

Control de Taps de Transformadores en Sistemas de Distribución
Usando Aprendizaje Reforzado

Christopheerth Anderson Bonilla León, Nicolás Augusto Marín Pinzón
y Aldo Marcel Rizo Casadiegos

Trabajo de Grado para Optar al Título de Ingenieros Electricistas

Director

Oscar Alexis Galvis Diaz

Ingeniero Electricista

Codirector

Juan Manuel Rey López

Doctor en Ingeniería Electrónica

Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones

Facultad de Ingenierías Físico-Mecánicas

Universidad Industrial de Santander

Bucaramanga

2024

Dedicatoria

Dedico este trabajo a mi familia, cuyo amor y apoyo incondicional han sido el pilar fundamental de mi camino académico. Agradezco a mis padres por su constante aliento y paciencia.

A mi querida pareja, gracias por ser mi compañera incondicional. Tu apoyo, comprensión y aliento han sido la luz que ha iluminado este camino.

A nuestros amigos, quienes han sido testigos de este viaje y han compartido risas y desafíos, agradezco su amistad invaluable.

Christopheerth Anderson Bonilla León

A mis padres, hermanos y amigos (Juan Manuel, Carlos y Cleto), han sido parte de mi formación, me han guiado a lo largo de cada etapa de mi vida, sin ellos no habría podido completar esta gran meta.

A mi amada novia Wendy por todo el apoyo incondicional que me ha dado, su comprensión y amor sin importar las situaciones que se presenten, que este logro sea un pilar para nuestro futuro.

A mis compañeros de tesis Aldo y Christopheerth por el gran esfuerzo que hicieron para llevar a cabo esta investigación y por todo el aprendizaje que llevamos a cabo juntos.

Y finalmente a Jizeth, Gabriel, Aldo, Oscar, Andrés y todos los compañeros que me acompañaron a lo largo de esta etapa, que me apoyaron y que siempre estuvieron presentes para que todos pudiésemos completar esta gran etapa en nuestras vidas.

Nicolás Augusto Marín Pinzón

A mis padres, que han sido el pilar de mi vida para formarme como persona y profesional y son el apoyo con el que siempre he contado y sé que siempre contaré. A mis directores que brindaron su guía y soporte para lograr culminar este trabajo de la mejor manera. Y a mi familia y amigos por el apoyo emocional para impulsarme a dar lo mejor de mí para lograr este objetivo y los que me han llevado a este momento.

Aldo Marcel Rizo Casadiegos

Agradecimientos

A nuestro director Oscar, por la paciencia y dedicación a lo largo de este proyecto. Muchas gracias por compartir sus conocimientos y por ayudarnos en este arduo proceso en el que sin su compañía no hubiese sido posible completar

A nuestro codirector Juan Manuel, por atender nuestras inquietudes y por darnos las herramientas para permitir el desarrollo de esta investigación.

A todos nuestros profesores que siempre estuvieron presentes en el proceso de aprendizaje y por su dedicación en compartir todo su conocimiento.

Tabla de Contenido

	Pág.
Introducción	11
1. Objetivos	13
<i>1.1 Objetivo general</i>	<i>13</i>
1.2 Objetivos específicos	13
2. Modelado de un sistema de distribución con <i>LTC</i>	14
2.1 Transformadores <i>LTC</i>	14
2.2 Sistema de distribución de PandaPower	15
2.3 Sistema de prueba del IEEE.....	16
<i>2.3.1 Balanceo de sistemas trifásicos</i>	<i>18</i>
3. Modelado del funcionamiento de los sistemas.....	19
3.1 Curvas de demanda y de radiación	19
3.2 Creación del día semilla.....	20
4. Estrategia de aprendizaje reforzado	22
4.1 Recompensa	27
4.2 Políticas.....	27
4.3 Entrenamiento	28
4.4 Generador Virtual de Transiciones (GVT)	29

5. Resultados	31
5.1 Resultados preliminares	32
5.2 Simulación en el sistema de 9 nodos	34
5.3 Sistema de 34 nodos.....	39
5.4 Estudio de sensibilidad	43
6. Conclusiones	46
7. Recomendaciones y trabajos futuros.....	48
Referencias Bibliográficas	50

Lista de Tablas

	Pág.
Tabla 1 Recompensa promedio y posiciones de taps de la prueba.	36
Tabla 2 Recompensa acumulada y resultados de la regulación en la barra 3.	38
Tabla 3 Recompensa promedio y posiciones de taps del entrenamiento 1.	41
Tabla 4 Recompensa acumulada y resultados de la regulación.	42
Tabla 5 Recompensa acumulada y resultados de la regulación.	45

Lista de Figuras

Figura 1 Modelo de un LTC	15
Figura 2 Topología del Sistema de 33 nodos de PandaPower	16
Figura 3 <i>Sistema de 9 nodos</i>	16
Figura 4 <i>Sistema de 34 nodos</i>	18
Figura 5 Gráfica de generación fotovoltaica (Irradiancia [W/m ²] vs tiempo[horas])	20
Figura 6 Gráfica de perfil de demanda para estrato 4 de la ESSA.....	21
Figura 7 Esquema de la estrategia de aprendizaje reforzado aplicada.....	23
Figura 8 Ejemplo de construcción de la matriz M.....	24
Figura 9 Lazos creados para el sistema de 34 nodos	24
Figura 10 Ubicación y modelado de los LTCs en la matriz M.....	26
Figura 11 Recompensas del sistema de 9 nodos con cada método.....	33
Figura 12 Resultados de tensión de cada escenario y comparación de la recompensa.....	35
Figura 13 Posiciones de los taps día 12 del sistema de 9 nodos	37
Figura 14 Pérdidas de potencia del sistema de 9 nodos para 21 días.....	39
Figura 15 Nomenclatura de las tensiones en las barras para el sistema de 34 nodos	39
Figura 16 Resultados de tensiones en los nodos para cada escenario y comparación de la recompensa	40
Figura 17 Pérdidas de potencia del sistema de 34 nodos para 21 días.....	43
Figura 18 Tensiones antes y después del control en el test.....	44
Figura 19 Posiciones de los LTC1 y LTC2 en el test	44
Figura 20 Pérdidas de potencia del sistema de 34 nodos para 21 días.....	46

Lista de Apéndices

Ver apéndices adjuntos y pueden ser consultados en la base de datos de la biblioteca UIS

Apéndice A. Parámetros Sistema de 9 Nodos

Apéndice B. Vectores semilla de las pruebas en ambos sistemas

Apéndice C. Resultados finales

Apéndice D. Sistemas sin entrenamientos por 7 días

Apéndice E. Prueba preliminar de entrenamientos en sistema de 9 nodos

Apéndice F. Prueba de 9 nodos recompensa de nodos antes y después del LTC

Apéndice G. Estudio de sensibilidad

Apéndice H. Código general

Resumen

Título: Control de Taps de Transformadores en Sistemas de Distribución Usando Aprendizaje Reforzado*

Autores: Christopheerth Anderson Bonilla León, Nicolás Augusto Marín Pinzón, Aldo Marcel Rizo Casadiegos**

Palabras clave: Aprendizaje reforzado, Transformadores cambia tap bajo carga, PandaPower.

Descripción: En este trabajo de grado se propuso una solución al problema de la regulación de tensión en sistemas de distribución de energía eléctrica. Un procedimiento que permite la correcta operación de esta variable, según los niveles establecidos por las normas legales, es la instalación de transformadores cambia *taps* bajo carga, también llamados *Load Tap Changer*. Estos utilizan las distintas posiciones en las que se pueden ajustar para modificar la conexión de sus devanados y, en consecuencia, la tensión de salida. Debido a esto, se planteó el objetivo de diseñar una estrategia para controlar la regulación de tensión en sistemas de distribución mediante el ajuste de *taps* en transformadores usando aprendizaje reforzado.

Con el fin cumplir con lo propuesto, se utilizó una estrategia de aprendizaje reforzado, la cual es una rama del *machine learning*; esta se ejecuta sobre el sistema analizado para que actúe sobre las posiciones de los *taps* convenientes según las variables a controlar. Como resultado de la presente investigación, se evidenció que la herramienta diseñada permite realizar el proceso de ajuste de los LTC, así como adaptarse a diferentes sistemas sin importar su topología o condiciones externas e internas que se requieran para converger y operar. Adicionalmente, con este resultado se demostró una mejora de la regulación con respecto al sistema no controlado, cumpliendo en todo momento con los rangos establecidos por la norma NTC 1340.

* Trabajo de Grado

** Facultad de Ingenierías Fisicomecánicas. Escuela de Ingenierías Eléctrica, Electrónica y Telecomunicaciones. Ingeniería Eléctrica. Director: Oscar Alexis Galvis Díaz. Ingeniero Eléctrico Codirector: Juan Manuel Rey López. Doctor en Ingeniería.

Abstract

Title: Control of Transformer Taps in Distribution Systems Using Reinforcement Learning*

Author(s): Christopheerth Anderson Bonilla León, Nicolás Augusto Marín Pinzón, Aldo Marcel Rizo Casadiegos**

Key Words: Reinforcement learning, Transformers load tap changers, PandaPower.

Description: This degree work proposed a solution to the voltage regulation problem in electric power distribution systems. A procedure that allows the correct operation of this variable, according to the levels established by the legal norms, is the installation of Load Tap Changers. These use the different positions that can be adjusted to modify the connection of their windings and, consequently, the output voltage. Due to this, the objective was to design a strategy to control voltage regulation in distribution systems by adjusting transformer taps using Reinforcement Learning.

A Reinforcement Learning strategy, which is a branch of Machine Learning, was used to comply with the proposed; this strategy is executed on analyzed systems to act on the positions of the taps according to the variables to be controlled. As a result, it was evidenced that the designed tool allows the LTC adjustment process to be performed and adapted to different systems regardless of their topology or external and internal conditions required to converge and operate. Additionally, this result demonstrated an improvement in the regulation concerning the uncontrolled system, complying at all times with the ranges established by the NTC 1340 standard.

* Degree Work

** Faculty of Physicomechanical Engineering. School of Electrical, Electronic and Telecommunications Engineering. Electrical Engineering. Advisor: Oscar Alexis Galvis Diaz. Electrical Engineer. Coadvisor: Juan Manuel Rey López. Doctor in Engineering.

Introducción

Hoy en día, la electricidad es un servicio ampliamente accesible gracias al desarrollo tecnológico conseguido a través de los años. Según la Ley 142 de 1994, esta tiene un papel fundamental en el desarrollo económico y social. Se destaca que, además de la provisión de la energía eléctrica a los usuarios, es imperativo asegurar tanto la calidad de la energía como la fiabilidad del sistema, siempre cumpliendo con los altos estándares normativos en el sector. (Congreso de la república, 1994).

Continuando con la misma línea, para garantizar una buena calidad en el suministro de energía eléctrica, es indispensable controlar la regulación de tensión (NTC 1340, 2004). Según la Norma ESSA, la regulación de tensión es la “razón en porcentaje (%) entre la diferencia de magnitudes de la tensión en el receptor en vacío y a plena carga, con respecto a la magnitud de la tensión en el receptor a plena carga”. Uno de los métodos para controlar esta variable es el uso de transformadores LTC que, a través de cambiar la relación de transformación de energía eléctrica con sus diferentes posiciones, pueden modificar la regulación de tensión con el fin de cumplir con los estándares establecidos, legal y técnicamente como la NTC 1340 y la Norma ESSA (Juarez, 1995). Esto permite el buen funcionamiento de los sistemas y la seguridad de los equipos pertenecientes a los usuarios finales (Gao & Redfern, 2011).

Con el fin de controlar la regulación, se utilizó una estrategia de aprendizaje reforzado, la cual es una rama del *machine learning* que consiste en “aprender qué hacer -cómo asignar situaciones a acciones- para maximizar la señal de recompensa numérica. No se le dice al agente qué acciones debe realizar, sino que debe descubrir, probándolas, qué acciones producen la mayor recompensa” (Sutton & Barto, 2015). Lo anterior permitió que, por medio de un agente, se

controlen las posiciones de los *taps* gracias a un entrenamiento que se realizó más adelante en dos sistemas de distribución, también conocidos como entorno; los cuales han sido adaptados de bases de datos del Instituto de Ingenieros Eléctricos y Electrónicos o IEEE, por sus siglas en inglés, y PandaPower, una librería de Python creada para resolver sistemas de potencia. Es importante destacar que los sistemas utilizados no se replicaron, se adaptaron, con el fin de tener redes eléctricas que se encuentran de manera recurrente en el entorno (Moya, 2014).

Lo expuesto, tiene la finalidad de enseñarle al agente a decidir cómo manipular los *taps* de manera que la regulación sea la mejor posible, independiente de los parámetros iniciales del sistema como la resistencia e inductancia de las líneas o de los parámetros variables como lo son la carga y la generación. Todas estas tareas fueron realizadas en Google Colaboratory (Google, 2023) y programadas en el lenguaje de Python, usando su librería PandaPower (Ernst et al., 2004a; Xu et al., 2020).

Finalmente, se analizaron las ventajas que existen al utilizar esta estrategia para controlar la regulación de tensión en sistemas de distribución. Posteriormente, los resultados obtenidos fueron expuestos por medio de tablas de datos y gráficas con el objetivo de verificar los parámetros de calidad de la energía, la regulación de tensión y las pérdidas de potencia, los cuales son importantes a la hora de examinar la operación de un sistema de energía eléctrica, así como las desventajas de utilizar este tipo de soluciones que aún se encuentran en proceso de desarrollo.

1. Objetivos

1.1 Objetivo general

Diseñar una estrategia para controlar la regulación de tensión en sistemas de distribución mediante el ajuste de *taps* en transformadores usando aprendizaje reforzado.

1.2 Objetivos específicos

- Modelar la operación de un sistema de distribución incluyendo transformadores ajustables para la regulación de tensión.
- Formular una estrategia de aprendizaje reforzado para controlar el ajuste de los *taps* de los transformadores.
- Definir un proceso de entrenamiento para la estrategia de aprendizaje reforzado propuesta.
- Simular un caso de estudio específico para analizar los resultados operativos de la estrategia diseñada.
- Estudiar la sensibilidad de las decisiones de diseño tomadas.

2. Modelado de un sistema de distribución con *LTC*

Los sistemas de distribución son un componente de los sistemas de potencia, cuya principal función es energizar las cargas conectadas al sistema, por medio de la transferencia de la energía que proviene de los generadores distribuidos y de los sistemas de transmisión. Entre los principales elementos que conforman a estos sistemas se encuentran las subestaciones de distribución, los alimentadores primarios y los transformadores de distribución (Juarez, 1995). Las primeras se definen como los elementos encargados de recibir la energía y convertirla en los niveles de tensión adecuados, las cuales generalmente manejan potencias en el orden de los MW. Los segundos, son circuitos que permiten el flujo de potencia desde la subestación hasta los transformadores de distribución, mientras que últimos son los que reducen el voltaje a los niveles de tensión del usuario final.

2.1 Transformadores *LTC*

Uno de los componentes fundamentales de los sistemas eléctricos de potencia son los transformadores, los cuales desempeñan un papel crucial al modificar los niveles de tensión para satisfacer las necesidades tanto en el transporte de la energía eléctrica, como en el suministro al usuario final. Una alternativa aún más llamativa es la implementación de transformadores *LTC*, que posibilitan la variación de la relación de transformación. Esto permite una adaptación más eficiente a las condiciones del sistema, considerando las fluctuaciones instantáneas provocadas por cambios en el consumo de cargas o la generación en los generadores distribuidos (GD).

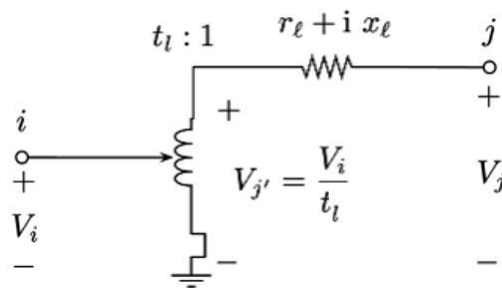
Estas variaciones de la relación de transformación ocurren al modificar el punto de conexión entre la línea de transmisión con el devanado del transformador. Dependiendo del número de espiras que se conecten, se representan las diferentes posiciones en que el *LTC* puede

ajustarse. Esta opción conlleva una mejora en la calidad de la energía suministrada, ya que, al reajustar la potencia activa y reactiva, se impacta en la modificación de la tensión del sistema. Asimismo, este proceso permite que la regulación se acerque más a los valores estipulados por la normativa (Gao & Redfern, 2010, 2011; NTC 1340, 2004; Kellerman et al., 1991).

En el escenario específico abordado en esta investigación, se empleó un modelo de LTC con 33 posiciones, operando en un ‘rango’ de 0.9 p.u a 1.1 p.u., donde cada paso del tap representaba una variación del 0.625 %. Además, la posición seleccionada del transformador se definió como ‘t’ en el modelo mostrado en la Figura 1.

Figura 1

Modelo de un LTC



Nota. Ilustración adaptada del artículo (Xu et al., 2020)

2.2 Sistema de distribución de PandaPower

Para modelar el sistema de distribución se empleó *PandaPower*, una herramienta de código abierto en forma de librería de Python diseñada para el análisis y la optimización de sistemas de distribución (PandaPower, 2023). Se seleccionó un sistema de 33 nodos con topología radial de su base de datos, como se presenta en la Figura 2. Este sistema, al ser preexistente del propio PandaPower, proporcionó todos los elementos necesarios para un sistema coherente, capaz de operar en las condiciones establecidas.

El caso de estudio se centró en un sistema de 9 nodos obtenido al reducir el sistema de 33 nodos, ajustando a conveniencia las líneas, las cargas (nodos 2, 4 y 7), la ubicación de dos transformadores *LTC* entre las líneas (2,3) y (6,5) (con el lado de alta de los *LTC* en el primer nodo) y la adición de generación en los nodos 6 y 8, simulando sistemas fotovoltaicos. El sistema de 9 nodos modificado se muestra en la Figura 3, donde las flechas negras representan las cargas y las amarillas, la generación distribuida. La información topológica del sistema se encuentra en el Apéndice A.

Figura 2

Topología del Sistema de 33 nodos de PandaPower

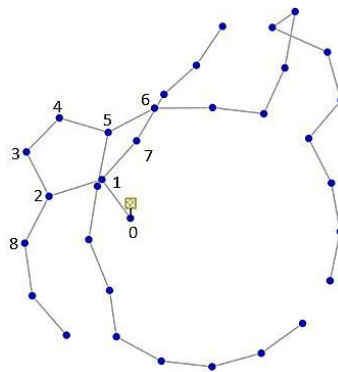
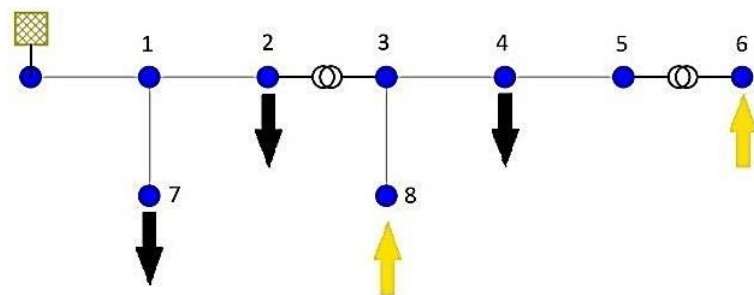


Figura 3

Sistema de 9 nodos



2.3 Sistema de prueba del IEEE

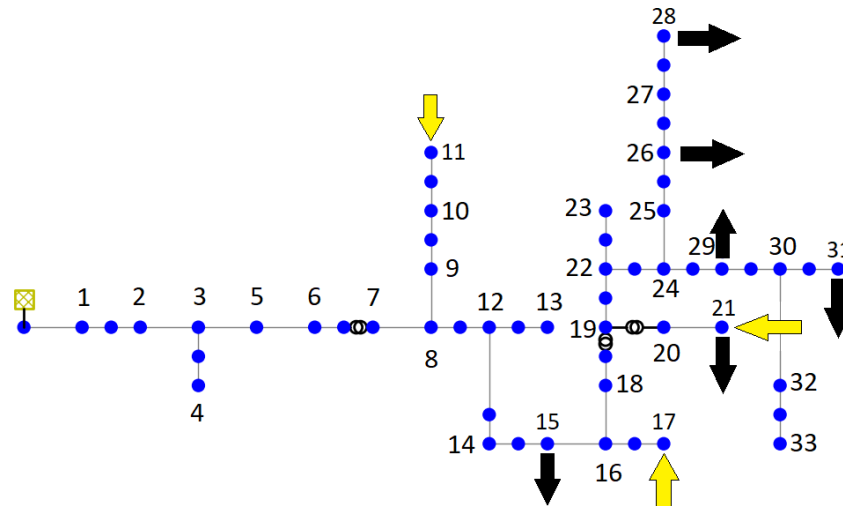
En respuesta al crecimiento de la distribución eléctrica a nivel mundial, el IEEE desarrolló los sistemas de prueba denominados *Test-Feeders*, con los cuales se buscó proveer modelos de

sistemas de distribución eléctrica disponibles abiertamente al público con una gran variedad de diseños, con el fin de lograr que los investigadores pudieran enfrentarse a los diferentes retos que proveen cada una de las características nominales que presentan estos sistemas (IEEE, 2017).

Los *Test-Feeders* van desde 4 hasta los 500 nodos; además, comparten un conjunto de datos en común, que incluye la presencia de cargas puntuales y cargas distribuidas, conectadas a lo largo de las líneas, todas estas con su respectivo tipo de conexión y potencia nominal. Entre las características cruciales para el estudio, también hacen parte el espaciamiento de conductores y el material del conductor, la resistencia, la capacidad amperimétrica y el diámetro; esto proporciona información tanto para conductores aéreos como subterráneos.

Considerando la naturaleza desbalanceada de los sistemas del IEEE, se aplicaron métodos de balanceo al sistema obtenido de su base de datos; esto por medio de los métodos detallados en la siguiente sección, en donde se busca lograr simetría en las cargas y las líneas (Anderson, 1995; Kersting, 2002).

Se optó por utilizar el sistema de pruebas del IEEE de 34 nodos, el cual mostró las características deseadas para el desarrollo de este proyecto. Este sistema radial incluyó cargas puntuales, dos transformadores LTC y cargas distribuidas. Con el fin de adaptar el sistema a las características adecuadas, se llevaron a cabo las siguientes modificaciones. En primer lugar, se incorporaron generadores al sistema, que fueron modelados como paneles solares. Posteriormente, se realizó un balance de cargas y líneas, como se detalla a continuación. Como se muestra en la Figura 4, tras estas adaptaciones, se obtuvo el sistema de 34 nodos con flechas negras que indican la ubicación de las cargas y flechas amarillas que representan los generadores distribuidos.

Figura 4*Sistema de 34 nodos*

2.3.1 Balanceo de sistemas trifásicos

En la tarea de modelar los sistemas de distribución en *PandaPower*, fue necesario balancear los sistemas trifásicos provenientes de los *Test-Feeders*. Esto debido a que los sistemas seleccionados presentan desbalance, lo cual se aparta del enfoque previsto en esta investigación. Para llevar a cabo el proceso de balanceo se usó la teoría de las componentes simétricas, según Anderson (1995), complementada con el método de reducción de Kron.

En la ejecución de este proceso, fue crucial que el sistema se reconociera como traspuesto, y que la tensión en el neutro fuera 0. Una vez verificado esto, se procedió tomando la matriz de impedancias Z del sistema y se eliminó la tensión del neutro mediante el método de reducción de Kron (Martínez, 2014). Posteriormente, se investigaron dos opciones de desarrollo: el primer método consistió en asumir que el sistema está traspuesto, permitiendo sumar la impedancia de las tres fases y luego dividirla en tres. En el segundo método se calcularon los radios medios geométricos, y a partir de este parámetro, se determinó la impedancia de cada fase (Kersting, 2002). Una vez realizados los cálculos correspondientes con ambos métodos, se obtuvo el mismo

resultado, lo que confirmó la correcta ejecución del procedimiento y la adecuación de los resultados para continuar con el modelado del sistema.

3. Modelado del funcionamiento de los sistemas

Para llevar a cabo esta investigación se efectuaron los cálculos de flujos de potencia por medio de la ejecución del comando de *PandaPower* ‘*runpp*’. La elección de este comando se sustentó en su capacidad para llevar a cabo cálculos con un bajo nivel de error, requiriendo un número reducido de iteraciones. Este comando utiliza el método de Newton-Raphson, el cual es una expansión de la serie de Taylor, que resuelve el cálculo de flujos a través de estimaciones de los parámetros eléctricos (Grainger et al., 2001).

3.1 Curvas de demanda y de radiación

Con el propósito de simular variaciones en las condiciones de cargabilidad y de generación en los sistemas de distribución analizados, se implementó el día semilla, el cual es un modelo de día creado para actuar como base de las curvas de demanda eléctrica y de generación fotovoltaica. El día semilla se originó a partir del promedio hora a hora de los parámetros mencionados anteriormente, extraídos a lo largo de un año desde datos reales. La primera de las dos curvas mencionadas es una gráfica que permite representar de manera estimada los hábitos de consumo de energía eléctrica de los usuarios residenciales, la cual depende de factores como el estrato social, que es suministrado por el distribuidor de energía eléctrica (Norma ESSA, 2020). En segundo lugar, se tiene a la curva de generación fotovoltaica, la cual es una gráfica que se obtuvo de la herramienta Nasa Power, que presenta la generación horaria de un sistema de paneles fotovoltaicos según su ubicación geográfica y varía adicionalmente según las condiciones climatológicas (NASA - National Aeronautics and Space Administration, 2021).

El objetivo de la creación de los dos días semilla fue obtener variabilidad sobre el consumo de las cargas y la generación de los generadores distribuidos, obligando al sistema a actuar en situaciones con condiciones cambiantes y así verificar que se cumplió con el objetivo de ajustar el tap de los LTC según las condiciones que puedan existir en un sistema real.

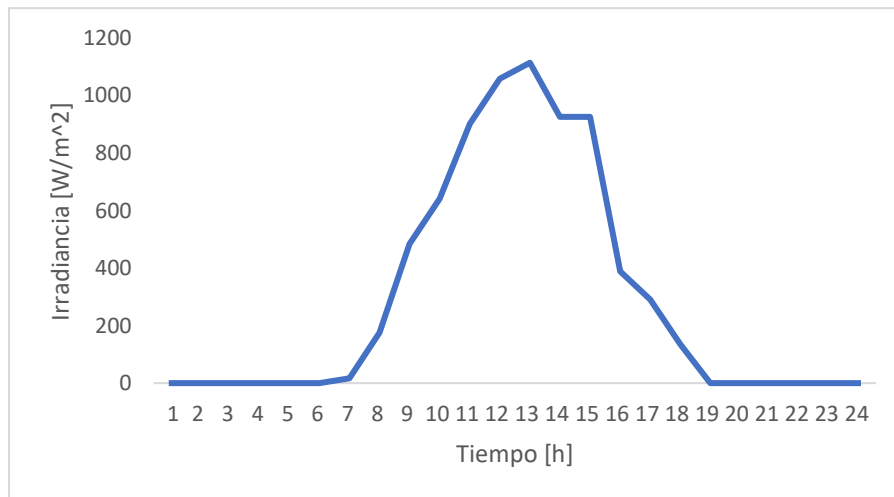
3.2 Creación del día semilla

Con el fin de obtener la variabilidad de los generadores distribuidos y de las cargas, se creó un vector a partir de los días semilla. Este fue una representación vectorial de los datos horarios obtenidos para contar con perfiles de la irradiancia solar y del consumo de energía eléctrica, presentado en un día promedio. La Figura 5 exhibe el vector semilla correspondiente a la generación fotovoltaica junto con su gráfica asociada, ilustrando dicho comportamiento.

Figura 5

Gráfica de generación fotovoltaica (Irradiancia [W/m²] vs tiempo[horas])

Horas	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Irradiancia [W/m ²]	0	0	0	0	0	0	16	175	484	640	900	1056	1112	925	925	387	290	135	0	0	0	0	0	0



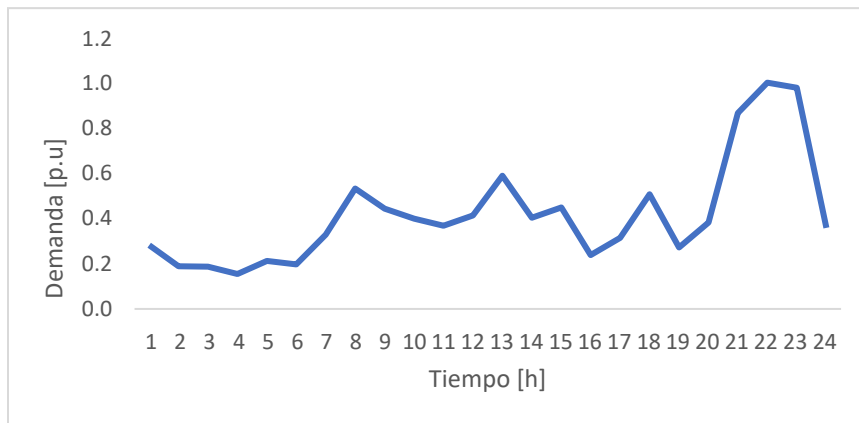
Asimismo, para modelar las cargas, se creó un vector semilla basado en la Norma ESSA, con el consumo en por unidad promedio de energía para un perfil de carga residencial estrato 4

(Norma ESSA, 2020; Hernández & Carrillo, 2020). La Figura 6 presenta esta información, mostrando tanto el vector como la gráfica correspondiente al día semilla.

Figura 6

Gráfica de perfil de demanda para estrato 4 de la ESSA

Horas	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Demanda [p.u]	0.28	0.19	0.19	0.15	0.21	0.20	0.33	0.53	0.44	0.40	0.37	0.41	0.59	0.40	0.45	0.24	0.31	0.51	0.27	0.38	0.87	1.00	0.98	0.36



Finalmente, se realizó un tratamiento probabilístico de distribución normal a estos vectores, con una desviación estándar según la conveniencia del estudio, en donde se contó con resultados diversos pero que se mantengan en un rango de operación real de un sistema eléctrico. Para este trabajo, la desviación estándar (σ) correspondió a 0.2 y el valor promedio a 1, con el fin de no presentar desviaciones poco razonables con respecto al promedio de las bases de datos. Los vectores finales tras haber realizado el tratamiento probabilístico se pueden visualizar en el Apéndice B.

4. Estrategia de aprendizaje reforzado

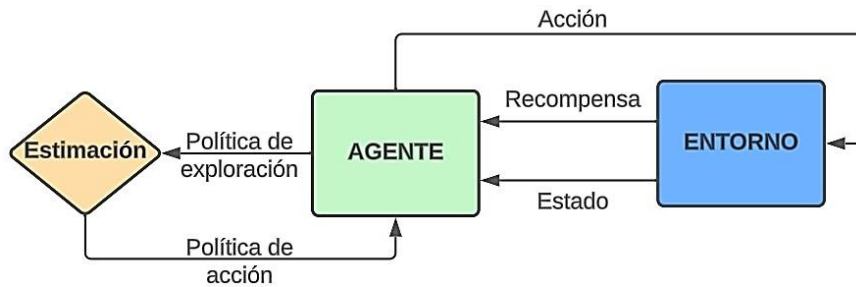
El aprendizaje reforzado es una rama de la inteligencia artificial aplicada a la solución de problemas por medio del aprendizaje de su **entorno**, el cual se caracteriza por ser un circuito cerrado de aprendizaje. A partir de este circuito se debe aprender qué acción tomar por medio de un mapeo de los posibles estados, determinando así, cuál presenta la mayor **recompensa**; para así influenciar las acciones realizadas y consecuentemente las próximas entradas al sistema de aprendizaje (Sutton & Barto, 2015).

Es necesario tener en cuenta que para poder determinar cuáles son las acciones con las mayores recompensas, el sistema debe descubrirlas por medio de la experimentación de todas las posibles situaciones en cada una de las iteraciones. Estas tres características, que sea un circuito cerrado, que aprenda por medio de la experimentación, y que incluya un sistema de recompensas, son las que la distinguen del resto de métodos posibles de aplicación (Ernst et al., 2004b).

Inicialmente, se definieron los aspectos claves de una estrategia de aprendizaje reforzado: el **entorno**, el **estado**, la **recompensa**, la **política** y el **agente**. Naturalmente, el **entorno** es el sistema de distribución modelado; el **agente**, es el propio algoritmo quién se encargó de modificar la posición del *LTC*; el **estado**, representado por la posición del *LTC* y, por último, la **política** se dividió en dos: una de **acción** y una de **exploración**; debido a que en un sistema real de operación el **agente** no contará con los datos generados por cada uno de los **estados** posibles. Por tal razón, se planteó una **política de exploración** para determinar qué **estados** estimar al **agente**. El esquema de la estrategia se observa en la Figura 7.

Figura 7.

Esquema de la estrategia de aprendizaje reforzado aplicada



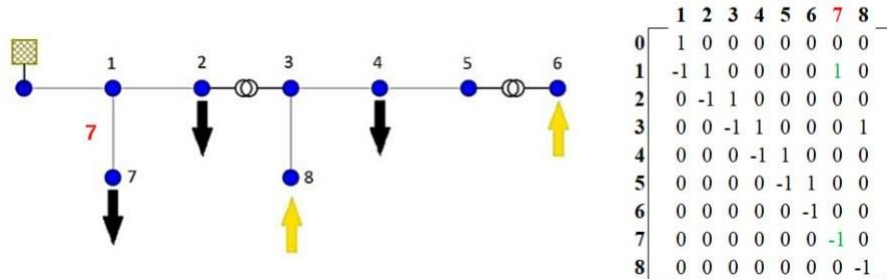
La **política de exploración** le indica al **agente** como realizar la estimación y evaluar el espacio de **estados**. Por otra parte, se tiene la **política de acción**, la cual indica qué decisión tomar tras la aplicación de la **política de exploración** (Sutton et al., 2015). En conjunto, estas políticas proporcionan un marco integral para que el **agente** aprenda y tome decisiones de manera efectiva en el **entorno** de aprendizaje reforzado.

En el artículo (Xu et al., 2020), se propone el uso de *Batch Reinforcement Learning*, una técnica de RL, para el ajuste de los taps en los transformadores de control de regulación. Para el análisis y desarrollo del método propuesto en el artículo es necesario representar en forma de matriz, en este caso denominada 'mt', al conjunto de líneas (l), y a su vez a las barras (i, j), que se identifican como los extremos de cada una de las líneas existentes en el sistema. En esta matriz la posición (i, l) se va a definir como 1, mientras que el -1 se utilizará para definir la posición (j, l), asimismo, es importante mencionar que las posiciones restantes se completarán con 0.

Tomando el sistema de 9 nodos como ejemplo, para agregar la 'línea 7' a la matriz M se localizan los nodos (i,j) que corresponden al (1,7) y se adicionan a la matriz como se muestra en la Figura 8, resaltado en verde. A partir del desarrollo de esta matriz, se extrae la primera fila, perteneciente a la barra *Slack* convirtiéndose en el vector 'm' y así la matriz resultante corresponde a 'M'.

Figura 8

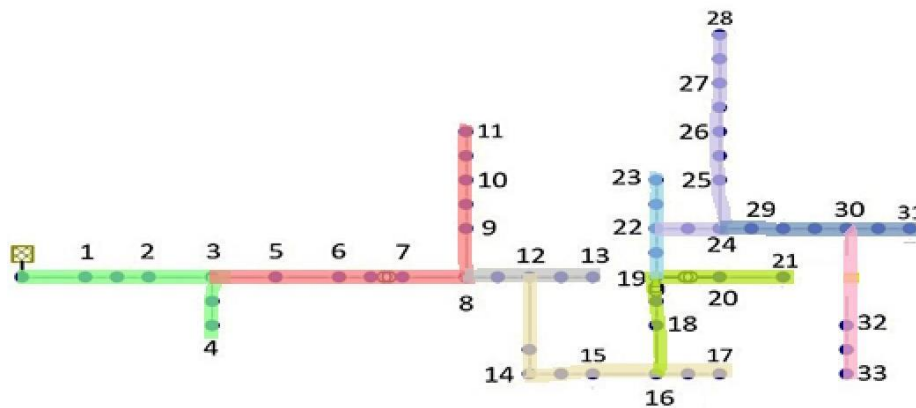
Ejemplo de construcción de la matriz M



La propuesta para la creación de esta matriz implica inicialmente la creación de dos matrices: una para la matriz M y la otra para almacenar la información de la conexión de las líneas que conforman el sistema. Lo anterior, haciendo uso de un conjunto de nodos conectados entre sí, denominados como lazo. La información sobre estos lazos se ingresó mediante un vector que contiene los nodos conectados de forma consecutiva. Por ejemplo, el lazo [1,3,4] representa una conexión entre el nodo 1 y 3 seguida de otra entre el 3 y 4. A continuación, en la Figura 9 se observa cómo se crearon los lazos para el sistema de 34 nodos, donde cada color representa un lazo.

Figura 9

Lazos creados para el sistema de 34 nodos



Una vez definidos los lazos y las matrices con las dimensiones correspondientes, la matriz `mt` , con n filas y $n-1$ columnas, donde n es el número total de nodos del sistema, y la matriz `posi_line` con 2 filas y $n-1$ columnas, se procede a ingresar la información de cada lazo a las matrices mediante la función `crearM` .

Algoritmo 1 Funcion crearM: Función para agregar los lazos a la matriz M

Entradas:

lazo: vector con los nodos que componen el lazo

matrizM: Matriz M del sistema (en su estado actual)

posi_line: Matriz para almacenar los nodos a lo que se conecta cada línea

línea: Número de la última línea agregada a la matriz.

Salidas:

matrizM: Matriz M del sistema

posi_line: Matriz para almacenar los nodos a lo que se conecta cada línea

línea: Número de la última línea agregada a la matriz.

For $i1 = 0, 1, \dots, (\text{número de nodos del lazo} - 1)$

Reemplaza el valor de la posición ($\text{lazo}(i1)$, línea) en matrizM por 1

Reemplaza el valor de la posición ($\text{lazo}(i1+1)$, línea) en matrizM por -1

Guarda el nodo inicial en la posición (0, línea) de posi_line

Guarda el nodo final en la posición (1, línea) de posi_line

línea=línea+1

End

La información de los LTC en la matriz M se ingresa de la siguiente manera: primero se identifica la línea en la cual está ubicado el LTC, esta línea tendrá asociados dos nodos de los cuales el lado correspondiente a alta tensión que se representa con un -1 y el lado de baja se cambia por el valor de $\frac{1}{t^2}$, como se explica en la Figura 10, siendo esto es interpretado por el algoritmo mediante la matriz 'posi_line'. Debido a esto, el algoritmo 2 expresa cómo se ingresan las posiciones de los taps en la matriz M; inicialmente, se necesita el vector para conocer la ubicación

del LTC mediante la matriz que contiene la información de la conexión de cada línea en el sistema, el número de la línea en que se encuentra el LTC, teniendo en cuenta que la numeración de estas arranca en uno y no en cero, la función automáticamente hace la corrección para usarse de forma correcta y por último, la posición del tap para calcular el valor de t.

Figura 10

Ubicación y modelado de los LTCs en la matriz M

	1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8
0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1	-1	1	0	0	0	0	1	0	0	1	-1	1	0	0	0	1	0
2	0	-1	1	0	0	0	0	0	0	2	0	-1	-1	0	0	0	0
3	0	0	-1	1	0	0	0	1	0	3	0	0	$1/t1^2$	1	0	0	1
4	0	0	0	-1	1	0	0	0	0	4	0	0	0	-1	1	0	0
5	0	0	0	0	-1	1	0	0	0	5	0	0	0	0	-1	-1	0
6	0	0	0	0	0	-1	0	0	0	6	0	0	0	0	0	$1/t2^2$	0
7	0	0	0	0	0	0	-1	0	0	7	0	0	0	0	0	0	-1
8	0	0	0	0	0	0	0	-1	0	8	0	0	0	0	0	0	-1

Algoritmo 2 Función fun_ltc: Función para reemplazar valores de ltc en la matriz M

Entradas:

matrizM: Matriz M del sistema

posi_line: Matriz para almacenar los nodos a lo que se conecta cada línea

ubic_ltc: Línea donde se encuentra el ltc.

posi_ltc: Valor de la posición donde se encuentra el tap

Salidas:

matrizM2: Matriz M del sistema con la posición de los taps indicados

Crea un vector Z igual a (Ubic_ltc - 1) debido a que las líneas se cuentan de 1 en adelante

Crea una variable valor_ltc igual $(1 + (\text{posi_ltc} * 0.00625))$ que corresponde a valor del tap

Crea una copia de la matrizM llamada matrizM2

Reemplaza el valor de la posición (nodo inicial de la línea del ltc, línea del ltc) en matrizM2 por

-1

Reemplaza el valor de la posición (nodo final de la línea del ltc, línea del ltc) en matrizM2 por $1/(\text{valor_ltc})^2$

4.1 Recompensa

La finalidad de la estrategia consiste en mantener la regulación de la tensión de acuerdo con los valores preestablecidos. El cálculo del valor de la **recompensa** se basó principalmente en la evaluación de las tensiones del sistema mediante la aplicación de la ecuación de **recompensa** propuesta en el artículo (Xu et al., 2020), realizando unas modificaciones mencionadas en la sección 5.1. En última instancia, el **agente** opta por la disposición que produzca la mayor **recompensa**, seleccionada entre aquellas generadas por el algoritmo.

$$Recompensa = -\frac{1}{N} \|V - Vref\| \quad (1)$$

En esta ecuación V representa el vector con las magnitudes al cuadrado de las tensiones en el **estado** evaluado, mientras $Vref$ es un vector con los valores de referencia al cuadrado para cada nodo, ambos excluyendo a la barra *slack* y, por último, N es el número de nodos, sin considerar la barra *slack*.

La **recompensa** se planteó con un valor negativo ya que el objetivo de la estrategia es que las tensiones del sistema sean cercanas a las tensiones de referencia planteadas. Por lo tanto, entre menor sea la diferencia de estos valores la recompensa será mayor. (Xu et al., 2020)

4.2 Políticas

La **política de exploración** más sencilla consiste en explorar cada uno de los **estados** posibles. Por ejemplo, si el sistema cuenta con un único *LTC* tendría que explorar los 33 **estados**, lo cual no es un número significativo de iteraciones para un caso, pero cuando el sistema cuenta con dos *LTC* este número se incrementa a 1089 (33^2). Esto aumenta la carga computacional del

algoritmo exponencialmente a medida que el sistema a estudiar cuente con una mayor cantidad de *LTC*. Por esta razón, se plantea una **política de exploración** que permite disminuir el número de **estados** a explorar reduciendo la carga computacional y el tiempo de respuesta del algoritmo. Adicionalmente, se planteó una **política de acción** que consistió en utilizar el **estado** que genera la mayor **recompensa** para cada iteración basado en los resultados de la exploración realizada.

4.3 Entrenamiento

El **entrenamiento** se refiere al proceso mediante el cual un **agente** interactúa con un **entorno** y ajusta su comportamiento en función de las **recompensas** y penalizaciones que recibe. El **entrenamiento** implementado para esta estrategia permitió mejorar la **política de exploración** con el enfoque de reducir el número de **estados** posibles a estudiar en cada iteración. Para esto, se precisó un periodo de prueba en el sistema, en el cual explore todos los **estados** posibles y progresivamente reduzca el rango de estados dependiendo donde se encontraron las mayores **recompensas** en el periodo de prueba establecido.

A su vez, este se podría llevar a cabo durante un periodo de tiempo predeterminado, o mediante el uso de un historial del sistema a evaluar de años anteriores; esto para efecto de un caso real donde no se disponga de un historial y con el propósito de evaluar la eficacia del **agente** ante un sistema como un conjunto de datos de **entrenamiento** relativamente limitado. Se planteó un **entrenamiento** dividido en 3 secciones, donde dos de estas fueron asignadas para los entrenamientos y una última designada como prueba del **agente**. Dado que el objetivo principal es mejorar la capacidad del **agente** para reducir el número de **estados** que debe estimar, la **política de exploración** se planteó de la siguiente manera.

En el primer **entrenamiento** el **agente** operó como en las pruebas anteriores, usando una **política de exploración** total entre las posiciones -16 y 16 para ambos *LTC* (1089 casos), tras esto

se fue creando un historial con las posiciones de los *taps* elegidos. Para el entrenamiento 2 la **política** se modificó de la siguiente manera.

$$\text{Posiciones de la politica LTC1} = (\min(\text{LTC1}) - \text{rango}, \max(\text{LTC1}) + \text{rango}) \quad (2)$$

Con base en la ecuación 2, se definieron los límites de la **política de exploración** para cada LTC de forma independiente, con un número mayor de **estados** para flujos en comparación con el Generador Virtual de Transiciones (método explicado en la siguiente sección), debido a que el tiempo de **entrenamiento** es corto y existía la posibilidad de presentar una posición mayor o menor.

Durante el entrenamiento 1 se almacenaron los máximos y mínimos alcanzados, con los cuales se define la **política de exploración** del **agente** y el entrenamiento 2 consistió en reducir el valor del ‘rango’ establecido inicialmente. Por último, el test desempeñó la función de verificar que nunca se excedieran los límites predefinidos por la **política**, existiendo la opción de reducir el ‘rango’ incluso a cero si se considera necesario, para reducir el número de operaciones realizadas por el **agente** y así disminuir el tiempo de compilación. En caso de que se excedieran los límites predefinidos, se podría considerar repetir el proceso de **entrenamiento** tomando en cuenta esta información o, alternativamente, el sistema podría entrar en operación utilizando la **política** establecida a discreción del usuario.

4.4 Generador Virtual de Transiciones (GVT)

Para la técnica de estimación del GVT se implementó como base el modelo de flujo de distribución linealizado (LinDisFlow, por su acrónimo en inglés), el cual es un modelo simplificado del análisis de flujos debido a que linealiza todos los términos cuadráticos dentro de sus ecuaciones con el fin de superar la complejidad de los cálculos de las ecuaciones cuadráticas (Xu et al., 2020). De igual forma, se simplifican las ecuaciones de flujos al eliminar los términos

relacionados con las pérdidas en las líneas. A partir de este modelo, se realizan modificaciones conforme a los autores (Xu et al., 2020). Estos ajustes consisten en adaptar las ecuaciones del modelo LinDisFlow para relacionarlas con la matriz M del sistema analizado. Este proceso ha permitido la obtención de modelos adecuados para el desarrollo de la investigación, permitiendo realizar estimaciones de los valores de tensión, sin conocer a profundidad los parámetros del sistema.

Una vez obtenida la matriz M y la capacidad de variar las posiciones de los *LTC*, falta realizar la estimación de los valores de tensión a partir de la ecuación 3. Donde la matriz M total se representa como ' mt ', la matriz ' M ' resultante de quitar la fila de la barra *slack*, ' m ' el vector correspondiente a la fila de la barra *slack*, ' t ' la posición de los *taps*, ' v ' el vector compuesto por los valores de la magnitud de tensión al cuadrado de cada barra sin contar la *slack* y ' v_0 ' el valor de tensión al cuadrado de la barra *slack*. Los apóstrofes indican el instante $k+1$, inmediatamente después de cambiar la posición del *LTC* y los demás el instante k o inicial de donde se van a estimar los valores de tensión.

$$v' = (M(t')^T)^{-1}(M(t)^T v + m(t)v_0 - m(t')v_0) \quad (3)$$

Teniendo en cuenta que para que la ecuación 3 los vectores deben ser vectores columnas, todo este procedimiento lo realiza la función *GVT*.

Algoritmo 3 Función *GVT*: Función para aplicar la técnica de estimación por *GVT*

Entradas:

N : número de nodos sin contar la barra *slack*

mt : Matriz M del sistema con la posición de los *taps* indicados en un instante k

$mt2$: Matriz M del sistema con la posición de los *taps* indicados en un instante $k+1$

$vx1$: Vector de tensiones sin la barra *slack* en el instante k

Salidas:

$v2_2$: Vector de tensiones al cuadrado sin la barra *slack* en el instante $k+1$

v2: Vector de tensiones sin la barra slack en el instante k+1
Crea matriz M1 igual a la matriz mt sin la primera fila
Crea un vector columna m1 igual a la primera fila de la matriz mt
Crea matriz M2 igual a la matriz mt2 sin la primera fila
Crea un vector columna m2 igual a la primera fila de la matriz mt2
Crea un vector columna vgvT con las tensiones vx1 al cuadrado
Crea una variable v0 con la magnitud al cuadrado de la barra slack
Crea un vector v2_2 para guardar el resultado de aplicar la ecuación 3
Crea un vector de ceros v2
For i=0,1..., (N-1)
 Reemplaza el valor de la posición (i) en v2 por la raíz cuadrada de v2_2(i)
End

5. Resultados

En esta sección se presentan los tres tipos de escenarios propuestos para el análisis. En primer lugar, está la técnica de estimación por flujos, la cual se llevó a cabo mediante la biblioteca PandaPower y sirve como punto de referencia para los demás escenarios. En segundo lugar, se abordó la técnica de estimación del GVT, como se describió detalladamente en la sección 4.4. En este método, los valores de tensión en el instante k se calculan a partir de la técnica por flujos. Por último, se encuentra el escenario denominado "Sin control", definido como la operación del sistema en una posición estática de los taps establecida en las posiciones cero para cada LTC.

Además, se definieron dos momentos clave en el sistema para interpretar los resultados, denominados como "antes" y "después" del control. El primer instante corresponde al estado del sistema previo a realizar los cambios en las posiciones de los taps, mientras que el segundo refleja la situación tras haber modificado las posiciones, lo que genera la obtención de dos resultados por cada hora. En cuanto a la política de exploración, se planteó para el entrenamiento 2, un 'rango'

de 4 posiciones del LTC para GVT y de 6 posiciones para la estimación por flujos. Este ‘rango’ se reduce durante el test a 2 y 4 posiciones de los taps respectivamente, aplicando esta política a ambos sistemas de prueba.

Por parte de los días semilla, se estableció un porcentaje de variación del 20% tras realizar las pruebas de los resultados preliminares. Con el enfoque claro del estudio a realizar en los sistemas y al agente, lo siguiente por definir fue el tiempo. Considerando lo expuesto en la sección 4.3, se llevaron a cabo las pruebas en un periodo total de 21 días, siendo una ventana de tiempo adecuada para obtener resultados concluyentes de la estrategia, asignando 7 días para cada sección del estudio. Todo esto proyectado a que en un futuro sea aplicable a un sistema de distribución en funcionamiento.

5.1 Resultados preliminares

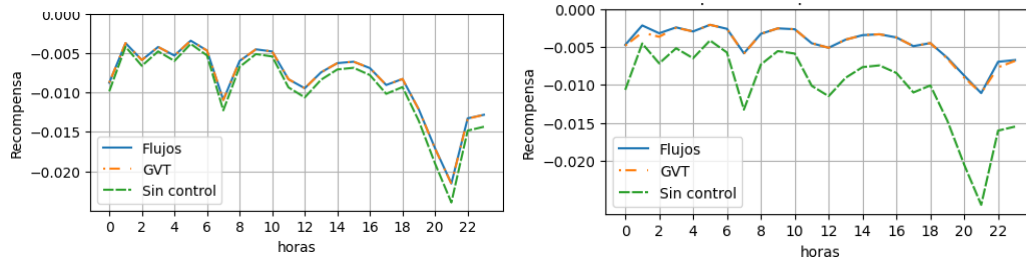
Inicialmente, los estudios de esta sección se realizaron en el sistema de 9 nodos, en el cual se implementó el valor de la recompensa, realizando pruebas en el sistema teniendo una variación del 10% al 20% en el día semilla y tomando todos los nodos para el cálculo de la recompensa del agente. Gracias a lo anterior se obtuvo unos resultados casi ineficientes, debido a que por más que se cambiaran las posiciones del *LTC*, alrededor de un tercio de los nodos tenían una afectación casi nula ante estas acciones.

Después de obtener estos resultados y realizar el estudio del control de la tensión de los nodos 1, 2 y 7, se planteó no utilizar estos tres nodos en el cálculo de la recompensa. En su lugar, se centró en mejorar la regulación de tensión en los nodos directamente relacionados con el control. Sin embargo, se optó por no eliminar estos nodos del sistema, ya que fueron tenidos en cuenta en la estimación de tensiones; teniendo presente que los resultados marcados como ‘TD’ representan el uso de todos los nodos para calcular la recompensa. Como resultado de esta estrategia, se

evidenciaron mejoras significativas en los valores de la regulación, recompensa y posiciones de los taps, con variaciones entre cada hora. El resultado total de esta prueba se puede observar en el Apéndice F.

Figura 11

Recompensas del sistema de 9 nodos con cada método



En la Figura 11, se presenta la recompensa obtenida por el agente mediante cada método. En la ilustración del lado izquierdo, se representa la recompensa utilizando todos los nodos (TD), mientras que al lado derecho se muestra el resultado al hacer uso exclusivamente de los nodos de interés. Dado que este valor está vinculado con la regulación de tensión, se destaca la eficiencia de cada método para disminuir su magnitud en el sistema. Antes de iniciar el entrenamiento, se llevó a cabo la prueba del funcionamiento del algoritmo con los dos entornos, haciendo uso de una política de exploración de $(-16,16)$ para ambos taps. Los resultados de esta prueba presentados en el Apéndice D, fueron favorables para ambos sistemas.

En vista de los datos calculados para esta sección, a continuación, se da a conocer como el agente opera el sistema en cada entrenamiento, haciendo uso únicamente de los nodos de interés para el cálculo de las recompensas. Inicialmente se hizo la prueba con el sistema de 9 nodos en un periodo de entrenamiento de 10 días, como se puede ver en el Apéndice E.

Algoritmo 4 Funcionamiento del agente con las dos técnicas de estimación

Entradas:

Matriz M
Ubicación de los LTC
Sistema diseño en panda power
Vectores semilla

Salidas:

Posiciones de los taps
Tensiones estimadas
Recompensas
Perdidas

Define el '*rango*' de posiciones para la política de exploración

For l=0, 1 ..., (número de días)

For j=0, 1 ..., 23

Cambia las condiciones de generados y carga según los vectores semilla

Hace un barrido de taps usando la técnica de estimación de flujos

Almacena los resultados del barrido

Aplica la política de acción el agente

Calcula con la técnica de estimación de flujos los valores de tensión de la posición inicial o de la hora anterior (instante k)

Realiza el barrido por la técnica de estimación GVT

Almacena los resultados del barrido

Aplica la política de acción

Almacena los resultados obtenidos tras realizar la operación con cada técnica

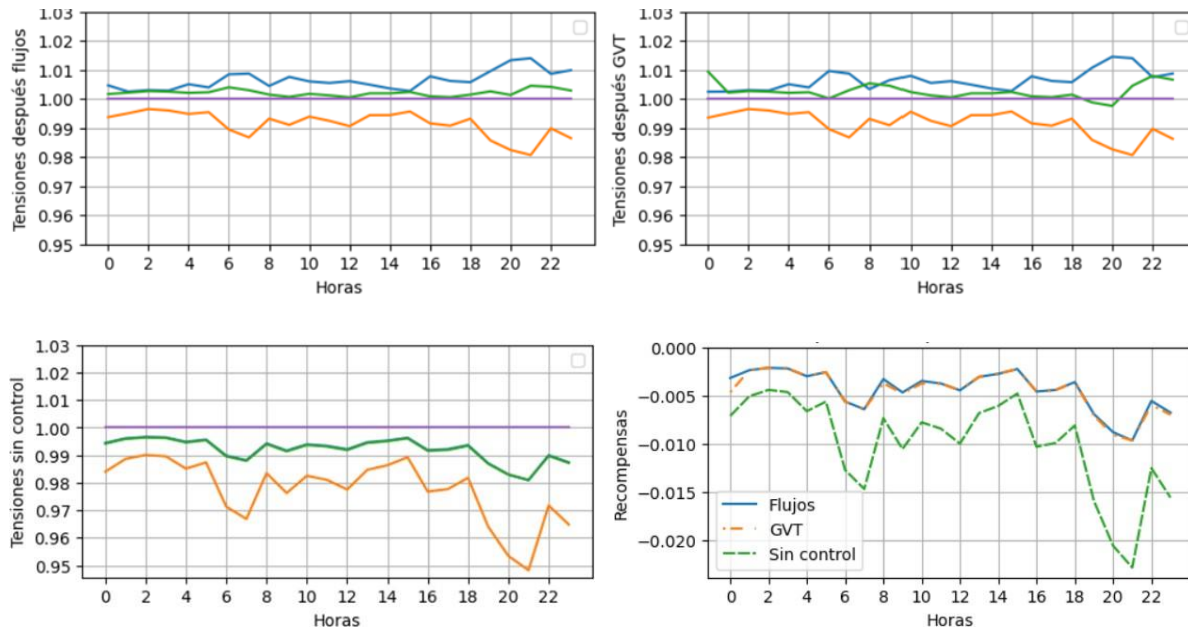
Teniendo en cuenta los resultados se reajusta la política de exploración

5.2 Simulación en el sistema de 9 nodos

Inicialmente, se puede presenciar la acción que tiene el control sobre los datos de tensión, y a su vez, se evidencia el gran impacto que tiene con respecto a los valores sin control en la Figura 12. Los resultados completos esta sección se encuentra en el Apéndice C.

Figura 12

Resultados de tensión de cada escenario y comparación de la recompensa.



Nota. Las gráficas corresponden a los valores de tensiones después de la aplicación de la política de acción por el agente.

Como se OBSERVA en la Figura 12, la curva morada representa un generador ubicado en la barra 8; por tal razón su tensión se mantiene constante. En contraste, las demás se ven significativamente afectadas por el control, aunque las curvas de recompensa presentan picos casi idénticos, se destaca la efectividad del agente en la mejora del sistema.

Respecto a las posiciones de los *taps*, se puede observar que los elegidos con las técnicas de estimación por flujos y GVT son bastante cercanos en la tabla 1, donde se muestra el promedio por cada hora durante los 21 días de prueba de la posición de los *taps* y la recompensa.

Tabla 1

Recompensa promedio y posiciones de taps de la prueba.

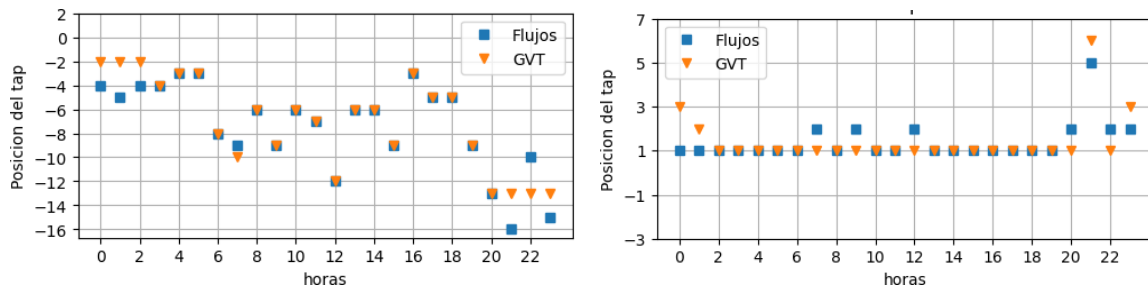
Hora	Recompensa GVT	Recompensa flujos	Tap 1 flujos	Tap 2 flujos	Tap 1 GVT	Tap 2 GVT
1	-0.00409	-0.00361	-6	1	-5	2
2	-0.00272	-0.00257	-4	1	-4	0
3	-0.00239	-0.00232	-3	1	-4	0
4	-0.00231	-0.00224	-3	1	-3	1
5	-0.00249	-0.00245	-4	1	-4	1
6	-0.00256	-0.00253	-4	1	-4	1
7	-0.00391	-0.00390	-6	1	-6	1
8	-0.00649	-0.00640	-10	2	-11	1
9	-0.00429	-0.00423	-7	1	-7	1
10	-0.00416	-0.00413	-7	1	-6	1
11	-0.00368	-0.00356	-6	1	-6	1
12	-0.00402	-0.00402	-7	1	-7	1
13	-0.00463	-0.00461	-7	1	-8	1
14	-0.00376	-0.00376	-6	1	-6	1
15	-0.00349	-0.00349	-6	1	-6	1
16	-0.00305	-0.00305	-5	1	-5	1
17	-0.00360	-0.00360	-6	1	-6	1
18	-0.00475	-0.00475	-8	1	-8	1
19	-0.00355	-0.00355	-6	1	-6	1
20	-0.00616	-0.00606	-10	2	-11	1
21	-0.00778	-0.00751	-13	2	-12	1
22	-0.01023	-0.00989	-15	3	-15	2
23	-0.00801	-0.00790	-13	2	-13	2
24	-0.00730	-0.00674	-11	2	-10	3

Con relación a los resultados de la política de exploración, el historial creado del entrenamiento 1 entregó los siguientes datos máximos y mínimos de las posiciones de los *taps* LTC1 (-16,0) y LTC2 (0,4).

Los resultados proporcionaron una política de exploración para el entrenamiento 2 de LTC1 (-16,4) y LTC2 (-4,8) para GVT (Dimensión de exploración de 273 estados). Lo que resultó de aplicar esta política, fueron unos resultados similares al entrenamiento 1. Las posiciones se mantuvieron en la ventana de estados establecida por la política. Un ejemplo se muestra en la Figura 13.

Figura 13

Posiciones de los taps día 12 del sistema de 9 nodos



Nota: La figura de la izquierda corresponde al LTC1 y la de la derecha al LTC2.

Los resultados del entrenamiento 1 y 2 proporcionaron una política de exploración para la prueba de LTC1 (-16,2) y LTC2 (-2,6) (Dimensión de exploración de 171 estados). Para resumir los resultados del entrenamiento, se presentó la Tabla 2. Al finalizar la prueba se rectificó que el agente mejoró su desempeño, disminuyendo la regulación de tensión para ambas técnicas de estimación, por lo cual estaba listo para entrar en la operación del sistema.

Tabla 2

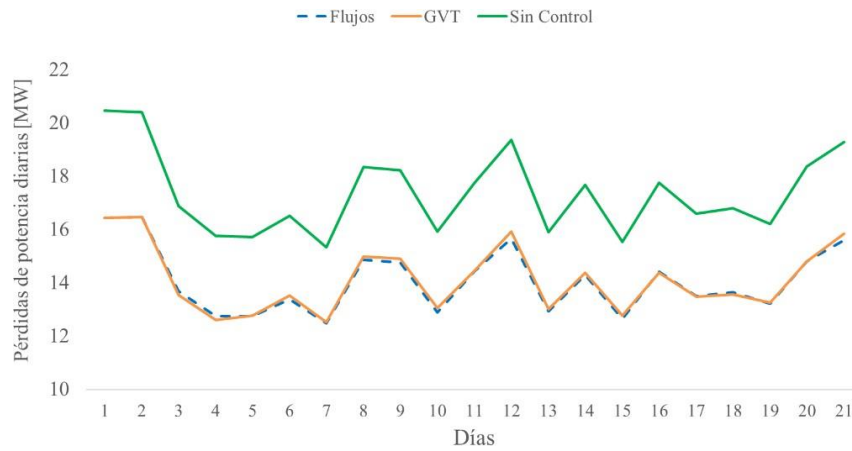
Recompensa acumulada y resultados de la regulación en la barra 3.

Día	Recompensa acumulada GVT	Recompensa acumulada flujos	RVmax Valor	RVmax Metodo
1	-0.00484	-0.0047	2.71%	GVT
2	-0.00476	-0.00471	2.46%	GVT
3	-0.00461	-0.00449	2.52%	GVT
4	-0.00429	-0.00421	2.80%	GVT
5	-0.00451	-0.00442	1.79%	Flujos
6	-0.00455	-0.00441	1.93%	GVT
7	-0.00434	-0.00422	1.82%	GVT
8	-0.00486	-0.00467	1.71%	Flujos
9	-0.00476	-0.00461	2.67%	GVT
10	-0.00454	-0.00438	1.79%	Flujos
11	-0.00465	-0.00459	1.90%	GVT
12	-0.00492	-0.00467	3.68%	GVT
13	-0.00454	-0.00436	1.84%	Flujos
14	-0.00454	-0.00438	2.78%	GVT
15	-0.00443	-0.00428	1.82%	GVT
16	-0.00458	-0.00448	2.32%	GVT
17	-0.00454	-0.00452	1.79%	GVT
18	-0.0045	-0.00443	2.09%	GVT
19	-0.00447	-0.00442	1.72%	GVT
20	-0.00472	-0.00462	2.40%	GVT
21	-0.00488	-0.00472	3.22%	GVT

Se apreció que la barra 3 registró la mayor desviación con respecto a la referencia. Esto se explica por la presencia de un generador y un transformador a un nodo de distancia, lo que representa una sobretensión debido a la influencia de las dos ramas que inyectan potencia en el nodo. Es importante destacar que, a pesar de esto, la barra 3 se mantiene dentro de los límites establecidos. La recompensa acumulada de la tabla corresponde a la suma del valor de la recompensa obtenida tras realizar la acción en cada hora del día, como se puede observar entre ambas técnicas de estimación, la acción realizada por el agente es eficaz y a su vez son cercanas entre sí.

Figura 14

Pérdidas de potencia del sistema de 9 nodos para 21 días



Para el sistema de 9 nodos, como se muestra en la Figura 14, se obtuvo una mejora significativa en la disminución de pérdidas eléctricas alrededor del 19% al aplicar control sobre el sistema, mejorando las pérdidas de potencia activa de la red con respecto a la solución sin control.

5.3 Sistema de 34 nodos

Inicialmente, se ingresaron los datos planteados al inicio del capítulo con la topología de la sección 2.3, tomando como nodos de interés los encontrados luego del primer LTC. Esto con el fin de constatar cómo se comportaba el agente y la recompensa al realizar esta modificación en el sistema, verificando que lo planteado en el sistema de 9 nodos es aplicable a otros sistemas. Los resultados del entrenamiento 1 se encuentran a continuación.

Figura 15

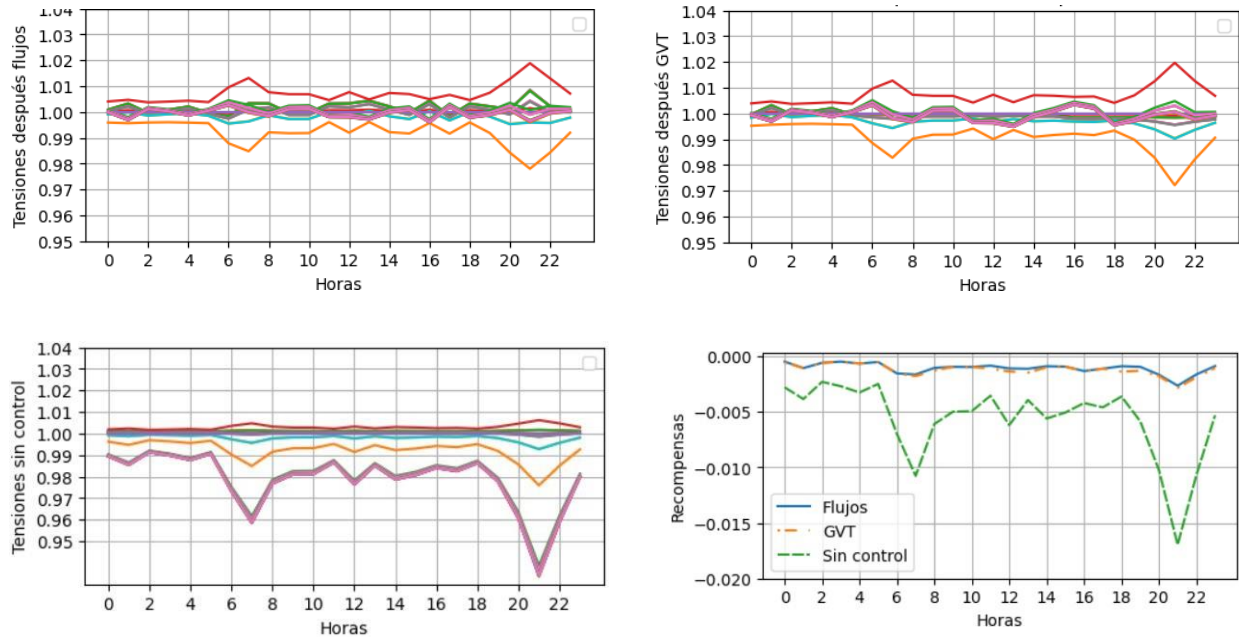
Nomenclatura de las tensiones en las barras para el sistema de 34 nodos



Nota. Esta nomenclatura aplica para todas las gráficas de tensiones correspondientes al Sistema de 34 nodos.

Figura 16

Resultados de tensiones en los nodos para cada escenario y comparación de la recompensa.



Nota. Las gráficas corresponden a los valores de tensiones después de la aplicación de la política de acción por el agente.

Como se puede apreciar en la Figura 16, una cantidad determinada de nodos experimentó un impacto significativo por el control, mientras que, en otros casos, la intervención de este no genera mejoras sustanciales, esto debido a su proximidad con generadores o LTC. Así pues, la recompensa aumentó considerablemente al aplicar el control y a su vez se redujo la regulación máxima, pasando de un 5% en el escenario sin control en distintas barras, a una menor al 3%, haciendo uso de las técnicas de estimación por flujos o GVT. Por último, los datos máximos y mínimos de las posiciones de los taps de GVT utilizados para la política de exploración son LTC1 (-3,4) y LTC2 (-0,11).

Tabla 3

Recompensa promedio y posiciones de taps del entrenamiento 1.

Hora	Recompensa GVT	Recompensa flujos	Tap 1 flujos	Tap 2 flujos	Tap 1 GVT	Tap 2 GVT
1	-0.00101	-0.00092	0	1	0	1
2	-0.00063	-0.00062	0	1	0	1
3	-0.00069	-0.00069	0	1	0	1
4	-0.00059	-0.00059	0	1	0	1
5	-0.00076	-0.00075	-1	1	0	1
6	-0.00073	-0.00070	0	1	0	1
7	-0.00115	-0.00111	0	2	0	2
8	-0.00151	-0.00137	0	3	0	3
9	-0.00126	-0.00112	1	2	-1	2
10	-0.00112	-0.00112	1	2	1	2
11	-0.00097	-0.00092	1	1	0	2
12	-0.00111	-0.00097	1	1	0	1
13	-0.00141	-0.00120	2	2	0	3
14	-0.00104	-0.00089	1	1	0	1
15	-0.00103	-0.00097	0	2	0	2
16	-0.00095	-0.00088	1	2	1	2
17	-0.00114	-0.00097	0	2	0	2
18	-0.00120	-0.00105	0	2	1	2
19	-0.00083	-0.00070	0	1	0	1
20	-0.00153	-0.00136	2	3	0	3
21	-0.00161	-0.00152	3	3	1	3
22	-0.00183	-0.00177	2	4	0	4
23	-0.00181	-0.00168	2	4	0	4
24	-0.00126	-0.00119	0	3	0	3

Finalmente, en la tabla 3 se sintetizaron los valores obtenidos para el entrenamiento 1, presentando el promedio de las recompensas y posiciones de los taps para cada hora, en donde el LTC1 varia alrededor de cero y el LTC2 es un poco más volátil. Además, se resalta que la hora con recompensa más baja en ambas estimaciones fue la 22.

El entrenamiento 1 entregó una política de exploración de LTC1 (-7,8) y LTC2 (-4,11) para GVT (Dimensión de exploración de 256 estados). Para el caso del primer día del entrenamiento 2, se obtuvo un nuevo máximo en el LTC1, en donde se presentó la posición 7 que es la más alta por parte de flujos sin llegar a salirse de la ventana de exploración de GVT. Sin embargo, por GVT

nunca se presentó dicha posición y a pesar de no coincidir en las tensiones al aplicar el control, se consiguió un resultado casi idéntico al de flujos.

Por último, el test fue exitoso debido a que la recompensa y las tensiones resultaron considerablemente mejores. Por lo tanto, el agente estaría entrenado para operar correctamente el sistema de 34 nodos. Una compilación de los resultados se presentó en la tabla 4.

Tabla 4

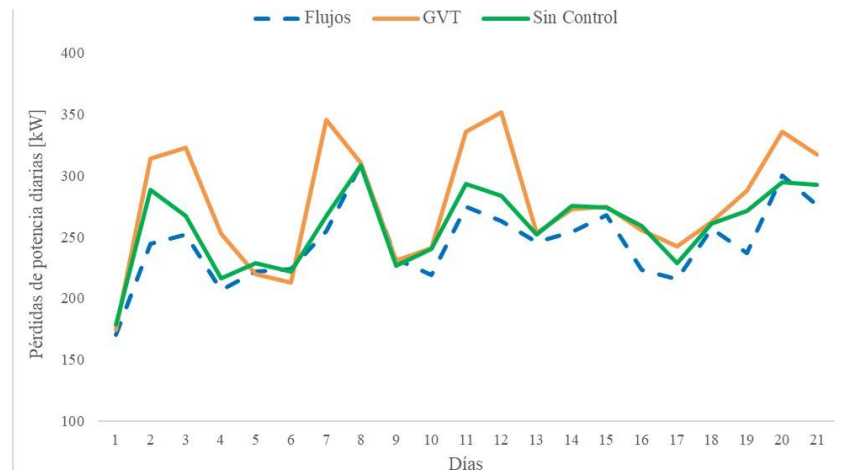
Recompensa acumulada y resultados de la regulación.

Día	Recompensa acumulada GVT	Recompensa acumulada flujos	RVmax Valor	RVmax Metodo
1	-0.000984	-0.000944	1.38%	GVT
2	-0.001287	-0.001164	2.14%	GVT
3	-0.001185	-0.001095	2.14%	GVT
4	-0.001142	-0.001020	2.02%	GVT
5	-0.001100	-0.001041	1.75%	GVT
6	-0.001035	-0.000985	2.26%	Flujos
7	-0.001189	-0.001061	2.79%	GVT
8	-0.001149	-0.001091	2.77%	GVT
9	-0.001115	-0.001066	2.05%	GVT
10	-0.001160	-0.001031	2.39%	GVT
11	-0.001196	-0.001093	1.95%	GVT
12	-0.001255	-0.001123	2.60%	GVT
13	-0.001115	-0.001012	2.47%	GVT
14	-0.001136	-0.001093	1.78%	GVT
15	-0.001178	-0.001074	2.58%	Flujos
16	-0.001195	-0.001107	1.96%	Flujos
17	-0.001152	-0.001048	1.90%	Flujos
18	-0.001157	-0.001088	2.88%	GVT
19	-0.001215	-0.001117	2.75%	GVT
20	-0.001231	-0.001123	3.24%	GVT
21	-0.001199	-0.001090	2.32%	GVT

De los resultados se exhibe que, a lo largo de los 21 días, el nodo 18 presenta la mayor desviación con respecto a la tensión de referencia. Esta situación puede justificarse considerando que este nodo es el más lejano del transformador, lo que provoca una mayor caída de tensión debido a las pérdidas en las líneas.

Figura 17

Pérdidas de potencia del sistema de 34 nodos para 21 días



En el caso del sistema de 34 nodos, se obtuvieron resultados variables en la reducción de pérdidas. Aunque se reflejaron mejoras en las simulaciones realizadas por flujos, la Figura 17 expone que el GVT no logra superar los resultados de la simulación “Sin Control”. Como era de esperar, el método de flujos generó una menor cantidad de pérdidas, al ser el método más riguroso de solución. Seguidamente, se encontraron las pérdidas sin control que son aproximadamente un 5% mayores que las pérdidas generadas por flujos, considerando el total de las pérdidas a lo largo de toda la simulación. En el caso del GVT, las pérdidas son aproximadamente 13% mayores a las halladas por flujos, pero obteniendo mejores perfiles de tensiones que en la simulación sin control.

5.4 Estudio de sensibilidad

El estudio de sensibilidad desempeñó un papel fundamental en el ámbito de la investigación y el análisis, con el fin de comprender cómo ciertos parámetros o variables influyen en los resultados de un modelo, experimento o sistema. Este tipo de estudio permitió evaluar la robustez y la respuesta del sistema ante cambios en las condiciones iniciales. En este estudio, se empleó el sistema de 34 nodos con un porcentaje de variación mayor en la probabilidad del día semilla,

empleando para este caso un 30%. Esta elección se basó en pruebas con distintos porcentajes, en donde este valor proporcionó el mejor comportamiento del agente ante las fluctuaciones de carga y generación. El escenario de prueba en las demás condiciones es idéntico a las dos secciones anteriores, condensando estos resultados en el Apéndice G.

Figura 18

Tensiones antes y después del control en el test

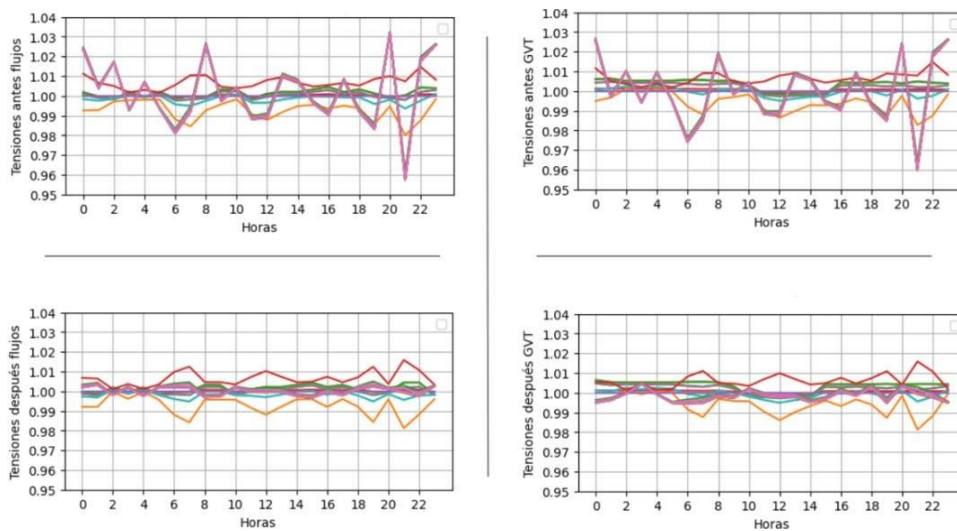
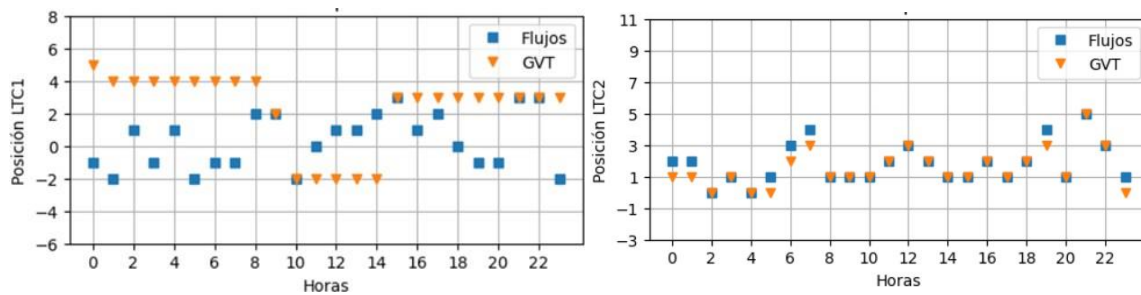


Figura 19

Posiciones de los LTC1 y LTC2 en el test



Como se evidencia en la Figura 18, el sistema tiende incluso a salirse de los valores de regulación establecidos por la norma previamente a que el agente realice la acción. Aunque la estimación por ambas técnicas no coincidió en la mayoría de los casos en las posiciones como se

observa en la Figura 19, se logró un control efectivo por parte del agente al hacer uso de ambos métodos, obteniendo resultados similares.

La recompensa obtenida por el agente a partir de las acciones realizadas es comparable a la obtenida con el escenario de la sección 5.3, como se muestra en la Tabla 5. Asimismo, se registró la máxima regulación en la barra 18 lo cual, en caso de tratarse de un problema real, podría ser un indicativo claro de la necesidad de futuras acciones de mejora en el sistema de control en esa barra y/o las circundantes.

Tabla 5

Recompensa acumulada y resultados de la regulación.

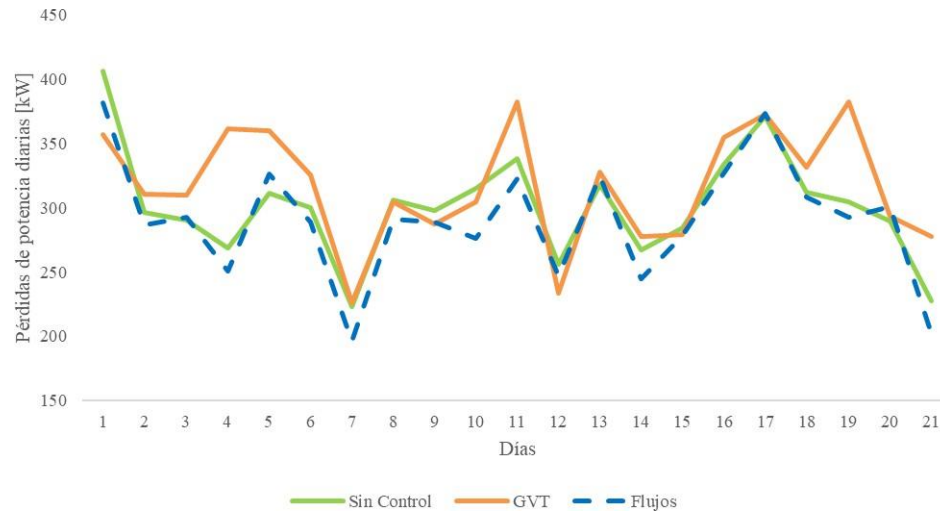
Día	Recompensa acumulada GVT	Recompensa acumulada flujos	RVmax Valor	RVmax Metodo
1	-0,00136	-0,00129	3,15%	GVT
2	-0,00121	-0,00109	2,64%	GVT
3	-0,00126	-0,00118	2,25%	GVT
4	-0,00130	-0,00110	1,90%	Flujos
5	-0,00129	-0,00115	2,29%	Flujos
6	-0,00120	-0,00112	2,28%	GVT
7	-0,00110	-0,00100	1,94%	GVT
8	-0,00132	-0,00118	2,25%	GVT
9	-0,00115	-0,00109	1,92%	GVT
10	-0,00130	-0,00119	2,32%	GVT
11	-0,00136	-0,00116	2,79%	GVT
12	-0,00108	-0,00102	2,60%	GVT
13	-0,00130	-0,00116	2,59%	GVT
14	-0,00130	-0,00115	2,54%	GVT
15	-0,00119	-0,00114	2,60%	GVT
16	-0,00126	-0,00121	2,35%	GVT
17	-0,00134	-0,00122	3,43%	GVT
18	-0,00126	-0,00117	1,93%	Flujos
19	-0,00141	-0,00109	2,75%	GVT
20	-0,00120	-0,00114	2,44%	GVT
21	-0,00135	-0,00106	1,87%	GVT

En el estudio de sensibilidad se obtuvieron pérdidas variables. Aunque la técnica de estimación por flujos logró disminuir las pérdidas en el sistema, los resultados nuevamente vuelven a ser volátiles y no definitivos con relación a la reducción de pérdidas. Como se constata en la

Figura 20, las pérdidas por GVT son mayores que las entregadas por flujos y sin control. No obstante, se destaca que los perfiles de tensión entregados por la técnica de estimación por flujos y GVT son los mejores.

Figura 20

Pérdidas de potencia del sistema de 34 nodos para 21 días



6. Conclusiones

A modo de conclusión, elaborar este proyecto abarcó temas relacionados a la convergencia de sistemas de potencia, balanceo de sistemas trifásicos y flujos de potencia; todos estos temas con bases teóricas sólidas y amplia bibliografía, los cuales, se relacionan en este trabajo de grado, con la aplicación de una estrategia de aprendizaje reforzado implementada con el fin de replicar una solución a un problema de distribución de la energía eléctrica, con el enfoque nuevo que brinda el aprendizaje de máquinas. La situación proporcionó una gama extensa de posibilidades en la automatización del control en sistemas de potencia reales que integran la matriz energética. Modelar un sistema de distribución resulta una tarea exhaustiva, debido a todas las variables que

tiene que entrar en sintonía para poder converger en una condición estable e inestable al momento de aplicar los cambios en los taps de los transformadores LTC de los sistemas estudiados. Este primer paso para realizar este trabajo de grado requirió además del uso de bases de datos de fuentes como el IEEE y PandaPower, la revisión de bibliografía para poder llevar a cabo balanceos de sistemas trifásicos, inyección de energía solar y comportamientos de cargas residenciales. Asimismo, la comprensión de las variables necesarias para que los sistemas pudieran ser modelados de acuerdo con cómo se explicó en los capítulos 2 y 3 de este proyecto.

Cabe destacar que, al aplicar una estrategia de aprendizaje reforzado, se adentra en un entramado de posibilidades, las cuales no son resultan convenientes para todos los tipos de problemas. Estas estrategias resultan fundamentales para obtener soluciones a desafíos cruciales que impactan directamente en la mejora de la calidad del servicio de energía eléctrica. En consecuencia, definir un proceso de entrenamiento para dar al agente una capacidad de predicción eficaz, también puede ser una labor para completarse de distintas formas, dependiendo el problema pueden existir caminos que resulten mejores en este caso de estudio. Por lo tanto, fue conveniente estudiar estos alcances para definir qué condiciones de entrenamiento podrían obtener los mejores resultados con enfoques sencillos, permitiendo que el agente sea generalizado y pueda tanto analizar cómo resolver distintos sistemas sin importar su topología o condiciones externas e internas que se requieran para funcionar.

En la mayoría de las simulaciones la regulación de tensión no superó el 3% y las mejores recompensas acumuladas, en torno al -0.002, se dieron en horas de la madrugada para ambos sistemas y ambas técnicas, esto se puede explicar por los bajos consumos y la nula generación fotovoltaica que pueda generar flujos de potencia adicionales que afecten la regulación de tensión.

Asimismo, se pudo observar que el LTC más alejado de la barra Slack cambia sus posiciones en un rango alrededor de 4.

Finalmente, resolver el sistema por las técnicas de flujos y GVT brindaron tensiones con una regulación dentro de la normativa (NTC 1340. Electrotecnia. Tensiones y frecuencia nominales en sistemas de energía eléctrica en redes de servicio público, 2004). Sin embargo, los valores de las tensiones antes de aplicar la estrategia también cumplen la norma, obteniendo como resultado una mejora de la regulación con respecto al sistema sin controlar en los distintos escenarios. Adicional a esto, el estudio y las pruebas realizadas referentes a la disminución de pérdidas no presentaron resultados que permitieran concluir que existieron mejoras al aplicar la técnica del GVT o flujos con certeza debido a la amplitud de datos, ya que, aunque para el sistema de 9 nodos obtuvimos mejoras de alrededor del 20% con ambos métodos, para el sistema de 34 nodos no se obtiene una mejora clara.

7. Recomendaciones y trabajos futuros

Recomendaciones:

- Validación Experimental: llevar a cabo experimentos en un entorno del mundo real para validar y verificar los resultados obtenidos en simulaciones, ya que esto proporcionaría una mayor confianza en la eficacia del enfoque propuesto.
- Optimización de Políticas de Control: explorar técnicas de optimización más avanzadas para mejorar las políticas de control del agente e incluir algoritmos de aprendizaje profundo o métodos de optimización más avanzados para ajustar las políticas de manera más efectiva.
- Estudio de Sensibilidad Avanzado: ampliar el estudio de sensibilidad para explorar cómo los resultados pueden variar en función de una amplia gama de parámetros y así ayudar a comprender mejor las limitaciones, oportunidades del sistema y el enfoque propuesto.

- Adaptación a Diferentes Topologías de Red: extender la investigación para que pueda adaptarse a una variedad de topologías de redes eléctricas, lo que permitiría aplicar el mismo enfoque a sistemas de diferentes tamaños y complejidades.
- Tiempo de Respuesta del Agente: continuar trabajando en la mejora del tiempo de respuesta del agente y explorar técnicas de paralelización o distribución para acelerar el proceso de toma de decisiones.

Trabajos Futuros:

- Aplicación en Redes Más Grandes: extender la aplicación del agente y el enfoque a redes eléctricas de mayor envergadura, como sistemas de transmisión de energía a mayor escala.
- Integración de Energía Renovable: considerar cómo el agente puede adaptarse para trabajar en sistemas eléctricos con una mayor participación de fuentes de energía renovable, lo que presenta desafíos adicionales en términos de regulación y control.
- Tecnologías de Medición Avanzadas: investigar cómo la integración de tecnologías de medición avanzadas, como medidores inteligentes, puede mejorar la precisión y eficiencia del control en sistemas eléctricos.
- Optimización del Consumo de Energía: ampliar el enfoque para incluir la optimización del consumo de energía por parte de los usuarios finales, para así contribuir a la gestión más eficiente de la carga en la red eléctrica.
- Seguridad Cibernética: considerar la seguridad cibernética en la gestión y control de sistemas eléctricos, especialmente en un mundo cada vez más conectado e investigar cómo proteger estos sistemas de amenazas y ataques cibernéticos.
- Aplicaciones en Redes Inteligentes (Smart Grids): explorar cómo el agente y la estrategia de control pueden aplicarse en redes eléctricas inteligentes para una gestión más eficiente.
- Desarrollo de Interfaces de Usuario: desarrollar interfaces de usuario amigables para que los operadores de sistemas eléctricos puedan interactuar y supervisar el agente y el control de la red de manera efectiva.
- Estudio de Costos y Beneficios: realizar un análisis de costos y beneficios para evaluar la viabilidad económica de la implementación de este enfoque en sistemas eléctricos reales.

Este trabajo de grado se desarrolló bajo la dirección del ingeniero electricista Oscar Alexis Galvis Diaz, estudiante de Maestría en Ingeniería Electrónica, como parte de su investigación, la cual presentará una continuación y profundización del uso del aprendizaje reforzado para la regulación de tensión en los sistemas de distribución con transformadores LTC.

Referencias Bibliográficas

- Anderson, P. M. (1995). *Analysis of Faulted Power Systems* (P. M. Anderson, Ed.).
- Ley 142 de 1994, Pub. L. No. 142, Por la cual se establece el régimen de los servicios públicos domiciliarios y se dictan otras disposiciones. 1 (1994).
- Ernst, D., Glavic, M., & Wehenkel, L. (2004a). Power Systems Stability Control: Reinforcement Learning Framework. *IEEE Transactions on Power Systems*, 19(1), 427–435. <https://doi.org/10.1109/TPWRS.2003.821457>
- Ernst, D., Glavic, M., & Wehenkel, L. (2004b). Power Systems Stability Control: Reinforcement Learning Framework. *IEEE Transactions on Power Systems*, 19(1), 427–435. <https://doi.org/10.1109/TPWRS.2003.821457>
- Norma ESSA, Pub. L. No. 3, 11 (2020).
- Gao, C., & Redfern, M. A. (2010). A review of voltage control techniques of networks with distributed generations using On-Load Tap Changer transformers. *45th International Universities Power Engineering Conference UPEC*, 1–6.
- Gao, C., & Redfern, M. A. (2011). Advanced Voltage Control Strategy for On-Load Tap-Changer Transformers with Distributed Generations. *46th International Universities' Power Engineering Conference (UPEC)*, 1–6.
- Google. (2023). *Colaboratory*.
- Grainger, J., & Stevenson, W. (2001). *Análisis de sistemas de potencia* (Vol. 1). McGraw-Hill.
- Hernández, K., & Carrillo, J. (2020). *Análisis de la curva de demanda eléctrica para usuarios residenciales estrato 4 en la ciudad de Bogotá ante diferentes escenarios de los hábitos de consumo*.
- NTC 1340. Electrotecnia. Tensiones y frecuencia nominales en sistemas de energía eléctrica en redes de servicio público, (2004).
- IEEE. (2017). *IEEE PES Test Feeder*.
- Juarez, J. D. (1995). *Sistemas de distribución de energía eléctrica* (Primera).
- Kellerman, W., El-Din, H. M., Graham, C. E., & Maria, G. A. (1991). Optimization of fixed tap transformer setting in bulk electric systems. *IEEE Transactions on Power*

Systems, 6(3), 1126–1132. <https://doi.org/10.1109/59.119256>

Kersting, W. H. (2002). *Distribution System Modeling and Analysis* (W. H. Kersting, Ed.; 1st ed.). CRC Press.

Martinez, A. (2014). *Método de las componentes simétricas* (1a ed.). Universidad Tecnológica Nacional.

Moya, O. (2014). El lenguaje de programación Python. *Redalyc*.

NASA - National Aeronautics and Space Administration. (2021, agosto 5). *NASA Power*. US Government.

PandaPower. (2023). *About PandaPower*.

Sutton, R., & Barto, A. (2015). *Reinforcement Learning: An Introduction* (2a ed.). Bradford.

Xu, H., Dominguez-Garcia, A. D., & Sauer, P. W. (2020). Optimal Tap Setting of Voltage Regulation Transformers Using Batch Reinforcement Learning. *IEEE Transactions on Power Systems*, 35(3), 1990–2001. <https://doi.org/10.1109/TPWRS.2019.2948132>