UNIVERSITY OF GÄVLE

FACULTY OF ENGINEERING AND SUSTAINABLE DEVELOPMENT
Department of Electrical Engineering, Mathematics and Science

Production 4.0 of Ring Mill 4 Ovako AB

Muhammad Hassan

September 2019

Student thesis, Advanced level (Master degree, two years), 30 HE
Electronics
Master Programme in Electronics/Automation

Supervisor: Marcus Svadling (Ovako Sweden AB)
Assistant supervisor: Prof. Niclas Björsell (University of Gävle)
Examiner: Dr. Per Mattsson (Uppsala University)

Abstract

Cyber-Physical System (CPS) or Digital-Twin approach are becoming popular in industry 4.0 revolution. CPS not only allow to view the real time status of equipment, but also allow to predict the health of tool, which lead it towards smart maintenance. CPS can contribute to sustainable environment as well as sustainable production, due to its real-time analysis on production. In this thesis, we analyzed the behavior of a tool of *Ringvalsverk 4*, at Ovako with its twin model (known as Digital-Twin) over a series of data. Initially, the data contained unwanted signals which is then cleaned in the data processing phase, and only before production signal is used to identify the tool's model. Matlab's *system identification* toolbox is used for identifying the system model, the identified model is also validated and analyzed in term of stability, which is then used in CPS. The Digital-Twin model is then used and its output being analyzed together with tool's output to detect when its start deviate from normal behavior.

Keywords: System Identification, Cyber-Physical System, Data analysis, Digital Twin, Cyber-Twin, Smart Maintenance, Modelling.

Preface

I would like to thanks to my supervisor Marcus Svadling, my additional supervisor Niclas Björsell, and everyone in Ovako who helped and support me. I considered this master thesis project as a great source of learning opportunity, in which I have used the tool that have learned from theoretical courses and implemented into the industrial environment. The three phase of project, data processing, identification of digital-twin model, and analysis of output using cyber physical system are quite interesting and challenging. Motivation from colleagues, guidance and suggestions from supervisor and professor are highly appreciable, which help me to boost my focus and make this project successful. I believe the successful results that achieved were only due to the guidance of professor and support from Ovako.

Table of contents

List of Abbreviations

CT	Cyber Twin
DT	Digital Twin
HHRR	Horizontal Hot Ring Rolling
CPS	Cyber-Physical System
CS	Cyber-Space
IoT	Internet of Things
PS	Physical System
LTI	Linear Time Invariant
LHP	Left Half Plane
RHP	Right Half Plane
BIBO	Bounded Input Bounded Output
LS	Least Square
RLS	Recursive Least Square
RUL	Remaining Useful Life
PEM	Prediction Error Method
SISO	Single Input Single Output
MIMO	Multiple Input Multiple Output
PLC	Programmable Logic Controller
ARx	Auto Regressive with exogenous Input
ARMAX	ARx with Moving Average
OE	Output Error model
GUI	Graphical User Interface

1 Introduction

Ovako is a leading manufacturer of engineering steel in European market for customers. The high quality product are used in various engineering applications e.g. in the bearing, transportation and manufacturing industries. With geographical presence in Europe, North America and Asia, and a steel product line that meet the high quality standard, satisfy the customer demand and provide customized solutions. Ovako is contributing to create value not only for its customers but also for their customers all over the world.

Ovako manufacture clean, high strength and sustainable steel. The steel production is sustainable due to the process, which is based on steel scrap and a Nordic low-carbon electricity mix. The unique process of Ovako, produce steel product with the carbon footprint 80 percent lower as compared to the global average. By minimizing the impurities during steel production process, the end product become clean and strong. This gives the Ovako’s steel high-quality properties that enable customers to create lighter, stronger and more durable end products.

High quality products that are use e.g., in wind turbine are produced from ring production. There are two way for ring production at Ovako. One way is to produce by sawing rings from a tube, and the second way is to roll the hot metal preform by horizontal hot ring rolling process (see Figure 1.1.1). However, this project is based on predictive maintenance of hot ring rolling process by using Digital-Twin approach. Which would then contribute to make production better. In this project, work has been done offline and result analysis has been done jointly with process people. There were numerous maintenance activities that have been performed during the offline period. The results in this report are explained well and satisfied the expectation of Cyber Physical System. However, some of the results are not clarified in detail due to lack of information and further analysis would be desirable.

This report is organized as follows. Section 1 gives the introduction and physical intuition of ring rolling process and CPS. Section 2 describe the method that used for system identification, description of physical system, system stability, and factors that affects the process. In Section 3, big data has been cleaned and extract only the desired data (Section 3.1). System identification method has applied to identify real plant that discussed in Section 2. In which, Matlab *System Identification* toolbox is used to find correct model order and evaluate different models (Section 3.3). However, it is also discussed to identify the model without Matlab *System Identification* toolbox (Section 3.4). The identified model is used as digital twin of real plant, and CPS technique is used to analyze the behavior of a tool in Ovako ring rolling process. The report ends with discussion and conclusion in Section 4 and 5.

1.1 Ring Rolling Process

Ring rolling is a complex hot forming process, and it is used for ring production of different shapes [1] with different diameters and cross sections, widely employed in energy, bearings, aircraft, and automobile industries. In the mill five major tools are used that are axial roller, support arms, drum roller, measuring roll, and mandrel. During the process these tools are involved to produce the desired shaped ring. During the rolling process, all of the tool work simultaneously, therefore it is important for a good ring production that all tools must be in good condition and maintained timely.

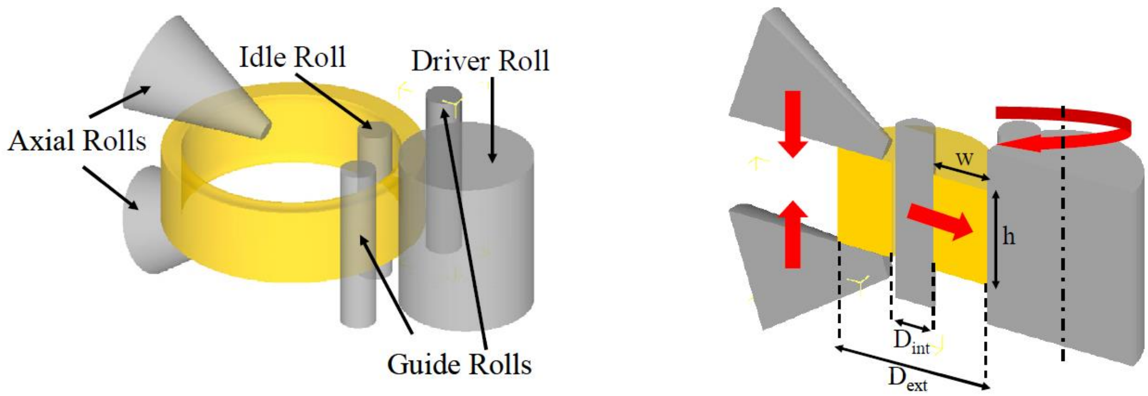

Fig 1.1.1 Process scheme of horizontal hot rolling ring production [1]

Process scheme is shown in Figure 1.1.1, a pierced circular preform is placed over Idle/Mandrel roll. The mandrel push the ring against Driver/Disk roller to decrease wall thickness of ring. With the rolling speed of disk roller, the ring starts expend and move outside. The two axial rollers apply the pressure, the upper axial roller move downward to define the height of the rotating ring, this movement is also illustrated in the upper sketch (8) in Figure 1.1.2. The measuring roll serves as to detect the current outer diameter [2]. Meanwhile, two guide rolls/support arms guide the ring to be centered by controlling the position of ring, they are the essential tool of ring rolling mill. The overall tools' motion are adaptively controlled in a closed-loop system according to the process and with preselected control strategies.

Horizontal hot ring rolling mills usually equipped with two support arms, which are controlled by hydraulic cylinder [3]. Figure 1.1.2 show the side and top view of horizontal hot ring rolling process (HHRR) scheme, in which arrows are representing the movement and direction of tool.

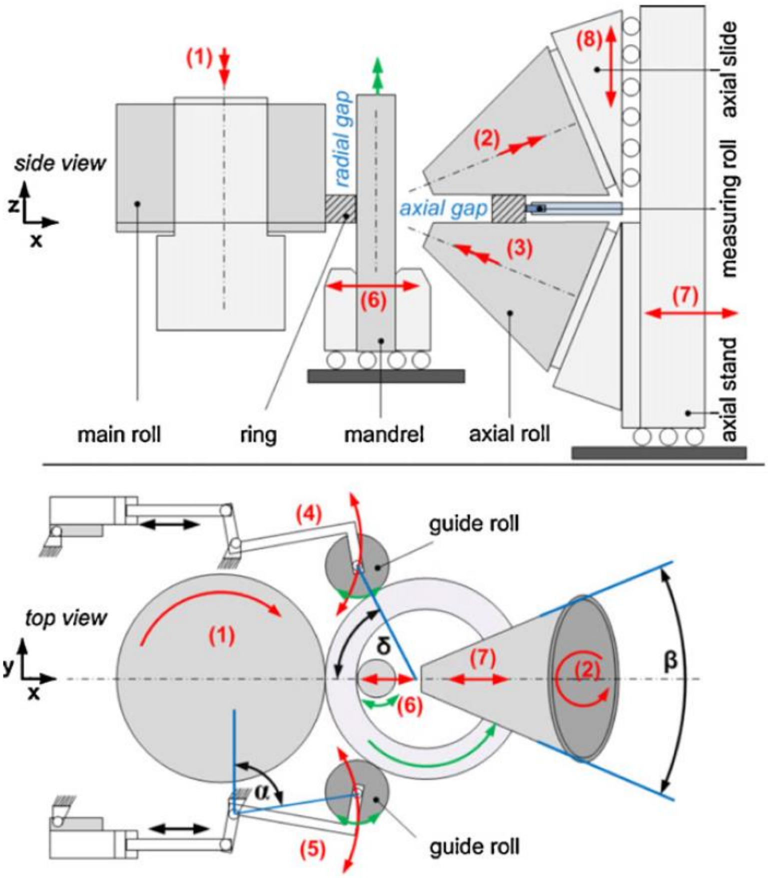

Fig 1.1.2 Process scheme of horizontal hot rolling ring production [2]

It can be observed from Figure 1.1.1, and 1.1.2 that each tool in the mill plays essential role in ring production. When the tools are not in a good condition, the production become unsatisfactory. It affect the product quality, production cost, and have high safety risk. Modern industries such as Ovako AB are moving towards predictive maintenance by using Digital-twin. *Cyber Twin (CT) or Digital Twin (DT) is basically a same thing*, which is the living model of physical plant [4]. This thesis is focus on predictive maintenance of one of the *Ringvalsverk 4* tool by using CPS [5], as a pilot project in Ovako AB, Hofors.

1.2 Digital-Twin Approach to Predictive Maintenance

In recent years, “Industry 4.0” has been widely discussed, and become an industrial revolution for most global industries. Modern industries are migrating from reactive maintenance to predictive maintenance. Reactive maintenance refer to repair of plant to restore it to normal operating condition when the equipment is damage or broken. It is also known as unplanned maintenance, it shorter the plant life expectancy, and require immediate action to rectify failures [6]. On the other hand, predictive or proactive maintenance is to predict future failure of equipment and allow to plan maintenance before the failure occur [7]. It allow to increase the production efficiency, plant efficiency, and plant’s lifespan. In predictive maintenance, process and theoretical knowledge of physical plant with data-driven analytics generate a new paradigm called “Digital-Twin” [8].

The Digital-twin is the mathematical model of physical system, which is continuously a learning system, powered by sensor data and machine learning algorithm. In other words, it is the virtual replica of actual physical system, which can be a production line, industrial machine, or the whole industrial plant [4]. The DT technique continuously adapted the operational changes based on the online data and information, and predict the future behavior of the real plant [8]. It is the essential component of industry 4.0, and allow centralized control and analysis of online production [10]. With this, it help to improve product design, and identify potential degradation by monitoring equipment health.

DT requires digitization of plant, and then use the plant’s digitalized data to power its twin model [11]. Internet of things (IoT) plays a crucial role in digitization of existing plant, different sensors are used to measure the state of physical plant. According to the survey, the expected level of enterprise digitization would be double in few years [12]. The high availability of sensor data generates a high volume of data, known as Big Data [13]. Which can be further processed to extract the additional information that provide the better insight of process. In industrial environment, physical and digital space can be connected via IoT [14].

The plant’s twin model is obtained or designed by the using system identification method, which is discussed in Section 3.3. To get the simulation model, modelling expert collaborate with process engineers that describe the physical properties of the plant. The twin characteristics can be summarized as an online situation reflection [15]. That reflect the physical system behavior and experiments are carried out from simulation model. The simulation-based decision making support tools are used for predictive maintenance by analyzing physical system output and simulated model output [16].

Ringvalsverk 4 (Plant) is running by the controlled input, which is sensed by the sensor. The sensed data is also used to power the DT. With the help of process knowledge, plant and twin output will analyze together to compare it with the historical data. It will help to predict future failure, help to avoid the critical damage, reduce the maintenance cost, and increase the life span of the plant. Figure 1.1.2 shows the concept of Cyber-Physical system.

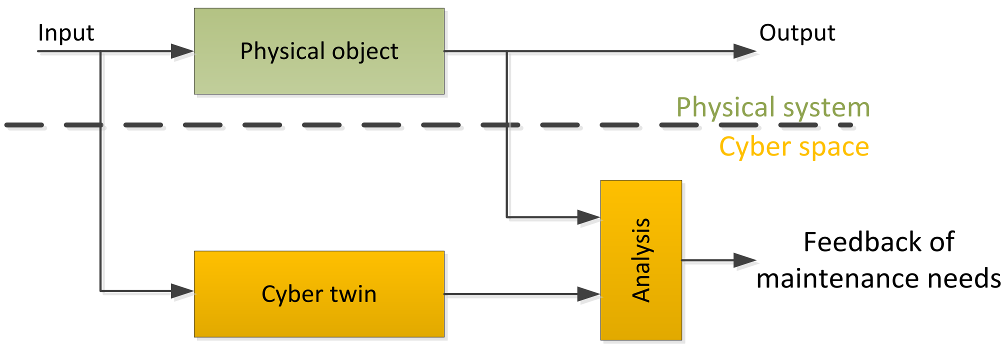

Fig 1.1.2 Cyber-Physical system concept for predictive maintenance

In the cyber space, information is being pushed from every connected machine to form network. With the massive information gathered from the sensors, specific analytics to be used to extract specific information, which would help to provide better insight [17]. The measured output from the physical system (PS) and digital twin are analyzed together by the analyzer block in cyber space (which is also known as digital space).

Currently in the *Ringvalsverk 4*, all tools are maintained by weekly maintenance and the target is to move from weekly maintenance to predictive maintenance. The overall theme of this project is to process the Big Data (sensor data) to extract the useful information, identify the DT model of plant, power the model with real plant’s sensor data, and analyze the physical and digital space output to predict the need of maintenance.

2 Theory

2.1 Logic Architecture use as Data Processing

Logical architecture techniques have been famous to address several process of sensor fusion. Such as logic based recognition of pattern. The general logical architecture is applied to implement the autonomous process that provide the interpretation and sensor fusion of data as well as logic for process control [18]. It is an essential tool in data processing and cleaning, where desired data can be extracted by the logical operation. The clean data is commonly use in system identification for example sensor data gathered from the plant is in raw form, contain additional information such as off data, preproduction data. Based on the interest, such data can be cleaned by using logical operators.

Input logical operation are depend upon set theory, in which intersection shares different logical conjunction properties such as associativity property of binary operation. AND operator is the intersection of element in set theory (collection of object). AND is the truth logical function that accept truth value as produce truth value as output. The output of AND logical true operand depend only upon all input true operand value. For example, the AND operation of A and B is true if both A and B are true. NOT logical operator converts true to false and false to true. For example in a cyclic process, digital signal can switch to high level during the calibration mode and another digital signal can switch to high during production mode. With the help of logical operators these digital signals can be used together to process the data easily. Figure 2.1.1, 2.1.2, and 2.1.3 shows an example how logical operator can be use in data processing.

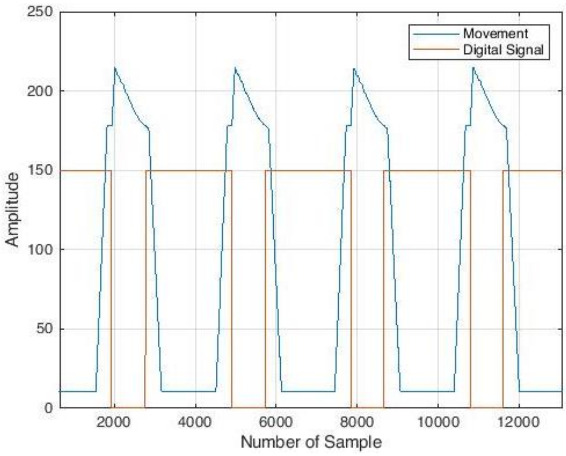

Figure 2.1.1 shows the cyclic movement of a tool (Blue graph), Digital signal become zero during process (Orange graph).

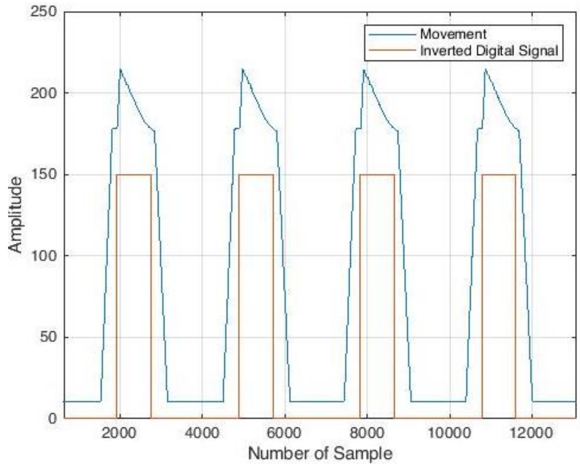

Figure 2.1.2 shows the cyclic movement of a tool (Blue graph), Inverted digital signal become high during process (Orange graph).

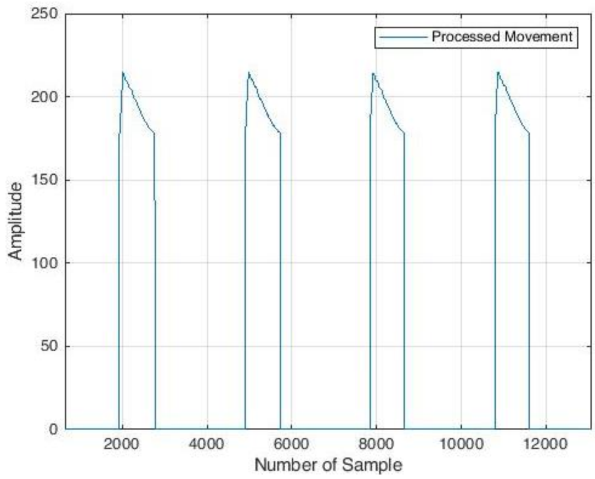

Figure 2.1.3 shows the processed cyclic movement of a tool movement (Blue graph).

We can see from the Figure 2.1.1 that non production digital signal can be inverted with the help of logical operator (NOT), and then it became high during the production shown by orange graph in Figure 2.1.2. One can use the resulted digital signal and the measured signal (shown in Figure 2.1.2) and compute together to acquire only the production signal, shown in Figure 2.1.3. Similarly, calibration signal and non-production signals can be removed in a single step by applying AND logical operator to calibration digital signal and process digital signal. In a complex process where multiple digital signal are generated, different logical operator such as AND or XOR can be used to process data easily and accurately.

2.2 Dynamic System (Plant)

System identification is a field of dynamic systems, modeling by using the experimental data. Figure 2.2.1 shows the conceptual description of a dynamic model. The system is feed by the input signal *u(t)* which is then processed and give the useful information about the system as output *y(t)*, however disturbance *w(t)* also influence the process and effect the output [19]. The user can control the input but the disturbance cannot be control.

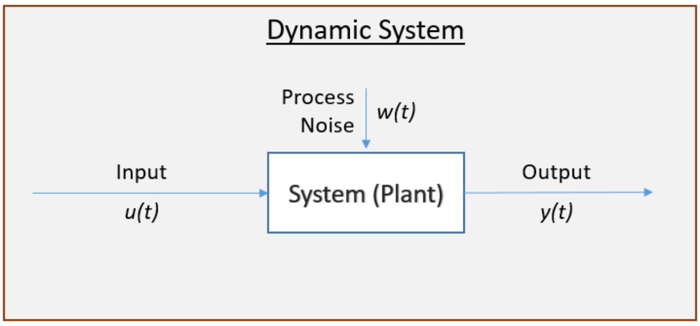

Figure 2.2.1 Dynamic system with process disturbances as $w(t)$*, feed the system with input as* $u(t)$*, and get output as* $y(t)$*.*

It can be observed from the block diagram in Figure 2.2.1 that process noise $w(t)$ is included in the system, which act as an additional unwanted input. Hence, the system process both input signal and unwanted signal. The output of a true world system always influenced by additional unwanted signals typically known as disturbances. These disturbances are discussed in Section 2.3.

2.2.1 Use of Mathematical Model in Maintenance

Mathematical model can plays a vital role in maintenance or in predictive maintenance. In industrial revolution, most of the human-interface process are obsoleting by automatic process or by robotic arms. Therefore, it is necessary to maintain such system properly and timely. An industrial robotic arm can be seen as an advanced servo motor system [20], and its mathematical model can be used in maintenance. Consider the robotic arm as a system. The system (robot arm) has a certain movement, for example for the purpose of lifting object from a specific position to another. It is obvious to consider position as an output $y(t)$. The movement of the arm is controlled by the electrical motors that are driven by control current. Therefore, it can be regarded as control input $u(t)$. The movement can be influenced by the varying load on the arm and by the friction, which can be regarded as process disturbances $w(t)$. By using the process and theoretical knowledge (like system identification), it is possible to identify the system mathematical model and use it as a twin. With the help of sensors, one can measure the amount of control input $u(t)$ and feed it to the identified mathematical model. The measured output position $y(t)$ of the robotic arm could be compared parallel to the identified twin model output. The performance deviation between the two systems can tell the need of maintenance. However, it can also help to estimate the remaining useful life (RUL) of the robotic arm. Hence, to maintain fast and reliable movement, it is important to identify an appropriate system.

The above example tells the need of modeling in industrial system. Many process in the industrial environment are designed by closed-loop control system in order to run the system fast, safe, and efficient. To create a digital twin, designer need a model that can also describes the disturbance properties acting on the process. The dynamics of the real world system might be unknown and contained numerous parameters. System identification technique is used in order to find the coefficient of such system. It is a tool to identify the mathematical model of real world system, which can be identified from the experimental data.

System identification is successfully used in many area either in the technical field or in the non-technical field [19]. In the technical field such as in control system, it can be used to identify the system model which can then be used in several applications such as, simulation, prediction etc. In signal processing application system identification technique can be used to identify the system model. Such model can be used for the analysis of fault detection, linear prediction etc. In non-technical field, system identification techniques are also successfully used [25]. Some successful applications are related to environmental science, medical science, and economics in which system identification tools are used to identify the model, then used it for prediction, adaptive control system, and to enhance the scientific knowledge on identified objects. Section 2.4 covers how system identification tools can be used to identify the mathematical model of real world system.

2.2.2 Discrete Time Systems (Sampled Data)

The real world system are continuous system, discrete time system is obtained by sampling of a continuous system. It only consider input and output at discrete time points. It is sometime necessary to describe a system in discrete time model due to availability for measurement signal values at certain time period. For example, signal values are available only at discrete time because the continuous system is sampled at discrete time. Another example is when the system naturally described in discrete form such as number of production, sales figures, are naturally associated with certain day or week. Nowadays majority of the control object that are continuous time system, are sampled and control by digital (discrete time) controllers.

Typical reason of describing continuous system by a discrete time system is; controller is calculated as continuous time system and the requirement is to implement it in a computer as a discrete time controller. As it is described in [22 Ch.2, 4], if the input of continuous time system can be reconstructed by its sampled time value, the discrete time output representation should be possible to obtain that exactly can describe the sampled output. In discrete time system, the sampled data is obtained by computing the continuous signal with sampling train, which can be describe as an approximation, or alternative of the system [22], where only the interested discrete time instances value of input and output are considered, often uniformly distributed in time as,

$$u(k) = \; ... u(-1), u(0), u(1), u(2), ...$$
$$y(k) = \; ... y(-1), y(0), y(1), y(2), ...$$
$$k \; = \; ... - 2, -1, 0, 1, 2, ...$$

where, $u(k)$ and $y(k)$ are discrete time input and output, and k is the discrete sample point. $\mathcal{Z}$-transform is convenient to deal with discrete time signals, the impulse response for discrete time can be written as,

$$G(z) = \mathcal{Z}\{g(k)\} = \sum_{k=0}^{\infty} g(k) z^{-k},$$
$$U(z) = \mathcal{Z}\{u(k)\} = \sum_{k=0}^{\infty} u(k)\, z^{-k},$$
$$Y(z) = \mathcal{Z}\{y(k)\} = \sum_{k=0}^{\infty} y(k)\, z^{-k}.$$

Here $G(z)$ is the discrete time transfer function. If the input values are zero before $k = 0$, then $\mathcal{Z}$-transform of the output, $Y(z)$, can be written as

$$Y(z) = G(z)U(z). \qquad (2.2.2.1)$$

It can be seen from (2.2.2.1) that the discrete time system has similarities to the continuous time case. Everything that concerns to continuous time system also applies to discrete time if Laplace transform table are exchangeable to $\mathcal{Z}$-transform table [22].

The system can also describe in state space representation. In **State space form** representation, it describe the state of a system. In discrete time system, state space form can represent as a system of first order difference equation as;

$$x(k+1) = Ax(k) + Bu(k) \qquad (2.2.2.2)$$
$$y(k) = Cx(k) + Du(k).$$

Here, x is the state vector of n-dimension, and k is representing discrete time. $A, B, C,$ and D are the matrices of compatible dimensions. The linear system representation in state space form has several advantages, such as it treat SISO and multivariable system in a similar way. It is also an advantages that the system model can obtained directly as a physical modelling result.

2.2.3 Linear Model Structure

A system is an object driven by number of input and as a response it generate number of output [21]. Typically, it is a real world object that need to control by choosing input in such a way that output follow the reference. Linear system model can be describe in several ways, such as transfer function, differential equation, state-space representation etc. In discrete system, input and output are sampled at discrete time point k. Figure 2.2.3.1 shows the model block diagram of a linear system.

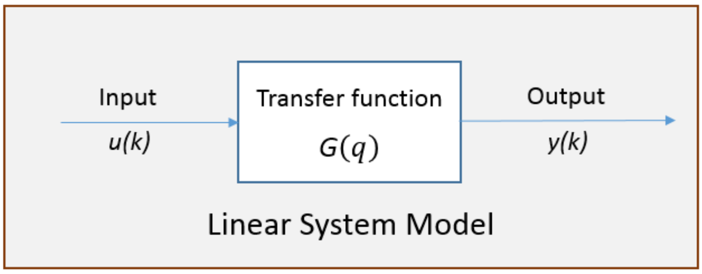

Fig 2.2.3.1 block diagram of a linear system

If the number of input and output is one then such system is called single input single output (SISO), if the input and output are multiple and output such system is known as multiple input multiple output (MIMO). A linear model can be often written using rational transfer function as,

$$y(k) = \underbrace{\frac{b_1 q^{-1} + b_2 q^{-2} + \cdots + b_m q^{-m}}{1 + a_1 q^{-1} + a_2 q^{-2} + \cdots + a_n q^{-n}}}_{G(q)} u(k), \qquad (2.2.3.1)$$

where, $y(k)$ is the output, $u(k)$ is the input, and $G(q)$ is the transfer function containing *m* and *n*, which are the degree of numerator and degree of denominator (represent $G(q)$ in q operator instead of $\mathcal{Z}$-transform as $G(z)$, the advantage of q-formulism is to express signal in time-domain instead of z-domain). The argument q^{-1} in (2.2.3.1) is a backward operator, which mean e.g., $q^{-1}u(t)$ contain the previous value that is, $u(t-1)$ etc. The numerator and denominator of (2.2.3.1) can also express as,

$$A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + \cdots + a_n q^{-n},$$
$$B(q) = b_1 q^{-1} + b_2 q^{-2} + \cdots + b_m q^{-m}.$$

The transfer function is also defined in term of degree of its numerator and denominator. It is called proper if the degree of numerator *m* does not exceed to the degree of denominator *n.* The transfer function is called strictly proper if the degree of denominator is greater than nominator. Therefore, 2.2.3.1 can be write as,

$$A(q)y(k) = B(q)u(k),$$
$$y(k) + a_1 y(k-1) + a_2 y(k-2) + \cdots + a_n y(k-n)$$
$$= b_1 u(k-1) + b_2 u(k-2) + \cdots + b_m u(k-m)$$

It can be rewritten as;

$$y(k) = -a_1 y(k-1) - \cdots - a_n y(k-n) + b_1 u(k-1) + \cdots + b_m u(k-m). \qquad (2.2.3.2)$$

The variable $a_1, \dots, a_n$ and $b_1, \dots, b_m$ plays an important role in order to understand the dynamics of the system. Section 2.2.4, and 2.2.5 shows how these variable can help to determine stability.

2.2.4 Poles and Zeros

To characterize the property of a system, poles and zeros plays an important role. Poles correspond to the general time response of the system, and essential for the behavior of the system. By definition, pole of a system means the eigenvalues of the system and pole polynomial means the characteristic polynomial for a matrix that is, $det(\lambda I - A)$. It shows the stability of the system, which is discussed in Section 2.3.4. In SISO system, poles are the zeros of denominator polynomial. According to theorem 3.5 [22], "*The pole polynomial for a system with transfer function matrix G(s) is the least common denominator of all minors to G(s). The poles of the system are the zeros of the pole polynomial."* Zeros is an s value in SISO system, which makes the transfer function zero. Thus in some sense, it describe the inverse dynamics of the system. The poles of the discrete time system are given by

$$e^{\lambda_i T},\ i = 1, \dots, n \tag{2.2.4.1}$$

$and,$

$$\lambda_i, \qquad i = 1, \dots, n.$$

Where, λ_i represent the pole of a continuous system, and T is the sampled time.

2.2.5 Stability

In control system, stability is the key concept [23]. It is also an important part to ensure either the system is stable of not. The system's stability can be ensured after getting the bounded output while applying bounded input, such system is called BIBO stable. Therefore, the eigenvalues of system's real parts plays an important role to check the BIBO-stability. Which can be identify as, if all eigenvalues lie in the left half plane (LHP), then the norm decrease exponentially and system is called stable system. However, if the eigenvalues lie in right half plane (RHP), the norm increase exponentially that represent an unstable system. According to definition 3.6 [22], stability region in a continuous time system is equal to the left half plane without imaginary axis. If all poles of a linear time invariant (LTI) system lie inside the stability region then the system is input-output stable. Figure 2.2.5.1 shows an example of a stable system, in which system's poles lie in the left half plane. For the initial value if the unstable is included, the output will however tend to infinity.

If the roots of the system are positive than the poles lie in LHP, which is the stable region, and it is consider as stable system. Otherwise, system is unstable if its pole value lie in RHP. These roots can be plotted in s-plane to see the stability. Figure 2.2.5.1 shows the pole placement by considering stable pole e.g., poles are -1 and -3.

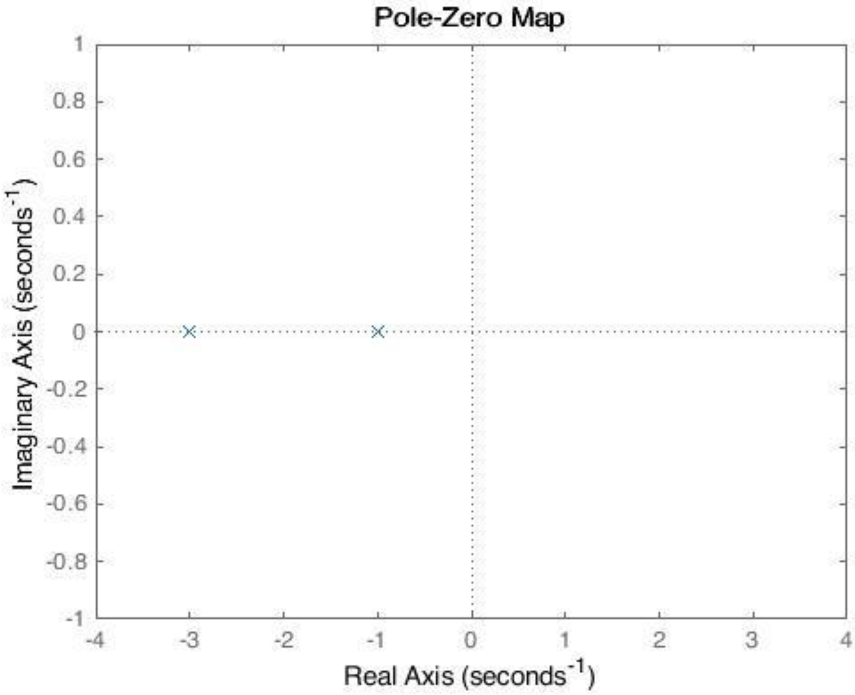

Fig 2.2.5.1, Left half plane (negative real axis) shows the stability region of continuous time system.

Mathematically, system's stability can easily understandable by converting it into time domain using inverse Laplace transform of transfer function G(s) as,

$$\mathcal{L}^{-1}\{G(s)\}.$$

Simplify the transfer function by using partial fraction of as,

$$G(s) = \frac{1}{(s+a)(s+b)} = \frac{A}{(s+a)} + \frac{B}{(s+b)} \tag{2.2.5.1}$$

$$\mathcal{L}^{-1}\{G(s)\} = Ae^{-at} + Be^{-bt},$$

where A and B are the gain of the system. If the impulse is feed when the roots are negative, then as time passes its response decrease exponentially which shows the stability of the system. It mean that after some time system would be at rest. However, if the pole are positive then as time passes, the response increase exponentially that lead the system towards instability, such system would never be at rest and become unstable till infinity.

In a discrete time system, the stability region is the interior of the unit circle. According to theorem 3.9 [22] "*A linear, time invariant system is input-output stable if its poles lie in the stability region"*. It is called BIBO-stable only if all the poles lie inside the unit circle. The pole of discrete time can be determined from the continuous time system. It is also described in Section 2.2.4 in (2.2.4.1). If (2.2.5.1) is converted into discrete time, by considering stable continuous time pole a and b as an example -1 and -3. The discrete time poles are 0.3678 and 0.0497. The discrete time plot is also shown in Figure 2.2.5.2.

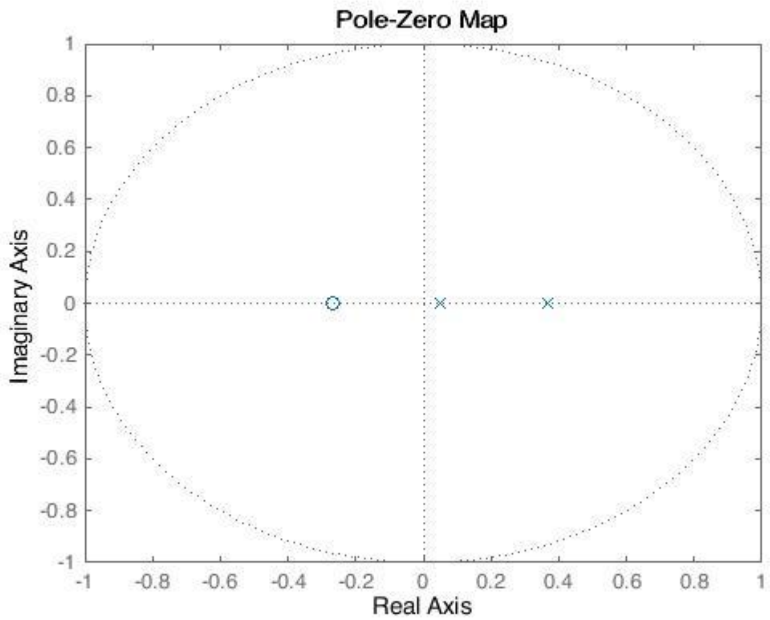

Fig 2.2.5.2, Poles(x) inside the unit circle show the stability region of discrete time system.

It has discussed above that, stable continuous time system has pole placement in *Left half plane*, and in discrete time pole placement inside unit circle. However, response of the system also depend on pole placement. As mentioned in [22 Ch. 3], if the pole placed away from origin (inside stability region), then the response would be fast. Otherwise, system will give slower response if their pole are close from origin (inside stability region).

2.3 Disturbances

A real system is always effected by disturbances, these disturbance could affect system's output. The task of controller is to feed the control input to the system so that output would be the same as desired, despite the influence of different disturbances. Mathematically, system having no external effects can be written as,

$$y = Gu. \tag{2.3.1}$$

In practice, it is fortunate to get measured signal from (2.3.1) without the influence of noise and disturbances. Figure 2.3.1 shows the block diagram of the system 'G' whose output y is influenced by disturbances w, and noise n.

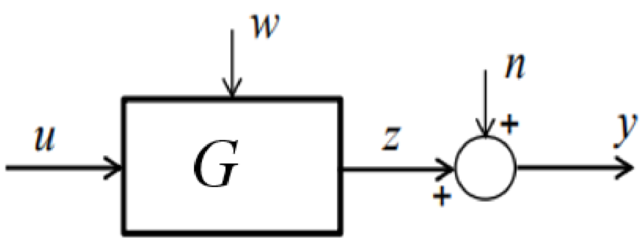

Fig 2.3.1 Block diagram of system influenced by external factors

Measured (output) signal deviate from the system having same input, in different ways. These output deviation depend upon disturbances. The real output can be written as,

$$y = Gu + w + n. \tag{2.3.2}$$

The additive term w is the disturbance signal, and n is the noise. From a system prospective, w is the additional input to the system. There are many different factors in addition to control input u that influence the output. It could be variation in plant and actuators, model errors, load variation, or the process is disturbed by variation in raw material. An example for a disturbance is; number of people in the room influence the room temperature, here the increase in people effect the room temperature, which is acting as disturbance. It could also be a measurement error that is, the output measurement influenced by noise, or the sensor that measure the output effect by external noises.

2.4 System Identification

It is described in earlier Sections about dynamic system (plant), and need of mathematical model in maintenance. Nowadays, majority of the industrial system work in a closed-loop, which is controlled by a feedback controller. Due to feedback, the controller vary (increase or decrease) the input by computing the reference and output signal. The variation in the input signal may also depend upon external and internal disturbances. In system identification, control signal and measured output signal can be used to identify the system. Figure 2.4.1 shows the block diagram for system identification.

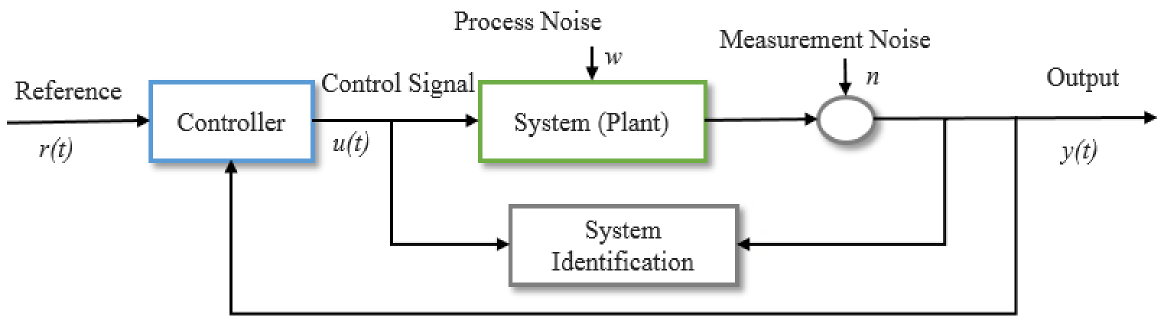

Fig 2.4.1 block diagram for system identification

The system (plant) generally process input data and transformed into a processed data. To perform theoretical analysis of identification, it is necessary to introduce assumption of the data. The word ‘plant’ or ‘system’ is use to denote the mathematical description of the process, it must be regarded as fixed, and its properties cannot be changed by user. In real scenario the system might be unknown. System Identification techniques are useful to apply in such scenario to identify the unknown system. Therefore, it is not necessary to know the dynamics of the system. However, the key problem in system identification is to identify suitable model structure, within which a best model is to be found [26].

Sometimes nonparametric model are used for identification, such model is described by a function, table, or by a graph. Frequency plots (Bode plots), and impulse response are the example of nonparametric model. Whereas, parametric model in general is used to determine the unknown parameter vector by mapping from the recorded data. Parametric models are relevant to deal in many cases, such models are characterize by parameter vector denote by θ. In this method, model structure is obtained when parameter vector varies over some set of feasible values. It is described in [24 Ch. 2] that in system Identification, the concept of experimental condition is important which determine when data are collected during the process. It include sampling interval, selection and generation of input signal, feedback in the process, preprocessing of data before parameter estimation etc. Once the data is collected, several choices of identification and model structure can be tried on the same data set until a satisfactory result is obtained.

As described in Section 2.2.1, system modeling is very important in many area and in industrial applications, there are two ways to build the mathematical models; white-box modeling, and black-box modeling [24 Ch. 2].

White-box modeling: It is a mathematical approach. White-box modeling is also known as mathematical modeling. In this modeling technique basic laws of physics are used to create the model (describe the dynamic behavior of a process). When the system is completely known, and possible to construct it completely from prior knowledge and physical insight [26]. For a complicated system, this modeling technique might be very time consuming, or even impossible.

Black-box modeling: It is an experimental approach, and also known as system identification. In this modeling technique, some experiments are performed on the system; measure the input and output data and find the mathematical model that is then fitted to the measured data. In black-box modeling it is typically assumed that there is no physical insight available or used. However, this method increase the computation for prediction and optimization [27].

In many real world application, the process of system is complex and not possible to obtain the mathematical model directly from the physical laws (using principles, e.g., Newton's laws, balance equations). To use a mathematical modeling approach to identify the real world system might not give the reasonable model due to different physical quantities (e.g., spring constants, mass etc.) link to each other. However, black-box modeling are easy to construct by measuring the data and identify the mathematical model. But the model obtained by system identification has limited validity such as limited to certain input data type, certain process etc. Also, recorded data might disturbed with measurement noise which must be taken into consideration. The identified model might give little physical insights, but parameters of the identified model typically have no direct physical meaning.

2.4.1 Linear Regression

There are different popular linear model structures that are used to represent unknown system such as, FIR, ARx (Auto Regressive with exogenous input), ARMAX (ARx with Moving Average), and OE (Output Error model), which differ with noise model. As discussed in Section 2.4, system identification increase the computation for prediction and optimization, it is worth taking to choose such model that is less complex and able to identify the model that is then fitted to the real system. Linear regression is the simplest type of parametric model. Its origin from Gauss (1809), which used this technique to calculate planets and orbits [24 Ch. 4].
The linear regression model structure can be write as

$$y_m(k) = \varphi^T(k)\theta, \tag{2.4.1.1}$$

where $y_m(k)$ is the measureable output (at sample k) and also known as regressed variable, $\varphi(k)$ is known quantities of n-vector it is also called regression variables (φ^T is the transpose of n-vector known quantities), and θ contain the unknown parameters of n-vector it is also called parameter vector. The estimate can improve by using residuals, also called equation errors as

$$\varepsilon(k) = y(k) - \varphi^T(k)\theta. \tag{2.4.1.2}$$

The result from Eq (2.4.1.2) is the calculated error of measured and estimated output. Therefore, it make sense to choose θ that make equation error $\varepsilon(t)$ as small as possible. Due to disturbance, noise, and model misfit, it is good to consider more data (greater than n-vector).
Therefore, by including error term (2.4.1.1) can be written as,

$$y_m(k) = \varphi^T(k)\theta + \varepsilon(k).$$

This can be write in matrix notation as

$$Y = \Phi\theta + \varepsilon. \tag{2.4.1.3}$$

The least square (LS) estimate of parameter vector θ is defined by vector $\hat{\theta}$ that minimize the loss function

$$V(\theta) = \frac{1}{2}\sum_{k=1}^{N} \varepsilon^T \varepsilon = \frac{1}{2}\|\varepsilon\|_2^2$$

If $\Phi^T\Phi$ is positive definite then $V(\theta)$ has unique minimum point;

$$\hat{\theta} = (\Phi^T\Phi)^{-1}\Phi^T y. \tag{2.4.1.4}$$

If $\Phi^T\Phi$ is not invertible then $V(\theta)$ has infinitely many minima.

2.4.2 ARx Model

Autoregressive model with exogenous input (ARx) is one of the model that is convenient and widely used in identifying the unknown system. It is also known as Autoregressive with exogenous variables, where exogenous variable is the input terms. It is a linear representation of dynamic system in discrete interval of time. Due to its ease, the results is not too difficult to obtain. To understand the model, consider a system driven by input signal $u(k)$, which generate its output $y(k)$. A general model of ARx model can be seen in Figure 2.4.2.1

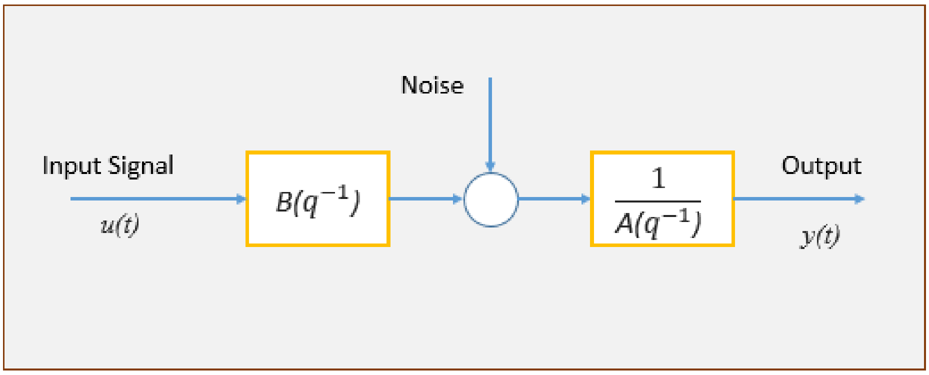

Fig 2.4.2.1 Block diagram of general ARx model

Mathematically, it can be modeled as

$$A(q^{-1})y(k) = B(q^{-1})u(k) + e(k), \qquad (2.4.2.1)$$

where for given, *n* and *m* >0, $q^{-1}$is the delay operator, and has,

$$A(q^{-1}) = 1 + a_1q^{-1} + a_2q^{-2} + \cdots + a_nq^{-n}$$

$$B(q^{-1}) = b_1q^{-1} + b_2q^{-2} + \cdots + b_mq^{-m},$$

it contain fixed *n* and *m*, but having unknown coefficients $\{a_1, \dots, a_n, b_1, \dots, b_m\}$. The residuals $e(k)$ are small in some sense, but it is consider as unknown otherwise. Therefore, the overall system can be written as,

$$y_m(k) = \varphi^T(k)\theta + e(k). \qquad (2.4.2.2)$$

In linear regression, regression variable $\varphi(k)$ contains the previously known values of input and output,

$$\varphi(k)^T = [-y(k-1), \dots, -y(k-n), u(k-1), \dots, u(k-m)]^T,$$

and the unknown parameters θ can be written as,

$$\theta = [a_1, \dots, a_n, b_1, \dots, b_m]^T.$$

With the help of Least Square, θ can be identified as described in Section 2.4.1.

2.5 Model Validation

In System Identification, the primary goal for modeling of a physical process is to predict the parameter from a given set of input, to determine how well the identified model fit the real system. It is commonly referred as model evaluation or model validation [28]. Model validation is an important aspect after the determination of model. A model with additional consideration of parameter can lead to unnecessary computation to find the parameter estimation value, as well as for using the identified model. The less parameterized model can lead to inaccuracy. A crude low order model is usually used in stabilizing regulator, whereas detailed or complex model are often necessary if the model is expected to give the physical insight into the system [24]. Correct model order are based on statistical assumption from which the data come. The two basic approach to sure that the identified model is large enough to represent the true world system are; plot the output of real plant and identified model. The second approach is to use prediction errors information.

From the first approach, the identified model is excited by measured input without the addition of disturbance, and the output is expected to be the same as real data output. Such model structure, can be written as;

$$y_m(t) = G\left(q^{-1}; \hat{\theta}\right) u(t), \tag{2.5.1}$$

where, $y_m(t)$ is the model output excited by measured input, without the addition of disturbance. $G\left(q^{-1}; \hat{\theta}\right)$ is the function of estimated parameter vector $\hat{\theta}$, and argument q^{-1} is the shift operator. The model structure in (2.5.1) describe a general linear function, and in most cases it will be of finite order [24 Ch. 6]. The model is consider to be good if it reassemble to the measured output. When the data is noisy, the measured and identified system output should differ from each other [24 Ch. 11].

From the second approach, the difference between estimated output and plant output is due to prediction error $\varepsilon(t)$. It is also called residuals which is evaluated at the parameter estimate $\hat{\theta}$. In this approach, model determination is tied with model structure and identification method, where disturbances are directly modeled. Therefore, this method can be used to validate identified system model [24 Ch. 11].

3 Process and Results

In Ovako's ring rolling process, we have worked on one of the system which we called as plant that equiped with *tool A and tool B*, these tools are in different shape. These tools can be replaceable, and the operator replace it according to the type of ring to be produced. The movement of the tools are derived with hydraulic system.

Tools are moved with piston movement attached to the hydraulic cylinders. During the ring production, two tools push the rotating preform to be in the ring shape. Otherwise, if the speed of any rotating tool is high then the rolling ring would be uncontrollable. These tools (*tool A and tool B*) are designed to help the process stability (*tool A* is further connected with *tool x, which is replaceable and replace according to the type of production*) However, the *tool A* utilize more as the rotation of ring force it to displace away. In our work, movement of *tool A* is considered as plant, in which input is a control signal that control the position, which is directly connected to hydraulic cylinder. The output of the plant is considered as position.

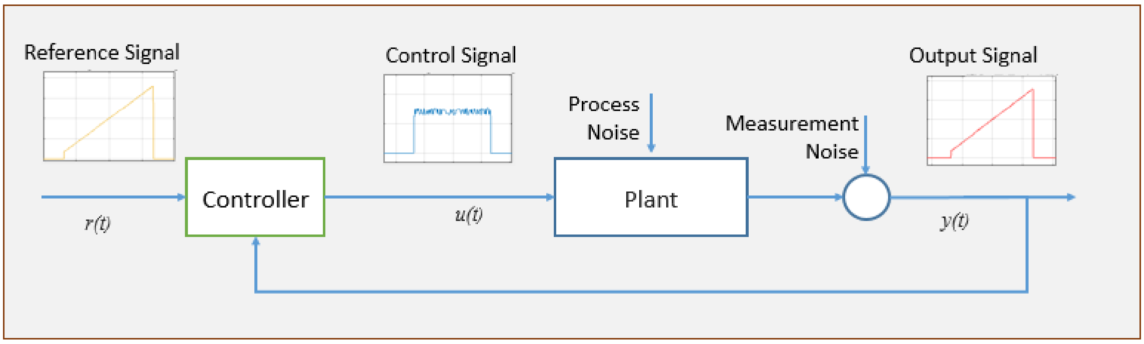

Fig 3.1 Block diagram of Plant

Figure 3.1 show the block diagram of plant. The whole system is a closed loop system, and is controlled by PLC controller. The PLC control the valve of hydraulic system, due to the movement of valve's spool position, the fluid flow and move the hydraulic actuator. *Tool A* is directly attached to the hydraulic cylinder, and its movement directly control the movement of *tool A*. The plant's output is measured with the help of position sensor, and feed it back to the controller. The controller adjust input by comparing the reference signal $r(k)$ and output signal $y(k)$ during the whole process. DT model would be identified with the help input signal, output signal, and system identification tool. Both digital-space and physical-space output would analyze together over a series of data and analyze the output deviation.

3.1 Data Cleaning

There are numerous signal have been generated during the ring production. Therefore, it is necessary to narrow down the area and focus on specific area or tool that play vital role in production such as, main roller, vertical roller etc., and analyze its twin model in CPS. In this thesis work, we emphasized on one of the tool. As mentioned in Section 3, *tool A* and *B* are used to push the rotating metal to be in the ring shape, and keep the process stable. In digitization mode, there movements are recorded by position sensors. There are also other sensors connected to the *tool A* and *B*, which give the real time process information such as, pressure, speed etc. To find the DT model of the plant, system identification tool would be used. In which input signal and output signal are required to identify the system. As the movement of *tool A* is derived by hydraulic cylinder. The sensor record all movement such as when it is at rolling, at waiting position, and stop position. These control and position signals contain value even when there is no movement, which need to be remove. Calibration is also performed before every new type of ring production, during the calibration mode, the whole ring mill is assumed to be in calibration mode.

Before the mill start ring production, the *tool A* move from initial position to the waiting position. Also when the rolling process start, other (rolling) tools effect *tool A* as process noise/disturbance. Therefore, to identify the plant mathematical model with less influence of other parameters, it is better to consider only before production signal (when *tool A* start from initial position to the waiting position). To study only desired signal, it is obligatory to remove the calibration, rolling process, and digital signals.

3.1.1 Remove the Calibration Signal

The calibration signal would not be used in the DT. Therefore, it is good to remove all the calibration signals from the available data. However, it is also necessary to sure that the time information should not remove. Otherwise, if the original time information removed, then it would be difficult to trace back if necessary (for example, to trace at what instance the tool starts deviate). The calibration procedure begin with the calibration button, a digital signal with constant amplitude also generate simultaneously, and it stop when the calibration mode end. Position sensor is used to record the tool movement. Figure 3.1.1 shows the plant output signal including calibration signal, as well as the corresponding digital calibration signal.

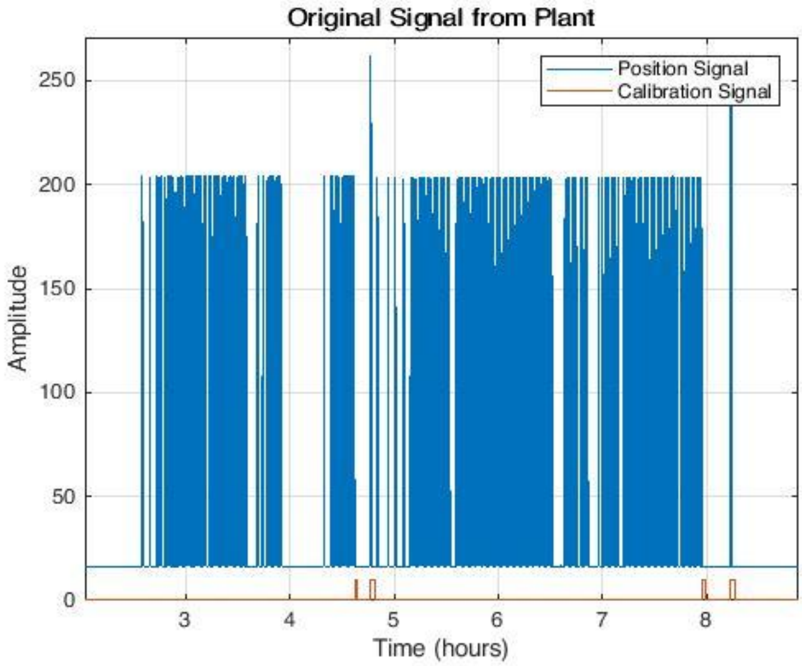

Fig3.1.1 Plant output position signals

Figure 3.1.1 shows the *tool A* position signal (blue graph) and digital calibration signal (red graph). The high peak signal in the blue graph represent the tools calibration signal, when operator pressed the calibration button, corresponding digital signal can be seen with red line. There are different way to remove the calibration signal, without losing the time information. Such as use threshold condition e.g., set the value to zero that are above, or in other way use digital signal information. However, we observed that by using calibration digital signal, it is more easy and accurate because the tool preproduction movement varies, depend upon production.

A calibration digital signal only generate during calibration otherwise it is become zero. Also the calibrated signal triggered the digital signal with constant amplitude, we divided it with the same amplitude to get unity signal. By the help of logical operator the calibration signal can be removed. We use the logical NOT operator to invert the digital signal, and then multiply it to the position signal. Figure 3.1.2 shows the resulting signal.

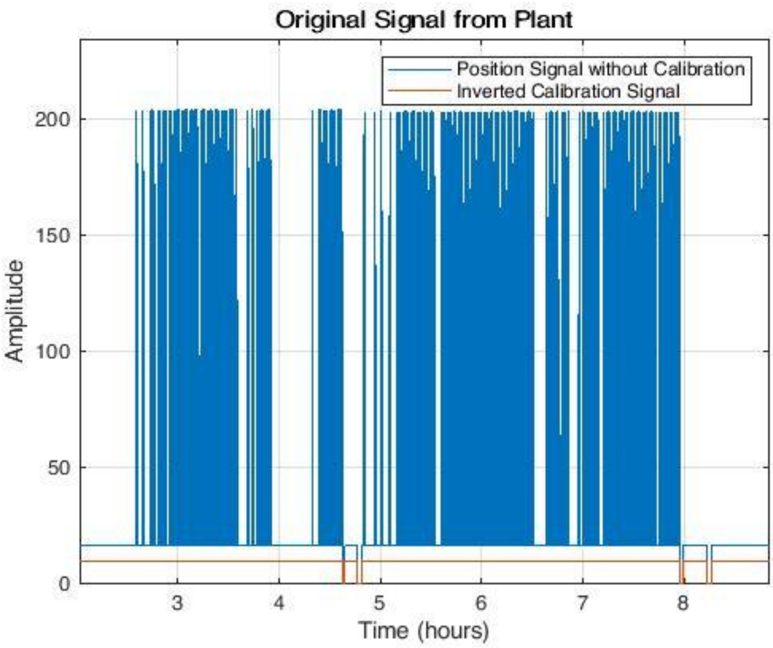

Fig 3.1.2 Plant position signals without calibration signal

It is shown in the Figure 3.1.2 that the calibrated signal has been removed, and the graph now only contain ring signal. However, the overall position signal also contain the digital signal (when *tool A* is at rest or it is at waiting position) and the production signal (when the ring starts rolling). It can be seen in Figure 3.1.3.

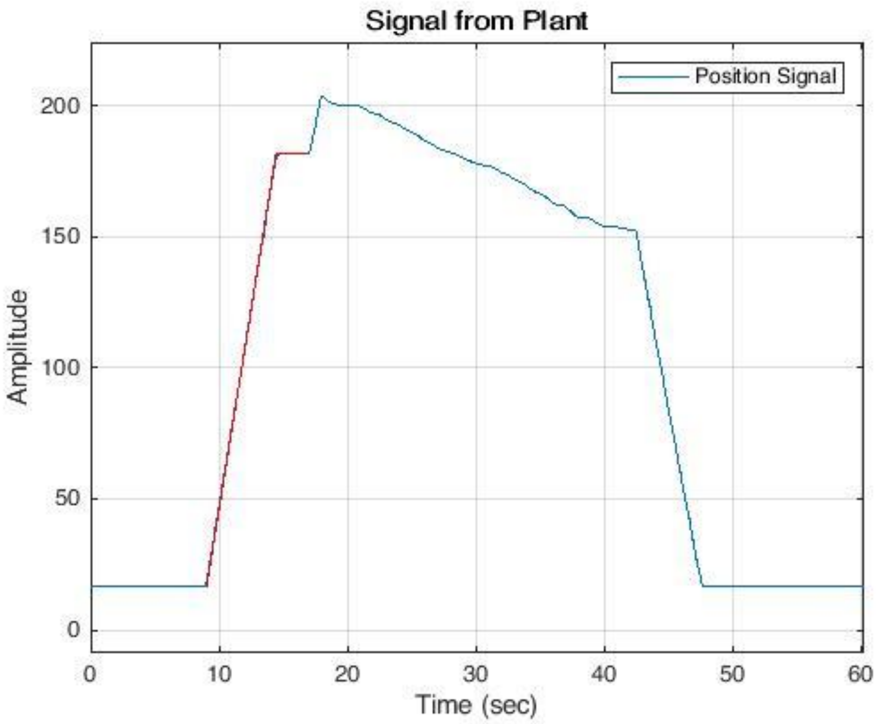

Fig 3.1.3 Plant position signals

Figure 3.1.3 shows the output position of the hydraulic cylinder that drive *tool A*. It has initial and waiting position, which are constant. The red line in the graph represent the preproduction signal, and blue line represent the overall signal. However, we are only interested in the preproduction signal, without having digital part.

3.1.2 Clean process and digital signal

There are different digital signals generate in the ring mill, such as calibration signal, insequence active, process active etc. Insequence active signal is useful as it became zero during the preproduction, otherwise it is one. Figure 3.1.2.1 shows the input, output, and insequence active digital signal.

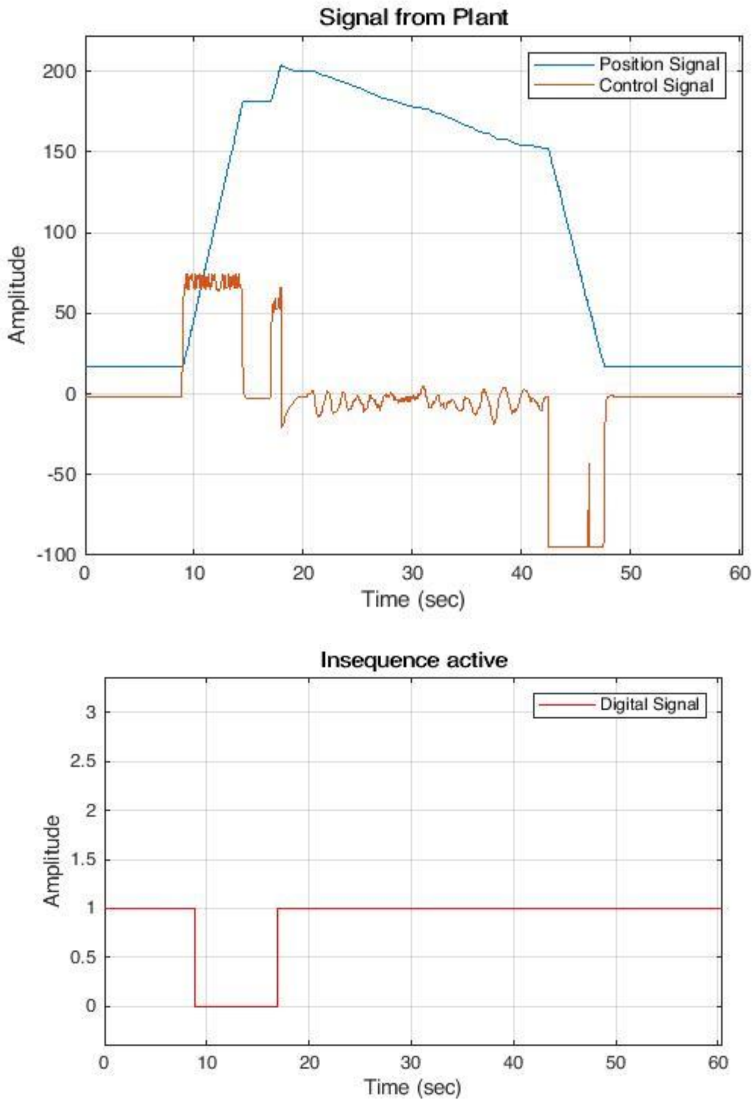

Fig 3.1.2.1 Upper graph shows position and control signal, lower graph shows the digital signal

The insequence active signal is at unity all the time except during the preproduction. As the target is to extract only the preproduction signals, insequence active digital signal is convenient. It can be inververted by using NOT logical operator, to get only the before production position and control signal. Once the inverted signal is generated, multiply it with the position and control signal. Hence, the resulting signal only contain preproduction signals. Figure 3.1.2.2 shows the resulting signal.

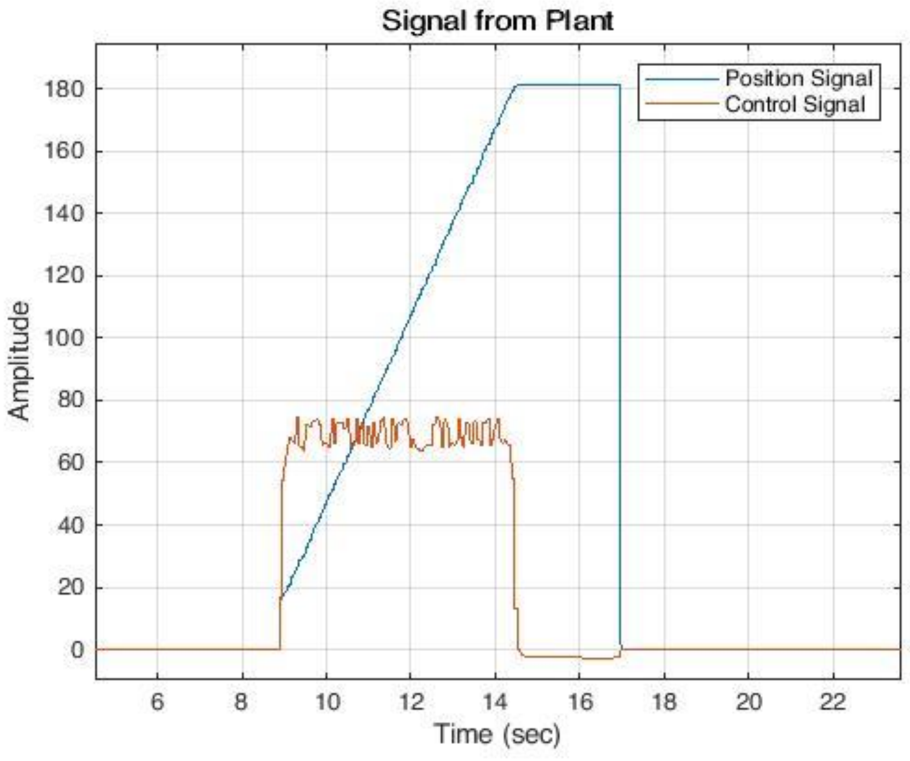

Fig 3.1.2.2 Before production position and control signal

It can be seen from the Figure 3.1.2.2 that the resulting signal also contain digital part, which would also need to remove. If it is removed by using the threshold value for each signal (e.g., pick value above zero) than each resulting signal would be in different length, due to different starting and ending value of each signal. It can be seen in Figure 3.1.2.3.

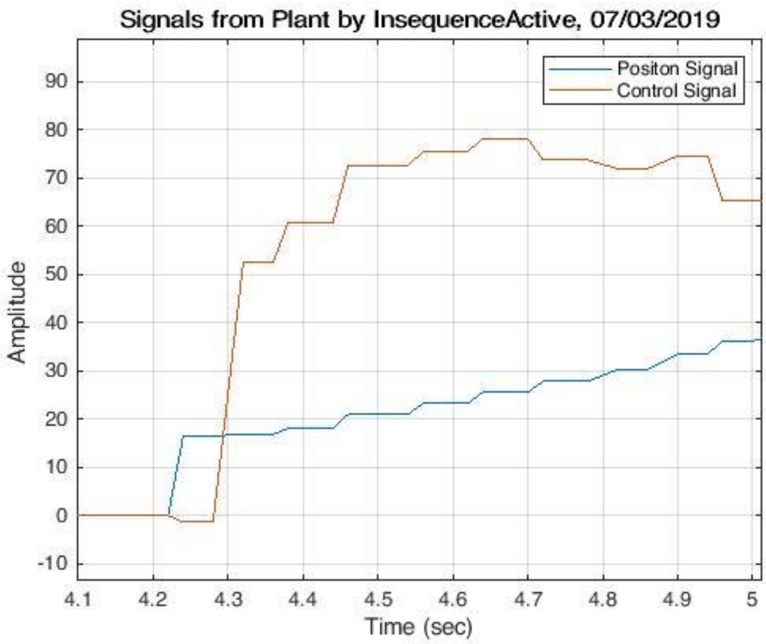

Fig 3.1.2.3 Different start value of position and control signal

Figure 3.1.2.3 shows the different starting value of position and control signal. However, it can be solved by considering one signal as a master signal, apply the threshold condition, and use same index value for other signal. In our case, we consider control signal as a master signal, apply the threshold condition to pick value above zero. Record the threshold index value in a separate array, and used those values for position signal to remove its digital part. Figure 3.1.2.4 shows the resulting signal.

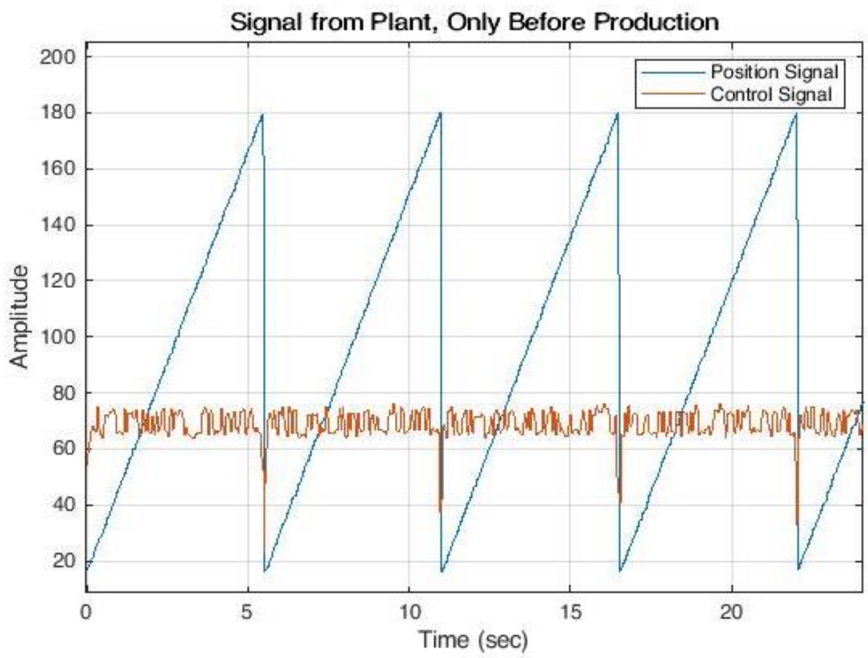

Fig 3.1.2.4 Different start value of position and control signal

Figure 3.1.2.4 contained only before production signal. These saw shaped output signal (in blue line) represent the pre-production tool movement of each ring production. Every depth in both signals (blue and orange line) represent the start point of tool movement and every edge peak (in blue line) represent the stop point of tool movement. The cleaned input and output signals would use to find the twin model of the plant. These signals are also used to power the DT in CPS system.

3.2 Sampled Data

Iba system is used to process data acquisition. Its analysis consist of smoothly adjusted hardware and software, and allow to record and analysis of measurement data. Signal from plant is served to iba system, which provide several features such as data analysis, data optimization, and data acquisition. The analyzer feature allow to export data in different format with desirable sampling rate.

The plant is controlled by programmable logic controllers (PLCs). The output is measured via sensor and feedback the information to controller, Figure 3.1 shows the block diagram of the plant. The controller data is transferred to iba system that allow to analyze, record and process measurement data [29]. The data transfer from controllers offers several interfaces such as RJ-45 Ethernet jack, which transfer data through packets [30]. Due to Ethernet nature, the transfer time for sending packet varies [31]. Ethernet does not reserve bandwidth, and temporary Ethernet delays may cause the PLCs to close the connection.

The recorded data can be analyze by ibaAnalyzer. The data is recorded when the plant is in running condition. It saved data file by file from controllers, and sensors to iba system. As the interesting part is to analyze longer series of data, it would consume much longer time to analyze signal file by file. Since IbaAnalyzer allow to export data to COMTRADE format, which is an IEEE C37.111 specified file format for power system. The exportation of data would also require data resampling, which is selectable by user.

The reason to choose COMTRADE is because this format is used for exchange medium and data files, use for interexchange of simulation, fault, and test data for electrical power system. The format provide easy interpretable form for use in exchanging data for storing oscillography and status data related to transient power system disturbance [32]. As the data files may contain both numerical and text data, the format allow to store either in binary or ASCII format. Since the advantage of ASCII numbers is that it can be interpreted by human, and by standard computer software and hardware. In our work we use ASCII format, to export the data with sampling interval of 0.02s as it give good description of recorded signal.

3.3 System Identification using MATLAB toolbox

As the tool is in moving condition during the production and pre-production mode, as well as it controlled in a closed loop environment. There are several parameters influence the ring rolling process (as process noise) and effect the output/performance. In the ring rolling, some of the process noise influenced as an additional input, which mostly due to variation in oil pressure, variation in the oil temperature, etc. The additive term of process and measurement noise account for the fact that control input and output may vary. Therefore, the goal is to identify the model that fit best to the tool's system. In the rolling process, each ring is produced one after another like a cyclic process. We have divided the data into small chunks, and pick a chunk/cycle and identify system. To get the system model for *tool A*, black box modeling is suitable to find the system equation as it is explained in Section 2.4. Figure 3.3.1 show the block diagram for system identification, in which plant is *tool A*.

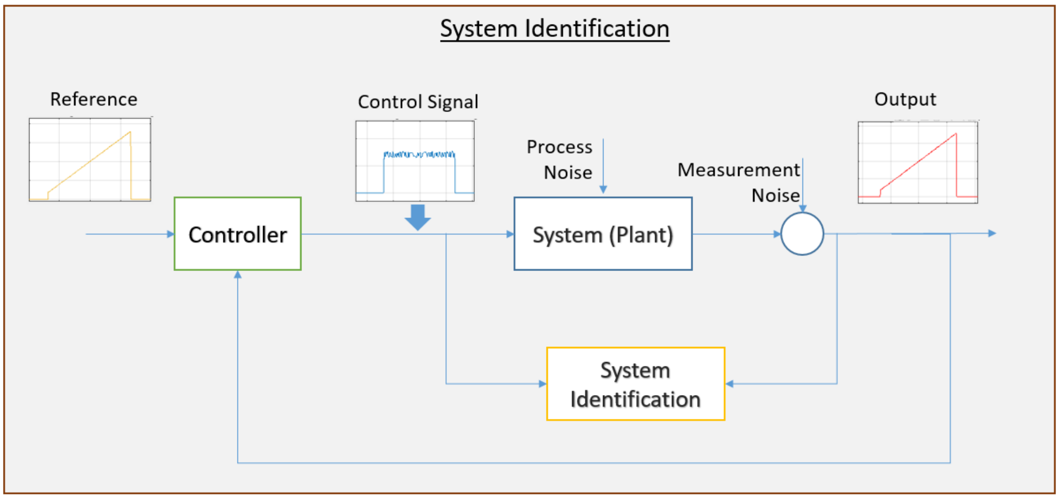

Figure 3.3.1, Block diagram of the system to find the system model.

There are different ways to find the system equation, Matlab *System identification* toolbox is one of them. It provides Matlab functions, Simulink blocks, and an app for construction mathematical models of dynamic systems from measured input-output data [33]. From the process knowledge, control and output signal were selected to identify tool's system.

As the data is recorded in iba system, it is also necessary to export the data that can be readable by Matlab. An automated process is written in Matlab, in which each file is processed and clean the unwanted signals (mentioned in Section 3.1). Once the data is cleaned, we have separated it into two files. One is for training the model, and other is for validate the model. Both files were fed to Matlab's *system identification* toolbox to identify the model of plant. If the identified model do not fit to the real system output, one can go back to data processing block to select another/better data. On the other hand, if the identified model fit best to plant output, than it can be used as a digital-twin model. Figure 3.3.2 show the flowchart that we have used for identifying the tool's model.

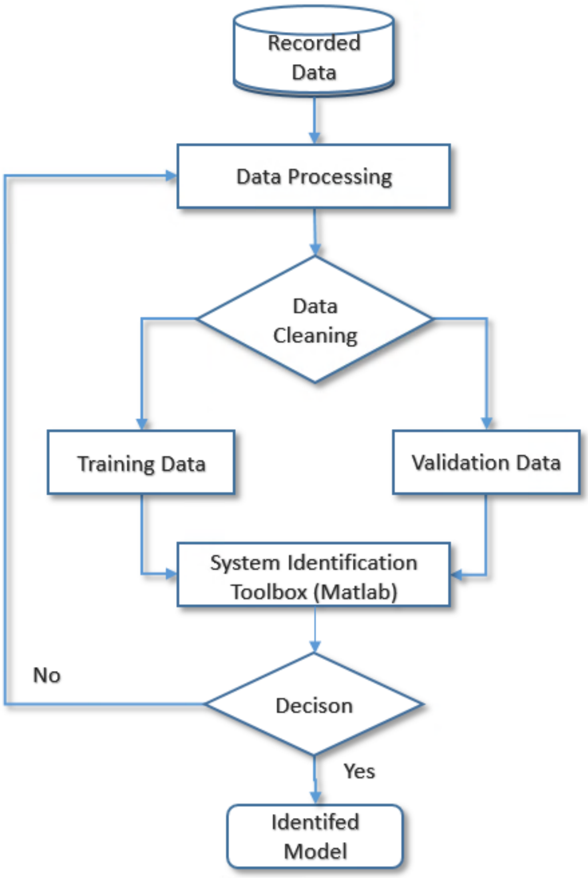

Figure 3.3.2, Flowchart to find the system model using System Identification.

The toolbox has a graphical user interface (GUI), which allow us to import data in time domain or in frequency domain. However, as we have recorded data over a period of time therefore time-domain option is appropriate in our project. The data that we have been processed and separated for validation and training, we can use GUI interface to import it. We have used training data to identify the mathematical model of *tool A*, whereas validation data is used to validate the identified tool model. We used different toolbox's option for identification. Figure 3.3.3 shows an overview of the *system identification* toolbox.

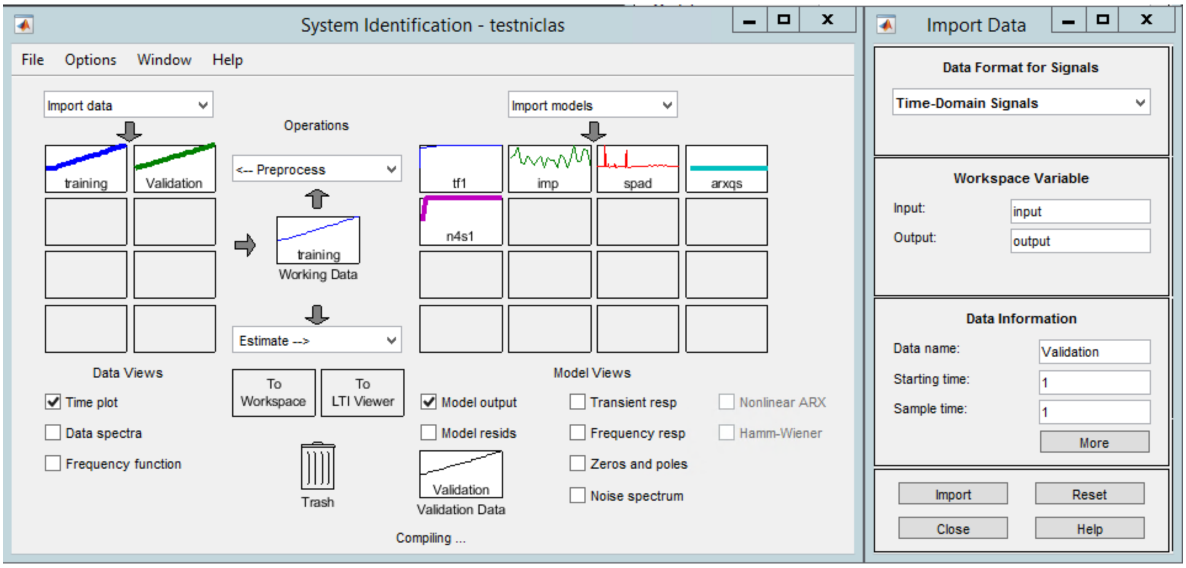

Figure 3.3.3 MATLAB app window for System Identification.

With the help of toolbox, we have identified the discrete time system by using different linear and non-linear model. As the given data is in discrete form, the estimated tool we found most suitable for system identification were; discrete time transfer function, discrete time state space model, and autoregressive exogenous (ARx) model. Figure 3.3.4 shows the estimation result of tool's output $y(k)$ and simulated model output $y_m(k)$.

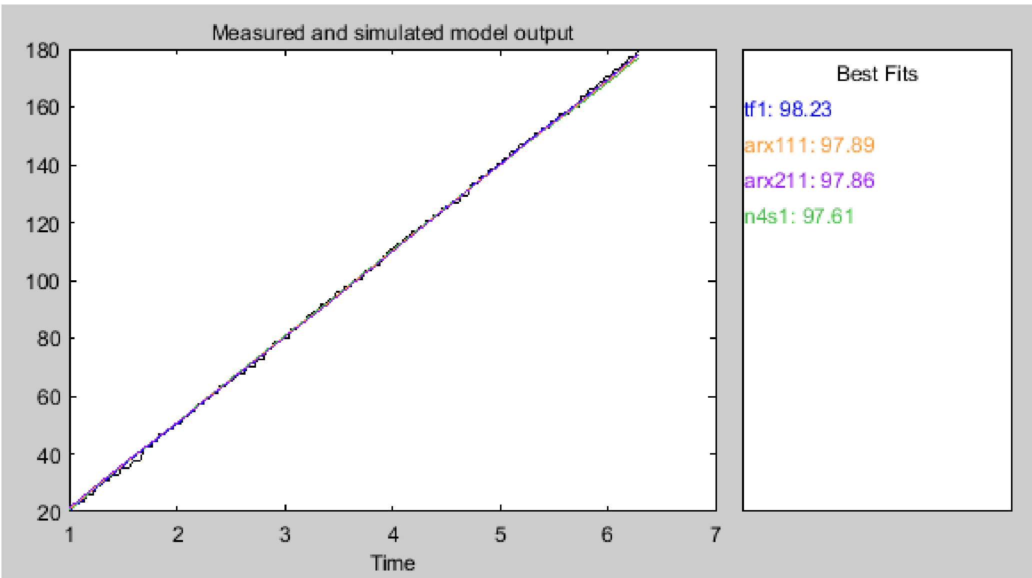

Figure 3.3.4 Simulated and measured result of Matlab system identification toolbox.

Figure 3.3.4 shows the estimated result and approximation percentage of state space, transfer function, ARx model. Table 3.3.1 shows the identified mathematical models.

Identified Model	Model Name	Mathematical Model	Fit Approx. (*Percentage*)
tf1	Transfer Function model	$G(z) = \frac{0.008572\ z^{-1}}{1 - z^{-1}}$	98.23%
arx111	ARx model *(polynomial form)*	$y_m(k) = y(k-1) + 0.008478u(k-1)$	97.89%
arx211	ARx model *(polynomial form)*	$y_m(k) = 0.764y(k-1) + 0.764y(k-2) + 0.008478u(k-1)$	97.86%
n4s1	State-Space model	$x(k+1) = 0.9998x(k) + 1.331e^{-5}$ $y_m(k) = 655.6x(k)$	97.61%

Table 3.3.1 Identified models from Matlab system identification toolbox

Table 3.3.1 shows the different models that we have used to identify tool's system, and their fit approximation in percentage. It can also observed from figure 3.3.4 that the identified models output are approximately fit to the real output. Therefore, we consider best fit model as transfer function, as it showed the fit estimation of 98.23% (Table 3.3.1). By using the identified system (DT), we feed it with the recorded input signal u, and compare the twin output y_m with the tool's output y. The simulated result of the identified model and the actual output can be seen in Figure 3.3.5.

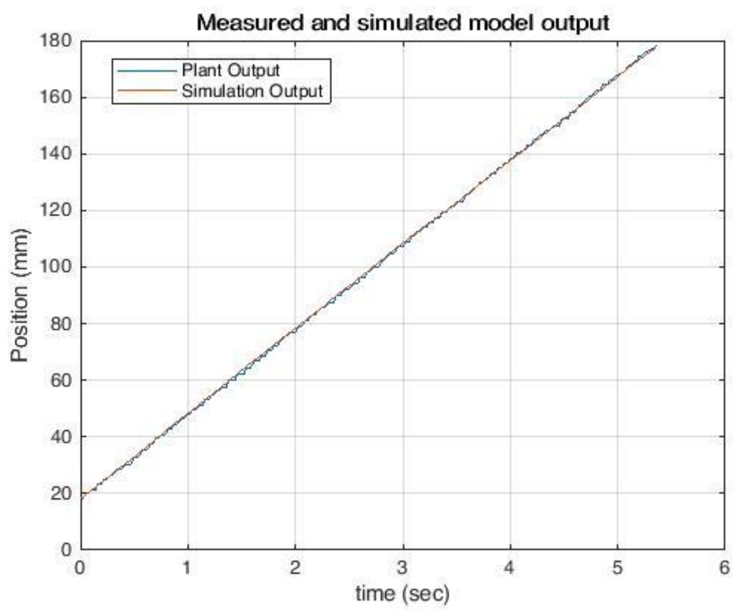

Figure 3.3.5 show the simulated and measured output

It can be seen from the Figure 3.3.5 that the output from the twin system has a good estimation fit, and it give quite accurate discription over a plant's output.

3.3.1 Analysis of Identified System

Once the system is identified using Matlab's toolbox, it is good to analyze the identified model, such as stability of the system. Since the input and output data are in discrete time, therefore the stability can be verified if the poles lie inside the unit circle, identified transfer function model can write as,

$$G(z) = \frac{0.008572z}{z-1}. \tag{3.3.1}$$

The denominator roots of the transfer function are the poles of the system. Hence, the determined pole is 1. In Matlab, the pole of the system can be determine by using eig(*sys*) command, where *sys* is the transfer function (3.3.1). The reason of having pole equal to 1 is due to the process. In other words, the identified system model is unstable as it is identified during an integrating process.

To analyze the system behavior, a step response is a useful tool. Mathematically, it can be analyze by giving a step input to the identified system. Unit step function in $\mathcal{Z}$-Transform can be written as,

$$U(z) = \frac{z}{(z-1)} \tag{3.3.2}$$

Multiply identified model $G(z)$ with the step $U(z)$ as,

$$Y_m(z) = G(z)U(z) \tag{3.3.3}$$

Therefore,

$$Y_m(z) = \frac{0.008572z}{(z-1)} * \frac{z}{(z-1)} \tag{3.3.4}$$

Using the property of inverse $\mathcal{Z}$-Transform to see the system in time domain, Eq (3.3.4) can be written as,

$$y_m(k) = 0.008572(k+1). \tag{3.3.5}$$

Here, the sample interval is k. It can be analyzed from (3.3.5) that when the identified discrete time model processed with the step input, the output will increase with a constant amplitude like a slope. As the discrete time k passes the output increases continuously. Figure 3.4.1.1 shows the output.

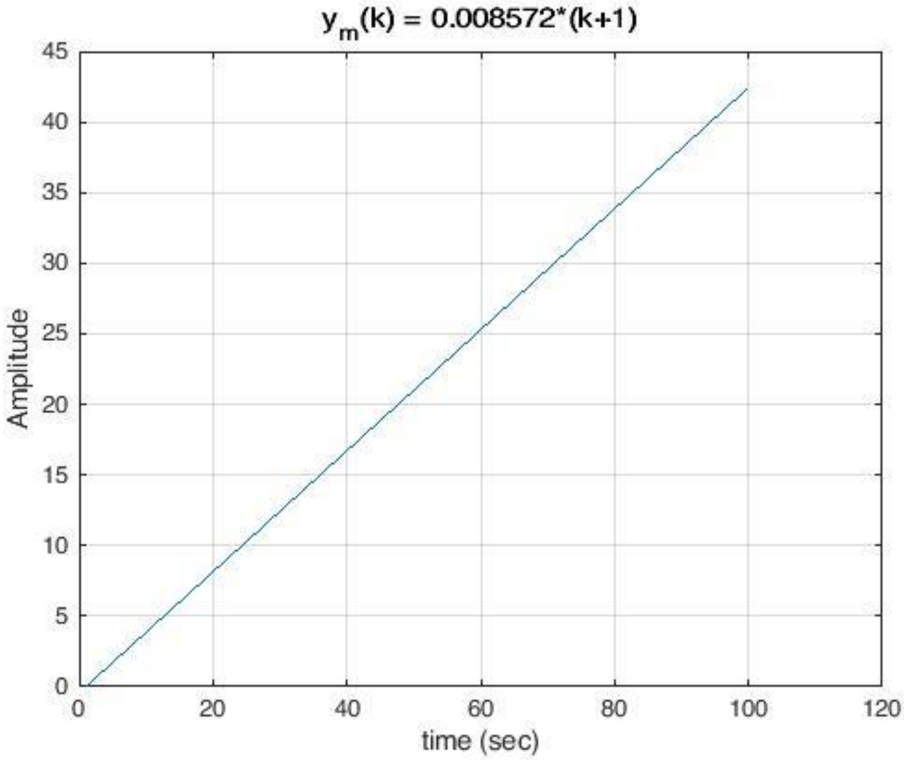

Fig 3.3.1.1, Step response by plotting Eq 3.3.5

Difference equation for the identified system (3.3.1) can be written as

$$y_m(k) = y(k-1) + 0.008572u(k-1). \qquad (3.3.6)$$

The discrete time interval is 0.02s, as it is mentioned in Section 3.2 that the recorded ibaAnalyzer data is exported to Matlab with the sampling rate of 0.02 sec. By computing the (3.3.6) in Matlab with the step input, we get the step response by using step function. Figure 3.3.1.s shows the step response of the identified system.

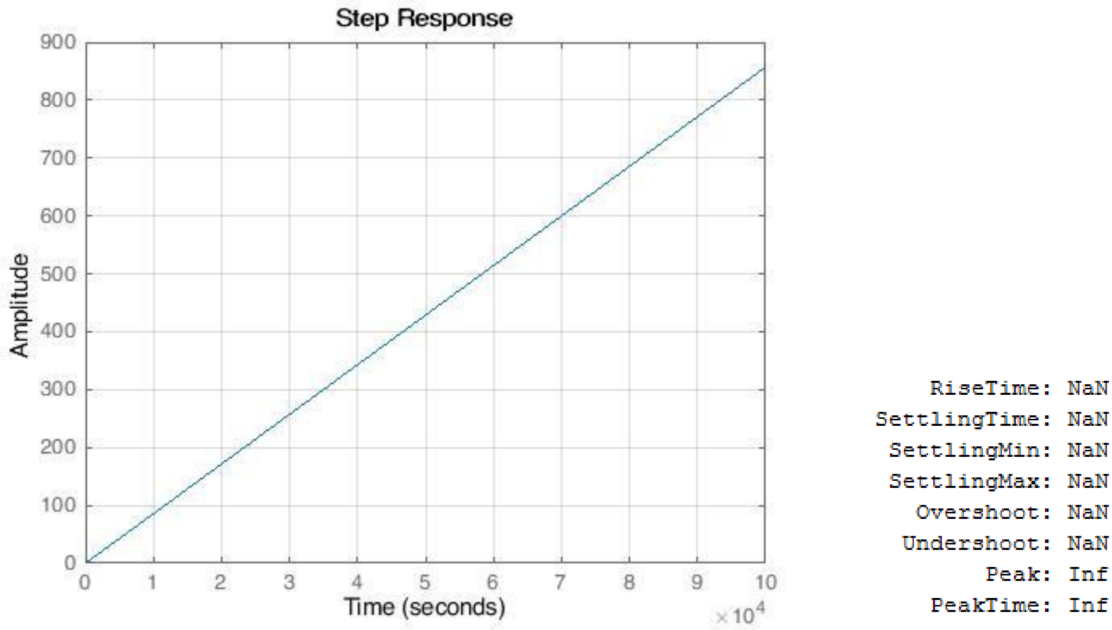

Fig 3.3.1.3, Step response plot and step response information of the identified system

The information about step response shows that the identified system is continuously increase with the step input. It is also observed from the identified model equation that one of the discrete time pole (see Figure 3.3.1.1) is one, causes the identified system response increased continuously. In general, as long as control signal feed the system, the position will increase and it will continue increase until the control signal stop.

3.4 Digital-Twin Model Analysis

The real system (plant) is a tool of a ring mill. Matlab *system identification* toolbox is used to find the correct model order and evaluate different model order as discussed in Section 3.3. However, for implementation on real plant it should be standalone and one should not depend only on Matlab. Since, we have longer series of data, contain thousands of ring productions/cycles. By using threshold on control signal (see in Figure 3.1.2.2), we can easily separate each cycle, and divided data into small chunks. Each chunk is representing each pre-production movement. In this section, system identification technique is used without using Matlab *system identification* toolbox. Data (chunk/cycle) in this section is different for model identification as compare to Section 3.3 (it is taken from another production unit).

The digital-twin model is identified when the plant work in normal condition, and the recorded control signal is used to power the twin model. The identified model process the input and generate the output. Figure 3.4.1 shows the block diagram of DT model. In this Section, we analyzed identified DT model over a longer series of data.

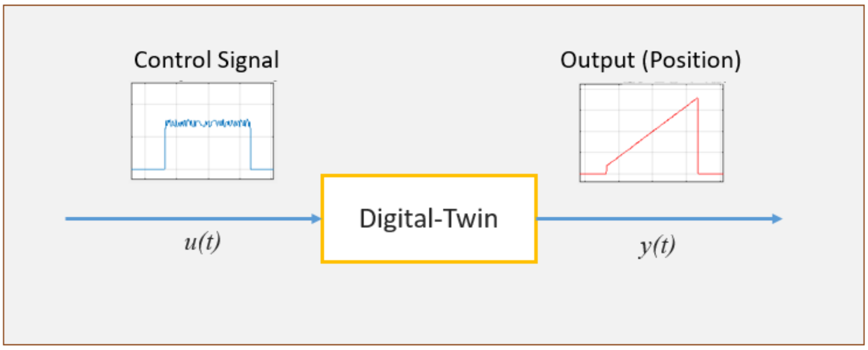

Fig3.4.1 Block diagram of Digital-Twin system

In system identification, a more often perform of identification is the increasing set of model orders, which help to know which set of model order is appropriate. This approach is typically known as principle of parsimony that is, consider the simplest solution, if it fit best use it [34]. The tool movement is linear, and it can be represent as an integrator. However, the process may also influence by external effect (e.g., depend upon production type, operator speed up the process of one or several tools). Therefore, based on the plot of identified model output and measured data output, also by the use of prediction errors. An ARx model seems suitable with two a and one b polynomial, as it accurately fit the real data. Hence, in this way one pole is used to explain the linear trend, and the second pole is to catch the rest of the system dynamics. Mathematical model of the system can be written as;

$$y_m(k) = \frac{b_1 q^{-1}}{1 + a_1 q^{-1} + a_2 q^{-2}} u(k) + \varepsilon(k) \tag{3.4.1}$$

$$y_m(k) = -a_1 y(k-1) - a_2 y(k-2) + b_1 u(k-1) + \varepsilon(k). \tag{3.4.2}$$

Where, $y_m(k)$ is the estimated output, $a_1, a_2, and\ b_1$ are the unknown parameters. The overall linear regression can be written by using Eq (2.4.1.3) as,

$$y_m(k) = \begin{bmatrix} -y(k-1) & -y(k-2) & u(k-1) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ b_1 \end{bmatrix} + \varepsilon(k). \tag{3.4.3}$$

Here, y_m is the estimated output, numerical values in Φ contain recorded input and output value from the plant, and θ contain unknown parameter. Recall the (2.5.1.4), theta θ that minimize the $\sum_{k=1}^{N} \varepsilon^2$ is,

$$\hat{\theta} = \begin{bmatrix} -0.585 \\ -0.414 \\ 0.012 \end{bmatrix}$$

The unknown parameters are now identified with $\hat{\theta}$, estimated output can be calculated as,

$$y_m(k) = 0.585y(k-1) + 0.414y(k-2) + 0.012u(k-1). \tag{3.4.4}$$

The overall system equation can written as,

$$y_m(k) = \frac{0.012q^{-1}}{1 - 0.585q^{-1} - 0.414q^{-2}} u(k) + \varepsilon(k). \tag{3.4.5}$$

Recall the theory from Sec 2.2.4, we identify the pole of the system by finding the root of denominator. That is,

$$1 - 0.585q^{-1} - 0.414q^{-2} = 0,$$

roots are,

$$q = 0.9993, and\ -0.4143.$$

Figure 3.4.2 shows the identified system's pole and zero.

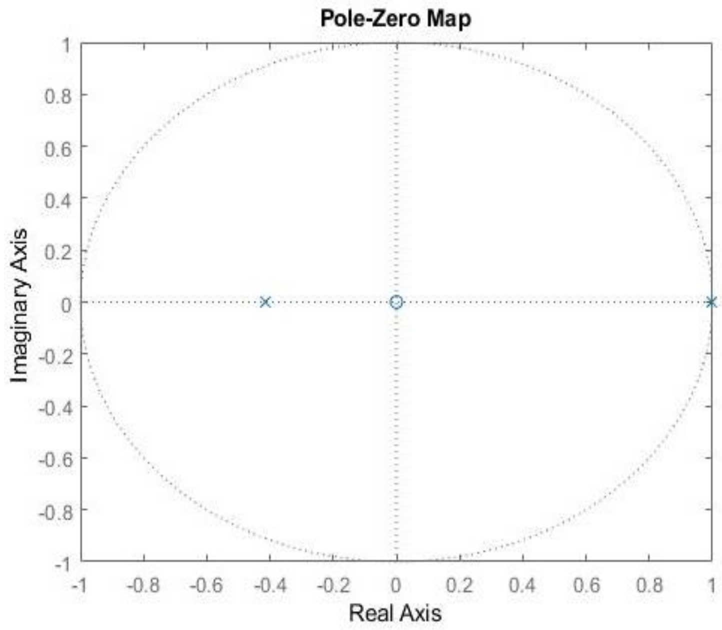

Fig 3.4.2 shows the poles (x) and zero (0) of the identified system

From the pole-zero plot, we can see that one of its pole is close to one. Since, it is a pole value of one chunk, and it do not explain in detail about the system behavior over different period of time. It is also interesting to calculate the system's pole over a longer series, and see how its pole react over time. Therefore, we identified the system's pole of each cycle of available data. Figure 3.4.3 shows how the identified system poles change between Feb 21 and Mar 14.

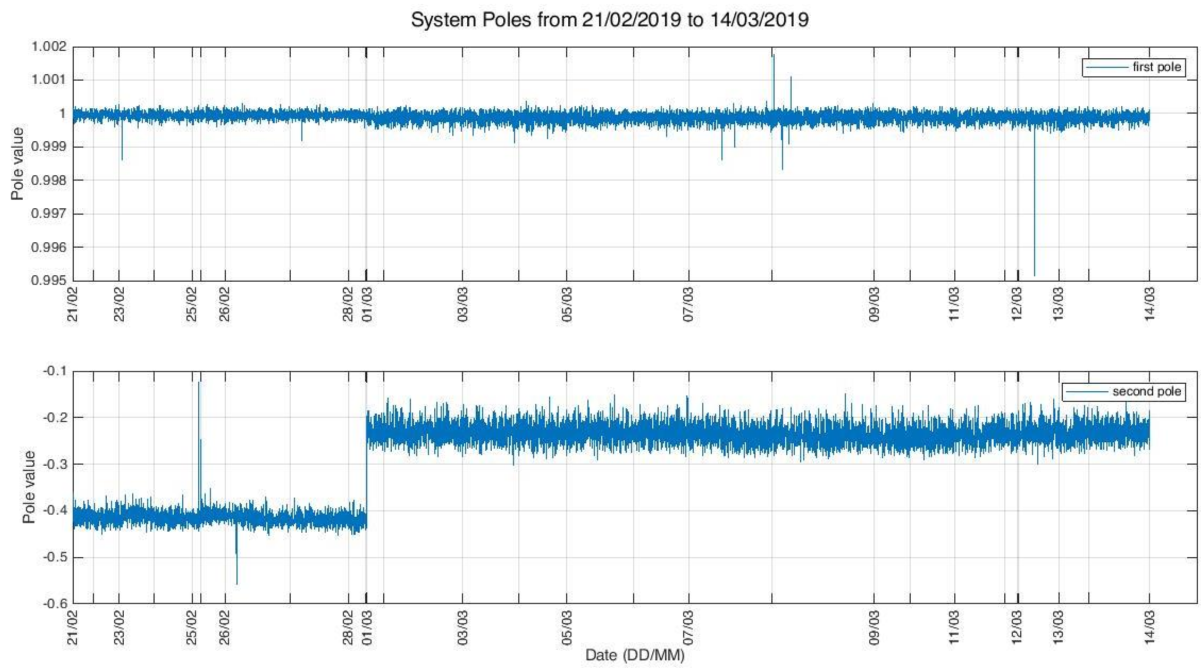

Fig 3.4.3 shows the poles of the identified system from Feb 21 to Mar 14.

We can see that one of the identified pole suddenly jump during Feb 28 and Mar 01, from lower value to higher value. If we look at the activities, we found that maintenance team has changed the hydraulic filters of several pumps, which are connected to servo valve. However, if we look at the other pole, its variation slightly increase after Mar 01 but it remain closed to 1, shows that the system behave like an integrator. From the process point of view it is also logical, as in this duration tool is moving from one point to another point, and as long as system gets control signal it moves. This movement is also known as preproduction movement, which is linear in direction like a slope. The prediction error is calculated by using (2.4.1.2) as,

$$\varepsilon(k) = y(k) - y_m(k),$$

$$\varepsilon(k) = \begin{bmatrix} -0.474 \\ 0.169 \\ \vdots \\ -0.771 \end{bmatrix}$$

Estimated output and the original output can be seen in the Figure 3.4.2, which shows that both outputs are close to each other. The x-axis shows time (sec) that *tool A* take to move from predefined origin position to the waiting position (position when *tool A* wait for the hot ring and other rolling tools to be aligned), and y axis shows the amplitude of tool's initial position to waiting position in mm.

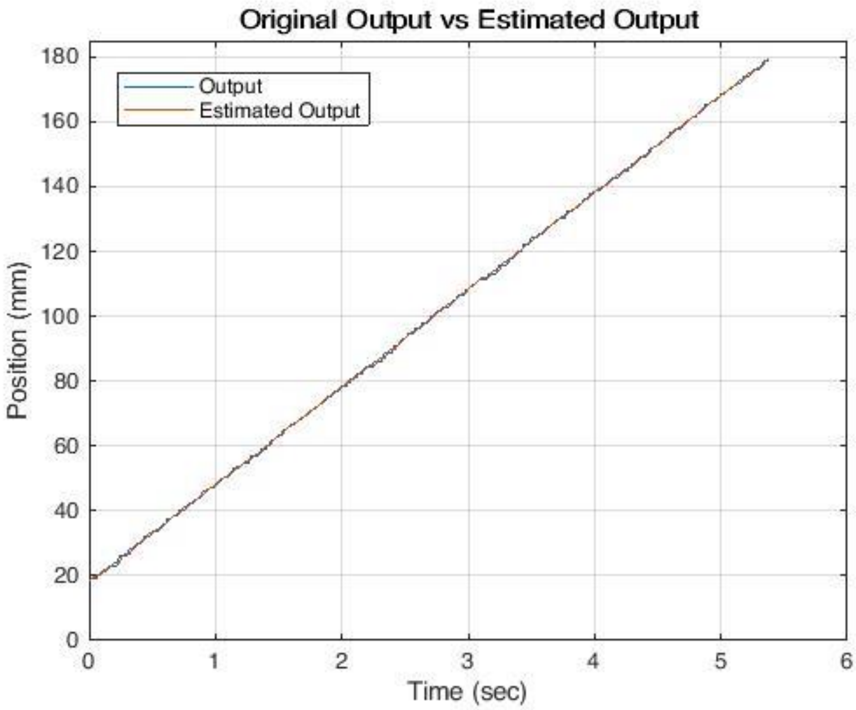

Fig 3.4.2 Original and estimated output using LS method

It can be observed from the Figure 3.4.2 that with the calculated parameter $\hat{\theta}$, estimated output is close to the real output. It is also interesting to see how the estimated parameters $\hat{\theta}$ react over a longer series. Therefore, apply the above least square method to the available series of data. Figure 3.4.3-5 shows how the estimated parameter value of each cycle changes during Feb 21 till Mar 14.

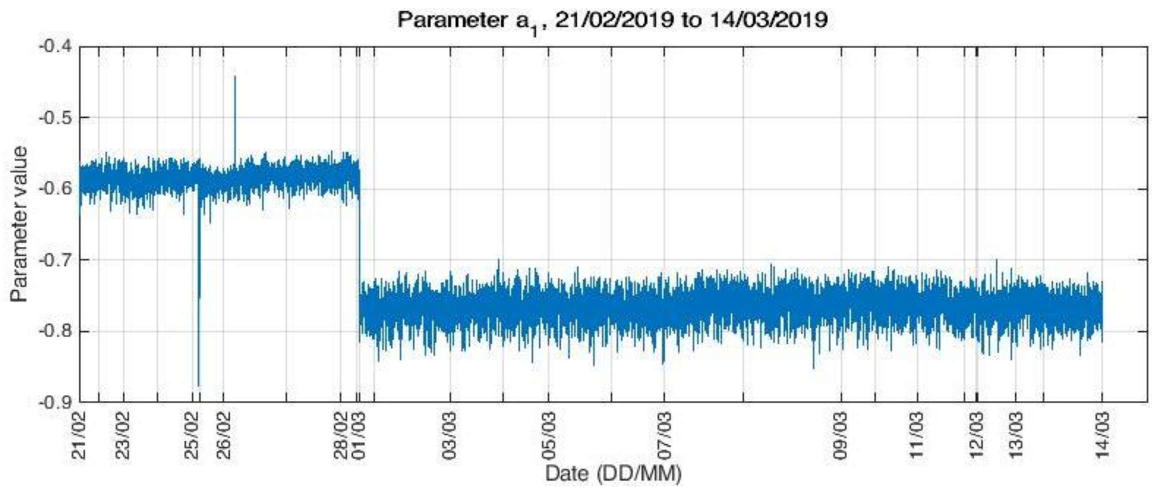

Fig3.4.3 Parameter a_1 from Feb 21 to Mar 14 using LS method

Figure 3.4.3 shows the parameter a_1 value over series of data (Feb 21 to Mar 14). The longer series data consist of 9573 rings. In which different maintenance activities has been performed. It can also observed that the parameter has abrupt change by a decrease of approx. 0.2 parameter value. There are two high positive and negative peaks between feb 25 and feb 26. those unique parameter value were calculated from bad production data.

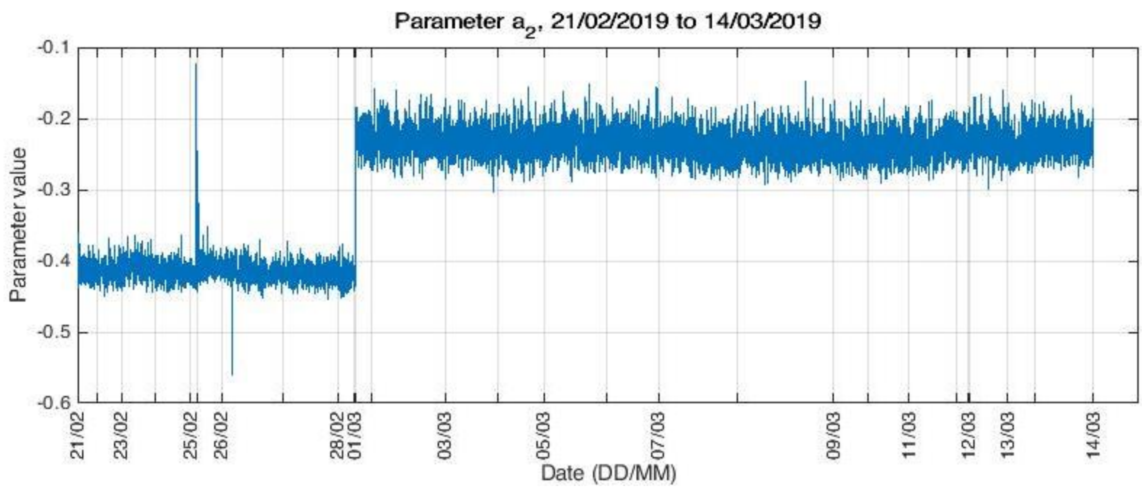

Fig3.4.4 Parameter a_2 from Feb 21 to Mar 14 using LS method

Figure 3.4.4 show the second parameter a_2 value over a longer series of data (Feb 21 to Mar 14). It is shown in the graph that the parameter value has been changed after Feb 28. It can be seen from the Figure that before March 01, the parameter value are approx. -0.4 and then it has been suddenly jumped by an increase of approx. 0.2 parameter value.

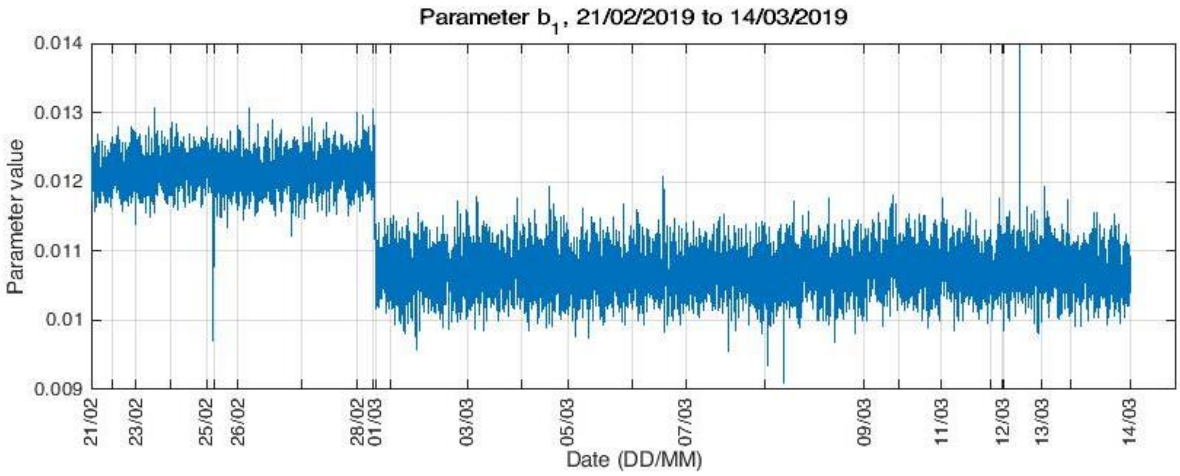

Fig 3.4.5 Parameter b_1 from Feb 21 to Mar 14 using LS method

Figure 3.4.3-5 shows the estimated parameter value of $a_1, a_2, and\ b_1$ have changed at same instance of time that is in between Feb 28 and March 01. To tracked the caused, we compared the parameter's graph with maintenance activities table. Table 3.4.1 shows the maintenance activity log.

Uppgiftsbeskrivning	Skapad	Avrapporteringsdatum
Byte av hydraulfilter (Pump 3)	2019-02-28 05:03	2019-03-07 10:52
Byte av hydraulfilter (Pump 2)	2019-02-28 05:03	2019-03-07 10:53
Byte av hydraulfilter (Pump 1.1)	2019-02-28 05:02	2019-03-07 10:54
Smörjning av stödarmar	2019-02-24 05:03	2019-03-12 10:22
FU Ringverk 4 Övre dornlagring	2019-02-24 05:03	2019-03-12 10:22
Smörjning av Ställskruv ENL R-418933	2019-02-24 05:03	2019-03-12 10:22
Smörning av valsverk ENL R-418933 Blad 3,4	2019-02-22 05:01	2019-02-28 07:30
Smörjning av stödarmar	2019-02-17 05:02	2019-02-28 07:28
FU Ringverk 4 Undre dornlagring	2019-02-17 05:02	2019-02-28 07:28
Smörjning av Ställskruv ENL R-418933	2019-02-17 05:02	2019-02-28 07:28
Mätning av horisontellt och vertikalt glapp på tärning	2019-02-17 05:01	2019-03-07 11:09
Det är glappt mellan löphjul och linjal	2019-02-15 10:57	2019-03-01 07:36
Smörjning av uttagare ENL R-418933 Blad 5,6	2019-02-14 05:02	2019-02-28 07:31
Efterfyllnad av fett i bågtandkoppling	2019-01-31 05:03	2019-02-28 09:28

Table 3.4.1 Maintenance activity created and reported date

It has been observed from the parameters graph that there is an immediate change in the parameter value after Feb 28, which could be due to maintenance activities. It can also noticed from the Table 3.4.1 that on Feb 28 the maintenance team created log (‘Skapad’) to change the hydraulic filters of different pumps. In addition, it is mentioned in the reported date column (‘Avrapporteringsdatum’) that they reported loose movement between ‘löphjul’ and ‘linjal’ on March 01. Since the exact time and data are not cleared, therefore it can be guess that due to loose movement between ‘löphjul’ and ‘linjal’ or due to hydraulic filter change the system parameter change. However, as the tool move with the movement of hydraulic cylinders it become more obvious from the process knowledge that these change appeared due to the change of hydraulic filters.

3.5 Digital-Twin of Ringvalsverk 4 (Tool analysis)

The Plant is moving from weekly maintenance to predictive maintenance. That will help to increase the plant efficiency, increase the plant's lifespan, and reduce the maintenance cost. The DT model has been identified and analyzed in earlier Sections. In this Section, we analyzed the output variation of both DT and Plant. Figure 3.5.1 shows the DT block diagram.

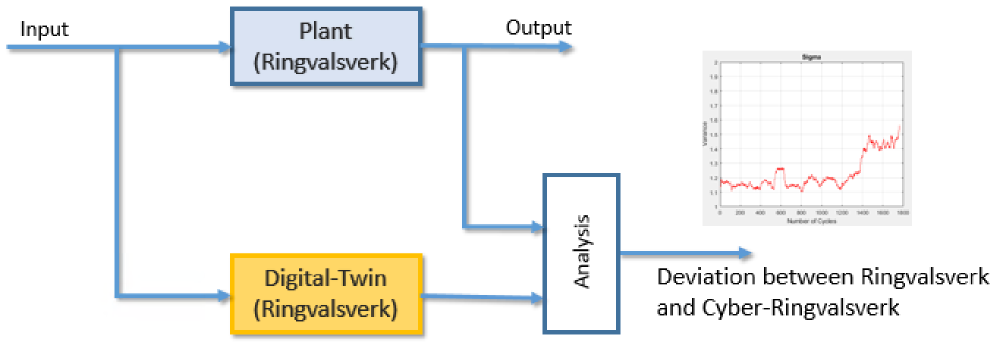

Fig 3.5.1 Digital-twin model block diagram for Ringvalsverk.

This thesis project mainly focused on a tool of *Ringvalsverk 4*. The twin of the plant is identified earlier in Section 3.2. Since the input and output data of plant has been recorded earlier by sensor data, we use that data and feed the input signal to DT model (same input that been fed to the actual plant). The plant and twin output were analyzed together to see the deviation. In the Figure 3.5.1, the analyzer block use variance (σ^2) to find the difference between plant output and DT output.

The recorded data has been processed by removing the unwanted signals, which mean that it contain only the preproduction signal and there is no digital part in it. The process of preprocessing data is mentioned earlier in Section 3.1. The DT and plant output were compared together in the analyzer block and then use the deviation to see when it differ from its normal state. Here, each output is representing as a unit (which mean one complete movement). Figure 3.5.2 shows the analyzer output.

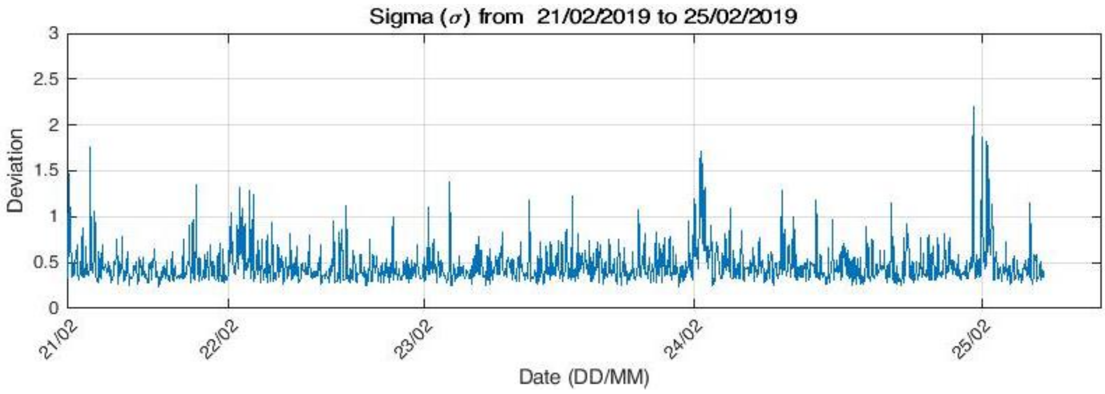

Fig 3.5.2 Digital-Twin analysis of Plant output and Twin output

Since the DT analyzed each unit one by one, which mean that if the tool start deviate from its original state then its deviations would be identified and troubleshoot in time. Therefore, it is interesting to analyze series of several units together instead of each unit/cycle. In other words, average value of each output over a sliding window of several units. In this way, it would help to draw a conclusion on several units instead of each unit. However, by considering less cycles could lead conclusion of less data and won't get rid of noise (may contain bad production). Similarly, by and considering more data it might neglect abnormal deviation as well as the overall algorithm would be slow. Figure 3.5.3 shows the moving average of less and more units.

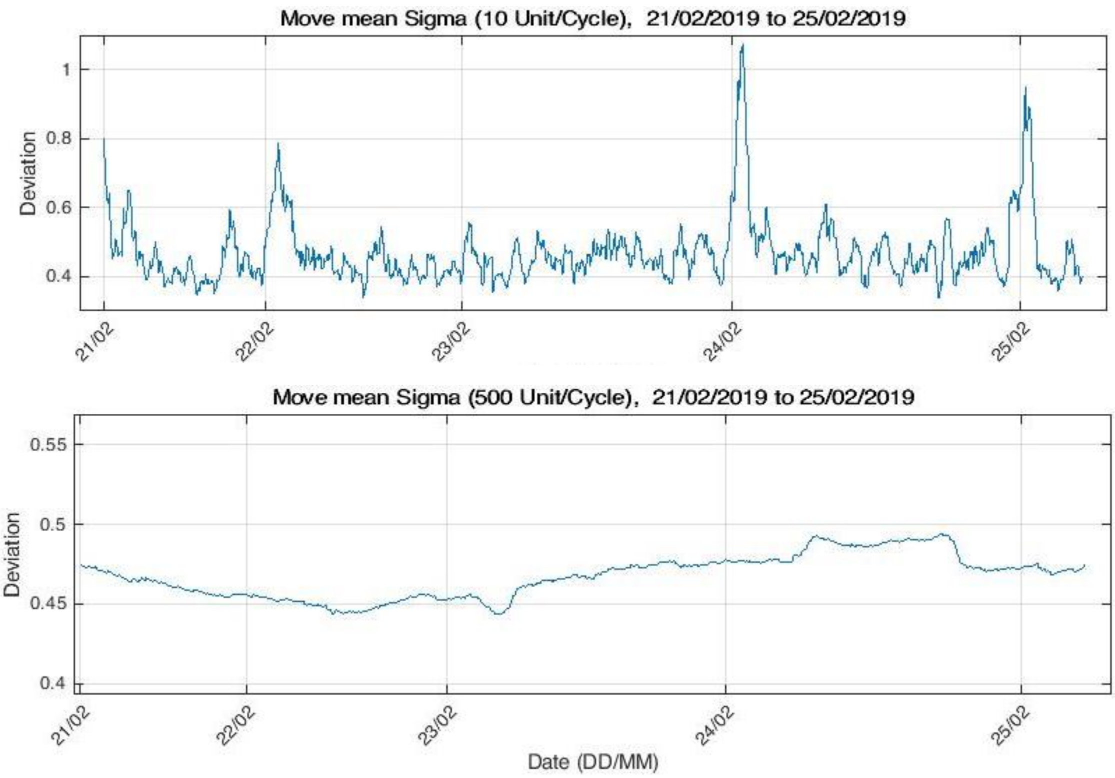

Fig 3.5.3, subplot (211) shows the moving average mean of 10, and subplot (212) shows the moving average mean of 500 units/cycles.

It can be observed from the Figure 3.5.3 that averaging over a small number could lead noisy data in the result, and too large value could neglect the deviation and changes won't be seen. Therefore, it is a tradeoff. In our work, each analyzed unit/cycle is a calculated averaging of 100 cycles/units. However, this averaging value can be adjustable. Figure 3.5.4 shows the moving average of 100 units.

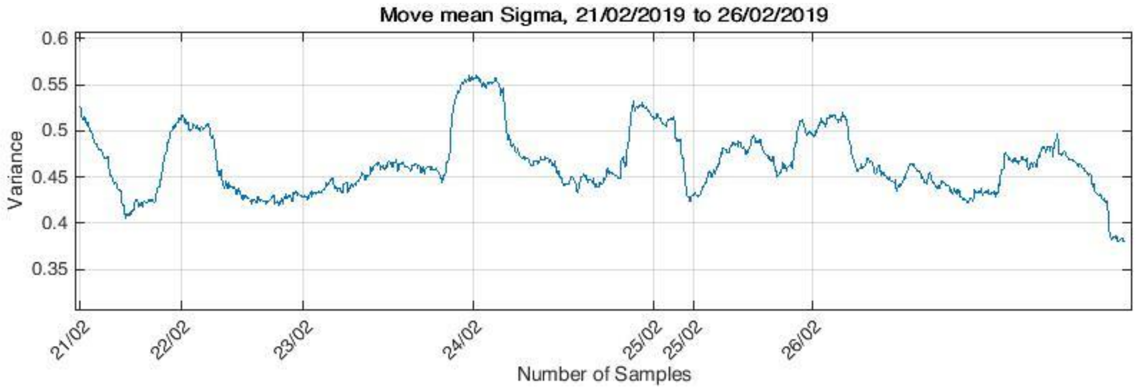

Fig 3.5.4 Digital-Twin analysis of Plant output and Twin output of six days

The graph showed in Figure 3.5.4 contained 1934 analyzed unit data. It is observed that there are some variation in the graph, which either due to maintenance activity or due to internal process disturbances. It is also mentioned in the maintenance activity that they have reported lubrication of rolling mill on Feb 22, it is also reported to perform the lubrication of different tools attached to *tool A* on Feb 24. Figure 3.5.5 shows the tool's pre-production output position.

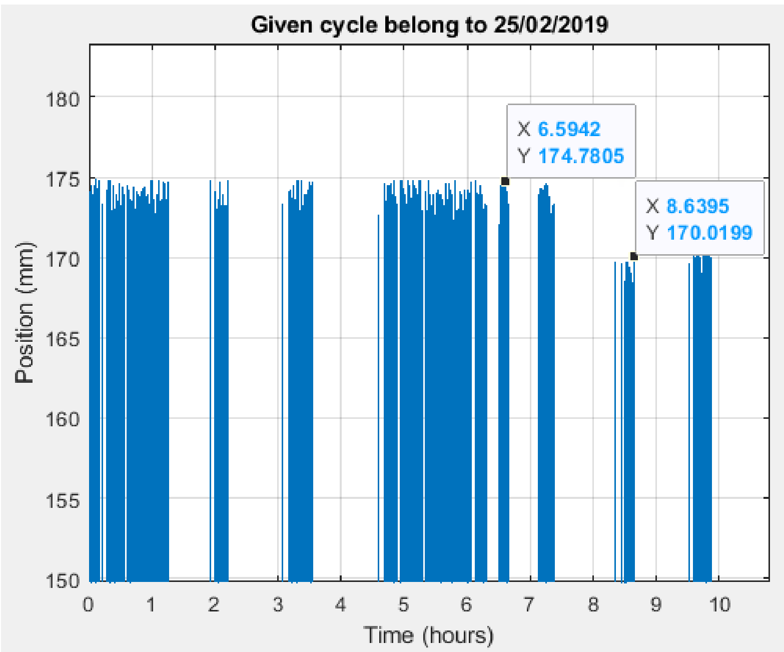

Fig 3.5.5 Plant output of day 5 (Feb 25)

It is showed in the Figure 3.5.5 that output position varies. The variation in position is normal, as it varies with respect to different type of ring production. As it is mentioned earlier that *tool A* is further attached with *tool x*. It could be the reason that tool x being changed. During the ring production, *tool x* is used to keep the rotating preform stable so that it transform according to desired shape. However, *tool x* is of different shape and size, which act as a process noise e.g., weight of *tool x* varies etc. From the control system point of view, if the weight vary then the controller need more input so that output follow the reference. Which is an external influence and act as a disturbances and these additional external noises were not considered in the DT model. Hence, the output from DT and actual plant also varies when external noise effect on the *tool A* movement.

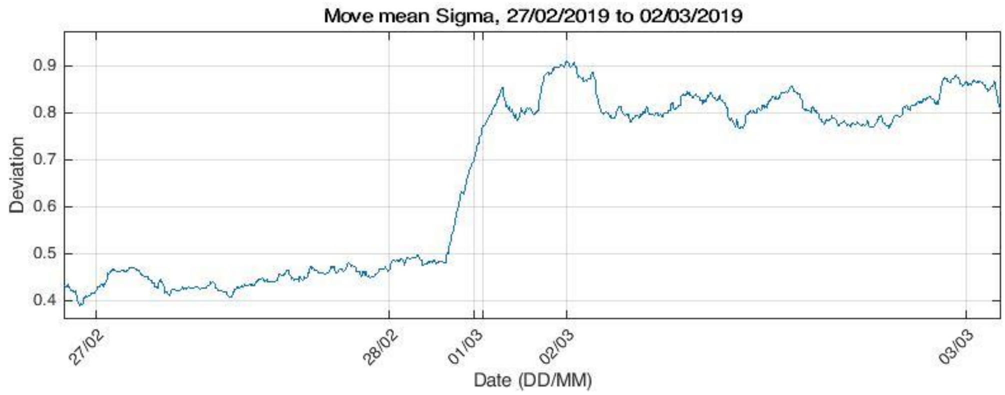

Fig 3.5.6 Digital-Twin analysis of Plant output and Twin output of four days

Figure 3.5.6 shows the dominant step change on Feb 28, and March 01. It is observed that deviation has been changed from approx. 0.48 to 0.9. Which could be due to maintenance activities. The similar change has also been observed in Section 3.2 where all the system's parameter changed instantly. In Figure 3.5.6, DT model analyzed 1530 ring/cycle in the mentioned four days (Feb 27 to Mar 02). Table 3.5.1 shows the maintenance activities.

Uppgiftsbeskrivning	Skapad	Avrapporteringsdatu
Byte av hydraulfilter (Pump 3)	2019-02-28 05:03	2019-03-07 10:52
Byte av hydraulfilter (Pump 2)	2019-02-28 05:03	2019-03-07 10:53
Byte av hydraulfilter (Pump 1.1)	2019-02-28 05:02	2019-03-07 10:54

Table 3.5.1 Maintenance activity created and reported date

From the maintenance activity shown in Table 3.5.1, it is mentioned on Feb 28 that the maintenance team has planned to change the hydraulic filters of several pumps. As *tool A* is directly attached and driven from hydraulic system, filter change could highly influenced on output deviation.

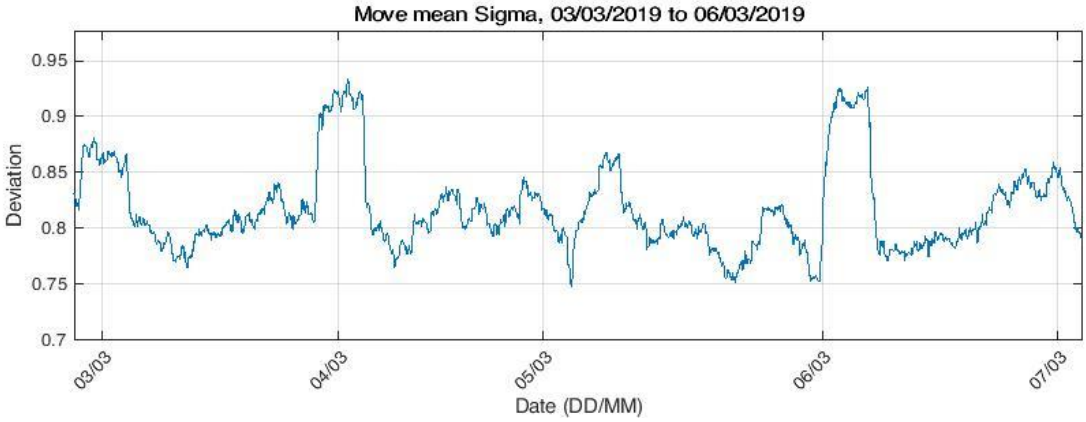

Fig 3.5.7 Digital-Twin analysis of Plant output and Twin output of two days

The graph in the Figure 3.5.7 showed the analyzed result of digital space and physical space between Mar 03 and Mar 06, which contained 2007 units. Several changes have been observed, on Mar 03 a lubrication activity has been reported for different tools. A dominating change can be observed on Mar 04 and Mar 06, which are slightly similar to changes shown in Figure 3.5.4 . The difference in variance (difference between real plant and twin output) has increased from approx. 0.75 to 0.92 respectively. It is also mentioned in the activity log that maintenance team reported the change/adjustment of 'spindlar', which can be seen in Table 3.5.2.

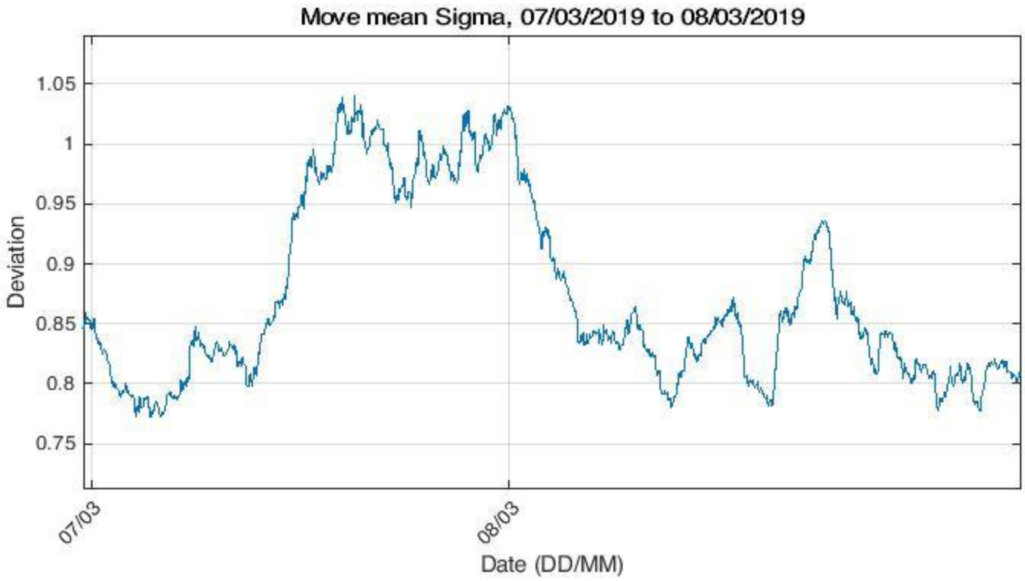

Fig 3.5.8 Digital-Twin analysis of Plant output and Twin output of two days

Figure 3.5.8 shows the DT analyzed output of plant and its twin. There were also a weekly maintenance activity on every Thursday, and March 07 was that day. In that activity, maintenance team check the condition of the plant, look for the defect, report what they have done, and what they will do in their next activity. It is also observed that on Mar 07, tool's movement varies after the first session of the day, which contain approx. 440 identical production units and were different from previous series. Figure 3.5.9 shows the plant's output of Mar 07.

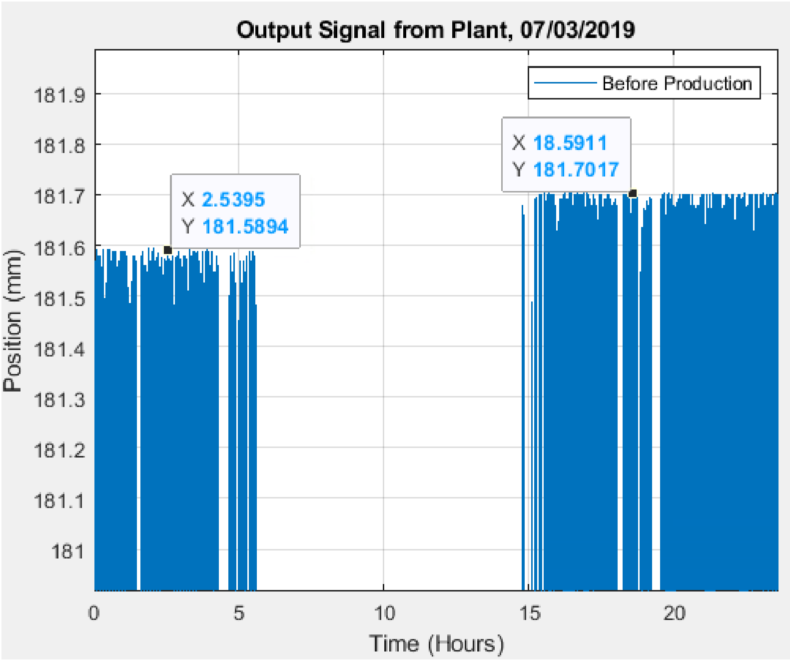

Fig 3.5.9 Plant output position of Mar 07

Figure 3.5.9 shows that before production movement of *tool A* varies. Due to absence of information, deviation in the Figure 3.5.8 may have different reasons. One reason could be due to maintenance activities that have been done before and reported later in the log. The other reason could be the new production series belong to another type of ring production, which is different from previous type of production. Since the system model is identified once, and used it as DT for all type of production. When there is a change in production type, certain tool (*tool x*) might need to be change. Therefore, change in production type would require change in control signal or change/update in system model. It might be more understandable by considering a car example, in which input is fuel and output is covering distance. By driving in snow condition, it need different tires and require different control strategy as compare to summer.

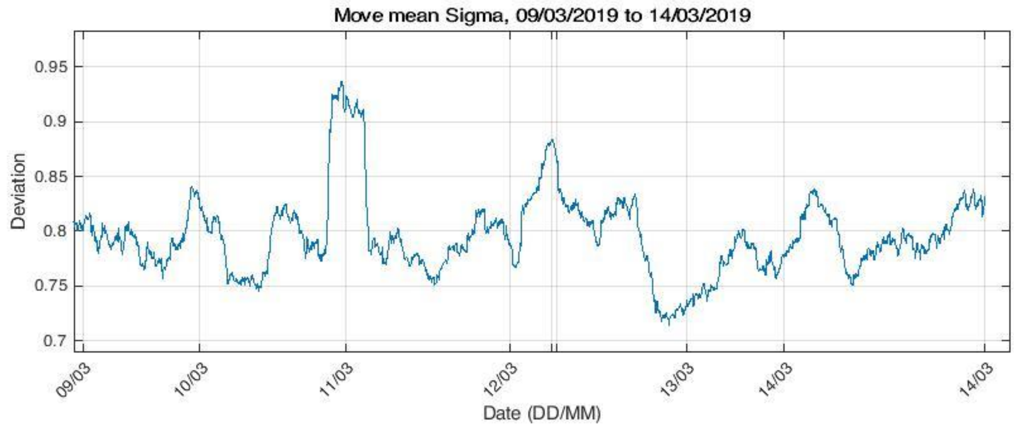

Fig 3.5.10 Digital-Twin analysis of Plant and Twin output from Mar 09 to Mar 14

Figure 3.5.10 shows the output deviation with respect to six consecutive days, the analyzed result consist of 2449 units. The deviation caused can be traced from activity log of Table 3.5.2 in which several maintenance activities reported on Mar 10, and Mar 11. On Mar 10, they reported lubrication of different tools, included piston, cylinder bracket, and bearings. Mar 11 activity; maintenance team check at different positions, by checking the hydraulic cylinders, look for oil leakage and hoses, check the loose movement between cylinder mount and pins. Another activity has been reported in which team checked and fixed the position sensors' cable connection, connected to hydraulic cylinders. The high deviation between Mar 10 and Mar 11 represent a series of 239 similar units, representing same set of production units. It is also observed from the Figure 3.5.10 that variation decreases on Mar 12. Figure 3.5.11 shows the preproduction output of the plant, in which marker shows different positions.

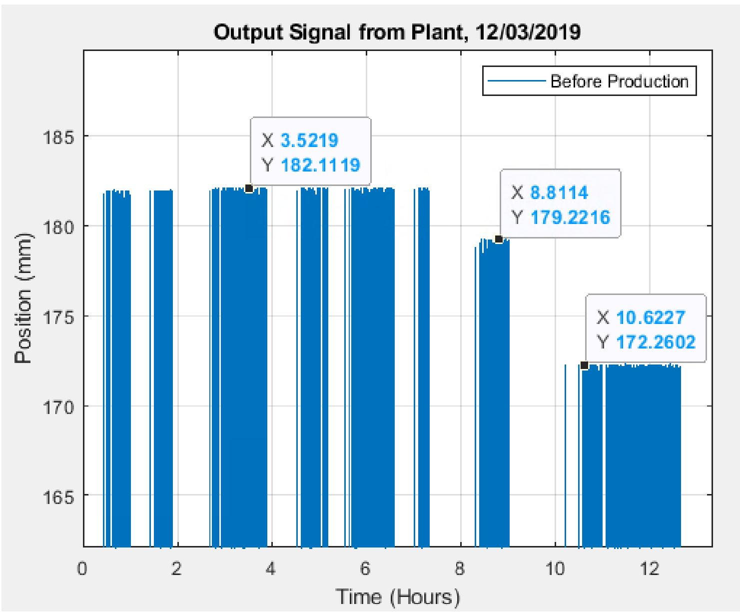

Fig 3.5.11 Plant output position of Mar 12

The output position of the plant was changed from 182 mm to 179.2 mm, and then to 172.2 mm. It can also observed from Figure 3.5.10 that on Mar 13, the deviation increases gradually. The similarities have also seen by analyzing the preproduction signal, which were also varies. The change in position might be due to change of *tool x*. It may be clear with the radiator heater example, the room temperature varies with the change (in number) of people inside the room (controller increase/decrease the flow of hot water) and control signal varies in order to meet the desired room temperature. Figure 3.5.12 shows *tool A* preproduction output position of March 13.

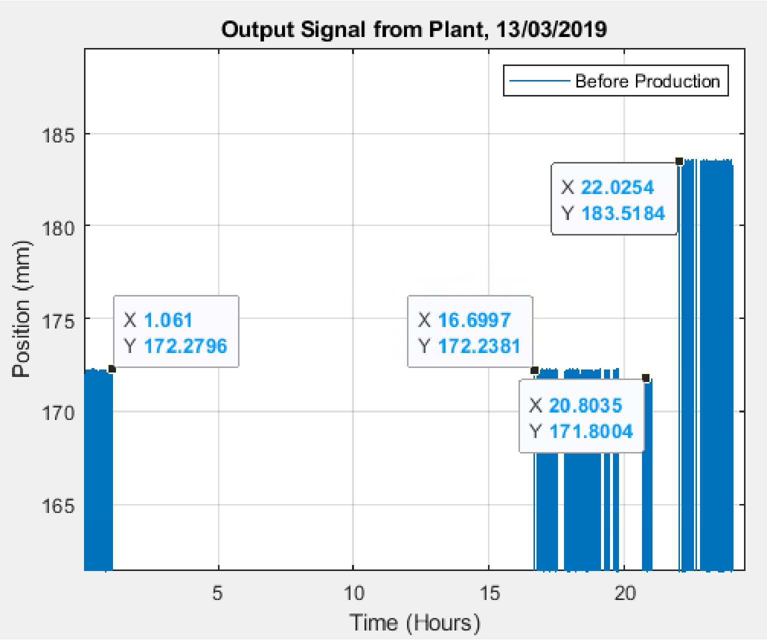

Fig 3.5.12 Plant output position of Mar 13

By analyzing the graph, it can be observed from the tool movements that it been changed several times during the day. In the initial series of production, the hydraulic cylinder's piston position decreased from 172.2 mm to 171.8 mm, and then increased to 183.5 mm. Also Mar 14 data shows that there is a gradual increase in the deviation. Therefore, the detected deviations in Figure 3.5.10 could either be due to maintenance activities or due to change of end tool of *tool A* or due to different type of series production.

Uppgiftsbeskrivning	Skapad		Avrapporteringsdatu
Smörjning av uttagare ENL R-418933 Blad 5,6	2019-03-14 05:02	S	2019-03-25 14:58
Rengörning av källaren "se text"	2019-03-11 05:01	M	2019-03-13 14:40
Kontroll av slitage på kilar och lister samt klämcylinder	2019-03-11 05:01	J	2019-03-13 14:41
Centrerarmar, Uttagare, Insättare	2019-03-11 05:01	C	2019-03-18 10:26
Smörjning av stödarmar	2019-03-10 05:03	S	2019-03-18 10:26
Kontroll av vaggan,leder, lister och skruv	2019-03-10 05:03	J	2019-03-13 14:45
Insättan	2019-03-10 05:03	C	2019-03-18 10:26
Smörjning av Ställskruv ENL R-418933	2019-03-10 05:02	S	2019-03-18 10:26
Byte/kontroll spindlar med jigg (rekl 140008)	2019-03-06 05:00	C	2019-03-12 10:22
Smörjning av stödarmar	2019-03-03 05:03	S	2019-03-12 10:22
FU Ringverk 4 Dornslid o dornstöd	2019-03-03 05:03	C	2019-03-12 10:22
Smörjning av Ställskruv ENL R-418933	2019-03-03 05:03	S	2019-03-12 10:22
Smörjning av Insättan ENL R-418933 Bald 1,2	2019-03-01 05:01	S	2019-03-12 10:22
Smörjning av uttagare ENL R-418933 Blad 1,2 och 5,6	2019-03-01 05:01	C	2019-03-07 13:04
Byte av hydraulfilter (Pump 3)	2019-02-28 05:03	U	2019-03-07 10:52
Byte av hydraulfilter (Pump 2)	2019-02-28 05:03	U	2019-03-07 10:53
Byte av hydraulfilter (Pump 1.1)	2019-02-28 05:02	U	2019-03-07 10:54
Smörjning av stödarmar	2019-02-24 05:03	S	2019-03-12 10:22
FU Ringverk 4 Övre dornlagring	2019-02-24 05:03	C	2019-03-12 10:22
Smörjning av Ställskruv ENL R-418933	2019-02-24 05:03	S	2019-03-12 10:22
Smörning av valsverk ENL R-418933 Blad 3,4	2019-02-22 05:01	S	2019-02-28 07:30

Table 3.5.2 Maintenance activity created and reported date

The Table 3.5.2 shows the maintenance activities, it can be seen that there is no exact date/time mentioned when the team performed the activity. Therefore, it is difficult to detect and trace the cause of exact change. It is also fact that mentioned in [10] that many of the planned operation are conducted manually, and any lack of theoretical evaluation renders the results difficult. It has been studied that some of the graph's analysis were based on assumption due to lack of detailed information. In the given scenario, Table 3.5.2 information is okay to report the activities, but for the future work when CPS system would implemented directly on real plant, more detailed information would be required. The detailed information would help to trace the causes of deviation, which could help to predict the need of future maintenance.

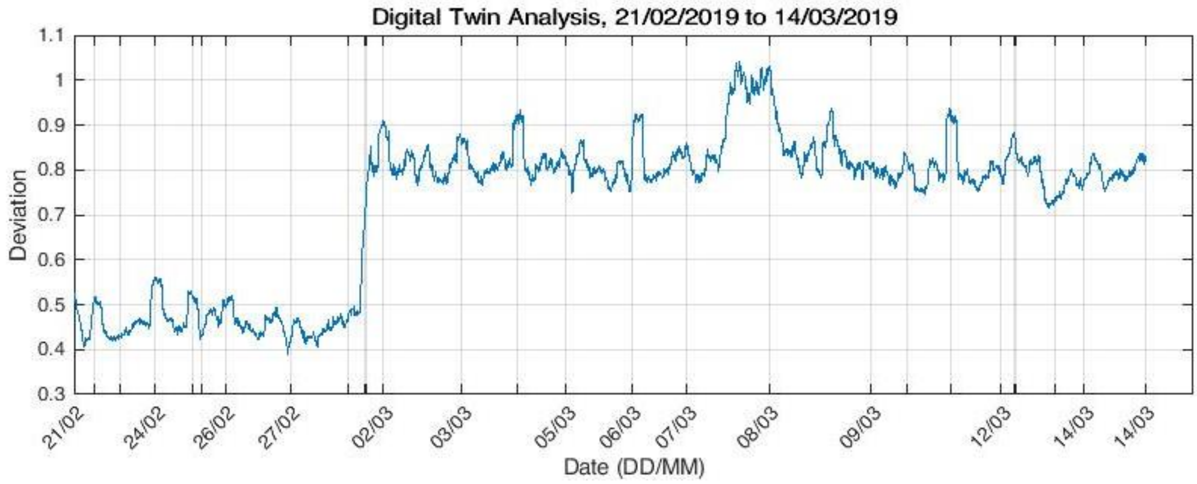

Fig 3.5.13 Plant's preproduction output analysis of three weeks

Figure 3.5.13 shows the overall DT analysis between plant's output and twin output. The three weeks DT analyzed data consists of 9500 units. It can be seen that deviation appeared during the time but it do not show any dominant trend that lead toward the poorest condition of *tool A*. One main reason is that *tool A* itself is a stable tool, and it do not need frequent maintenance. The second reason is that the given data was collected during the normal production process, and we do not have such data when *tool A* do not work in its conventional condition.

In this project, the identified system (DT) catches the reflection of real system by showing change in deviation level. Whenever there is an update in the real system, the control signal of the updated system may varies. As the sensor data is also feed to the twin system, twin process such input, and compare its output with real system output. The lower deviation shows that system is performing well, and higher deviation shows that there is a change in real system. One can also consider the deviation level and use it for predictive maintenance. As this is the pilot project, further study would require to take decision and to set alarm for maintenance need.

4 Discussion

We have seen that the available data is in discrete raw form, and data file can be read by ibaAnalyzer. We have exported file to COMTRADE format and wrote a Matlab script to combine discrete data into longer series, which is a bit time consuming. However, in the data cleaning part due to discrete input and output signal, it was a bit tricky to pick logic to extract both preproduction signals of same variable length as shown in Figure (3.2.3) but we came up to a solution to pick one signal as a master and used its index value to extract other signals.

Process knowledge is very important. It has seen while analyzing the DT model that identified system parameters have changed in a longer data series due to different maintenance activities, but due to lack of maintenance information (exact maintenance date are not mentioned, created and reported date has one or two weeks gap). Therefore, it is difficult to guess which maintenance activity caused the deviation. The method we used to identify unknown parameter without Matlab toolbox is Linear Least Square. In this method, φ depend upon previous output value, which is indirectly depend upon error. The available data is a recorded series of data, and has possibility to process further. However, in an online system when previous output value would be unknown, other methods would be useful such as Recursive Least Square etc.

It is obvious from the tool movement that its pre-production movement of each cycle is linear like an integrator (see Figure 3.1.2.4). It is actually a continuous step up signal due to zero-order hold, we have processed the data and detrend (removed the linear trend of) the output signal, but the resulting signal contained nonlinearites. Then the input control signal could not accurately able to identify the detrended output signal. It might require a high order complex system. Therefore, we decided not to detrend the signal and used it directly.

While analyzing the output variation between *tool A* and its DT model in CPS, it has been observed from the results that CPS detect the variation, which tells that tool A varied from its normal state. These changes were due to several maintenance activities, as well as due to change of production such as *tool x* need to change. It has been noticed from Table (3.4.1 and 3.5.2) that maintenance team have done several maintenance activities, and it would be difficult to track which activities caused the variation. In the future, when this will implement as a system then there will be a need for better description of activities.

Furthermore, this project is also focused on United Nation’s 2030 agenda, goal 9.4 representing to upgrade infrastructure, and retrofit industries to make them sustainable. The CPS technique would be same for all other tool in *Ringvalsverk* 4, as shown in block diagram (Figure 3.5.1). Successful implementation of CPS in an online system will help to maintain sustainable environment, as well as sustainable production. It will also help to reduce the frequent maintenance and maintenance cost.

5 Conclusions

Digital-twin is quite useful in industry upgradation. We have seen from the results that, twin has detected deviation in *tool A*. Since *tool A* is a stable tool and maintained frequently, therefore the graph do not show any gradual change from one maintenance activity to another. However, it shows the abrupt variation when there is a change in *tool A* such as hydraulic filter change etc. Interestingly, from result it is noticed that with the help of digital twin, it become easy to see the real time status of the tool. The graph showed that DT captured the system change, and the corresponding maintenance activity testify the detection.

It is also fact that there are different type of ring produced in *Ringvalsverk 4*. Some production require to change the *tool x*, which is a part of *tool A*. However, there is no information available when the operator/maintenance team change those tool. The resulting graph are challenging to track either the change is due to regular maintenance activity or due to the change of *tool x*, or due to wear of tool. Therefore, it is advantageous to know which day maintenance team perform activity, which can help to get sure about resulting graph and help to troubleshoot if production become unsatisfactory. If the maintenance activity require change in plant, such as change of *tool x* then the corresponding twin model would also need to update. The result from this project are promising but it require further verification to eliminate cyclic deviations.

Result from CPS technique show that, this method can be applied for predictive maintenance of the whole plant process (ring mill). This will require digitization of whole ring mill, and system identification of each tool. After that DT model of each tool correspond to the real time status of ring mill. Hence, it would be possible to see how each tool parameter change over time, and thus detect when the tool start deviate from its normal condition. However, process knowledge is still very important aspect to determine which signal should consider in order to detect the need of maintenance.

For the future work, it is useful to study the influence of pressure as it is observed that it effect directly on the movement. Since the tool (system) is controlled by hydraulic actuators (sub-system), connected to proportional or servo valve (system) and some other systems. Therefore, it is also worth taking to explore system of system and create twin(s) of necessary system. In this way, when any system meet the maintenance threshold level. An alarm generate, indicate that system need maintenance, and help maintenance team to plan accordingly.

References

[1] L. Giorleo, E. Ceretti, and C. Giardini, “Speed Roll Laws Influence in a Ring Rolling Process,” *Key Eng. Mater.*, vol. 554–557, no. January 2016, pp. 337–344, 2013.

[2] V. Jenkouk, G. Hirt, M. Franzke, and T. Zhang, “Finite element analysis of the ring rolling process with integrated closed-loop control,” *CIRP Ann. - Manuf. Technol.*, vol. 61, no. 1, pp. 267–270, 2012.

[3] X. Wang and L. Hua, “Analysis of guide modes in vertical hot ring rolling and their effects on the ring’s dimensional precision using FE method,” *J. Mech. Sci. Technol.*, vol. 25, no. 3, pp. 655–662, 2011.

[4] B. Shiklo, “A Digital Twin Approach to Predictive Maintenance.” [Online]. Available: https://www.informationweek.com/big-data/ai-machine-learning/a-digital-twin-approach-to-predictive-maintenance/a/d-id/1333331.

[5] R. Baheti and H. Gill, “The impact of control technology-Part3-02CyberphysicalSystems,” no. 1, 2011.

[6] K. Nita Ali, M. Sun, G. Petley, and P. Barrett, “Improving the business process of reactive maintenance projects,” *Facilities*, vol. 20, pp. 251–261, 2002.

[7] K. Anrell, “Uppkopplingen digitaliserar gamla fabriker,” 2017. .

[8] Z. Liu, N. Meyendorf, and N. Mrad, “The role of data fusion in predictive maintenance using digital twin,” *AIP Conf. Proc.*, vol. 1949, no. April, 2018.

[9] “Virtual testing of Siemens Siprotec 5 protective devices with a digital twin,” 2019. [Online]. Available: www.siemens.com/press/PR2019030209EMEN.

[10] T. H. J. Uhlemann, C. Lehmann, and R. Steinhilper, “The Digital Twin: Realizing the Cyber-Physical Production System for Industry 4.0,” *Procedia CIRP*, vol. 61, pp. 335–340, 2017.

[11] “Siemens, Turing corn into clothing, digital bio-manufacturing.” [Online]. Available: https://assets.new.siemens.com/siemens/assets/public.1534243025.0ce3e3b1b6db60e4acaad86d6ce6f289b77ca7b4.turning-corn-into-clothing-en.pdf.

[12] “Industry 4.0: Building the digital enterprise,” *Global Industry 4.0 Survey*, 2016. [Online]. Available: https://www.pwc.com/gx/en/industries/industries-4.0/landing-page/industry-4.0-building-your-digital-enterprise-april-2016.pdf.

[13] J. Lee, H. A. Kao, and S. Yang, “Service innovation and smart analytics for Industry 4.0 and big data environment,” *Procedia CIRP*, vol. 16, pp. 3–8, 2014.

[14] I. Ungurean, N. C. Gaitan, and V. G. Gaitan, “An IoT architecture for things from industrial environment,” *IEEE Int. Conf. Commun.*, 2014.

[15] F. Tao *et al.*, “Digital twin-driven product design framework,” *Int. J. Prod. Res.*, vol. 57, no. 12, pp. 3935–3953, 2019.

[16] W. Yang, K. Yoshida, and S. Takakuwa, “Digital Twin-Driven Simulation for a Cyber-Physical System in Industry 4.0 Era,” pp. 227–234, 2017.

[17] J. Lee, B. Bagheri, and H. A. Kao, “A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems,” *Manuf. Lett.*, vol. 3, pp. 18–23, 2015.

[18] J. A. Stover, D. L. Hall, and R. E. Gibson, “A fuzzy-logic architecture for autonomous multisensor data fusion,” *IEEE Trans. Ind. Electron.*, vol. 43, no. 3, pp. 403–410, 1996.

[19] D. W. Clarke, “Application of Generalized Predictive Control to Industrial Processes,” *IEEE Control Syst. Mag.*, vol. 8, no. 2, pp. 49–55, 1988.

[20] G. N. Saridis, “Intelligent Robotic Control,” *IEEE Trans. Automat. Contr.*, vol. 28, no. 5, pp. 547–557, 1983.

[21] B. Godin, “The Linear Model of Innovation,” *Sci. Technol. Hum. Values*, vol. 31, no. 6, pp. 639–667, 2006.

[22] Torkel Glad and Lennart Ljung, *Control Theory Multivariable and Nonlinear Methods*. 2008.

[23] C. W. Taylor, D. C. Erickson, K. E. Martin, R. E. Wilson, and V. Venkatasubramanian, “WACS - Wide-area stability and voltage control system: R&D and online demonstration,” *Proc. IEEE*, vol. 93, no. 5, pp. 892–906, 2005.

[24] T. Soderstrom and S. Petre, *System Identification*, 2001st ed. Prentice Hall International (UK) Ltd, 1989.

[25] T. Wagener and H. V. Gupta, “Model identification for hydrological forecasting under uncertainty,” *Stoch. Environ. Res. Risk Assess.*, vol. 19, no. 6, pp. 378–387, 2005.

[26] J. Sjöberg *et al.*, “Nonlinear Black-box Modeling in System Identification: a Unified Overview*,” *Automatica*, vol. Volume 31, no. 12, pp. 1691–1724, 1995.

[27] K. Ažman and J. Kocijan, “Application of Gaussian processes for black-box modelling of biosystems,” *ISA Trans.*, vol. 46, no. 4, pp. 443–457, 2007.

[28] G. J. Mccabe and D. R. Legates, “Evaluating the use of ‘goodness-of-fit’ measures in hydrologic and hydroclimatic model validation,” *Water Resour. Res.*, vol. 35, no. 1, pp. 233–241, 1999.

[29] “ibaPDA-PLC-Xplorer at a glance.” [Online]. Available: https://www.iba-ag.com/en/products/ibapda-plc-xplorer/.

[30] M. Part, “SIMATIC NET S7 − CPS for Industrial Ethernet Manual Part B,” *Displays*, pp. 1–59, 2009.

[31] K. U. Kasemir and L. R. Dalesio, “Interfacing the ControlLogix PLC over Ethernet/IP,” *arXiv*, vol. cs.NI, pp. 1–3, 2001.

[32] Power Systems Relay Committee of the IEEE Power Engineering Society, *IEEE Standard Common Format for Trainsient Data Exchange (COMTRADE) for Power Systems*. 1999.

[33] “Create linear and nonlinear dynamic system models from measured input-output data.” [Online]. Available: https://se.mathworks.com/products/sysid.html.

[34] P. Mattsson, D. Zachariah, and N. Bjorsell, “Flexible models for smart maintenance,” *Proc. IEEE Int. Conf. Ind. Technol.*, vol. 2019-February, pp. 1772–1777, 2019.