Optimal Raw Material Inventory Analysis Using Markov Decision Process with Policy Iteration Method

Ikhsan Maulidi[1], Radhiah[2], Chintya Hayati[3], Vina Apriliani[4*]

[1,2,3]Department of Mathematics, Syiah Kuala University, Indonesia
[4]Department of Mathematics Education, UIN Ar-Raniry Banda Aceh, Indonesia
ikhsanmaulidi@unsyiah.ac.id[1], vina.apriliani@ar-raniry.ac.id[4]

ABSTRACT

Inventory of raw materials is a big deal in every production process, both in company production and home business production. In order to meet consumer demand, a business must be able to determine the amount of inventory that should be provided. The purpose of this research is to choose an alternative selection of ordering raw materials that produce the maximum amount of raw materials with minimum costs. The raw material referred to in this study is pandan leaves used to make pandan mats. Analysis of raw material inventory used in this research was the Markov decision process with the policy iteration method by considering the discount factor. From the analysis conducted, it is obtained alternative policies that must be taken by producers to meet raw materials with minimum costs. The results of this study can be a consideration for business actors in the study location in deciding the optimal ordering policy that should be taken to obtain the minimum operational cost.

—————————— ◆ ——————————

A. INTRODUCTION

Materials or goods stored that will be used to fulfill the production process are called inventories. Inventory of raw materials is a major factor that is important to companies in supporting the smooth operation of the production process, both in large companies and in small companies such as home-based business products or small "home made" products. Raw material is a material that is used in the manufacture of a product, then processed so that it becomes a finished product for sale (Ristono, 2009).

In optimal control of raw material inventory, the company must provide an amount of certain raw material at a certain time. This kind of procurement is due to the fact that the number of arrivals of requests cannot be known with certainty, so it will lead to a less than optimal inventory level and cause costs that should be measured. In an effort to overcome the fluctuating demand for raw material prices, a method is needed that can link current demand with previous demand. One of them is by using the Markov decision process (Mani et al., 2021).

The Markov decision process is a system that can move one particular state to another possible state by considering several alternative policies (Hasan & Iqbal, 2004). The decision-maker must take an action from those set of alternatives. This action affects the probability of the transition to the next move and brings a profit or loss after that. The transition probability matrix and the income (cost) matrix depend on the decision alternatives that can be used. The goal is to determine decisions that optimize the expected revenue or costs of that process.

The applications of Markov decision process can be seen in the field of medical science (Bennett & Hauser, 2013) which studies how to determine optimal decisions in health services. Furthermore, in the field of telecommunication (Ksentini et al., 2014; Y. J. Liu et al., 2017), each of which studies the transfer of telecommunications services and election for machine type communications using reinforcement learning based Markov decision process. Research in other fields can also be seen such as in the field of transportation (Iversen et al., 2014; Ong & Kochenderfer, 2017; Shou et al., 2020), social science (Rong et al., 2016), game application (Zheng & Siami Namin, 2018), and Internet of Things (Yousefi et al., 2020).

One method of making decisions in the Markov decision process is the policy iteration method. In this method, the policies taken are evaluated first, and then policy improvements are made so that convergent improvement results are obtained (Feinberg & Shwartz, 2002). An example of an analysis using the policy iteration method in previous research can be seen in (Fürnkranz et al., 2012), (D. Liu & Wei, 2013), (Luo et al., 2014), (Alla et al., 2015), (Pérolat et al., 2016), (Wu & Shen, 2017), and (Yang & Wei, 2018).

One of the cases related to the problem in this research is the inventory of raw materials. The raw material used in this research is pandan leaves. Pandan leaves are one of the raw materials for woven pandanus mats or in Acehnese called "Tika Seukee" which are generally obtained from self-cultivated gardens and can be cultivated quickly and easily in Aceh, especially in North Aceh Regency. If the raw materials are not sufficient, then the craftsmen can obtain raw materials from wholesale results obtained from other regional farmers. Therefore, this study provides an application of the Markov decision process to determine the optimal inventory of raw materials in the pandanus mat weaving business. Research related to inventory with Markov decision process can be seen in (Noorida, 2003), (Sarjono et al., 2011), (Layla, 2016), (Ferreira et al., 2018), and (Oktaviyani et al., 2018). For example, (Noorida, 2003) studies an analysis of fire tube boiler stock inventory using the policy iteration method without considering the discount factor. The thing that distinguishes this research from what has been done is that in this study an inventory analysis was carried out by involving the discount factor. This means that the price of raw materials is expected to increase in the future and this is more realistic in real terms.

B. METHODS

The data used are primary data obtained directly from woven pandanus mat craftsmen in Meunasah Aron Village, Muara Batu District, North Aceh Regency. The data in Table 1 is the amount of raw material inventory (Pandan leaves) for 2 years, starting from January to December 2018 and 2019, as shown in Table 1.

Table 1. The Amount of Raw Material Inventory in 2 Years

Month	Year 2018 (Kg)	Year 2019 (Kg)
January	24	34
February	16	24
March	22	30
April	26	25
May	30	35
June	35	45
July	20	42
August	28	18
September	24	20
October	26	25
November	18	41
December	27	32

In addition to the raw material inventory data, data on raw material price, raw material ordering cost, and raw material storage cost are given as follows:
1. The raw material price for pandan leaves is IDR 50.000/Kg.
2. The average of raw material ordering cost (IDR/Month) in 2 years (2018 and 2019) is presented in Table 2.

Table 2. The Average of Raw Material Ordering Costs in 2 Years

No	Cost Type	Cost (IDR/Month)
1	Shipping cost to the destination	350.000
2	Inspection cost	30.000
3	Miscellaneous cost	50.000
	Total	430.000

3. The average of raw material storage cost (IDR/Kg) in 2 years (2018 and 2019) is presented in Table 3.

Table 3. The Average of Raw Material Storage Costs in 2 Years

No	Cost Type	Percent of Raw Material Prices (%)	Cost (IDR/Kg)
1	Warehouse facility cost (maintenance cost)	4	2.000
2	Miscellaneous cost	2	1.000
	Total	6	3.000

The steps for solving the problem with the Markov decision process are given as follows:
1. Determine the possible state for the initial inventory (i) and determine the ordering-level alternative decisions (x) from the results of the frequency distribution.
2. Determine the transition probability matrix of the obtained state. The probability of each initial inventory state i to the initial inventory state j at each ordering-level alternative (x) is $p_{ij}(x)$, where the initial inventory state i will transition to the initial inventory state j for $j = i + x - d$ and the result of probabilities is arranged in a square matrix P with the elements in the P matrix being $p_{ij}(x)$, where $p_{ij}(x) = p(d)$, $p(d)$ represents the probability of raw materials ordering demand. The basis for determining this probability

can be studied in (Taylor & Samuel, 1998), (Ching & Ng, 2006), and (Ross, 2010).

3. Find the shortage cost. In each state i and each decision x, the shortage cost can be calculated using the equation:

$$E = a + b\left(\sum_{d>i+x}(d-i-x)p(d)\right), i = 0, \cdots n \tag{1}$$

with:

a: ordering costs, can be seen in Table 2, which is IDR 430.000,

b: storage costs, can be seen in Table 3, which is IDR 3.000,

d : raw materials ordering demand, provided that $d > i + x$,

i : initial inventory state,

x: ordering-level alternative decisions.

4. Determine the total cost matrix. The total cost of inventory for each initial inventory (i) and each ordering-level alternative (x) is given by the equation:

$$C_i(x) = a + bi + E \tag{2}$$

with:

$C_i(x)$: total cost of each initial inventory (i) at each ordering-level alternative (x),

a : ordering costs, can be seen in Table 2, which is IDR 430.000,

b : storage costs, can be seen in Table 3, which is IDR 3.000,

E : shortage cost.

Furthermore, for each calculation of the total cost of inventory at each initial inventory (i) and each ordering-level alternative (x), the smallest total cost will be chosen. The total cost matrix is arranged in a column C matrix with the entries being $C_i(x)$.

5. Calculating the optimal solution using the policy iteration method with a discount factor. The discount factor (α) is a multiplier to calculate the future value of money when valued in the present time. The presence of a discount factor ($\alpha < 1$) can result in a change in the optimal policy, compared to the case without a discount ($\alpha = 1$) (Littman et al., 2013). The steps for the solution are as follows:

a. Determine the initial policy $x_i^{(k)}$ marked with the initial policy ($k = 0$) and take any ordering-level alternative $x = x_i^{(k)}$ for each initial inventory (i), and construct a probability matrix and its cost matrix.

b. Evaluate the routine policy is to determine $f_i^{(k)}$ for each initial inventory (i) which is the solution to the linear equation, namely in the equation:

$$f_i^{(k)} = C_i(x) + \alpha\sum_j p_{ij}(x)f_j^{(k)}, i,j = 0, \cdots, n \tag{3}$$

with:

$\sum_j p_{ij}(x)f_j^{(k)}$: the probability sum of each initial inventory state (i) to the initial inventory state (j) in each ordering-level alternative (x) for the initial inventory state (j),

$C_i(x)$: minimum total cost,

α : discount factor, discount factor must be $\alpha < 1$, while the discount factor here is 0.98,

k : initial policy, that is $k = 0$.

To solve the system of linear equations, LINDO (Linear Interactive Discrete Optimizer) program is used.

c. Improvements to the routine policy $x_i^{(k+1)}$ is determining a new policy $(k+1)$ by finding an ordering-level alternative (x), $x = x_i^{(k+1)}$ for each initial inventory (i) that can be seen in the equation:

$$\left(C_i(x) + \alpha \sum_j p_{ij}(x) f_j^{(k)} \right), i, j = 0, \cdots n. \tag{4}$$

with:

$\sum_j p_{ij}(x) f_j^{(k)}$: the probability sum of each initial inventory state (i) to the initial inventory state (j) in each ordering-level alternative (x) for the initial inventory state (j),

$C_i(x)$: total cost for each initial inventory (i) with each ordering-level alternative (x),

α : discount factor, discount factor must be $\alpha < 1$, while the discount factor here is 0.98,

k : new policy $k + 1 = 1$.

d. If the new policy $k + 1$ differs from the initial policy $k = 0$ by at least one state, increase the count of k by one and return to step two, as shown in Figure 1.
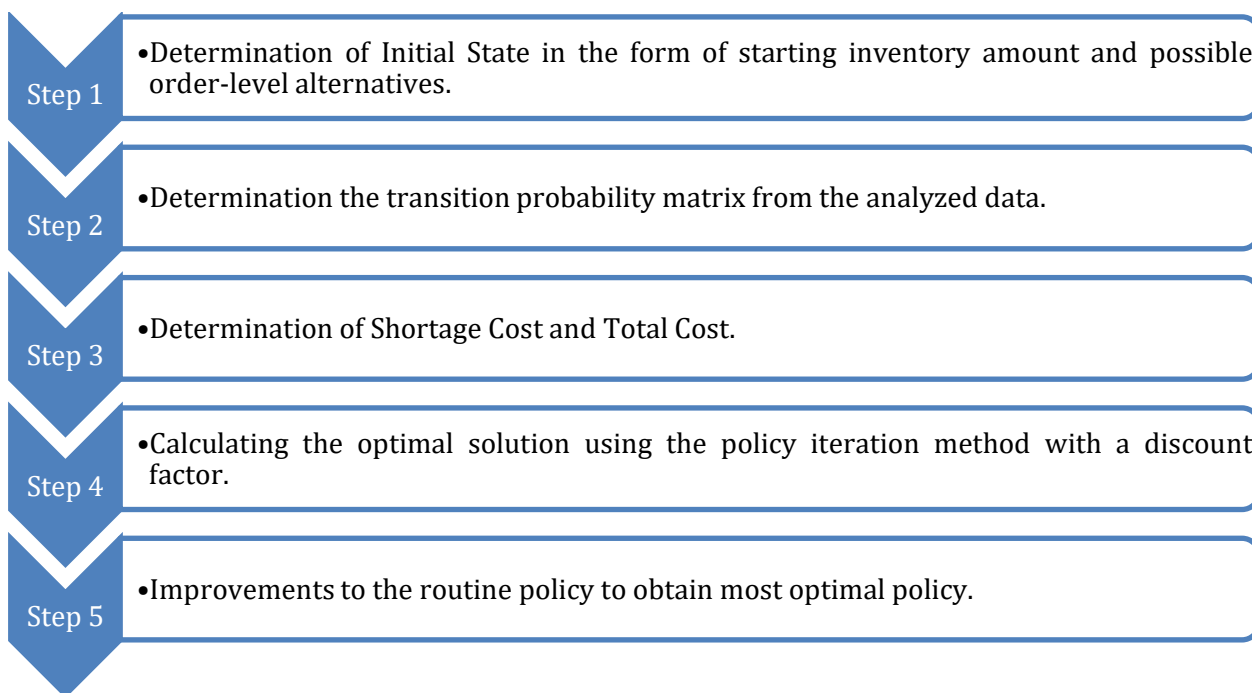
•Determination of Initial State in the form of starting inventory amount and possible order-level alternatives.
Step 1

•Determination the transition probability matrix from the analyzed data.
Step 2

•Determination of Shortage Cost and Total Cost.
Step 3

•Calculating the optimal solution using the policy iteration method with a discount factor.
Step 4

•Improvements to the routine policy to obtain most optimal policy.
Step 5

Figure 1. Problem-Solving Flow Using the Policy Iteration Method

C. RESULT AND DISCUSSION

1. Frequency Distribution Table

Data on the amount of raw material inventory for two years in Table 1 is presented in a frequency distribution Table 4.

Table 4. Frequency Distribution of Raw Material Inventory Used

Amount Used	Raw Materials Ordering Demand (d)	Frequency	Probability $P(D = d)$
16 − 20 kg	20 kg	5	0,21
21 − 25 kg	25 kg	6	0,25
26 − 30 kg	30 kg	6	0,25
31 − 35 kg	35 kg	4	0,17
36 − 40 kg	40 kg	0	0
41 − 45 kg	45 kg	3	0,12

Table 4 describes about the frequency distribution of raw material inventory used, which contains raw materials ordering demand, its frequency, and the probability for each interval class. Frequency is the amount of data included I n each interval class, namely the number of raw materials used and the number of frequencies is the total number of data n. Raw materials ordering demand (d) is taken from the number of upper edges of each interval class. While the probability is $\frac{frequency}{amount\ of\ data\ (n)}$. The probability of raw materials ordering demand (d) for the first interval class (16-20 kg) is $5/24 = 0.21$. The same way is obtained for the other intervals.

2. Analysis of the Markov Decision Process

The Markov decision process has elements, namely state (i) and decision (x), in this case the state is referred to as initial inventory (i) and the decision is referred to as an ordering-level alternative (x). From Table 4, two assumptions are obtained, where the first assumption is the possible state for initial inventory (i) as many as 6 states are obtained starting from 0 kg, 5 kg, 10 kg, 15 kg, 20 kg, 25 kg, where the distance between the top edges each interval class is 5 kg. Second, it can be assumed that the ordering-level alternatives (x) are 20 kg, 25 kg, 30 kg, 35 kg, 40 kg, 45 kg which is the upper edge of each interval class, and the demand for raw material ordering (d) corresponds to the ordering-level alternatives. The order is 20 kg, 25 kg, 30 kg, 35 kg, 40 kg, or 45 kg. Table 5 is a table of the state of the initial inventory (i) and the ordering-level alternative (x), as shown in Table 5.

Table 5. Initial Inventory (i) and Ordering-Level Alternative (x)

Initial Inventory (Kg)	Ordering-Level Alternative (Kg)					
0	20	25	30	35	40	45
5	20	25	30	35	40	-
10	20	25	30	35	-	-
15	20	25	30	-	-	-
20	20	25	-	-	-	-
25	20	-	-	-	-	-

For example, if the initial inventory (i) is 0 kg, then there are 6 alternative choices for the ordering-level (x), namely orders of 20 kg, 25 kg, 30 kg, 35 kg, 40 kg, and 45 kg. The total for each line of raw material inventory must be less or equal with 45 kg, because in Table 1 it can be seen that the maximum amount of raw material inventory is 45 kg. The empty entry from Table 5 describes that there is no alternative to order in that ordering-level alternatives.

3. Transition Probability, Shortage Cost, and Total Cost

The probability of each initial inventory state i to the initial inventory state j at each ordering-level alternative (x) is $p_{ij}(x)$, where the initial inventory state i will transition to the initial inventory state j for $j = i + x - d$ and the result of probabilities is arranged in a square matrix P with the elements in the P matrix being $p_{ij}(x)$, where $p_{ij}(x) = p(d)$, for $p(d)$ represents the probability of raw materials ordering demand. For each initial inventory (i) and each ordering-level alternative (x), the shortage cost can be calculated using equation (1). The total cost (C) for each initial inventory (i) at each ordering-level alternative (x) can be calculated using equation (2).

The following steps are given to calculate the transition probability (p_{ij}), shortage cost (E), and total cost $\left((C_i(x)\right)$ using Table 4 and Table 5 for **initial inventory (i) 0 kg**. From Table 5, if the initial inventory is 0 kg, then there are 6 alternative choices for the ordering-level (x), namely the order of 20 kg, 25 kg, 30 kg, 35 kg, 40 kg, and 45 kg, and the raw materials ordering demand (d) begins of 20 kg, 25 kg, 30 kg, 35 kg, 40 kg, and 45 kg. Calculation of the raw materials ordering demand (d) stops when $j = 0$, where $i = j$. The probability value can be seen in Table 4. For the ordering-level (x) 20 kg, then:

a. Transition Probability $p_{ij}(x)$

If $x = 20$ kg and $j = 0$ kg then $j = i + x - d \rightarrow d = i + x - j = 0 + 20 - 0 = 20$.
$p_{ij}(x) \quad = p(D \geq d)$
$p_{0,0}(20) = p(D \geq 20)$
$p_{0,0}(20) = 0{,}21 + 0{,}25 + 0{,}25 + 0{,}17 + 0 + 0{,}12 = 1.$

b. Shortage Cost

For $i = 0$ kg, $x = 20$ kg, with $d > i + x$ then d starts from 25 kg, 30 kg, 35 kg, 40 kg, and 45 kg, and the value of $p(d)$ can be seen in Table 4.

$$E = a + b\left(\sum_{d>i+x} (d - i - x)p(d)\right)$$

$E = $ IDR 430.000 + IDR 3.000[(25 - 0 - 20)(0{,}25) + (30 - 0 - 20)(0{,}25)$
$E = +(35 - 0 - 20)(0{,}17) + (40 - 0 - 20)(0) + (45 - 0 - 20)(0{,}12)]$
$E = $ IDR 457.900.

c. Total Cost

For $i = 0$ kg, $x = 20$ kg,
$C_i(x) = a + bi + E$
$C_0(20) = a + bi + E$
$C_0(20) = $ IDR 430.000 + IDR 3.000 (0) + IDR 457.900
$C_0(20) = $ IDR 887.900.

So, the total cost required for initial inventory (i) 0 kg at the ordering-level alternative (x) 20 kg is IDR 887.900. In the same way, the results of the calculation of the transition probability, shortage cost, and total cost for each initial inventory with ordering-level alternative (x) of 20 kg, 25 kg, 30 kg, 35 kg, 40 kg, and 45 kg can be summarized in Table 6, Table 7, and Table 8.

Table 6. Transition Probability of Initial Inventory State for Each Policy that can be taken $\left(p_{i,j}(x)\right)$

Initial Inventory State	Policy	Initial Inventory State					
		0	5	10	15	20	25
0	20	1	0	0	0	0	0
	25	0,79	0,21	0	0	0	0
	30	0,54	0,25	0,21	0	0	0
	35	0,29	0,25	0,25	0,21	0	0
	40	0,12	0,17	0,25	0,25	0,21	0
	45	0,12	0	0,17	0,25	0,25	0,21
5	20	0,79	0,21	0	0	0	0
	25	0,54	0,25	0,21	0	0	0
	30	0,29	0,25	0,25	0,21	0	0
	35	0,12	0,17	0,25	0,25	0,21	0
	40	0,12	0	0,17	0,25	0,25	0,21
10	20	0,54	0,25	0,21	0	0	0
	25	0,29	0,25	0,25	0,21	0	0
	30	0,12	0,17	0,25	0,25	0,21	0
	35	0,12	0	0,17	0,25	0,25	0,21
15	20	0,29	0,25	0,25	0,21	0	0
	25	0,12	0,17	0,25	0,25	0,21	0
	30	0,12	0	0,17	0,25	0,25	0,21
20	20	0,12	0,17	0,25	0,25	0,21	0
	25	0,12	0	0,17	0,25	0,25	0,21
25	20	0,12	0,17	0,25	0,25	0,21	0

Table 6 illustrates the transition probability from an initial state stock to an initial state stock based on the policies selected. A value of zero indicates that the transition is not possible. This is because the value of ordering raw materials is a positive variable. It is not possible that the transitional state will have a smaller value than the initial supply state, as shown in Table 7.

Table 7. Shortage Cost of Initial Inventory for Each Policy that can be taken

Initial Inventory (Kg)	Shortage Cost due to the Policies Taken (IDR)					
	20	25	30	35	40	45
0	457.900	446.050	437.950	433.600	431.800	430.000
5	446.050	437.950	433.600	431.800	430.000	
10	437.950	433.600	431.800	430.000		
15	433.600	431.800	430.000			
20	431.800	430.000				
25	430.000					

Table 7 illustrates the amount of Shortage Cost for each decision taken if the initial condition of the inventory is known. The decisions taken for each initial inventory have been given in Table 5. It can be seen that the Shortage Cost value is getting smaller for larger alternative orders, as shown in Table 8.

Table 8. Total Cost of Initial Inventory for Each Policy that can be taken

Initial Inventory (Kg)	Total Cost due to the Policies Taken (IDR)					
	20	25	30	35	40	45
0	887.900	876.050	867.950	863.600	861.800	860.000
5	891.050	882.950	878.600	876.800	875.000	
10	897.950	893.600	891.800	890.000		
15	908.600	906.800	905.000			
20	921.800	920.000				
25	935.000					

Table 8 illustrates the total cost value for each policy taken if the current inventory state is known. The value of this total cost looks bigger if the initial inventory state is higher. This can happen because of the costs needed to handle inventory items that are still stored to be used in the future and also there is a fixed ordering cost that presented in Table 2.

4. Policy Iteration Method with Discount Factor

The following are some steps to complete the optimal solution using the policy iteration method with discount factor.

a. Determine the Initial Policy

Determining the initial policy $\left(x_i^{(k)}\right)$ marked with the initial policy $(k = 0)$ and take any ordering-level alternative (x) i.e. $x = x_i^{(k)}$ for each initial inventory (i) and construct a probability matrix and its cost matrix.

For initial inventory of 0 kg, 5 kg, 10 kg, 15 kg, 20 kg, and 25 kg, and ordering-level alternative of 20 kg, 25 kg, 30 kg, 35 kg, 40 kg and 45 kg, then we choose any initial policy $x = x_i^{(k)}$ as follows:

$$x_0^{(0)} = 45, x_5^{(0)} = 40, x_{10}^{(0)} = 35, x_{15}^{(0)} = 30, x_{20}^{(0)} = 25, x_{25}^{(0)} = 20.$$

The meaning of $x = x_i^{(k)}$ is that if the initial inventory is at level i kg, with an initial policy of k, then the optimal of ordering-level alternative is x kg. For this initial policy, the probability matrix is as follows:

$$P = p_{i,j}(x) = \begin{bmatrix} 0,12 & 0 & 0,17 & 0,25 & 0,25 & 0,21 \\ 0,12 & 0 & 0,17 & 0,25 & 0,25 & 0,21 \\ 0,12 & 0 & 0,17 & 0,25 & 0,25 & 0,21 \\ 0,12 & 0 & 0,17 & 0,25 & 0,25 & 0,21 \\ 0,12 & 0 & 0,17 & 0,25 & 0,25 & 0,21 \\ 0,12 & 0 & 0,17 & 0,25 & 0,25 & 0,21 \end{bmatrix}.$$

The minimum total cost matrix is as follows:

$$C = \begin{bmatrix} C_0(45) \\ C_5(40) \\ C_{10}(35) \\ C_{15}(30) \\ C_{20}(25) \\ C_{25}(20) \end{bmatrix} = \begin{bmatrix} \text{IDR 860.000} \\ \text{IDR 875.000} \\ \text{IDR 890.000} \\ \text{IDR 905.000} \\ \text{IDR 920.000} \\ \text{IDR 935.000} \end{bmatrix}.$$

b. Evaluate the Routine Policy

Evaluate the routine policy, which is to determine $f_i^{(k)}$ for each initial inventory (i) which is the solution to a linear equation using equation (3). We obtained 6 linear equations as follows:

$f_0^{(0)} = 860.000 + (0,1176)f_0^{(0)} + (0)f_5^{(0)} + (0,1666)f_{10}^{(0)} + (0,245)f_{15}^{(0)} + (0,245)f_{20}^{(0)} + (0,2058)f_{25}^{(0)}$,

$f_5^{(0)} = 875.000 + (0,1176)f_0^{(0)} + (0)f_5^{(0)} + (0,1666)f_{10}^{(0)} + (0,245)f_{15}^{(0)} + (0,245)f_{20}^{(0)} + (0,2058)f_{25}^{(0)}$,

$f_{10}^{(0)} = 890.000 + (0,1176)f_0^{(0)} + (0)f_5^{(0)} + (0,1666)f_{10}^{(0)} + (0,245)f_{15}^{(0)} + (0,245)f_{20}^{(0)} + (0,2058)f_{25}^{(0)}$,

$f_{15}^{(0)} = 905.000 + (0,1176)f_0^{(0)} + (0)f_5^{(0)} + (0,1666)f_{10}^{(0)} + (0,245)f_{15}^{(0)} + (0,245)f_{20}^{(0)} + (0,2058)f_{25}^{(0)}$,

$f_{20}^{(0)} = 920.000 + (0,1176)f_0^{(0)} + (0)f_5^{(0)} + (0,1666)f_{10}^{(0)} + (0,245)f_{15}^{(0)} + (0,245)f_{20}^{(0)} + (0,2058)f_{25}^{(0)}$,

$f_{25}^{(0)} = 935.000 + (0,1176)f_0^{(0)} + (0)f_5^{(0)} + (0,1666)f_{10}^{(0)} + (0,245)f_{15}^{(0)} + (0,245)f_{20}^{(0)} + (0,2058)f_{25}^{(0)}$.

From the 6 equations above, the solution for $f_0^{(0)}, f_5^{(0)}, f_{10}^{(0)}, f_{15}^{(0)}, f_{20}^{(0)}, f_{25}^{(0)}$ obtained by using the LINDO application. The final solutions have been summarized as follows:

$$f_0^{(0)} = \text{IDR 45.308.000},$$
$$f_5^{(0)} = \text{IDR 45.323.000},$$
$$f_{10}^{(0)} = \text{IDR 45.338.000},$$
$$f_{15}^{(0)} = \text{IDR 45.353.000},$$
$$f_{20}^{(0)} = \text{IDR 45.368.000},$$
$$f_{25}^{(0)} = \text{IDR 45.383.000}.$$

c. Improvements to the Routine Policy

Improvements to the routine policy $\left(x_i^{(k+1)}\right)$ is determining a new policy $(k + 1)$. In the first step, it is stated that the initial policy is $k = 0$, then for the new policy $k + 1 = 1$. Then find an ordering-level alternative (x) that is $x = \left(x_i^{(k+1)}\right)$ for each initial inventory (i) using equation (4).

After getting the value of $f_0^{(0)}, f_5^{(0)}, f_{10}^{(0)}, f_{15}^{(0)}, f_{20}^{(0)}, f_{25}^{(0)}$ from the second step then substitute it into equation (4). The results of the calculations can be summarized as follows.

For $f_5^{(0)} = \text{IDR 45.323.000}$, which is $i = 5$ kg, when $x = 40$ kg we get $\left(C_5(x) + \alpha \sum_j p_{5j}(x)f_j^{(k)}\right) = \text{IDR 45.323.000}$,

For $f_{10}^{(0)} = \text{IDR 45.338.000}$, which is $i = 10$ kg, when $x = 35$ kg we get $\left(C_{10}(x) + \alpha \sum_j p_{10j}(x)f_j^{(k)}\right) = \text{IDR 45.338.000}$,

For $f_{15}^{(0)} =$ IDR 45.353.000, which is $i = 15$ kg, when $x = 30$ kg we get $\left(C_{15}(x) + \alpha \sum_j p_{15j}(x) f_j^{(k)} \right) =$ IDR 45.353.000,

For $f_{20}^{(0)} =$ IDR 45.368.000, which is $i = 20$ kg, when $x = 25$ kg we get $\left(C_{20}(x) + \alpha \sum_j p_{20j}(x) f_j^{(k)} \right) =$ IDR 45.368.000,

For $f_{25}^{(0)} =$ IDR 45.383.000, which isi $i = 25$ kg, when $x = 20$ kg we get $\left(C_{25}(x) + \alpha \sum_j p_{25j}(x) f_j^{(k)} \right) =$ IDR 45.383.000,

then we obtained a new policy with an ordering-level alternative $(x) = x_i^{(1)}$:

$$x_0^{(1)} = 45, x_5^{(1)} = 40, x_{10}^{(1)} = 35, x_{15}^{(1)} = 30, x_{20}^{(1)} = 25, x_{25}^{(1)} = 20.$$

Because the new policy $(k = 1)$ is the same as the initial policy $(k = 0)$, the iteration is stopped. The decision results of each initial inventory (i) with each ordering-level alternative (x) are the optimal results with optimal costs as follows:

$$C_0(45) = \text{IDR } 860.000,$$
$$C_5(40) = \text{IDR } 875.000,$$
$$C_{10}(35) = \text{IDR } 890.000,$$
$$C_{15}(30) = \text{IDR } 905.000,$$
$$C_{20}(25) = \text{IDR } 920.000,$$
$$C_{25}(20) = \text{IDR } 935.000.$$

This means:

1) If the initial inventory (i) is at the level of 0 kg, then the optimal ordering-level alternative (x) is 45 kg with a minimum cost of IDR 860.000.
2) If the initial inventory (i) is at the level of 5 kg, then the optimal ordering-level alternative (x) is 40 kg with a minimum cost of IDR 875.000.
3) If the initial inventory (i) is at the level of 10 kg, then the optimal ordering-level alternative (x) is 35 kg with a minimum cost of IDR 890.000.
4) If the initial inventory (i) is at the level of 15 kg, then the optimal ordering-level alternative (x) is 30 kg with a minimum cost of IDR 905.000.
5) If the initial inventory (i) is at the level of 20 kg, then the optimal ordering-level alternative (x) is 25 kg with a minimum cost of IDR 920.000.
6) If the initial inventory (i) is at the level of 25 kg, then the optimal ordering-level alternative (x) is 20 kg with a minimum cost of IDR 935.000.

The calculation results show that for each initial inventory state (i), the alternative order level that must be selected in order to produce a minimum cost is $45 - i$ kg. However, the resulting minimum cost for a smaller starting inventory is greater than the greater initial inventory one. This can happen because of the storage costs and shortage costs as a result of a shortage of inventory. This result has a behavior similar to what has been done in previous studies by (Noorida, 2003).

D. CONCLUSION AND SUGGESTIONS

This research has reached the optimal solution using the Markov decision process with the policy iteration method. This policy iteration method can achieve an optimal solution in a small number of iterations. In this study, the calculation results obtained from the policy iteration method prove that no iteration occurs, it means that for the data in this research case the optimal solution has been reached using the Markov decision process with the policy iteration method. Suggestions that can be conducted in further research are to consider other operational factors such as the length of work on a product and the honorarium expenses to produce.

REFERENCES

Alla, A., Falcone, M., & Kalise, D. (2015). An efficient policy iteration algorithm for dynamic programming equations. *SIAM Journal on Scientific Computing*, *37*(1), A181–A200. https://doi.org/https://doi.org/10.1137/130932284

Bennett, C. C., & Hauser, K. (2013). Artificial intelligence framework for simulating clinical decision-making: A Markov decision process approach. *Artificial Intelligence in Medicine*, *57*(1), 9–19. https://doi.org/10.1016/j.artmed.2012.12.003

Ching, W. K., & Ng, M. K. (2006). *Markov Chains : Models, Algoritms and Application*. Springer Science.

Feinberg, E. A., & Shwartz, A. (2002). Handbook of Markov Decision Processes: Handbook and Applications. In *Kluwer's International Series* (Issue December).

Ferreira, G. O., Arruda, E. F., & Marujo, L. G. (2018). Inventory management of perishable items in long-term humanitarian operations using Markov decision processes. *International Journal of Disaster Risk Reduction*, *31*, 460–469. https://doi.org/https://doi.org/10.1016/j.ijdrr.2018.05.010

Fürnkranz, J., Hüllermeier, E., Cheng, W., & Park, S. H. (2012). Preference-based reinforcement learning: a formal framework and a policy iteration algorithm. *Machine Learning*, *89*(1), 123–156. https://doi.org/https://doi.org/10.1007/s10994-012-5313-8

Hasan, & Iqbal, M. (2004). *Pokok-Pokok Teori Pengambilan Keputusan*. Ghalia Indonesia.

Iversen, E. B., Morales, J. M., & Madsen, H. (2014). Optimal charging of an electric vehicle using a Markov decision process. *Applied Energy*, *123*, 1–12. https://doi.org/https://doi.org/10.1016/j.apenergy.2014.02.003

Ksentini, A., Taleb, T., & Chen, M. (2014). A Markov decision process-based service migration procedure for follow me cloud. *2014 IEEE International Conference on Communications (ICC)*, 1350–1354. https://doi.org/https://doi.org/10.1109/ICC.2014.6883509

Layla, M. (2016). Analisis persediaan bahan baku yang optimal menggunakan rantai Markov di PT. PDM Indonesia. *Jurnal Mantik Penusa*, *20*(1), 17–23. https://e-jurnal.pelitanusantara.ac.id/index.php/mantik/article/view/206

Littman, M. L., Dean, T. L., & Kaelbling, L. P. (2013). On the complexity of solving Markov decision problems. *The Eleventh Conference on Uncertainty in Artificial Intelligence*, 394–402. https://doi.org/https://doi.org/10.48550/arXiv.1302.4971

Liu, D., & Wei, Q. (2013). Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems. *IEEE Transactions on Neural Networks and Learning Systems*, *25*(3), 621–634. https://doi.org/https://doi.org/10.1109/TNNLS.2013.2281663

Liu, Y. J., Cheng, S. M., & Hsueh, Y. L. (2017). eNB selection for machine type communications using reinforcement learning based Markov decision process. *IEEE Transactions on Vehicular Technology*, *66*(12), 11330–11338. https://doi.org/https://doi.org/10.1109/TVT.2017.2730230

Luo, B., Wu, H. N., Huang, T., & Liu, D. (2014). Data-based approximate policy iteration for affine nonlinear continuous-time optimal control design. *Automatica*, *50*(12), 3281–3290. https://doi.org/https://doi.org/10.1016/j.automatica.2014.10.056

Mani, A., Bakar, S. A., Krishnan, P., & Yaacob, S. (2021). Markov decision process approach in the estimation of raw material quality in incoming inspection process. *Journal of Physics: Conference Series*, *2107*(1), 012025. https://doi.org/10.1088/1742-6596/2107/1/012025

Noorida, T. (2003). Analisa persediaan material pada fire tube boiler stock yang optimal dengan

menggunakan Markov chain. *Jurnal Teknik Industri*, *4*(2), 40–49. https://doi.org/https://doi.org/10.22219/jtiumm.vol4.no2.40-49

Oktaviyani, O., Dwijanto, D., & Supriyono, S. (2018). Optimization of production scheduling and planning of raw materials supply using Markov chains (a case study of Kinken Cake & Bakery Kutoarjo). *Unnes Journal of Mathematics*, *7*(2), 165–180. https://doi.org/https://doi.org/10.15294/ujm.v7i2.13894

Ong, H. Y., & Kochenderfer, M. J. (2017). Markov decision process-based distributed conflict resolution for drone air traffic management. *Journal of Guidance, Control, and Dynamics*, *40*(1), 69–80. https://doi.org/https://doi.org/10.2514/1.G001822

Pérolat, J., Piot, B., Geist, M., Scherrer, B., & Pietquin, O. (2016). Softened approximate policy iteration for Markov games. *International Conference on Machine Learning*, 1860–1868. http://proceedings.mlr.press/v48/perolat16.html

Ristono, A. (2009). *Manajemen Persediaan*. Graha Ilmu.

Rong, H., Zhou, X., Yang, C., Shafiq, Z., & Liu, A. (2016). The rich and the poor: a Markov decision process approach to optimizing taxi driver revenue efficiency. *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, 2329–2334. https://doi.org/https://doi.org/10.1145/2983323.2983689

Ross, S. M. (2010). *Introduction to Probability Models* (10th ed.). Academic Press.

Sarjono, H., Edwin, E., Sentosa, H., & Bong, F. (2011). Analisis Markov chain terhadap persediaan: studi kasus pada CV Sinar Bahagia Group. *Binus Business Review*, *2*(2), 1071–1076. https://doi.org/https://doi.org/10.21512/bbr.v2i2.1249

Shou, Z., Di, X., Ye, J., Zhu, H., Zhang, H., & Hampshire, R. (2020). Optimal passenger-seeking policies on E-hailing platforms using Markov decision process and imitation learning. *Transportation Research Part C: Emerging Technologies*, *111*, 91–113. https://doi.org/https://doi.org/10.1016/j.trc.2019.12.005

Taylor, H. M., & Samuel, K. (1998). *An Introduction to Stochastic Modelling* (3rd ed.). Academic Press.

Wu, Y., & Shen, T. (2017). Policy iteration algorithm for optimal control of stochastic logical dynamical systems. *IEEE Transactions on Neural Networks and Learning Systems*, *29*(5), 2031–2036. https://doi.org/https://doi.org/10.1109/TNNLS.2017.2661863

Yang, X., & Wei, Q. (2018). An off-policy iteration algorithm for robust stabilization of constrained-input uncertain nonlinear systems. *International Journal of Robust and Nonlinear Control*, *28*(18), 5747–5765. https://doi.org/https://doi.org/10.1002/rnc.4342

Yousefi, S., Derakhshan, F., Karimipour, H., & Aghdasi, H. S. (2020). An efficient route planning model for mobile agents on the internet of things using Markov decision process. *Ad Hoc Networks*, *98*(102053). https://doi.org/https://doi.org/10.1016/j.adhoc.2019.102053

Zheng, J., & Siami Namin, A. (2018). A Markov decision process to determine optimal policies in moving target. *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, 2321–2323. https://doi.org/https://doi.org/10.1145/3243734.3278489