

**Diana Yaziel, Ortiz-
Muñoz**

*Universidad Autónoma
de Ciudad Juárez,
al216921@alumnos.uacj.
mx*

**David, Luviano-
Cruz***

*Universidad
Autónoma de
Ciudad Juárez,
david.luviano@uacj.
mx*

**Luis A.,
Rodriguez-
Picón**

*Universidad
Autónoma de
Ciudad Juárez,
luis.picon@uacj.
mx*

**Luis A., Perez-
Dominguez**

*Universidad
Autónoma de Ciudad
Juárez,
luis.dominguez@uacj.
mx*

Resumen: En un futuro las fuentes de energía renovables tendrán un rol más importante debido a la necesidad de combatir los efectos del cambio climático derivado del uso de combustibles fósiles. Entre las opciones renovables tenemos la energía solar, como una fuente prácticamente inagotable de energía limpia. En ciudad Juárez, México se presentan más de 300 días soleados al año, por lo que en este documento mostramos los resultados de la aplicación de aprendizaje por reforzamiento para el control de orientación de un panel fotovoltaico.

Palabras clave: aprendizaje por reforzamiento, control inteligente, panel fotovoltaico.

1 INTRODUCCIÓN

Uno de los principales problemas actuales que se tienen en el mundo es el calentamiento global. Una de las fuentes de energía renovable que ha tenido un gran impulso ha sido la energía solar, el cual se define como la generación de energía eléctrica a partir de la recepción de los rayos solares por medio de paneles fotovoltaicos. Esta ha mostrado su gran

versatilidad en lugares en donde la distribución de energía eléctrica convencional no está disponible o es intermitente. En Ciudad Juárez México se presentan al menos 300 días soleados al año, lo que ofrece una gran oportunidad de usar la energía solar mediante paneles fotovoltaicos. Por lo que la orientación de los paneles interviene de manera significativa en la eficiencia fotovoltaica. La gran mayoría de los controladores en lazo cerrado que manipula la orientación están basados en esquemas de control clásico los cuales usan un modelo dinámico de la planta, sin embargo, en este artículo proponemos el uso de una estrategia de control inteligente basada en aprendizaje por reforzamiento para controlar la orientación de un panel solar de 2 grados de libertad.

* Citation: Autor 1., Autor 2, Autor 3, and Autor 4 (2022). Título. *Conference Proceedings of the International Congress on Innovation and Sustainable*, Barcelona-España, October 27-28, 2022, p.p. XX-XX.

2 MARCO TEÓRICO Y CONCEPTUAL

En la actualidad, los sistemas de paneles solares se utilizan en muchos campos de la vida, especialmente para la generación y almacenamiento de energía eléctrica, en donde el rendimiento obtenido por los paneles depende de la cantidad de radiación solar recibida por el sistema, esta radiación solar recibida varía en función con el ángulo de los rayos del sol que inciden sobre la superficie del panel [1].

La mayoría de los controles utilizados en estos sistemas se basan en estrategias clásicas como controladores PID [2], sin embargo, en nuestra propuesta abordaremos el enfoque de aprendizaje por reforzamiento (RL) basado en diferencias temporales. Bajo este enfoque se selecciona las variables de estado que conformaran el marco de aprendizaje basado en un proceso de Markov. Las acciones de control serán discretas las cuales manipularan los movimientos angulares de los servomotores que rigen la orientación de panel fotovoltaico. Un proceso finito de Markov para un agente es una tupla (S, A, f, ρ) :

$$f : S \times A \times S \rightarrow [0,1]$$

$$\rho : S \times A \times S \rightarrow \mathbb{R}$$

Donde S es el conjunto finito de estados en el entorno, A es el conjunto finito de acciones disponibles para el agente, f es la función de probabilidad de transición de estados y ρ es la función de recompensa la cual se asume acotada [3].

El objetivo final del aprendizaje por reforzamiento es encontrar una política π , para todo estado s que maximice el retorno R :

$$R^\pi = E \left\{ \sum_{t=0}^{\infty} \gamma^k r_{t+1} \mid s_0 = s, \pi \right\}$$

Por lo tanto, el objetivo de control es maximizar la captación de rayos solares a largo plazo caracterizado por el retorno, solo recibiendo la realimentación acerca de su inmediata actuación en forma de la señal de recompensa.

3 METODOLOGÍA

La propuesta se centra en usar el algoritmo Q-learning [4], para diseñar un controlador que regule la orientación del panel solar con respecto a los rayos solares. En la terminología de RL un agente es el controlador, para ello se debe de realizar los siguientes pasos:

1) Se captura los estados de cada ciclo de entrenamiento del algoritmo de aprendizaje por reforzamiento: el estado de los agentes con respecto al entorno es capturados a través de sensores.

2) Establecer las acciones disponibles para los agentes: En cada momento los agentes son requeridos de realizar una acción con un grado de coordinación, por lo tanto, es necesario seleccionar de antemano las acciones más fiables a realizar por los agentes.

3) Estimar los valores estado-acción de cada agente: La recompensa

numérica de cada acción es calculada y dada al agente después de que se realiza una acción en conjunto, este cálculo se realiza por medio del algoritmo de aprendizaje por reforzamiento Q-learning [5].

4) Repetir los pasos 2-4 hasta que los valores converjan: Esto sucede cuando los valores permanecen sin cambios o ellos se encuentran debajo de alguna cota establecida de antemano.

5) Obtención de Q-tabla final de estados-acciones: La tabla final de estados-acciones óptimas es puesta a punto para la selección de las acciones óptimas por medio de la localización de la acción que generara el máximo valor en cada estado.

4 RESULTADOS

Para la etapa de entrenamiento se toman los datos que proporciona el INEGI [6], el cual es un organismo federal descentralizado del gobierno mexicano que ofrece datos estadísticos de libre acceso, los datos graficados se muestran en la figura 1.

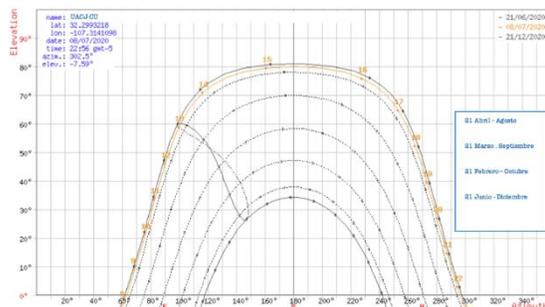


Figura 1: visualización de datos de entrenamiento

En una etapa adicional se procedió a la extracción de características a partir del conjunto de datos de entrenamiento, aquí es donde se convierte los datos brutos en información que es adecuado de procesar por los algoritmos de aprendizaje, los cuales se muestran en la tabla 1.

Hora	Elevación	Acimut
08:10	-0.833	62.63
09:00	8.87	69.07
10:00	20.96	76.11
11:00	33.41	82.97
12:00	46.06	90.42
13:00	58.67	99.96
14:00	70.74	116.43
15:00	79.54	161.2
16:00	75.83	228.18
17:00	64.63	253.52
18:00	52.17	265.35
19:00	39.5	273.52
20:00	26.92	280.54
21:00	14.63	287.42
22:00	2.8	294.76
22:20	-0.833	297.28

Tabla1. Datos de entada-salida para entrenamiento

En la figura 2 se muestra el diagrama de fase del sistema de orientación del panel, en el eje X se muestra la posición angular del servo 1 y en el Y del servo 2.

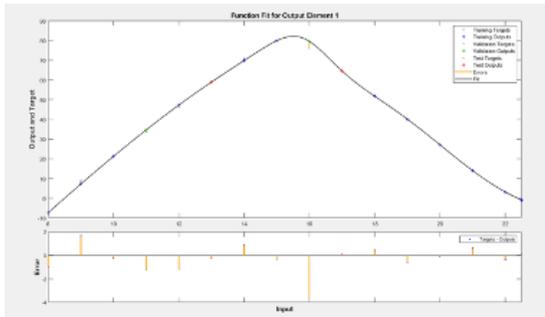


Figura 2 Diagrama de fase

La respuesta en estado transitorio del sistema electromecánico ante una entrada escalón muestra una forma sub amortiguada, lo que ofrece una respuesta rápida en tiempo de levantamiento y una sobre elongación menor al 5%. Se muestra en la figura 3

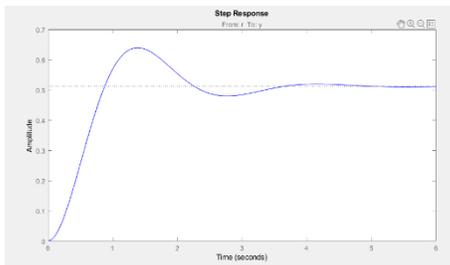


Figura 3. Respuesta transitoria al escalón unitario de la salida del sistema

5 CONCLUSIONES

La integración del aprendizaje por reforzamiento al control de orientación del panel solar permitió recibir una mayor cantidad de radiación solar en la superficie del panel, en esta estrategia no se uso un modelo del sistema, ya que el aprendizaje se dio con el algoritmo Q-learning el cual es libre de modelo.

Las graficas muestran una respuesta sub amortiguada en la salida del sistema, queda pendiente implementar una red neuronal profunda para agregar más características al modelo de aprendizaje.

6 REFERENCIAS

- 1-Hoyos, D., Gea, M., Cadena, C., & Echazú, R. (2008). Diseño y construcción de seguidores solares. Avances en energías renovables y medio ambiente, 12.
- 2-Sheikh, S. S., Javed, A., Anas, M., & Ahmed, F. (2018, September). Solar based smart irrigation system using PID controller. In IOP Conference Series: Materials Science and Engineering (Vol. 414, No. 1, p. 012040). IOP Publishing.
- 3-Luviano-Cruz, D., & Yu, W. (2017). Path planning of multi-agent systems in unknown environment with neural kernel smoothing and reinforcement learning. Neurocomputing, 233, 34-42.
- 4- Watkins, C. J., & Dayan, P. (1992). Q-learning. Machine learning, 8(3), 279-292.
- 5- Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.
- 6-INEGI, P. W. (2003). INEGI. Recuperado el, 25.