ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 6, June 2017

Statistical Data Analysis and Performance Evaluation of Batch Process Pilot Plant

Rahul.B.Dawre[1], Sudhir.D.Agashe[2]

PG Student [Process Instrumentation], Dept. of I&C Engineering College of Engineering, Pune Maharashtra India[1]

Professor, Dept. of I&C Engineering College of Engineering, Pune Maharashtra India[2]

ABSTRACT: The objective of presenting this paper is designing and predicting the activities related to data analysis of the batch process pilot plant using R-tool on data acquired through plant sensors with the help of plant local PLC. This pilot plant helps us to analyse the effects on variables at different process conditions and in optimization of the process. The optimization of the process depends upon the efficiency in terms of energy, usage of raw material, time consumed and manpower required to run the process. To optimize the process it is necessary to find opportunities which save on the above mentioned parameters. To accomplish this objective, it becomes essential to collect data regarding the process. This data needs to be processed and analysed so as to understand the performance of the process. This knowledge of the process will be used to improve the efficiency. This plant is workbench for experimentation and data collection for better understanding of analytical concepts like co-relation, regression, data modelling and visualization. Predictive Analytics and Parameter forecasting is widely used for equipment maintenance. From the data analytics we can predict the performance of the plant in the future.

KEYWORDS: Data Analysis, R tool, Pilot Plant, Efficiency, Optimization, Energy efficiency, PLC, Collect data, Process.

I. INTRODUCTION

Industrial manufacturing processes can be generally classified as continuous, discrete or batch. A process is said to be batch in nature if the physical structuring of the process consists of a sequence of one or more steps (or phases) that must be performed in a given order. The completion of this sequence of steps creates a finite quantity of finished product. If more of the product is to be created, then the sequence can be repeated. Batch processes are neither discrete nor continuous but they have the characteristics of both. In the process industries like pharmaceutical and many other batch production companies, it has become quite important to statistically analyze data of the plants. From data analysis, predictive model for the plants can be designed that will enhance their performance.

Previously, industries used to perform maintenance, only when breakdown used to occur. Thus, this statistical data analysis will help in predictive maintenance. In general, the batch process exhibits some batch-to-batch variations because of errors in the charging of the recipe of materials, deviations of the process control between operating point and set point, and disturbances to the manipulated variables. To achieve consistent product quality from a batch process, data analysis of batch-to-batch variability is important.

The acquired sensor data is of much value from the data analytics perspective. The enormous data collected from the pilot-plants is used to understand the statistical correlation between Dependant and Independent variables which can in-turn give insights into domains like reliability and risk management, HAZOP & predictive maintenance.
The main aim of this work is to make the batch process pilot-plant facility run continuously and efficiently from local as well as remote controls. Asset mapping and realization for implementation of business models can be done from the data analysis.

II. OVERVIEW OF BATCH PROCESS PILOT PLANT

The hardware consists of four tanks, three feed & one reactor tank, various sensors and final control elements. All tanks are made of SS-304 due to food grade application of the plant. The reactor tank is jacketed to prevent heat loss

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 6, June 2017

and maintain the internal temperature desired for the batch product. There are four tanks named as Tank A, Tank B, Tank C and Reactor Tank. Tanks A, B and C are of same dimension and same capacity.

Fig.1 Overview of Batch Process Pilot Plant

III. MANAGEMENT OF BATCH PROCESS PILOT PLANT

A control philosophy is a detailed narrative that makes us aware of the various operations in Batch Processes. For the pilot plant setup at advanced process lab COEP this narrative is a set of instructions to operate pilot plant remotely and locally, initial equipment check list, start-stop sequence and emergency stop procedures following are the steps.

1. Verify all check lists.
2. Open solenoid valve 1 & then start feed pump.
3. Once high level switch is triggered of Tank A, then close Solenoid valve 1 & open Solenoid valve 2.
4. Once high level switch is triggered of Tank B, then close Solenoid valve 2 & open Solenoid valve 3 till high level in Tank C reached.
5. Close Solenoid valve 3 and stop feed pump.
6. Once the ingredient is filled into all the ingredient tank then as per requirement select the recipe.
7. To fill A ingredient from tank A we have to open the angle valve 1 & angle valve 4 to fill the reactor tank as per the selected recipe. Ingredient A is filled into the reactor tank close the angle valve 1 & angle valve 4.
8. To fill B ingredient from tank B we have to open the angle valve 2 & angle valve 4 to fill the reactor tank as per the selected recipe. Ingredient B is filled into the reactor tank close the angle valve 2 & angle valve 4.
9. To fill C ingredient from tank C we have to open the angle valve 3 & angle valve 4 to fill the reactor tank as per the selected recipe. Ingredient C is filled into the reactor tank close the angle valve 3 & angle valve 4.
10. Start the agitation process with help of agitator motor.
11. Regulate and maintain the temperature loop of reactor tank as per the selected recipe.
12. For maintaining & controlling the temperature of the reactor tank, circulate the cold water & hot water inside the reactor jacket.

IV. IMPORTANCE OF DATA ANALYTICS

Sensors used in process plants generates large amount of data. This data is processed and used for the benefit of industries. Data analysis helps us to optimize the raw materials and time needed to execute a process.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 6, June 2017

Data Analytics - Techniques and their Application

In this section we will understand in depth the trendiest technological concepts that have swept the industrial world like a hurricane. Many experts believe that these technologies will embark the new industrial revolution. We will be restricting our discussion to: *Data Analytics.*

Data analytics, also known as analysis of data is a process of inspecting, cleansing, transforming and modelling data with the goal of discovering useful information, suggesting conclusions, and supporting decision-making. Data analysis has multiple facets and approaches, encompassing diverse techniques under a variety of names, in different business, science, and social science domains. Data analytics analyses large amounts of information to decode hidden patterns, correlations and other insights. Data analysis breaks the total information into the separate components for individual analysis. Data is collected and examined to solve questions, test hypotheses or disprove theories. The main steps in which data collection-to- data analysis is achieved are given as follows:

1. Measurement
2. Analysis
3. Reporting/enterprise
4. Collaboration/collaboration platform
5. Knowledge management

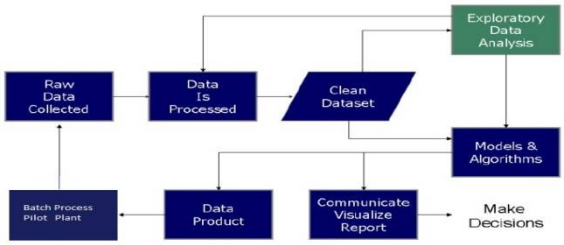

Fig.2 Processes involved in Data Analysis

V. INTRODUCTION TO R TOOL OPEN SOURCE (SOFTWARE ENVIRONMENT FOR DATA ANALYTICS)

Data analysis is a way to summarize and visualize the important characteristics of a data set. Learning R is same as learning an Application program interface (API). An API is made of operation routines, protocols and tools for building software applications. An API specifies the interaction between software components. Additionally, API's can also be used as programming graphical user interface (GUI) components.

Exploratory data analysis (EDA) comes before formal hypothesis testing and data modelling and makes use of visual methods to analyse and summarize data sets. R will be our tool for generating those visuals and conducting analyses. We will install R Studio and packages, learn the layout and basic commands of R , inspect data sets and generate data visuals. We Perform EDA to understand the distribution of a variable and to check for anomalies and outliers using R. We quantify and visualize individual variables within a data set to make sense of a random data set. We create Histograms/box plots/charts for better data visualizations. Getting Started with R

We will begin our journey with taking the first steps:

• *Download & Installation:*

Downloading a suitable binary distribution of R for your operating system.

• *Get R studio:*

R Studio is a leading IDE for R development. This will help you to code more productively with all the plots, package management and the editor in one place.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 6, June 2017

A.Finding Co-relation between Temperature and level sensors parameters using Linear Regression

Simple Linear Regression is used to predict the value of a dependent variable using an independent variable; also the alpha value gives us information about the relation between the two. In the case of Batch Process pilot plant I will be using this technique to find the co-relation between the three most important parameters which are:

- Temperature-(TT-4)
- Temperature-(TT-5)
- Level (LT-4)

The data acquired from the sensors is usually stored in EXCEL format. We will first convert the .XLS into a CSV format to load it in the R tool. Following figure shows us how data acquired from the sensors is stored in excel sheet. Once the CSV file is ready we will load this file into R using the R Studio workbench. Once the data is successfully imported from the CSV file to R, it will reflect in the R script, R console & R environment windows of the R Studio. We will now generate results to find the co-relation between the parameters depending upon the alpha & p values. While comparing the alpha and the P-values two conditions that can emerge are:

- The p-value is less than or equal to alpha (p-value < 0.05)

In this case we reject the null hypothesis.. When This happens we say that the result is statistically significant.

- The p-value is greater than alpha (p-value >0.05)

In this case we fail to reject the null hypothesis. When this happens we say that the result is not statistically significant. Now we find out the p-value of TT-5 against LT-4 & TT-4 P-values of Temperature & Level collected from the R tool give us a clear idea about the correlation between the three & their statistical significance.

B.PREDICTION BASED ON HISTORICAL DATA

The main objective of prediction is to be equipped to take an informed decision based on the past or historical data that we have. Parameter forecasting is the under structure on which algorithms of Predictive Analytics are written. Organizations achieve their strategic maintenance objectives by going beyond regular maintenance standards. By replacing reactive and preventive maintenance with a predictive maintenance strategy, any organization can prevent breakdowns and costly repairs reduce travel time and the costs associated which inturn improves the asset life and shortening the return on investment.

A database of values for the two parameters was collected for an entire year (approximately). The on-site Level and Temperature transmitters send the real-time data to the host PLC. From the PLC this data is archived in the Historian. The archived data can be made available on the Cloud or Mongo-DB platforms through data algorithms.

Time based data is a specific type of time-series. A time series is a collection of measurements of variable with respect to time. For example, measuring the temperature value from an in-line sensor during execution of a single batch would comprise a time series.

Objective of analysing a time series data set is to:

- *Identify the patterns in the data set*
- *Prediction from previous patterns*

VI.RESULTS

Taking the results of linear regression from the R tool, we can conclude that level of the reactor tank (LT-4) and reactor temperature (TT-3) is strongly co-related.

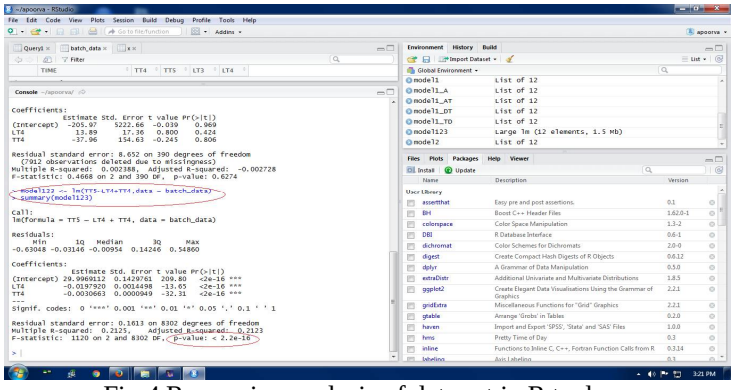

Fig.4 Regression analysis of data set in R tool

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijareeie.com

Vol. 6, Issue 6, June 2017

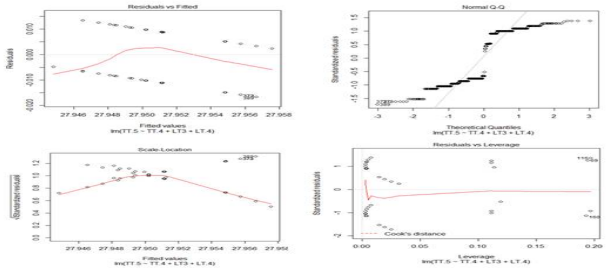

Fig.5 Analysis Plots

VII. CONCLUSION

The results of the analysis allowed to focus on the critical variables. By performing data analysis on data obtained from various sensors installed on Batch Process Pilot Plant we can determine the most important process parameters which affects the quality of end product. Also the knowledge obtained from data analysis can be useful for the predictive maintenance of the plant in near future.

REFERENCES

[1] L. H. Chiang, R. Leardi, R. J. Pell, and M. B. Seasholtz, "Industrial experiences with multivariate statistical analysis of batch process data," Chemometrics and Intelligent Laboratory Systems, vol. 81, no. 2, pp. 109–119, 2006.

[2] E. Keogh and S. Kasetty, "On the need for time series data mining benchmarks: a survey and empirical demonstration," Data Mining and knowledge discovery, vol. 7,no. 4, pp. 349–371, 2003.

[3] Y. Su and F. Yu, "Data mining applications for finding golden batch benchmarks and optimizing batch process control," in Intelligent Control and Automation (WCICA),2016 12th World Congress on. IEEE, 2016, pp. 1058–1063.

[4] P. Eichinski and P. Roe, "Datatrack: An r package for managing data in a multi-stage experimental workflow," 2016.

[5] N. Naik, P. Jenkins, N. Savage, and V. Katos, "Big data security analysis approach using computational intelligence techniques in r for desktop users," in Computational Intelligence (SSCI), 2016 IEEE Symposium Series on. IEEE, 2016, pp. 1–8.

[6] D. James, \Batch process automation - applying the principles of batch automation and selecting a suitable control system and supplier is only part of the story.managing a batch automation project requires considerations over and above those needed for the automation of a continuous process plant.,"Manufacturing Engineer, vol. 85, no. 6, pp. 36{41, 2006.