



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

**METODOLOGÍA DE AGREGACIÓN TEMPORAL PARA LA SIMULACIÓN
DE LA OPERACIÓN UTILIZADA EN EL PROCESO DE PLANIFICACIÓN
DE LA TRANSMISIÓN REALIZADA POR LA COMISIÓN NACIONAL DE
ENERGÍA (CNE)**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL ELÉCTRICA

FERNANDA SOFÍA ATALA HERMOSILLA

PROFESOR GUÍA:
EDUARDO PEREIRA BONVALLET

MIEMBROS DE LA COMISIÓN:
RODRIGO MORENO VIEYRA
PATRICIO MENDOZA ARAYA

SANTIAGO DE CHILE
2023

METODOLOGÍA DE AGREGACIÓN TEMPORAL PARA LA SIMULACIÓN DE LA OPERACIÓN UTILIZADA EN EL PROCESO DE PLANIFICACIÓN DE LA TRANSMISIÓN REALIZADA POR LA COMISIÓN NACIONAL DE ENERGÍA (CNE)

En la actualidad, uno de los grandes objetivos a nivel mundial es la inclusión de una mayor cantidad de fuentes renovables a la matriz energética. En particular, para cumplir con este objetivo, en Chile se está trabajando en el “Plan de Retiro y/o Reconversión de Unidades a Carbón” desde el año 2019, con el fin de alcanzar la carbono neutralidad al año 2050. Uno de los procesos importantes para conseguir este objetivo es la planificación de las redes de transmisión, la cual es realizada de manera anual por la Comisión Nacional de Energía. Luego, la motivación del presente trabajo es investigar la representación de la variabilidad temporal que se presenta en este proceso, la cual implica una alta carga computacional, con el fin de representarla de manera agregada sin distorsionar los resultados del modelo.

En este trabajo, el objetivo es encontrar una metodología para definir una mejor representación temporal de la generación eólica y solar del Sistema Eléctrico Nacional respecto a la actual utilizada por la CNE. Cabe mencionar que esta metodología consiste en definir un día hábil y un día no hábil representativo para cada mes que se considere en la simulación.

Para esto, se aplican 3 métodos de *clustering* a la generación separada por tecnología, mes y con o sin división geográfica, para luego comparar los métodos de *clustering* y decidir cuál de los 3 métodos es mejor para el objetivo en cuestión, y escoger si se aplicará o no la división geográfica de los datos de generación. Además, al analizar el ajuste de los días representativos a los datos reales de los meses, y ver la correlación entre zonas, se escoge la cantidad de días representativos necesarios a utilizar en la simulación de la operación para la planificación de la transmisión.

Como resultado, se obtiene que el mejor método de *clustering* con este fin es el método *k-medoids*, aplicando la división geográfica por subestación asociada, y considerando al menos 2 días representativos por mes, tanto para la generación eólica como la solar. De esta manera se logran ajustar de mejor forma los días representativos a los datos originales en comparación con la metodología actual, además de mantener la correlación entre zonas.

Además, al realizar la simulación de la operación en Ameba con la nueva representación temporal obtenida y analizar distintos indicadores, se tiene un buen ajuste al aplicar el método *k-medoids* en cuanto a la generación renovable. Sin embargo, no se puede decir lo mismo de la generación convencional, induciendo diferencias en el costo de operación del sistema respecto a la realidad esperada. Finalmente, es importante mencionar que, si bien el tiempo de simulación requerido es mayor al actual de la CNE, de igual manera se logra reducir en más de un 90% el tiempo de simulación respecto al caso horario.

A mis padres, hermanos, pololo y amigos, por su incondicional apoyo y amor durante todo este camino. Sin ustedes, este logro no sería posible.

Gracias.

Agradecimientos

Agradecer a todas las personas que han sido fundamentales en mi camino hacia la culminación de mi carrera universitaria.

En primer lugar, a mis padres por su apoyo y amor incondicional. Gracias por su dedicación, paciencia y confianza en mí. Han sido mi mayor motivación en todo momento y sin duda alguna, han sido parte fundamental en mi éxito.

A Matías, gracias por su apoyo y amor durante toda la carrera universitaria. Gracias por estar siempre a mi lado, por escucharme y motivarme. Sin ti, este logro no habría sido posible.

A mis amigos, Paula, Janus, Vicky, Myra, Dani, Ule, gracias por su gran apoyo y por estar ahí en los momentos más difíciles. Gracias por ser mi escape cuando necesitaba un respiro y por escucharme cuando necesitaba consejos. No puedo expresar lo agradecida que estoy por tenerlos a mi lado.

A mis hermanos, Cony, Salmi, Luis, gracias por siempre estar conmigo en cada paso que he dado. Su apoyo y amor han sido una fuente constante de motivación y me han dado fuerzas para seguir adelante.

Por último, pero no menos importante, agradezco a mis compañeros de voleibol. Gracias por su amistad, apoyo y por haberme enseñado el valor del trabajo en equipo y la disciplina.

Sin el apoyo de todas estas personas, no habría logrado este importante hito en mi vida. Una vez más, gracias por estar ahí en cada paso del camino.

Tabla de Contenido

1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	2
1.2.1. Objetivo general	2
1.2.2. Objetivos específicos	2
1.3. Estructura del trabajo	2
2. Marco teórico y estado del arte	3
2.1. Marco teórico	3
2.1.1. Modelos de despacho eléctrico y operación del sistema eléctrico	3
2.1.2. Modelos de planificación de expansión de la transmisión	5
2.1.3. Métodos de <i>clustering</i>	6
2.1.3.1. <i>Clustering</i> jerárquico aglomerativo	7
2.1.3.2. <i>Clustering k-means</i>	8
2.1.3.3. <i>Clustering k-medoids</i>	8
2.1.4. Métrica de error RMSE	9
2.1.5. Plataforma de simulación Ameba	10
2.2. Estado del arte	11
3. Antecedentes	15
3.1. Descripción del caso de estudio	15
3.2. Método de agregación temporal de la CNE	18
3.3. Adaptación de los días representativos hábil y no hábil de generación renovable a los datos horarios	19
3.4. Correlación entre zonas	21
4. Metodología y desarrollo	23
4.1. Esquema metodológico	23
4.2. Desarrollo de la propuesta metodológica	25
5. Resultados y análisis	27
5.1. Elección del método de <i>clustering</i>	27
5.2. División geográfica por subestación	28
5.3. Cantidad de días representativos por mes	29
5.3.1. Correlación entre zonas	32
5.4. Simulación de la operación en Ameba	36
5.4.1. Costo de operación del sistema	37
5.4.2. Costos marginales	37

5.4.3.	Generación por tecnología	38
5.4.3.1.	Generación a carbón, GNL y diésel	39
5.4.3.2.	Generación eólica	40
5.4.4.	Corrección del método de <i>clustering k-medoids</i> basado en series de tiempo	43
5.4.5.	Tiempo de simulación	46
6.	Conclusiones y trabajos futuros	47
6.1.	Conclusiones generales	47
6.2.	Trabajos futuros	49
	Bibliografía	50
	Anexo	53

Índice de Tablas

2.1.	Clasificación de Modelos de Despacho Eléctrico obtenida de [9].	3
2.2.	Características de la literatura revisada.	12
3.1.	Proyección de la potencia instalada y generación eólica por subestación a los años 2025, 2030, 2035 y 2040, en conjunto con el número de centrales asociadas a cada una de ellas.	16
3.2.	Proyección de la potencia instalada y generación solar fotovoltaica por subestación a los años 2025, 2030, 2035 y 2040, en conjunto con el número de centrales asociadas a cada una de ellas.	17
3.3.	Asignación de bloques y etapas según la CNE.	18
5.1.	Rango de días representativos necesarios por mes tanto para la generación eólica como la solar, para cumplir con los criterios definidos.	30
5.2.	Días representativos necesarios por mes, tanto para la generación eólica como la solar, para que al menos 24 de los 48 meses cumplan con los criterios definidos (el porcentaje de meses que cumplen con el criterio queda explicitado entre paréntesis).	31
5.3.	Tiempo de simulación de la operación en la plataforma web Ameba para el caso horario real y las metodologías de agregación temporal, además de la reducción del tiempo de simulación para estas metodologías en comparación con el caso real.	46
A.1.	Potencia instalada por año de las centrales eólicas, junto a la subestación asociada.	53
A.2.	Potencia instalada por año de las centrales solares fotovoltaicas, junto a la subestación asociada.	56

Índice de Ilustraciones

3.1.	Mapa del SEN donde se muestran las S/E consideradas, en conjunto con la tecnología asociada.	15
3.2.	Capacidad anual de generación prevista para cada subestación asociada, tanto para las centrales eólicas como las solares fotovoltaicas.	17
3.3.	Adaptación de los días representativos hábil y no hábil de la generación eólica y solar a los datos reales utilizando el método actual de la CNE, medida con el indicador RMSE.	19
3.4.	Adaptación de los días representativos hábil y no hábil de la CNE a los datos horarios reales de generación eólica para la subestación Parinas (enero y junio 2040).	20
3.5.	Adaptación de los días representativos hábil y no hábil de la CNE a los datos horarios reales de generación eólica para la subestación Río Malleco (enero y junio 2040).	20
3.6.	Adaptación de los días representativos hábil y no hábil de la CNE a los datos reales de generación solar fotovoltaica para la subestación Kimal (enero y junio 2040).	21
3.7.	Adaptación de los días representativos hábil y no hábil de la CNE a los datos reales de generación solar fotovoltaica para la subestación Polpaico (enero y junio 2040).	21
3.8.	Matriz de correlación entre zonas de los datos horarios de 2040 versus la correlación entre zonas utilizando la metodología de la CNE, para los datos de viento.	22
3.9.	Matriz de correlación entre zonas de los datos horarios de 2040 versus la correlación entre zonas utilizando la metodología de la CNE, para los datos de sol.	22
4.1.	Diagrama de bloques para la metodología propuesta.	23
5.1.	RMSE anual de la generación eólica y solar fotovoltaica según el número de clústeres, año 2040	27
5.2.	RMSE anual de la generación eólica y solar fotovoltaica según el número de clústeres, para el método de <i>clustering k-medoids</i> , año 2040.	28
5.3.	Adaptación de los días representativos con y sin división por subestación, aplicando el método <i>k-medoids</i> a los datos reales de generación eólica para las subestaciones Parinas y Río Malleco, y de generación solar fotovoltaica para las subestaciones Kimal y Polpaico (enero 2040).	29
5.4.	RMSE anual de la generación eólica según el número de clústeres, para el método de <i>clustering k-medoids</i> , para enero y mayo del año 2040, aplicando distintos requerimientos de error.	30

5.5.	RMSE anual de la generación solar fotovoltaica según el número de clústeres, para el método de <i>clustering k-medoids</i> , para enero y mayo del año 2040, aplicando distintos requerimientos de error.	31
5.6.	Matriz de correlación entre zonas de los datos horarios de 2040 versus la correlación entre zonas utilizando la metodología de la CNE y el método de <i>clustering k-medoids</i> considerando dos y tres días representativos al mes, para los datos de viento.	32
5.7.	Matriz de correlación entre zonas de los datos horarios de 2040 versus la correlación entre zonas utilizando la metodología de la CNE y el método de <i>clustering k-medoids</i> considerando uno y dos días representativos al mes, para los datos de sol.	33
5.8.	Adaptación de los días representativos de la generación eólica obtenidos con el <i>clustering k-medoids</i> , en contraste con los días hábil y no hábil de la CNE, para la subestación Parinas (enero y junio 2040).	34
5.9.	Adaptación de los días representativos obtenidos con el <i>clustering k-medoids</i> a los datos reales de generación eólica, en contraste con los días hábil y no hábil de la CNE, para la subestación Río Malleco (enero y junio 2040).	34
5.10.	Adaptación de los días representativos obtenidos con el <i>clustering k-medoids</i> a los datos reales de generación solar fotovoltaica, en contraste con los días hábil y no hábil de la CNE, para la subestación Kimal (enero y junio 2040).	35
5.11.	Adaptación de los días representativos obtenidos con el <i>clustering k-medoids</i> a los datos reales de generación solar fotovoltaica, en contraste con los días hábil y no hábil de la CNE, para la subestación Polpaico (enero y junio 2040).	35
5.12.	Generación total anual del sistema.	36
5.13.	Costo de operación anual del sistema considerando los escenarios de hidrología seca y húmeda, para todas las metodologías analizadas.	37
5.14.	Costos marginales por barra, considerando el escenario de hidrología seca, para los datos reales no agregados.	38
5.15.	Costos marginales por barra para las distintas metodologías de agregación temporal versus el caso horario, considerando el escenario de hidrología seca para el año 2040.	38
5.16.	Generación por tecnología para el año 2040, considerando el escenario de hidrología seca, para todas las metodologías analizadas.	38
5.17.	Generación anual convencional a carbón, GNL y diésel considerando el escenario de hidrología seca, para todas las metodologías analizadas.	39
5.18.	Generación convencional mensual para el año 2040, considerando el escenario de hidrología seca, para todas las metodologías analizadas.	40
5.19.	Generación eólica anual considerando los escenarios de hidrología seca y húmeda, para todas las metodologías analizadas.	40
5.20.	Generación ERNC mensual para el año 2040, considerando el escenario de hidrología seca, para todas las metodologías analizadas.	41
5.21.	Generación eólica anual considerando los escenarios de hidrología seca y húmeda, para todas las metodologías analizadas.	41
5.22.	Costo de operación versus generación eólica diaria para el año 2040, considerando el escenario de hidrología seca, para los datos reales no agregados.	42

5.23.	Costo de operación versus generación eólica diaria para marzo y agosto del 2040, sin y con agregación temporal utilizando el método <i>k-medoids</i> , considerando el escenario de hidrología seca.	42
5.24.	Costo de operación versus generación eólica diaria para marzo y agosto del 2040, con y sin agregación temporal utilizando los métodos propuestos, considerando el escenario de hidrología seca.	44
5.25.	Costo de operación anual del sistema considerando los escenarios de hidrología seca y húmeda, para todas las metodologías analizadas (incluyendo las propuestas).	44
5.26.	Generación anual convencional a carbón, GNL y diésel considerando el escenario de hidrología seca, para todas las metodologías analizadas (incluyendo las propuestas).	45
5.27.	Generación anual convencional a carbón, GNL y diésel considerando el escenario de hidrología húmeda, para todas las metodologías analizadas (incluyendo las propuestas).	45

Capítulo 1

Introducción

1.1. Motivación

Actualmente, una de las grandes prioridades en todo el mundo es incluir una mayor cantidad de fuentes renovables -como la eólica y la solar- a la matriz energética, con el fin de obtener un sistema sostenible en el tiempo [1]. Así lo deja en evidencia la Agencia Internacional de la Energía (IEA) en su base de datos “Explorador de Datos de Energías Renovables 2021” [2], donde se observa que en los últimos 10 años, la generación eléctrica a partir de fuentes renovables a nivel mundial, aumentó un 84 %, y se prevé un aumento del 30 % en los próximos 4 años.

En particular, para cumplir con este objetivo, en Chile desde el año 2019 se está trabajando en el “Plan de Retiro y/o Reconversión de Unidades a Carbón”, cuyo fin principal es alcanzar la carbono neutralidad al año 2050, fijando el cese de la generación a carbón antes del 2040 [3]. Es más, hoy en día la Política Energética Nacional [4] está buscando generar los espacios que permitan retirar y/o reconvertir totalmente las centrales a carbón al año 2030.

Uno de los procesos importantes para conseguir este objetivo corresponde a la planificación de las redes de transmisión, las cuales se encargan del suministro de electricidad desde los puntos de generación hacia los puntos de consumo. En Chile, este proceso es realizado anualmente por la Comisión Nacional de Energía (CNE), la cual utiliza un horizonte de al menos 20 años, donde se consideran los datos de demanda eléctrica, proyección de precios, costo de tecnologías, entre otros, provenientes de la Planificación Energética de Largo Plazo (PELP), realizada por el Ministerio de Energía cada 5 años [5], la cual entrega los resultados con una resolución horaria.

Luego, la motivación de este trabajo corresponde a investigar la representación de la variabilidad temporal que se presenta en la simulación de la operación, empleada en el modelo de planificación de las redes de transmisión a largo plazo, la cual puede estar presente tanto del lado de la demanda como del lado de la generación renovable. Esto, dado que la variabilidad temporal, en conjunto con otras variables, provoca que exista una alta carga computacional al momento de resolver el modelo, complejizando y haciendo más lento el proceso. De aquí nace la importancia de representar la variabilidad temporal de manera agregada, teniendo en consideración que las señales o resultados que entrega el modelo como, por ejemplo, la visión económica, no se vean distorsionadas en gran medida.

1.2. Objetivos

1.2.1. Objetivo general

El objetivo general de este trabajo corresponde a encontrar una metodología que permita definir una mejor representación temporal de los datos respecto a la que actualmente utiliza la Comisión Nacional de Energía (CNE) para la simulación de la operación utilizada en el proceso de planificación de las redes de transmisión, de tal manera de aminorar la brecha al comparar los resultados de ambas metodologías, considerando distintos indicadores, tales como el costo de operación del sistema, la generación a partir de fuentes renovables y la generación convencional, entre otros.

1.2.2. Objetivos específicos

En particular, los objetivos específicos que se buscan cumplir son los siguientes:

1. Encontrar una representación temporal que, bajo ciertas métricas de error, se adapte de mejor manera a los datos horarios que la representación utilizada por la CNE.
2. Simular la operación en Ameba bajo los mismos supuestos y base de datos utilizados actualmente por la CNE.
3. Simular la operación en Ameba utilizando la nueva representación temporal.
4. Definir los indicadores que se considerarán para medir la brecha entre los resultados de las distintas simulaciones (ejemplo: costo de operación, vertimientos, congestiones de transmisión, etc.), además de establecer cómo se definirá tal brecha.
5. Medir la brecha entre ambas simulaciones utilizando los indicadores definidos.
6. Analizar el *trade-off* existente entre la carga computacional versus el resultado obtenido.

1.3. Estructura del trabajo

El presente trabajo contiene 6 capítulos, los cuales se describen brevemente a continuación:

- El capítulo 1 corresponde a la introducción, donde se presenta tanto la motivación como los objetivos del trabajo realizado.
- El capítulo 2 contiene el marco teórico y estado del arte, lo que le permite al lector entender el contexto del trabajo, al brindarle la información necesaria para ello.
- El capítulo 3 entrega los antecedentes utilizados en este trabajo de título.
- El capítulo 4 presenta la metodología propuesta y su desarrollo, donde se describe, paso a paso, el procedimiento realizado.
- El capítulo 5 contiene los resultados obtenidos, donde se analizan las diferentes etapas desarrolladas (tanto la etapa previa a la simulación como la post simulación).
- Finalmente, en el capítulo 6, se exponen las principales conclusiones del trabajo realizado, y el trabajo a futuro de éste.

Capítulo 2

Marco teórico y estado del arte

2.1. Marco teórico

2.1.1. Modelos de despacho eléctrico y operación del sistema eléctrico

Los modelos de despacho eléctrico corresponden a algoritmos de optimización mediante los cuales se pretende obtener el despacho eléctrico óptimo para alguna característica como, por ejemplo, a mínimo costo, con mínimas emisiones, etc. Usualmente, su formulación consta de una función objetivo y las restricciones que acotan las posibles soluciones del sistema.

Dependiendo del objetivo, las restricciones, el horizonte, el método de solución o del sistema a utilizar, los modelos de despacho eléctrico se pueden clasificar de diversas maneras [6–8], tal como se observa en la Tabla 2.1:

Tabla 2.1: Clasificación de Modelos de Despacho Eléctrico obtenida de [9].

Clasificación	Ejemplo
Según Objetivo	Minimización de Costos, Emisiones, Pérdidas
	Maximización de Utilidades
	Optimización Multi-objetivo
Según Restricciones y Horizonte	Corto y Largo Plazo: Estáticos ¹
	Corto Plazo: Dinámicos ²
Según Formulación	Enfoques basados en Control Óptimo
	Enfoques Dinámicos Económicos
Según Método de Resolución	Métodos basados en Programación Matemática
	Métodos basados en Heurísticas
	Métodos Híbridos
Según Sistema a Optimizar	Sistemas Térmicos, Hidráulicos, Hidrotérmicos

En la actualidad, los modelos de despacho eléctrico utilizados en el sistema eléctrico chileno tienen como objetivo principal minimizar los costos variables de generación. El Sistema Eléctrico Nacional (SEN), al poseer grandes embalses, necesita contar con un modelo que sea capaz de representar de forma adecuada el uso de los mismos, los cuales se conocen como

¹ Sin restricciones de rampas de generación

² Con restricciones de rampas de generación

modelo de operación hidrotérmica. De esta manera, el Coordinador Eléctrico Nacional (CEN) emplea un modelo de Programación de Largo Plazo (PLP), el cual se basa en la Programación Dinámica Dual Estocástica (SDDP, por sus siglas en inglés) cuya ventaja principal consiste en reducir los tiempos de cómputo derivados de la gran extensión del problema, mediante la utilización de variables duales [10].

Por otro lado, desde el año 2001 la Comisión Nacional de Energía utiliza un modelo llamado OSE2000 para obtener la simulación de la operación del SEN, el cual también se basa en la SDDP [11]. Este modelo representa múltiples embalses, con una representación temporal mínima de cierto número de bloques por mes (usualmente cinco). La aleatoriedad hidrológica se representa a través de series históricas ingresadas por archivos. Se ignora la indisponibilidad de las centrales hidráulicas y, para el caso de las centrales térmicas, se considera una reducción de la capacidad. Los resultados del modelo son: la producción de centrales, la disponibilidad del sistema y de cada central, los costos de operación, el racionamiento del sistema, estadígrafos de los embalses y de los consumos de combustibles y los flujos de potencia en las líneas. Es a partir de todos estos resultados que se puede, entre otras cosas, obtener la planificación de la red de transmisión [12].

La simulación de la operación de la CNE considera, entre otros, los siguientes puntos respecto al modelamiento de la demanda y de las unidades solares y eólicas [13]:

- La demanda mensual se representa mediante 12 bloques de horas consecutivas para los días hábiles (de lunes a viernes) y 12 bloques para los días no hábiles (sábados, domingos y festivos). Cabe mencionar que cada uno de los bloques agrupa dos horas consecutivas dentro de cada tipo de día.
- En un día hábil, la duración total de los bloques es mayor que la duración de los bloques de un día no hábil, ya que en cada mes la cantidad de días hábiles es mayor que la de días no hábiles.
- En cada bloque se incluye la generación solar en forma horaria para todos los meses del año. Luego, los bloques se separan para los niveles de mayor demanda del sistema.
- Para determinar los perfiles de demanda por bloque para cada barra, se utiliza la información de retiros horarios en cada mes del año anterior, obteniendo así los promedios de demanda por bloque en cada nudo. Luego, estos valores se dividen por la demanda promedio en el mes, obteniéndose un factor correspondiente a cada bloque y mes para todas las barras de consumo.
- Respecto a los datos de radiación solar, se utilizan perfiles de generación tipo obtenidos del Explorador de Energía Solar de la Universidad de Chile, desarrollado para el Ministerio de Energía. Además, se consideran perfiles de generación de centrales existentes.

Además, la CNE define una metodología para la representación tanto de las centrales solares como de las centrales eólicas en el modelo de despacho económico. La metodología para las centrales solares se estructura en tres etapas: la determinación de perfiles solares referenciales para cada zona geográfica, la representación de perfiles solares en estructura de bloques y el desarrollo de perfil para tecnología de Concentración Solar de Potencia (CSP). Por otro lado, la metodología empleada para las centrales eólicas se divide en las siguientes

3 etapas: serie de tiempo del recurso primario, la transformación del recurso primario en potencia eléctrica, y la representación de la potencia eólica en bloques [13].

2.1.2. Modelos de planificación de expansión de la transmisión

Un modelo de planificación para expandir la transmisión, se formula como un problema de optimización que posee múltiples etapas y agentes, los cuales toman decisiones tanto de inversión como de operación en distintos momentos y con distintos objetivos, considerando incertidumbre. Este proceso se realiza de manera anual, con un horizonte de al menos 20 años [14].

En el caso chileno, para llevar a cabo la planificación de la transmisión, es necesario que la Comisión Nacional de Energía anticipe mediante estimaciones tanto las decisiones de crecimiento de la generación como las decisiones de operación del mercado eléctrico, lo cual convierte el proceso en uno bastante complejo.

Según el Artículo 87° de la Ley General de Servicios Eléctricos (LGSE) [5], el cual se encarga de regular este proceso anual, los antecedentes que se requieren son los siguientes: la PELP, el informe realizado por el Ministerio de Energía sobre variables ambientales y territoriales (VAT) y criterios de eficiencia energética y las propuestas de los promotores y del CEN. También, pueden requerirse otros documentos determinados por la CNE.

Por otro lado, respecto a los criterios considerados para la planificación según el mismo artículo, se encuentra la minimización de riesgos en el abastecimiento, en conjunto con la creación de condiciones que promuevan oferta y faciliten la competencia, propendiendo a un mercado eléctrico común (MEC). Además, es necesario contar con instalaciones económicamente eficientes y necesarias para el desarrollo del sistema eléctrico. De igual manera se debe estar abiertos a una posible modificación de instalaciones existentes.

De esta manera, los objetivos de la planificación de expansión de la transmisión son: poseer eficiencia económica, que exista competencia y obtener un sistema seguro y diversificado.

A grandes rasgos, según el Decreto 37 [14], la metodología del proceso de planificación se divide en 2 etapas principales: la primera etapa, consta de la conformación de los escenarios de generación necesarios y, la segunda etapa, consiste en el análisis y la propuesta de las necesidades de expansión. En particular, esta segunda etapa está compuesta de los siguientes análisis sucesivos, antes de conformar el Plan de Expansión:

1. Análisis Preliminar,
2. Análisis de Suficiencia y Eficiencia Operacional,
3. Análisis de Seguridad y Calidad de Servicio,
4. Análisis de Factibilidad Técnica y Valorización de los Proyectos de Expansión,
5. Análisis Económico de los Proyectos de Expansión,
6. Análisis de Resiliencia, y
7. Análisis de Mercado Eléctrico Común.

2.1.3. Métodos de *clustering*

El análisis de clústeres tiene una extraordinaria importancia en distintas áreas de investigación científica. Consiste en una tarea descriptiva que pretende agrupar los N objetos disponibles en grupos o clústeres de comportamiento similar, donde a priori los grupos son desconocidos y son precisamente lo que se quiere determinar. Es decir, el objetivo consiste en ser capaces realizar una clasificación de manera que los objetos pertenecientes a un grupo sean tan similares entre sí como sea posible, siendo los distintos grupos entre ellos tan disimilares como sea posible.

Cabe mencionar que estas N observaciones u objetos pueden ser matrices de dimensión $n \times p$, donde, por ejemplo, n puede corresponder a los 365 días del año, y p a las 24 horas del día. El resultado de este algoritmo corresponde a un vector de dimensión $n \times 1$, entregando los índices de cada clúster para cada observación [15].

Los métodos de análisis de clústeres pueden clasificarse de la siguiente manera [16]:

- **Clustering jerárquico:** estos métodos van generando grupos en cada una de las fases del proceso buscando el número de clúster que genera una agrupación óptima. Son capaces de fijar por sí solos el número de clústeres.
 - **Clustering jerárquico aglomerativo:** se parte de un conjunto de elementos individuales y van juntando los elementos que más se parezcan hasta quedarse con un número de clústeres que se considere óptimo.
 - **Clustering jerárquico divisivo:** se parte del conjunto de elementos completo y se van separando los grupos que más diferentes sean entre ellos hasta quedarse con un número de clústeres que se considere óptimo.
- **Clustering no jerárquico:** estos métodos categorizan los elementos según un número de clúster dado. Necesitan que el número de particiones esté fijado a priori.
 - **Métodos de reasignación:** permiten reasignar los individuos a los clústeres en los casos en que la clasificación haya sido dudosa en las primeras etapas del análisis. El proceso acaba cuando no quedan individuos cuya reasignación permita optimizar el resultado que se ha conseguido. Entre ellos se encuentran los métodos *k-means* y *k-medoids*.
 - **Búsqueda de densidad:** se basa en la detección de en qué áreas existen concentraciones de puntos y dónde están separados por áreas vacías o con escasos puntos. Los puntos que no forman parte de un clúster se etiquetan como ruido. Entre estos métodos se encuentra el método Taxmap.
 - **Métodos directos:** permiten clasificar de manera simultánea a los individuos y a las variables. El algoritmo más conocido dentro de este grupo es el *Block-Clustering*.
 - **Métodos reductivos:** consisten en la búsqueda de unos factores en el espacio de los individuos, donde cada factor corresponde a un grupo. Se les conoce como Análisis Factorial tipo Q.

A continuación se ahondará en algunos de estos métodos, los cuales son importantes para el desarrollo de este trabajo.

2.1.3.1. *Clustering* jerárquico aglomerativo

Se comienza con los objetos de modo individual, de tal manera de obtener tantos clústeres iniciales como objetos. Luego, los grupos o clústeres se van agrupando de modo que los primeros en hacerlo son los más similares y al final, todos los subgrupos se unen en un único clúster que contiene todos los objetos. Es importante mencionar que, si un objeto ha sido asignado a un grupo, ya no cambia más de grupo. Finalmente, se deriva un dendograma, que es un gráfico que ilustra cómo se van haciendo las subdivisiones o los agrupamientos, etapa a etapa [17, 18], y a partir del cual se puede obtener la cantidad de clústeres necesarios para tener un nivel de similitud determinado.

Existen diferentes métodos de unión o *linkage methods* para ir formando los clústeres, donde los más importantes son [19]:

- *Linkage* simple, mínima distancia o vecino más próximo: mide la proximidad entre dos grupos calculando la distancia entre sus objetos más próximos o la similitud entre sus objetos más semejantes.
- *Linkage* completo, máxima distancia o vecino más lejano: mide la proximidad entre dos grupos calculando la distancia entre sus objetos más lejanos o la similitud entre sus objetos menos semejantes.
- Distancia media o *average distance*: mide la proximidad entre dos grupos calculando la distancia entre las medias de ambos grupos.
- Método de Ward o mínima varianza: en cada etapa se unen los dos clústeres para los cuales se tenga el menor incremento en el valor total de la suma de los cuadrados de las diferencias de cada individuo dentro de cada cluster, al centroide del mismo. Es el método menos sensible a *outliers*.

Además, son distintas las medidas de distancia que se pueden ocupar para calcular la distancia entre dos objetos. Entre las distancias más habituales se encuentran: la distancia euclidiana, distancia de Manhattan, distancia de Chebyshev, entre otras.

De esta manera, el algoritmo básico dado N objetos es el siguiente [18]:

1. Empezar con N clústeres (el número inicial de elementos) y una matriz $N \times N$ simétrica de distancias o similitudes $D = [d_{ik}]_{ik}$.
2. Dentro de la matriz de distancias, buscar aquella entre los clústeres U y V (más próximos, más distantes o en media más próximos) que sea la menor entre todas, d_{uv} .
3. Juntar los clústeres U y V en uno solo. Actualizar la matriz de distancias:
 - a) Borrando las filas y columnas de los clústeres U y V
 - b) Formando la fila y columna de las distancias del nuevo clúster (UV) al resto de clústeres.
4. Repetir los pasos (2) y (3) un total de $(N - 1)$ veces.

Una consideración importante del *clustering* jerárquico es que es muy sensible a la presencia de objetos muy diferentes del resto (*outliers*). Además, se utiliza cuando no se conoce el número de clústeres a priori y cuando el número de objetos no es muy grande.

2.1.3.2. *Clustering k-means*

Este algoritmo corresponde al método de *clustering* no jerárquico más famoso. Puede aplicarse cuando todas las variables observadas son cuantitativas y lo que hace es buscar la mejor partición de las observaciones en k clústeres.

La idea básica de la agrupación *k-means* es que, dada una agrupación inicial no óptima, se reubica cada punto en su nuevo centro más cercano, se actualizan los centros de agrupación calculando la media de los puntos miembros y se repite el proceso de reubicación y actualización hasta que se satisfagan los criterios de convergencia (como, por ejemplo, un número predefinido de iteraciones). Para calcular la distancia a cada centro, por lo general se utiliza la métrica de distancia euclídea al cuadrado, donde cada centroide es la media de los puntos de ese clúster, y para la inicialización del centro del grupo se utiliza el algoritmo *k-means++* o también se puede realizar de manera aleatoria. Cabe mencionar que, cuando se utiliza la distancia euclídea, la idea del algoritmo consiste en elegir los centroides que minimicen la suma de errores al cuadrado.

El algoritmo procede como sigue:

1. Primero se utiliza el algoritmo *k-means++* para la inicialización de los centros de clúster (también se podría realizar de manera aleatoria).
2. Luego se calculan las distancias euclídeas punto-centroide del clúster de todas las observaciones a cada centroide.
3. Se realiza una actualización por lotes, es decir, se asigna cada observación al clúster con el centroide más cercano.
4. Se calcula la media de las observaciones de cada clúster para obtener k nuevas ubicaciones del centroide.
5. Se repiten los pasos 2 a 4 hasta que las asignaciones de los clústeres no cambien, o hasta que se alcance el número máximo de iteraciones definido.

Cabe mencionar que este método suele ser muy sensible a la solución inicial dada, debido a que se utiliza la media para calcular los centroides. Además, es necesario elegir a priori el valor de k .

2.1.3.3. *Clustering k-medoids*

En este método, en lugar de utilizar la media convencional, se utiliza el medoide para representar a los clústeres, el cual corresponde al elemento más centralizado perteneciente al conjunto de datos. Es decir, a diferencia del método *k-means* que utiliza el vector de medias como centroide, en este caso se usa el vector correspondiente a un dato real (medoide).

El algoritmo *k-medoids* inicia con la selección aleatoria de k elementos de datos como centros iniciales para representar los k clústeres (también se puede utilizar el algoritmo *k-means++*), y los elementos restantes se incluyen en el grupo que tiene el medoide más cercano a ellos. Posteriormente se determina un nuevo centro que puede representar mejor al grupo. En cada iteración, todos los elementos distintos a los centros se asignan nuevamente a los

clústeres que tienen el medoide más cercano, provocando que los medoides alteren su ubicación.

El algoritmo minimiza la suma de las distancias entre cada elemento de datos y su correspondiente medoide, es decir, minimiza la suma de las disimilitudes entre éstos. Este ciclo se repite hasta que ningún medoide cambie su colocación, lo cual marca el final del proceso y se obtienen los clústeres finales.

Cabe mencionar que este algoritmo fue diseñado para no depender del orden de las observaciones o una semilla inicial, debido a que prueba todas las posibles combinaciones, por lo que siempre converge en la misma solución. Además, este corresponde al método más robusto a los *outliers* y al ruido.

2.1.4. Métrica de error RMSE

La Raíz del Error Cuadrático Medio o *Root Mean Squared Error* (RMSE) es un indicador clave de desempeño (KPI, por sus siglas en inglés) bastante útil [20], que se define como la raíz cuadrada del error cuadrático medio, tal como se observa en la ecuación (2.1):

$$RMSE = \sqrt{\frac{1}{n} \sum e_t^2} \quad (2.1)$$

donde e_t corresponde al error en el instante t y, dado el contexto del presente trabajo, se define en la ecuación (2.2) como:

$$e_t = f_t - g_t \quad (2.2)$$

con

- f_t : generación reconstruida a partir de los días representativos en el instante t .
- g_t : generación real en el instante t .

Un problema de este KPI es que no se encuentra escalado a la generación real media. Por ejemplo, si se tiene un RMSE de 10 MWh, no se puede saber si eso es bueno o malo. Si la generación media es de 1000 MWh, sería un buen error, pero si la generación media fuese igual a 1 MWh, entonces sería un error muy grande. Luego, para solucionar esto, generalmente se divide el valor del RMSE por la generación media, para así obtener un porcentaje, tal como se muestra en la ecuación (2.3):

$$RMSE \% = \frac{\sqrt{\frac{1}{n} \sum e_t^2}}{\frac{1}{n} \sum g_t} \quad (2.3)$$

Es importante mencionar que el RMSE no trata cada error de la misma manera, ya que le da mayor importancia a los errores más significativos, por lo que basta con tener un gran error para tener un valor RMSE muy malo. Además, una predicción de los datos más cercana a la media da como resultado un mejor RMSE respecto a una predicción más cercana a la mediana.

2.1.5. Plataforma de simulación Ameba

Ameba corresponde a una plataforma web creada por la consultora SPEC el año 2016, la cual permite simular el comportamiento futuro de los actores que integran el mercado eléctrico y analizarlo mediante múltiples módulos, utilizando diversas escalas de tiempo, ya sea en el corto, mediano o largo plazo, pudiendo considerar múltiples escenarios [21].

De esta manera, la plataforma es capaz de, por ejemplo, proyectar precios de la electricidad, observar el comportamiento en las distintas instalaciones del sistema, evaluar opciones de inversión y determinar la forma óptima de operar los sistemas eléctricos.

Los modelos con los que cuenta Ameba corresponden al modelo de predespacho o Unit Commitment, el modelo PLP (coordinación hidrotérmica), el modelo de expansión o de inversión, y el modelo de autodespacho. A continuación, se explica la función principal de cada uno de estos modelos:

- Modelo de predespacho: permite la evaluación de la operación con una precisión horaria, obteniendo, entre otros, los precios, las restricciones en la operación real, y diversos análisis técnicos.
- Modelo PLP: este modelo de programación a largo plazo permite determinar el valor estratégico del agua, obteniendo tendencias de precios y la evolución de las cotas de los embalses.
- Modelo de expansión: permite determinar los planes de obras, entregando la ubicación de centrales a construir y los refuerzos del sistema de transmisión.
- Modelo de autodespacho: se encarga de maximizar las utilidades para proyectos tomadores de precios.

Para realizar simulaciones en la plataforma, es necesario realizar 4 etapas previas:

1. La primera, consiste en describir, mediante un diagrama, la topología eléctrica e hidráulica a considerar en el caso de estudio.
2. Luego, es necesario ingresar los parámetros de simulación.
3. La tercera etapa corresponde a la definición temporal y realizaciones de escenarios.
4. Por último, se selecciona el modelo y la configuración a utilizar.

Una vez seguidos estos 4 pasos, se puede simular, para así obtener y analizar los resultados, los cuales se pueden ver resumidos en la ventana “Dashboard” o se pueden descargar como archivos .csv.

Respecto a la interfaz, Ameba posee una interfaz gráfica fácil de usar, permitiendo trabajar con grandes volúmenes de datos, evitando cometer errores y perder horas de trabajo. Además, como es una plataforma web, corre las simulaciones en la nube, permitiendo el acceso desde un laptop, tablet o teléfono.

2.2. Estado del arte

Existen diversas formas de reducir la resolución temporal de los datos que se utilizan para la simulación de la operación de un sistema eléctrico. En general, se pueden identificar tres enfoques principales [22]: la selección aleatoria de los datos, la selección heurística y la agregación basada en análisis de clústeres, con una complejidad metodológica creciente.

En primer lugar, la idea de una selección aleatoria es obtener un subconjunto de datos horarios basado en una elección arbitraria o aleatoria. Si bien sus requisitos de implementación son muy bajos, el resultado se ve muy afectado por el número y la selección de los datos horarios del sistema.

Por otro lado, los enfoques heurísticos utilizan procedimientos de selección más avanzados para determinar las horas representativas del sistema, basándose en las características fundamentales del conjunto de datos de referencia, como la estacionalidad dentro del año, la semana o el día. Los datos horarios también pueden seleccionarse basándose en índices de validez externos, como, por ejemplo, métricas de error estadístico.

Por último, los métodos de *clustering* tienen como objetivo identificar subconjuntos similares entre sí dentro de un conjunto de datos, según una medida de similitud (objetiva) entre todas las observaciones, generando grupos homogéneos llamados clústeres. Cabe mencionar que la ponderación de los datos horarios está implícita en función del tamaño del clúster, lo que permite una representación adecuada tanto de los eventos extremos como de los más comunes en la aproximación. Aunque la aplicación tiene un coste más elevado en comparación con otros enfoques, son superiores en cuanto a la resolución temporal de la agregación, y son fundamentales en las investigaciones relacionadas en los últimos años.

Con el fin de estudiar las metodologías existentes de agregación temporal aplicadas en el contexto de la simulación de la operación de un sistema eléctrico, a continuación se describe brevemente la selección de literatura relevante, con una descripción detallada en la Tabla 2.2.

Dentro del enfoque de selección aleatoria, se encuentra el método realizado en [23], el cual consiste en seleccionar repetidamente un subconjunto aleatorio de días (*Random Selection* o RS, por sus siglas en inglés) y retener, de entre todos estos subconjuntos, el que haya obtenido los errores más bajos. En este caso, se puede presentar el inconveniente de que calcular las métricas de error para todos los posibles subconjuntos de días de un mismo año, puede ser inviable desde el punto de vista computacional si la cardinalidad del subconjunto es superior a 3. Es más, en [23] se presenta un novedoso enfoque basado en la optimización para seleccionar períodos representativos, donde se combinan características del enfoque RS y del modelo de optimización MILP (*Mixed-Integer Linear Problem* o Programación Lineal Entera Mixta). Además, en [24] se utiliza un algoritmo de muestreo diario, donde se selecciona una muestra de días que minimice la suma de las diferencias cuadradas de las medias, desviaciones estándar y correlaciones entre los días muestreados y el conjunto de datos completo, en particular, para 10.000 réplicas aleatorias.

Tabla 2.2: Características de la literatura revisada.

Enfoque de selección de datos	Cita	Algoritmo	Nivel de agregación	Nº de muestras o conglomerados	Criterio	Aplicación	Variables de agrupación
Aleatorio	[23]	Selección repetida de un subconjunto aleatorio de días (50.000 subconjuntos aleatorios)	Días representativos	Número creciente	Minimizar los errores (error energético relativo, media del error cuadrático medio normalizado), correlación, curva de duración de rampa	Planificación de expansión de la generación utilizando como referencia el sistema eléctrico de Bélgica	Demanda, energía eólica y solar
	[24]	Selección repetida de un subconjunto aleatorio de días (10.000 subconjuntos aleatorios)	Días representativos	Número creciente	Minimizar la suma de las diferencias cuadradas de las medias, desviaciones estándar y correlaciones	Planificación de inversión en generación	Energía eólica y solar
Heurístico	[25]	Escoger 2 días del mes en función de la demanda (máxima y mediana) utilizando periodos de muestreo de 4 horas dentro del día	Días, horas	24 días con 6 periodos c/u (144 horas/año)	NA	Planificación de la capacidad eléctrica utilizando como referencia el sistema eléctrico de Texas	Demanda, energía eólica y solar
	[25]	Media de los datos anuales	Año	1 periodo	NA	Planificación de la capacidad eléctrica utilizando como referencia el sistema eléctrico de Texas	Demanda, energía eólica y solar
	[26]	Búsqueda exhaustiva de todas las combinaciones de 4 semanas seleccionadas	Semanas representativas	4 semanas (672 horas/año)	Minimizar la suma de errores al cuadrado entre la curva de duración de la carga meta y su aproximación	Expansión de la capacidad	Energía eólica y solar
Análisis de clústeres	[27]	k-Means	Días representativos	5 a 25 días (120 a 600 horas/año)	Minimizar la distancia euclidiana	Operación del sistema eléctrico de Gran Bretaña	Demanda
	[28]	k-Means	Horas	Número creciente (hasta 500 clústeres)	Mínima varianza	Planificación de la transmisión y la generación aplicada al modelo WECC (simplificado) de Estados Unidos	Demanda, energía eólica, solar e hidráulica
	[24]	k-Means	Días representativos	Número creciente (de 5 a 2.000 clústeres)	Minimizar la suma de las diferencias cuadradas de las medias, desviaciones estándar y correlaciones	Planificación de inversión en generación	Energía eólica y solar
	[29]	Agrupación jerárquica de Ward	Días, horas	Días representativos con 8 periodos c/u	Minimizar la suma de errores al cuadrado entre todos los días observados y su representante	Decisiones de inversión y operación (LIMES-EU)	Demanda, energía eólica y solar
	[25]	Clustering jerárquico aglomerativo	Horas	Número creciente	Mínima varianza para buscar las horas duplicadas o extremadamente similares	Planificación de la capacidad eléctrica utilizando como referencia el sistema eléctrico de Texas	Demanda, energía eólica y solar
	[30]	Clustering jerárquico aglomerativo	Días representativos, semanas representativas	Número creciente	Mantener la información cronológica, tiempo de resolución computacional	Planificación de la ampliación de la capacidad con almacenamiento aplicado a la red eléctrica europea	Demanda, energía eólica y solar

Luego, uno de los enfoques heurísticos consiste en obtener distintas resoluciones al escoger uno o más días del mes en función de la demanda y utilizando distintos periodos de muestreo dentro del día. En el caso de [25] se escogen 2 días de cada mes, uno el día de máxima demanda y el otro la mediana, obteniendo un total de 24 días al año, muestreados en periodos de 4 horas, obteniendo de esta manera 144 periodos anuales para representar los datos.

Otra forma de reducir la temporalidad con enfoque heurístico es utilizar la media de los datos horarios de la demanda y la disponibilidad eólica y solar, obteniendo 365 periodos al año al calcular la media diaria, 12 periodos en caso de usar la media mensual y, tal como lo hacen en [25], obteniendo un periodo al usar la media anual de los datos, lo cual no logra representar de manera adecuada la realidad.

Siguiendo con el enfoque anterior, en [26] se realiza una búsqueda exhaustiva de todas las combinaciones de 4 semanas seleccionadas para encontrar la combinación que minimice una métrica de error, considerando tanto la variabilidad de la demanda, como de las energías renovables no convencionales (solar y eólica).

Uno de los métodos más utilizados con el fin de agregación temporal en el contexto de este trabajo, tal como se mencionó anteriormente, es el algoritmo de agrupación o *clustering*. En [27] utilizan el algoritmo de *clustering k-means* para realizar una agregación diaria sobre un modelo de operación del sistema eléctrico de Gran Bretaña, donde cada vector que representaba un día tenía 24 horas, permitiendo considerar las restricciones que necesitan mantener la relación cronológica. En este caso, utilizan el método *k-means* por su tendencia a producir grupos de igual tamaño, evitando que los eventos aislados y muy distintos (el día de Navidad, la final de la Copa del Mundo o fenómenos meteorológicos extraños) formen su propio clúster. Cabe mencionar que en este caso solo consideraron la variabilidad de la demanda.

Al igual que el caso anterior, [28] y [24] también adoptan métodos de agrupación *k-means* para elegir sus horas representativas. En particular, en [28], se utiliza el método sobre el modelo WECC de Estados Unidos, considerando niveles de energía eólica, solar e hidráulica en varias regiones que representan las condiciones de funcionamiento de un año típico y, en [24], donde se utiliza un modelo propuesto. Se trata de minimizar la distancia entre cada observación y los centroides, donde el centroide de cada clúster se utiliza para crear horas representativas, que se ponderan en la función objetivo en proporción al tamaño de cada clúster.

Por otra parte, en [29] se aplica el método de agrupación jerárquica de Ward sobre el modelo LINES-EU (*Long-term Investment Model for the Electricity Sector of Europe*), donde se toma en cuenta la variabilidad a corto plazo de la demanda de electricidad y de las fuentes de energía renovable eólica y solar. El método en cuestión se encarga de agrupar los días con patrones similares de demanda diurna y de entrada de energías renovables, con el fin de que cada grupo de días se refleje en un día representativo en el modelo del sistema eléctrico. De manera similar, en [25] se plantea un método de *clustering* jerárquico aglomerativo, aplicado a una representación reducida del sistema eléctrico de Texas, donde se fusionan los clústeres que producen la mínima varianza en el clúster fusionado, utilizando el método Ward hasta que se alcanza un número deseado de clústeres, con el fin de buscar las horas duplicadas o extremadamente similares entre sí.

En un estudio más reciente [30], también se utiliza el método de *clustering* jerárquico aglomerativo, aplicado esta vez a una simplificación de la red eléctrica europea. En este caso, a diferencia de los anteriores, se propone un método para agrupar los periodos de tiempo manteniendo, en la medida de lo posible, la información cronológica de los parámetros dependientes del tiempo a lo largo de todo el horizonte de planificación.

Todos los métodos presentados corresponden a soluciones propuestas por otros autores para la reducción temporal de los datos de entrada de la simulación de la operación para distintos Sistemas Eléctricos. Sin embargo, hasta el momento no se ha realizado ningún estudio que pruebe el rendimiento de las distintas metodologías existentes, o que proponga alguna nueva metodología de agregación aplicada al Sistema Eléctrico Nacional chileno, de tal manera de comparar los resultados con los obtenidos en la simulación de la operación para la planificación de la transmisión, utilizando la agregación temporal empleada en la actualidad.

Capítulo 3

Antecedentes

3.1. Descripción del caso de estudio

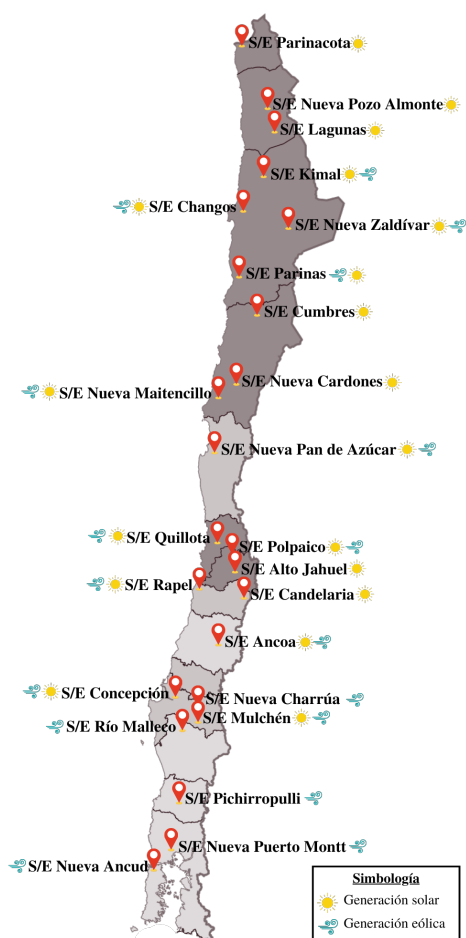


Figura 3.1: Mapa del SEN donde se muestran las S/E consideradas, en conjunto con la tecnología asociada.

Los datos empleados en el desarrollo de este trabajo corresponden a los utilizados en el Plan de Expansión Anual de la Transmisión del año 2022, realizada por la Comisión Nacional de Energía [13], en particular para el escenario E de la Planificación Energética de Largo Plazo (PELP) 2018-2022 realizada por el Ministerio de Energía [31]. Este escenario considera, entre otras cosas, lo siguiente:

- Altos costos de inversión para todas las tecnologías de generación, excepto la solar fotovoltaica (referencial).
- Busca alcanzar la descarbonización al año 2030.
- Proyección de demanda de energía eléctrica alta.

De esta manera, los datos principales utilizados en este trabajo corresponden a la potencia nominal de las centrales eólicas y solares a lo largo de todo el SEN, considerando posibles expansiones entre los años 2021 y 2041 y fechas de puesta/salida de servicio, además de la subestación asociada y sus respectivos perfiles de generación horaria, todos estos definidos por el Ministerio de Energía. En las Tablas A.1 y A.2, adjuntas en Anexos, se observan, respectivamente, las centrales eólicas y fotovoltaicas con los datos recién mencionados.

Por otro lado, en la Figura 3.1 se presenta un mapa de Chile donde se destacan las subestaciones en cuestión, con el fin de tener una idea de la ubicación geográfica de cada una de

ellas. Este mapa está basado en el mapa del Coordinador Eléctrico Nacional, actualizado en noviembre del año 2019.

Luego, se generan las Tablas 3.1 y 3.2, donde se muestra la potencia nominal y la generación total (obtenida al multiplicar la potencia por los perfiles horarios) de cada subestación, en conjunto con el número de centrales asociadas, según el tipo de tecnología, proyectadas a los años 2025, 2030, 2035 y 2040.

Tabla 3.1: Proyección de la potencia instalada y generación eólica por subestación a los años 2025, 2030, 2035 y 2040, en conjunto con el número de centrales asociadas a cada una de ellas.

Subestación	N° de centrales eólicas	Potencia instalada [MW]				Generación [10^6 MWh]			
		2025	2030	2035	2040	2025	2030	2035	2040
Kimal	7	375	525	1.995	3.558	1,05	1,52	5,85	10,67
Changos	1	156	156	156	156	0,42	0,42	0,42	0,42
Nueva Zaldívar	5	424	424	430	657	1,10	1,10	1,11	1,74
Parinas	8	1.078	3.163	4.795	5.526	3,54	10,50	15,57	17,78
Nueva Maitencillo	10	902	1.062	1.196	1.264	2,53	3,02	3,41	3,60
Nueva Pan de Azúcar	17	962	1.607	1.640	1.640	1,93	3,86	3,96	3,96
Quillota	2	-	197	197	197	-	0,55	0,55	0,55
Polpaico	7	128	153	153	153	0,32	0,40	0,40	0,40
Rapel	1	-	45	45	45	-	0,13	0,13	0,13
Ancoa	6	-	713	835	899	-	2,15	2,49	2,67
Concepción	24	732	1.682	2.280	3.392	2,03	4,79	6,62	10,14
Nueva Charrúa	6	72	925	925	960	0,20	2,59	2,59	2,69
Mulchén	9	640	1.037	1.688	1.807	1,71	2,88	4,89	5,24
Río Malleco	9	-	190	2.569	3.218	-	0,56	7,81	9,74
Pichirropulli	8	156	205	1.004	1.143	0,47	0,63	3,08	3,54
Nueva Puerto Montt	8	379	725	742	773	1,27	2,31	2,36	2,44
Nueva Ancud	1	-	868	1.701	1.930	-	2,74	5,37	6,09
TOTAL SEN		6.004	13.677	22.351	27.318	16,58	40,15	66,61	81,79

De la Tabla 3.1 se tiene que la subestación donde se conectará la mayor potencia eólica instalada en el SEN es Parinas, mientras que de la Tabla 3.2 se tiene que la subestación donde se conectará la mayor potencia solar fotovoltaica instalada es Kimal. También se observa que habrá un aumento sostenido en el tiempo de instalación eólica, mientras que para la tecnología solar fotovoltaica, se tendrá un aumento significativo entre los años 2035 y 2040. Cabe mencionar que, la relación entre potencia instalada y generación, es directa, ya que el perfil de generación utilizado es el mismo para cada año.

Luego, en la Figura 3.2 se observa de manera gráfica la proyección de la generación por subestación, tanto eólica como solar fotovoltaica, para cada uno de los años en cuestión, donde se puede comprobar lo anteriormente descrito. Además, se ve que a futuro habrá una gran capacidad de generación eólica en todo el país, a excepción de la Región de Valparaíso y Metropolitana. Respecto a la generación solar fotovoltaica, se tendrá una mayor predominancia en la zona norte de Chile, entre la subestación Parinacota, en la Región de Arica y Parinacota, y la subestación Cumbres, ubicada en la Región de Atacama.

Tabla 3.2: Proyección de la potencia instalada y generación solar fotovoltaica por subestación a los años 2025, 2030, 2035 y 2040, en conjunto con el número de centrales asociadas a cada una de ellas.

Subestación	N° de centrales solares FV	Potencia instalada [MW]				Generación [10 ⁶ MWh]			
		2025	2030	2035	2040	2025	2030	2035	2040
Parinacota	2	124	124	124	124	0,32	0,32	0,32	0,32
Nueva Pozo Almonte	11	386	386	386	386	1,13	1,13	1,13	1,13
Lagunas	11	992	992	992	2.679	2,88	2,88	2,88	7,78
Kimal	26	4.480	4.480	4.480	4.480	13,09	13,09	13,09	13,09
Changos	1	72	72	72	72	0,22	0,22	0,22	0,22
Nueva Zaldívar	10	710	710	710	710	2,10	2,10	2,10	2,10
Parinas	4	128	128	128	1.301	0,39	0,39	0,39	3,85
Cumbres	24	2.261	2.261	3.457	3.789	6,37	6,37	9,93	10,92
Nueva Cardones	15	624	624	624	624	1,70	1,70	1,70	1,70
Nueva Maitencillo	8	786	982	982	982	1,99	2,50	2,50	2,50
Nueva Pan de Azúcar	6	189	189	189	244	0,46	0,46	0,46	0,60
Quillota	20	244	320	320	875	0,55	0,73	0,73	1,87
Polpaico	27	563	739	739	786	1,28	1,69	1,69	1,78
Alto Jahuel	18	151	151	151	151	0,37	0,37	0,37	0,37
Rapel	1	-	147	147	147	-	0,34	0,34	0,34
Candelaria	1	-	-	-	13	-	-	-	0,03
Ancoa	20	123	123	123	1.041	0,27	0,27	0,27	2,29
Concepción	17	102	102	102	102	0,23	0,23	0,23	0,23
Mulchén	4	12	12	12	58	0,03	0,03	0,03	0,12
TOTAL SEN		11.947	12.542	13.738	18.564	33,36	34,79	38,35	51,24

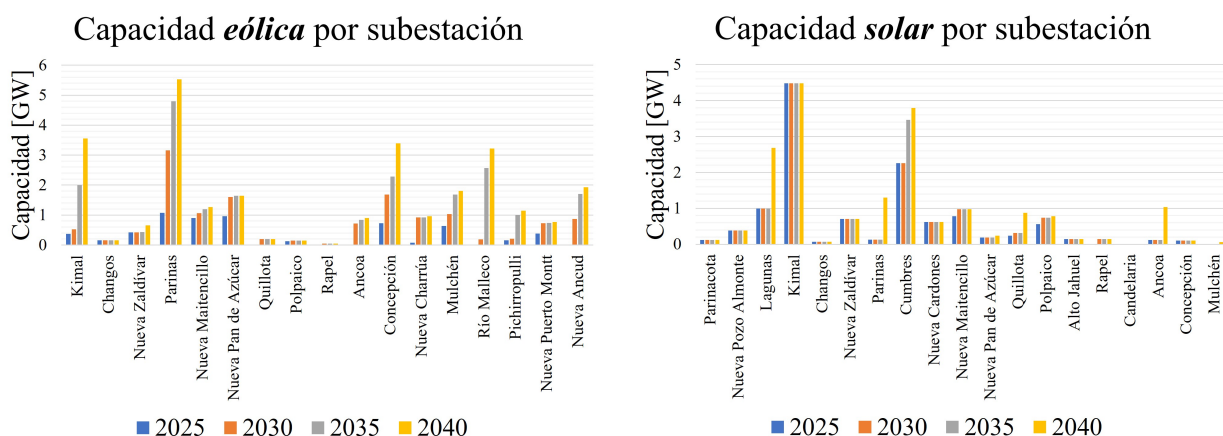


Figura 3.2: Capacidad anual de generación prevista para cada subestación asociada, tanto para las centrales eólicas como las solares fotovoltaicas.

Cabe mencionar que, debido a la relevancia actual que tienen las siguientes subestaciones en el rubro de la generación a partir de ERNC en el país, en este trabajo de título se mostrarán sólo los resultados correspondientes a Parinas y Río Malleco para la generación eólica, y los relacionados con Kimal y Polpaico para la generación solar fotovoltaica, además del Sistema Eléctrico Nacional en su conjunto, a menos que sea necesario mostrar el comportamiento de otra subestación. En cuanto a los años, sólo se considerará el año 2040, ya que es en este año donde se encuentran instaladas todas las expansiones para ambas tecnologías. Sin embargo, a pesar de mostrar sólo los resultados para estos casos, durante el desarrollo de este trabajo se ha estudiado el comportamiento para todas las subestaciones mostradas en la Figura 3.1 y cada uno de los 4 años mencionados.

3.2. Método de agregación temporal de la CNE

Para obtener una reducción temporal de los datos, la Comisión Nacional de Energía asigna bloques y etapas de manera horaria, tal como se observa en la Tabla 3.3, donde los bloques van de 1 a 24 según la hora del día y si corresponde a un día hábil o no hábil, y las etapas toman valores de 1 a 252 según el mes en cuestión para cada año, entre el año 2021 y 2041.

Por ejemplo, el día hábil representativo de la etapa 1, que corresponde a enero 2021, se obtiene al calcular el promedio de generación, ya sea solar o eólica, o de la demanda, de los bloques 1 al 12 considerando todos los días hábiles de la etapa 1, y luego repitiendo el promedio asociado a cada bloque 2 veces de manera consecutiva, ya que son bi-horarios. Esto, debido a que con el método se obtienen 12 promedios para representar un día, pero un día consta de 24 horas. Este procedimiento es análogo para el resto de las etapas, y para el día no hábil, donde en este último caso la única diferencia es que se utilizan los bloques 13 al 24.

De esta manera, con el método de la CNE se obtienen dos días representativos para cada uno de los meses considerados, uno que representa al día hábil de cada mes y, el otro, que representa al día no hábil. Además, es importante mencionar que, en este caso, como la asignación de bloques se basa únicamente en los días hábiles y no hábiles, esta metodología es independiente de si existe o no división geográfica de los datos, y de si se realiza o no de manera conjunta para los datos de viento y sol, ya que se obtienen los mismos resultados.

Tabla 3.3: Asignación de bloques y etapas según la CNE.

Asignación de bloques			Asignación de etapas												
Hora	Día hábil	Día no hábil	Año	ene	feb	mar	abr	may	jun	jul	ago	sept	oct	nov	dic
0:00	1	13	2021	1	2	3	4	5	6	7	8	9	10	11	12
1:00	1	13	2022	13	14	15	16	17	18	19	20	21	22	23	24
2:00	2	14	2023	25	26	27	28	29	30	31	32	33	34	35	36
3:00	2	14	2024	37	38	39	40	41	42	43	44	45	46	47	48
4:00	3	15	2025	49	50	51	52	53	54	55	56	57	58	59	60
5:00	3	15	2026	61	62	63	64	65	66	67	68	69	70	71	72
6:00	4	16	2027	73	74	75	76	77	78	79	80	81	82	83	84
7:00	4	16	2028	85	86	87	88	89	90	91	92	93	94	95	96
8:00	5	17	2029	97	98	99	100	101	102	103	104	105	106	107	108
9:00	5	17	2030	109	110	111	112	113	114	115	116	117	118	119	120
10:00	6	18	2031	121	122	123	124	125	126	127	128	129	130	131	132
11:00	6	18	2032	133	134	135	136	137	138	139	140	141	142	143	144
12:00	7	19	2033	145	146	147	148	149	150	151	152	153	154	155	156
13:00	7	19	2034	157	158	159	160	161	162	163	164	165	166	167	168
14:00	8	20	2035	169	170	171	172	173	174	175	176	177	178	179	180
15:00	8	20	2036	181	182	183	184	185	186	187	188	189	190	191	192
16:00	9	21	2037	193	194	195	196	197	198	199	200	201	202	203	204
17:00	9	21	2038	205	206	207	208	209	210	211	212	213	214	215	216
18:00	10	22	2039	217	218	219	220	221	222	223	224	225	226	227	228
19:00	10	22	2040	229	230	231	232	233	234	235	236	237	238	239	240
20:00	11	23	2041	241	242	243	244	245	246	247	248	249	250	251	252
21:00	11	23													
22:00	12	24													
23:00	12	24													

Luego, se reproducen los datos horarios de cada mes asignando los valores del día hábil representativo a los días hábiles, y los valores del día no hábil representativo a los días no hábiles. De esta manera, es posible calcular distintas métricas de error de esta representación al compararla con los datos horarios.

3.3. Adaptación de los días representativos hábil y no hábil de generación renovable a los datos horarios

A continuación, en la Figura 3.3 se presenta la adaptación de los días representativos obtenidos con la metodología de la CNE a los datos horarios reales. En particular, se mostrará la adaptación para 2 meses, enero y junio del año 2040, ya que son meses que presentan un comportamiento distinto. Cabe mencionar que esta adaptación está medida mediante la Raíz del Error Cuadrático Medio o RMSE por sus siglas en inglés, la cual se calcula utilizando la reproducción horaria de los datos, obtenida utilizando el método actual de la CNE, respecto a los datos horarios reales.

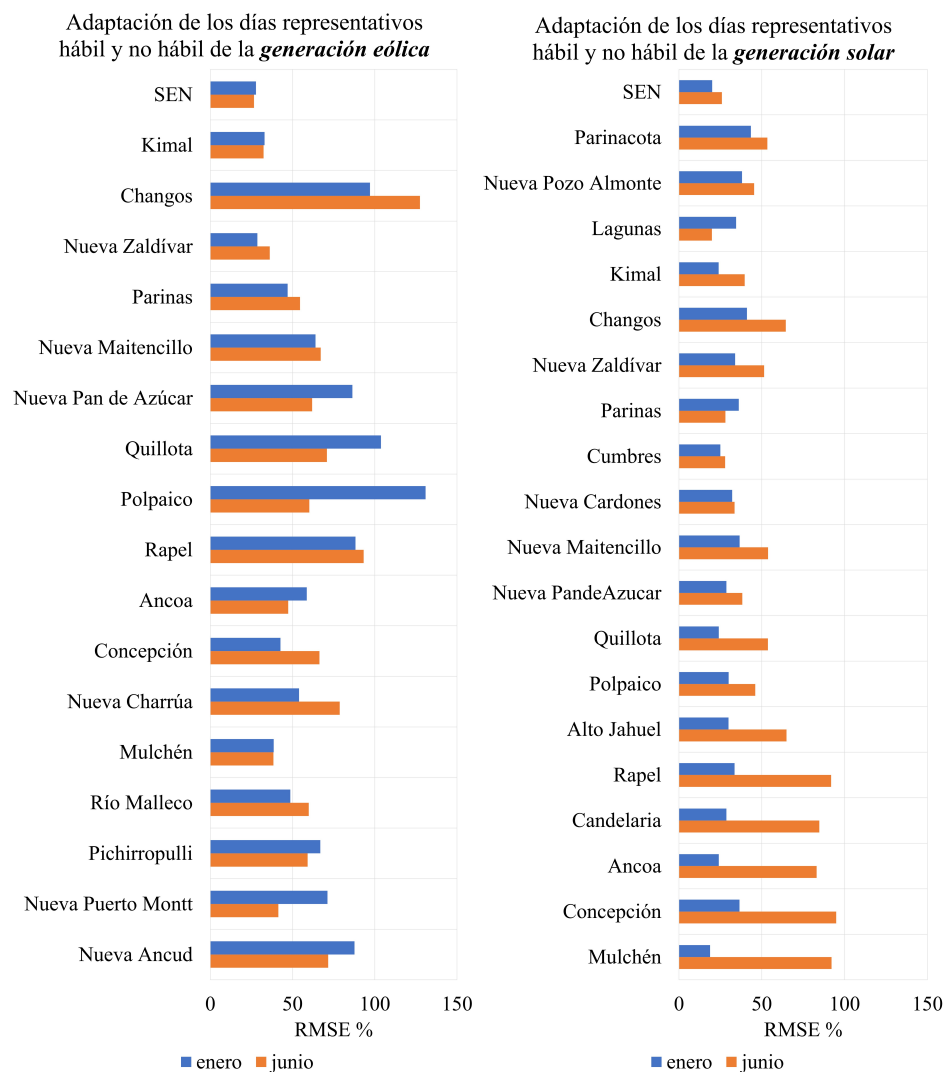


Figura 3.3: Adaptación de los días representativos hábil y no hábil de la generación eólica y solar a los datos reales utilizando el método actual de la CNE, medida con el indicador RMSE.

A partir del gráfico de barras para la generación eólica se ve que, si bien a nivel de Sistema Eléctrico no hay gran diferencia entre ambos meses, en algunas zonas es bastante mayor el error en enero, como, por ejemplo, en la zona norte-centro del país, entre las subestaciones

Nueva Pan de Azúcar y Polpaico, donde destaca el gran error que hay en esta última, el cual se genera debido a que en enero la generación es nula o casi nula la mayoría de los días, provocando un error muy alto en caso de que los días representativos tomen un valor no nulo. Además, se observa que, en general, los errores alcanzados en ambos meses son bastante grandes, encontrándose en su mayoría entre un 50 % y 100 %.

En particular, en la subestación Parinas, el error es mayor en junio debido a la variabilidad de los días, lo cual queda representado en la Figura 3.4. Lo mismo ocurre en la subestación Río Malleco, tal como se ve en la Figura 3.5, lo cual no es visible a simple vista. Sin embargo, se observa que en junio los valores máximos alcanzados son mayores que en enero, por lo que los días representativos quedan más alejados de estos valores extremos.

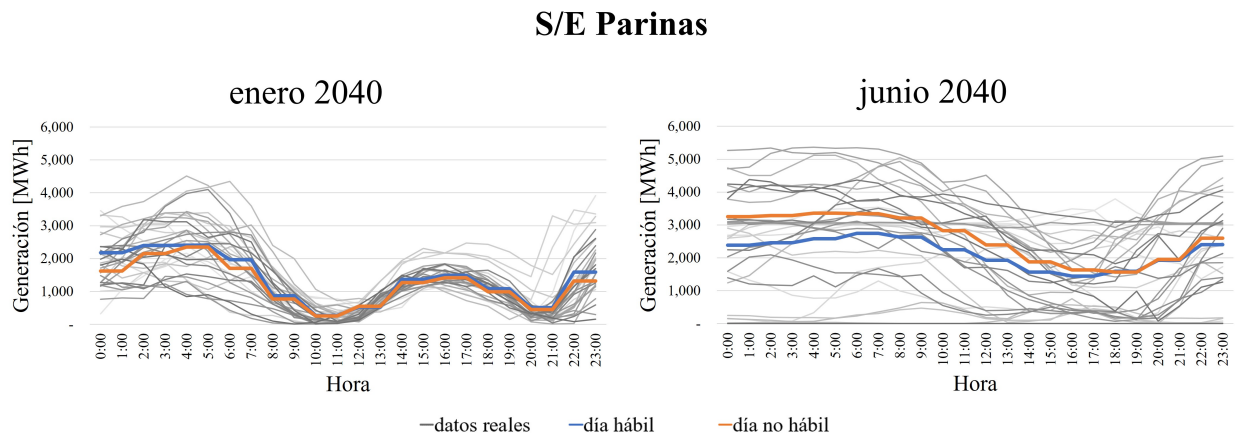


Figura 3.4: Adaptación de los días representativos hábil y no hábil de la CNE a los datos horarios reales de generación eólica para la subestación Parinas (enero y junio 2040).

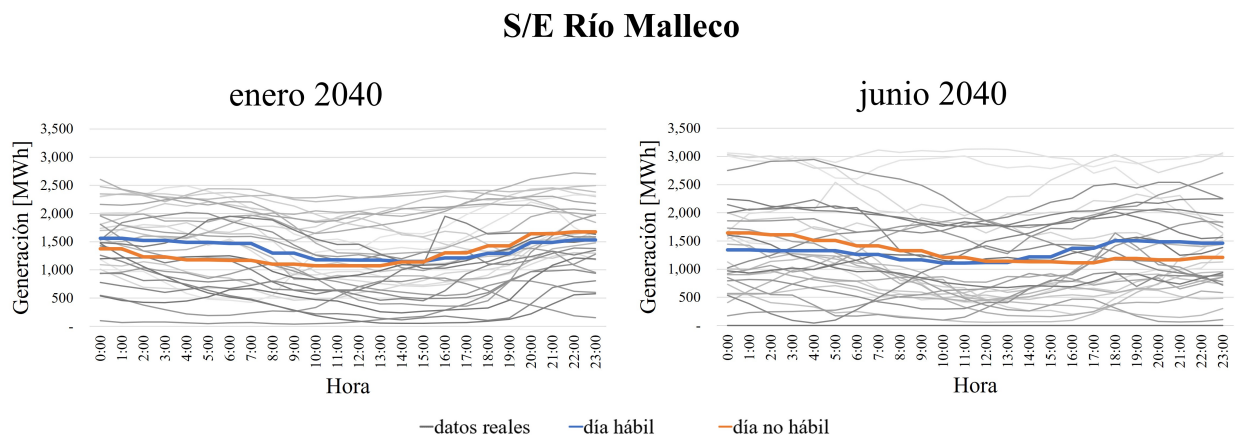


Figura 3.5: Adaptación de los días representativos hábil y no hábil de la CNE a los datos horarios reales de generación eólica para la subestación Río Malleco (enero y junio 2040).

Luego, del gráfico de barras para la generación solar fotovoltaica se ve claramente que, en junio, el error es mucho mayor en la zona centro-sur del país, entre las subestaciones Quillota y Río Malleco, alcanzando valores cercanos al 75 %-100 % de error, debido a que se presentan

días u horas aisladas donde la generación disminuye de manera considerable por la nubosidad característica de invierno, alejando los días representativos de lo generado la mayoría de los días. Esto se puede observar de manera gráfica en las Figuras 3.6 y 3.7.

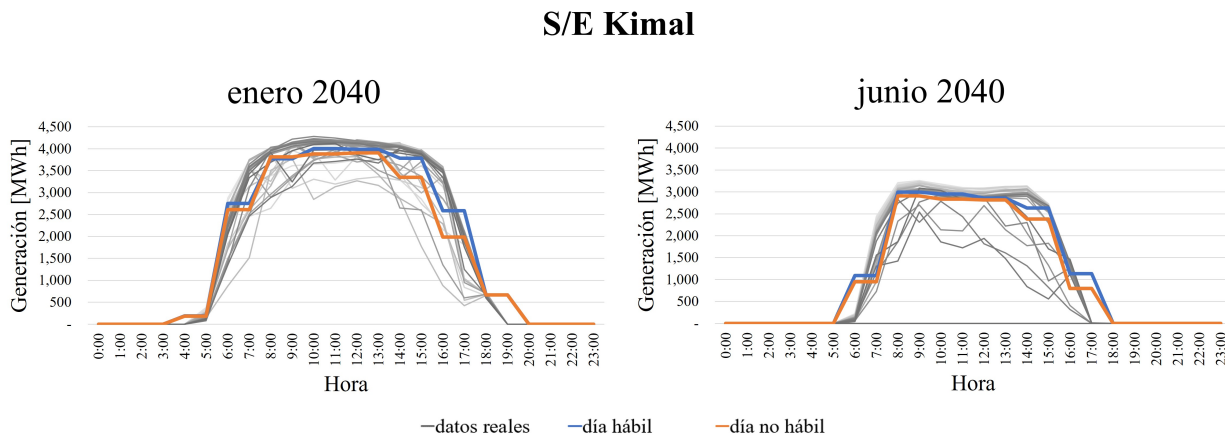


Figura 3.6: Adaptación de los días representativos hábil y no hábil de la CNE a los datos reales de generación solar fotovoltaica para la subestación Kimal (enero y junio 2040).

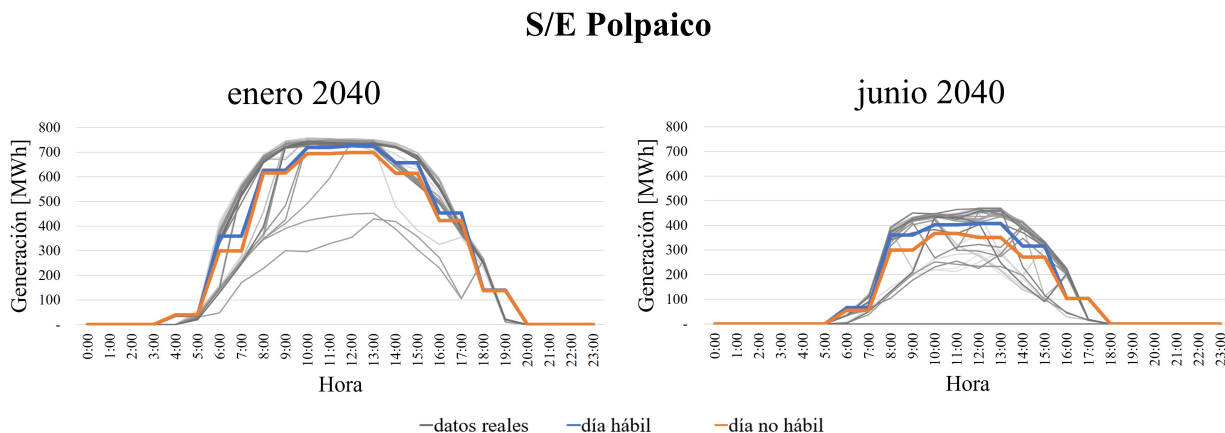


Figura 3.7: Adaptación de los días representativos hábil y no hábil de la CNE a los datos reales de generación solar fotovoltaica para la subestación Polpaico (enero y junio 2040).

Es por estos grandes errores inducidos con la metodología de la CNE, que uno de los objetivos de este estudio es reducir la magnitud del error, que es lo mismo que mejorar la adaptación de los días representativos a la data real. Además, se observa que no hay gran diferencia entre un día representativo hábil y uno no hábil, por lo que no aporta mucha información tener 2 días representativos por mes en este caso para la generación ERNC.

3.4. Correlación entre zonas

En esta sección se analiza la correlación entre zonas al utilizar la metodología actual de la CNE, donde las zonas quedan definidas por la subestación asociada a cada central

generadora. En particular, para el viento, se tiene lo mostrado en la Figura 3.8, donde, al aplicar la metodología de la CNE, se observan fuertes correlaciones entre zonas que antes no estaban fuertemente correlacionadas, como, por ejemplo, lo que ocurre entre la subestación Quillota y el centro-sur del país, o lo que ocurre desde Ancoa hacia el sur.

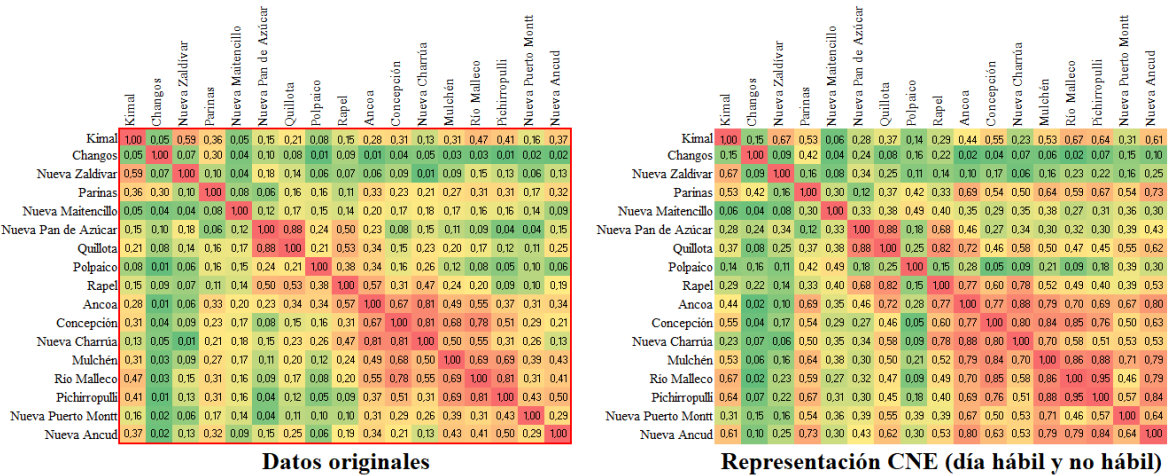


Figura 3.8: Matriz de correlación entre zonas de los datos horarios de 2040 versus la correlación entre zonas utilizando la metodología de la CNE, para los datos de viento.

Respecto a la generación solar fotovoltaica, se observa que, en general, todos los valores son mayores al aplicar la metodología de la CNE, en particular en la zona norte-centro, desde la subestación Parinas hasta Nueva Pan de Azúcar, tal como se observa en la Figura 3.9.

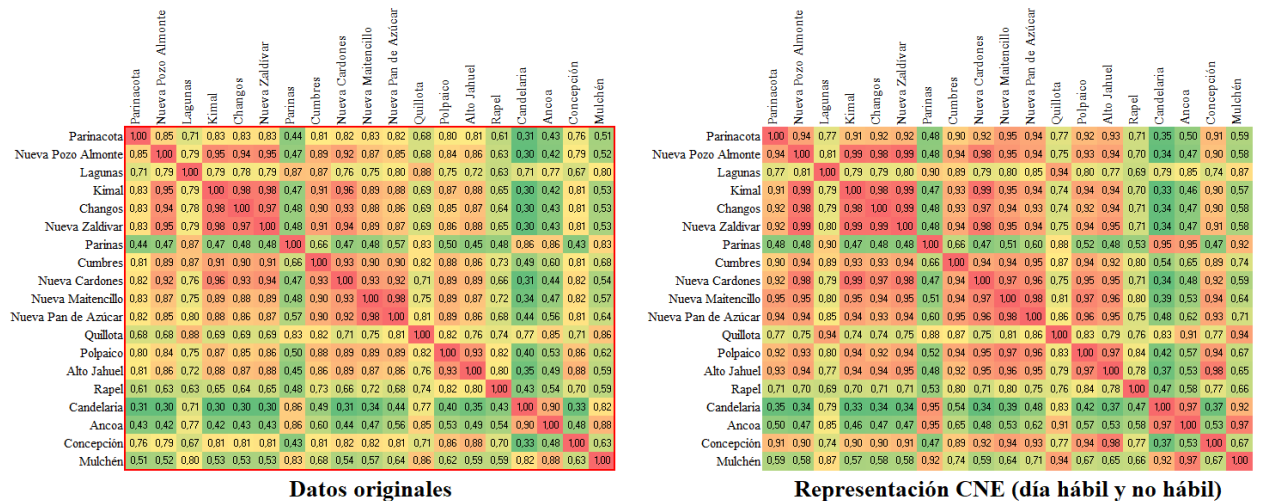


Figura 3.9: Matriz de correlación entre zonas de los datos horarios de 2040 versus la correlación entre zonas utilizando la metodología de la CNE, para los datos de sol.

Cabe recalcar que mantener la correlación original entre subestaciones es importante, debido a las congestiones en las líneas de transmisión que se pueden presentar entre zonas. De esta manera, otro de los temas a analizar al momento de proponer una nueva metodología de agregación temporal, es ver si se mejora o no la correlación entre zonas.

Capítulo 4

Metodología y desarrollo

4.1. Esquema metodológico

En la Figura 4.1 se presenta un esquema que especifica las etapas que componen la metodología de este trabajo de título y que son necesarias para cumplir con los objetivos establecidos.

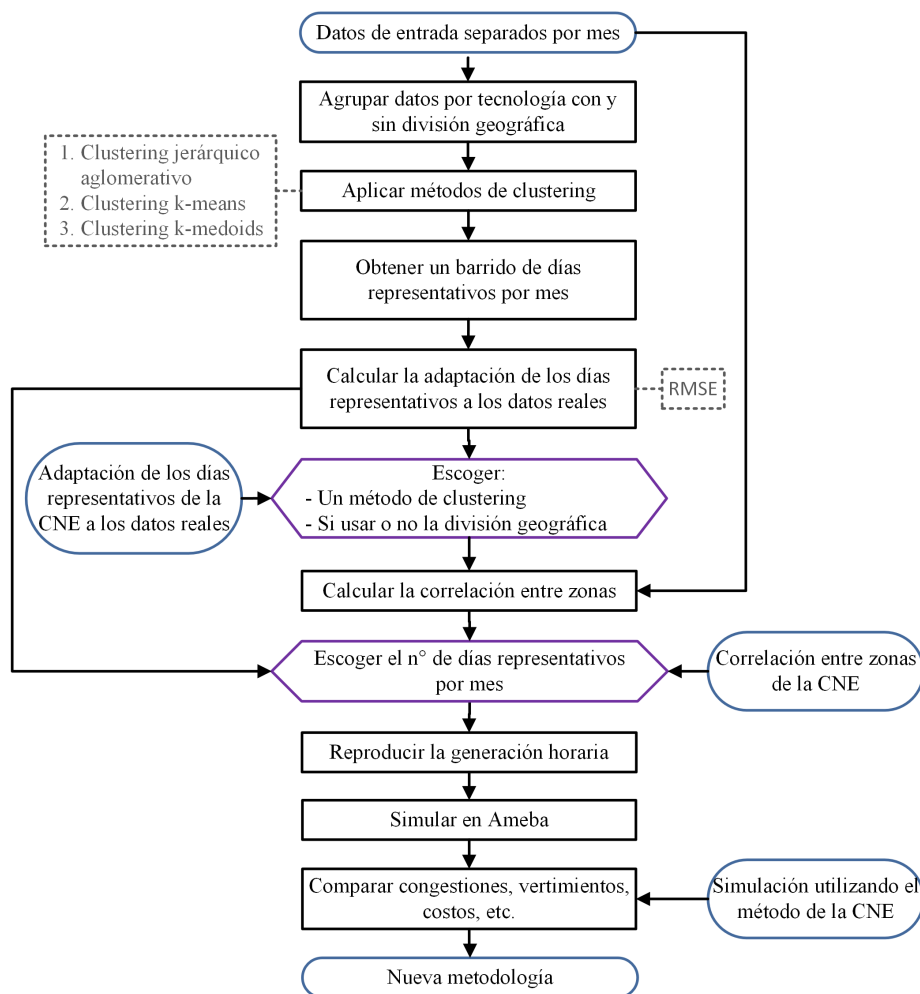


Figura 4.1: Diagrama de bloques para la metodología propuesta.

La metodología empleada considera los siguientes puntos:

- Los datos de entrada deben estar separados por mes, ya que así es como los utiliza actualmente la Comisión Nacional de Energía.
- Además, estos datos deben agruparse de dos maneras distintas:
 1. Sin división geográfica, es decir, deben estar agrupados para todo el Sistema Eléctrico Nacional en su conjunto.
 2. Con división geográfica, que en este caso, por simplicidad, está dada por la subestación relacionada con las centrales generadoras.

Es importante mencionar que, debido a que no todas las subestaciones tienen necesariamente generación eólica y solar, los datos también se dividen por tecnología.

- Los métodos de agregación temporal utilizados corresponden a los siguientes 3 métodos de *clustering* revisados en el Estado del Arte: el *clustering* jerárquico aglomerativo, *clustering k-means* y *clustering k-medoids*.
- Se obtiene un barrido de días representativos por mes, de 1 a 10, para así poder ver el comportamiento a medida que aumenta la densidad de datos utilizados, y poder determinar posteriormente el número de días representativos necesarios por mes para obtener los resultados que se deseen.
- Además, es necesario calcular la correlación entre zonas dadas por la subestación, para ver si se mantiene o no, y tener otro punto de vista para determinar la cantidad de días representativos necesarios por mes.
- Para calcular la adaptación de los días representativos a la data original, ya sea anual o mensual, se utiliza la Raíz del Error Cuadrático Medio medida en porcentaje, ya que es la métrica más utilizada en este tipo de problemas, y porque al estar medida en porcentaje no es necesario tener conocimientos de la generación media, ya que está escalado.
- Es necesario tener los antecedentes de los resultados que se obtienen al aplicar el método actual de reducción de datos de la Comisión Nacional de Energía, para poder realizar las comparaciones correspondientes.
- Las decisiones a tomar en el proceso, para poder realizar la simulación, son las siguientes:
 1. Escoger 1 de los 3 métodos de *clustering* aplicados.
 2. Determinar si se considerará o no la división geográfica por subestación.
 3. Definir el número de días representativos por mes.
- Para poder realizar la simulación, es necesario tener una reproducción horaria de la generación al utilizar los días representativos obtenidos con la nueva metodología propuesta.
- La simulación se realiza en la plataforma Ameba, la cual tiene las ventajas de tener fácil acceso al ser una plataforma web, y de permitir realizar distintos análisis debido a los múltiples módulos que posee.
- Con el fin de definir si efectivamente la nueva metodología propuesta es mejor que la de la CNE, se analizan distintos indicadores, tales como los costos operacionales, el tiempo de resolución computacional, la generación obtenida por tecnología, entre otros.

4.2. Desarrollo de la propuesta metodológica

En primer lugar, se tienen los datos de entrada, entre ellos, la generación eólica y solar del Sistema Eléctrico Nacional prevista para los años 2021-2041, los cuales se separaron por mes, tal como lo realiza la Comisión Nacional de Energía. Luego, el primer paso fue agrupar estos datos por tecnología, tanto eólica como solar fotovoltaica, y separarlos por subestación asociada (caso 1), o sumarla y tener un total para todo el SEN (caso 2). Además, fue necesario normalizar los datos agrupados, para así poder aplicar los métodos de *clustering*. Luego, a cada uno de estos grupos de datos, se les aplicó los 3 métodos de *clustering* mencionados en el esquema metodológico, el *clustering* jerárquico aglomerativo, *clustering k-means* y *clustering k-medoids*, con el fin de obtener un barrido de 1 a 10 días representativos por cada mes. Cabe mencionar que, tanto para el *clustering* jerárquico como para el *k-means*, cada día representativo está dado por el promedio de cada clúster, mientras que para el método *k-medoids* cada día representativo corresponde a un día de la data real.

Una vez obtenidos los días representativos con los métodos de *clustering*, se procedió a reproducir la generación horaria anual de todo el SEN, tanto para el caso 1 como para el caso 2, y se calculó el RMSE según el número de días representativos, es decir, se midió el ajuste de los datos “clusterizados” a los datos horarios reales, o lo que es lo mismo, se midió el rendimiento de los métodos de *clustering*. Luego, se compararon y analizaron estos errores obtenidos, en conjunto con el error dado al utilizar la metodología de la CNE, y se procedió a escoger uno de los 3 métodos de *clustering* empleados.

Luego, al comparar el RMSE anual del SEN del método de *clustering* seleccionado del caso 1 versus el del caso 2, se escogió si se utiliza o no la división por subestación, utilizando como criterio obtener un mejor ajuste a los datos reales. De manera paralela, se procedió a calcular la correlación entre subestaciones según el número de días representativos para la metodología definida, además de calcularla para los datos horarios y tener presente la correlación utilizando la metodología de la CNE, para así compararlas y analizar si se mantiene o no la correlación entre zonas para ambos métodos, y tener una idea de los días representativos necesarios en el caso del método de *clustering*.

Seguido de esto, se calculó nuevamente el RMSE según el número de días representativos, pero esta vez de manera mensual para el método y la modalidad (o caso) seleccionado, para así compararlo con el error del método de la CNE y escoger el o los números de días representativos necesarios por mes para que se cumpla algún o algunos criterios a determinar (como por ejemplo, que el RMSE de la CNE disminuya un 40 %, o que el RMSE sea menor a un 12 %, entre otros).

Hecho esto, se obtuvieron los datos de entrada a la simulación en Ameba, que corresponden a todos los perfiles de generación horarios, tanto eólicos como solares, obtenidos al utilizar el método de *clustering* escogido con el número de días representativos definidos, además del archivo de bloques, el cual está asociado al “peso” de cada día representativo en el mes correspondiente (es decir, al porcentaje de representación de cada clúster), y a la asignación de día hábil y no hábil de la CNE.

Para obtener los perfiles de generación horarios con el método de *clustering* escogido, una

vez conocido el número de días representativos necesario por mes, se utilizó la asignación de clústeres mensual obtenida para el año 2040, la cual se aplicó a los perfiles según su respectiva subestación asociada. Luego, dependiendo del método de *clustering*, se obtuvo el promedio de cada clúster mensual, o se encontró el medoide de cada clúster de manera mensual, obteniendo así los días representativos por mes de cada perfil. De esta manera se crea el archivo de perfiles, el cual contiene $n \cdot 24 \cdot 12 \cdot 2$ filas, siendo n el número de clústeres necesarios por mes, 24 las horas del día, 12 los meses del año, y todo se multiplica por 2 debido a que es necesario repetir cada día representativo para cruzar los datos con la demanda agregada según el día hábil o no hábil de la CNE.

Por otro lado, para obtener el archivo de bloques, en primer lugar fue necesario calcular el porcentaje de representación de cada día representativo obtenido con el método de *clustering* en cada mes, para cada uno de los perfiles. Por simplicidad, una vez obtenido estos valores para cada perfil, se calculó el promedio de los “pesos” para cada día representativo de cada mes, con el fin de obtener sólo un vector que sirva para todos los perfiles. Realizado lo anterior, y para obtener el archivo de bloques, fue necesario cruzar estos pesos con los pesos del día hábil y no hábil de la CNE. Por ejemplo, si el mes de enero contiene 9 días no hábiles y 22 días hábiles según la asignación de la CNE, y si la asignación de clústeres es de 18 días para el clúster 1 y 13 días para el clúster 2 (al emplear 2 días representativos por mes), entonces los bloques asignados son los siguientes:

$$\begin{aligned} \text{bloque 1} &= \frac{9 \cdot 18}{31} = 5 & \text{bloque 2} &= \frac{22 \cdot 18}{31} = 13 \\ \text{bloque 3} &= \frac{9 \cdot 13}{31} = 4 & \text{bloque 4} &= \frac{22 \cdot 13}{31} = 9 \end{aligned}$$

Es decir, el bloque 1 representa a 5 de los 31 días del mes de enero, el bloque 2 a 13 de 31, etc. De esta manera los bloques impares representan a los días no hábiles, mientras que los días pares a los días hábiles. Es necesario hacer este cruce de información, dado que en este estudio el método de *clustering* sólo se aplica para la generación solar fotovoltaica y eólica, pero la demanda y la generación dada por otras tecnologías, entre otros datos, siguen la representación de la CNE. Cabe mencionar que esta asignación de bloques se realiza para todos los meses de cada uno de los años, por lo que en un año, si se utilizan n clústeres al mes, se obtienen $2 \cdot n$ bloques al mes, lo que equivale a $24 \cdot n$ bloques al año.

Por último, se procedió a realizar la simulación tanto para la representación actual de la CNE, así como para la nueva representación propuesta, y se compararon distintos indicadores de salida, entre ellos, los costos operacionales, el tiempo de resolución computacional, la generación por tecnología, entre otros. Además, se simuló la operación con los datos horarios sin agregar, con el fin de realizar comparaciones. Es de esta manera que se pudo analizar si efectivamente se obtienen mejores resultados, y así queda propuesta una nueva metodología a emplear para la simulación de la operación utilizada en el proceso de planificación de la transmisión.

Capítulo 5

Resultados y análisis

5.1. Elección del método de *clustering*

El primer resultado corresponde a la Raíz del Error Cuadrático Medio anual tanto de la generación eólica como solar del Sistema Eléctrico Nacional, según el número de clústeres empleado, para los 3 métodos de *clustering* seleccionados, con y sin división geográfica por subestación.

A partir de la Figura 5.1 se observa que, tanto para la generación eólica como para la solar fotovoltaica, los 3 métodos se comportan muy similar, ya sea considerando o no la división por subestación. Además se ve que, considerando 2 días representativos por mes para la generación eólica, el error es bastante menor que el error obtenido al utilizar la metodología de la CNE, y lo mismo al considerar cualquier número de días representativos por mes para la generación solar fotovoltaica.

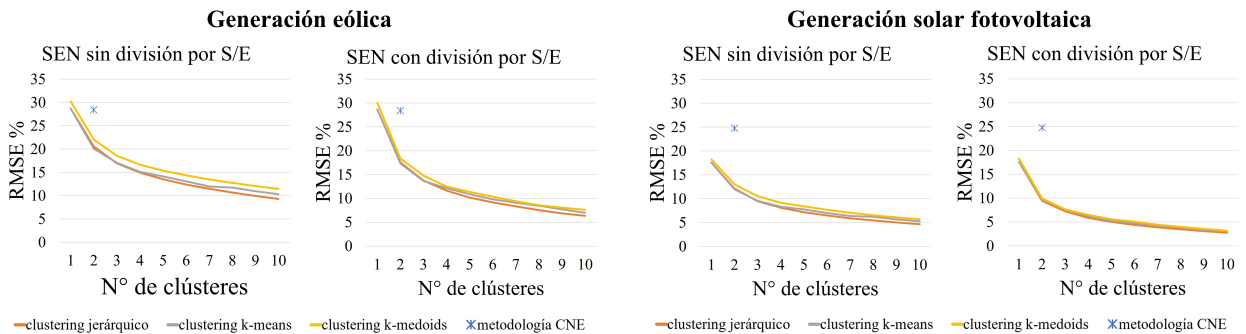


Figura 5.1: RMSE anual de la generación eólica y solar fotovoltaica según el número de clústeres, año 2040

Además, se analizaron de manera separada las subestaciones en el caso de considerar la división geográfica, obteniendo de manera independiente el mismo resultado que para todo el Sistema Eléctrico Nacional.

Si bien el rendimiento de los 3 métodos de *clustering* es bastante similar, tal como ya se mencionó, el peor rendimiento se obtiene a través del método *k-medoids* por su definición, ya que, para cada clúster, se escoge el día de la data real que estén más cerca del resto de los datos, mientras que en los otros dos métodos se calcula el promedio de cada clúster. Por otra

parte, según lo revisado en el Estado del Arte, el método de *clustering* jerárquico se utiliza, la mayoría de las veces, para ver las horas que son extremadamente similares entre sí, pudiendo provocar que se generen clústeres con una dimensión muy grande/pequeña, mientras que el método *k-means* se suele escoger por sobre el método jerárquico porque su tendencia es de producir clústeres de igual tamaño, evitando que los eventos aislados y muy distintos (como, por ejemplo, el día de Navidad, o cualquier día festivo) creen su propio clúster. A pesar de esto, se escoge el método de *clustering k-medoids*, ya que, para efectos de este trabajo, se requiere obtener como día representativo un día perteneciente a los datos reales, para evitar el efecto de “suavizamiento” de la curva. Además, según lo presentado en el Marco Teórico, este algoritmo es el más robusto a los *outliers* y al ruido.

5.2. División geográfica por subestación

A continuación, en la Figura 5.2, se presenta la Raíz del Error Cuadrático Medio anual obtenida con y sin división por subestación para el método de *clustering k-medoids*, además del error original con el método actual de la CNE. De aquí se observa que, tanto para la generación eólica como la solar fotovoltaica, se tiene un mejor rendimiento al considerar la división geográfica por subestación, por lo que se opta por este método.

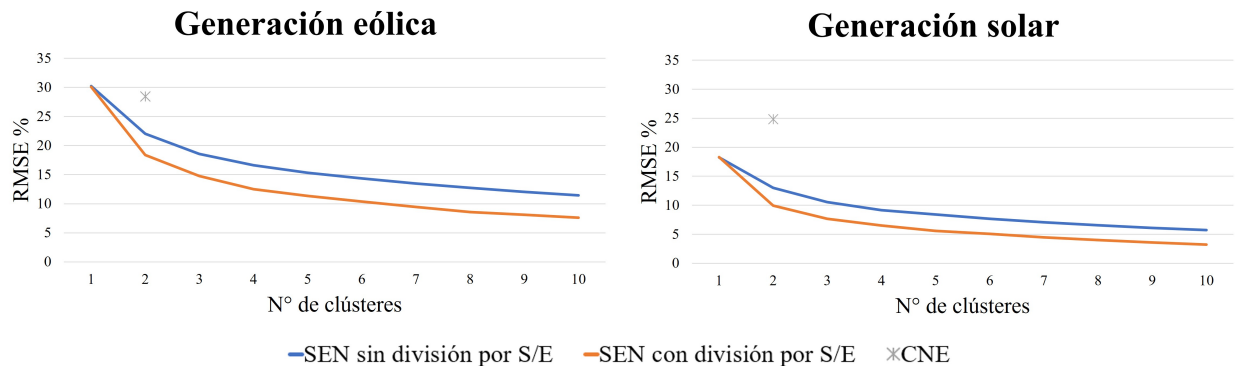


Figura 5.2: RMSE anual de la generación eólica y solar fotovoltaica según el número de clústeres, para el método de *clustering k-medoids*, año 2040.

Al momento de evaluar el Sistema Eléctrico Nacional en su conjunto, la asignación de clústeres está guiada por la generación total del sistema, por lo que no se asegura que esta asignación sea la mejor para todas las subestaciones. Por ejemplo, para el caso del viento, se esperaría un mejor ajuste a la subestación Parinas que a la subestación Río Malleco, ya que la primera representa hasta el 23 % del SEN, mientras que la segunda representa sólo hasta el 12 %. Por otro lado, para la generación solar fotovoltaica, la asignación de clústeres se ajustaría más a la subestación Kimal que a la subestación Polpaico, ya que la primera representa hasta el 38 % del SEN versus la segunda que representa sólo hasta el 6 %.

Para ejemplificar lo anterior, en la Figura 5.3 se presenta la adaptación de los días representativos para el mes de enero de 2040 para ambas tecnologías, con y sin división por subestación, al aplicar el método de *clustering k-medoids* en las subestaciones previamente mencionadas. A partir de estos gráficos se ve que los días representativos obtenidos con y sin división geográfica para las subestaciones Parinas y Kimal son más similares entre ellos que

los obtenidos para las subestaciones Río Malleco y Polpaico, incluso obteniendo un mismo día representativo con ambas metodologías para la subestación Kimal. Además, se ve de manera gráfica que los días representativos obtenidos sin división por subestación tienen un peor ajuste a los datos reales respecto a los días representativos obtenidos con división geográfica por subestación, ya que, por ejemplo, en la subestación Río Malleco las curvas naranjas representan menos a los días con baja generación eólica que las curvas azules, ocurriendo lo mismo en la subestación Polpaico, donde los días con baja generación solar fotovoltaica no quedan bien representados al no aplicar la división por subestación.

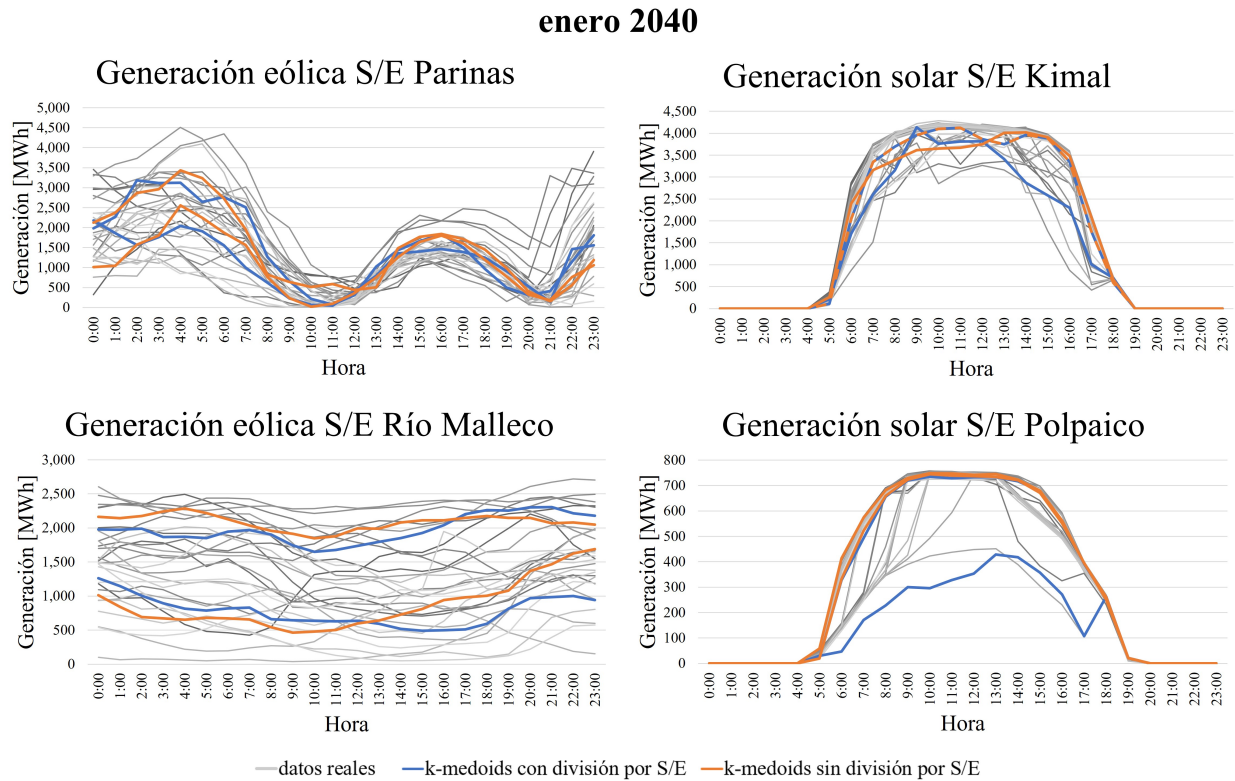


Figura 5.3: Adaptación de los días representativos con y sin división por subestación, aplicando el método *k-medoids* a los datos reales de generación eólica para las subestaciones Parinas y Río Malleco, y de generación solar fotovoltaica para las subestaciones Kimal y Polpaico (enero 2040).

5.3. Cantidad de días representativos por mes

Luego, al calcular la RMSE mensual del Sistema Eléctrico Nacional utilizando la división por subestación para el método del *clustering k-medoids*, se procede a escoger el número de clústeres por mes que se utilizará posteriormente en la simulación. Para esto, se compara el error calculado versus dos casos:

- El error mensual obtenido con la metodología actual de la CNE, de tal manera que se reduzca un 40 %, 50 % y un 60 %.
- Los siguientes valores fijos de RMSE: 15 %, 20 % y 25 %.

Tabla 5.1: Rango de días representativos necesarios por mes tanto para la generación eólica como la solar, para cumplir con los criterios definidos.

Criterio RMSE	Fijo			Proporcional al RMSE de la CNE		
	<25 %	<20 %	<15 %	-40 %	-50 %	-60 %
Días representativos generación eólica	1-3	2-4	3-8	2-4	3-7	4-10
Días representativos generación solar fv	1-2	1-2	1-3	1-2	1-3	2-4

En la Tabla 5.1 se muestra, de manera resumida, el rango de días representativos necesarios en cada mes considerado para cumplir con los objetivos mencionados, donde se observa que se necesitan más días representativos a medida que es más exigente el requerimiento y, en general, para la generación eólica por sobre la generación solar fotovoltaica. Además, se evidencia que se requiere distinta cantidad de días representativos por mes para tener el mismo rendimiento.

A modo de ejemplo, en las Figuras 5.4 y 5.5 se observa lo descrito anteriormente para los meses de enero y mayo de 2040, para la generación eólica y solar fotovoltaica, respectivamente.

En particular, para la generación eólica mostrada en la Figura 5.4, se tiene que, para disminuir un 40 % el RMSE actual de la CNE, se necesitan 2 días representativos para el mes de enero, mientras que para tener el mismo rendimiento, para el mes de mayo se necesita 1 día representativo adicional. Lo mismo ocurre para disminuir un 60 % el RMSE actual de la CNE, dado que en enero se necesitan 3 días representativos, mientras que en mayo se necesitan 4.

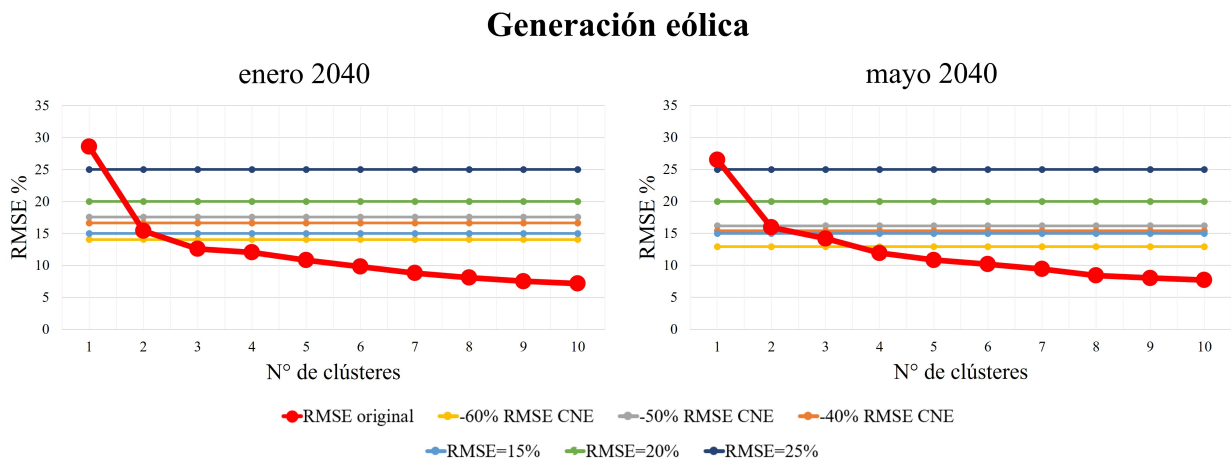


Figura 5.4: RMSE anual de la generación eólica según el número de clústeres, para el método de *clustering k-medoids*, para enero y mayo del año 2040, aplicando distintos requerimientos de error.

Generación solar

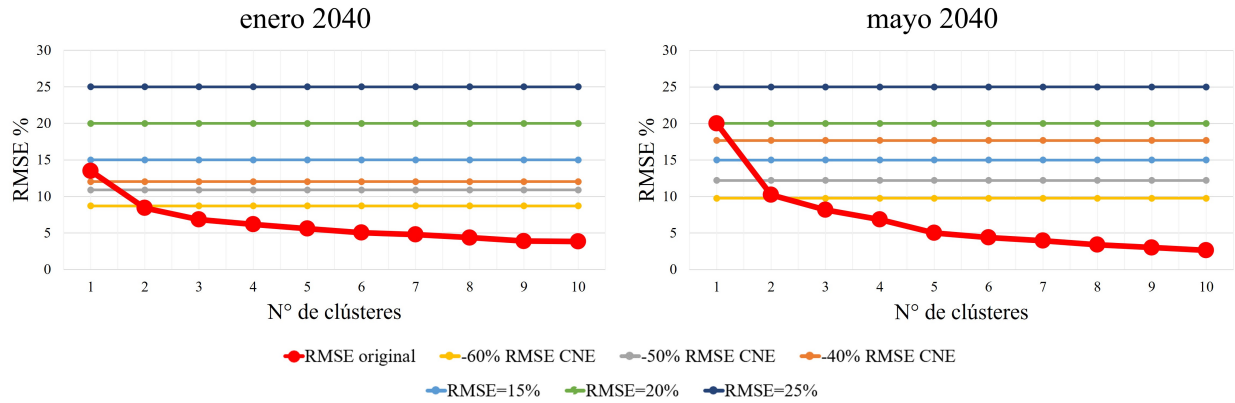


Figura 5.5: RMSE anual de la generación solar fotovoltaica según el número de clústeres, para el método de *clustering k-medoids*, para enero y mayo del año 2040, aplicando distintos requerimientos de error.

Respecto a la generación solar fotovoltaica de la Figura 5.5 se observa que, para cumplir con todos los requerimientos, se necesitan entre 1 a 2 días representativos para ambos meses, a excepción del requerimiento de disminución de un 60 % el RMSE actual de la CNE, donde en mayo del año 2040 se necesitan 3 días representativos versus enero del mismo año, donde se necesitan sólo 2 días representativos.

A pesar de que la cantidad de días representativos por mes, para cumplir con los requerimientos establecidos, puede variar de un mes a otro, por simplicidad, se considerará sólo una cantidad de días representativos para cumplir con cada objetivo, la cual consiste en la que tenga mayor frecuencia (si la frecuencia acumulada no es al menos 24 de los 48 meses, se considera el número siguiente de días representativos). Esto, porque podría ocurrir que, tal como ocurre con la generación eólica, para cumplir con un error menor al 25 %, sólo 3 de los 48 meses requieran de 1 día representativo, y que, a la vez, sólo 1 de los 48 meses requieran de 3 días representativos al mes, por lo que considerar 2 días representativos se podría considerar una buena aproximación para cumplir con ese criterio.

Tabla 5.2: Días representativos necesarios por mes, tanto para la generación eólica como la solar, para que al menos 24 de los 48 meses cumplan con los criterios definidos (el porcentaje de meses que cumplen con el criterio queda explicitado entre paréntesis).

Criterio RMSE	Fijo			Proporcional al RMSE de la CNE		
	<25 %	<20 %	<15 %	-40 %	-50 %	-60 %
Días representativos generación eólica	2 (98 %)	2 (67 %)	4 (81 %)	3 (81 %)	4 (60 %)	7 (71 %)
Días representativos generación solar fv	1 (73 %)	1 (58 %)	2 (81 %)	2 (100 %)	2 (92 %)	2 (52 %)

Luego, en la Tabla 5.2 se muestra la cantidad de días representativos necesaria por mes,

separada por tecnología, para que al menos un 50% de los meses cumplan con los requerimientos establecidos. Cabe mencionar que la cantidad de meses que cumplen con el criterio se presenta entre paréntesis, medida en porcentaje. De aquí se puede concluir, a modo general, que con 1 a 3 días representativos por mes se puede obtener un buen índice de error RMSE para ambos tipos de tecnología, al utilizar el método de *clustering k-medoids*.

5.3.1. Correlación entre zonas

Por último, en cuanto a la correlación entre zonas, y según lo establecido en la Tabla 5.2, en la Figura 5.6 se presenta la correlación para la generación eólica considerando 2 y 3 días representativos para cada uno de los 48 meses, mientras que en la Figura 5.7 se muestra lo mismo para la generación solar fotovoltaica, pero considerando 1 y 2 días representativos por mes. Además, en ambas Figuras se presenta la correlación para los datos horarios y al utilizar la metodología actual de la CNE, con el fin de realizar una comparación.

Luego, en la Figura 5.6 se observa que, al considerar 2 o 3 días representativos de generación

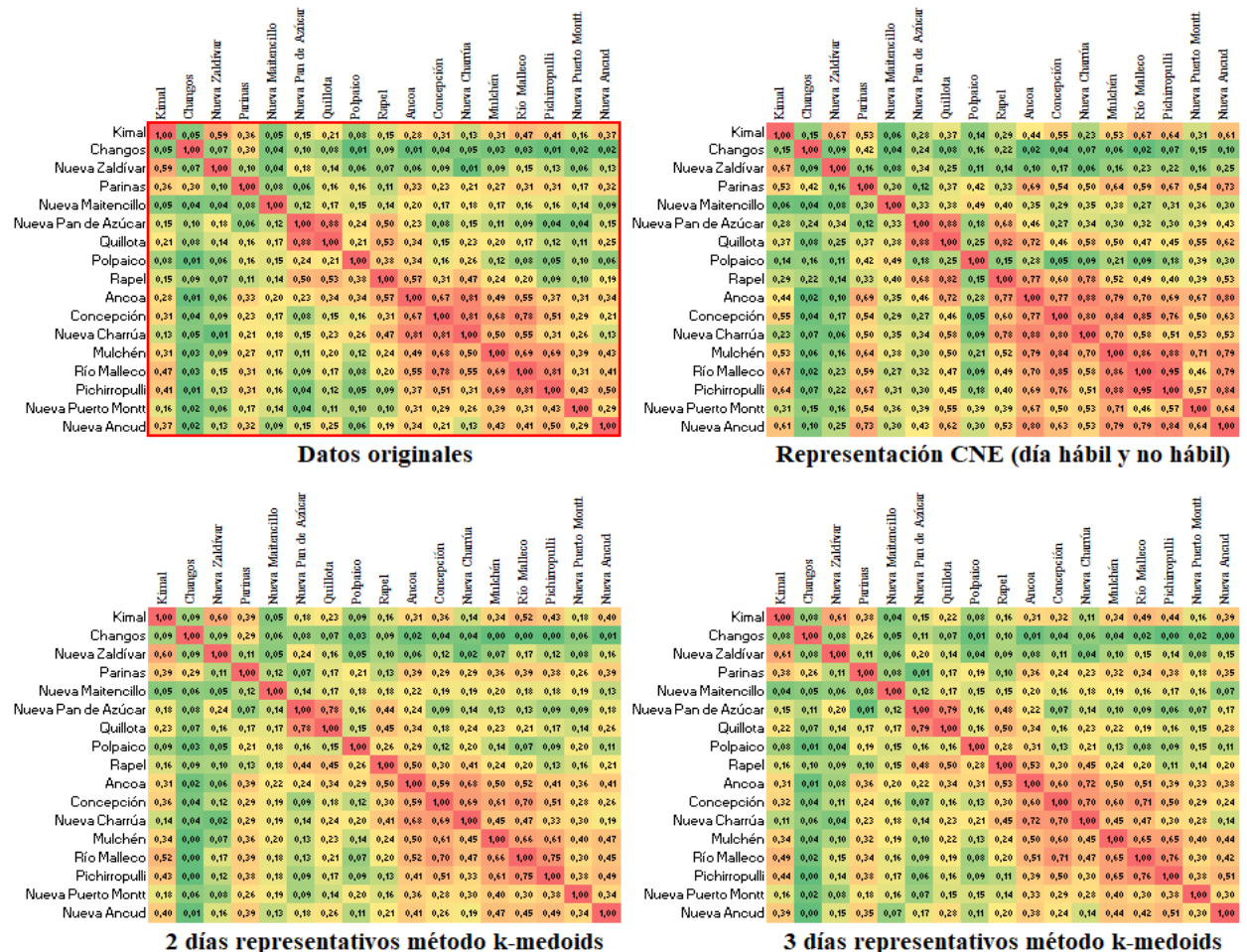


Figura 5.6: Matriz de correlación entre zonas de los datos horarios de 2040 versus la correlación entre zonas utilizando la metodología de la CNE y el método de *clustering k-medoids* considerando dos y tres días representativos al mes, para los datos de viento.

eólica por mes, la correlación entre subestaciones es muy similar a la de los datos horarios de generación eólica, presentando un mejor comportamiento respecto a la metodología de la CNE, con una tendencia a suavizar la correlación entre zonas más que acentuarla. Por otro lado, en la Figura 5.7, se observa que, al considerar sólo un día representativo de generación solar fotovoltaica al mes, sigue una tendencia más similar a la correlación entre zonas obtenida utilizando la representación actual de la CNE, mientras que, al considerar una mayor cantidad de días representativos, la correlación es similar a la de los datos horarios de generación solar.

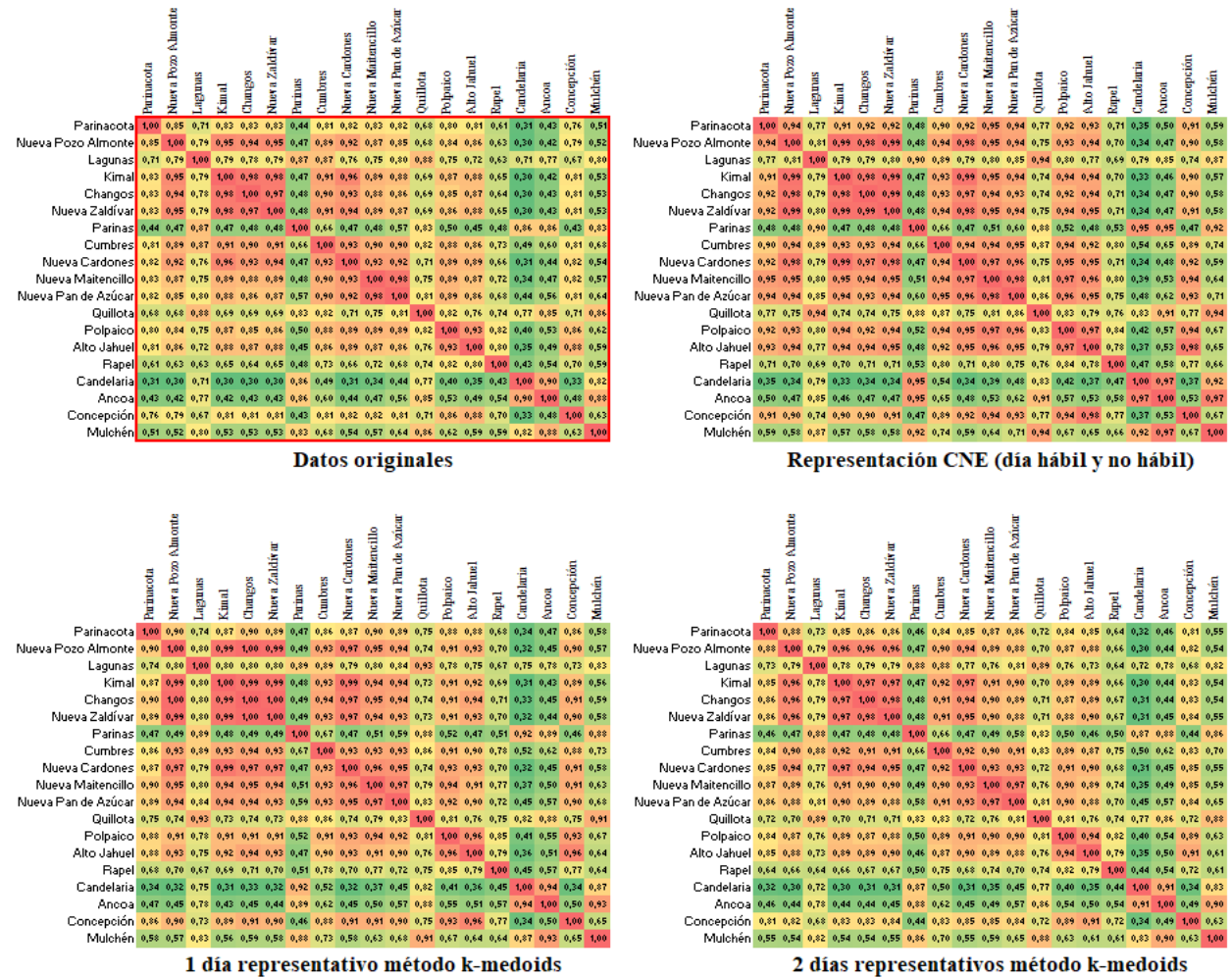


Figura 5.7: Matriz de correlación entre zonas de los datos horarios de 2040 versus la correlación entre zonas utilizando la metodología de la CNE y el método de *clustering k-medoids* considerando uno y dos días representativos al mes, para los datos de sol.

Cabe recalcar que mantener la correlación original entre subestaciones es importante, debido a las congestiones en las líneas de transmisión que se pueden presentar entre zonas, ya que puede ocurrir que si una línea está congestionada, por ejemplo la línea que conecta Ancoa con Alto Jahuel, al estar correlacionada alguna de estas subestaciones con otra subestación, con la cual no tenía una fuerte correlación originalmente, podrían presentarse congestiones en la transmisión que en realidad no existen.

De esta manera se puede definir la cantidad de días representativos necesarios por tecnología para proceder con la simulación de la operación del SEN en Ameba, al tener como objetivo, entre otros, mantener la correlación entre zonas lo más similar posible a la obtenida para los datos horarios. De esta manera, para ambas tecnologías, se necesitan al menos 2 días representativos por mes.

Luego, en las Figuras 5.8, 5.9, 5.10 y 5.11 se visualiza la adaptación de los días representativos utilizando el método de *clustering k-medoids* con división geográfica por subestación, considerando que el indicador RMSE respecto a los datos horarios sea menor al 25 % en, al menos, un 50 % de los meses, y manteniendo la correlación original, en particular para las subestaciones Parinas y Río Malleco para la generación eólica, y las subestaciones Kimal y Polpaico para la generación solar fotovoltaica, respectivamente, para los meses de enero y junio de 2040. Cabe mencionar que, además, se grafica el comportamiento de los días representativos hábil y no hábil de la CNE, con el fin de realizar una comparación del rendimiento.

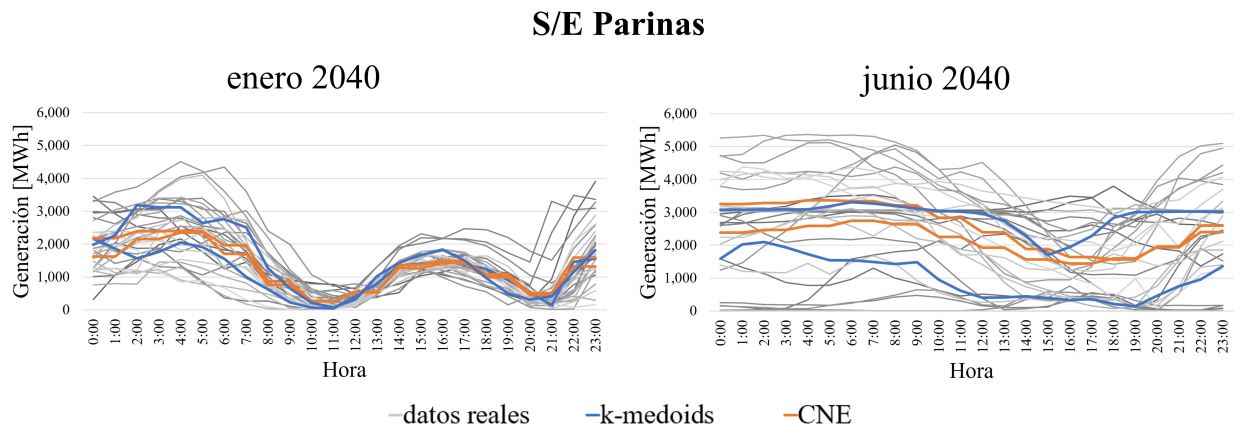


Figura 5.8: Adaptación de los días representativos de la generación eólica obtenidos con el *clustering k-medoids*, en contraste con los días hábil y no hábil de la CNE, para la subestación Parinas (enero y junio 2040).

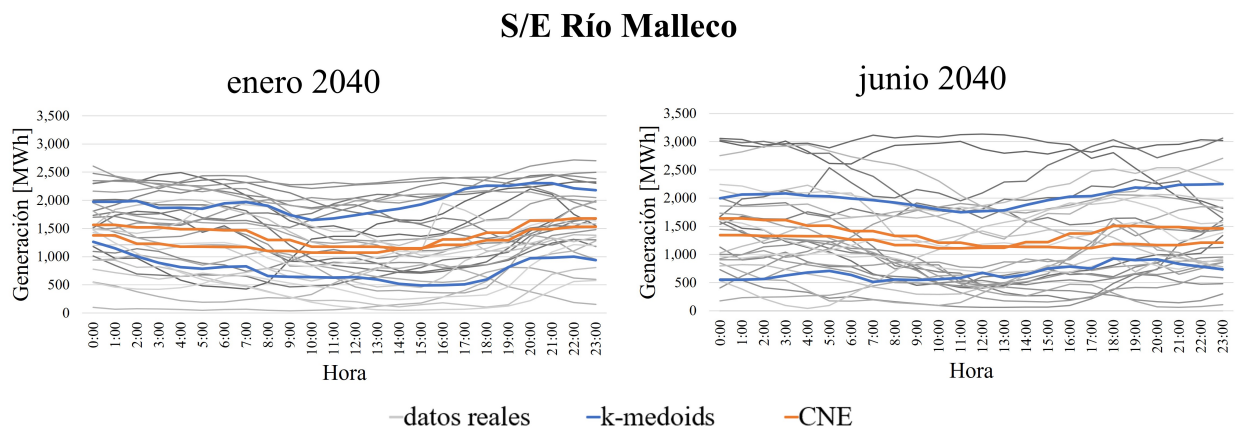


Figura 5.9: Adaptación de los días representativos obtenidos con el *clustering k-medoids* a los datos reales de generación eólica, en contraste con los días hábil y no hábil de la CNE, para la subestación Río Malleco (enero y junio 2040).

A partir de la Figura 5.8 se observan los dos días representativos necesarios para cumplir con los criterios recién mencionados, para la subestación Parinas, a partir de la cual se infiere que efectivamente hay una mejor adaptación de los días representativos a la data real de generación eólica respecto al día hábil y no hábil de la CNE, y lo mismo ocurre con la subestación Río Malleco en la Figura 5.9, donde se observa que la metodología actual de la CNE no representa de buena manera los días de alta y baja generación eólica en enero del año 2040, centrándose en los días de generación eólica promedio.

Luego, en la Figura 5.10 se observan los dos días representativos necesarios para cumplir con los criterios recién mencionados para la subestación Kimal, a partir de la cual se puede ver una mejor adaptación de los días representativos a la generación solar fotovoltaica real obtenidos con el método de *clustering k-medoids*. Si bien la metodología actual de la CNE

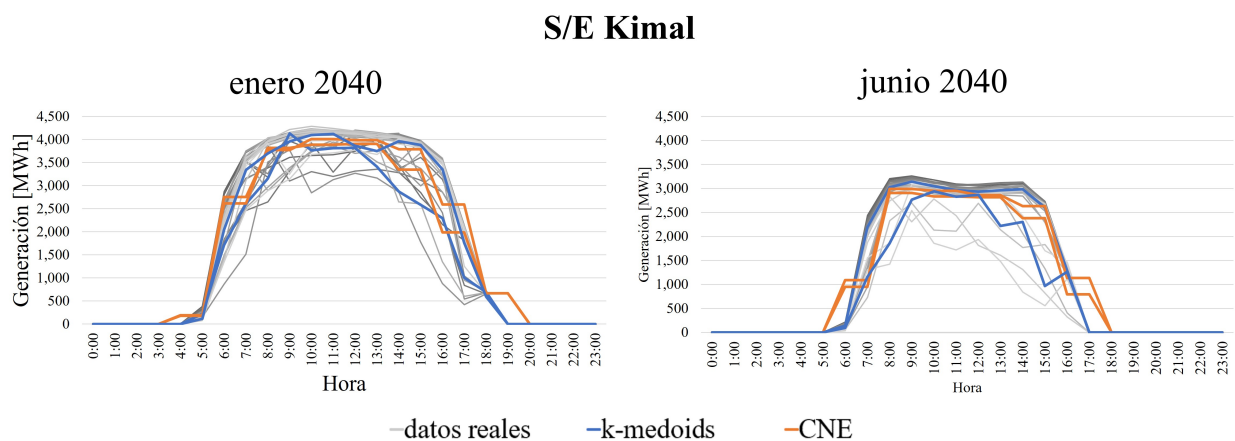


Figura 5.10: Adaptación de los días representativos obtenidos con el *clustering k-medoids* a los datos reales de generación solar fotovoltaica, en contraste con los días hábil y no hábil de la CNE, para la subestación Kimal (enero y junio 2040).

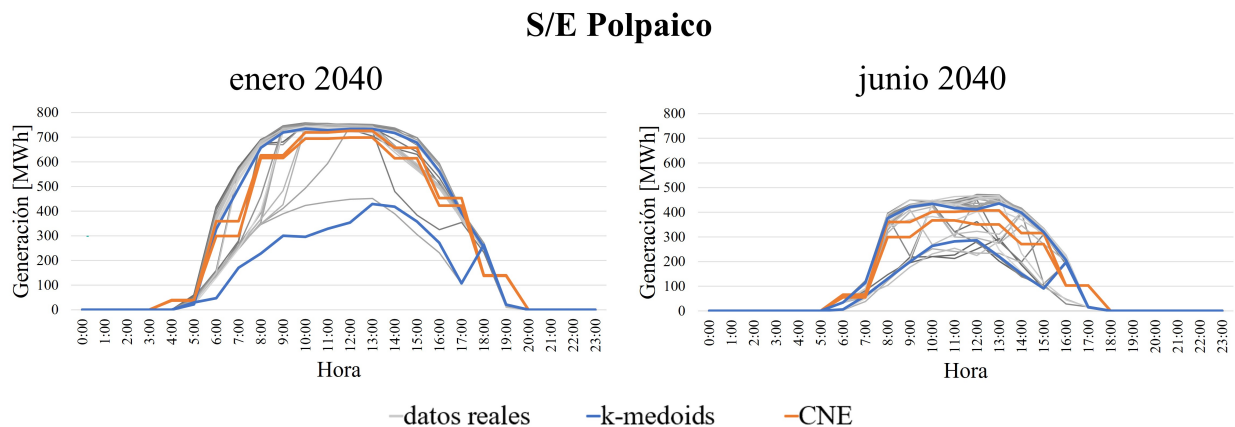


Figura 5.11: Adaptación de los días representativos obtenidos con el *clustering k-medoids* a los datos reales de generación solar fotovoltaica, en contraste con los días hábil y no hábil de la CNE, para la subestación Polpaico (enero y junio 2040).

representa de buena forma la mayoría de los días (donde se ve una mayor concentración de curvas grises), deja de lado los días “extremos”, mientras que los días representativos obtenidos con el *clustering k-medoids* son capaces de representar de mejor manera los días de baja generación solar fotovoltaica presentes tanto en enero como junio de 2040. Lo mismo ocurre en la subestación Polpaico mostrada en la Figura 5.11, donde el día representativo de baja generación solar fotovoltaica obtenido con el *clustering k-medoids* permite representar estos días extremos, lo cual es importante considerar para la posterior simulación de la operación.

5.4. Simulación de la operación en Ameba

Al realizar la simulación de la operación utilizando la plataforma web Ameba, se obtienen distintas salidas, entre ellas los costos marginales y el comportamiento del sistema, que incluye la generación horaria a partir de distintas tecnologías para cada año y los costos de operación, entre otros. Cabe mencionar que en la simulación se consideran 3 escenarios distintos de hidrología: un escenario seco (H1), medio (H2) y húmedo (H3).

Para efectos de este trabajo, y con el fin de obtener un análisis más detallado del efecto de considerar distinta cantidad de días representativos mensuales para la generación eólica y solar fotovoltaica, y poder comparar las distintas metodologías, es que se procedió a simular los siguientes casos, considerando:

- 1 día representativo al mes³ para el método de *clustering k-medoids* (24 bloques).
- 2 días representativos al mes para el método de *clustering k-medoids* (48 bloques).
- 3 días representativos al mes para el método de *clustering k-medoids* (72 bloques).
- El día representativo hábil y no hábil mensual para el método de la CNE (24 bloques).
- Los datos horarios no agregados.

Un punto importante a tener en consideración es, que la generación total (sin considerar vertimientos) será la misma para todas las metodologías e hidrologías, ya que sigue la demanda, la cual es un dato de entrada del modelo de simulación. Lo único que se esperaría que cambie es la generación eólica y solar fotovoltaica, lo cual trae como consecuencia que se vea afectada la generación a partir de otras tecnologías para cumplir con este total de generación, y por ende, que cambien los costos marginales, los flujos en las líneas de transmisión, etc. Esta generación total del sistema por año se observa en la Figura 5.12.

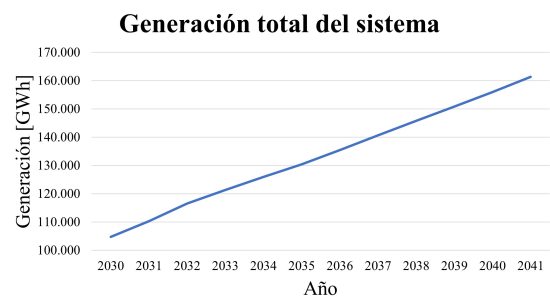


Figura 5.12: Generación total anual del sistema.

³ Hay que tener en consideración que con 1 día representativo al mes no se mantiene bien la correlación entre zonas estudiada anteriormente.

5.4.1. Costo de operación del sistema

En la Figura 5.13 se presenta el comportamiento esperado del costo de operación del sistema eléctrico a lo largo de los años, tanto para un escenario de hidrología seca como para un escenario de hidrología húmeda, considerando la metodología actual de agregación temporal de la CNE, el método de *clustering k-medoids* con 1, 2 y 3 días representativos por mes, y el caso horario real. De aquí se observa que, para todas las metodologías, la tendencia, es decir, la forma de la curva, es similar, a excepción de la curva real horaria (curva roja), la cual, con el pasar de los años, presenta un aumento significativo, llegando a aumentar más del doble en 10 años. Además, se observa que el menor costo se obtiene con la metodología actual de la CNE, correspondiendo así a la metodología que más se aleja de la realidad.

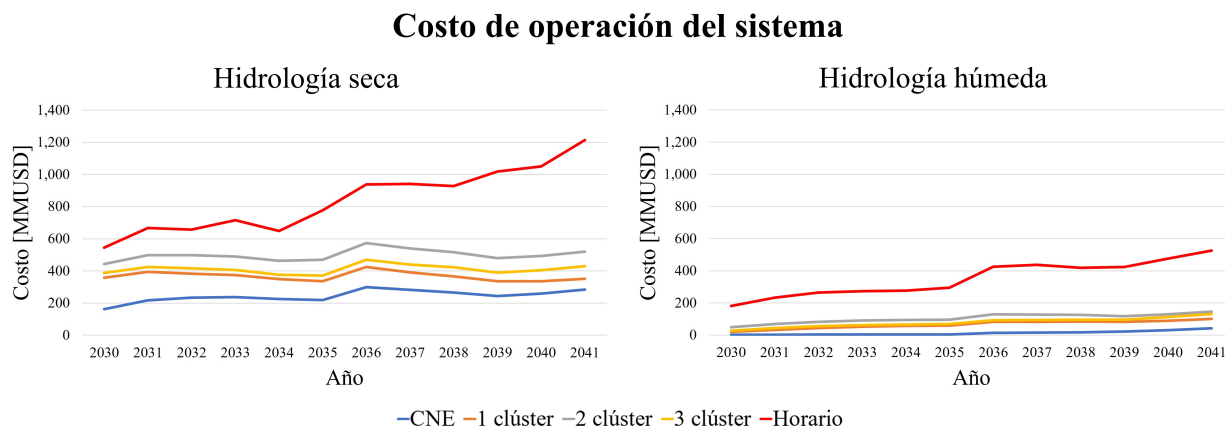


Figura 5.13: Costo de operación anual del sistema considerando los escenarios de hidrología seca y húmeda, para todas las metodologías analizadas.

Este comportamiento distinto entre la metodología actual de la CNE y el *clustering k-medoids*, respecto al caso horario, se puede deber a varias razones, las cuales se revisarán a continuación. Entre ellas, puede estar ocurriendo que la generación eólica prevista a lo largo de los años sea mucho mayor en estos métodos respecto al caso real, aminorando los costos, o que la generación convencional sea menor, entre otros.

5.4.2. Costos marginales

En cuanto a los costos marginales, en la Figura 5.14 se presenta una evolución de los costos marginales en las barras a medida que pasan los años para la simulación horaria, en particular considerando un escenario de hidrología seca. De aquí se puede observar cómo éstos aumentan con el tiempo, lo cual es consistente con el aumento del costo de operación del sistema. Luego, en la Figura 5.15 se tienen los costos marginales por barra para el año 2040, obtenidos con las distintas metodologías de agregación temporal, y comparados con el caso horario, obteniendo lo esperado, donde los costos marginales son menores al agregar temporalmente los datos, independiente de la metodología utilizada. Cabe mencionar que, si bien los resultados para las otras hidrologías no son exactamente los mismos, sí es igual la conclusión obtenida, que corresponde a que el comportamiento de los costos marginales es un reflejo de la evolución del costo operacional del sistema eléctrico con el tiempo, pero no explica el origen de esta diferencia.

Costos marginales por barra - hidrología seca (simulación horaria)

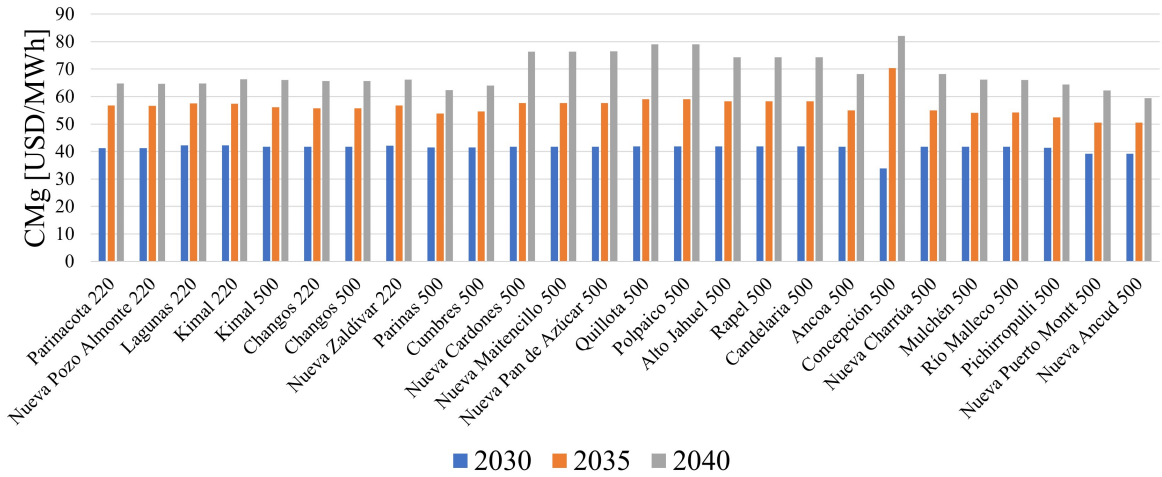


Figura 5.14: Costos marginales por barra, considerando el escenario de hidrología seca, para los datos reales no agregados.

Costos marginales por barra - hidrología seca año 2040 (simulación horaria)

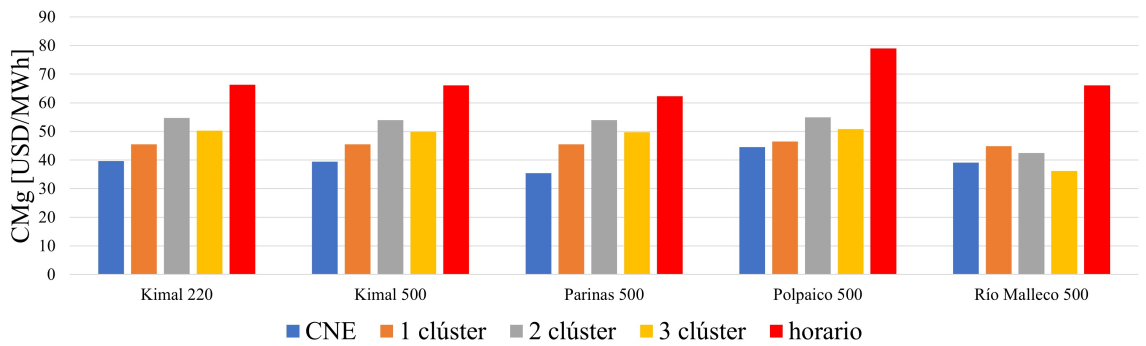


Figura 5.15: Costos marginales por barra para las distintas metodologías de agregación temporal versus el caso horario, considerando el escenario de hidrología seca para el año 2040.

5.4.3. Generación por tecnología

Generación por tecnología para el año 2040, hidrología seca

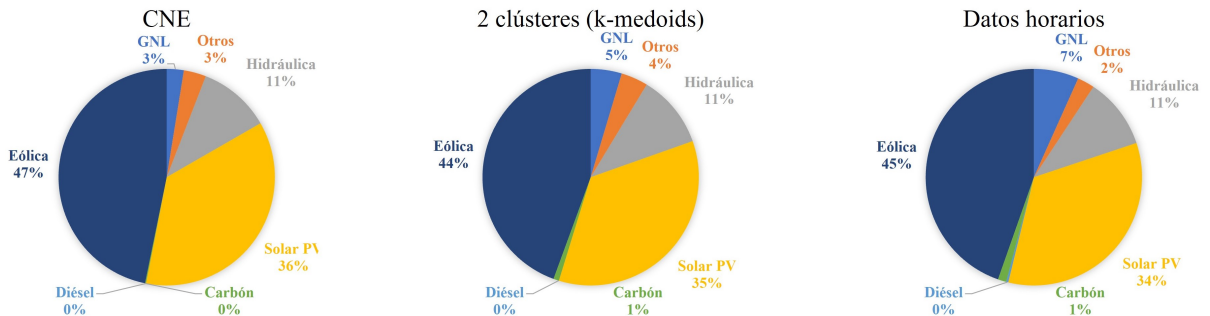


Figura 5.16: Generación por tecnología para el año 2040, considerando el escenario de hidrología seca, para todas las metodologías analizadas.

En la Figura 5.16, se presenta la generación dividida por tecnología para el año 2040, en particular para un escenario de hidrología seca, a partir de lo cual se observa que, a simple vista, no hay gran diferencia en la distribución de la generación para los distintos casos.

Sin embargo, para entender la diferencia producida en el costo de operación, se procede a analizar los sectores del gráfico de manera separada. Cabe mencionar que en la Figura 5.16 se muestra solamente el caso con dos días representativos para el método de *clustering k-medoids*, debido a que los resultados son muy similares al considerar 1 o 3 días representativos al mes.

5.4.3.1. Generación a carbón, GNL y diésel

En la Figura 5.17 se observa la generación a carbón, GNL (gas natural licuado) y diésel prevista a lo largo de los años, donde se ve que, tanto para la metodología de la CNE como para la metodología *k-medoids*, la generación con tecnología convencional es menor al caso horario. En particular, destaca la generación a partir de diésel, tecnología con la cual no se está generando al considerar las metodologías de agregación temporal, pero sí se genera de manera creciente en el caso horario a medida que pasan los años, lo cual, a pesar de no ser una gran cantidad de energía, podría explicar el aumento de los costos de operación, ya que corresponde a la tecnología con mayor costo variable de generación (seguida por el GNL y luego por el carbón). Esto se puede comprobar al comparar la Figura 5.13 con la Figura 5.17, donde se ve claramente que el aumento de los costos de operación a lo largo de los años sigue una tendencia similar a la generación a partir de diésel.

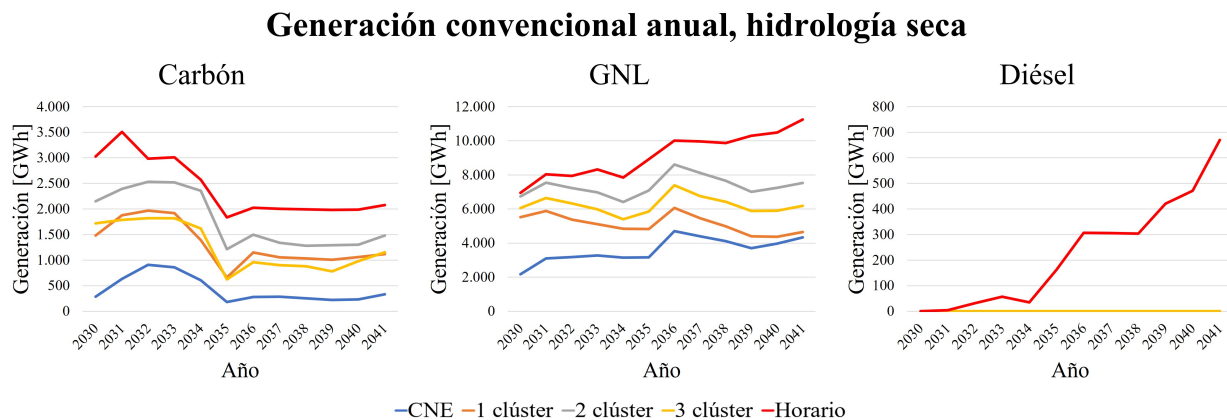


Figura 5.17: Generación anual convencional a carbón, GNL y diésel considerando el escenario de hidrología seca, para todas las metodologías analizadas.

Al graficar en la Figura 5.18 el comportamiento mensual de la generación convencional versus el costo de operación del sistema, en específico para el año 2040, en el cual se tiene una gran diferencia de costo de operación al comparar el caso horario con las metodologías de agregación temporal, se observa que el costo de operación tiene un comportamiento similar a la generación a partir de GNL para los métodos de agregación temporal, la cual corresponde a la tecnología convencional con la que más se genera, siendo la segunda más cara después del diésel. Sin embargo, para el caso horario, se puede ver que existe una leve diferencia, producto de la generación a partir de diésel previamente mencionada.

Costo de operación del sistema versus generación convencional, hidrología seca año 2040

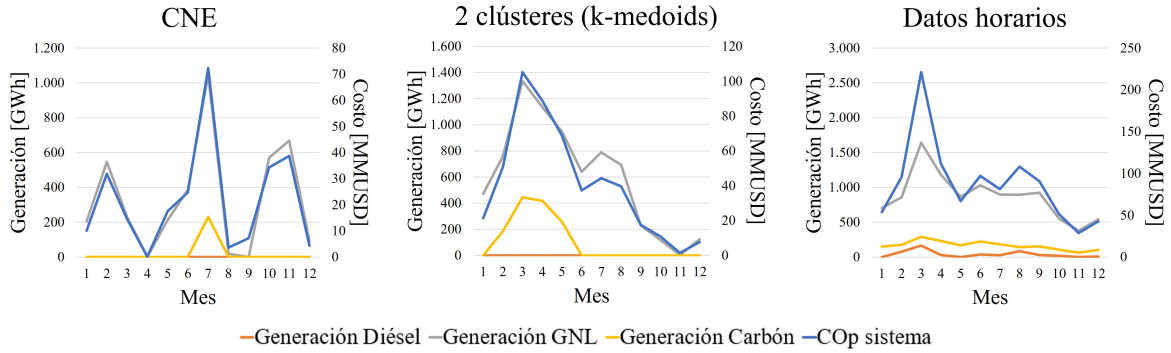


Figura 5.18: Generación convencional mensual para el año 2040, considerando el escenario de hidrología seca, para todas las metodologías analizadas.

5.4.3.2. Generación eólica

En la Figura 5.19 se presenta el comportamiento de la generación eólica a lo largo de los años, tanto para un escenario de hidrología seca como para un escenario de hidrología húmeda, considerando la metodología actual de agregación temporal de la CNE, el método de *clustering k-medoids* con 1, 2 y 3 días representativos por mes, y el caso horario real. De aquí se ve cómo constantemente la generación eólica aumenta con los años, y es mayor la obtenida con la metodología de la CNE en todos los años, por lo que se esperaría que el costo de operación sea menor con esta metodología (lo cual, tal como se vio anteriormente, si se cumple). Sin embargo, con los datos reales, la generación eólica no es significativamente menor como para que los costos sean tan altos. Es más, es muy similar a la generación eólica obtenida con 1 y 3 clústeres al mes.

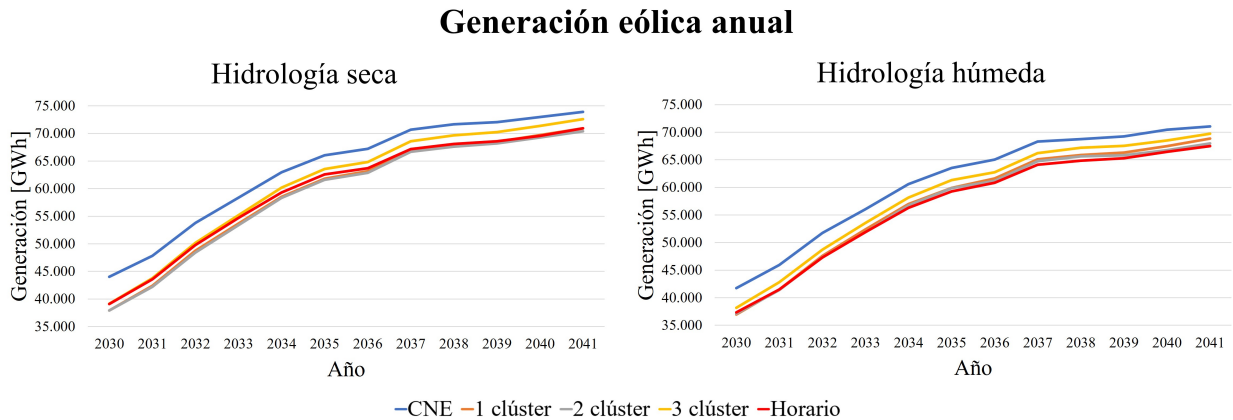


Figura 5.19: Generación eólica anual considerando los escenarios de hidrología seca y húmeda, para todas las metodologías analizadas.

Luego, en el gráfico de la Figura 5.20 se muestra el costo de operación del sistema versus la generación ERNC (que corresponde a la generación eólica más la generación solar fotovoltaica) para el año 2040, considerando una hidrología seca. A partir de aquí se ve que se mantiene una relación inversa entre estos dos vectores, que es lo esperado según la teoría.

Además, en la Figura 5.20 se observa que el comportamiento, tanto de los costos dentro

del año, así como el de la generación ERNC, es mucho más similar al real al aplicar el método *k-medoids* respecto al método actual de la CNE. Con los datos reales, se tiene un mínimo de generación ERNC en abril, lo cual coincide con el método *k-medoids*, mientras que el mínimo de generación con el método de la CNE, y por ende, el peak de costo, se obtiene en el mes de julio, alejándose de la realidad. Además, se nota que los valores de costo que toman los datos horarios son mayores al método *k-medoids*, a pesar de seguir una forma bastante similar, es decir, a pesar de estar bien correlacionados.

Costo de operación del sistema versus generación ERNC, hidrología seca año 2040

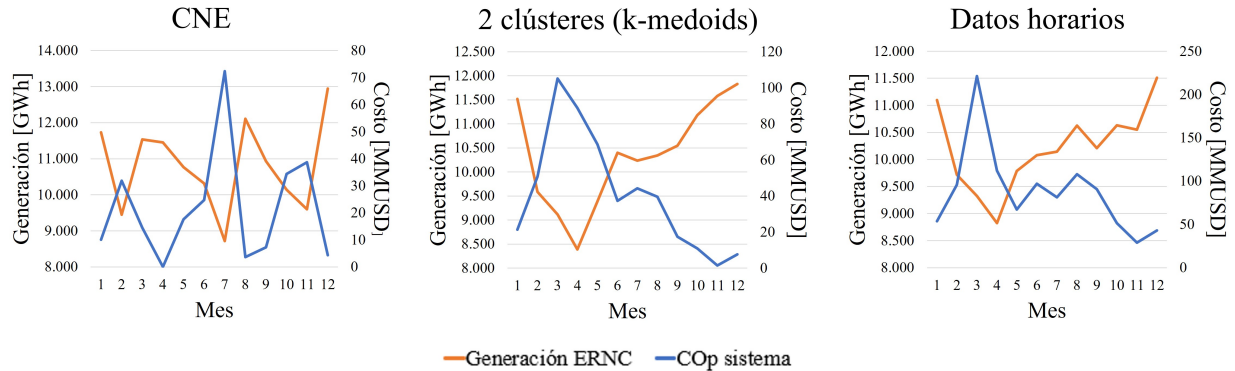


Figura 5.20: Generación ERNC mensual para el año 2040, considerando el escenario de hidrología seca, para todas las metodologías analizadas.

Luego, en la Figura 5.21 se ve la generación eólica horaria para el mes de enero, abril, agosto y diciembre de 2040. De aquí se observa que, para todos los meses, la generación eólica horaria de todos los métodos es bastante similar, a excepción de la obtenida por la CNE. En particular, para los meses de abril, agosto y diciembre del año 2040, la CNE obtiene una generación eólica bastante mayor a la de los otros métodos, lo cual efectivamente se refleja en los costos, ya que es el método con menor costo de operación en esos meses.

Generación eólica horaria, hidrología seca

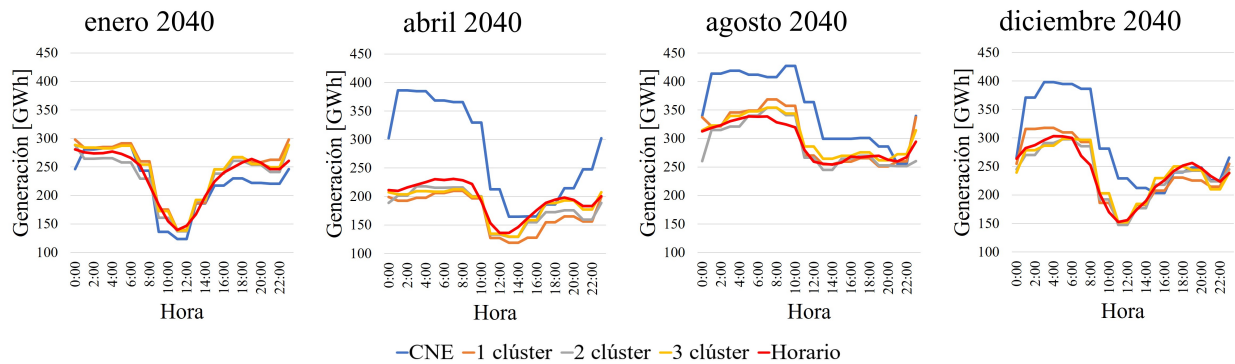


Figura 5.21: Generación eólica anual considerando los escenarios de hidrología seca y húmeda, para todas las metodologías analizadas.

A pesar de las diferencias analizadas para la metodología de la CNE, la previsión de generación eólica obtenida para el método de *clustering k-medoids*, independiente del número de días representativos escogido, es muy similar a la obtenida con los datos horarios reales,

por lo que esto no estaría explicando la gran diferencia de costos de operación.

Luego, en la Figura 5.22 se observa gráficamente que los días donde realmente hay menor generación eólica, hay un mayor costo de operación del sistema eléctrico, el cual aumenta con una tendencia exponencial a medida que la generación eólica es menor.

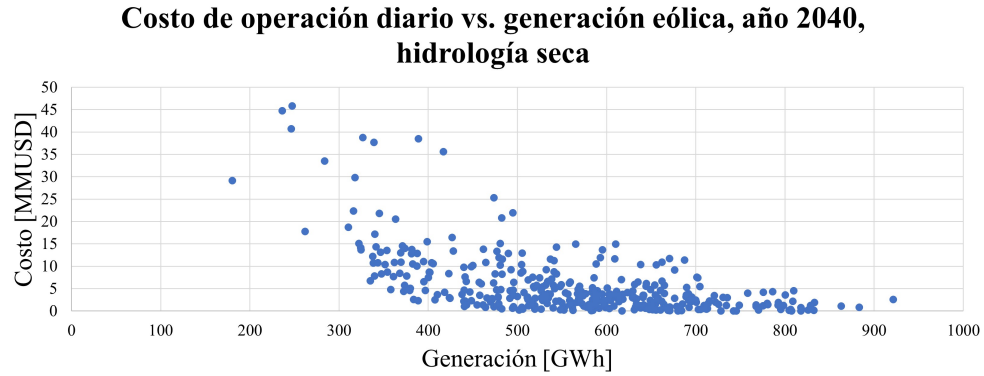


Figura 5.22: Costo de operación versus generación eólica diaria para el año 2040, considerando el escenario de hidrología seca, para los datos reales no agregados.

Por otro lado, en la Figura 5.23 se observa la misma comparación anterior, pero de manera mensual para los meses de marzo y agosto del año 2040, donde además se marcan los días representativos utilizados al reducir temporalmente los datos con el método de *clustering k-medoids*. De aquí se ve cómo los días representativos no son capaces de captar los días de los datos reales donde los costos de operación son altos, lo cual se debe a que no son capaces de representar los días de mínima generación renovable, tanto eólica como solar fotovoltaica, tal como se presentó anteriormente en las Figuras 5.8, 5.9, 5.10 y 5.11, donde podría ser necesario generar con tecnología convencional más costosa, como el diésel.

Costo de operación diario vs. generación eólica, con y sin agregación temporal, hidrología seca

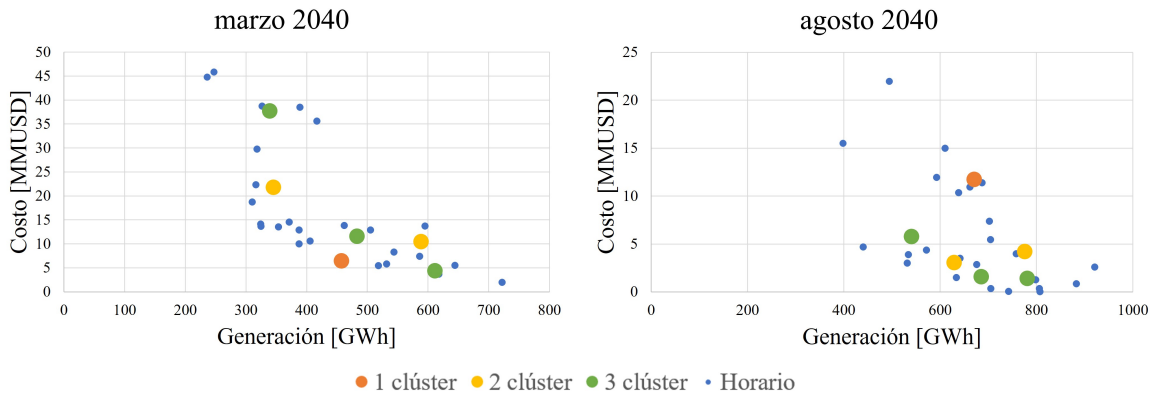


Figura 5.23: Costo de operación versus generación eólica diaria para marzo y agosto del 2040, sin y con agregación temporal utilizando el método *k-medoids*, considerando el escenario de hidrología seca.

5.4.4. Corrección del método de *clustering k-medoids* basado en series de tiempo

Con la intención de lograr representar de mejor manera la generación convencional, y por ende los costos operacionales, al agregar temporalmente los datos, es que se plantean nuevas metodologías compuestas de 3 días representativos, donde los 2 primeros correspondan a los obtenidos con el método de *clustering k-medoids* al considerar 2 clústeres al mes, y para el tercer día se consideran los siguientes casos:

- Caso 1: Se agrega el día de menor generación renovable de cada mes para cada perfil, el cual representa la misma cantidad de días al mes que el clúster 3 del método *k-medoids*.
- Caso 2: Nuevamente se agrega el día de menor generación renovable de cada mes para cada perfil, pero esta vez representa a un día no hábil y 2 días hábiles del mes.
- Caso 3: Se busca el día donde se obtiene un mayor costo operacional del sistema eléctrico, y se agrega este día para la generación renovable, siendo independiente de la división geográfica. Se utiliza el mismo porcentaje de representación mensual obtenido para el clúster 3 del método *k-medoids*
- Caso 4: Se busca el día donde se obtiene un una menor generación eólica a nivel de sistema eléctrico, y se agrega este día para la generación renovable, siendo independiente de la división geográfica. Al igual que el caso 1 y 3, se utiliza el mismo porcentaje de representación mensual obtenido para el clúster 3 del método *k-medoids*.

Es importante mencionar que se espera que el caso 1 tenga el costo de operación más alto, debido a que es aplicado considerando la división geográfica, y a diferencia del caso 2, tiene un porcentaje de representación más alto. Además, para poder aplicar el caso 3, es necesario tener previamente los resultados de la simulación horaria, lo cual no es viable en la realidad si se quiere evitar los altos tiempos de simulación y teniendo en consideración el objetivo de aplicar un método de agregación temporal. También hay que tener en cuenta que para los casos 1 y 2, donde se considera el mínimo de generación renovable por subestación, se asume algo que no ocurriría en la realidad, dado que es poco probable que exista un día al mes donde se produzca el mínimo de generación de manera simultánea en todas las zonas del país. De esta forma, estos casos corresponden sólo a casos de estudio para dar puntapié inicial al planteamiento de una metodología que no se base puramente en métodos de *clustering*, sino que considere otros indicadores y criterios adicionales.

Dicho esto, en la Figura 5.24 se muestra cuál es el día que se estaría agregando al aplicar el caso 3 (día de máximo costo de operación) y 4 (día de mínima generación eólica), en particular para los meses de marzo y agosto de 2040. Este procedimiento es análogo para el resto de los meses y años.

Costo de operación diario vs. generación eólica, con y sin agregación temporal, hidrología seca

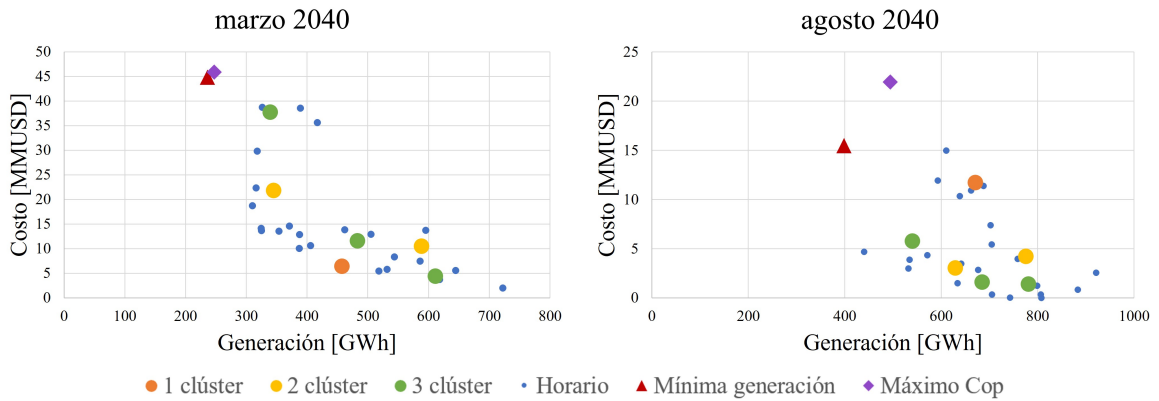


Figura 5.24: Costo de operación versus generación eólica diaria para marzo y agosto del 2040, con y sin agregación temporal utilizando los métodos propuestos, considerando el escenario de hidrología seca.

Luego, en la Figura 5.25 se presenta el costo de operación anual del sistema eléctrico al aplicar las nuevas metodologías propuestas, además de seguir considerando los casos principales analizados. De aquí se ve una mejor aproximación en cuanto a costos de los casos 1-4 al compararlos con los resultados horarios, obteniendo los mayores costos para el caso 1 tal como se mencionó anteriormente, seguido de los casos 2, 4 y 3, pero teniendo un distinto rendimiento según la hidrología que se considere (ya que, para una hidrología seca, el caso 1 tiene costos de operación mucho más altos que el caso horario, mientras que para una hidrología húmeda son muy similares). Además, se ve que entre los casos 3 y 4 no hay gran diferencia.

Costo de operación del sistema

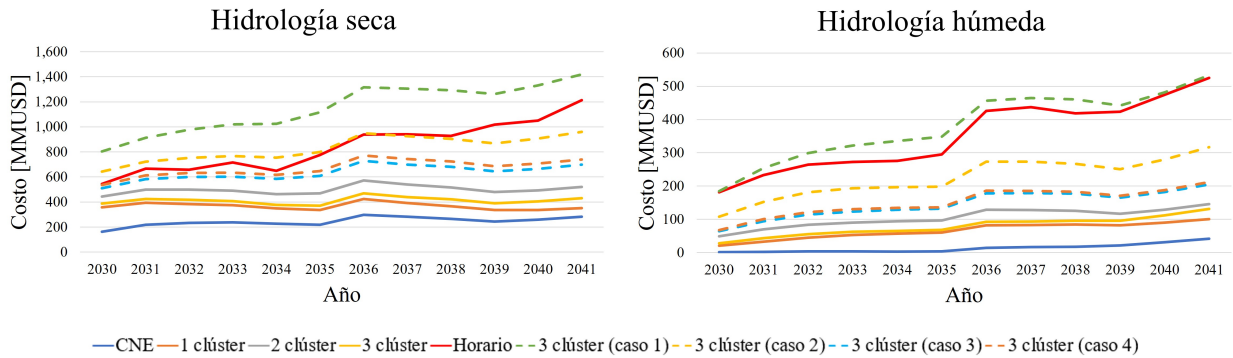


Figura 5.25: Costo de operación anual del sistema considerando los escenarios de hidrología seca y húmeda, para todas las metodologías analizadas (incluyendo las propuestas).

Para revisar de qué está compuesto este mayor costo operacional, y ver si realmente se logra mejorar la generación a diésel, se grafica la generación convencional anual en la Figura 5.26 para una hidrología seca. A partir de esta imagen se tiene que el aumento del costo de operación del sistema se debe a un aumento en la generación a carbón y GNL, sobrepasando al caso horario, mientras que no se produce un aumento en la generación a diésel. Algo similar ocurre para la hidrología húmeda mostrada en la Figura 5.27, donde se explica el aumento

en el costo de operación debido a una mayor generación convencional a carbón y GNL, pero no a diésel.

Generación convencional anual, hidrología seca

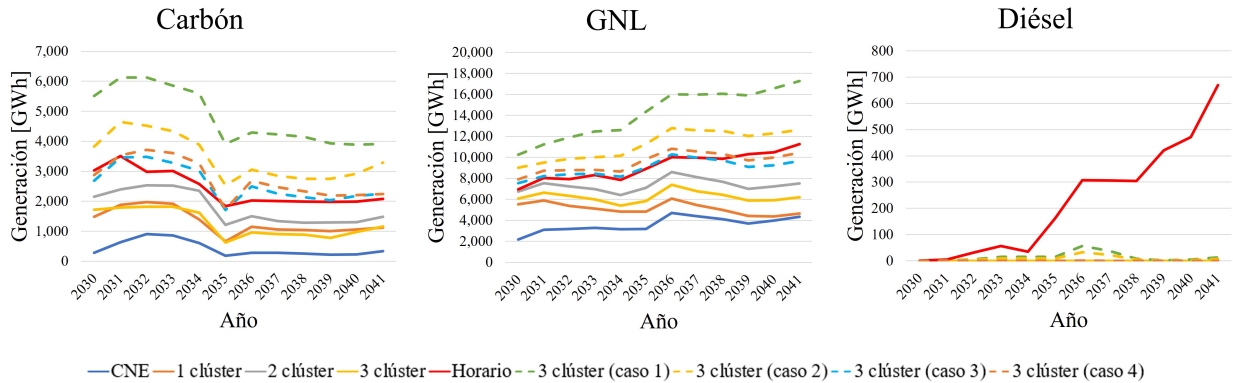


Figura 5.26: Generación anual convencional a carbón, GNL y diésel considerando el escenario de hidrología seca, para todas las metodologías analizadas (incluyendo las propuestas).

Generación convencional anual, hidrología húmeda

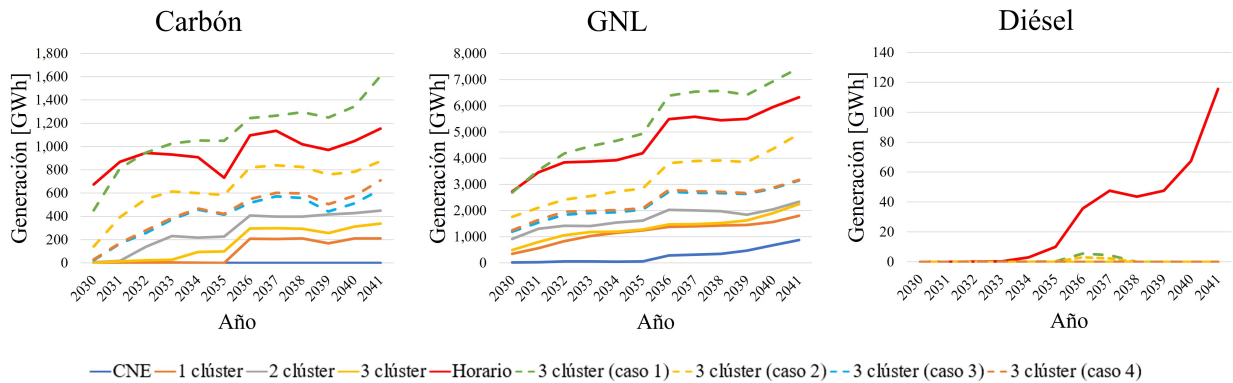


Figura 5.27: Generación anual convencional a carbón, GNL y diésel considerando el escenario de hidrología húmeda, para todas las metodologías analizadas (incluyendo las propuestas).

Luego, el hecho que no se genere con centrales a diésel se debe a la agregación temporal propiamente tal, que no permite captar en detalle la variabilidad de la generación renovable, lo cual ocurre al usar días representativos en vez de, por ejemplo, semanas representativas, las cuales permitirían captar la variabilidad dentro de la semana. Esto, ya que, por ejemplo, podría estar ocurriendo en la realidad que hayan 2-3 días consecutivos con baja generación eólica, lo cual no se estaría considerando al aplicar la agregación temporal. A esto se le suma que, al momento de realizar *clustering k-medoids* y escoger los medoides dentro de cada clúster, por definición el día representativo nunca tomará el menor valor dentro de un clúster, al menos que éste sea el único elemento del clúster, no permitiendo la representación de los días “malos” en cuanto a generación renovable.

5.4.5. Tiempo de simulación

Por último, un punto importante a tener en consideración al momento de comparar las metodologías con la simulación horaria, es el tiempo de simulación. En la Figura 5.3 se presenta este valor, además de la reducción del tiempo de simulación que se puede obtener al considerar una metodología de agregación temporal versus utilizar todas las horas del mes, para cada uno de los meses en cuestión. De aquí se desprende que con el método actual de la CNE se obtiene el menor tiempo de simulación. Sin embargo, para todas las metodologías la reducción es considerable, encontrándose por sobre un 90 % la reducción de horas requeridas para simular la operación del SEN en la plataforma web Ameba, por lo que se podría sacrificar carga computacional al considerar una nueva metodología de agregación temporal, con el fin de obtener mejores resultados que los obtenidos con la metodología actual de la CNE.

Tabla 5.3: Tiempo de simulación de la operación en la plataforma web Ameba para el caso horario real y las metodologías de agregación temporal, además de la reducción del tiempo de simulación para estas metodologías en comparación con el caso real.

	CNE	<i>Clustering k-medoids</i>			Caso horario
		1 clúster	2 clústeres	3 clústeres	
Tiempo de simulación (Ts) ^a	1 hora	2 horas	6 horas	9 horas	174 horas
Reducción del Ts	99,43 %	98,85 %	96,55 %	94,83 %	-

^a Tener en consideración que estos son los tiempos de simulación obtenidos en un computador, y que mediante más hardware éstos se podrían ver reducidos.

Capítulo 6

Conclusiones y trabajos futuros

6.1. Conclusiones generales

En el presente capítulo se presentan las conclusiones generales obtenidas durante la realización de este trabajo de título.

En primer lugar, respecto a la adaptación de los días representativos a los datos reales para ambas tecnologías, se tiene que, al aplicar la metodología de la CNE, no hay gran diferencia entre un día hábil y no hábil, por lo que no aporta mucha información considerar dos días representativos al mes para la generación renovable. Esto provoca la necesidad de encontrar alguna metodología que permita definir de otra forma los días representativos, de manera de tener un mejor ajuste y que permita capturar la variabilidad de las ERNC.

Luego, los métodos de *clustering* analizados permiten obtener una menor Raíz del Error Cuadrático Medio en comparación a la obtenida con la metodología actual de la Comisión Nacional de Energía. En otras palabras, con los métodos de *clustering* se obtiene una mejor adaptación de los días representativos mensuales a los datos reales de generación, tanto para la generación eólica como para la generación solar fotovoltaica. Lo anterior da indicios de que los métodos de *clustering* pueden ser una opción viable para encontrar una nueva metodología de agregación temporal para la generación renovable. Además, entre ellos no existe una gran diferencia en rendimiento, por lo que optar por el método *k-medoids* parece ser la mejor opción al tener la particularidad de seleccionar un día de los datos reales para evitar el “suavizamiento” de las curvas, y ser el algoritmo más robusto a *outliers*.

Otro punto es que, en particular para el método de *clustering k-medoids*, se obtienen mejores resultados al considerar la generación dividida geográficamente, en términos del indicador RMSE, ya que de no considerarla, la asignación de los clústeres, es decir, la asignación de los días representativos, queda definida por la generación total del sistema, la cual no representa en más de un 40% la generación de las subestaciones de manera individual, induciendo mayores errores (o peores ajustes de los días representativos a la generación real). Cabe mencionar que este porcentaje de representación es para la generación específica utilizada en este estudio, no obstante, la tendencia es similar en el Sistema Eléctrico Nacional bajo cualquier escenario.

En cuanto a la cantidad de días representativos a considerar por mes para la simulación de la operación, se tiene que, a partir de 2 días al mes, se obtiene una mejor correlación

entre zonas al utilizar el método de *clustering*, considerando que la metodología actual de la Comisión Nacional de Energía no mantiene la correlación entre zonas al considerar un día representativo hábil y no hábil. Es importante recordar que mantener la correlación es fundamental para no inducir congestiones en las líneas de transmisión donde no corresponde, lo cual tendrá gran impacto en el Plan de Expansión de las Líneas de Transmisión. Además, al utilizar el indicador RMSE para medir la adaptación de los días representativos de manera mensual, se concluye que se necesita una cantidad distinta de días representativos según las características del mes en cuestión, y que, en general, esta cantidad es mayor para la generación eólica que para la generación solar fotovoltaica. Sin embargo, por temas de simplicidad, es posible definir solo un número de días necesarios para todos los meses, el cual consiste en el que tenga mayor frecuencia, de tal manera que la frecuencia acumulada sea al menos del 50 %.

Al realizar la simulación de la operación en la plataforma web Ameba, se tiene como resultado que el método de *clustering k-medoids* es capaz de representar de mejor manera la generación eólica y solar fotovoltaica respecto a la representación actual de la CNE, obteniendo resultados muy similares a los obtenidos en la realidad al no agregar temporalmente los datos. Sin embargo, ninguno de los métodos de agregación temporal es capaz de representar la generación convencional tal como sucedería en la realidad, puesto que no son capaces de capturar los días extremos de generación renovable, que podrían estar ocasionando que sea necesario producir energía eléctrica utilizando las centrales generadoras más costosas, tal como es el diésel, generando por ende diferencias notorias en el costo de operación del sistema eléctrico. Esto podría traer como consecuencia que, al utilizar métodos de agregación temporal, las señales de planificación de las redes de transmisión a futuro no sean tan certeras, por lo que es importante encontrar una metodología que sí sea capaz de captar las horas donde sea necesario generar con tecnología más cara.

En resumen, la metodología actual de la Comisión Nacional de Energía tiene un gran potencial de mejora, donde, en particular para este estudio, se plantea la aplicación de un método de *clustering k-medoids* utilizando como mínimo 2 días representativos por mes, tanto para la generación eólica como la solar fotovoltaica, al separarla por subestación asociada a cada central generadora, lo cual permite una mejor adaptación de los días representativos a los datos horarios de generación, además de mantener la correlación entre zonas, lo cual no sucede con la metodología actual. Sin embargo, este planteamiento basado en series de tiempo no es suficiente, ya que no se logra captar el comportamiento real de los costos operacionales del sistema eléctrico, por lo que se plantea analizar otros indicadores al momento de realizar el *clustering* y usar criterios adicionales tal como los vistos en la sección 5.4.4, donde se obtuvieron costos de operación más similares a los reales.

Por último, y en relación a los tiempos de simulación, se destaca que con los métodos de agregación temporal se logra reducir en al menos un 90 % la carga computacional. Si bien el método actual de la CNE reduce en un 99 % el tiempo de simulación, el método de *clustering k-medoids* considerando 3 días representativos al mes logra reducir en un 95 % este tiempo, pero obteniendo resultados más similares a los reales, siendo también un tiempo compatible con el proceso.

6.2. Trabajos futuros

Partiendo por el análisis de otros métodos de *clustering*, en el desarrollo de este trabajo se aplicaron tres métodos, y se determinó que el método *k-medoids* es el más efectivo para la representación temporal. Sin embargo, como trabajo futuro se plantea analizar otros métodos de *clustering* y comparar su efectividad con el método *k-medoids*.

Por otro lado, se plantea considerar otros factores en la simulación. Además de la generación renovable y convencional, se podría considerar otros factores tales como la demanda eléctrica, el almacenamiento de energía, la capacidad de la red de transmisión, entre otros, para obtener una representación más completa del sistema eléctrico y mejorar la precisión de los resultados de la simulación.

Un punto importante como trabajo futuro es la evaluación de incertidumbres. Se podría realizar un análisis de incertidumbres en la metodología propuesta para evaluar su robustez ante diferentes escenarios y condiciones climáticas, y cuantificar la variabilidad de los resultados.

En cuanto a la generación convencional, un trabajo futuro es analizar y mejorar la representación temporal de la generación convencional para lograr una mejor simulación de la operación, para así obtener un costo de operación del sistema más similar al real.

Otro trabajo futuro corresponde a validar la metodología en otros sistemas eléctricos para comprobar su efectividad y generalidad, ya que la metodología propuesta se probó únicamente utilizando datos del Sistema Eléctrico Nacional de Chile.

Finalmente, se plantea a futuro analizar los impactos económicos y ambientales de la incorporación de la nueva metodología propuesta en la planificación de la transmisión, además de evaluar cómo la inclusión de la nueva metodología propuesta en la planificación de la transmisión afecta la confiabilidad y seguridad del sistema eléctrico.

Bibliografía

- [1] Park, S., Xu, Q., y Hobbs, B., “Comparing scenario reduction methods for stochastic transmission planning,” *IET Generation, Transmission & Distribution*, vol. 13, pp. 1–2, 2019, [doi:10.1049/iet-gtd.2018.6362](https://doi.org/10.1049/iet-gtd.2018.6362).
- [2] International Energy Agency, “Renewables Data Explorer,” 2022. www.iea.org/renewables-data-explorer.
- [3] Ministerio de Energía de Chile, “Plan de Retiro y/o Reconversión de Unidades a Carbón,” 2020. www.energia.gob.cl.
- [4] Ministerio de Energía de Chile, “Política Energética Nacional, Transición Energética de Chile,” 2022. www.energia.gob.cl/energia2050.
- [5] Ministerio de Economía, Fomento y Reconstrucción, “DFL 4/20.018 Ley General de Servicios Eléctricos,” 2022, <http://bcn.cl/32jr0>. Art. 87°.
- [6] Chowdhury, B. H. y Rahman, S., “A review of recent advances in economic dispatch,” *IEEE Transactions on Power Systems*, vol. 5, pp. 1248–1259, 1990, [doi:10.1109/59.99376](https://doi.org/10.1109/59.99376).
- [7] Xia, X. y Elaiw, A. M., “Optimal dynamic economic dispatch of generation: a review,” *Electric Power Systems Research*, vol. 80, pp. 975–986, 2010, [doi:10.1016/j.epsr.2009.12.012](https://doi.org/10.1016/j.epsr.2009.12.012).
- [8] Ciornei, I. y Kyriakides, E., “Recent methodologies and approaches for the economic dispatch of generation in power systems,” *International Transactions on Electrical Power*, vol. 23, 2013, [doi:10.1002/etep.1635](https://doi.org/10.1002/etep.1635).
- [9] González, J., “Impacto económico y ambiental de la capacidad de las líneas de transmisión en los sistemas eléctricos de potencia,” tesis de magíster, Pontificia Universidad Católica de Chile, Facultad de Ingeniería, Departamento de Ingeniería Eléctrica, Santiago de Chile, 2013.
- [10] Palacios, E., “Modelo para manejo de la incertidumbre hidrológica en la planificación de la operación del sic,” tesis de pregrado, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, Departamento de Ingeniería Eléctrica, Santiago de Chile, 2004.
- [11] Kas Ingeniería, “Ose2000.” www.ose2000.kasing.cl.
- [12] Avilés, N., “Estudio y mejoramiento de un modelo de predespacho aplicado a la operación de embalses hidroeléctricos del sistema interconectado central,” tesis de pregrado, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, Departamento de Ingeniería Eléctrica, Santiago de Chile, 2011.
- [13] Comisión Nacional de Energía, “Informe Técnico Preliminar Plan de Expansión Anual de Transmisión Año 2022,” 2023. www.cne.cl.

- [14] Ministerio de Energía de Chile, “Decreto 37 Aprueba Reglamento de los Sistemas de Transmisión y de Planificación de la Transmisión,” 2021.
- [15] Universitat de Valencia, “Análisis Cluster,” 2018. www.uv.es/cluster.
- [16] Universidad de Granada, “Introducción al Análisis Cluster. Consideraciones generales,” 2012. www.ugr.es/cluster-g.
- [17] Universidad de Granada, “Métodos de Análisis Multivariante: Análisis Clúster,” 2015. www.wpd.ugr.es/practica8.
- [18] Universidad Carlos III, “Análisis de Cluster y Árboles de Clasificación,” 2015. www.halweb.uc3m.es/tema6.
- [19] Universidad de Granada, “Métodos Jerárquicos de Análisis Cluster,” 2012. www.ugr.es/cluster-3.
- [20] Towards Data Science, “Forecast KPIs: RMSE, MAE, MAPE & Bias,” 2019. www.towardsdatascience.com/forecast.
- [21] SPEC, “Ameba.” www.spec.cl/AmebaCloud.
- [22] Kittel, M., Hobbie, H., y Dierstein, C., “Temporal aggregation of time series to identify typical hourly electricity system states: A systematic assessment of relevant cluster algorithms,” *Energy*, vol. 247, 2022, doi:10.1016/j.energy.2022.123458.
- [23] Poncelet, K., Höschle, H., Delarue, E., Virag, A., y D’haeseleer, W., “Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems,” *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 1936–1948, 2017, doi:10.1109/TPWRS.2016.2596803.
- [24] Munoz, F. y Mills, A., “Endogenous assessment of the capacity value of solar pv in generation investment planning studies,” *IEEE Transactions on Sustainable Energy*, vol. 6, pp. 1–12, 2015, doi:10.1109/TSTE.2015.2456019.
- [25] Merrick, J., “On representation of temporal variability in electricity capacity planning models,” *Energy Economics*, vol. 59, pp. 261–274, 2016, doi:10.1016/j.eneco.2016.08.001.
- [26] De Sisternes, F. J., Webster, M. D., y Pérez-Arriaga, I. J., “The impact of bidding rules on electricity markets with intermittent renewables,” *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1603–1613, 2015, doi:10.1109/TPWRS.2014.2355775.
- [27] Green, R., Staffell, I., y Vasilakos, N., “Divide and conquer? k -means clustering of demand data allows rapid and accurate simulations of the british electricity system,” *IEEE Transactions on Engineering Management*, vol. 61, no. 2, pp. 251–260, 2014, doi:10.1109/TEM.2013.2284386.
- [28] Munoz, F. D., Hobbs, B. F., y Watson, J. P., “New bounding and decomposition approaches for milp investment problems: Multi-area transmission and generation planning under policy constraints,” *European Journal of Operational Research*, vol. 248, no. 3, pp. 888–898, 2016, doi:10.1016/j.ejor.2015.07.057.
- [29] Nahmmacher, P., Schmid, E., y Knopf, B., “Documentation of limes-eu -a long-term electricity system model for europe,” rep. tec., Potsdam Institute for Climate Impact Research, Potsdam, 2014. www.pik-potsdam.de.
- [30] Pineda, S. y Morales, J. M., “Chronological time-period clustering for optimal capacity expansion planning with storage,” *IEEE Transactions on Power Systems*, vol. 33, no. 6,

pp. 7162–7170, 2018, [doi:10.1109/TPWRS.2018.2842093](https://doi.org/10.1109/TPWRS.2018.2842093).

- [31] Ministerio de Energía de Chile, “Informe de Actualización de Antecedentes 2022, Planificación Energética de Largo Plazo,” 2022. www.energia.gob.cl.

Anexo

Tabla A.1: Potencia instalada por año de las centrales eólicas, junto a la subestación asociada.

Central eólica	Subestación	Potencia instalada [MW]			
		2025	2030	2035	2040
Antofagasta 3	Kimal	-	-	-	533
Antofagasta 4	Kimal	-	-	1.470	1.470
Antofagasta 5	Kimal	-	-	-	1.030
Ckani	Kimal	108	108	108	108
Marmoleras	Kimal	-	150	150	150
Sierra Gorda Este	Kimal	112	112	112	112
Tchamma	Kimal	155	155	155	155
Llanos del Viento	Changos	156	156	156	156
Antofagasta 18	Nueva Zaldívar	-	-	-	227
Antofagasta 2	Nueva Zaldívar	-	-	6	6
Calama	Nueva Zaldívar	150	150	150	150
Cerro Tigre	Nueva Zaldívar	185	185	185	185
Valle de Los Vientos	Nueva Zaldívar	89	89	89	89
Antofagasta 12	Parinas	-	-	-	364
Antofagasta 15	Parinas	-	-	1.632	2.000
Horizonte	Parinas	980	980	980	980
Lomas de Taltal	Parinas	-	353	353	353
Nolana	Parinas	-	280	280	280
Pampa Fidelia	Parinas	-	920	920	920
Pampa Yolanda	Parinas	-	532	532	532
Tal Tal rbc	Parinas	98	98	98	98
Atacama	Nueva Maitencillo	165	165	165	165
Atacama 21	Nueva Maitencillo	-	160	160	160
Cabo Leones I	Nueva Maitencillo	116	116	116	116
Cabo Leones I (Extensión)	Nueva Maitencillo	60	60	60	60
Cabo Leones II	Nueva Maitencillo	204	204	204	204
Cabo Leones III - Fase I	Nueva Maitencillo	78	78	78	78
Cabo Leones III Fase 2	Nueva Maitencillo	110	110	110	110
Coquimbo 25	Nueva Maitencillo	-	-	110	110
Coquimbo 30	Nueva Maitencillo	-	-	24	92

Sarco	Nueva Maitencillo	169	169	169	169
Canela 01	Nueva Pan de Azúcar	18	18	18	18
Canela 02	Nueva Pan de Azúcar	59	59	59	59
Coquimbo 24	Nueva Pan de Azúcar	-	-	21	21
Coquimbo 26	Nueva Pan de Azúcar	-	329	329	329
Coquimbo 27	Nueva Pan de Azúcar	-	-	11	11
Coquimbo 28	Nueva Pan de Azúcar	-	317	317	317
El Arrayán	Nueva Pan de Azúcar	115	115	115	115
Los Cururos	Nueva Pan de Azúcar	110	110	110	110
Monte Redondo	Nueva Pan de Azúcar	48	48	48	48
Punta Colorada	Nueva Pan de Azúcar	20	20	20	20
Punta de Talca	Nueva Pan de Azúcar	86	86	86	86
Punta Palmeras	Nueva Pan de Azúcar	45	45	45	45
Punta Sierra	Nueva Pan de Azúcar	80	80	80	80
San Juan	Nueva Pan de Azúcar	185	185	185	185
Talinay Oriente	Nueva Pan de Azúcar	90	90	90	90
Talinay Poniente	Nueva Pan de Azúcar	60	60	60	60
Total	Nueva Pan de Azúcar	46	46	46	46
Coquimbo 29	Quillota	-	182	182	182
Valparaiso 32	Quillota	-	14	14	14
Cardonal	Polpaico	33	33	33	33
La Estrella	Polpaico	50	50	50	50
Manantiales	Polpaico	27	27	27	27
Metropolitana 33	Polpaico	-	19	19	19
Metropolitana 35	Polpaico	-	6	6	6
Ucuquer I	Polpaico	7	7	7	7
Ucuquer II	Polpaico	11	11	11	11
Ohiggins 36	Rapel	-	45	45	45
Maule 37	Ancoa	-	188	188	188
Maule 41	Ancoa	-	160	160	160
Maule 42	Ancoa	-	75	75	75
Maule 43	Ancoa	-	-	123	187
Maule 45	Ancoa	-	119	119	119
Maule 46	Ancoa	-	171	171	171
Alena	Concepción	84	84	84	84
Biobío 63	Concepción	-	259	259	259
Biobío 69	Concepción	-	-	427	1.539
Biobío 71	Concepción	-	360	360	360
Biobío 72	Concepción	-	330	330	330
El Arrebol	Concepción	9	9	9	9
El Maitén	Concepción	9	9	9	9
El Nogal	Concepción	9	9	9	9
Huajache	Concepción	6	6	6	6

La Esperanza	Concepción	11	11	11	11
La Flor	Concepción	32	32	32	32
Las Peñas	Concepción	8	8	8	8
Lebu	Concepción	10	10	10	10
Lebu III	Concepción	5	5	5	5
Lomas de Duqueco	Concepción	57	57	57	57
Los Buenos Aires	Concepción	24	24	24	24
Mesamávida	Concepción	60	60	60	60
Negrete	Concepción	36	36	36	36
Negrete Cuel	Concepción	32	32	32	32
Nuble 48	Concepción	-	-	171	171
Raki	Concepción	9	9	9	9
Renaico	Concepción	88	88	88	88
Renaico II	Concepción	144	144	144	144
San Rarínco	Concepción	99	99	99	99
Biobío 60	Nueva Charrúa	-	106	106	106
Campo Lindo	Nueva Charrúa	72	72	72	72
Colinas	Nueva Charrúa	-	188	188	188
Dañicalqui	Nueva Charrúa	-	68	68	68
Nuble 50	Nueva Charrúa	-	-	-	36
Nuble 54	Nueva Charrúa	-	491	491	491
Biobío 65	Mulchén	-	-	31	150
Biobío 68	Mulchén	-	154	430	430
Los Olmos	Mulchén	100	100	100	100
Los Ríos 98	Mulchén	-	122	467	467
Malleco	Mulchén	135	135	135	135
Malleco II	Mulchén	138	138	138	138
San Andrés	Mulchén	-	120	120	120
San Gabriel	Mulchén	183	183	183	183
Tolpán Sur	Mulchén	84	84	84	84
Araucanía 75	Río Malleco	-	-	-	33
Araucanía 76	Río Malleco	-	-	430	430
Araucanía 86	Río Malleco	-	-	-	399
Biobío 58	Río Malleco	-	-	-	25
Los Ríos 88	Río Malleco	-	43	43	43
Los Ríos 89	Río Malleco	-	-	709	709
Los Ríos 90	Río Malleco	-	-	174	367
Los Ríos 92	Río Malleco	-	-	793	793
Los Ríos 99	Río Malleco	-	147	420	420
Caman	Pichirropulli	150	150	150	150
El Cruce	Pichirropulli	3	3	3	3
Los Lagos 106	Pichirropulli	-	49	181	181
Los Lagos 114	Pichirropulli	-	-	176	231

Los Ríos 100	Pichirropulli	-	-	418	418
Los Ríos 87	Pichirropulli	-	-	-	6
Los Ríos 97	Pichirropulli	-	-	73	151
OCHS	Pichirropulli	3	3	3	3
Aurora rbc	Nueva Puerto Montt	126	126	126	126
Los Lagos 111	Nueva Puerto Montt	-	-	18	18
Los Lagos 127	Nueva Puerto Montt	-	-	-	31
Los Lagos 130	Nueva Puerto Montt	-	221	221	221
Puelche Sur	Nueva Puerto Montt	152	152	152	152
San Pedro	Nueva Puerto Montt	36	36	36	36
San Pedro II	Nueva Puerto Montt	65	65	65	65
Vientos del Lago	Nueva Puerto Montt	-	125	125	125
Los Lagos 117	Nueva Ancud	-	868	1.701	1.930

Tabla A.2: Potencia instalada por año de las centrales solares fotovoltaicas, junto a la subestación asociada.

Central solar fotovoltaica	Subestación	Potencia instalada [MW]			
		2025	2030	2035	2040
Arica Solar	Parinacota	26	26	26	26
Willka	Parinacota	98	98	98	98
Almonte	Nueva Pozo Almonte	75	75	75	75
La Huayca 2	Nueva Pozo Almonte	25	25	25	25
Los Puquios	Nueva Pozo Almonte	2	2	2	2
Pampa Camarones	Nueva Pozo Almonte	6	6	6	6
PAS1	Nueva Pozo Almonte	17	17	17	17
PAS3	Nueva Pozo Almonte	16	16	16	16
Pica	Nueva Pozo Almonte	90	90	90	90
Pintados	Nueva Pozo Almonte	77	77	77	77
Salar de Huasco	Nueva Pozo Almonte	30	30	30	30
Tamarugo	Nueva Pozo Almonte	3	3	3	3
Wara III	Nueva Pozo Almonte	45	45	45	45
Atacama Solar	Lagunas	144	144	144	144
Atacama Solar II	Lagunas	150	150	150	150
El Águila	Lagunas	2	2	2	2
Granja Solar	Lagunas	105	105	105	105
Huatacondo	Lagunas	98	98	98	98
Nuevo Quillagua	Lagunas	100	100	100	100
Sol del Loa Fase I	Lagunas	110	110	110	110
Sol del Loa Fase II	Lagunas	190	190	190	190
Sur Viejo	Lagunas	93	93	93	93
Tarapacá 12	Lagunas	-	-	-	74
Tarapacá 15	Lagunas	-	-	-	1,613

Alfa Solar	Kimal	854	854	854	854
Bolero	Kimal	146	146	146	146
Capricornio	Kimal	88	88	88	88
CEME1	Kimal	480	480	480	480
Cerro Dominador	Kimal	100	100	100	100
Domeyko	Kimal	186	186	186	186
Elena - Primera Etapa	Kimal	270	270	270	270
Finis Terrae	Kimal	137	137	137	137
Finis Terrae Etapa 2 (Extensión)	Kimal	18	18	18	18
Finis Terrae Etapa I (Ampliación)	Kimal	126	126	126	126
La Cruz Solar	Kimal	50	50	50	50
La Sierra (Ex PMG El Tesoro)	Kimal	9	9	9	9
Las Salinas	Kimal	364	364	364	364
Lascar	Kimal	65	65	65	65
María Elena	Kimal	68	68	68	68
PMG Caliche	Kimal	9	9	9	9
Proyecto Coya	Kimal	180	180	180	180
Quillagua	Kimal	90	90	90	90
San Pedro	Kimal	106	106	106	106
Santa Isabel Etapa I - Fase I	Kimal	159	159	159	159
Sol del Desierto Fase I	Kimal	175	175	175	175
Sol del Desierto Fase II	Kimal	55	55	55	55
Tamaya Solar	Kimal	114	114	114	114
Tocopilla	Kimal	428	428	428	428
Uribe Solar	Kimal	53	53	53	53
Valle del Sol	Kimal	150	150	150	150
Aguas Blancas 2	Changos	72	72	72	72
Andes 220 III	Nueva Zaldívar	36	36	36	36
Andes 220 IV	Nueva Zaldívar	45	45	45	45
Andes IIB	Nueva Zaldívar	113	113	113	113
Andes Solar	Nueva Zaldívar	21	21	21	21
Andes Solar IIA	Nueva Zaldívar	80	80	80	80
Azabache	Nueva Zaldívar	60	60	60	60
JAMA	Nueva Zaldívar	51	51	51	51
Pampa Tigre	Nueva Zaldívar	100	100	100	100
Sol de Lila	Nueva Zaldívar	152	152	152	152
Usya	Nueva Zaldívar	52	52	52	52
Antofagasta 29	Parinas	-	-	-	1,173
Lalackama 2 rbc	Parinas	17	17	17	17
Lalackama rbc	Parinas	55	55	55	55
Solar Toro	Parinas	56	56	56	56
Almeyda	Cumbres	52	52	52	52
Antofagasta 24	Cumbres	-	-	1,196	1,528

Campos del Sol	Cumbres	381	381	381	381
Campos del Sol II	Cumbres	370	370	370	370
Carrera Pinto	Cumbres	93	93	93	93
Carrera Pinto II	Cumbres	50	50	50	50
Chañares	Cumbres	35	35	35	35
Conejo I	Cumbres	104	104	104	104
Covadonga	Cumbres	9	9	9	9
DAS	Cumbres	8	8	8	8
Diego de Almagro	Cumbres	32	32	32	32
Diego de Almagro Sur	Cumbres	208	208	208	208
El Salvador 110 I	Cumbres	6	6	6	6
Guanaco Solar	Cumbres	77	77	77	77
Inca de Varas I	Cumbres	50	50	50	50
Inca de Varas II	Cumbres	50	50	50	50
Javiera	Cumbres	65	65	65	65
Luz del Norte	Cumbres	141	141	141	141
Malgarida I	Cumbres	28	28	28	28
Malgarida II	Cumbres	163	163	163	163
Salvador	Cumbres	68	68	68	68
Sol de Atacama	Cumbres	81	81	81	81
Sol de Los Andes	Cumbres	89	89	89	89
Sol de Varas	Cumbres	101	101	101	101
Calderaza	Nueva Cardones	9	9	9	9
Cardones	Nueva Cardones	35	35	35	35
Cardones 220 I	Nueva Cardones	9	9	9	9
El Ñandú	Nueva Cardones	3	3	3	3
H. Fuentes 023 I	Nueva Cardones	18	18	18	18
Llano de Llampos	Nueva Cardones	101	101	101	101
Los Loros	Nueva Cardones	46	46	46	46
Los Loros 023 I	Nueva Cardones	18	18	18	18
Pampa Solar Norte	Nueva Cardones	69	69	69	69
PMG Castilla	Nueva Cardones	3	3	3	3
PMGD Bramada	Nueva Cardones	9	9	9	9
Río Escondido	Nueva Cardones	145	145	145	145
San Andrés	Nueva Cardones	51	51	51	51
Solar Las Terrazas	Nueva Cardones	3	3	3	3
Valle Escondido	Nueva Cardones	105	105	105	105
Don Carlos	Nueva Maitencillo	-	196	196	196
El Romero	Nueva Maitencillo	196	196	196	196
La Huella	Nueva Maitencillo	84	84	84	84
La Silla	Nueva Maitencillo	2	2	2	2
Pelícano	Nueva Maitencillo	100	100	100	100
PMGD Astillas	Nueva Maitencillo	9	9	9	9

Sol de Vallenar - Fase II	Nueva Maitencillo	250	250	250	250
Tamarico	Nueva Maitencillo	145	145	145	145
Coquimbo 87	Nueva Pan de Azúcar	-	-	-	55
Incahuasi 023 I	Nueva Pan de Azúcar	5	5	5	5
Llanos del Potroso	Nueva Pan de Azúcar	9	9	9	9
PMGD Gaviotín	Nueva Pan de Azúcar	9	9	9	9
Punta del Viento	Nueva Pan de Azúcar	165	165	165	165
SDGx01	Nueva Pan de Azúcar	1	1	1	1
Caracas II	Quillota	9	9	9	9
Casablanca 012 IV	Quillota	25	25	25	25
Combarbalá 13.2 III	Quillota	15	15	15	15
Coquimbo 79	Quillota	-	-	-	72
Coquimbo 83	Quillota	-	76	76	76
Doña Carmen PV	Quillota	40	40	40	40
El Boco	Quillota	3	3	3	3
El Divisadero PV	Quillota	3	3	3	3
Panquehue 012 II	Quillota	9	9	9	9
PMGD El Boco (Ampliación)	Quillota	5	5	5	5
PMGD Las Majadas	Quillota	9	9	9	9
PMGD Pitra	Quillota	3	3	3	3
PMGD Villa Alemana	Quillota	3	3	3	3
Quilpué 012 II	Quillota	3	3	3	3
San Rafael 023 II	Quillota	15	15	15	15
Santiago Solar	Quillota	98	98	98	98
TilTil Solar	Quillota	3	3	3	3
UTFSM Viña del Mar	Quillota	1	1	1	1
Valparaíso 110	Quillota	-	-	-	474
Valparaíso 115	Quillota	-	-	-	10
Alcones	Polpaico	50	50	50	50
Cintac	Polpaico	3	3	3	3
Curacaví 012 III	Polpaico	12	12	12	12
El Manzano	Polpaico	87	87	87	87
El Monte 13.2 II	Polpaico	6	6	6	6
El Paico 13.2 III	Polpaico	13	13	13	13
La Esperanza 13.2 III	Polpaico	15	15	15	15
La Manga 13.2 IV	Polpaico	20	20	20	20
Las Arañas 1 13.2 II	Polpaico	6	6	6	6
Mandinga 13.2 III	Polpaico	15	15	15	15
Marchigue 13.2 III	Polpaico	9	9	9	9
Marchigue II	Polpaico	9	9	9	9
Meseta de los Andes	Polpaico	138	138	138	138
Metropolitana 127	Polpaico	-	-	-	23
PMG Parque Doña Carmen RM	Polpaico	9	9	9	9

PMG Solar Palermo	Polpaico	9	9	9	9
PMGD Nahuén	Polpaico	9	9	9	9
Polpaico Chilectra 023 II	Polpaico	6	6	6	6
Portezuelo 023 III	Polpaico	15	15	15	15
Punta de Peuco 110 II	Polpaico	12	12	12	12
Quilapilún	Polpaico	103	103	103	103
Quilapilún (Ampliación)	Polpaico	7	7	7	7
RLA	Polpaico	3	3	3	3
Santa Marta 023 II	Polpaico	6	6	6	6
Solar Loma Los Colorados	Polpaico	1	1	1	1
Tagua Tagua	Polpaico	-	176	176	176
Valparaíso 107	Polpaico	-	-	-	24
BU GR Rovian	Alto Jahuel	7	7	7	7
Colbún 13.8 II	Alto Jahuel	18	18	18	18
Don Mariano	Alto Jahuel	3	3	3	3
Doñihue	Alto Jahuel	8	8	8	8
El Picurio	Alto Jahuel	3	3	3	3
Graneros 015 III	Alto Jahuel	26	26	26	26
Isla de Maipo 012 II	Alto Jahuel	18	18	18	18
Las Cabras 015 II	Alto Jahuel	13	13	13	13
Las Cachañas	Alto Jahuel	9	9	9	9
PMG Llançay	Alto Jahuel	9	9	9	9
PMGD Antonia Solar	Alto Jahuel	3	3	3	3
PMGD Curicura	Alto Jahuel	9	9	9	9
PMGD El Guanaco Solar I	Alto Jahuel	3	3	3	3
PMGD Quinta	Alto Jahuel	8	8	8	8
PMGD Techos Solares Watts	Alto Jahuel	1	1	1	1
PMGD Watt's Lonquén	Alto Jahuel	1	1	1	1
San Isidro	Alto Jahuel	3	3	3	3
Tricahue	Alto Jahuel	9	9	9	9
Andino Occidente	Rapel	-	147	147	147
Metropolitana 120	Candelaria	-	-	-	13
Cordillerilla	Ancoa	1	1	1	1
El Chincol	Ancoa	3	3	3	3
El Chincol rbc	Ancoa	3	3	3	3
El Cuervo	Ancoa	3	3	3	3
La Foresta	Ancoa	3	3	3	3
Los Patos	Ancoa	3	3	3	3
Maule 149	Ancoa	-	-	-	918
Molina	Ancoa	9	9	9	9
Panguilemo 015 II	Ancoa	11	11	11	11
Panguilemo 015 III	Ancoa	11	11	11	11
Panimávida	Ancoa	8	8	8	8

Peralillo	Ancoa	3	3	3	3
Piduco 015 II	Ancoa	6	6	6	6
PMG Teno Solar	Ancoa	7	7	7	7
PMGD Lemu	Ancoa	5	5	5	5
PMGD Los Tordos	Ancoa	5	5	5	5
PMGD Quinantu Solar	Ancoa	9	9	9	9
Villa Alegre	Ancoa	9	9	9	9
Villa Prat 13.2 II	Ancoa	6	6	6	6
Yerbas Buenas 13.2 II	Ancoa	18	18	18	18
Bicentenario	Concepción	3	3	3	3
Cabrero 023 IV	Concepción	30	30	30	30
Chillan 066 II	Concepción	18	18	18	18
Chillán Confluencia	Concepción	3	3	3	3
El Chucao	Concepción	3	3	3	3
El Pitio	Concepción	3	3	3	3
Hualte 13.2 I	Concepción	3	3	3	3
Las Palomas	Concepción	3	3	3	3
Los Gorriones	Concepción	3	3	3	3
PMGD Alcaidesa	Concepción	6	6	6	6
PMGD Bulnes Los Barones	Concepción	3	3	3	3
PMGD Coihueco San Carlos	Concepción	3	3	3	3
PMGD Coihueco Vicente Méndez	Concepción	3	3	3	3
PMGD Dadinco	Concepción	3	3	3	3
PMGD Playero Etapa II	Concepción	3	3	3	3
PMGD Puelche	Concepción	3	3	3	3
Tres Esquinas 066 II	Concepción	9	9	9	9
Araucanía 208	Mulchén	-	-	-	46
PMGD Cancura II Solar	Mulchén	3	3	3	3
PMGD Quillón I	Mulchén	3	3	3	3
Victoria 13.2 I	Mulchén	6	6	6	6