MARKOVIAN DECISION PROCESS TO FIND OPTIMAL POLICIES IN THE MANAGEMENT OF AN ORANGE FARM

Juan Manuel Izar Landeta*, Carmen Berenice Ynzunza Cortés**, Héctor Méndez Azúa***
*Facultad de Ingeniería, Universidad Autónoma de San Luis Potosí, Manuel Nava # 8, Zona Universitaria Poniente, C. P. 78290, San Luis Potosí, S. L. P., México
**Universidad Tecnológica de Querétaro, Av. Pie de la Cuesta # 2501, Col. Unidad Nacional, C. P. 76148, Querétaro, Qro., México
*** Facultad de Ingeniería, Universidad Autónoma de San Luis Potosí, Manuel Nava # 8, Zona Universitaria Poniente, C. P. 78290, San Luis Potosí, S. L. P., México

ABSTRACT
This paper presents the application of a Markovian decision process to define the optimal policy in the case of an orange farm management.
The problem has been proposed to be solved with two approaches: policy iteration and linear programming, which have been formulated in first instance without discount factor and then applying various discount factors ranging from 0.5 to 0.9. The policy iteration methodology has been more efficient than linear programming, since it requires fewer calculations to achieve the optimal solution.
In all cases the result has been that the optimal policy is to better take care the orange farm to maximize economic benefits, which would reach an amount of 431 Mexican pesos per orange tree each year.

KEYWORDS*:* Markovian decision process, Markov chains, Optimization, Policy iteration, Linear programming

MSC: 60J10

RESUMEN
Este trabajo presenta la aplicación de un proceso de decisión Markoviano para definir la política óptima en el caso del manejo de un huerto de naranjos.
El problema se ha planteado para solucionarse con dos metodologías: iteración de política y programación lineal, que se han formulado en primera instancia sin descuento y luego aplicando varios factores de descuento en un rango de 0.5 a 0.9. La metodología de iteración de política ha resultado más eficiente que la programación lineal, pues requiere de un número menor de cálculos para alcanzar la solución óptima.
En todos los casos el resultado ha sido que la política óptima es dar una buena atención al huerto de naranjos, a fin de maximizar los beneficios económicos, los cuales alcanzarían un monto de 431 pesos mexicanos por naranjo por año.

1. INTRODUCTION

Markovian decision processes can be used to find optimal operating policies of a given system, which once raised, can be solved by some of the optimization techniques of operations research like dynamic or linear programming.
Their birth dates back to 1960, when Howard combined Markov chains with dynamic programming to find optimal policies for the system, he called this technique Markovian decision process (MDP). Likewise Howard (1960) helped finding solutions to problems of endless steps, leading to the policy iteration method, which is efficient finding a solution with few calculations.
Later on, other researchers found that MDP could be solved with linear programming, however it is a less efficient methodology than the policy iteration technique.
These processes have been applied in a variety of problems, such as finding optimal policies for replacement and maintenance of equipment, inventory management, animal breeding, cash flow management and administration of water tanks, among other applications (Taha 2010).
In an orange orchard, a farmer faces a series of problems that are often out of control. Some of them are: the possibility of a climatic setback, such as a winter frost, summer drought or the risk of hail, but also the presence of pests that attack the fruit and / or tree, and, ultimately, the market conditions, which sometimes make the price of the fruit low.

In those circumstances, it's up to the farmer to keep providing the proper attention to the orchard, since this will make the investment per tree to rise. However, with better care the production per tree increases, and if this is combined with a good market price at the time of sale, it can yield very beneficial economic results.
It is precisely the aim of this paper to define an optimal policy for managing an orange farm, using a technique such as Markovian decision process.

1.1 Literature Review

There are many MDP applications in many diverse fields. Here are some of the most known in academia.
White (1993) lists a number of applications of Markovian processes in many areas, the most common are equipment maintenance, water resources, agriculture, animal breeding, purchasing management, inventories, production programming, queues studies, sales promotion, finance and investment, epidemic control, lending, sports, space location, design of experiments and others.
The author classifies them according to the following criteria:

- The realism of application. If the case of study was something used to make any real decision or implementing any action as a result thereof, which means for White realism.
- The horizon. The Markovian model can be finite or infinite.
- Performance measures. Usually represented by an objective function to be optimized and can be a benefit, an income or an expected cost.
- The underlying assumptions of the model. One of the biggest barriers is that model transition states be truly representative of a Markovian model.

As mentioned before, the area where MDP is more frequently applied is in replacement and maintenance of machinery and equipment. Hernandez (2012) describes the application of MDP to replacement of equipment under a finite horizon, which is solved with dynamic programming to find the optimal replacement policy. Chen and Trivedi (2005) have applied MDP to define optimal policies for preventive maintenance of equipment. Goulionis (2010) introduced the use of MDP to obtain an optimal policy for predictive maintenance and long-term replacement of a system. Lisnianski and colleagues (2008) have applied it for evaluating the cost of maintenance contracts for older systems.
Other authors have used MDP for animal's replacement, as Kristensen and Jorgensen (1995), who have applied policy iteration and linear programming in the replacement of milch-cows with insemination and veterinary treatment. Nielsen and colleagues (2010) have used MDP to find optimal policies for milch-cows replacement based on performance measures of milk produced daily. Pla et al (2004) have used a semi Markov model to design facilities for breeding piglets, which have produced a better result than that obtained with the classical methodology.
Some authors have used MDP on medical issues, such as the assistance of patients with dementia when they wash their hands (Hoey et al 2007), or Shechter and colleagues (2008), who applied it as a means for programming medical therapies optimally.
Other researchers have used MDP in the field of construction, as Farran and Zayed (2009) who present their application to determine the life cycle cost of rehabilitating a damaged tile in the Montreal subway, or Yang et al (2009), who presented a study in which they discuss the wide use of MDP in predicting the performance of bridge management systems and suggests some options to make them more reliable.
In the administrative field, it is common to find multiple applications of MDP, as the work of Aviv and Pazgal (2005), who have used a partially observed MDP to price fashion items dynamically. In planning inventories, Yin and colleagues (2002) have applied it to the case of a major producer of paper, seeking to obtain an optimal policy for handling it. Pardo and De la Fuente (2010) have used it in the case of diffuse states to get the best policy decisions regarding advertising in queues and Higginson and Bookbinder (1995) have applied it in consolidation of shipments in a logistics strategy that combines two or more orders or shipments, so that you can send a larger amount in the same vehicle. By MDP they determine when to release the consolidated loads.
Rong and Pedram (2006) have applied these models to maximize the usability of a power portable electronic system under latent periods and speed loss constraints.
Song and Liu (2000) have used MDP in the electricity supply sector to find optimal policies for pricing of the market, which have been solved by policy iteration.

2. PROBLEM STATEMENT

The first thing to formulate a problem as a Markovian decision process is to define the number of states the system can be and the number of possible decisions that can be taken.

The next step is to define the transition matrix for each decision, which is denoted as P^i, where i is the decision, which is defined by the following matrix expression:

$$P^i = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1K} \\ p_{21} & p_{22} & \dots & p_{2K} \\ \dots & \dots & \dots & \dots \\ p_{K1} & p_{K2} & \dots & p_{KK} \end{bmatrix} \tag{1}$$

Where K is the number of states of the system and each p_{ij} is the probability of moving from state *i* to *j*.

Also it has to be defined for each transition matrix, the benefits matrix, which is expressed as follows:

$$R^i = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1K} \\ r_{21} & r_{22} & \dots & r_{2K} \\ \dots & \dots & \dots & \dots \\ r_{K1} & r_{K2} & \dots & r_{KK} \end{bmatrix} \tag{2}$$

Where r_{ij} is the corresponding benefit of moving from state *i* to *j* and v_i^k is the expected benefit from state *i* if decision *k* has been taken, which is calculated by the following equation:

$$v_i^k = \sum_{j=1}^{K} p_{ij}^k r_{ij}^k \tag{3}$$

An optimal policy is one that optimizes the performance of a system, denoted by *S* and can be interpreted as a guideline mandating a decision $d_i(S)$ if the system is in state *i*, so that *S* can take different values, such as:

$$[d_0(S), d_1(S), d_2(S), \dots\dots, d_K(S)] \tag{4}$$

The use of decision processes is formulated with two of the best known methodologies, such as policy iteration and linear programming.

2.1 Policy iteration

This methodology is solved by dynamic programming where the expected revenue is calculated recursively, keeping the optimal values of the previous stage and getting the value of the current stage by the following equation:

$$f_n(i) = v_i + \sum_{j=1}^{K} p_{ij} f_{n+1}(j), \quad \mathrm{j} = 1,2,....,\mathrm{K} \tag{5}$$

Where $f_n(i)$ is the expected optimum income in stage *n*, when *i* is the system state.
The best decision for each state is reached, which optimizes the benefit and the process ends when the same policy is obtained in two successive iterations, which usually happens with a small number of iterations.
In the case of using this methodology without a discount factor, a process of two steps is used (Taha 2010):

1. Determination of value. An S policy is chosen arbitrarily with the matrices P^s and R^s, assuming an f value to solve the following equations system:

$$E^s + f^s(i) - \sum_{j=1}^{K} p_{ij} f^s(j) = v_i^s, \quad \mathrm{i} = 1,2,\ldots,\mathrm{K} \tag{6}$$

Which has as unknown variables E^s and *(m-1)* values of f, being E^s the expected income with policy S.

2. Improving the policy. For each state i, the best policy is determined, corresponding to the optimal benefit.

$$Max \quad v_{\mathrm{i}}^{\mathrm{k}} + \sum_{j=1}^{K} p_{ij}^{k} f^s(j), \quad \mathrm{i} = 1,2,\ldots,\mathrm{K} \tag{7}$$

This procedure stops when the policy of choosing the best decisions for each state does not change between the current and previous iteration, thus defining the optimal policy.
For the case of including a discount factor α, determination of value is made by the following equation:

$$f^s(i) - \alpha \sum_{j=1}^{K} p_{ij}^{s} f^s(j) = v_i^s, \quad \mathrm{i} = 1,2,\ldots,\mathrm{K} \tag{8}$$

Which produces K unknown variables with the same number of equations.

To improve the policy, the following expression is used:

$$Max \quad v_{\mathrm{i}}^{\mathrm{k}} + \alpha \sum_{j=1}^{K} p_{ij}^{k} f^s(j), \quad \mathrm{i} = 1,2,\ldots,\mathrm{K} \tag{9}$$

2.2 Linear programming

To express the system as a linear programming problem, it is convenient to define a new variable w_{ik} which is the joint probability that the system is in state i and the k decision is taken.
With this approach without discount factor, the formulation is as follows (Taha 2010):
Maximize the objective function, denoted as Z with the following equation:

$$Z = \sum_{i=1}^{M} \sum_{k=1}^{K} v_i^k w_{ik} \tag{10}$$

The term p_{ij}^k is the element in row i and column j of transition matrix P^k if decision k is made and r_{ij}^k is the element of benefits matrix R^k located in row i and column j if k decision is taken.
Meanwhile w_{ik} is the joint probability that being in state i, k decision is made.
The constraints are:

$$\sum_{i=1}^{M} \sum_{k=1}^{K} w_{ik} = 1 \tag{11}$$

$$\sum_{k=1}^{K} w_{jk} - \sum_{i=1}^{M} \sum_{k=1}^{K} w_{ik} p_{ij}^k = 0 \qquad \text{para j} = 1, 2, \ldots, \mathrm{M} \tag{12}$$

The resulting model is linear in w_{ik}. Its optimal solution guarantees that q_i^k is *1* for one k of each i. q_i^k is obtained with the following expression:

$$q_i^k = \frac{w_{ik}}{\sum_{k=1}^{K} w_{ik}} \tag{13}$$

There are M independent equations and M basic variables and w_{ik} must be positive and greater than zero for at least one k of each i.
In case of including a discount, the linear programming approach is as follows:

Maximize

$$Z = \sum_{i=1}^{M}\sum_{k=1}^{K} v_i^k w_{ik} \tag{14}$$

Subject to

$$\sum_{k=1}^{K} w_{jk} - \alpha \sum_{i=1}^{M}\sum_{k=1}^{K} w_{ik}\, p_{ij}^k = \mathrm{b}_{\mathrm{j}} \qquad \text{para j} = 1, 2, \ldots, \mathrm{M} \tag{15}$$

Being b_j a constant arbitrarily defined with a nonzero value.

3. RESULTS

3.1 Application to an orange farm

This study applies a decision process to the case of an orange farm, which once formulated in terms of its economic benefits, which depend on the costs and revenues and its transition matrix, is solved by iteration policy and the simplex method of linear programming, to make a comparison of both techniques. In both cases it is calculated without discount factor and doing the same with a value of discount factor in the range between *0.5* and *0.9*.
The MDP is used in this case to define the optimal policy of orchard management, for which it is necessary to identify the states of the system, in this case a particular orange tree and the possible decision options.
In the case of the states, in an orchard are identified *3*, which are:

Table 1. States in which an orange tree can be found**.**

State	Orange tree condition
1	Good
2	Medium
3	Bad

Source: Authors.

The management decision options for each tree are:

Table 2. Decision options in orchard management**.**

Decision	Description
1	Minimal care
2	Medium care
3	Good care

Source: Authors.

The care of the orchard includes activities such as land preparation, irrigation, fertilization, pruning, weed control, pest control and others, so that depending of the extent these actions are carried out, attention is good, normal or bad.
Based upon the experience of *28* years of a farmer in the middle region of the state of San Luis Potosi, Mexico, information regarding benefits and costs of care of an orange tree has been obtained. With proper care, an adult

orange tree can produce up to *350* kilograms of fruit, being the normal production an amount ranging between *150* and *200* kg.
In a previous study, we applied strategic SWOT analysis (Strengths, Weaknesses, Opportunities and Threats) to the case of an orange orchard, concluding that one of the best strategies for success is the good care of the garden, consisting of a proper fertilization program (Izar-Landeta et al 2012).
To obtain a given production, weather conditions influence largely on it, since adverse situations may occur in winter as frost, or hail and drought in the summer season.
In order to put the results in Mexican pesos, the proposed values of economic benefits are based on historical average prices of fruit, which often vary depending on conditions of supply and market demand.
With this perspective, the following three transition matrixes, P^k are proposed[1]:

$$P^1 = \begin{pmatrix} 0 & 0.65 & 0.35 \\ 0 & 0.20 & 0.80 \\ 0 & 0 & 1 \end{pmatrix} \quad P^2 = \begin{pmatrix} 0.35 & 0.50 & 0.15 \\ 0.30 & 0.60 & 0.10 \\ 0.10 & 0.40 & 0.50 \end{pmatrix} \quad P^3 = \begin{pmatrix} 0.60 & 0.35 & 0.05 \\ 0.50 & 0.45 & 0.05 \\ 0.20 & 0.65 & 0.15 \end{pmatrix}$$

This means that if you take the second decision, of giving the normal care to an orange tree, and if it is found in fair condition, the probability of passing to good condition is *0.30*, to stay in the same condition is *0.60*, and passing to a bad condition is *0.10*, just the values of the second row of the matrix P^2.
As it can be seen, the sum of transition probabilities of each row is one unit and the matrix correspondent to the first decision has an absorbing state, that of an orange tree in poor condition, which once reaching this state, must only be replaced by a new one.
We must also define arrays of economic benefits, which are produced from the revenues minus the costs of management of the orchard, expressed in Mexican pesos per year per tree. These matrixes are arranged as follows[2]:

$$R^1 = \begin{pmatrix} 320 & 160 & 90 \\ 50 & 150 & 0 \\ 0 & 50 & -60 \end{pmatrix} \quad \mathrm{R}^2 = \begin{pmatrix} 460 & 280 & -50 \\ 500 & 240 & 0 \\ 300 & 140 & -100 \end{pmatrix} \quad \mathrm{R}^3 = \begin{pmatrix} 600 & 400 & -80 \\ 500 & 300 & -30 \\ 240 & 200 & -160 \end{pmatrix}$$

What you do now is calculate v_i^k for each state *i* and decision *k.* This is illustrated for the case of the first stage (*i = 1)* and the first decision (*k = 1):*

$$v_1^1 = (0)(320) + (0.65)(160) + (0.35)(90) = 135.5$$

If similar calculations are made for each decision and each state, the following results are obtained:

Table 3. Values of v_i^k for each state and decision**.**

State, i	v_i^1	v_i^2	v_i^3
1	135.5	293.5	496.0
2	30.0	294.0	383.5
3	-60.0	36.0	154.0

Source: Authors.

By raising the Markov process under the policy iteration method, applying equation (6) to the case without discount factor and beginning with the policy of giving minimal attention to an orange tree, regardless of its condition, S (1, 1, 1), the following system of equations is obtained:

[1] The values in the matrixes were collected during a period of 28 years in a local orange orchard, taking into account yearly average values from 1984.
[2] The values in the matrixes were collected during a period of 28 years in a local orange orchard, taking into account yearly average values from 1984 updated for inflation.

$$E + f(1) - 0.65f(2) - 0.35f(3) = 135.5$$
$$E + f(2) - 0.2f(2) - 0.8f(3) = 30$$
$$E + f(3) - f(3) = -60$$

If f(3) = 0 is arbitrarily chosen, the following values are obtained:

$$E = -60$$
$$f(1) = 268.625$$
$$f(2) = 112.5$$

By applying the policy improving equation for the three states, the following results are obtained:

Table 4. Values of benefits with policy improvement**.**

State	v_1	v_2	v_3	Optimum	k
1	135.5+0.65x112.5 =208.625	293.5+0.35x268.625 +0.5x112.5=443.77	496+0.6x268.625 +0.35x112.5=696.55	696.55	3
2	30+0.2x112.5 =52.5	294+0.3x268.625 +0.6x112.5=442.09	383.5+0.5x268.625 +0.45x112.5=568.44	568.44	3
3	-60+0 =-60	36+0.1x268.625 +0.4x112.5=107.86	154+0.2x268.625 +0.65x112.5=280.85	280.85	3

Source: Authors.

The best policy was the third decision, regardless of the condition of the tree, *S (3,3,3)*, that being different than the previous one, it implies at least another iteration.

Repeating the procedure for determining whether this policy values are repeated and according to the equation (6) for the benefit of the third decision, it is obtained:

$$E + f(1) - 0.6f(1) - 0.35f(2) - 0.05f(3) = 496$$
$$E + f(2) - 0.5f(1) - 0.45f(2) - 0.05f(3) = 383.5$$
$$E + f(3) - 0.2f(1) - 0.65f(2) - 0.15f(3) = 154$$

If again f (3) is zero, the equations system yields the following results:

$$E = 431.167$$
$$f(1) = 421.667$$
$$f(2) = 296.667$$

By applying the policy improvement of equation (7), the results of Table 5 are obtained, which being the same that the previous iteration, it means that the optimal solution has been reached.

Table 5. Values of benefits in this iteration.

State	v_1	v_2	v_3	Optimum	k
1	135.5+0.65x296.667 =328.33	293.5+0.35x421.667 +0.5x296.667=589.42	496+0.6x421.667 +0.35x296.667=852.83	852.83	3
2	30+0.2x296.667 =89.33	294+0.3x421.667 +0.6x296.667=598.5	383.5+0.5x421.667 +0.45x296.667=727.83	727.83	3
3	-60+0 =-60	36+0.1x421.667 +0.4x296.667=196.83	154+0.2x421.667 +0.65x296.667=431.17	431.17	3

Source: Authors.

If you use the method with a discount factor *α* of *0.9*, starting again with the policy of giving minimal care to the orange tree in whatever condition it is, *S (1,1,1*), and implementing determination of benefits of equation (8), the following system of equations is obtained:

$$f(1) - 0.9\left[0.65f(2) + 0.35f(3)\right] = 135.5$$
$$f(2) - 0.9\left[0.2f(2) + 0.8f(3)\right] = 30$$
$$f(3) - 0.9\left[f(3)\right] = -60$$

It yields the following results:

$$f(1) = -340.29$$

$$f(2) = -490.24$$

$$f(3) = -600$$

Then the benefits are calculated using equation (9) and the results are:

Table 6. Values of benefits.

State	v_1	v_2	v_3	Optimum	k
1	135.5+0.9[0.65 (-490.24)+0.35 (-600)]=-340.29	293.5+0.9[0.35 (-340.29)+0.5 (-490.24)+0.15 (-600)]=-115.3	496+0.9[0.6(-340.29) +0.35(-490.24)+0.05 (-600)]=130.82	130.82	3
2	30+0.9[0.2(-490.24)+ 0.8(-600)]=-490.24	294+0.9[0.3(-340.29) +0.6(-490.24)+0.1 (-600)]=-116.61	383.5+0.9[0.5 (-340.29)+0.45 (-490.24)+0.05 (-600)]=4.82	4.82	3
3	-60+0.9(-600) =-600	36+0.9[0.1(-340.29) +0.4(-490.24)+0.5 (-600)]=-441.11	154+0.9[0.2(-340.29) +0.65(-490.24)+0.15 (-600)]=-275.05	-275.05	3

Source: Authors.

The best policy is to give the best care in any given condition of the tree, and being this policy different than the former one, therefore another iteration must take place with the policy *S(3,3,3)*. If equation (8) is applied, it yields the following equations system:

$$f(1) - 0.9[0.6f(1) + 0.35f(2) + 0.05f(3)] = 496$$

$$f(2) - 0.9[0.5f(1) + 0.45f(2) + 0.05f(3)] = 383.5$$

$$f(3) - 0.9[0.2f(1) + 0.65f(2) + 0.15f(3)] = 154$$

Solving for the system, the following results are obtained:

$$f(1) = 4384.95$$

$$f(2) = 4261.32$$

$$f(3) = 3972.44$$

With these values, the benefits of each decision for each state are:

Table 7. Values of benefits with the new policy.

State	v_1	v_2	v_3	Optimum	k
1	3879.69	4128.63	4384.95	4384.95	3
2	3657.20	4136.57	4261.32	4261.32	3
3	3515.20	3752.32	3972.44	3972.44	3

Source: Authors.

Since it has been the same policy, it means that it is the optimum.

If this procedure is repeated for discount factor values from *0.5* to *0.9*, the results are the same: best policy is to give good care to the orchard, regardless of the condition of the tree.

If the problem is posed as linear programming without discount, according to equation (10) the equations system would be:

Maximize

$$Z = 135.5w_{11} + 293.5w_{12} + 496w_{13} + 30w_{21} + 294w_{22} + 383.5w_{23} - 60w_{31} + 36w_{32} + 154w_{33}$$

Subject to

$$w_{11} + w_{12} + w_{13} + w_{21} + w_{22} + w_{23} + w_{31} + w_{32} + w_{33} = 1$$

$$w_{11} + w_{12} + w_{13} - (0.35w_{12} + 0.6w_{13} + 0.3w_{22} + 0.5w_{23} + 0.1w_{32} + 0.2w_{33}) = 0$$

$$w_{21}+w_{22}+w_{23}-(0.65w_{11}+0.5w_{12}+0.35w_{13}+0.2w_{21}+0.6w_{22}+0.45w_{23}+0.4w_{32}+0.65w_{33})=0$$
$$w_{31}+w_{32}+w_{33}-(0.35w_{11}+0.15w_{12}+0.05w_{13}+0.8w_{21}+0.1w_{22}+0.05w_{23}+w_{31}+0.5w_{32}+0.15w_{33})=0$$
Solving by the simplex method it yields the solution:

$$w_{11}=0,\ w_{12}=0,\ w_{13}=0,\ w_{21}=0,\ w_{22}=0,\ w_{23}=0,$$
$$w_{13}=0.5370,\ w_{23}=0.4074,\ w_{33}=0.0556,$$
$$Z=431.17$$

It can be noted that, regardless of the condition of the tree, the best decision is the third, namely to give good care to the tree, which coincides with the solution obtained by policy iteration, yielding thus profits for *431.17* pesos annually per orange tree.
If this problem is solved using a discounted value α of *0.9* and using values b_j of *1*, the approach is:

Maximize
$$Z=135.5w_{11}+293.5w_{12}+496w_{13}+30w_{21}+294w_{22}+383.5w_{23}-60w_{31}+36w_{32}+154w_{33}$$
Subject to
$$w_{11}+w_{12}+w_{13}-0.9(0.35w_{12}+0.6w_{13}+0.3w_{22}+0.5w_{23}+0.1w_{32}+0.2w_{33})=1$$
$$w_{21}+w_{22}+w_{23}-0.9(0.65w_{11}+0.5w_{12}+0.35w_{13}+0.2w_{21}+0.6w_{22}+0.45w_{23}+0.4w_{32}+0.65w_{33})=1$$
$$w_{31}+w_{32}+w_{33}-0.9(0.35w_{11}+0.15w_{12}+0.05w_{13}+0.8w_{21}+0.1w_{22}+0.05w_{23}+w_{31}+0.5w_{32}+0.15w_{33})=1$$
Whose solution is:

$$w_{11}=0,\ w_{12}=0,\ w_{13}=0,\ w_{21}=0,\ w_{22}=0,\ w_{23}=0,$$
$$w_{13}=15.17,\ w_{23}=12.25,\ w_{33}=2.58,$$
$$Z=12{,}618.7$$

In this case the *Z* value represents nothing, since it depends on the values arbitrarily selected for $b_{j;}$ what is important is that the values of w_{ik} are different than zero, which being exactly the same as in the case without discount, it indicates that the best policy is to give good care to the trees regardless of its condition.
If the same calculations are made, using other values for the discount factor, (in this case the values used were from *0.5* to *0.9*), the solution is the same: take the third decision in any state of the system.

4. CONCLUSIONS

It is interesting that the results obtained are the same with or without considering discount factor and this conclusion is valid for a wide range of values of α factor. This has been identical with the two most common methodologies for solving Markovian decision processes.
There is a considerable utilization of much less calculations when policy iteration methodology is used compared to linear programming. With this methodology, the problem was solved using Lindo software and the solution was achieved in five iterations.
In the application to the orange farm, the best policy is to give good care regardless of the condition of the trees, which is basically due to the fact that with good care the amount of fruit produced per tree is significantly higher, leading to markedly improved profits, since the wholesale price of the fruit outweighs largely the cost invested per tree. Moreover, this is even further favored by the use of new agricultural technologies that have emerged in the last two decades, as the micro sprinkler irrigation, organic fertilization and biological pest control.
This problem has been formulated as a Markovian decision process with a simple solution, confirming its application in a wide variety of fields.

RECEIVED OCTOBER, 2013
REVISED DECEMBER, 2013

REFERENCES

[1] AVIV, Y., PAZGAL, A. (2005): A Partially Observed Markov Decision Process for Dynamic Pricing. **Management Science**, 51(9), 1400-1416.
[2] CHEN, D., TRIVEDI, K. S. (2005): Optimization for condition-based maintenance with semi-Markov decision process. **Reliability Engineering and System Safety**, 90, 25-29.
[3] FARRAN, M., ZAYED, T. (2009): Comparative Analysis of Life-Cycle Costing for Rehabilitating Infrastructure Systems. **Journal of Performance of Constructed Facilities**, 23, 320-326.

[4] GOULIONIS, J. E. (2010): A generalized model of optimal maintenance strategy. **Pakistan Journal of Statistics**, 26, 357-364.
[5] HERNÁNDEZ, E. S. (2012): A reward functional to solve the replacement problem. **Working Paper, XVI Congreso Internacional de la Academia de Ciencias Administrativas AC (ACACIA), ITESM, Atizapán de Zaragoza, México**.
[6] HIGGINSON, J. K., BOOKBINDER, J. H. (1995): Markovian Decision Processes in Shipment Consolidation. **Transportation Science**, 29, 242-255.
[7] HOEY, J., VON BERTOLDI, A., POUPART, P., MIHAILIDIS, A. (2007): Assisting persons with Dementia during Handwashing Using a Partially Observable Markov Decision Process. **Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007), Bielefeld University, Germany**.
[8] HOWARD, R. A. (1960): **Dynamic Programming and Markov processes**. The Technology Press of the MIT and John Wiley & Sons, Inc. New York.
[9] IZAR-LANDETA, J. M., YNZUNZA-CORTÉS, C. B., CASTILLO-RAMÍREZ, A. (2012): Aplicación del análisis FODA y el proceso analítico de jerarquización en la citricultura. **Revista Vértice Universitario**, 44, 43-51.
[10] KRISTENSEN, A. R., JORGENSEN, E. (1995): Applicational perspectives of recent developments in dynamic programming methods for herd management support. **Dina Nota No. 33**, URL: ftp://ftp.dina.kvl.dk/pub/Dina-reports/notat33.ps.
[11] LISNIANSKI, A., FRENKEL, I., KHVATSKIN, L., DING, Y. (2008): Maintenance Contract Assessment for Aging Systems. **Quality and Reliability Engineering International**, 24, 519-531.
[12] NIELSEN, L. R., JORGENSEN, E., KRISTENSEN, A. R., OSTERGAARD, S. (2010): Optimal replacement policies for dairy cows based on daily yield measurements. **Journal of Dairy Science**, 93, 75-92.
[13] PARDO, M. J., DE LA FUENTE, D. (2010): Fuzzy Markovian decision processes: Application to queuing systems. **Computers and Mathematics with Applications**, 60, 2526-2535.
[14] PLÁ, L. M., BABOT, D., POMAR, J. (2004): A mathematical model for designing and sizing sow farms. **International Transactions in Operational Research**, 11, 485-494.
[15] RONG, P., PEDRAM, M. (2006): Battery-Aware Power Management Based on Markovian Decision Processes. **IEEE Transactions on Computer-Aided Design of Integrated Circuits & Systems**, 25, 1337-1349.
[16] SHECHTER, S. M., BAILEY, M. D., SCHAEFER, A. J. (2008): A modeling framework for replacing medical therapies. **IIE Transactions**, 40, 861-869.
[17] SONG, H., LIU, CH. (2000): Optimal Electricity Supply Bidding by Markov Decision Process. **IEEE Transactions on Power Systems**, 15, 618-624.
[18] TAHA, H. A. (2010): **Operations Research: An Introduction, 9th edition.** Pearson Prentice Hall New York.
[19] WHITE, D. J. (1993): A Survey of Applications of Markov Decision Processes. **The Journal of the Operational Research Society**, 44, 1073-1096.
[20] YANG, Y. N., PAM, H. J., KUMARASWAMY, M. M. (2009): Framework Development of Performance Prediction Models for Concrete Bridges. **Journal of Transportation Engineering**, 135, 545-554.
[21] YIN, K.K., LIU, H., JOHNSON, N. E. (2002): Markovian inventory policy with application to the paper industry. **Computers & Chemical Engineering**, 26, 1399-1413.