Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio

doi: 10.3850/978-981-18-8071-1_P630-cd

Mathematical Formulation of Markov Decision Process to Address Maintenance Policy in Photovoltaic Farms

Eduardo Tadeu Bacalhau
Pontal do Parana Campus, Center of Marine Studies, Federal University of Parana, Brazil
E-mail: bacalhau@ufpr.br

Flávia Barbosa
INESC TEC, Faculdade de Engenharia, Universidade do Porto, Portugal. E-mail: flavia.barbosa@inesctec.pt

Luciana Casacio
Pontal do Parana Campus, Center of Marine Studies, Federal University of Parana, Brazil
E-mail: lucianacasacio@ufpr.br

Felipe Yamada
INESC TEC, Faculdade de Engenharia, Universidade do Porto, Portugal. E-mail: felipe.yamada@inesctec.pt

Luís Guimarães
INESC TEC, Faculdade de Engenharia, Universidade do Porto, Portugal. E-mail: luis.guimaraes@inesctec.pt

In order to face the challenges incurred by climate change, the industry has been striving to improve the overall performance of photovoltaic (PV) systems. Unsolved challenges remain concerning reliability, numerous unforeseen outages, and high operation and maintenance (O&M) costs. In this context, this work increases the operational performance of PV plants by improving current methodologies for O&M in PV systems. We develop a maintenance approach based in a Markov decision process model to analyze the data from PV power plants, prioritize actions, advise asset replacement, and schedule preventive maintenance tasks based on past experiences and the PV system condition. The results allow economic improvement through downtime reduction and early detection of system under performance.

Keywords: Maintenance policy, Markov decision process, photovoltaic farms, Reliability, Operation and maintenance cost, Optimisation.

1. Introduction

The Paris Agreement, an international treaty on climate change, has defined the necessary targets to limit global warming to 1.5°C with a massive contribution by renewable energy. Even considering the drastic decrease in solar panel cost in the last ten years and the exponential increase in photovoltaic (PV) installation, we remain below expectations regarding goals for solar energy implementations. The United Nations Climate Change Conference's goal of global solar capacity by 2025 remains 4,500 GW above our current forecast.

In order to face these challenges, the industry has been striving to improve the overall performance of PV systems. Solar PV energy has grown tremendously due to technological and business maturities. Solar has evolved rapidly from being a subsidies-based model to being competitive in auctions at wholesale electricity markets with traditional fossil fueled energies. However, unsolved challenges remain concerning reliability, numerous unforeseen outages, and high operation and maintenance (O&M) costs. In the scenario, the most critical aspect is to define the appropriate time to perform maintenance actions.

The primary focus of this work is to provide recommendations for effective maintenance actions. In cases where a decrease in performance is observed, preventive or corrective maintenance may be necessary. This decision involves assessing trade-offs between the most cost-effective approach for the PV system and the costs associated with maintenance, such as labor, equipment, and downtime. These trade-offs are evaluated based on the expected future rewards. The main goal of this project is to enhance the operational performance of PV plants by improving current methodologies for O&M in PV systems. To achieve this, we have developed a Markov Decision Process that estimates the probability of each system component reaching a given degradation state. Maintenance actions determine the transition between states and are planned based on future rewards.

2. Markov decision processes

Markov Decision Process has been often used to model maintenance management. Maintenance involves technical and administrative actions aimed at preserving an item or system, or restoring it to a functional state. A Markov model is a dynamic framework that allows for probabilistic modeling of the evolution of a system over time. It is a dynamic stochastic process that consists of a sequence of states with a discrete state space and transition matrix. The transitions between states depend only on the current state, and not on any past history, making it a memoryless process. Such a model is called a Markov chain. A Markov decision chain is an extension of a Markov chain that allows for optimal decision-making based on a sequence of actions. It is a stochastic process governed by the state space, the action set in each state, the transition probabilities, and the immediate rewards obtained when a given action is taken. A Markov Decision Process describes such a dynamic model, which can be used to identify optimal decisions over time. MDPs are controlled by policies and decision rules, which specify how to choose actions in each state of the system. A policy is a sequence of decision rules that guides the selection of actions in a way that maximizes the expected reward.

Control policies can be highly dependent on the age of the machine, leading to a large problem dimension when considering additional state variables such as machine age and preventive maintenance states within the Markov chain model (Kenné and Gharbi, 2004). To address this issue, Park and Yoon (2012) proposed a modified Markov chain model that accounts for the deterioration process, inspection, and maintenance of equipment. Tomasevicz and Asgarpoor (2009) employed a semi-Markov decision process to evaluate the need for maintenance in power equipment at different stages of deterioration, and determine the appropriate maintenance type. Zhang et al. (2017) proposed a Markov chain model to estimate the time interval for the degradation of offshore structures and investigated the potential of this stochastic model in predicting maintenance schedules. Chan and Asgarpoor (2006) proposed an approach to determine the optimal maintenance strategy for a component. By employing Markov processes, the authors computed the state probabilities and the optimal mean time for preventive maintenance while maximizing the component's availability.

3. Mathematical Formulation

This section introduces a Markov Decision Process to model the problem, where the states are updated at each stage based on transition matrices and depend only on the current state. The model considers the discrete time $t = 1, \ldots, T$ and a planning horizon of T.

Variables

The decision variables in this model are binary:

- $x^t_{k_{a_m}}$ takes the value 1 if maintenance action a_m $(m = 1, .., N - 1)$ is performed for equipment k at time t, and 0 otherwise.

Equations

Maintenance actions determine the transition between states. The transition matrix of the Markov process, denoted as $\mathbf{P}^t_{ka_n}$, provides the probabilities of transitioning from one state to another for equipment k at time t, based on each maintenance action a_n. Here, $n = 0, \ldots, N$ represents the types of maintenance actions available, where 0 denotes no maintenance action, and N corresponds to corrective maintenance action involving equipment replacement.

In the first step of the process, the initial probability state of each component is calculated based on the equipment's degradation level. To accomplish this, the current deterioration of each piece of equipment must be determined. A binary indicator, w_k^{tp}, is used where 1 denotes maintenance action performed for equipment k at past time tp, and 0 indicates otherwise. The value of tp can range from 0 to TP, which represents the total time past for each equipment. A vector is created to represent the probabilities of the deterioration states according to Eq.(1), where $d_{k_i}^t$ represents the probability of the equipment being in the deterioration state i at time t.

$$d_k^0 = \lambda \cdot \prod_{tp=1}^{TP} P_{ka_0}^{\left(1 - w_k^{tp}\right)} \cdot P_{ka_1}^{\left(w_k^{tp}\right)}, \quad (1)$$

where λ is the probability vector for new equipment.

Next, using the Markov process, the vector $\mathbf{d}_k^t$ for each equipment k is updated at each time t. This is done by multiplying it with the transition matrices $\mathbf{P}_{ka_n}^t$ according to Eq.(2) (on the next page). Therefore, at each time, this vector indicates the probabilities of the equipment being in each state of deterioration.

Objective Function

Maintenance actions are planned based on the expected future rewards, which in this case, depend on the efficiency of the PV farm. To measure the farm's efficiency, the irradiance is used as a metric. Therefore, the objective function incorporates a reward for the irradiance, which is a function of both the equipment efficiency and the maintenance action costs. The objective function is represented in Eq.(3).

$$\max Z = \sum_{t \in T} R(I^t | a_n^t) \sum_{k \in K} E(\mathbf{d}_k^t) - C(a_n^t), \quad (3)$$

where $E(\mathbf{d}_k^t)$ is the equipment efficiency function regarding the probability vector of equipment k at time t, $C(a_n^t)$ represents the cost of performing maintenance action a_n at time t, and $R(I^t)$ is the reward that takes into account the maintenance action schedule and the irradiance level at time t.

4. Optimal Maintenance Policy

In recent years, optimal maintenance policy solution techniques have been sought using only one maintenance state. It ignores the possibility that different types of maintenance can be done to correct specific problems. Including more than one maintenance state, a maintenance model can be more sufficiently applied to real-life situations. The optimal maintenance policy is determined using Markov Decision Processes, which describe the action to be taken at each state to yield minimal cost and ensure high equipment availability. Since the system may have numerous states with different alternatives for each state, the number of total possible policies can be very large. To validate the proposed model and the optimal policy, three experiments are performed, each considering time in months, three maintenance actions (where a_0 represents no maintenance, a_1 represents low maintenance, and a_2 represents high maintenance or equipment replacement), and three deterioration states (where D_0^k represents a working equipment, D_1^k represents a failed equipment, and D_2^k represents a broken equipment). These experiments use the following transition matrices:

$$P_{ka_0}^t = \begin{bmatrix} 0.93 \; 0.06 \; 0.01 \\ 0.00 \; 0.85 \; 0.15 \\ 0.00 \; 0.00 \; 1.00 \end{bmatrix}$$

$$P_{ka_1}^t = \begin{bmatrix} 0.98 & 0.015 & 0.005 \\ 0.75 & 0.20 & 0.05 \\ 0.45 & 0.30 & 0.25 \end{bmatrix} \quad P_{ka_2}^t = \begin{bmatrix} 1 \; 0 \; 0 \\ 1 \; 0 \; 0 \\ 1 \; 0 \; 0 \end{bmatrix}.$$

To perform the experiments, the matrices $P_{ka_n}^t, n = 1, 2, 3$, are used for all equipment k and all time periods t.

The rewards for irradiance are calculated by considering the average number of daily sunlight hours per month and the maintenance action a_n^t planned at time t. The resulting matrix is of size 3×12 and is represented as follows:

$$d_k^{t+1} = d_k^t \cdot \left\{ P_{ka_0}^{\left(x_{a_0}^t\right)} \cdot \left[\cdots P_{ka_{N-1}}^{\left(x_{ka_{N-1}}^t\right)} \cdot P_{ka_N}^{\left(1 - x_{ka_{N-1}}^t\right)} \cdots \right]^{\left(1 - x_{ka_0}^t\right)} \right\} \tag{2}$$

$$R(I^t|\mathbf{a}_n) = \begin{bmatrix} 5 & 10 & 15 & 20 & 30 & 50 & 50 & 30 & 20 & 15 & 10 & 5 \\ 30 & 40 & 50 & 40 & 30 & 10 & 10 & 30 & 40 & 50 & 40 & 30 \\ 50 & 40 & 30 & 10 & 10 & 5 & 5 & 10 & 10 & 30 & 40 & 50 \end{bmatrix}$$

The maximum and minimum reward values were determined through empirical tests, taking into account the months with the highest solar incidence. The analysis considers the following possibilities: not planning maintenance during months with high sun exposure hours as the maintenance actions may cause system interruption, and planning the high maintenance level only during months with the lowest hours of sun exposure.

As mentioned, the equipment efficiency function is based on the probability vector of equipment k at time t. However, for these experiments, a constant vector ($3\times$ 1) is used to represent the equipment efficiency values for each deterioration state, which are then multiplied by the corresponding probabilities in the actual probability vector of equipment.

$$E(\mathbf{d}_k^t) = \mathbf{d}_k^t \cdot \begin{bmatrix} 100 \\ 70 \\ 20 \end{bmatrix} \tag{4}$$

An efficiency of 100% corresponds the equipment in the deterioration state D_0; 70% of the efficiency corresponds the equipment in deterioration state D_1; and 20% of the efficiency corresponds the equipment in deterioration state D_2.

The maintenance cost is determined by a matrix that relates the reduction in equipment efficiency associated with deterioration states to the planned maintenance actions a_n^t. This matrix specifies no reduction for the action a_0, a 20% reduction for the action a_1, and a 90% reduction for the action a_2. It is defined as follows:

$$C(a_n^t) = \begin{bmatrix} 0\%(100) & 20\% & 90\% \\ 0\%(70) & 20\% & 90\% \\ 0\%(20) & 20\% & 90\% \end{bmatrix} = \begin{bmatrix} 0 & 20 & 90 \\ 0 & 14 & 63 \\ 0 & 4 & 18 \end{bmatrix}$$

Finally, the equipment analyzed in the experiments is a cable. Table 1 presents the components k along with their respective lifetime (TP), the number of times the low maintenance level action has been performed (w_k^{tp}), and the time at which this action has been taken (tp).

Table 1. Components

k	TP	w_k^{tp}	tp
1	10	1	3
2	12	1	8
3	6	1	3
4	12	0	0
5	12	1	12

Additionally, all experiments were validated using an algorithm developed in Matlab® Language that simulates an exact method, thus obtaining an optimal solution by exhaustively exploring all solution space.

4.1. *Experiment 1*

In this experiment, only one equipment, namely $k = 1$, was considered. It has been operating for 10 months, and a low maintenance level action was carried out in month 3. The updated deterioration performance of this equipment, denoted as d_k^0, is calculated using Eq.(1), where $\lambda = [1 \;\; 0 \;\; 0]$:

$$d_k^0 = \lambda \cdot [P_{a_0}]^2 \cdot [P_{a_1}] \cdot [P_{a_0}]^7 = [xx \quad xy \quad yy].$$

The vector d_k^0 indicates the probabilities of the equipment being in deteriorating states 0, 1, and 2, represented by xx%, xy%, and yy%, respectively.

This information is subsequently utilized in Eq.(2) to simulate the equipment's deterioration states.

Furthermore, in this experiment, maintenance action planning is performed for a period of one year, starting from January. The objective is to demonstrate the effectiveness of the proposed model in achieving strategic maintenance planning considering irradiation, while simultaneously maximizing the efficiency of the equipment.

Table 2. Experiment 1 – Results

k	Total Time (s)	Z	Optimal Solution (months) 1	2	3	4	5	6	7	8	9	10	11	12
1	1663	1295.4	2	2	2	2	1	1	1	1	2	1	1	1

Table 2 shows the results obtained from the algorithm for the experiment. The objective function value Z was calculated using Eq.(3), which represents the accumulated sum of the rewards and the equipment efficiency over the year, discounted by the maintenance costs. For the case with only one equipment and two levels of deterioration, the optimal maintenance policy suggests that maintenance actions should not be performed in months 5,6,7,8,10,11, and 12. In contrast, minor performance actions should be carried out in months 1,2,3,4, and 9. The optimal solution avoids performing maintenance actions during months with higher irradiance levels, which include May, June, July, and August. It is noteworthy that the equipment has not been replaced, and the total reward obtained was 1295.4.

4.2. *Experiment 2*

The second experiment aims to analyze the impact of increasing the number of equipment. For this purpose, all components presented in Table 1 are considered for a planning horizon of one year, starting in January.

Table 3 shows the suggested maintenance policy for each equipment k, comprising the objective function values and the maintenance action performed each month for each equipment. The executions were split since there was no dependency on functionality among the components.

Table 3. Experiment 2 – Results

k	Total Time (s)	Z	Optimal Solution (months) 1	2	3	4	5	6	7	8	9	10	11	12
1	1630	1295.4	2	2	2	2	1	1	1	1	2	1	1	1
2	2050	1306.2	1	2	2	2	2	1	1	1	2	1	1	1
3	2188	1311.7	1	2	2	2	2	1	1	1	2	1	1	1
4	2105	1281.5	3	1	1	2	2	1	1	1	2	1	1	1
5	2247	1330.5	1	2	2	2	1	1	1	1	2	1	1	1

-: The total time for the execution was 10221 seconds (2.84 hours)

Thus, the total time is calculated by the sum of the computational times of each execution. Furthermore, the non-dependence of functionality among the components substantially reduces the number of solutions, thereby decreasing the complexity of the problem. The number of solutions can be calculated as $5 \cdot 3^12$, i.e., the same number of solutions multiplied by the number of components of the system.

4.3. *Experiment 3*

This experiment aims to evaluate the model's ability to avoid maintenance actions during periods of high irradiation, even when the equipment is highly deteriorated, by analyzing the starting month of the planning process. The maintenance planning horizon for both simulations is six months, with one starting in April and the other starting in June.

Table 4. Experiment 3 – Results from April start

k	Total Time (s)	Z	Optimal Solution (months) 4	5	6	7	8	9
1	0.06	703.6	2	2	1	1	1	2
2	0.04	715.7	2	2	1	1	1	2
3	0.04	719.3	2	2	1	1	1	2
4	0.05	597.6	3	1	1	1	1	2
5	0.06	736.2	1	2	1	1	1	2

-: The total time for the execution was 0.25 seconds

Table 4 displays the recommended maintenance policy for each equipment k over a planning horizon starting in April, while Table 5 shows the

Table 5. Experiment 3 – Results from June start

k	Total Time (s)	Z	Optimal Solution (months) 6	7	8	9	10	11
1	0.08	624.5	3	1	1	2	1	1
2	0.04	644.2	1	1	2	2	1	1
3	0.04	657.7	1	1	2	2	1	1
4	0.04	624.5	3	1	1	2	1	1
5	0.04	709.2	1	1	1	2	1	1

-: The total time for the execution was 0.24 seconds

suggested maintenance policy for each equipment k over a planning horizon starting in June. The obtained results indicate that the suggested maintenance actions are directly influenced by the irradiance rewards. However, the model does not completely prevent highly deteriorated equipment from undergoing maintenance actions in months with higher solar incidence, leading to a lower objective function value.

In this experiment, the initialization of the process does not update the values relative to the equipment's life, regardless of the month it starts. This means that the initial probability vector remains the same, as shown in the following table:

Table 6. Initial probability vector

k	d_k^0		
1	[0.5359	0.2143	0.2498]
2	[0.6755	0.1929	0.1316]
3	[0.7240	0.1762	0.0998]
4	[0.1752	0.1162	0.7885]
5	[0.9030	0.0763	0.0207]

4.4. *Overall Analysis*

The results of this study showed that the proposed model is complex and can be classified as NP-complete, as it is an NP problem and the initial solutions calculations are polynomial. However, introducing functionality dependence among the components and increasing the number of equipment types could make it an NP-hard problem. These scenarios should be further explored and presented in future research.

To interface with this algorithm operators are required to input two crucial sets of information. Firstly, a matrix is provided where each row specifies details for component k, encompassing the current lifetime (TP), the number of times the low maintenance level action (w_k^{tp}) has been performed, and the corresponding time (tp). Table 1 illustrates a comprehensive example of this matrix. Secondly, operators are required to indicate the initial and final months of analysis. The algorithm incorporates essential data concerning transition matrices, rewards, efficiency, and cost. Using the exact method, the algorithm is capable of effectively handling data for a duration of approximately 12 months. The outputs are maintenance actions recommended by the algorithm, enabling operators to evaluate their suitability. For each equipment and month, a value between 1 and 3 is assigned, indicating actions as follows: 1 (no maintenance action), 2 (low maintenance level action), and 3 (equipment replacement).

The proposed model is capable of simulating the specific characteristics of the maintenance planning problem presented. This demonstrates the importance of involving subject matter experts, such as those knowledgeable about system interruptions during periods of high irradiance, in the development of the model. However, further improvements can be made to the model by upgrading the objective function to incorporate real-world operating scenarios, with a particular emphasis on the equipment's lifespan and efficiency.

Moreover, the study indicates that exploring alternative transition matrices is essential, as taking any preventive maintenance actions may trigger a new equipment deterioration process. In addition, it is important to note that different types of equipment require distinct maintenance actions, suggesting that the model could be further enhanced by considering equipment-specific maintenance strategies. Future research could focus on addressing these limitations to improve the model's applicability and effectiveness in practical scenarios.

5. Conclusion and Future Works

This work presents a mathematical model based on Markov Process to support maintenance decision-making in PV systems. The model was validated through numerical experiments that simulated different scenarios in PV systems. Results demonstrated that the model is capable of adequately planning preventive or corrective maintenance actions by analyzing the current state of the equipment and estimating its future state.

However, the scalability of the model was found to be limited, as the solution space increases exponentially with the size of the system and the planning horizon. To address this issue, a reinforcement learning approach is being developed to deal with the so-called "curse of dimensionality".

Based on the analysis of the obtained results, several improvements have been proposed to enhance the model's applicability to real operating scenarios. These improvements include incorporating various types of equipment into the maintenance plan for the entire system, varying the transition matrices for each equipment and period, and considering a wider range of maintenance actions. Additionally, constraints related to the system's average interruption duration and frequency will be included to improve the model's performance. Future research will focus on evaluating and implementing these improvements.

Acknowledgement

The authors thank the National Council for Scientific and Technological Development (CNPq), for their financial support to this research, under grant processes 402068/2022-0 and 200875/2022-1.

References

Chan, G. and S. Asgarpoor (2006). Optimum maintenance policy with markov processes. Electric Power Systems Research 76(6), 452–456.

Kenné, J. and A. Gharbi (2004). Stohastic optimal production control problem with corrective maintenance. Computers & Industrial Engineering 46(4), 865–875. Computers and Industrial EngineeringSpecial Issue on Selected papers form the 29th.International Conference on Computers and Industrial Engineering.

Park, G.-P. and Y. T. Yoon (2012). Application of ordinal optimization on reliability centered maintenance of distribution system. European Transactions on Electrical Power 22(3), 391–401.

Tomasevicz, C. L. and S. Asgarpoor (2009). Optimum maintenance policy using semi-markov decision processes. Electric Power Systems Research 79(9), 1286–1291.

Zhang, Y., C.-W. Kim, and K. F. Tee (2017). Maintenance management of offshore structures using markov process model with random transition probabilities. Structure and Infrastructure Engineering 13(8), 1068–1080.