MODELING OF STOCK MANAGEMENT BY PARAMETRIC MODELS ARX, ARMAX, BJ AND OE

1 HICHAM FOURAIJI, 2 BANAYAD NSIRI, 3 BAHLOUL BENSASSI

1,3 GITIL laboratory, faculty of sciences Ain Chock Hassan II University, Casablanca Morocco
2 Departement of Applied Mathematics and informatics, Ecole Normale Supérieure de l’Enseignement Technique Mohamed V University, Rabat, Morocco

E-mail: 1 hicham.fouraiji@gmail.com, 2 benayad.nsiri@telecom-bretagne.eu, 3 b.bensassi@fsac.ac.ma

ABSTRACT

In this paper, we propose a new approach to the modeling of storage systems. This method is based on the knowledge of the automatic domain, in order to construct a mathematical model that accurately formalizes the behavior of the system studied. The approach adopted for this study is the parametric identification of ARX, ARMAX, Box Jenkins and OE linear systems. The logistic system studied will be considered as a black box, which means that the input and output data of the system will be used to identify internal system parameters and propose a mathematical model.

Keywords: Warehousing system, Modeling, Identification, ARX, ARMAX, OE, BJ

1. INTRODUCTION

During last decades, the industrial companies are confronted with great changes in their environment, a world competition, a dubious market, and increasingly demanding customers. These various constraints impose a reactivity and flexibility of the companies with an aim of adapting the capacity of the systems of production to the changes of the request and the internal and/or external risks of the line productions. What requires a thorough knowledge of the characteristics of the supply chain.

Difficult to control, these systems continue to pose serious problems of design, modeling and control. Indeed, the logistic study of the systems, like any type of dynamic system, proves to be a task very difficult to realize, and very often requires that one has mathematical models of these systems. These models can be deduced directly from the physical laws which govern the behavior of the system, but it is often impossible to obtain a knowledge a priori supplements and precise of all the parameters of the model.

In this case, to refine and specify this knowledge, one resorts to a behavioral model estimated starting from the input-outputs observed of the system. This refers to the identification approach, which is the set of methodologies for the mathematical modeling of systems, based to real measurements from the process [1].

In our context we are interested in identifying logistics systems based on parametric models, in this context we can mention the work of K.LABADI and al. presented a modeling and performance analysis of logistics systems, based on a new model of stochastic Petri nets. This model is suitable for modeling flows evolving in discrete amounts (lots of different sizes). It also allows taking into account the more specific activities such as customer orders, inventory the supply, production and delivery in batch mode [2]. N.SAMATA and al. presented a model of global supply chain using Petri networks with variable speed. They transposed the developed traffic concepts to production type of supply chain (manufacturing). They also proposed a modular approach for modeling the different actors in the supply chain is still based on the formalism of Petri nets with variable speed [3]. F.PETITJEAN and al. develop their work in a global modeling methodology of the supply chain from an audit of the company. Then using the UML model they
realized a simulation platform and they also proposed pilotage principles of integrated logistics chain [4]. BROHEE and al. used hybrid Petri networks to offer an offline simulation approach that supports multiple constraints (control change, time among developments, friction ....). The originality of their work lies in the fact that they proposed a simulation of continuous part production and the study of interactions between the continuous model and discrete data exchanged with the control part. This approach allows simulating and controlling the system without using the actual operative part, these contributions are based on a schematic modeling, without interested to the internal parameters of the system studied [5]. H.SARIR and al. presented a modeling approach and regulatory work in progress stocks in the macroscopic analogy of production lines with the control model of a hydraulic tank. They used the concepts of automatic control for monitoring and mastering of in-process inventories [6]. H.SARIR and al. who also presented a model of a production line using the behavioral identification in discrete time by the transfer functions, they used the PEM algorithm for the construction of models and the simulation was performed on the graphical interface (IDENT) in MATLAB© [7]. K.TAMANI and al. proposed a process for controlling product flow, where they broke down the system studied as basic production modules. They proposed thereafter, the control of flow through each output module and supervision which was based on fuzzy logic [8].

This literature study, we show that the logistics systems today are a focus for scientific research in the field of modeling.

Modeling approaches are many and varied, but it appears that the methods of analysis and production system design combining different approaches are preferred. Indeed the latter bring an ease of analysis or increased use or opportunities to put in simplified work.

We found that modeling logistics systems, based on parametric models is rarely used, and that much of the work focuses on the schematic or analytical modeling (Petri nets, UML ....), while parametric and behavioral modeling is rarely discussed.

This article propose a method for modeling logistics systems based on parametric models.

2. IDENTIFICATION METHOD

The purpose of any modeling system is to build a model, that is to say, a mathematical representation of its operation. Dynamical systems, we are interested in this work, are at the heart of the supply chain (storage system and storage). It is often necessary to build a model to understand, simulate, controlled and steered these systems.

When modeling a process, two approaches are possible:

The first is to build a knowledge model. Designing knowledge models stems from a physical analysis of phenomena involved in the system; when necessary, the system is broken down into simpler components studied, for which already has a proven knowledge model. Experimental data are then used, first to numerically estimate the values of model parameters, then obtained to validate the model. In particular, scientific research essential purpose the construction of models of this type, which allow not only to understand but also to extrapolate the behavior of any system.

The second approach is to build a model type "black box". Specifically, the aim is a mathematical expression that reflects faithfully the behavior "input-output" system studied in an area of operation defined use. The parameters usually have no physical meaning. The numerical estimation of these parameters based primarily on a set of experimental observations that are available on the system; The models 'black box' is usually in times of economic calculation. Their validity is limited to an operating range determined by the set of measured input-output, while that of knowledge models is determined by the accuracy of assumptions and relevance of the approximations made in the physical analysis of phenomena and equation of their layout. As part of the design of "black box" models of dynamic system, the parametric models are an excellent candidate that is typically used to approximate the dynamic behavior of the system in a satisfactory manner.

In this work, we focus on the development of models of the type "black box" of dynamical systems. We only consider the case of stationary systems, that is to say, such as the laws that govern their behavior does not change over time. The models we consider are discrete-time models.

The models 'black box' are well suited to the design models for the control of dynamic systems, where one often needs relatively simple models to be able to perform many calculations, and adjust if necessary the parameters of the model new sets of experimental data. Knowledge of the models, although based on a thorough analysis of the process and having validity ranges typically less restricted, are often too complex to perform
calculations quickly. In addition, their design is closely related to the particular process that is desired to be modeled, and the knowledge available on the physical thereof.

A mathematical model is always an approximation of the real system. In practice, system complexity, limited prior knowledge of the system and incomplete observed data prevents an accurate mathematical description of the system. However, even if we have a complete knowledge of the system and enough data, an accurate description often is not desirable because the model would become too complex to be used in an application. Therefore, identification of the system is regarded as an approximate model for a specific application on the basis of the observed data and knowledge of the prior system. The identification procedure in order to reach an appropriate mathematical model of the system, is described in detail in Figure 1.

2.1 The acquisition of Data set

The first step in the identification process is the acquisition of data, this step is very important, why we must choose the data that will be used to model the system, for that must use data covering the entire range system operating in normal conditions.

The input and output data should be divided into two data sets, the first data portion to estimate while the second part of data validation purposes.

2.2 Choosing the model structure

The parametric model describes a system in terms of differential transfer function. There are few models of structures that can be used to represent certain system. In general, the structure of the parametric model used is based on the equation (1) [9].

\[ y(t) = q^{-nk}G(q)u(t) + H(q)e(t) \]  (1)

Where \( q^{-nk}G(q)u(t) \) represents the output without disturbance, and \( H(q)e(t) \) designates the disturbance [1]. \( q \) is the argument of \( G(q) \) and \( H(q) \), it is the offset operator, which is equivalent to \( q^{-1} \) represented by \( q^{-nk} \) and can be demonstrated by \( q^{-1}x(t) = x(t-1) \), \( nk \) is the delay time in the sampling time between the input and the output of the process. In modeling process, always \( nk \geq 1 \) to ensure causality [2].

Four linear models are considered in this study. These models are ARX, ARMAX, BJ, and OE model. In this article, the studied system is a multi inputs single-output (MISO).

2.2.1 ARX model

The structure of the ARX model is given by:

\[ y(t) + a_1 y(t-1) + \ldots + a_{na} y(t-n_a) = b_1 u(t-1) + \ldots + b_{nb} u(t-nb - nk + 1) + e(t) \]  (2)

when \( n_a \) and \( n_b \) are the orders of the ARX model, \( n_a \) is the number of poles and \( n_b \) is the number of zeros plus 1, \( nk \) is the delay (number of samples of the input that occur before the input affects the output, also called the dead time in the system), \( y(t) \) is output as a function of time, \( u(t) \) is the input of system, is function of time, and \( e(t) \) is the term that represents the disturbance in the form of white noise.

We present this function with compact form as (1):

\[ y(t) = q^{-nk}\frac{B(q)}{A(q)}u(t) + \frac{1}{A(q)}e(t) \]  (3)

The polynomials \( A(q) \) and \( B(q) \) are given by :

\[ A(q) = 1 + a_1 q^{-1} + \ldots + a_{na} q^{-na} \]  (4)

\[ B(q) = b_0 + b_1 q^{-1} + \ldots + b_{nb} q^{-nb} \]  (5)

\( A(q) \) and \( B(q) \) represent the dynamic system, \( q^{-1} \) is the delay operator, this description of \( q \) is equivalent to the transformed \( z \) [10].
2.2.2 ARMAX model

The structure of the ARMAX model is given by:

\[ y(t) + a_1 y(t-1) + \ldots + a_n y(t-n_a) = b_1 u(t-n_b) + \ldots + b_{nb} u(t-n_b - n_k + 1) + c_1 e(t-1) + \ldots + c_{nc} e(t-n_c) + e(t) \]

where \( n_a \) and \( n_b \) are the orders of the ARMAX model, \( n_a \) is the number of poles and \( n_b \) is the number of zeros plus 1, \( n_k \) is the delay (number of samples of the input that occur before the input affects the output, also called the dead time in the system), \( y(t) \) is output as a function of time, \( u(t) \) is the input of system, function of time, and \( e(t) \) is the term that represents the disturbance in the form of white noise.

\[ y(t) = q^{-n_b} B(q) u(t) + \frac{C(q)}{A(q)} e(t) \]  

(7)

The polynomials \( A(q) \), \( B(q) \) and \( C(q) \) are given by:

\[ A(q) = 1 + a_1 q^{-1} + \ldots + a_n q^{-n_a} \]

\[ B(q) = b_0 + b_1 q^{-1} + \ldots + b_{nb} q^{-nb} \]  

(8)

\[ C(q) = 1 + c_1 q^{-1} + \ldots + c_{nc} q^{-nc} \]

\( A(q) \) and \( B(q) \) represent the dynamic system, \( C(q) \) represent the model of disturbance, \( q^{-1} \) is the delay operator, this description of \( q \) is equivalent to the transformed \( z \).

2.2.3 OE model

The structure of the OE model is given by:

\[ y(t) + f_1 y(t-1) + \ldots + f_{nf} y(t-n_f) = b_1 u(t-n_b) + \ldots + b_{nb} u(t-n_b - n_k + 1) + v(t) \]

when \( n_f \) and \( n_b \) are the orders of the OE model, \( n_f \) is the number of poles and \( n_b \) is the number of zeros plus 1, \( n_k \) is the delay (number of samples of the input that occur before the input affects the output, also called the dead time in the system), \( y(t) \) is output as a function of time, \( u(t) \) is the input of system, function of time, and \( v(t) \) is the term that represents the disturbance.

\[ y(t) = q^{-n_b} F(q) u(t) + v(t) \]  

(10)

The polynomials \( F(q) \) and \( B(q) \) are given by:

\[ F(q) = 1 + f_1 q^{-1} + \ldots + f_{nf} q^{-nf} \]  

(11)
B(q) = b₀ + b₁q⁻¹ + ... + bᵦq⁻ⁿᵇ

F(q) and B(q) represent the dynamic system, q⁻¹ is the delay operator, this description of q is equivalent to the transformed z [10], f₁ ... fᵦ and b₁ .... bᵦ are the parameters of the polynomials. The flow of the signal can be represented by the following figure:

2.2.4 Box Jenkins model

BJ model belongs to the class of output error models, it is a model OE with additional degrees of freedom for the noise model while the OE model assumes a white additive disturbance at the output of the process, allows the BJ modeling of any disturbance. It can be generated by filtering white noise through a linear filter with numerator and denominator arbitrary.

BJ model is illustrated in Figure 5 by:

\[ y(t) = \frac{B(q)}{F(q)} u(t) + \frac{C(q)}{D(q)} v(t) \]  

where

\[ F(q) = 1 + f₁q⁻¹ + ... + fₙq⁻ⁿᶠ \]
\[ B(q) = b₀ + b₁q⁻¹ + ... + bᵦq⁻ⁿᵇ \]
\[ C(q) = 1 + c₁q⁻¹ + ... + cₙcq⁻ⁿᶜ \]
\[ D(q) = 1 + d₁q⁻¹ + ... + dₙdq⁻ⁿᵈ \]

2.3 Parameter estimation

It is choosing an estimation algorithm or a criterion to be minimized. For parametric models, we use least squares algorithm to minimize prediction error ε between the estimated output \( \hat{y}(t) \) and real output \( y(t) \) Figure 2 illustrate the estimation model adopted for this study.

Parameter estimation by least squares (MC) is the approach most commonly used for identification systems [10]. Introducing a vector regression vector \( \varphi(k) \), equation (1) can be rewritten as a linear regression:

\[ y(k) = [\varphi(k)]^T \theta + \varepsilon(k) \]  

Where \( \varphi(k) \) is the observation vector:

\[ [\varphi(k)]^T = [-y(k-1), ..., -y(k-na), u(k-nk), ..., u(k-nk-nb+1)] \]  

The criterion function used for the least squares estimation is given by:

\[ V_n(\theta) = \frac{1}{N} \sum_{k=1}^{N} (y(k) - [\varphi(k)]^T \theta)^2 \]  

Minimizing the function of the test, we can get an estimate of \( \theta \):

\[ \hat{\theta} = \text{argmin} V_n(\theta) \]  

Once a series of input-output information is given, the model parameters are estimated using the least square method. In addition, the recursive modeling technique is a fast technique for estimating the model parameters, after each sampling of a data set. Thus, in this study, the least squares method and the recursive algorithm with forgetting factor used for the identification of line parameters of the model presented in (9).

This method gives an estimate of parameter model studied minimizing recursively (MC) defined by the criterion (14).
to check whether the model identified meet the requirement of modeling for a particular application. In achieving good estimated model, it is necessary to distinguish between the lack of fit between the model and the data due to random processes and, due to the lack of model complexity. In most statistical tools, measuring the fit of a model is determined by the coefficient of determination $R^2$ [12].

The $R^2$ is given by:

$$R^2 = 1 - \frac{\text{sum of squared residuals}}{\text{total sum of squares}}$$

(18)

The RMSE (root mean square error) is used to evaluate the performance prediction of a preacher, like the prediction accuracy increases, the mean square error decreases [12].

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\bar{y} - y)^2}$$

(19)

3. STUDY CASE

To better understand this method of identification and verify its effectiveness, we propose to model the stock level of a distribution company, that stock is divided into two families, high turnover and average rotation, for items low rotation, are not a focus for this study.

The warehouse of this company is managed according to a procurement policy, receptions are based on estimates prepared by the purchase and supply department, while orders are random (according to customer demand), which requires good management stock never dropped out or have a costly overstock. The procurement method followed for the management of this warehouse is an anticipation of demand (forecasts are not included in this study).

The stock will be considered a MIMO (multiple input, single output) or entries are the receipts and shipments while the output is the stock level for this study were asked a few assumptions:

We have considered the stock as an invariant system in time (the stock characteristics do not change over time).

We took the time unit equal to one day.

After receiving the items are entered into the management system and immediately appear on the system.

In the first place we consider the stock as a black box with two inputs and one output (entries are: corresponds to the currently tuned merchandise, and represents the daily flow of shipping, while the output is the level of stock holding ).

![Figure 7: representation of the system as "black box"](image)

The input and output data were extracted from the warehouse information system. These data are used to model the stock by comparing models of structures ARX, ARMAX, OE and Box Jenkins [12].

The data is divided into two parts. The first part is used to determine the model of the system and the second is applied for the purpose of validating the model.

3.1. FOR FAST MOVING PRODUCTS:

![Figure 8: representation of inputs and outputs](image)
Figure 8 shows the change in the level of stock based on monthly shipments and receptions for a product with high rotation.

Figures 9, 10, 11, 12 representing simulation models found with parametric models, we find that the OE and BJ models give satisfactory results; the best fit is that of the OE model (fewer parameters and the best fit).

At the end of ensuring the results found previously, we conducted a correlation test for the 4 models, these results confirm that the best model for items with a high turnover is the OE.

The table below shows the settings and information for each model (the number of parameters, the percentage adjustment and FPE).
<table>
<thead>
<tr>
<th>Model</th>
<th>Best fit %</th>
<th>FPE</th>
<th>Polynomial equation</th>
</tr>
</thead>
</table>
| ARX   | 80,32%     | 2255| $A(z) = 1 - 1.003Z^{-1} + 0.01241Z^{-2} - 0.00195Z^{-3}$  
|       |            |     | $B_1(z) = 0.2954Z^{-1}$  
|       |            |     | $B_2(z) = 0.9856$  |
| ARMAX | 72,54%     | 2356| $A(z) = 1 - 1.509Z^{-1} - 0.04706Z^{-2} + 0.552Z^{-3} + 0.00639Z^{-4}$  
|       |            |     | $B_1(z) = -0.2786Z^{-1} + 0.5027Z^{-2} - 0.007711Z^{-3}$  
|       |            |     | $B_2(z) = 0.9842 - 0.4906Z^{-1} + 0.5725Z^{-2}$  
|       |            |     | $C(z) = 1 - 1.19Z^{-1} + 0.1903Z^{-2}$  |
| OE    | 93,11%     | 2081| $A(z) = -0.3264Z^{-1} - 0.3748Z^{-2} - 0.1426Z^{-3}$  
|       |            |     | $B_1(z) = 0.9982 - 0.9306Z^{-1}$  
|       |            |     | $F_1(z) = 1 + 0.008749Z^{-1} + 0.9447Z^{-2}$  
|       |            |     | $F_2(z) = 1 - 1.931Z^{-1} + 0.9314Z^{-2}$  |
| BJ    | 88,43%     | 2056| $B_1(z) = 0.3433Z^{-1} - 0.4435Z^{-2} + 0.1853Z^{-3}$  
|       |            |     | $B_2(z) = 0.9881 - 0.8789Z^{-1}$  
|       |            |     | $C(z) = 1 + 1.137Z^{-1} + 0.179Z^{-2}$  
|       |            |     | $D(z) = 1 + 0.4123Z^{-1} - 0.3263Z^{-2}$  
|       |            |     | $F_1(z) = 1 - 1.88Z^{-1} + 0.8898Z^{-2}$  
|       |            |     | $F_2(z) = 1 - 1.891Z^{-1} + 0.8935Z^{-2}$  |

3.2 For medium rotation products:

Figures 14 shows the change in the level of stock based on monthly shipments and receptions for a product with high rotation.

Figures 15 representing simulation models found with parametric models, we find that the OE and BJ models give satisfactory results, the best fit is that of the OE model (fewer parameters and the best fit).
To ensure the results found previously, we conducted a correlation test for the 4 models, these results confirm that the best model for items to medium rotation is the OE.

The table below shows the settings and information for each model (the number of parameters, the percentage adjustment and FPE).

Figure 15: the comparison of the actual stock level and the estimated stock level for the ARX, ARMAX, OE and BJ models

Figure 16: result of the correlation tests of the four models
<table>
<thead>
<tr>
<th>Model</th>
<th>Best fit %</th>
<th>FPE</th>
<th>Polynomial equation</th>
</tr>
</thead>
</table>
| ARX    | 71.38%     | 6897.104 | \[A(z) = 1 - 1.102z^{-1} - 0.1436z^{-2} + 0.02692z^{-3} - 0.123z^{-4} + 0.1096z^{-5} - 0.03473z^{-6}\]  
|        |            |       | \[B_1(z) = 0.9795z^{-1} - 0.08847z^{-2}\]  
|        |            |       | \[B_2(z) = -4.133z^{-1} + 6.349z^{-2} - 3.324z^{-3} - 0.4717z^{-4} + 4.732z^{-5}\] |
| ARMAX  | 72.74%     | 2356  | \[A(z) = 1 - 0.9235z^{-1} - 0.1506z^{-2} + 0.089z^{-3}\]  
|        |            |       | \[B_1(z) = 0.957z^{-1} + 0.09434z^{-2}\]  
|        |            |       | \[B_2(z) = -4.244z^{-1} + 6.226z^{-2}\]  
|        |            |       | \[C(z) = 1 - 0.2277z^{-1} - 0.1233z^{-2} - 0.2191z^{-3}\] |
| OE     | 88.89%     | 2081  | \[B_1(z) = -0.3264z^{-1} - 0.3748z^{-2} - 0.1426z^{-3}\]  
|        |            |       | \[B_2(z) = 0.9982 - 0.9306z^{-1}\]  
|        |            |       | \[F_1(z) = 1 + 0.008749z^{-1} + 0.9447z^{-2}\]  
|        |            |       | \[F_2(z) = 1 - 1.931z^{-1} + 0.9314z^{-2}\] |
| BJ     | 82.89%     | 2056  | \[B_1(z) = 0.3433z^{-1} - 0.4435z^{-2} + 0.1853z^{-3}\]  
|        |            |       | \[B_2(z) = 0.9881 - 0.8789z^{-1}\]  
|        |            |       | \[C(z) = 1 + 1.137z^{-1} + 0.179z^{-2}\]  
|        |            |       | \[D(z) = 1 + 0.4123z^{-1} - 0.3263z^{-2}\]  
|        |            |       | \[F_1(z) = 1 - 1.88z^{-1} + 0.8898z^{-2}\]  
|        |            |       | \[F_2(z) = 1 - 1.891z^{-1} + 0.8935z^{-2}\] |

### 4. CONCLUSION

In this paper we proposed a new method of identifying a logistics system (warehouse) with parametric models, based on receipts and shipments as inputs and output of the system is the level of stock. We chose quad models known for their simplicities (ARX, ARMAX, OE and BJ), the comparative study mounted that the best model for the two categories of products (high and medium rotation) are given by the OE model, with 93.11% for products with a high turnover and 88.89% for the items to medium rotation, these results will allow us to have an idea of the dynamics of the stock and be able to do simulations to better manage the warehouse.
REFERENCES:


