,123,124,125,126,127,128,129,130,131,132,133,134,135
1	54,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100,101,102,103,104,105,106,108
2	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,55,56,57,58,59,60,61

TABLE VI. DATA POINTS FOR K=3: RAW MATERIAL TYPE 2 (SILHOUETTE SCORE: 0.601)

Cluster	Data Points
0	109,110,111,112,113,114,115,116,117, 118,119,120,121,122,123,124,125,126,127,128,129,130,131,132,133,134,135
1	59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99, 100,101,102,103,104,105,106,107,108
2	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18 19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,55,56,57,58

TABLE VII. CONTROL PARAMETER WITH CLUSTER

Cluster	Time (Minutes)	TEMPERATURE POSITION LEVEL	ROTATIONAL STIRRER POSITION LEVEL	MODE
2	1-58	5 (Full)	2	Heating Up
1	59-108	3-4	3-4	Stabilizing
0	109-135	2	4	Finishing

From Tables III-VI, it shows the conditions and characteristic of the esterification process in the laboratories, the data clustering carried out with K=3 gave a silhouette score of 0.601-0.604, and with K=4 a silhouette score: 0.489. The silhouette score with K=3 with two type of feed composition was higher than K=4 and K=5, so we decided to use 3 cluster for temperature and rotational stirrer speed control parameter as in Table VII. and the comparison result from existing methods as in Table VIII.

TABLE VIII. COMPARISON RESULT

COMPARISON	METHODS		
	Existing (Traditional)	Adaptive Type 1	Adaptive Type 2
Process Time	120 minutes	105 minutes	100 minutes
%Yield	85	93	92
Stirrer (rpm)	200-500	300-450	300-450
Final Temp(°C)	230	210	200

As shown in experiments of this research, we find the advantage of this proposed model are: The esterification process finished in a shorter time than with existing methods, an increase in the % yield of product (mono and di-glycerides), a lower final temperature in the reactor (reduce consumption of energy), lower range of stirrer rotation using two type of feed composition (Glycerol) type 1 and type 2. On the other hand, the disadvantage of this proposed model is: this esterification optimization system works well in reaction with oleic acid but need different sensor for different reactant, this system still require learning batch and several precise laboratory test for calibration.

VI. CONCLUSION

In this paper, we have presented a new approach for optimization process of Glycerol esterification with Relief feature selection to select the sensors at specified intervals and Pillar K-means data clustering algorithm to determine the set control parameters like temperature and rotational speed of stirrer in real time adaptive control. The experimental results show that our proposed approach for this research is able to identify the conditions of the process, increase product yield (7%-8%), reduce process time by 15-20 minutes, lower range of stirrer rotation (300rpm-450rpm) and lower final temperature at 200^{0}C -210^{0}C. For recommendation: this proposed model can developed using IR sensors with variable wavelength and another computational method for optimizing process.

ACKNOWLEDGMENT

The authors would like to acknowledge the financial support of Kementerian Pendidikan dan Kebudayaan Republik Indonesia (DIKTI) with contract number 095/K3/KM/2015

REFERENCES

[1] D. Isabel, M. Federico,P. Joaqun,S. Enrique, "Synthesis of MCM-41 materials dialkyl silane groups and their catalytic activity in functionalized with the esterification of glycerol with fatty acids," Applied Catalysis A: General, 242,2003, pp.161-169.

[2] J.Cramer,M. Dueholm.,S. Nielsen,D. Pedersen, and J.F.Cramer, "Controlling the degree of esterification in lipase catalyzed synthesis of xylitol fatty acid esters," Enzyme and Microbial Technology, 41, 2007,pp. 346-352.

[3] .V. Adetola, D. DeHaan, M. Guay," Adaptive model predictive control for constrained nonlinear systems," Systems and Control Letters 58 (5) ,2009,pp. 320–326.

[4] J.Wieslander and B. Wittenmark, "An Approach adaptive control using real time identification," Automatica Vol 7. Pergamon Press, 1971,pp. 211-217.

[5] R.B Sepe and J.H.Lang," Real time adaptive control of permanent magnet synchronous motor," IEEE Transaction on Industry Applications.Vol 27 no 4,1991, pp.706-714.

[6] I.A Soenandi and T.Djatna, "A real time simulation model of production system of Glycerol Esterification with self optimization," in Advanced Computer Science and Information Systems (ICACSIS), 2014 International Conference, pp.41-46, 18-19 Oct. 2014.doi: 10.1109/ICACSIS.2014.7065864.

[7] I.A.Soenandi,T.Djatna,A.Suryani&Irzaman "Real Time Monitoring Glycerol Esterification Process with Mid IR Sensors using Support Vector Machine Classification," in 3rd Interational Conference of Adaptive Agro-Industry (ICAIA),Bogor,3rd-4th Ags.2015,pp. 350-355.

[8] G. Chryssolouris, M. Domroese, P. Beaulieu ,"Sensor Synthesis for Control of Manufacturing Processes," Journal of Engineering for Industry, ASME, Volume 114, No. 2,1992, pp. 158-174 .

[9] A.Edelmann, J.Diewok,K.C Schuster & B.Lendl," Rapid method for the determination of red wine cultivars based on mid-infrared spectroscopy of phenolic wine extracts," Journal of Agriculture and Food Chemistry, 49,2001,pp.1139–1145.

[10] I.Kononenko,et al,"Overcoming the myopia of inductive learning algorithms with RELIEF," Applied Intelligence, 7(1),1997, pp.39-55.

[11] A.R. Barakbah, A. Helen, "Optimized K-means: an algorithm of initial centroids optimization for K-means," Proc.Seminar on Soft Computing, Intelligent System, and Information Technology (SIIT), Surabaya, 2005.

[12] A.R. Barakbah, Y. Kiyoki, "A Fast Algorithm for K-means Optimization using Pillar Algorithm", The 2nd International Workshop with Mentors on Database, Web and Information Management for Young Researchers. August 2-4,2010, Tokyo, Japan.

[13] P J. Rousseeuw . "Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis". Computational and Applied Mathematics 20,1987, pp. 53–65. doi:10.1016/0377-0427(87)90125-7.