Facultad de Ingeniería

Localización, dimensionamiento y conexión de sistemas de energía: Un marco metodológico para un desarrollo sustentable

David Alfredo Olave Rojas

Grado: Doctor en Sistemas de Ingeniería
Director de Tesis: Dr. Eduardo Álvarez Miranda
Co-director de Tesis: Dr. Claudio Tenreiro Leiva

Abril 2019

CONSTANCIA

La Dirección del Sistema de Bibliotecas a través de su unidad de procesos técnicos certifica que el autor del siguiente trabajo de titulación ha firmado su autorización para la reproducción en forma total o parcial e ilimitada del mismo.

Talca, 2021

A la amada familia,
de la cual provengo

"Phantasie ist wichtiger als Wissen, denn Wissen ist begrenzt"
"La fantasía es más importante que la ciencia, pues la ciencia es limitada"
Albert Einstein

Título de la Tesis	Localización, dimensionamiento y conexión de sistemas de energía: Un marco metodológico para un desarrollo sustentable.
Thesis Title	Location, sizing and connexion of energy systems: A methodological framework for a sustainable development.
Nombre Estudiante	David Alfredo Olave Rojas
Fecha Inicio Tesis	Octubre 2016
Fecha Término Tesis	Abril 2019
Director de Tesis	Dr. Eduardo Álvarez Miranda. Universidad de Talca – Facultad de Ingeniería Departamento de Ingeniería Industrial ealvarez@utalca.cl
Co-director de Tesis	Dr. Claudio Tenreiro Leiva. Universidad de Talca – Facultad de Ingeniería Departamento de Tecnologías Industriales ctenreiro@utalca.cl

Dr. Javier Muñoz Vidal

Dr. Alfredo Iriarte García

Dr. Eduardo Álvarez Miranda

Agradecimientos

Al profesor Eduardo Álvarez Miranda por su invaluable apoyo y por convertirse en el Sr. Miyagi, según sus propias palabras. Su dedicación a la hora de formarme como científico y enseñarme los secretos del arte de la ciencia es algo que siempre valoraré y agradeceré, al igual que sus cursos introductorios a *la vera cucina italiana*.

Al profesor Claudio Tenreiro Leiva por su ayuda, paciencia y sus certeras palabras en distintos momentos de mi vida profesional, primero como estudiante de pregrado y luego en esta etapa de doctorado como co-director de tesis. Estaré eternamente agradecido por confiar en mi siempre y darme las oportunidades que he recibido de su parte.

A los profesores Stefan Nickel y Francisco Saldanha da Gama por sus consejos, por ponerme a prueba y alentarme a descubrir nuevas áreas del saber humano. Espero poder seguir el camino que me han mostrado. Gracias por todo.

A la Escuela de Graduados de la Universidad de Talca y su programa de becas para posgrado, al igual que al Consejo Nacional de Ciencia y Tecnología (CONICYT) mediante los proyectos FONDECYT N.1180670 y FONDECYT N.11140060, al igual que a través del Instituto de Sistemas Complejos de Ingeniería - ISCI (ICM-FIC:P-05-004-F, CONICYT:FB0816) por permitirme una dedicación completa a este trabajo doctoral.

A la familia de la cual provengo, por apoyarme incondicionalmente siempre a lo largo de mi vida.

Y a cada una de las personas en el mundo, grandes o pequeñas, de aquí o de allá, locas o cuerdas, que me han alentado, inspirado y apoyado a lo largo de estos años.

Índice general

Resumen

El principal desafío de la humanidad en los siguientes años es desarrollar un estilo de vida económica, social y ambientalmente sustentable. Uno de los principales caminos para lograr esto es mediante la generación y uso de la energía de forma sustentable, pues afecta a todas la áreas de la humanidad. Esto implica la necesidad de cambiar nuestra matriz actual basada en combustibles fósiles, especialmente si pensamos que las reservas actuales de combustible tienen un horizonte máximo ubicado entre el año 2050 y el 2100. En este contexto, las energías renovables no convencionales son una solución sumamente atractiva y competitiva, presentando desafíos técnicos, económicos y operativos. Parte de estos desafíos se pueden enfrentar, por ejemplo, mediante sistemas de almacenamiento de energía, pero su incorporación a los sistemas eléctricos conlleva afrontar cuestiones que requieren el desarrollo de herramientas de apoyo a la toma de decisiones en las etapas de diseño y dimensionamiento de los sistemas de potencia. Si bien existen trabajos anteriores que han dedicado esfuerzos desde la investigación de operaciones, con el fin de aportar mejoras que avancen hacia una solución, ninguno de ellos ha incluido todos los actores de forma integrada, tomando en cuenta la interacción entre ellos. Persiguiendo ese fin, en esta tesis doctoral se estudia y presenta un marco metodológico que permita mejorar la sustentabilidad de una región determinada desde el punto de vista de los sistemas de energía.

Para esto se han utilizado herramientas de análisis y pronóstico de datos, con el objetivo de confeccionar una base de datos curada sobre el sistema eléctrico chileno. Luego se ha confeccionado un modelo de programación matemática mixta con el fin de apoyar la toma de decisiones de nivel estratégico, correspondiente a la inversión en proyectos de energía renovable. Las decisiones incluidas son las relativas a la localización, el dimensionamiento y la conexión de componentes, además de la generación, el almacenamiento y la transmisión de energía. Finalmente se ha desarrollado una formulación heurística que no solo incluya las complejidades estratégicas estudiadas en el trabajo anterior, sino que también, las complejidades existentes a nivel operativo y que afectan una evaluación de proyecto.

Los resultados muestran que el marco metodológico presentado no solo contribuye para apoyar la toma de decisiones desde el punto de vista del inversionista, sino que también permite evaluar aspectos territoriales con el objetivo de determinar su potencial para proyectos de energía por parte de instituciones reguladoras u organismos gubernamentales, mejorando los niveles de sustentabilidad del área. Además permite explorar posibles tecnologías en fase de desarrollo como lo son algunas asociadas al almacenamiento de energía, determinando parámetros de configuración e incluso conocer direcciones de desarrollo. Por otro lado, el marco metodológico permite conocer el impacto de herramientas de incentivo, como lo son los subsidios a la inversión, permitiendo focalizar y establecer una dirección para el desarrollo de políticas asociadas a este tema.

Finalmente, si bien esta tesis permite apoyar procesos de toma de decisión para desarrollar una generación y uso de energía más sustentable, también es claro ver que es necesario profundizar en esta metodología, incluyendo mas agentes o nuevas aproximaciones al problema, como lo son una combinación de fuentes de energía o modelos de simulación-optimización.

Palabras clave Energía, Dimensionamiento, Localización, Sustentabilidad, Optimización

Abstract

The greatest challenge for the humankind today is the develop of a lifestyle sustainable, from a social, economical and environmental point of view. One of the main tasks to accomplish this goal is the sustainable generation and use of energy. From an energy point of view, this implicates a need for changing our fossil fuel-based energy systems, as it affects every field of society, especially if we notice today that current oil reserves have a maximum expected lasting horizon between years 2050 and 2100. In this context, non conventional renewable sources are a highly attractive and competitive solution, however their use results in technical, economical and managerial issues. Some of these challenges can be faced, for example, through energy storage systems, but their incorporation to electric systems does not only involve technological challenges but also calls for decision-making tools that ensure appropriate design and sizing of the corresponding power systems. Although there are previous works that have devoted efforts from operations research, in order to provide improvements towards a solution, none of them has included all the actors in an integrated manner, taking into account the interaction between them. Following this idea, in this thesis, a methodological framework is studied and presented, with the aim of enhancing the sustainability of a particular territory from an energy system point of view.

Taking this in mind, analytics and data forecasting tools are used for building a carefully-curated database about the Chilean power grid. Thereafter, a mixed integer programming model is presented for aiding the evaluation of the strategic design of complex renewable power systems. Included decisions are related to the location, sizing and connexion of components, as the generation, storage and transmission of energy as well. Finally, an heuristic formulation is developed with the aim of enhancing the solution applicability and economic evaluation accuracy. The approach includes not just the complexity from strategical decisions (addressed also in the previous work), but also from operational decisions as well.

From results it is possible to conclude that the presented methodological framework is not just useful for supporting decision processes from an investment point of view, but also is useful for policymakers with the aim of enhancing sustainability and the competitiveness of a certain region. Such an approach made possible by exploring new zones and determining the potential for energy projects. Furthermore, the presented framework supports testing novel technologies, such as those related to energy storage systems, in a real-world scenario. Additionally, the framework allows knowing and studying the impact of incentive tools, such as investment subsidies or investment bonuses, targeting and establishing a direction for new policies or tools related to this issue.

Finally, although this thesis is an effort in the right direction in order to discover new opportunities for developing a more sustainable energy use and generation process, it is also clear that it exists a need, an urgency, to go further with this doctoral thesis,

Keywords Energy, Sizing, Location, Sustainability, Optimization

Capítulo 1

Introducción

Los humanos enfrentamos actualmente uno de los desafíos mas grandes de nuestra historia: hacer el estilo de vida del hombre moderno sustentable desde una dimensión económica, social y ambiental, con el fin de mejorar las opciones de sobrevivencia futura como raza humana. Hasta ahora hemos aprovechado recursos energéticos provenientes de procesos de descomposición ejecutados por la tierra durante millones de años, obteniendo petróleo, gas y otros derivados. Esto nos ha permitido mejorar nuestras capacidades de producción, por ejemplo al usar tractores en vez de caballos o personas para arar la tierra, gracias a fuentes de energía baratas pero que contribuyen también negativamente, tanto a la contaminación como a la emisión de gases de efecto invernadero. Además, los pronósticos actuales indican reservas de combustibles fósiles hasta un horizonte ubicado entre el año 2050 a 2100 [1–3]. Lo anterior es crítico si se toma en cuenta que un 45.7 % de la electricidad generada a nivel mundial durante el año 2013 se obtuvo de carbón y petroleo, un 21.7 % se obtuvo de gas natural. De hecho, 81.4 % de la energía primaria del mundo proviene de combustibles fósiles [4].

Además, en el año 2015, las naciones unidas aprobó la Agenda 2030 sobre el Desarrollo Sostenible, la cual cuenta con 17 objetivos [5] correspondientes a distintas aristas de la vida humana, entre los cuales el número 7 se refiere a la energía y establece la necesidad de "energía asequible y no contaminante" [6], de donde se desprende que la energía debe ser de acceso universal, fiable, local y no contaminante.

En este sentido, se han investigado, desarrollado e implementado distintas opciones de generación de energía, buscando sustentabilidad ambiental, tecnológica y de disponibilidad en el tiempo. En particular, las energías renovables (ER) se presentan como una, o tal vez la única, posible respuesta al problema [7, 8], aunque algunos expertos sostienen que existen escenarios posibles en el futuro donde deberemos usar otras fuentes de energía, como la nuclear, para generar una transición exitosa [8], con el fin de abastecer nuestra creciente demanda. Debido a esto, cualquier esfuerzo es bienvenido, no solo en cuestiones de generación, sino también en las relativas al uso eficiente de la energía. En este sentido, transmitir energía desde lugares lejanos siempre supondrá pérdidas del recurso, en vez de desarrollar soluciones locales, acorde con los recursos y posibilidades existentes en un territorio determinado. Esto implica evaluar las fuentes existentes localmente, siempre pensando en lograr un desarrollo sustentable [9].

Sin embargo, el uso de energías renovables no convencionales genera problemas o desafíos de orden económico, tecnológico y administrativo, los cuales aparecen al momento de diseñar, planificar y operar sistemas de energía con incorporación de fuentes renovables. Evidentemente, a una mayor participación de estas fuentes en un sistema dado, mayor también es el impacto de estos desafíos. Esto puede abordarse a través de los sistemas de almacenamiento de energía (*Energy Storage System*, ESS), cuyas tecnologías se han desarrollado intensamente durante las últimas décadas, facilitando el balance de energía como también el desacoplamiento entre tiempo de generación y consumo.

No obstante, los ESS siguen siendo soluciones costosas y con patrones de operación restringidos (en algunos casos deben incluso estar asociadas a un terreno con características particulares), por tanto una incorporación adecuada a los sistemas de energía. conlleva no solo desafíos tecnológicos, sino que también la necesidad de herramientas de decisión que aseguren el diseño y dimensionamiento apropiado de estos sistemas. En otras palabras, la implementación de este tipo de soluciones debe evaluarse tomando en cuenta sus características particulares, y cómo se acoplan a los sistemas existentes. Es decir, los proyectos de generación requieren ser diseñados, desde el punto de vista territorial, estratégico y táctico, considerando la disponibilidad del recurso, la demanda de energía, las características topológicas y funcionales de la red de generación existente, y el marco regulatorio. Dicho diseño requiere el uso de herramientas de análisis de datos y pronósticos, evaluación económica, modelamiento matemático, resolución algorítmica, y análisis de resultados. Estas herramientas deben permitir a los gestores proyectar políticas de generación que no sólo resuelvan una coyuntura de oferta-demanda, si no que mejoren a nivel sistémico el mercado de energía de un país, desde un punto de vista sustentable.

1.1. Estudios previos

La aplicación de herramientas de evaluación provenientes de la investigación de operaciones, en el diseño de sistemas de generación que incorporan ESS no es nuevo. No obstante, los trabajos previos se han enfocado en contestar, por separado, las problemáticas asociadas al dimensionamiento de generadores de energía renovable [10, 11], de sistemas de almacenamiento de energía [12–15], localización de generadores de energía renovable [16–18], y de sistemas de almacenamiento de energía [15, 19, 20], además de la operación y planificación de estos sistemas en la red a la que pertenecen [21–23], junto con el diseño del layout en situaciones de expansión o desde cero [15, 17, 23]. Complementariamente, se han desarrollado modelos para caracterizar la operación de redes de generación y transmisión que posean este tipo de sistemas, además del impacto de estos sistemas en la estabilidad de las redes donde son integrados [11, 18].

Los modelos desarrollados consideran diferentes objetivos, tecnologías y escalas. Sin embargo, no existe algún marco metodológico general que permita responder, en forma simultánea, cuestiones como dónde localizar unidades de generación, qué capacidad se debe instalar, dónde es posible localizar unidades de ESS, qué dimensión deben tener cómo y dónde integrar un proyecto de este tipo, cómo invertir y operar con tal de asegurar que dicho proyecto respete las condiciones de operación al mismo tiempo que, por ejemplo, maximiza el beneficio económico para un inversionista o cómo generar políticas públicas que incentiven la inversión en proyectos sustentables cuando el escenario de evaluación es complejo. Esto plantea las siguientes preguntas de investigación (RQ):

RQ 1. ¿Son las soluciones tecnológicas basadas en energías renovables competitivas en un contexto de libre competencia, presentando una solución realmente sustentable?.

RQ 2. ¿Qué nivel de calidad de datos es necesaria para representar situaciones de evaluación de proyectos en un contexto real?.

RQ 3. ¿Qué nivel de impacto tienen las tecnologías de ESS dentro de un proyecto de energías renovables?.

RQ 4. ¿Qué características matemáticas debe tener un modelo que combine los requerimientos de los entes reguladores y de los inversionistas en proyectos de energía?

RQ 5. ¿Cómo debería ser un sistema de ESS en términos de su capacidad para considerarse ideal?

RQ 6. ¿Cómo debería incentivar el organismo regulador o generador de políticas las soluciones sustentables para su implementación?

RQ 7. ¿Dónde debe realizarse la instalación de proyectos renovables?

RQ 8. ¿Cómo debe ser el layout o distribución de componentes de un sistema integrado?

Estas preguntas se pueden agrupar entre preguntas de diseño (RQ 1, 2, 7 y 8), desde el punto de vista del inversionista y preguntas de políticas (RQ 3, 4, 5 y 6) desde el punto de vista del organismo regulador. Además son de carácter estratégico y representan 2 importantes visiones que interactúan en los sistemas de energía. Finalmente, las preguntas dirigen el esfuerzo de investigación, cuyos resultados se presentan en esta tesis doctoral. Además estas preguntas inspiran los objetivos de este trabajo.

1.2. Objetivos

1.2.1. Objetivo general

Desarrollar un marco metodológico que permita diseñar políticas de desarrollo y estrategias de evaluación (económica, técnica, de sustentabilidad) de proyectos de generación que utilicen ER e incorporen tecnologías de almacenamiento. Dicha herramienta tendrá como componentes centrales: (i) métodos sofisticados de análisis de datos y pronóstico; (ii) técnicas de modelización de programación matemática (lineal, entera, no lineal, mixta, etc.); (iii) estrategias algorítmicas exactas y heurísticas; (iv) métodos de análisis relativos a la caracterización de nuevas tecnologías de almacenamiento.

1.2.2. Objetivos específicos

- Entender el comportamiento histórico de una región desde el punto de vista del mercado eléctrico.

- Caracterizar los elementos que toman parte de los sistemas de generación y almacenamiento de energía, mediante la representación de las características funcionales y topológicas, el marco regulatorio correspondiente,la disponibilidad de los recursos y la demanda.

- Formular modelos de programación matemática que permitan optimizar el nexo (económico, operativo, ambiental, etc.) entre proyectos de energía con almacenamiento y la infraestructura existente en un territorio determinado.

- Determinar la aplicabilidad del marco metodológico en la generación de políticas públicas para el desarrollo sustentable de un territorio.

- Diseñar estrategias algorítmicas de resolución, que permitan obtener soluciones en un tiempo razonable según la naturaleza de la problemática a desarrollar.

- Diseñar un esquema de análisis, cuantitativo y cualitativo, de las soluciones que permita la validación de los modelos y los algoritmos propuestos.

Estos objetivos constituyen el núcleo principal de la metodología desarrollada, la cual fue aplicada para resolver las preguntas de investigación presentadas previamente.

Capítulo 2

Marco Teórico

En esta tesis se utiliza teoría que proviene de distintas áreas de las ciencias, como son evaluación de proyectos, sistemas de información, estadística aplicada, econometría, programación, entre otras. Sin embargo, este capítulo se enfoca en la teoría que representa el corazón del marco metodológico presentado en esta tesis doctoral, es decir, la teoría relativa a la optimización matemática. Debido a que existen distintos tipos de optimización matemática, este capítulo se concentra en aquellos tipos que son utilizados en algunos de los trabajos incluidos en esta tesis doctoral. Se incluyen también otras áreas de suma relevancia para esta tesis como lo son el desarrollo sustentable, además del análisis y pronóstico de datos. Esta última se utiliza especialmente en la primera publicación asociada a esta tesis.

2.1. Desarrollo sustentable

El desarrollo sustentable implica un crecimiento o una mejora que puede sostenerse en el tiempo, es decir, mejorar las condiciones de vida de las personas mediante recursos, herramientas y procesos sostenibles o, dicho de otra forma, conectados con los sistemas circundantes. El desarrollo sustentable tiene claramente al menos una dimensión social, una económica y una ambiental. Estas dimensiones se pueden conectar al referirse al desarrollo sustentable como la búsqueda de altas tasas de crecimiento económico, considerando el bienestar de los menos privilegiados y sin dañar el medio ambiente.

La primera aparición de este concepto fue en un reporte para las naciones unidas sobre el futuro realizado por la Comisión Brundtland [24]. Esta comisión define el desarrollo sustentable como "el desarrollo que satisface las necesidades del presente sin comprometer la habilidad de generaciones futuras de abastecer sus propias necesidades". Uno de los principales desafíos asociado a este concepto es la correcta medición del impacto de los procesos humanos en las dimensiones antes nombradas. En ese sentido se han creado distintas herramientas y protocolos como lo son los estudios de ciclo de vida y el concepto de exergía. Esta última se define de una manera sencilla como la porción de energía que puede ser transformada a trabajo mecánico. Dos libros al respecto son el de Dincer y Rosen [25] y el de Wall [26]. Por otro lado, los estudios de ciclo de vida son estudios que buscan determinar el impacto y el uso de recursos de un producto a lo largo de su vida mediante distintos indicadores. Generalmente estos indicadores abarcan áreas como energía, emisiones, uso de suelo, uso de agua, entre otros. Dos de las metodologías mas conocidas son la *Eco-indicator 99* [27] y la *IMPACT 2002+* [28].

2.2. Optimización matemática

Una primera idea sobre la optimización es dada por la definición existente en el diccionario de la Real Academia Española (RAE) [29], donde es posible leer lo siguiente:

> **optimizar** (de óptimo e -izar, *verbo transitivo*) Buscar la mejor manera de realizar una actividad.

Bajo esta idea y considerando que la toma de decisión corresponde a definir cómo se realiza una actividad, es posible pensar que un proceso de toma de decisiones ideal es aquel que no solo permite encontrar una manera de resolver una situación, sino la mejor manera de hacerlo, es decir, encontrar la mejor solución, según un indicador determinado o un grupo de estos.

La formalidad y lenguaje que posee la matemática, permiten utilizarla como una herramienta para modelar y resolver situaciones en donde se requiere optimizar. Según lo anterior por tanto, en estas situaciones se puede utilizar el conocimiento matemático con el fin de encontrar la mejor solución para una situación dada. Estas situaciones, cuando se aplica este conocimiento, suelen llamarse comúnmente problemas de optimización matemática. La solución obtenida proviene de un grupo o conjunto de soluciones que resuelven el problema, las cuales son evaluadas para determinar cual es la mejor. Esta evaluación se hace matemáticamente, es decir, mediante una fórmula, la obtención de un indicador o un grupo de ellos. La mejor de todas estas soluciones posibles se considera la solución óptima del problema. Matemáticamente esto es posible expresarlo de la siguiente forma:

$$P = \begin{cases} \text{Optimizar} & f(\mathbf{x}) \\ \text{Sujeto a:} & \mathbf{x} \in S \subseteq \mathbb{K}^n; \end{cases} \tag{2.1}$$

donde P es un problema de optimización que busca minimizar o maximizar una función objetivo $f(\mathbf{x})$, la cual asigna un valor a cada solución $\mathbf{x} = \{x_1, x_2, x_3, ..., x_n\}$. Las soluciones $\mathbf{x}$ pertenecen a un conjunto de soluciones factibles S, que a su vez pertenece o vive en un conjunto $\mathbb{K}^n$, el cual usualmente puede ser el conjunto $\mathbb{N}$ o $\mathbb{R}$ de n dimensiones. La idea principal de este tipo de problemas es encontrar una solución que pertenezca a S, tal que maximice o minimice el valor de $f(\mathbf{x})$, es decir, encontrar un $\mathbf{x}^* \in S : f(\mathbf{x}^*) \geq f(\mathbf{x}), \forall \mathbf{x} \in S$ si el problema es de maximización o en su defecto, encontrar un $\mathbf{x}^* \in S : f(\mathbf{x}^*) \leq f(\mathbf{x}), \forall \mathbf{x} \in S$ si el problema es de minimización.

En optimización matemática, existen diferentes tipos de problemas, los cuales pueden ser clasificados según sus características, como por ejemplo la presentada en la Figura 2.1. Esto es importante, pues define el tipo de técnica o estrategia a utilizar para resolver el problema que se espera resolver.

A un nivel más detallado, los tipos de optimización que se han utilizado en esta tesis doctoral son los que se presentan a continuación.

2.2.1. Optimización lineal

La optimización lineal (también conocida como programación lineal), es un tipo de optimización donde las restricciones (o inecuaciones) y la función objetivo son funciones lineales en sus variables. Típicamente, este tipo de formulaciones, se pueden escribir de la siguiente forma para un problema de minimización:

$$minimizar \;\; z = \mathbf{c}^{\mathbf{T}}\mathbf{x} \tag{2.2}$$

$$sujeto\ a: \;\; A\mathbf{x} \leq \mathbf{b} \tag{2.3}$$

$$\mathbf{x} \in \mathbb{R}^n_{\geq 0}; \tag{2.4}$$

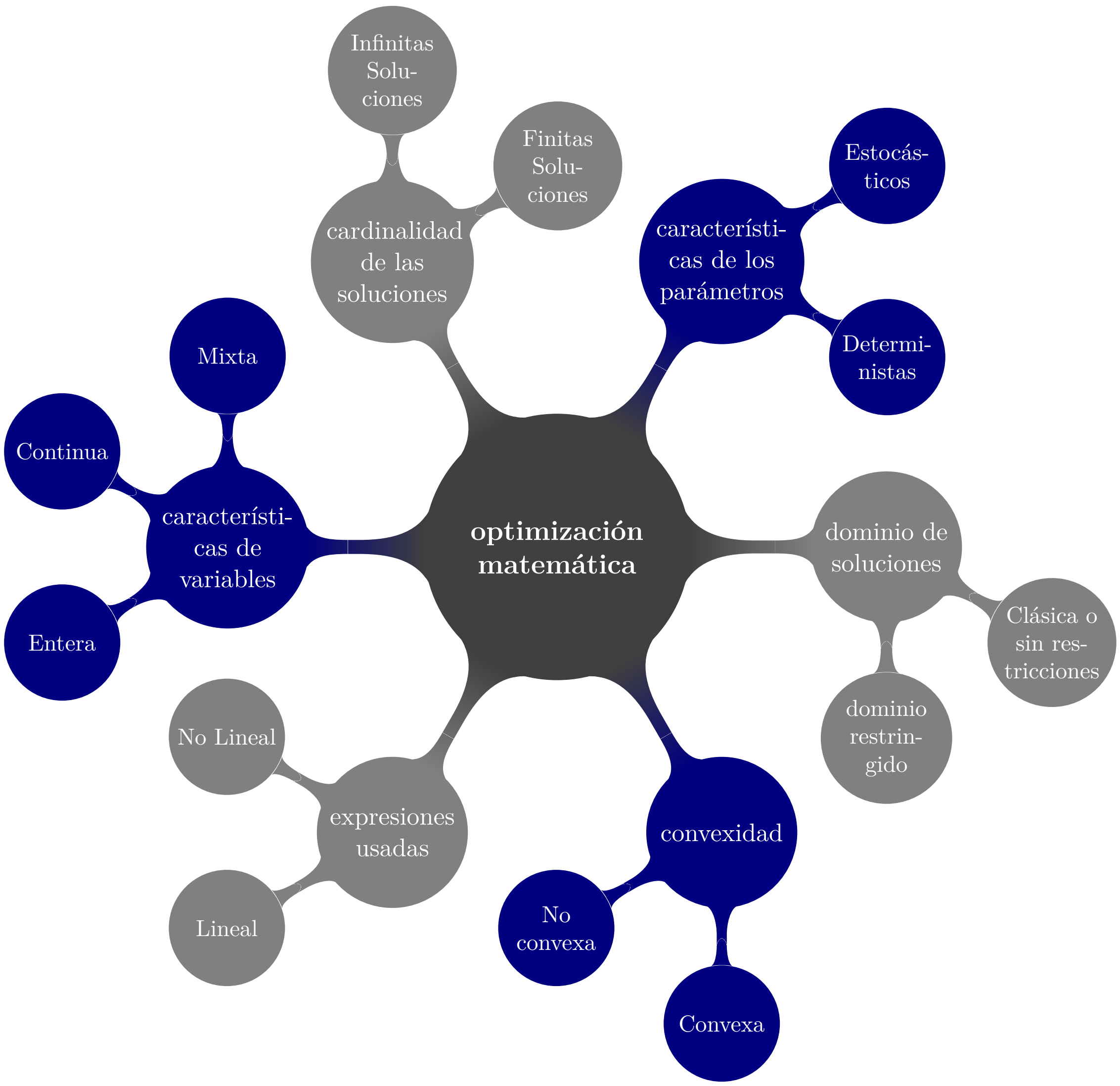


Figura 2.1. Categorías que describen un problema de optimización

donde z es el valor de una función objetivo que depende de las variables en $\mathbf{x}$, $\mathbf{x}$ es el vector que contiene las variables que son parte del problema, $\mathbf{c^T}$ es el vector de costos asociados a cada una de las variables, A es la matriz con los coeficientes correspondientes a cada variable en cada restricción, y $\mathbf{b}$ es el vector que contiene el valor de cada recurso que es modelado por cada restricción. La ecuación 2.4 determina el dominio donde habitan las variables del problema. Además, las restricciones 2.3 y 2.4 definen el espacio de soluciones factibles, el cual, en el caso de estar acotado, también recibe el nombre de poliedro [30].

Siguiendo lo presentado en la figura 2.1, un problema puede ser descrito por más de una categoría. A continuación se presentan las dos mas importantes en esta tesis.

2.2.2. Programación entera y programación estocástica

En algunos casos, los problemas de programación lineal incluyen problemas de inversión de maquinarias o elementos que no pueden ser modelados por cantidades decimales, como lo son la cantidad de plantas de energía que debo construir para maximizar mis utilidades. Este tipo de problemas, las soluciones son caracterizadas por la naturaleza de $\mathbf{x}$, pues incluyen variables que pertenecen al conjunto $\mathbb{N}$, es decir que la ecuación 2.4 se transforma en $\mathbf{x} \in \mathbb{N}^n_{\geq 0}$ y es posible resolverlos mediante programación entera o mixta.

La programación entera o mixta se encarga de resolver problemas en donde todas las variables (programación entera) o algunas (programación mixta) deben tomar valores enteros. Un caso particular de este tipo de optimización es aquella en donde las variables pueden tomar valores en el conjunto $\{0, 1\}$, i.e. $\mathbf{x} \in \{0, 1\}^n$. Este tipo de variables se llaman binarias y comúnmente se encuentran en problemas en donde se debe decidir entre 2 opciones, generalmente entre sí o no, como lo son el instalar o no una subestación, el conectar o no una central de generación a una subestación y otras similares.

En el arte de modelar una situación y hacer que esta sea representativa de la realidad, generalmente se deben incorporar en el modelo parámetros que no son estables en el tiempo, como pueden ser, precios de combustible, costos de algún insumo, precios de venta de energía, disponibilidad de algún recurso, disrupciones en el sistema a modelar, demanda, etcétera. Este tipo de entradas se llaman parámetros estocásticos y el tipo de optimización que se encarga de trabajar con la incertidumbre asociada a ella es la programación estocástica [31].

Existen diferentes formas de modelar la variabilidad de los datos. A continuación se presentan las mas relevantes.

escenarios Consiste en generar un conjuntos de escenarios posibles, cada uno de ellos asociado a una probabilidad de ocurrencia. Esta probabilidad generalmente sigue una función de probabilidades, relacionando los escenarios.

modelamiento por etapas En ciertos casos, entre algunas variables existe una dependencia jerárquica, siendo necesario encontrar el valor de unas primero, para poder determinar el correspondiente a otras. Un claro ejemplo de esto en sistemas de energía es (i) la decisión de instalar o no una central de energía y (ii) la decisión de cuánta capacidad de generación instalar. Determinar la capacidad a instalar depende claramente de la decisión (i), pues si esta es “no”, la decisión (ii) es igual a cero.

grados de arrepentimiento o manejo del riesgo Dentro de esta categoría entran variadas técnicas o formulaciones como lo son el valor en riesgo (*value at risk* o *VaR*) , el mínimo máximo arrepentimiento (*min-max regret*) e incluso la optimización robusta (*robust optimization*). La filosofía detrás de la optimización robusta es la existencia de una aversión total al riesgo, por tanto se utiliza esta aproximación cuando se necesita una solución que sea la mejor en el peor de los escenarios posibles. El mínimo máximo arrepentimiento se basa en la idea de encontrar una solución que minimize el arrepentimiento en el peor caso posible. Por otro lado el valor en riesgo, como su nombre lo indica, es la probabilidad de que un parámetro o valor varíe de forma negativa en una cantidad dada durante un horizonte temporal dado. Si bien este concepto proviene del mercado de valores, es posible aplicarlo a diferentes áreas en donde exista un riesgo de variabilidad de un parámetro o valor.

2.3. Problemas de optimización

Debido al simple hecho de que este trabajo está motivado por un sistema complejo de ingeniería, en esta tesis doctoral se utiliza teoría desarrollada para distintos problemas de programación lineal, los cuales se procede a describir a continuación.

2.3.1. Problemas de localización

Los problemas de localización o *Facility Location Problems* consisten en ubicar de forma óptima una o varias instalaciones, con el fin de cubrir una demanda particular. Generalmente se minimizan los costos de instalación, los costos de transporte o similares. Dos libros al respecto son el de Laporte, Saldanha da Gama y Nickel [32], y el de Chen, Batson y Dang [33] en donde se pueden revisar en profundidad distintas versiones de este problema como las siguientes:

***p*-mediano** conocidos en inglés como *p-Median Problems* son un tipo de problemas en los cuales se busca localizar p instalaciones (*facilities*) en un grafo, con el fin de minimizar las distancias ponderadas promedio entre los nodos de demanda y las instalaciones. Este tipo de problemas son el caso mas simple de los problemas de localización, especialmente cuando se resuelven en un árbol.

cobertura conocidos en inglés como *Covering Location Problems* son problemas en los cuales se busca determinar la mejor posición para un grupo de instalaciones con el fin de maximizar la cobertura sobre una población o región determinada.

capacitado conocidos en inglés como *Capacitated Facility Location Problems* son problemas en los cuales la capacidad de la instalaciones está limitada por un recurso. Este recurso puede ser aquel que se demanda, o bien uno del tipo presupuestario.

no-capacitado conocidos en inglés como *Uncapacitated Facility Location Problems*, son problemas en los cuales no existe un límite de recurso o presupuesto, por lo tanto el resultado son todas aquellas ubicaciones que logran satisfacer la demanda y que lo hacen de una forma óptima.

centros de distribución conocidos en inglés como *Hub Location Problems*, son problemas que consisten en localizar de forma óptima, centros de distribución, que permitan concentrar los elementos que se transportan en una red y distribuirlos a el o los destinos correspondientes. Un ejemplo de esto en el área de redes de energía son las subestaciones.

2.3.2. Problemas de dimensionamiento

Los problemas de dimensionamiento en el área de energía, están relacionados con determinar la capacidad de plantas generadoras, subestaciones y sistemas de almacenamiento de energía. La idea principal es determinar la capacidad óptima de generación (plantas), almacenamiento o concentración (subestaciones) de energía. Si bien la teoría correspondiente a los problemas de dimensionamiento comenzó a desarrollarse principalmente mediante aplicaciones a almacenes (en inglés *warehouses*), existen distintos ejemplos en el campo de la energía en donde se resuelven este tipo de problemas [12–14]. Además, generalmente están inspirados por la existencia de parámetros estocásticos, como pueden ser la disponibilidad del recurso, valores de venta de energía o valores de demanda.

2.3.3. Problemas de pre-despacho y despacho

Los problemas de pre-despacho y despacho o *Unit Commitment Problems* (UCP) en inglés, son generalmente problemas asociados a la gestión de una red de energía, donde el interés es minimizar los costos de operación y generación de energía, buscando satisfacer una demanda en un periodo de tiempo a estudiar. Si se tiene una matriz de energía, con un grupo de generadores a operar, el problema de optimización es posible describirlo mediante el siguiente esquema:

- Minimizar los costos
  - de encendido y apagado de las centrales de generación y,
  - de generación de energía
- sujeto a:
  - restricciones sobre la toma de decisión de qué centrales operar.
  - restricciones sobre cómo operar las centrales de generación.

Como es posible intuir, el UCP sigue una aproximación de un problema de 2 etapas. Además, en el UCP se busca minimizar los costos de pre-despacho y despacho de energía mediante una formulación, en la cual existen unos costos de operación, según una estrategia de encendido y apagado determinada. Esto se determina para un grupo de generadores en una red o sistema interconectado. Existen diferentes versiones de este problema, como es posible ver en [10–12, 34, 35], en donde siempre se ha resuelto este problema desde el punto de vista del operador de la red de energía.

2.4. Análisis y pronóstico de datos

Al afrontar un problema, se hace necesario comprender sus características principales, con el fin de caracterizar la situación y poder realizar modelos acorde con ella. De esta forma, se pueden saber cuáles son las variables independientes o explicativas del problema, cuáles son las relaciones entre las variables y si existen tendencias que permitan pronosticar el comportamiento futuro de aquellas variables. Si bien existen distintas opciones para analizar datos, como lo son la minería de datos o el aprendizaje automático (conocido mundialmente como *Machine Learning*), en esta tesis se utilizan herramientas relacionadas con la estadística y la econometría. En este contexto, algunas herramientas características son la aproximación por mínimos cuadrados [36], estadística descriptiva y técnicas de determinación de parámetros para distribuciones estadísticas o series de tiempo, las cuales son conocidas como técnicas de *bootstrapping* [37].

Capítulo 3

Metodología

En el presente capítulo se procederá a describir las herramientas metodológicas utilizadas para desarrollar esta tesis, las cuales son relativas a las matemáticas debido a la naturaleza de este trabajo.

3.1. Proceso de modelización matemática

Si bien es lógico pensar que el principal método seguido en un proceso investigativo es el método científico, los problemas de investigación motivados por algún sistema complejo de ingeniería presentan desafíos particulares. Estos desafíos se pueden abordar, además del método científico, mediante la modelización matemática.

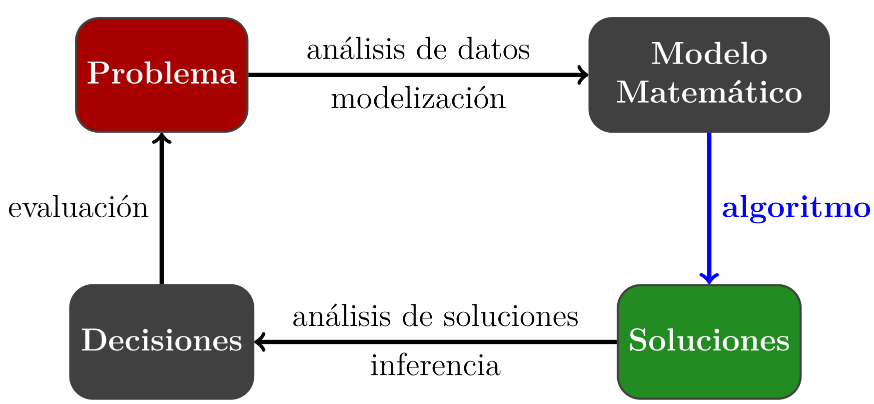


Figura 3.1. Proceso para resolver un problema mediante modelización matemática

En la figura 3.1 se pueden apreciar 4 etapas que es posible describirlas de la siguiente forma:

Problema Es la situación del mundo real que debe ser solucionada. Casi siempre la problemática pertenece a un sistema complejo con muchas variables interactuando entre ellas.

Modelo matemático Es el resultado del proceso de modelización, el cual consiste en realizar una abstracción representativa de la realidad. Se deben contemplar las principales características del problema o sistema a intervenir. Además se debe realizar un compromiso entre la representatividad y la facilidad de resolución del modelo, pues a mas representación, mayor complejidad de resolución en términos de recursos teóricos, computacionales y de tiempo.

Soluciones Son el resultado del proceso de resolución del modelo matemático, mediante un algoritmo. Las soluciones deben interpretarse con el fin de entender que información nos entregan.

Decisiones Son el resultado del proceso de análisis, el cual consiste en entender las soluciones y sus limitaciones en cuanto a la implementación de estas en el mundo real. Determinar la validez de las soluciones, desde el punto de vista técnico, económico y temporal se hace clave en este punto del método, permitiendo la evaluación final en el mundo real para aplicarlas posteriormente al problema.

Ahora bien, el trabajo resumido en la sección 4.1 engloba la primera etapa y el proceso de análisis de datos, el segundo y el tercer trabajo, resumidos en las secciones 4.2 y 4.3 respectivamente, están ubicados entre la primera etapa y los procesos de análisis de soluciones e inferencia. Debido a la imposibilidad de asociar esta tesis con una empresa del sector, la última parte (decisiones y evaluación) no forma parte de este trabajo.

3.2. Recolección de datos

Debido a que este trabajo se centra en una problemática relativa a los sistemas de energía, la gran mayoría de las fuentes de datos provienen del mercado energético. Esto permite incluir en la etapa de modelización, las características principales del sistema que se estudia en esta tesis.

Los costos de inversión y las características técnicas de la red, al igual que los componentes a considerar, fueron recolectados de trabajos encontrados en la literatura [38–41] como también en fuentes oficiales del operador de la red eléctrica [42–44] , que en este caso es la red chilena. Estos valores fueron contrastados con una empresa del sector, mediante conversaciones con los ingenieros y personal técnico [45]. Otros parámetros técnicos de componentes, como también valores relativos al diseño de plantas de energía renovable, fueron obtenidos también de expertos que trabajan en el sector [45]. Además, con el fin de asegurar una correcta topología de la red en la cual se realizaron los estudios, se recolectaron mapas unilineares de la red eléctrica chilena [42].

3.2.1. Tratamiento de datos

Una vez recolectados, se procedió a limpiar y seleccionar los datos. Para esto se generaron códigos en lenguaje de programación C++, con el fin de leer archivos tipo ".clx" y ".txt", limpiarlos de posibles errores de escritura y pasarlos a archivos con un formato unificado tipo ".txt"para su lectura por los códigos que contienen las formulaciones matemáticas.

Seguidamente, en algunos puntos donde los resultados parecían extraños, se contrastó con los mapas de conexión eléctricos mediante inspección ocular de estos.

Parte de este trabajo se resume en la sección 4.1, de donde resulta la base para generar instancias de prueba utilizadas en los trabajos resumidos en las secciones 4.2 y 4.3, según la topología y características de la red eléctrica.

Para generar estas instancias se utilizaron los siguientes métodos:

Costos de instalación mediante regresión lineal, se obtuvieron distintas curvas lineales que relacionan capacidad a instalar con costo. Esto se realizó tanto para los sistemas de generación, como los de transmisión y almacenaje de energía.

Precios de venta de energía mediante técnicas de muestreo (*boostraping*) [37, 46, 47], se generaron perfiles anuales para el precio de venta por hora, usando datos de precios de venta comprendidos entre los años 2015-2016 [43]. Estos años fueron seleccionados debido a que se necesitan al menos datos de 2 años para aplicar las técnicas de muestreo, siendo estos los mas actuales al momento de comenzar este trabajo.

Demanda al igual que con los precios de venta, se desarrollaron perfiles de demanda con el fin de obtener la actividad existente en la red. Para este punto se utilizaron técnicas de muestreo con datos entre 2015-2016, debido a los mismos motivos presentados para los precios de venta.

3.3. Formulación matemática

Para la formulación matemática, se utilizaron métodos relativos a la teoría de programación matemática mixta (*mixed integer linear programming* o *MILP*), descrita principalmente el capítulo 2. Además, se han incluido métodos relativos a evaluación de proyectos e ingeniería económica [48], como son los conceptos de movilidad de capitales a través del tiempo y el valor anual neto (*Net present value* o *NPV*). Para la resolución de los modelos, se ha propuesto una heurística, programada en lenguaje C++, como la presentada en el trabajo de la sección 4.3, además del uso del software ILOG® CPLEX® 12.6.3.

3.4. Análisis de resultados

Los trabajos presentados en las secciones 4.1, 4.2 y 4.3 tienen como resultados principales una base de datos, una formulación matemática y un método de resolución que mejora la exactitud de la evaluación económica. Sin embargo, el rendimiento de distintas instancias posibles se analiza mediante indicadores como son la tasa interna de retorno (*Internal rate of return* o *IRR*), el horizonte temporal de retorno de la inversión (*Payback*) y el valor actual neto (*Net present value* o *NPV*). Estos indicadores se representan gráficamente con el fin de comparar las distintas instancias y escenarios posibles para ayudar a la toma de decisiones asociada a una estrategia de inversión. Finalmente, se realiza también un análisis de estructura de costos con el fin de ver como estos varían ante diferentes estímulos.

Capítulo 4

Descripción de Publicaciones

En este capítulo, se presentan 3 trabajos que conforman el núcleo esta tesis. Siguiendo un orden lógico, la primera publicación envuelve a la base de datos topológica que sirve para la generación de instancias de este trabajo [49]. La segunda publicación corresponde al primer modelo desarrollado en el marco de este trabajo doctoral [50]. El tercer trabajo utiliza el modelo de la segunda publicación, con el fin de desarrollar un modelo mas complejo y completo en cuando a la cantidad de variables que puede manejar para la toma de decisiones, además de presentar una estrategia de resolución.

Como es posible ver a continuación, la denominación de las secciones corresponde al título de cada uno de los trabajos presentados en el apéndice de esta tesis.

4.1. In-depth data on the network structure and hourly activity of the central chilean power grid

Introducción Para poder realizar investigación y ciencia aplicada, es necesario poseer datos de buena calidad, esto significa, que la procedencia de estos sea de fuentes de información reales, con el fin de replicar y modelar de la mejor forma los fenómenos a estudiar. En este sentido, poseer una base de datos basada en una red eléctrica real permite entender cómo ciertas disrupciones o nuevos elementos, como proyectos de energía, impactan en la estructura existente y se influencian mutuamente desde el punto de vista del servicio, la estabilidad o la sustentabilidad.

Objetivos El principal objetivo es confeccionar una base de datos de una red real curada, que represente la red de generación y distribución de energía eléctrica chilena, para usos principalmente científicos. La base de datos debe ser capaz de recoger las características técnicas y topológicas reales existentes en la matriz energética.

Métodos Para realizar este trabajo, se consultaron 3 fuentes de información distinta, para luego utilizar técnicas de análisis de datos además de teoría de bases de datos, con el fin de lograr una representación real y unificada de todo el sistema chileno. Además se utilizaron técnicas de programación para evitar cualquier potencial error humano en el proceso.

Resultados y conclusiones Entre los resultados principales de este trabajo, se encuentran la base de datos que representa correctamente las características técnicas y económicas de topología y funcionamiento de la red eléctrica chilena. Esta base de datos consiste en 3 instancias desarrolladas para ser usada con fines diferentes. Parte de los posibles usos de esta base de datos consiste en el análisis de modelos de estabilidad de sistema, como también modelos de inversión en el área de energía. Otros usos son el análisis de clústers, estudios asociados a la robustez de la red, modelos de ayuda a la toma de decisiones de nivel estratégico, etc.

4.2. An optimization framework for investment evaluation of complex renewable energy systems

Introducción Segun lo presentado en el capítulo 1, los esfuerzos para generar un marco de evaluación de proyectos de energía renovable, han incluido las características técnicas asociadas a decisiones estratégicas de forma separada. Esto último significa que las estrategias presentadas no están optimizadas entre si, al no estar integradas en un mismo modelo. Además, estas aproximaciones se han trabajado, típicamente, desde el punto de vista del operador del sistema y no desde el inversionista.

Objetivos El objetivo de este trabajo es contribuir con un modelo de programación lineal mixta (MILP), que permita ayudar la toma de decisiones de nivel estratégico, correspondientes a la inversión de proyectos de energía renovable. El modelo debe contemplar de forma integrada y simultánea, las distintas decisiones técnicas que afectan la rentabilidad de un proyecto y que lo transforman en un sistema complejo. Finalmente se evalúa la manejabilidad resolutiva del problema desde el punto de vista computacional.

Métodos Programación mixta, técnicas de *boostraping*, regresión lineal, evaluación de proyectos mediante escenarios y consulta a expertos.

Resultados y conclusiones El principal resultado de la publicación es un MILP que permite hacer evaluaciones de proyectos complejos de forma preliminar. Además se presentan como resultados las relaciones de costos de componentes, la instancia con la cual se prueba el modelo y los perfiles que caracterizan los precios de venta de energía para la zona estudiada según los escenarios de evaluación. Algunas conclusiones destacables son que el problema presentado permite utilizar perfiles agregados de datos con el fin de obtener un mejor tiempo de computo, sin afectar los resultados de manera relevante. Además se determina que los sistemas de almacenamiento de energía, tienen una mejor oportunidad, si la decisión de inversión se determina conjuntamente con los sistemas de generación asociados, así como su tamaño y posición optimizados para ello. Además se explora el impacto de distintas tecnologías de almacenamiento de energía, dentro de escenarios de inversión, mediante instancias curadas gracias a los datos del trabajo presentado en la sección 4.1. Igualmente, se presenta la capacidad del modelo para comparar tecnologías de almacenamiento de energía y de realizar análisis de sensibilidad para apoyar la toma de decisiones.

4.3. Towards a complex investment evaluation framework for renewable energy systems: A 2-stage-heuristic approach

Introducción Siguiendo el trabajo presentado en la sección 4.2, Los proyectos de energía generan un impacto en la red en la cual serán instalados. Este impacto es no solo de carácter técnico, al aumentar la capacidad de toda la red, sino que también de caracter económico, al tener una influencia directa en los costos operativos del sistema y por ende en los precios de venta de energía de la red. Lo anterior implica que se hace necesario contemplar la componente operacional de la red donde se instalará un proyecto al momento de su evaluación, permitiendo un cálculo mas preciso y realista en cuando a las condiciones de funcionamiento del sistema a instalar.

Objetivos El principal objetivo es desarrollar una metodología de evaluación que contemple, además de las complejidades técnicas de nivel estratégico, las complejidades económicas de nivel operativo en la evaluación de proyectos de energías renovables y almacenamiento de energía.

Métodos Lo anterior se realiza mediante un marco heurístico que resuelve de manera iterativa 2 modelos en una configuración de 2 etapas: uno de inversión, presentado en el trabajo de la sección 4.2 y otro comúnmente utilizado para determinar la operación de una red eléctrica. Las herramientas a utilizar son principalmente programación mixta, técnicas de *boostraping*, evaluación de proyectos y formulación heurística.

Resultados y conclusiones El principal resultado es la metodología desarrollada para contemplar las características operacionales de la red en la evaluación de una estrategia de inversión en sistemas de energía. Además se presenta el uso de la metodología aplicándola a un caso de estudio obtenido del trabajo presentado en la sección 4.1. Luego se exploran las posibilidades de uso bajo diferentes escenarios como lo son la existencia de un presupuesto acotado (desde el punto de vista de un inversor) y la determinación de un nivel de subsidio con el fin de hacer mas atractiva una región con el fin de mejorar los niveles de sustentabilidad presentes en ella (desde el punto de vista de un generador de políticas públicas).

Capítulo 5

Conclusiones Generales y Dirección Futura de Investigación

En esta tesis se han estudiado sistemas complejos de energía, específicamente los relativos a las energías renovables, abordando las siguientes características:

- Levantamiento de una base de datos topológica de la red eléctrica chilena.
- Dimensionamiento de sistemas de generación y de sistemas de almacenamiento.
- Localización de sistemas de almacenamiento, generación y de subestaciones.
- Interconexión de componentes de una red de energía.
- Operación de una red de energía eléctrica basada en energía renovable.
- Determinación de los precios de venta de energía.
- Inclusión de parámetros técnicos y tecnológicos de componentes de sistemas de energía renovable.

La lista anterior permite visualizar la complejidad presente en los sistemas de energía y el tipo de desafío que implica modelarlos con el fin de obtener respuestas aplicables al mundo real. De esta forma se ha logrado determinar que las tecnologías renovables están en un nivel de madurez tal, que permite incluirlas de forma atractiva en proyectos sin bonificación ni subvenciones especiales. Esto es muy importante, pensando que se han evaluado en un contexto de libre competencia. La clave para lograr esto es incluir en el proceso de evaluación, modelos de optimización que incluyan todos los componentes a considerar en la inversión. Los modelos presentados en este trabajo por tanto, son de una gran importancia para permitir el desarrollo de las soluciones renovables, permitiendo mejorar la sustentabilidad de una región. Lo anterior se hace aún mas importante para el desarrollo de las soluciones de almacenamiento de energía, puesto que muchas están aún en etapa de desarrollo.

Por otro lado, las instancias a evaluar, deben contemplar las características particulares de terreno, como lo son la cantidad de recurso energético disponible y topología del lugar, además de la red existente. Una caracterización adecuada de una zona, permite hacer evaluaciones acertadas y que son más fáciles de implementar y tomar en cuenta a la hora de decidir.

Desde el punto de vista económico, se ha demostrado el impacto de los proyectos de energía renovable en los costos de operación y precios de venta. Esto implica que este impacto debe considerarse en la evaluación de proyectos de esta índole con el fin de aumentar las posibilidades de éxito de la inversión. Además, esta situación desemboca en una red mas limpia, sustentable y una energía de precios mas accesibles.

Por otra parte, se ha demostrado la utilidad del marco propuesto con fines asociados a políticas de desarrollo, mediante la exploración de zonas con el objetivo de determinar polos de desarrollo energético y si fuese necesario, tomar medidas para mejorar el atractivo de estas zonas a la hora de invertir, por ejemplo, por medio de subsidios o mediante una planificación territorial particular con ese fin. Se ha demostrado también la posibilidad de usar este marco para determinar etapas de inversión a largo plazo. Esto aumenta las posibilidades de ejecución de proyectos de energía renovable.

De acuerdo a lo expuesto hasta ahora, podemos visualizar que se hace necesaria una extensión de la base de datos del primer trabajo presentado en esta tesis, incluyendo mapas de recursos renovables, permitiendo generar instancias completas y complejas en cuanto a contemplar las características principales para el desarrollo de proyectos de energías renovables. En este sentido, con el concepto "mapas de recursos renovables" se quiere expresar que en una misma base de datos se contemple no solo la configuración topológica de una red, sino también la cantidad de recurso de energía disponible proveniente de más de una fuente renovable a la vez.

Además, debido a lo complejo del sistema, un modelo basado en una estrategia de simulación-optimización parece ser el siguiente paso lógico, con el fin de contemplar más parámetros y su correspondiente componente estocástica. Un esfuerzo en esta línea permitiría estudiar un mayor grado de complejidad del problema.

Luego, también se ve como un camino atractivo, realizar estudios sobre la estabilidad de la red y la respuesta del sistema al incluir disrupciones mediante la creación adecuada de instancias de estudio. Un ejemplo de esto es el estudio a nivel operativo de los efectos de nubes sobre los generadores fotovoltaicos. Al existir sistemas de almacenamiento de energía, esto puede controlarse adecuadamente, pero no existe claridad sobre el impacto de estos fenómenos ni de como optimizar la operación de los sistemas para tales efectos sin perder el rendimiento esperado, desde el punto de vista económico y de satisfacción de demanda.

Subsecuentemente, contemplar otras fuentes de energía como la eólica o la mareomotriz se ve como un posible campo de investigación futura debido a que posiblemente las conclusiones asociadas a esas tecnologías pueden ser distintas en cuanto a la competitividad de estas, presentando desafíos de modelamiento que requieren otros modelos diferentes a los presentados en esta tesis.

Finalmente, se hace necesario desarrollar investigación en todas las áreas del conocimiento, que influencian de alguna forma los sistemas de energía, con el fin de generar un modelo de sociedad sustentable. Los sistemas complejos de energía están presentes en todas las áreas del quehacer humano y conforman un área prioritaria a la hora de visualizar y definir el futuro de la humanidad.

Bibliografía

[1] R. Carlyle. How Big Are The Currently Known Oil Reserves And What Are The Chances Of Finding New Ones? - Forbes. https://www.forbes.com/sites/quora/2013/03/27/how-big-are-the-currently-known-oil-reserves-and-what-are-the-chances-of-finding-new-ones/#a907ca2593da. visited 2016-07-21.

[2] A. Tully. How long will world's oil reserves last? 53 years, says BP. http://www.csmonitor.com/Environment/Energy-Voices/2014/0714/How-long-will-world-s-oil-reserves-last-53-years-says-BP. visited 2016-07-21.

[3] BP British Petroleum. Oil reserves — About BP — BP Global. http://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy/oil/oil-reserves.html. visited 2016-07-21.

[4] IEA. Key World Energy Statistics 2015. Technical report, International Energy Agency, 2015.

[5] UN - United Nations. Energy - United Nations Sustainable Development. https://www.un.org/sustainabledevelopment/energy/. visited 2019-02-09.

[6] UN - United Nations. About the Sustainable Development Goals - United Nations Sustainable Development. https://www.un.org/sustainabledevelopment/sustainable-development-goals/. visited 2019-02-09.

[7] I. Dincer. Renewable energy and sustainable development: a crucial review. *Renewable and sustainable energy reviews*, 4(2):157–175, 2000.

[8] N. Nakićenović, A. Grübler, and A. McDonald. *Global energy perspectives*. Cambridge University Press Cambridge, 1998.

[9] S. Ahmad and R.M. Tahar. Selection of renewable energy sources for sustainable development of electricity generation system using analytic hierarchy process: A case of malaysia. *Renewable energy*, 63:458–466, 2014.

[10] J. Kim and W. Powell. Optimal energy commitments with storage and intermittent supply. *Operations research*, 59(6):1347–1360, 2011.

[11] N. Koltsaklis and M. Georgiadis. A multi-period, multi-regional generation expansion planning model incorporating unit commitment constraints. *Applied Energy*, 158:310–331, 2015.

[12] C. Suazo-Martínez, E. Pereira-Bonvallet, and R. Palma-Behnke. A simulation framework for optimal energy storage sizing. *Energies*, 7(5):3033–3055, 2014.

[13] K. Bradbury, L. Pratson, and D. Patiño-Echeverri. Economic viability of energy storage systems based on price arbitrage potential in real-time us electricity markets. *Applied Energy*, 114:512–519, 2014.

[14] A. Berrada and K. Loudiyi. Operation, sizing, and economic evaluation of storage for solar and wind power plants. *Renewable and Sustainable Energy Reviews*, 59:1117–1129, 2016.

[15] P. Xiong and C. Singh. Optimal planning of storage in power systems integrated with wind power generation. *IEEE Transactions on Sustainable Energy*, 7(1):232–240, 2016.

[16] S. Dehghan and N. Amjady. Robust transmission and energy storage expansion planning in wind farm-integrated power systems considering transmission switching. *IEEE Transactions on Sustainable Energy*, 7(2):765–774, 2016.

[17] C. Rahmann and R. Palma-Behnke. Optimal allocation of wind turbines by considering transmission security constraints and power system stability. *Energies*, 6(1):294–311, 2013.

[18] D. Jakus, J. Krstulovic, and J. Vasilj. Algorithm for optimal wind power plant capacity allocation in areas with limited transmission capacity. *International Transactions on Electrical Energy Systems*, 24(10):1505–1520, 2014.

[19] J. Asamer, M. Reinthaler, M. Ruthmair, M. Straub, and J. Puchinger. Optimizing charging station locations for urban taxi providers. *Transportation Research Part A: Policy and Practice*, 85:233–246, 2016.

[20] S. Li, Y. Huang, and S. Mason. A multi-period optimization model for the deployment of public electric vehicle charging stations on network. *Transportation Research Part C: Emerging Technologies*, 65:128–143, 2016.

[21] R. Jiang, J. Wang, and Y. Guan. Robust unit commitment with wind power and pumped storage hydro. *IEEE Transactions on Power Systems*, 27(2):800–810, 2012.

[22] E. Álvarez-Miranda, C. Campos-Valdés, and C. Rahmann. Two-stage robust UC including a novel scenario-based uncertainty model for wind power applications. *Energy Conversion and Management*, 101:94–105, 2015.

[23] C. MacRae, M. Ozlen, A. Ernst, and S. Behrens. Locating and sizing energy storage systems for distribution feeder expansion planning. In *TENCON 2015-2015 IEEE Region 10 Conference*, pages 1–6. IEEE, 2015.

[24] G. Brundtland, M. Khalid, S. Agnelli, S. Al-Athel, P. Casanova, B. Chidzero, L. Padika, V. Hauff, I. Lang, M. Shijun, M.M. de Botero, N. Singh, P. Nogueira-Neto, S. Okita, S.S. Ramphal, W.D. Ruckelshaus, M. Sahnoun, E. Salim, B. Shaib, V. Sokolov, J. Stanovnik, M. Strong, and J. MacNeill. Report of the World Commission on Environment and Development: Our common future ('Brundtland report'). Technical report, UN - United Nations, 1987. http://www.un-documents.net/our-common-future.pdf visited 2019-02-09.

[25] Ibrahim Dincer and Marc A Rosen. *Exergy: energy, environment and sustainable development*. Newnes, 2012.

[26] G. Wall. Exergetics. http://exergy.se/ftp/exergetics.pdf. visited 2019-02-09.

[27] M.J. Goedkoop. The eco-indicator 99 a damage oriented method for life cycle impact assessment methodology report. *Pre Concultants*, 1999.

[28] O. Jolliet, M. Margni, R. Charles, S. Humbert, J. Payet, G. Rebitzer, and R. Rosenbaum. Impact 2002+: a new life cycle impact assessment methodology. *The international journal of life cycle assessment*, 8(6):324, 2003.

[29] RAE Real Academia Española. optimizar — Definición de optimizar. http://dle.rae.es/srv/fetch?id=R7YxPPp. visited 2018-09-16.

[30] I. Griva, S.G. Nash, and A. Sofer. *Linear and nonlinear optimization*, volume 108. Siam, 2009.

[31] J.R. Birge and F. Louveaux. *Introduction to stochastic programming*. Springer Science & Business Media, 2011.

[32] G. Laporte, S. Nickel, and F. Saldanha da Gama, editors. *Location science*. Springer, 1 edition, 2015.

[33] D. Chen, R.G. Batson, and Y. Dang. *Applied integer programming: modeling and solution*. John Wiley & Sons, 2011.

[34] A. Tuohy, P. Meibom, E. Denny, and M. O'Malley. Unit commitment for systems with significant wind penetration. *IEEE Transactions on Power Systems*, 24(2):592–601, 2009.

[35] F. Pérez-Illanes, E. Álvarez-Miranda, C. Rahmann, and C. Campos-Valdés. Robust unit commitment including frequency stability constraints. *Energies*, 9(11):957, 2016.

[36] S.M. Stigler. Gauss and the invention of least squares. *The Annals of Statistics*, pages 465–474, 1981.

[37] B. Efron. Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics*, pages 569–593. Springer, 1992.

[38] ABB AB, SE-771 80 Ludvika, Sweden. http://www.abb.com/hvdc. *It's time to connect - HVDC Light ©*, 7 edition, 2013. POW-0038.

[39] European Network of Transmission System Operators for Electricity. Offshore Transmission Technology. Technical Report prepared by the regional group north sea for the NSCOGI, ENTSOE, Brussels - Belgium, November 2011.

[40] C. Lancheros. Transmission systems for offshore wind farms: a technical, enviromental and economic assessment. Master's thesis, Hamburg University of Technology, Hamburg, April 2013.

[41] B. Van Eeckhout, D. Van Hertem, M. Reza, K. Srivastava, and R. Belmans. Economic comparison of VSC, HVDC and HVAC as transmission system for a 300 MW offshore wind farm. *European Transactions on Electrical Power*, 20(5):661–671, 2010.

[42] Dirección de planificación y desarrollo CDEC-SIC. Determinación de puntos de conexión STT. Technical report, Coordinador Eléctrico Nacional, Santiago, Chile. https://sic.coordinadorelectrico.cl/wp-content/uploads/2015/07/Determinación-de-puntos-de-conexión-STT.pdf, July 2015. "[Online; accessed 04-May-2016]".

[43] Coordinador Eléctrico Nacional. Programación Semanal CDEC SIC. https://sic.coordinadorelectrico.cl/informes-y-documentos/fichas/operacion-programada-2/, 2016. [Online; accessed 03-May-2016].

[44] Coordinador Eléctrico Nacional. Información técnica centrales SING. http://cdec2.cdec-sing.cl/pls/portal/cdec.pck_web_coord_elec.sp_pagina?p_id=5187, 2016. [Online; accessed 20-May-2016].

[45] R. Fuentes-Ryks and C. Wander. Technical information. personal communication, 2016. [Vivest Energías Renovables s.a.].

[46] L. Pascual, J. Romo, and E. Ruiz. Bootstrap predictive inference for arima processes. *Journal of Time Series Analysis*, 25(4):449–465, 2004.

[47] L. Pascual, J. Romo, and E. Ruiz. Bootstrap prediction intervals for power-transformed time series. *International Journal of Forecasting*, 21(2):219–235, 2005.

[48] W.G. Sullivan, E.M. Wicks, and J.T. Luxhoj. *Ingeniería económica de DeGarmo*. Pearson Educación, 2004.

[49] H. Kim, D. Olave-Rojas, E. Álvarez-Miranda, and SW. Son. In-depth data on the network structure and hourly activity of the central chilean power grid. *Scientific data*, 5:180209, 2018.

[50] D. Olave-Rojas, E. Álvarez-Miranda, A. Rodríguez, and C. Tenreiro. An optimization framework for investment evaluation of complex renewable energy systems. *Energies*, 10(7):1062, 2017.

Apéndice A

In-depth data on the network structure and hourly activity of the central chilean power grid

Contiene:

- Correo electrónico de aceptación de la revista
- Artículo original obtenido según DOI: 10.1038/sdata.2018.209

Email

Manuscript #	SDATA-18-00184B
Title	In-depth data on the network structure and hourly activity of the Central Chilean power grid
Corresponding Author	Prof. Seung-Woo Son (Hanyang University)
Date:	10th August 18 11:53:10
Last Sent:	10th August 18 11:53:10
Created By:	Redacted
From:	scientificdata@nature.com
To:	sonswoo@gmail.com
BCC:	Redacted
Subject:	Scientific Data: Decision letter for SDATA-18-00184B
Email	Manuscript Number: SDATA-18-00184B Manuscript Title: In-depth data on the network structure and hourly activity of the Central Chilean power grid Corresponding Author: Prof Son Dear Prof Son, We are now delighted to accept your manuscript entitled "In-depth data on the network structure and hourly activity of the Central Chilean power grid" for publication in Scientific Data. Thank you for choosing to publish your work with us. Thank you for improving the language in this version. The Editor would like to note two very minor remaining language issues: 1. In the abstract, you write: "We harmonize three data sources across diverse sources to generate a unified dataset." I personally find this sentence somewhat confusing as written. Perhaps it would be better written as "We harmonized data from three diverse sources to generate a unified dataset." 2. In the Background & Summary, you write "those networks are unweighed and undirected". I think the more common term, in this context, may be "unweighted". If you agree with the suggested changes above, please send us a revised tex file as an attachment to a reply email within the next week. <B>Licence to Publish</B> - The corresponding author of an accepted manuscript is required to complete an Open Access Licence to Publish on behalf of all authors; a link to the online portal through which you can submit this licence agreement will be sent in a separate email. <B>Article-processing charge</B> - If applicable, you will also receive a link to our payment application where you can provide your billing information and pay your article-processing charge (APC) via credit card or by requesting an invoice. <B>Please note that your paper cannot be sent for typesetting to our production team until we have received the above information.</B> An in-house curator will now produce the ISA-Tab metadata that will accompany your Data Descriptor upon publication. You may be contacted and asked to provide additional information or files, as needed. Please be sure to respond to the curator promptly. The metadata record must be completed before proofs can be generated and your submission

published online.

You will receive a PDF proof for checking once the aforementioned metadata has been completed and your manuscript sent for typesetting. Please check the author list and affiliations to ensure that they are correct. For the main text, only errors that have been introduced during the production process or those that directly compromise the scientific integrity of the paper may be corrected at this stage. Please ensure that only one author communicates with us and that only one set of corrections is returned. The corresponding (or nominated) author is responsible on behalf of all co-authors for the accuracy of all content, including spelling of names and current affiliations.

We would welcome the submission of a 'cover image' to highlight your Data Descriptor on the Scientific Data homepage. Images should relate to the content of your Data Descriptor, but need not be contained within the paper. Images should be at least 1400px x 400px in size and of as high resolution as possible (ideally 300dpi), with the focus of the image (if applicable) to the right of centre. Unfortunately, we cannot promise that your suggestion will be used.

Acceptance of your manuscript is conditional on all authors' agreement with our publication policies (see http://www.nature.com/sdata/policies/editorial-and-publishing-pol icies). In particular, your manuscript must not be published elsewhere and there must be no announcement of this work to any media outlet until the publication date. Advice about media relations is available from Nature Research's press office at press@nature.com; we strongly suggest that you seek our advice should you wish to bring your work to the media's attention. We will inform you by email as soon as your manuscript is scheduled for publication, which will be after we have received your proof corrections.

A form to order reprints of your article is available at htt p://www.nature.com/reprints/author-reprints.html. To obtain the special author reprint rate, orders must be made within a month of the publication date. After that, reprints are charged at the normal (commercial) rate.

We look forward to publishing your paper.

Sincerely,

Andrew Hufton

--
Andrew L. Hufton, PhD
Chief Editor, Scientific Data
Nature Research
www.nature.com/scientificdata

The Editorial Board Member for this manuscript was Wided Medjroubi.

If you no longer wish to receive any email correspondence from this journal, please click the link below to unsubscribe Link Not Available ?
el=A7CZ7Bbis5C5bnT1BX6A9ftdQBMhOFrKnLr08n55IRoD9gZ

www.nature.com/scientificdata

SCIENTIFIC DATA

OPEN

Data Descriptor: In-depth data on the network structure and hourly activity of the Central Chilean power grid

Received: 3 April 2018
Accepted: 10 August 2018
Published: 23 October 2018

Heetae Kim[1,2], David Olave-Rojas[1], Eduardo Álvarez-Miranda[1] & Seung-Woo Son[2,3,4]

Network science enables us to improve the performance of complex systems such as traffic, communication, and power grids. To do so, it is necessary to use a well-constructed flawless network dataset associated with the system of interest. In this study, we present the dataset of the Chilean power grid. We harmonized data from three diverse sources to generate a unified dataset. Through an intensive review on the raw data, we filter out inconsistent errors and unrealistic faults, making the data more trustworthy. In contrast to other network dataset for power grids, we especially focus on preserving the physical structure of nodes' connection incorporating the 'tap' structure. As a result, we provide three different versions of the dataset: 'with-tap', 'without-tap', and 'reduced versions'. Along with structure, we incorporate various attributes of the nodes and edges such as the geo-coordinates, voltage of transmission lines, and the time series data of generation or consumption. These data are useful for network scientists to analyze the performance and dynamic stability of power grids.

Design Type(s)	database creation objective • process-based data transformation objective • network analysis objective
Measurement Type(s)	Connectivity
Technology Type(s)	network graph construction
Factor Type(s)	
Sample Characteristic(s)	Chile • electric power system

[1]Department of Industrial Engineering, Universidad de Talca, Curicó 3341717, Chile. [2]Asia Pacific Center for Theoretical Physics, Pohang 37673, Korea. [3]Department of Applied Physics, Hanyang University, Ansan 15588, Korea. [4]Department of Physics and Astronomy, University of Calgary, Calgary, Alberta T2N 1N4, Canada. Correspondence and requests for materials should be addressed to H.K. (email: hkim@utalca.cl) or S.-W.S. (email: sonswoo@hanyang.ac.kr)

Background & Summary

With the recent development of network science, power grids have been studied to prevent large-scale blackouts, maintaining the stability of high-voltage transmission systems. For instance, studies on synchronization typically analyze the cascading consequences of malfunction in power-grid systems[1–7]. These studies rely on network datasets that embody the information of power grids. This information includes the connection topology of the transmission lines and the characteristics of each power facility: power plants or substations.

Usually, real-world power grids are mapped into network structures by performing aggregations and simplifications in order to produce convenient datasets for network analyses. As a matter of fact, in most cases such as the U.S.[7] and EU countries including Italy[8], France[8], Spain[8] and Great Britain[9,10], the resulting datasets have only one type of transmission line and the amount of power consumption or generation is assumed to be uniform in all power-grid nodes. In terms of network terminology, those networks are unweighted and undirected, which are favorable only for analyzing general characteristics in the average sense.

Nowadays, power systems are comprised by a diverse pool of generation and transmission technologies, which respond to different operation schemes and to much more complex demand and revenue patterns. Therefore, it is important to use rather realistic network dataset in order to analyze more sophisticated (or realistic) operation issues on the power grids. Such a realistic network datasets should keep the detailed characteristics of power components such as capacity of each transmission lines (weight of edges), the amount of electricity consumption or generation of demand and power nodes, respectively, and geological coordinates (attributes of nodes). The attributes of network components will enable us to investigate the stability on the new-type of power grids, such as smart grid system, distributed power generation, and renewable power generation technologies.

There exist various methods for mapping a given power grid into a network representation[11]; this allows to generate power-grid networks with different degrees of accuracy. This means that, according to the scope, the power-grid network can be more detailed and realistic. It is obvious that the more accurate and realistic the performed analyses are, the more detailed network modeling is required[7].

Indeed, some projects and protocols aim at constructing well-digitalized power-grid information. For instance, the European Network of Transmission System Operators for Electricity (ENTSO-E) is a global project among thirty-five European countries devoted to collecting network data of Pan-European power grids[12]; likewise, the Spatial Optimization Model of the Electricity Sector (ELMOD) is a model to digitize power grid data of Europe developed in 2012[13]. ELMOD covers the entire European electricity market. Continuing the effort of ELMOD, the Institute for Economic Research and the Technical University of Berlin recently developed ELMOD-DE, which particularizes the German case[14]. ELMOD-DE data package includes detailed power-grid information of German power grid is opened to the public. Complementary examples correspond to Sci-grid[15] and GridKit[16], which are open-source tool kits for automatically extracting power-grid topology information from OpenStreetMap[17]. Despite of these efforts, most of the well-structured data are concentrated in Europe and North America.

With the aim of addressing this issue, in this study we provide a curated network dataset of a power grid from South America. In particular, we map the Chilean power grid into three different network versions. Each of them is prepared by adopting or omitting one of two conversion schemes—*tapping* and *reduction*. These versions range from the most simplified topology to the most realistic one, enabling the users to selectively use a proper version considering the purpose of their analyses.

In the first version, the one with the most realistic structure, power facilities are connected to the nearest transmission line by means of so-called tap connections. A tap connection is a short extension line from a facility to the main grid. Embedding this tap structure produces a complex connection structure, encompassed by four types of nodes with power plants, substations, taps, and junctions (Fig. 1b). This leads to a total of 347 nodes: 124 plants, 94 substations, 85 junctions (branch points), and 44 tap nodes. The second version, with intermediate degree of complexity, ignores these tap connections (Fig. 1c), and it encompasses 318 nodes: 124 power plants, 94 substations, and 100 junction nodes. Finally, the third, and most simplified version, aggregates nodes to a power plant and substation level, i.e., only *activity* nodes, which leads to a network comprised by 218 nodes (Fig. 1d). Besides the set of nodes, and the corresponding interconnection layout, we also provide the geo-coordinates and activity information of power-grid nodes that can be used for scenario-based analysis. This first, second, and third version of the networks will be referred to as *with-tap* network (WT), *without-tap* network (WOT), and *without-tap reduced* network (WOR).

Methods

In Chile, the *Centro de Despacho Económico de Carga del Sistema Interconectado Central* (CDEC-SIC) is the agency devoted to the coordination and operation of the Central Chilean power grid. CDEC-SIC provides, in the website of Coordinador Eléctrico Nacional[18], an extensive collection of raw data regarding the structure and operation of the generation and transmission network.

In particular, for this investigation, we merged three different sets of raw data into a unified dataset. Each source contains the geo-coordinates of power grid facilities (Mapa Sistemas Eléctricos de Chile[19]), the connection structure as a circuit diagram (Diagrama Unilineal SIC[20]), and the activity records

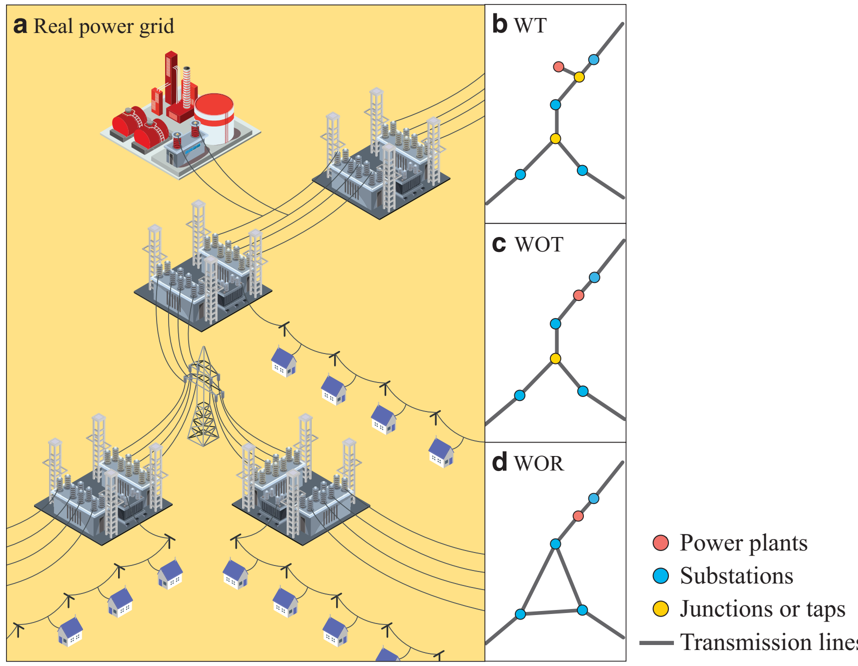


Figure 1. Three conversion schemes from a power grid to a network topology. A real power grid (**a**) can be represented in different versions. 'With tap (WT)' version in (**b**) has all power-grid components such as power plants, substations, taps, and junctions while 'Without tap (WOT)' version in (**c**) does not include taps. 'Without tap reduced (WOR)' version in (**d**) has the simplest structure having only power plants and substations.

regarding production and consumption (Operación real[21]), respectively. In this section, we describe the detailed method used for the data cleaning process.

Geographic coordinates

The data source containing geo-coordinates includes the longitude and latitude information of power plants and substations. The geo-reference system in the raw data is based on three time zones: EPSG 32717, 32718, and 32719, which correspond to UTM zone 17 S, 18 S, to 19 S. We convert the coordinate of 286 power plants, 860 substations, and 108 taps to EPSG 4326 system to have a standard geo-coordinate system.

In the process of integrating, cleaning, correcting, and processing the source data, we faced several difficulties due to imprecisions, lack of consistency, and missing information. In the following, we describe the main difficulties faced in the preparation process.

In the raw data, a relatively common situation corresponds to a mismatch among nodes' ID and geo-coordinates when comparing different files. Hence, we have to individually inspect each problematic entry, in order to correct the geographical information, or ensure a correct association with respect to the node ID. Another troublesome situation occurs when mapping the connection layout associated with large substation complexes. For such situations, several transmission lines are connected to and from various directions, and each connection is at multiple locations within and around the complex. In this case, we merge substations into an *artificial* substation, whose location is given by the average value (center of mass) of the coordinates of the different substations comprising the complex.

Additionally, due to the fact that generators, substations, and other facilities belong to different owners, it is frequent that some of them have the same name. For instance, following the name of the area where the facility is located; e.g., two substations, owned by different operators, might be named *Curicó* because they are both located in a village named *Curicó*, and there could be more than one village with this name. Since part of the layout source data only associates names to facilities, having elements with replicated name produces a significant number of inconsistencies; most of them are manually corrected in a case-by-case analysis.

Furthermore, for some nodes of the power grid, the information regarding geo-coordinates is missing in the raw data. In these cases, we first manually checked each conflicting entry and attempted to assign coordinates by looking at the node ID, facility name, facility location, exiting connections, or any other relevant information at hand. Through this process, we corrected the information of 81 nodes in the case of WT (18.2% of the total nodes), and 59 nodes in the case of WOT (14.4%) and WOR (11.2%). When further manual assignment of coordinates was not possible, we indirectly set the locations by the following scheme. For the case of nodes with degree 1, we looked at the only connection of each of them, if the length of such connection is less or equal than a threshold value, then we set the missing node's coordinates similar to its neighbor node. Complementary, for the case of nodes connected to several other nodes, we set the average value of the neighbors' positions as the missing node's geo-coordinates. These two steps are recursively applied to the data, until all nodes acquire geo-coordinates; the information of 59 nodes is corrected in the case of WT, and of 50 nodes in the case of WOT and WOR.

Activity

For a given node, its *activity* corresponds to the amount of (net) power generation. If the node is a substation, the activity is non-positive; while if it is a generation node, the activity is non-negative. In particular, we collect the activity data for one year starting from 16th September 2015 until 15th September 2016, from the official web page of CDEC-SIC[18]. In order to minimize the human error during the data handling, we developed a Python code to automatically parse the data from the raw file.

Although all raw data are provided from the same organization, the information does not exactly coincide among the raw files. For example, 'Sauzal' is a unique node in the topology data. However, in the activity data, there exist 'Sauazal 1' and 'Sauzal 2'. Because the activity data is made for the purpose of pricing and billing, each subunit is recorded individually. In this manner, we manually inspect the list of all raw data sources, and create a consolidated node list. In addition, for those nodes that do not have any activity data, we assume that the power facilities are not functioning due to maintenance. In that case, we set zero value for activity.

In reality the amount of power generation is always larger than the amount of consumption. The gap between the generation and the consumption is the so-called reserve capacity that compensates a sudden peak power demand preventing a blackout. In power studies, however, with the assumption of satisfying Kirchhoff's law, it is common to set the net activity to zero. Therefore, by ensuring the grid to verify the Kirchhoff law, we are enabled to exploit electric principles associated with the grid. In this study, for the sake of applicability, we prepare two different versions of activity data: the original version and the Kirchhoff-adjusted version.

In the Kirchhoff-adjusted version, the net activity of each node is zero, i.e.,

$$\sum_{i=1}^{N} a_i' = 0, \tag{1}$$

where a_i' is the adjusted input or output of node i and N is the total number of nodes. In order to adjust the original values for making the total sum zero, we cancel out the reserve capacity from the generation activity. We do this by multiplying the activity values of generation with the ratio of the total consumption with respect to the total generation:

$$a_i^{'} = r a_i, \text{where} \begin{cases} r = 1 & \text{if } i \in S, \\ r = \frac{\sum_{j \in S} a_j}{\sum_{j \in P} a_j} & \text{if } i \in P. \end{cases} \tag{2}$$

Here S is the set of substation nodes, P is the set of power plant nodes, a_i is the original activity of node i, and a_i'Here S is the set of substation nodes, P is the set of power plant nodes, a_i is the original activity of node i, and a_i' is the adjusted activity of node i.

Since the product operation only affects to the nodes having non-zero activity value, the zero-activity nodes still remain inactive after the adjustment (See Equation (2)). This enables the adjustment process to still keep the unique characteristics of electricity generation profile of Chile.

Connection structure

A power grid is a giant electrical circuit. In such circuits, power plants send electric power to substations through transmission lines. The single line diagram (SLD)[20] of the Chilean power grid, which is one of the raw data sources, shows how the power plants and substations are connected.

The SLD is the ground truth of the Chilean power grid; based on it, we are able to construct the edge list between power-grid nodes. The SLD illustrates each power-grid component even for the individual generators in a power plant. The detailed information is useful to understand the connection layout structure of the power grid. For example, we can distinguish the capacity of transmission lines as 550 kilovolts (kV), 220 kV, 154 kV, 110 kV, 66 kV, and less than 66 kV, which can be used for the edge weight.

When mapping the SLD into a network structure, we apply, in particular, two conversion principles in this study: tap-embedding and node reduction. The practical conversion process is described in each subsection below.

Tap-embedding and further network characterization

Tapping is a connecting method of a node to the middle of a transmission line. An extended line from a node hangs on a spanning transmission line, which enables the node to connect the main grid. The tapping has an advantage in system's stability point of view. When an accident or a failure happens at the

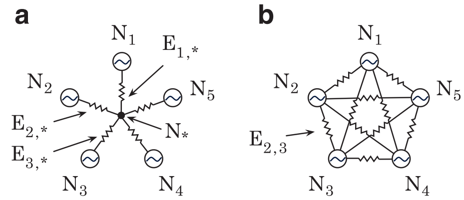


Figure 2. Star-mesh conversion results in the electrically equivalent circuit removing the center node. Conversion from (**a**) a star-shape circuit to (**b**) a mesh shape. *i* is the node index and * represents the central node that is being removed by the transformation process.

tapped node, it does not affect the current flow of the main grid. However, when the node is placed between two nodes without tapping, all three nodes are connected in order. In this case, a single failure on the middle node could directly break the current flowing through the node, which makes the grid vulnerable. The tapping connection is to preserve the connectivity between other nodes regardless of the functionality of the tapped node.

One can find tap structures often in real transmission grids. However, for the sake of simplicity, their structure is usually simplified when it is converted into a network topology. The topological structure is one of the main ingredients for the network analysis on power grids. Whether we ignore the tap or not, directly affects the result. Therefore, one should carefully decide to include or exclude the tap structure based on the purpose of the analysis. In order to characterize the tap connection, we keep all tapped structures from the SLD, which could be identified either by their name (explicitly identifying a tap role) or they connection arrangement.

Besides tap connections, *junction points* also matter from a structural point of view. At a junction point, transmission lines are merged (or diverged) not incorporating any power facility such as generators or substations. We consider the junctions as nodes when we find from the SLD of the Chilean power grid. This enables us to keep the characteristic of the power grid.

From the SLD, we classify the role of nodes from the corresponding symbol. The nodes with transformer symbols are substations; generator symbols, power plants. As a result, we collect 818 nodes from the SLD: 463 substations, 168 plants, 51 junctions, and 136 tap nodes.

Among the original 818 power-grid nodes, we only keep those appearing in all data sources. To do so, when we remove the nodes that exist only in some data sources, we keep the connectivity between the adjacent nodes making alternate edges to prevent network segregation. However, in order not to make any duplicated edge, we remove the multi-edges between the same nodes. The final network data include only the nodes commonly exist in all data sources: 347 nodes in tap-embedded version and 318 nodes in without-tap version.

Node reduction

So far, we have explained two versions of the Chilean power grid: WT and WOT. In order to construct the most simplified version (WOR)—but most structurally similar to other power-grid dataset—we further reduce nodes, using the so-called star-mesh transformation principle[22].

Star-mesh transformation is a useful technique to convert a circuit to other electrically equivalent configuration with less number of nodes (See Fig. 2). However, the voltages applying to outer nodes are still identical for both configurations. Therefore, it is useful to simplify power grids keeping the electrical relationship between nodes. When the reduction process is conducted to only four nodes, the conversion process is called Kron reduction[23–25] or Y-Δ transformation.

This transformation implies that edge characteristics shall be mapped from the information of the original setting. For the case of power grids, the resulting edges impedance is based on the impedance of edges connecting the central node to the peripheral nodes, according to

$$E_{i,j} = E_{i,*}E_{j,*}\sum_{i\,\in\, n.n(*)}\frac{1}{E_{i,*}} \tag{3}$$

where * is the index of the central node, i, j are the node indices connected to the central node, and $E_{i,j}$ corresponds to the impedance of the transmission line between node i and j. Applying the Star-mash transformation to WOT version, we construct WOR version of dataset with total 218 nodes.

Code availability

We used command line tools to avoid any potential human error during manual work. All library and modules used in this study are freely available open access tools. During the data process, we mainly used *Python 2.7* along with *Numpy 1.9.1* and *Pandas 0.19.2* for parsing and cleaning data. To visualize the network we used *Basemap 1.0.7* and *Matplotlib 1.4.2*.

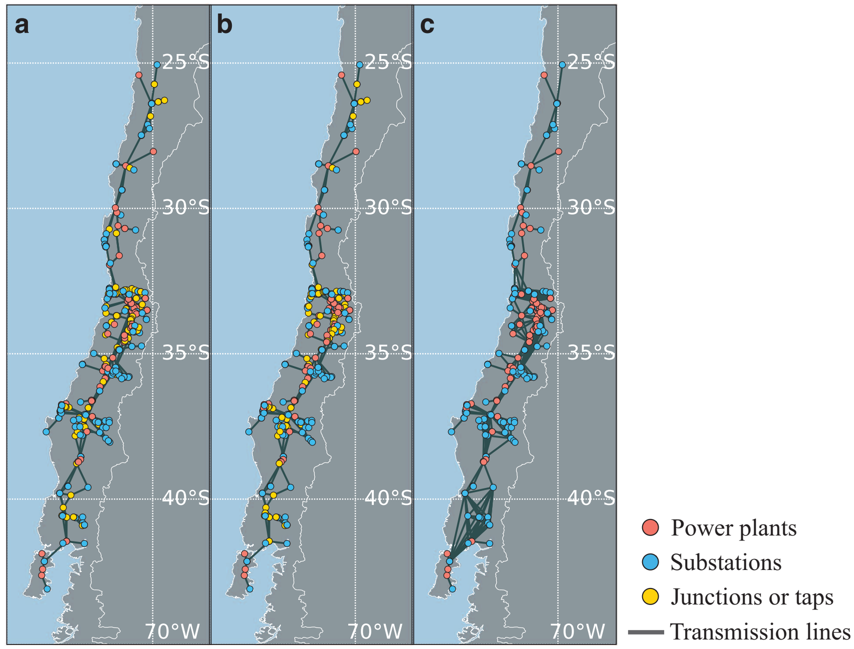


Figure 3. Network representation of Chilean power grid on the map of Chile. The network components in Chilean power grid are converted according to the concept of (**a**) WT, (**b**) WOT, or (**c**) WOR, respectively. Nodes are located on the actual geo-coordinates such that some nearby nodes are illustrated overlapping.

	Node					Edge
	Power plant	Substation	Tap	Junction	Total	
WT	124	94	44	85	347	444
WOT	124	94	0	100	318	409
WOR	124	94	0	0	218	527

Table 1. The number of nodes and edges in each version.

Data Records

We prepared three versions of power-grid dataset by selectively combining different conversion techniques that are described before. These versions are shown in Fig. 3a–c, corresponding to WT, WOT, and WOR, respectively, i.e., ranging from a connection structure including all tap and junction points, to a structure including only activity nodes.

All versions of the network preserve the functional relationship between nodes in terms of dynamics. The brief information regarding the three versions is in Table 1.

The network dataset are prepared in a data repository (Data Citation 1). Files are organized by versions having the name of each version in the front of file name. The information with respect to nodes, edges, and activity are prepared as *.csv* files and the name of the file tells what it includes. In the node list, each row corresponds to a node and columns show the detail attributes of the node. To clarify the data, the node attributes include following information:

- ID: the identification index of the node that is also used in the edge list
- Label: the name of the node
- Owner: the company that owns the power facility
- Role: the kind of the node (power plant, substation, tap, or junction)
- Longitude: the longitude of the node
- Latitude: the latitude of the node

The edge list has the detailed transmission connection status of the power grid. 'Source' and 'Target' columns in the file mean end points of an edge. Note that all edges are indirect connections. Edge information in the file includes:

- Source: the ID of the end point of an edge
- Target: the ID of another end point of the edge
- Distance: the distance of the corresponding transmission line (km)
- Voltage: the capacity of the transmission line (kV)

Activity data is encoded into a large spreadsheet with time series information about the amount of consumption or generation (MWh) of activity nodes. For users' convenience, we also leave the node attributes in the same file. The file includes 8788 rows in total: 4 rows for role, ID, label, and owner of a node plus 8784 rows for time stamp from 16th September 2015 to 15th September 2016 (366 days) on an hourly basis. Note that within the period, there is no importation nor exportation of electricity associated with SIC. The central Chilean power grid is self-sustaining independently to neighborhood as a closed system.

Power grid with tap
As explained above, the tap-embedded network data reflects the physical connection structure of power grids. We leave out tap connections and junctions as they exist as a part of the real power-grid structure. This version is particularly useful to simulate the network response against physical attacks. In this version of the data, the network is comprised by 444 edges, and by a total 347 nodes consisting of 124 plants, 94 substations, 85 junctions (branch points), and 44 tap nodes.

Power grid without tap
In this version power plants and substations are no longer connected via taps. The tap nodes are removed except for the case that they are necessary from the structural point of view. For the exceptional cases, we set the attribute of the tap nodes as junctions since they are making branches. The network topology in this version still reveals the unique physical structure of the power grid; it is the intermediate version between the most realistic tap-embedded version and the most simplified Star-mesh reduced one. In this case, there are 409 edges and there exist 318 nodes in total (124 power plants, 94 substations, and 100 junction nodes without any tap node).

Reduced power grid
The most simplified network form of power grid is the reduced network. It eliminates the redundant nodes that do not affect the dynamical interaction between power producers and consumers. For instance, network studies about synchronization problems usually consider only power producers and consumers as oscillators. In this case, junctions and taps are ignored because they do not affect the dynamic interaction between the oscillators. Most network studies utilize power-grid networks that contain only plant and substations, the reduced version is useful especially to compare with other studies. This reduced version has solely 527 edges with 124 power plant nodes and 94 substations.

Activity data
The total power production for an year is about 47,005 GWh in the original data with 3,176 GWh of reserve capacity that is in good agreement with the normal 10–20 percent of reserve ratio. In the Kirchhoff-adjusted dataset, the total power generation is scaled to be equal to the total amount of power consumption, which results in 43,829 GWh of power production. The maximum power producer during the period is *Guacolda* with 4,499 GWh (4,195 GWh after scaling) of power generation and the maximum consumer is *Polpaico* with 3,054 GWh.

Activity data, which is associated to generators and sub-stations, is useful to analyze power grids considering the actual demand and supply pattern of consumers and power plants, respectively. Figure 4 shows heat maps of activity data of four power plants in the Chilean power grid. The activity data clearly shows the distinct pattern of power generation according to the season, time, and technology. It enables us to do a scenario-based case study with historical records.

For example, Campiche (Fig. 4a) is one of the base power facilities in Chile and constantly supplies electric power. On the other hand Teno (Fig. 4b) generates electricity during only a specific period. The activity pattern also reveals the technological and seasonal characteristics of power activity of nodes. Loma Las Colorados (Fig. 4c) is a solar power plant such that it generates only during day time. The electric power generation of hydro power plant increases during the rainy season as Pangue shows in Fig. 4d.

Technical Validation
As the constructed network corresponds to a unified interconnected power system, all nodes shall comprise a single large connected component. Therefore, in the three data versions we ensure that connectivity was verified by the resulting network, through a depth-first search algorithm.

While the connectivity can be checked by a standard procedure, the remaining characteristics can be only verified by an exhaustive brute-force method. For instance, the correct removal of a tap or junction node, can be only verified by analyzing candidates one by one. A tap node (yellow) in Fig. 5a, for example, is correctly eliminated after applying our conversion procedure resulting in a reduced network Fig. 5b. A redundant central node (yellow) in Fig. 5c, along with its connected edges, is also removed in

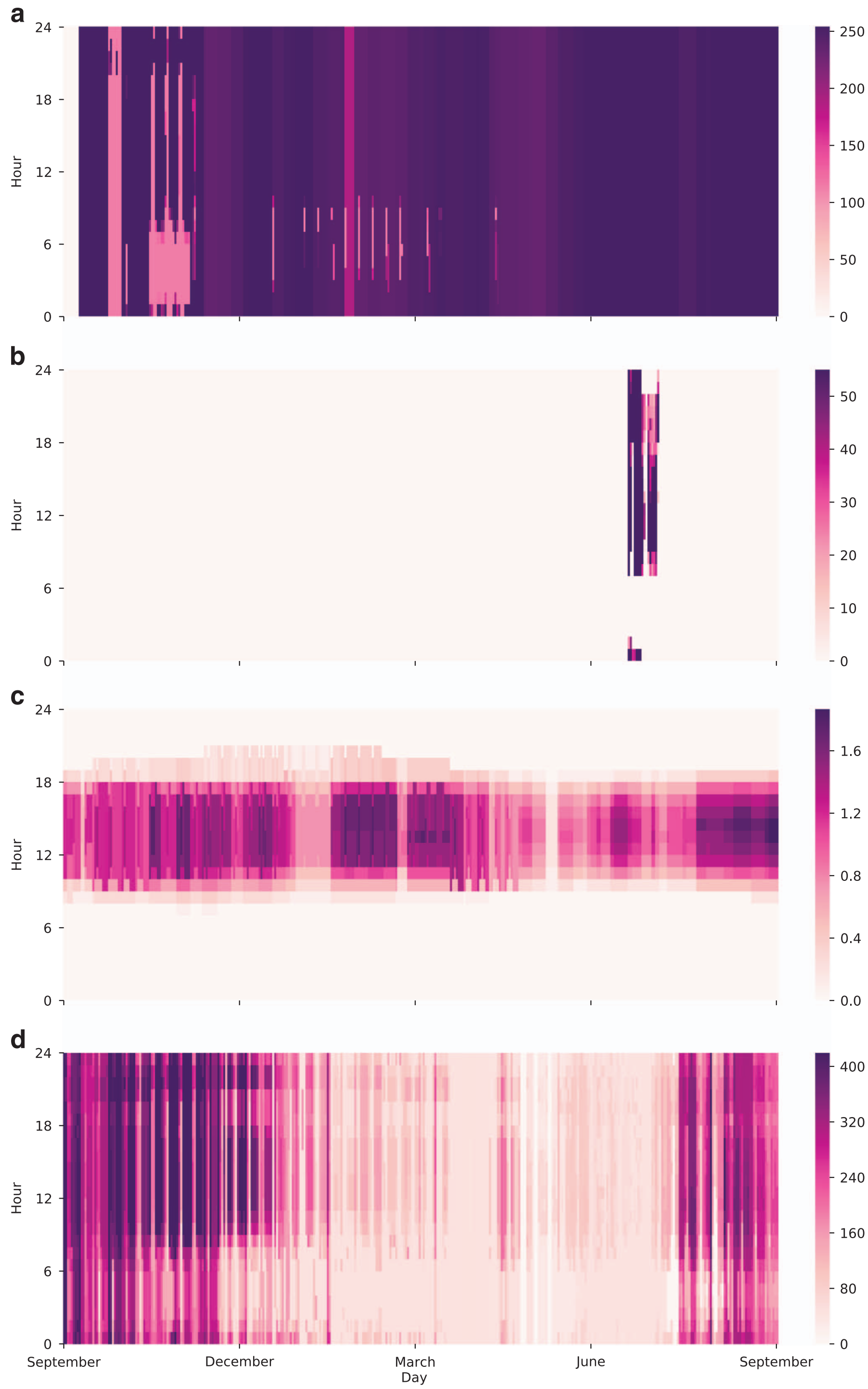


Figure 4. Hourly activity record of few power plants in Chile from 16th September in 2016 to 15th September in 2017. The power plants show various generation patterns according to the generation technology of (**a**) Campiche, (**b**) Teno, (**c**) Loma Las Colorados, and (**d**) Pangue (in the unit of MWh).

the reduced version in Fig. 5d. Note that the resulting networks shown in Fig. 5a and c correspond to WT, Fig. 5b to WOT, and Fig. 5d to WOR, respectively. We also check that the sum of activity of all nodes in Kirchhoff-adjusted versions is less than -3.72×10^{-7}.

Usage Notes

One of the easiest ways to use the activity data provided by this study is to import with the *Pandas* library in *Python*. For the sake of users' convenience, we include the attribute of nodes–ID, name, and owner–as the multiple indexes in *.csv* file along with the activity data. Since the activity data is prepared by using

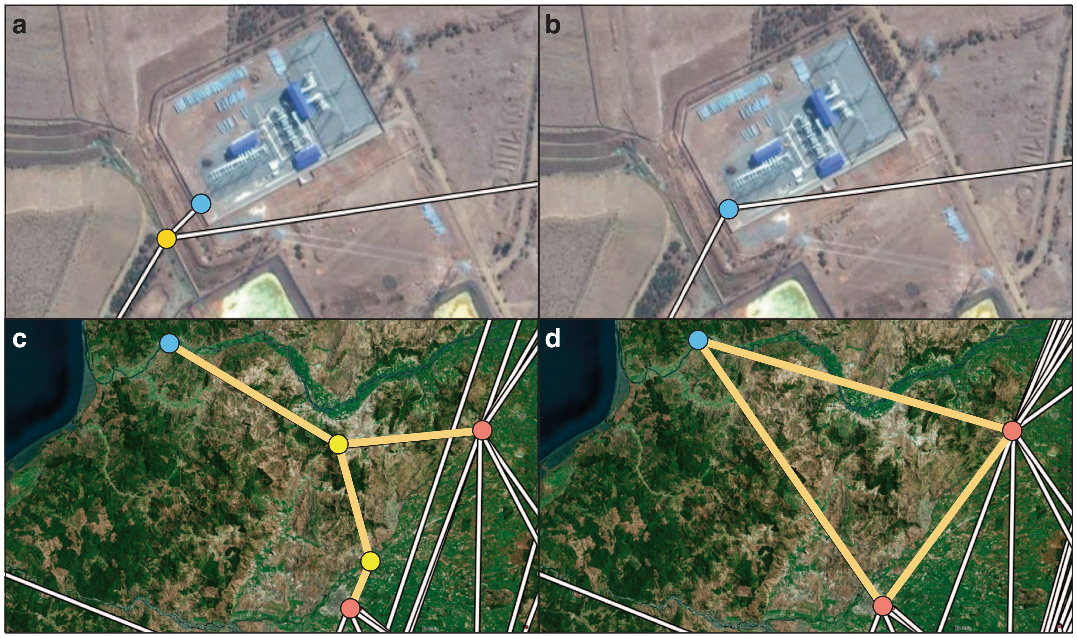


Figure 5. Visual comparison between WT, WOT, and WOR versions. The different structure between WT (**a**) and WOT (**b**) shows the role of tap connection. The tap node (yellow) in WT (**a**) is the tapping point of the power plant, *Lo Aquirre* (blue) to the transmission line. The redundant junction nodes (yellow) in WT (**c**) are eliminated in WOR (**d**). The edge connection also changes (yellow lines) in order to generate electrically equivalent circuit.

standard packages and recorded with UTF-8 text encoding, it is available in any programming language without sensitive development environment issues.

Here we describe usage examples of our dataset. For those who want to use this data for the purpose of comparison to other power grids, it is recommended to use the reduced version. It is because of the fact that most current network dataset of power grids in many studies contain only power plant and substation nodes[2,6,26–30]. The reduced version is the general type of power-grid data at the moment.

On the other hand, when one needs a detailed structure of power grids for analyses related with the physical connection topology or optimal power flow[30–33], the 'with-tap' or 'without-tap' version is a good choice. The 'with-tap' version is a dedicated version for those who are interested in zooming in the networks. The 'without-tap' version ignores tapping but still preserves the overall topological shape such that it is appropriate for the analysis in which the difference due to tapping is negligible.

References

1. Motter, A. E. & Lai, Y.-C. Cascade-based attacks on complex networks. *Phys. Rev. E* **66,** 065102 (2002).
2. Crucitti, P., Latora, V. & Marchiori, M. Model for cascading failures in complex networks. *Phys. Rev. E* **69,** 045104 (2004).
3. Kinney, R., Crucitti, P., Albert, R. & Latora, V. Modeling cascading failures in the North American power grid. *Eur. Phys. J. B* **46,** 101–107 (2005).
4. Dobson, I., Carreras, B. A., Lynch, V. E. & Newman, D. E. Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos* **17,** 026103 (2007).
5. Pahwa, S., Scoglio, C. & Scala, A. Abruptness of cascade failures in power grids. *Sci. Rep* **4,** 3694 (2014).
6. Rohden, M., Jung, D., Tamrakar, S. & Kettemann, S. Cascading failures in ac electricity grids. *Phys. Rev. E* **94,** 032209 (2016).
7. Yang, Y., Nishikawa, T. & Motter, A. E. Small vulnerable sets determine large network cascades in power grids. *Science* **358,** 6265 (2017).
8. Rosato, V., Bologna, S. & Tiriticco, F. Topological properties of high-voltage electrical transmission networks. *Electr. Power Syst. Res* **77,** 99–105 (2007).
9. Rohden, M., Sorge, A., Witthaut, D. & Timme, M. Impact of network topology on synchrony of oscillatory power grids. *Chaos* **24,** 013123 (2014).
10. Rohden, M., Sorge, A., Timme, M. & Witthaut, D. Self-Organized Synchronization in Decentralized Power Grids. *Phys. Rev. Lett.* **109,** 064101 (2012).
11. Nishikawa, T. & Motter, A. E. Comparative analysis of existing models for power-grid synchronization. *New J. Phys.* **17,** 015012 (2015).
12. ENTSO-E, European Network of transmission system operators for electricity https://www.entsoe.eu/.
13. Leuthold, F. U., Weigt, H. & Hirschhausen, C. von A large-scale spatial optimization model of the European electricity market. *Netw. Spat. Econ.* **12,** 75–107 (2012).
14. ELMOD-DE, DIE. An open source model for Germany http://www.diw.de/de/diw_01.c.528493.de/forschung_beratung/nachhaltigkeit/umwelt/verkehr/energie/modelle/elmod.html (2016).
15. Medjroubi, W., Matke, C. & Kleinhans., D. SciGRID - An Open Source Reference Model for the European Transmission Network http://www.scigrid.de (2015).
16. Wiegmans, B. Gridkit 1.0 'for scientists' https://zenodo.org/record/47263 (2016).

17. The OpenStreetMap Foundation. Openstreetmap https://www.openstreetmap.org (2017).
18. Coordinador Eléctrico Nacional. *Coordinador Eléctrico Nacional* https://www.coordinador.cl (2018).
19. Coordinador Eléctrico Nacional. *Mapa Sistema Eléctricos de Chile* https://infotecnica.coordinador.cl (2018).
20. Coordinador Eléctrico Nacional. *Diagrama Unilineal SIC* https://infotecnica.coordinador.cl (2018).
21. Coordinador Eléctrico Nacional. *Operación Real* https://sic.coordinador.cl/informes-y-documentos/operacion-real/ (2018).
22. Bedrosian, S. Converse of the Star-Mesh Transformation. *IRE Trans. Circuit Theory* **8,** 491–493 (1961).
23. Kron, G. *Tensor Analysis of Networks*. (Wiley, 1939).
24. Kennelly, A. E. The equivalence of triangles and three-pointed stars in conducting networks. *Electrical world and engineer* **34,** 413–414 (1899).
25. Dörfler, F. & Bullo, F. Kron Reduction of graphs with applications to electrical networks. *IEEE Trans. Circuits Syst. I, Reg. Papers* **60,** 150–163 (2013).
26. Lozano, S., Buzna, L. & Díaz-Guilera, A. Role of network topology in the synchronization of power systems. *Eur. Phys. J. B* **85,** 231 (2012).
27. Schultz, P., Heitzig, J. & Kurths, J. Detours around basin stability in power networks. *New J. Phys.* **16,** 125001 (2014).
28. Menck, P. J., Heitzig, J., Kurths, J. & Joachim Schellnhuber, H. How dead ends undermine power grid stability. *Nat. Commun.* **5,** 3969 (2014).
29. Kim, H., Lee, S. H. & Holme, P. Community consistency determines the stability transition window of power-grid nodes. *New J. Phys.* **17,** 113005 (2015).
30. Nitzbon, J., Schultz, P., Heitzig, J., Kurths, J. & Hellmann, F. Deciphering the imprint of topology on nonlinear dynamical network stability. *New J. Phys.* **19,** 033029 (2017).
31. Tchuisseu, E. B. T., Gomila, D., Brunner, D. & Colet, P. Effects of dynamic-demand-control appliances on the power grid frequency. *Phys. Rev. E* **96,** 022302 (2017).
32. Rohden, M., Witthaut, D., Timme, M. & Meyer-Ortmanns, H. Curing critical links in oscillator networks as power flow models. *New J. Phys.* **19,** 013002 (2017).
33. Yang, Y., Nishikawa, T. & Motter, A. Small vulnerable sets determine large network cascades in power grids. *Science* **358,** 6365 (2017).

Data Citation

1. Kim, H., Olave-Rojas, D., Álvarez-Miranda, E. & Son, S.-W. *figshare* https://doi.org/10.6084/m9.figshare.c.4053374 (2018).

Acknowledgements

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2017R1D1A1B03032864). S.-W.S. acknowledges the support and hospitality of the University of Calgary, Canada. We thank Jae Ryoung Park for the contribution on the infographic figures. D. O.-R. acknowledges the support of the graduate school of the U. de Talca. E.A.-M. and D. O.-R. acknowledge the support of the Chilean Council of Scientific and Technological Research, CONICYT, through the grant FONDECYT N.1180670 and through the Complex Engineering Systems Institute (ICM-FIC:P-05-004-F, CONICYT:FB0816).

Author Contributions

H.K., E.A.-M., and S.-W.S. designed, conceptualized, and wrote the manuscript. H.K. and E.A.-M. searched out and digitized the network data. D.O.-R. assisted with the early generation of the datasets. All the authors contributed to reviewing and editing the manuscript.

Additional Information

Competing interests: The authors declare no competing interests.

How to cite this article: Kim, H. *et al.* In-depth data on the network structure and hourly activity of the Central Chilean power grid. *Sci. Data*. 5:180209 doi: 10.1038/sdata.2018.209 (2018).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Apéndice B

An optimization framework for investment evaluation of complex renewable energy systems

Contiene:

- Correo electrónico de aceptación de la revista
- Artículo original obtenido según DOI: 10.3390/en10071062

Gmail

David Olave <olavecl@gmail.com>

Fwd: [Energies] Manuscript ID: energies-203167 - handling fees invoice sent

1 mensaje

Eduardo Álvarez-Miranda <eduardo.alvmir@gmail.com> 12 de septiembre de 2018, 16:17
Para: David Olave Rojas <d.olave.rojas@gmail.com>, David Olave <olavecl@gmail.com>

---------- Forwarded message ---------
From: **Eduardo Alvarez Miranda** <ealvarez@utalca.cl>
Date: mié., 12 de sep. de 2018 11:17
Subject: RV: [Energies] Manuscript ID: energies-203167 - handling fees invoice sent
To: Eduardo.alvmir@gmail.com <Eduardo.alvmir@gmail.com>

De: billing@mdpi.com [billing@mdpi.com]
Enviado: martes, 11 de julio de 2017 5:24
Para: Eduardo Alvarez Miranda
Cc: Anna Xiao; Billing Dpt; Xun Yin; Energies Editorial Office
Asunto: [Energies] Manuscript ID: energies-203167 - handling fees invoice sent

Dear Dr. Eduardo Álvarez-Miranda,

Your paper has been accepted for publication in the journal Energies.

The invoice is related to the following manuscript:
Manuscript ID: energies-203167
Type of manuscript: Article
Title: An optimization framework
for investment evaluation of complex renewable energy systems
Authors: Eduardo Álvarez-Miranda *, David Olave, Alejandro Rodríguez,
Claudio Tenreiro
Received: 31 May 2017
E-mails: ealvarez@utalca.cl, d.olave.rojas@gmail.com, alrodriguez@utalca.cl,
ctenreiro@utalca.cl
Energy Production Systems
http://www.mdpi.com/journal/energies/special_issues/energy_prodction_systems

We will publish your accepted paper as soon as we have received the article processing charge (APC) for publication in open access and the editing process is completed. Please follow the link to complete your information for the APC invoice of 1500,00 CHF:
https://susy.mdpi.com/user/manuscript/71aa63226d946a765f665ae5e311ae17/invoices

For immediate payment by credit card follow https://payment.mdpi.com/146375.

For detailed payment instructions or more payment methods, visit
http://www.mdpi.com/about/payment.

Kindly remember to include the invoice ID (energies-203167) in all payment related documentation.

Thank you.

Kind regards,
Aleksandra Cuculovic
Financial Assistant

MDPI AG - Multidisciplinary Digital Publishing Institute
St. Alban-Anlage 66
4052 Basel, Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
E-mail Accounting: billing@mdpi.com
E-mail: aleksandra.cuculovic@mdpi.com
http://www.mdpi.com/

Invoice_MDPI-AG_energies-203167_1500,00CHF.pdf
36K

energies

MDPI

Article

An Optimization Framework for Investment Evaluation of Complex Renewable Energy Systems

David Olave-Rojas [1], Eduardo Álvarez-Miranda [2,*], Alejandro Rodríguez [2] and Claudio Tenreiro [3]

1. DSc Program on Complex Engineering Systems, Universidad de Talca, Curicó 3341717, Chile; d.olave.rojas@gmail.com
2. Department of Industrial Engineering, Universidad de Talca, Curicó 3341717, Chile; alrodriguez@utalca.cl
3. Department of Industrial Technologies, Universidad de Talca, Curicó 3341717, Chile; ctenreiro@utalca.cl

* Correspondence: ealvarez@utalca.cl; Tel.: +56-9-5018-6692

Received: 31 May 2017; Accepted: 11 July 2017; Published: 22 July 2017

Abstract: Enhancing the role of renewable energies in existing power systems is one of the most crucial challenges that society faces today. However, the high variability of their generation potential and the temporal disparity between the demand and the generation potential represent technological and operational gaps that burden the massive incorporation of renewable sources into power systems. Energy storage technologies are an alternative to tackle this gap; nonetheless, their incorporation within large-scale power grids calls for decision-making tools that ensure an appropriate design and sizing of power systems that exploit the benefits of incorporating storage facilities along with renewable generation power. In this paper, we present an optimization framework for aiding the evaluation of the strategic design of complex renewable power systems. The developed tool relies on an optimization problem, the generation, transmission, storage energy location and sizing problem, which allows one to compute economically-attractive investment plans given by the location and sizing of generation and storage energy systems, along with the corresponding layout of transmission lines. Results on a real case study (located in the central region of Chile), characterized by carefully-curated data, show the potential of the developed tool for aiding long-term investment planning.

Keywords: renewable energies; energy storage systems; power system planning; mathematical optimization

1. Introduction and Motivation

One of the main challenges of today's society corresponds to evolving the energy generation matrix due to economical and environmental arguments. According to [1], 81.4% of the world's primary energy comes from fossil fuels, which exposes us to high economic and environmental risks; on the one hand, current forecasts indicate that fossil fuel reserves might reach critical levels in 2050–2100 [2–4], and on the other, fossil fuels are responsible for most of greenhouse gas emissions. The penetration of generation technologies based on renewable sources, such as solar and wind power, seems as the only sustainable alternative for a world with a constant increase of energy demand; see, e.g., [1]. One of the main technological gaps associated with these energy sources is the high variability of their generation potential and the temporal disparity between the demand and the generated energy. These two issues can be addressed through energy storage systems (ESSs); however, these are quite costly solutions and respond to quite strict operative patterns. Therefore, their incorporation within large-scale power grids is not only associated with technological challenges, but also calls for decision-making tools that ensure an appropriate design and sizing of power systems that exploit the benefits of incorporating ESSs along with renewable generation power.

Different efforts have been devoted to the development of decision-making tools for strategic power system design. These approaches address different long-term decisions such as the location and capacity of generators and substations, the layout of transmission lines and the incorporation of ESSs. As we will show in the literature review in Section 2, these tools do not incorporate all decision dimensions simultaneously, and they are typically devised from the perspective of the system operator. On the one hand, this implies that they do not take into account the fact that power grid components work collectively and simultaneously; and on the other, they dismiss the fact that in the current energy market conditions, design tools must respond to the financial expectations of investors. Hence, we believe that there is a need for designing an investor-oriented decision-making framework that incorporates generation, transmission and storage decisions; such a tool aims at exploiting the competitiveness, the attractiveness and the profitability of new renewable energy-based systems.

Contribution and Paper Outline

The main contribution of our paper is an optimization modeling framework, based on a mixed integer linear programming (MILP) model, for supporting decision-making in (renewable-based) complex energy systems' planning. By means of a curated case study, we show the potential of the proposed approach in providing economically-attractive investment plans that exploit the use of ESSs. Furthermore, we also investigate the use of the proposed tool for analyzing the impact of technological advances in ESSs.

The paper is organized as follows. A literature review on related applications is presented in Section 2. The optimization model that embodies our modeling framework is presented in Section 3. In Section 4, we present the results obtained when applying the designed optimization model to our case study. Finally, conclusions and paths for future work are drawn in Section 5.

2. Literature Review

As mentioned in the previous section, over the last few years, several efforts have been devoted to providing a modeling and algorithmic framework for aiding strategic decision-making in energy system design; see [5] for a general reference on design by optimization of electrical energy systems. In this section, we review some of the most recent and relevant examples associated with our planning setting. Our review focuses on decision-making tools concerning decisions on: (i) the location and sizing of renewable-based generators; (ii) the layout and sizing of transmission grids; and (iii) the location and sizing of ESSs. After a detailed analysis, we have selected 18 articles, all of them published over the last six years, where optimal energy system planning approaches, concerning these decisions, are investigated. In Table 1, we summarize the characteristics of these articles according to three criteria: storage, generation and transmission. With respect to storage, we differentiate among those papers where capacity is decided ("capacity"), where location is decided ("location"), or where other characteristics ("others"), such as operational profiles ("Op.") or technological assessment ("Tech."), are within the decision scope. A similar classification is considered for generation decisions. In the case of transmission decisions, we identify which papers consider layout design ("layout"), which consider decisions on lines' capacity ("capacity") and which consider decisions on power flow over the designed or existing lines ("flow").

Table 1. Summary of the literature related to energy systems' planning incorporating renewable and conventional generation, transmission and storage sizing and design. Op., operational profiles; Tech., technological assessment.

Reference	Storage			Generation			Transmission		
	Capacity	Location	Others	Capacity	Location	Others	Layout	Capacity	Flow
[6]	✓								
[7]			Op.			Op.			
[8]						Op.			
[9]			Op.						
[10]	✓	✓	Op.				✓		
[11]	✓	✓	Op.				✓		
[12]	✓		Tech.						
[13]	✓		Op. and Tech.						
[14]	✓	✓	Op.						
[15]					✓	Tech.			
[16]				✓	✓		✓	✓	
[17]	✓	✓					✓	✓	
[18]		✓							✓
[19]		✓	Op.						✓
[20]	✓	✓							✓
[21]				✓		Op. and Tech.			
[22]	✓		Op.	✓		Op. and Tech.			
[23]			Op.	✓	✓	Op. and Tech.			
our work	✓	✓	Op. and Tech.	✓	✓	Connect.	✓	✓	✓

As can be seen from the table, some of the existing works mainly focus on the design of optimal storage systems; for instance, the work presented in [11]. In that paper, the authors provide a mathematical optimization model where the solutions correspond to the design of a storage system that defines its storage capacity, its location within the existing power system, its operative regime and how this system will be integrated, through new transmission lines, into the existing grid. On the contrary, other papers, such as [16], do not take into account the implementation of ESSs, and their core corresponds to the design of optimal renewable-based generation strategies (capacity and location), along with the definition of a corresponding grid (layout and capacity). A similar design setting as the one discussed in this paper is approached in [21]; in that work, an optimal investment plan of hybrid renewable generation technologies is studied in the context of microgrids. Similarly, a joint investment and operation framework is studied in [22] for the sizing of solar energy, wind energy and energy storage units using realistic microgrid data. Furthermore, the location and sizing of hybrid renewable generators over multiple locations in distributed microgrid systems is studied in [23]. In addition to these cases, the other articles summarized in Table 1 focus on different aspects of the decision-making setting motivated in the Introduction; nonetheless, none of them offers a general modeling framework integrating all of the decisions in a systematic fashion. Moreover, none of them emphasizes the need for considering the investor perspective as a fundamental characteristic of any real-world-oriented decision-making tool for large power systems. Our work, as can be seen from the last line of the table, offers an integrated decision-making framework that closes this methodological gap. Furthermore, we will show in the following section that our framework aims at designing economically-attractive energy systems by optimizing the economic value of the sought investment strategies.

3. Problem Definition

The core of our energy system planning tool corresponds to an optimization problem, the given MILP problem, that we coined as the Generation, Transmission, Storage Energy Location and Sizing Problem (GTSELSP). The GTSELSP involves technological and economical considerations, and it is associated with a quite complex mathematical formulation, which will be described in this section.

3.1. Problem Notation: Parameters and Variables

3.1.1. Sets

The sets that we use in our problem definition are the following:

- V_1 set of candidate locations where generation plants can be located.
- V_2 set of candidate locations where substations or line junctures can be located.
- O set of selling points, i.e., sub-stations of the existing power grid where energy is associated with a spot price.
- E_1 set of potential lines connecting generation locations from V_1 with locations in V_2 (i.e., $E_1 \subseteq (V_1 \times V_2) \cup (V_1 \times O)$).
- E_2 set of lines connecting locations from E_2 with both, other elements from E_2 and selling points from O (i.e., $E_2 \subseteq (V_2 \times V_2) \cup (V_2 \times O)$).

The designed tool seeks for strategic decisions, which are evaluated through a long-term evaluation of the corresponding operational decisions. The strategic evaluation time span corresponds to LT years. Each of these years is characterized by a representative year (equal for all years). This representative year is divided into cluster-days that represent shrunk periods (e.g., one month is shrunk into one day); index t will represent one hour of each of these cluster-days. As will be shown in Section 4, the composition of such cluster-days has a strong influence on the computational difficulty of the underlying mathematical optimization model; however, it does not really affect the structure of the corresponding solutions.

3.1.2. Generation Parameters

From a strategic design point of view, please consider the following generation parameters:

- P^{max} maximum generation capacity to be installed, at a single location (MWp).
- P^B_{it} available radiation for a given generation location $i \in V_1$ and a given hour t (MW/m^2).
- P_{stc} nominal power of a PV module under standard test conditions (STC) (MWp/m^2) of a photovoltaic (PV) module with a surface of a (m^2)). Hence, a PV generation plant will be comprised by a group (of say Q) of these modules, and its capacity will be given by QaP_{stc} (MWp).
- η_{inv} performance of the inverter that is planned to be used in the installed PV plants.
- ϕ_{dust} loss factor of the PV modules due to the presence of dust.

Additionally, from an economic point of view, we use the following parameters:

- A_{IG} fixed cost of installing a PV power plant (USD).
- B_{IG} variable cost of installing a PV power plant (USD/MW).
- $C_{IS/E}$ variable cost of installing a sub-station (USD/MW).
- A_{IS} variable cost of ESS installation (USD/MWh).
- P_{ot} sale price in node o at period t (USD/MW).

3.1.3. Transmission Parameters

Transmission takes place in lines between generation, transformation (i.e., substation) and the existing grid (where the energy is injected to the system). These lines are characterized by several technological and economic parameters, which are listed below:

- η_{DC} transmission loss factor associated with direct current (DC).
- η_{AC} transmission loss factor associated with alternate current (AC).
- D_{max} maximum extension of a single line segment due regulations (km).
- D_{ij} distance between point i and point j (km).
- η_{trans} loss factor associated with power transmission.
- A_{IL} fixed installation cost of installing one kilometer of transmission line (USD/km).
- B_{IL} variable installation cost of installing, per MW, one kilometer of transmission line (USD/MW/km).

3.1.4. Energy Storage Parameters

Energy storage systems operate according to the following technological and economical parameters:

Λ^c ESS' charge efficiency.
Λ^d ESS' discharge efficiency.
SOC minimum level of the state of charge that must be preserved in the ESS.
Φ^c maximum ESS' charge power flow (MW).
Φ^d maximum ESS' discharge power flow (MW).
S^{max} maximum storage capacity (MWh).
A_{IS} variable cost of ESS installation (USD/MWh).

3.1.5. Problem Variables

From a structural point of view, we use the following decision variables:

x_i binary variable, so that $x_i = 1$ if a generator (i.e., a PV plant) is installed at location $i \in V_1$, and $x_i = 0$ otherwise.
w_i installed fraction of P^{max} of installed generator at a given $i \in V_1$ ($w_i \in [0, 1]$).
P_i^n installed nominal power (PV) at a given $i \in V_1$. ($P_i^n = w_i P^{max}$).

Operation is characterized by the following decision variables:

ρ_{it} generated power from the generator located at $i \in V_1$ in period t.
f_{it} dispatched power from the generator located at $i \in V_1$ in period t.
δ_{it} spilled power in the generator located at $i \in V_1$ in period t.
z_j binary variable so that $z_j = 1$ if a substation is installed at location $j \in V_2$, and $z_j = 0$ otherwise.
κ_j^s capacity of installed substation, at a given $j \in V_2$.

The layout of transmission lines is associated with the following decision variables:

y_{ij}^1 binary variable, so that $y_{ij}^1 = 1$ if a line between i and j is built ($\{i, j\} \in E_1$), and $y_{ij}^1 = 0$ otherwise.
y_{jk}^2 binary variable, so that $y_{jk}^2 = 1$ if a line between j and k is built ($\{j, k\} \in E_2$), and $y_{jk}^2 = 0$ otherwise.
κ_{ij}^1 capacity of installed transmission line, between i and j ($\{i, j\} \in E_1$).
κ_{jk}^2 capacity of installed transmission line, between j and k is built ($\{j, k\} \in E_2$).

Furthermore, the operation on transmission lines associated with set E_1 is modeled by the decision variables listed below:

h_{ijt}^1 power flow on line $\{i, j\} \in E_1$ in period t.
δ_{ijt}^1 non-transmitted power through the line $\{i, j\} \in E_1$ in period t.
$Tloss_{ijt}^1$ power losses at line $\{i, j\} \in E_1$ in period t.
h_{ijt}^2 power flow on $\{i, j\} \in E_2$ in period t.
δ_{ijt}^2 non-transmitted power through the line $\{i, j\} \in E_2$ in period t.
$Tloss_{ijt}^2$ power losses on $\{i, j\} \in E_2$ in period t.

Additionally, for the location and sizing of the ESS' facilities, we consider the following variables:

s_i^n capacity of the ESS installed at $i \in V_1 \cup V_2$.
σ_i installed portion of S^{max} of installed ESS at $i \in V_1 \cup V_2$ ($\sigma_i \in [0, 1]$).

Note that an ESS facility can be installed only on nodes where a generator or a substation is located. Finally, the operation of the ESS is defined by the following variables:

s_{it} charge status, in period t, of the ESS located at i.
ϕ_{it}^c charging flow, in period t, at the ESS located at i.
ϕ_{it}^d discharging flow, in period t, at the ESS located at i.
z_{it}^ϕ binary variable, so that $z_{it}^\phi = 1$, if the ESS in i is discharged in period t, and $z_{it}^\phi = 0$, otherwise.

3.2. An MILP for the GTSELSP

Considering the notation presented above, we can now formulate the mixed integer program that encodes the strategic design tool proposed in this paper. The different types of constraints encompassing the GTSELSP are described in the remainder of this subsection.

3.2.1. Generation

The first group of constraints aims at characterizing the total energy produced at the (installed) generators; i.e., generation constraints, and they are given by:

$$P_i^n = P^{max} w_i \ \wedge \ w_i \leq x_i, \qquad \forall i \in V_1 \qquad (1)$$

$$\rho_{it} = P_{it}^B \frac{P_i^n}{P_{stc}} \eta_{DC} \eta_{inv} \eta_{AC} (1 - \phi_{dust}), \qquad \forall i \in V_1, \ \forall t \in T \qquad (2)$$

$$f_{it} + \delta_{it} = \rho_{it}, \qquad \forall i \in V_1, \ \forall t \in T \qquad (3)$$

$$\begin{array}{l} \mathbf{f} \in \mathbb{R}_{\geq 0}^{|V_1| \times T}, \boldsymbol{\rho} \in \mathbb{R}_{\geq 0}^{|V_1| \times T}, \\ \boldsymbol{\delta} \in \mathbb{R}_{\geq 0}^{|V_1| \times T}, \mathbf{P^n} \in \mathbb{R}_{\geq 0}^{|V_1|}, \ \mathbf{w} \in \mathbb{R}_{\geq 0 \wedge \leq 1}^{|V_1|} \end{array} . \qquad (4)$$

Constraint (1) ensures that generation capacity is installed at given site $i \in V_i$ ($w_i \geq 0$; thus $P_i^n \geq 0$), if and only if it is decided that at such site, a generator is installed ($x_i = 1$); moreover, the same constraint ensures that the installed capacity will not exceed the predefined maximum capacity P^{max}. The actual energy flow that leaves a given generation site i at a given time t (ρ_{it}) is characterized by the constraints in (2); such energy flow depends not only on the installed capacity (P_i^n) the nominal power (P_{stc}) or the effect of the dust ($(1 - \phi_{dust})$), but also on the losses in the line from the panels to the inverter (η_{DC}), the efficiency of the inverter (η_{inv}) and the losses from the inverter to the connection point into the grid (η_{AC}). Constraint (3) allows one to split the energy generated at site $i \in V_1$ in time $t \in T$ (ρ_{it}), into two components; f_{it} which is the energy that will be injected into the grid, and δ_{it}, which represents the (eventually) spilled energy. Finally, the nature of the involved variables is expressed in (2).

3.2.2. Transmission Topology

The second group of constraints defines the topology and size of the transmission layout, and it is given as follows:

$$\sum_{\{i,j\} \in \delta(i)} y_{ij}^1 \geq x_i, \qquad \forall i \in V_1 \qquad (5)$$

$$y_{ij}^1 \leq x_i, \qquad \forall \{i,j\} \in \delta(i), \ \forall i \in V_1 \qquad (6)$$

$$\sum_{\{i,o\} \in \delta(o)} y_{io}^1 + \sum_{\{j,o\} \in \delta(o)} y_{jo}^2 \geq x_i, \qquad \forall i \in V_1 \qquad (7)$$

$$y_{ij}^1 \leq z_j, \qquad \forall \{i,j\} \in \delta(j), \forall j \in V_2 \cup O \qquad (8)$$

$$y_{ij}^2 \leq z_i \ \wedge \ y_{ij}^2 \leq z_j, \qquad \forall i \in V_2 \cup O, \ \forall j \in V_2 \cup O, \ \forall \{i,j\} \in E_2 \qquad (9)$$

$$y_{ij}^1 D_{ij} \leq D_{max}, \qquad \forall \{i,j\} \in E_1 \qquad (10)$$

$$y_{ij}^2 D_{ij} \leq D_{max}, \qquad \forall \{i,j\} \in E_2 \qquad (11)$$

$$\begin{array}{l} \mathbf{y^1} \in \{0,1\}^{|E_1|}, \mathbf{x} \in \{0,1\}^{|V_1|}, \\ \mathbf{y^2} \in \{0,1\}^{|E_2|} \ \mathbf{z} \in \{0,1\}^{|V_2 \cup O|} \end{array} . \qquad (12)$$

The constraints in (5) ensure that at least one line must by built from the generation site $i \in V_1$ in the case that a power plant is installed at such a site (i.e., if $x_i = 1$, then $\sum_{\{i,j\} \in \delta(i)} y_{ij}^1 \geq 1$). Additionally, Constraints (6) and (7), ensure the right topology of the connections associated with generation points. The type and topology of the connections from (or towards) installed substations are characterized by Constraints (8) and (9). Constraints (10) and (11) imply that due to technical requirements, the sought layout must be such that no direct connection shall by longer than D_{max}. The constraints in (12) characterize the nature of the variables.

3.2.3. Energy Storage Systems

The functioning of the (eventually installed) energy storage systems must fulfill several technological rules. Such rules are characterized by the following set of constraints;

$$\sigma_i \leq x_i, \qquad \forall i \in V_1 \tag{13}$$

$$\sigma_i \leq z_i, \qquad \forall i \in V_2 \tag{14}$$

$$s_i^n = S^{max}\sigma_i, \qquad \forall i \in V_2 \cup V_1 \tag{15}$$

$$s_{it} \leq s_i^n, \qquad \forall i \in V_2 \cup V_1,\ \forall t \in T \tag{16}$$

$$s_{it} \geq s_i^n \cdot SOC, \qquad \forall i \in V_2 \cup V_1,\ \forall t \in T \tag{17}$$

$$s_{it} = s_{i(t-1)} + \Lambda^c \phi_{it}^c - \frac{\phi_{it}^d}{\Lambda^d}, \qquad \forall i \in V_2 \cup V_1,\ \forall t \in T \tag{18}$$

$$0 \leq \phi_{it}^c \leq \Phi^c z_{it}^{\phi}, \qquad \forall i \in V_2 \cup V_1,\ \forall t \in T \tag{19}$$

$$0 \leq \phi_{it}^d \leq \Phi^d (1 - z_{it}^{\phi}), \qquad \forall i \in V_2 \cup V_1,\ \forall t \in T \tag{20}$$

$$s_{i0} = s^n \cdot SOC, \qquad \forall i \in V_2 \cup V_1 \tag{21}$$

$$\begin{array}{r} \mathbf{z}^{\boldsymbol{\phi}} \in \{0,1\}^{|V_2 \cup V_1| \times T},\ \boldsymbol{\sigma} \in \mathbb{R}_{\geq 0\ \wedge\ \leq 1}^{|V_2 \cup V_1|}, \\ \mathbf{s^n} \in \mathbb{R}_{\geq 0}^{|V_2 \cup V_1|},\ \mathbf{s} \in \mathbb{R}_{\geq 0}^{|V_2 \cup V_1| \times T}, \\ \boldsymbol{\Phi}^{\boldsymbol{c}} \in \mathbb{R}_{\geq 0}^{|V_2 \cup V_1| \times T}\ \ \text{y}\ \ \boldsymbol{\Phi}^{\boldsymbol{d}} \in \mathbb{R}_{\geq 0}^{|V_2 \cup V_1| \times T} \end{array} . \tag{22}$$

An ESS can be installed at a given site $i \in V_1 \cup V_2$ ($\sigma_i \geq 0$), only if at that site, a generator ($x_i = 1$, $i \in V_1$) or a substation is installed ($z_i = 1$, $i \in V_2$); this requirement is encoded by Constraints (13) and (14), respectively. The constraints in (15) ensure that the nominal installed capacity of an ESS at a given site $i \in V_1 \cup V_2$ (s_i^n) is a portion (given by σ_i) of the maximum reference capacity S^{max}. The amount of energy that can be stored at a given ESS located at $i \in V_1 \cup V_2$, at a period $t \in T$ (s_{it}), cannot exceed its corresponding capacity s_i^n; this is enforced by the constraints in (16). Likewise, the minimum energy that shall be kept at the same ESS shall not be less than the state of charge threshold $s_i^n \cdot SOC$; this is modeled by the constraints in (17). The constraints in (18) model the fact that the energy stored in the ESS installed at $i \in V_1 \cup V_2$, at a given period $t \in T$, depends on energy stored in the previous period ($s_{i(t-1)}$), the energy that is charged in t ($\Lambda^c \phi_{it}^c$) and the energy that is discharged in t ($\frac{\phi_{it}^d}{\Lambda^d}$). Complementary to this, the amount of energy that is charged and discharged at a given period t, in an ESS located in $i \in V_1 \cup V_2$, is characterized by Constraints (19) and (20), respectively. Furthermore, the initial conditions of the installed ESS are modeled by the constraints in (21). The nature of the associated variables is characterized by Constraint (22).

3.2.4. Power Balance

So far, we have presented the constraints that define the feasibility conditions of the generation and of the energy storage; however, no linkage among them has been established. This linkage, or coupling, is given by the power balance at each node (generation, storage, and/or substation) and involves the corresponding transmission lines associated with it (and of course, the energy flow through them). The following constraints ensure a correct interdependence among the different energy flows:

$$\sum_{\{i,j\}\in\delta(i)} \left(h^1_{ijt} + \delta^1_{ijt} + Tloss^1_{ijt}\right) = f_{it} + \left(\phi^d_{it} - \phi^c_{it}\right), \qquad \forall i \in V_1,\ \forall t \in T \tag{23}$$

$$\sum_{\{i,j\}\in\delta(j)} h^2_{ijt} + \sum_{\{i,j\}\in\delta(j)} h^1_{ijt} + \left(\phi^d_{jt} - \phi^c_{jt}\right) = \sum_{\{j,k\}\in\delta(j)} \left(h^2_{jkt} + \delta^2_{jkt} + Tloss^2_{jkt}\right), \qquad \forall j \in V_2,\ \forall t \in T \tag{24}$$

$$\begin{aligned} &\mathbf{h^1} \in \mathbb{R}^{|E_1|\times T}_{\geq 0},\ \boldsymbol{\delta^1} \in \mathbb{R}^{|E_1|\times T}_{\geq 0},\ \mathbf{Tloss^1} \in \mathbb{R}^{|E_1|\times T}_{\geq 0},\\ &\mathbf{h^2} \in \mathbb{R}^{|E_2|\times T}_{\geq 0},\ \boldsymbol{\delta^2} \in \mathbb{R}^{|E_2|\times T}_{\geq 0},\ \mathbf{Tloss^2} \in \mathbb{R}^{|E_2|\times T}_{\geq 0}. \end{aligned} \tag{25}$$

The constraints in (23) model the following phenomenon: at a given generation node $i \in V_1$ and in given period $t \in T$, the total energy that leaves i through the corresponding lines ($\sum_{\{i,j\}\in\delta(i)} \left(h^1_{ijt} + \delta^1_{ijt} + Tloss^1_{ijt}\right)$) must be equal to the energy that is injected from the corresponding generator (f_{it}), plus the energy that is discharged from the corresponding ESS (ϕ^d_{it}), minus the energy that is charged into the corresponding ESS (ϕ^c_{it}).

Similarly, the constraints in (24) ensure a correct energy balance within the intermediate nodes V_2. Basically, these constraints ensure that the total energy that is injected into a given $j \in V_2$ ($\sum_{\{i,j\}\in\delta(j)} h^2_{ijt} + \sum_{\{i,j\}\in\delta(j)} h^1_{ijt}$), plus the energy that is discharged from the ESS located at j (ϕ^d_{jt}), minus the energy that is charged into the ESS located at j (ϕ^c_{jt}), must be equal to the energy that leaves j through the associated transmission lines ($\sum_{\{j,k\}\in\delta(j)} \left(h^2_{jkt} + \delta^2_{jkt} + Tloss^2_{jkt}\right)$).

Following (23) and (24), we have that the energy that is injected into a selling point $j \in O$, in $t \in T$, is modeled by $\sum_{\{i,j\}\in\delta(j)} h^2_{ijt} + \sum_{\{i,j\}\in\delta(j)} h^1_{ijt}$. Note that we do not allow generation nor the installation of ESS in the points comprising O.

3.2.5. Transmission and Substation Capacities

The power dispatched from the generation and substation (or line juncture) nodes must respect the capacity of the transmission lines. In the GTSELSP, we not only decide on the layout of the transmission grid, i.e., which lines are constructed, but also the capacity of these lines. The relation among layout, capacity and transmission decisions is encoded by the following constraints;

$$h^1_{ijt} \leq \kappa^1_{ij} \leq My^1_{ij}, \qquad \forall \{i,j\} \in E_1,\ \forall t \in T \tag{26}$$

$$h^2_{ijt} \leq \kappa^2_{ij} \leq My^2_{ij}, \qquad \forall \{i,j\} \in E_2,\ \forall t \in T \tag{27}$$

$$\boldsymbol{\kappa^1} \in \mathbb{R}^{|E_1|}_{\geq 0},\ \boldsymbol{\kappa^2} \in \mathbb{R}^{|E_2|}_{\geq 0}. \tag{28}$$

The constraints in (26) ensure, on the one hand, that the power flow through line $\{i,j\} \in E_1$, at a given period $t \in T$ (h^1_{ijt}), must be less than the installed capacity κ^1_{ij}, and on the other, that we will install such capacity ($\kappa^1_{ij} \geq 0$) if an only if a line is built on $\{i,j\}$, i.e., $y^1_{ij} = 1$ (note that M is a sufficiently large number). The constraints in (27) are equivalent to (26), but on the connections comprising E_2. The nature of the capacity variables is given by (28).

In addition to the transmission lines, the capacity of the installed substations is also part of the decision-making process; the associated constraints correspond to:

$$\sum_{\{i,j\}\in\delta(j)} h^1_{ijt} + \sum_{\{i,j\}\in\delta(j)} h^2_{ijt} \leq \kappa^s_j \leq Mz_j, \qquad \forall j \in V_2,\ \forall t \in T, \tag{29}$$

$$\sum_{\{j,k\}\in\delta(j)} \left(h^2_{jkt} + Tloss^2_{jkt}\right) \leq \kappa^s_j \leq Mz_j, \qquad \forall j \in V_2,\ \forall t \in T \tag{30}$$

$$\boldsymbol{\kappa^s} \in \mathbb{R}^{|V_2|}_{\geq 0}. \tag{31}$$

The constraints in (29) ensure that the total power that is injected in the substation located at $j \in V_2$ in period $t \in T$ ($\sum_{\{i,j\}\in\delta(j)} h^1_{ijt} + \sum_{\{i,j\}\in\delta(j)} h^2_{ijt} \leq \kappa^s_j$) must be less than or equal to the capacity of the substation (κ^s_j); moreover, the same constraints also ensure that if no substation is located at j ($z_j = 0$), then the corresponding capacity must be zero, as well. Equivalently, the constraints in (30), which are very similar to (29), ensure that the total power dispatched from j, including the corresponding losses ($\sum_{\{j,k\}\in\delta(j)} \left(h^2_{jkt} + Tloss^2_{jkt}\right)$), must also respect the corresponding substation capacity. Finally, Constraint (31) ensures the nature of the substation installation variables.

Note that Constraints (26), (27), (29) and (30) ensure a proper coupling of the grid elements by modeling a correct power flow balance.

3.2.6. Losses and Boundary Conditions

As in any real-world-oriented application, we need to take into account the losses in the transmission lines. These losses depend on the power flow and on the efficiency factor, which is inherent to the transmission lines. The constraints:

$$Tloss^1_{ijt} = h^1_{ijt}\eta_{trans}, \qquad \forall\{i,j\} \in E_1,\ \forall t \in T \tag{32}$$

$$Tloss^2_{ijt} = h^2_{ijt}\eta_{trans}, \qquad \forall\{i,j\} \in E_2,\ \forall t \in T, \tag{33}$$

allow us to express the transmission losses, $Tloss^1_{ijt}$ and $Tloss^2_{ijt}$, as a portion (given by η_{trans}) of the power that flows through the lines (h^1_{ijt} and h^2_{ijt}, respectively).

Along with the constraints presented so far, we also impose boundary conditions that must be preserved: (i) transmission lines cannot be built from a node and towards the same node; (ii) there are no lines outgoing from a selling point $o \in O$ (i.e., the grid ends at the selling points); and (iii) we assume that in all selling points $o \in O$, there must be a substation (which switches from the power transmission towards the power distribution). Such conditions are modeled by the following constraints:

$$y^2_{ii} = 0, \qquad \forall i \in \delta(i) \tag{34}$$

$$y^2_{oi} = 0, \qquad \forall\{o,i\} \in \delta(o) \tag{35}$$

$$z_o = 1, \qquad \forall o \in O. \tag{36}$$

3.2.7. The GTSELSP

So far, we have presented all of the constraints that characterize the set of feasible solutions of the GTSELSP. For a full representation of the problem, we shall consider that the objective function corresponds to the maximization of the economical profit of the sought solutions; for a given evaluation time span, such profit is given by the difference between the incomes obtained from selling the produced energy, minus the installation and operation costs. The total income is given by:

$$income = \sum_{t=1}^{T}\sum_{o\in O}\left[P_{ot}\left(\sum_{\{i,o\}\in\delta(o)} h^1_{iot} + \sum_{\{i,o\}\in\delta(o)} h^2_{iot}\right)\right],$$

and it corresponds to the sum, over the planning horizon, of the gains induced by the power that flows towards the selling points. The total generation installation costs are given by:

$$C_{PV} = \sum_{i\in V_1}\left(A_{IG}x_i + B_{IG}P^n_i\right).$$

Likewise, the total cost of the grid layout construction is given by

$$C_{trans1} = \sum_{\{i,j\}\in E_1} \left(A_{IL}y^1_{ij} + B_{IL}\kappa^1_{ij}\right) D_{ij}$$
$$C_{trans2} = \sum_{\{i,j\}\in E_2} \left(A_{IL}y^2_{ij} + B_{IL}\kappa^2_{ij}\right) D_{ij};$$

as can be seen, the grid installation cost does not only depend on the whether a line is built or not, but also on its capacity. Another crucial investment cost corresponds to the ESS installation, which is given by:

$$C_{ESS} = \sum_{i\in V_2\cup V_1} A_{IS}s^n_i;$$

as can be seen, the total cost associated with the ESS infrastructure, mainly depends on the size of the installed units. Finally, the latter installation cost that must be considered corresponds to the substations installation cost, given by:

$$C_{SE} = \sum_{i\in V_2} C_{IS/E}\kappa^s_i;$$

as well as for the ESS, the substations' installation cost basically depends on the size of the installed infrastructure. Consequently, we can model the GTSELSP objective (profit) function $profit(\mathbf{f}, \boldsymbol{\rho}, \delta, \mathbf{P^n}, \mathbf{w}, \mathbf{y^1}, \mathbf{x}, \mathbf{y^2}, \mathbf{z}, \mathbf{z}^{\Phi}, \sigma, \mathbf{s^n}, \mathbf{s}, \mathbf{\Phi}^c, \mathbf{\Phi}^d, \mathbf{h^1}, \delta^1, \mathbf{Tloss^1}, \mathbf{h^2}, \delta^2, \mathbf{Tloss^2}, \boldsymbol{\kappa}^1, \boldsymbol{\kappa}^2, \boldsymbol{\kappa}^s) = profit$ as:

$$profit = (income - C_{PV} - C_{trans1} - C_{trans2} - C_{ESS} - C_{S/E})\,.$$

Note that in this profit formula, we have not included generation costs since they are assumed to be negligible when compared to the installation costs and the total incomes. Since future incomes strongly depend on the future realizations of energy market, we present in Section 4.2 an econometric strategy to address this variability. Please note that in the case that generation costs are relevant, one should adopt complementary strategies as the cost benefits analysis discussed in [24].

Considering all of the elements presented so far, a MILP formulation of the GTSELSP is given by:

$$\text{(GTSELSP)} \qquad profit^* = \max profit$$
$$\text{s.t.} \quad (1)\text{–}(4), (5)\text{–}(12), (13)\text{–}(22), (23)\text{–}(25)$$
$$(26)\text{–}(31), (32)\text{–}(33), (34)\text{–}(36).$$

This MILP formulation has a polynomial number of variables and constraints; therefore, one could use any of the state-of-the-art solvers, e.g., CPLEX, to tackle it. This enhances the practical use of our approach, since there is no need for developing sophisticated algorithmic strategies to tackle the GTSELSP.

4. Results and Discussion

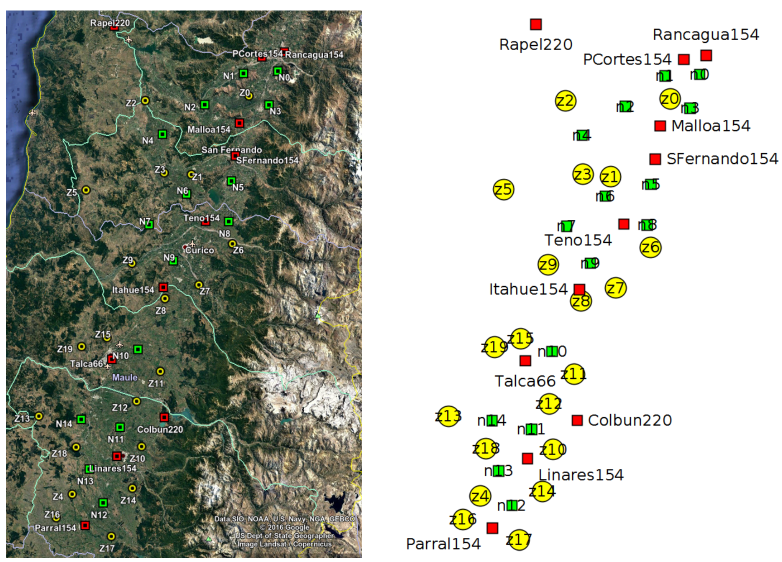
In this section, we will report results on a case study in which the GTSELSP is used as the core element for the design and evaluation of a large-scale (multi-site) PV-based generation infrastructure. We first contextualize the case study; later, we present an exhaustive analysis of how the problem parameters were estimated; and finally, we present and discuss the obtained results.

4.1. Case Study: Chilean Central Region

Chile has one of the best solar potential in the world; in particular, in the northern region, the average radiation is as high as 2055 (kWh/m^2) and with a quite stable profile during the day. However, the main sources of energy demand are located in the central region, where the radiation potential fluctuates more. Moreover, the current Chilean power system is comprised by a northern grid and a central grid, which are not connected, making it impossible to exploit the northern solar radiation to satisfy the power demand in the central region. Therefore, a need arises to be able to efficiently exploit the solar potential of the Chilean central region in order to contribute, by a clean and sustainable source, to satisfying the corresponding energy demand.

The following characteristics of the central region play an important role in our decision setting: (i) the average radiation is above 1555 (kWh/m^2) , which is still higher than the radiation of countries like Germany (1261 (kWh/m^2)), with high penetration of solar generation; (ii) this area incorporates 51.51% of the Chilean population, implying quite large urban areas, making it impossible to install very large power plants combining substations or ESS facilities; (iii) besides the domiciliary demand, there is also a vast need for electricity for industrial activities such as agriculture, mining and retail; (iv) the orography in this sector is more complex than in the northern region, which incorporates another level of complexity in the design process.

The geographical zone where the solar power generation system is to be designed corresponds to the 6th and 7th regions (see Figure 1a). This zone fulfills two important operative requirements: it is practically attached to the metropolitan area (where most of the demand is concentrated), and there is a quite complex existing power grid infrastructure where our designed power system can eventually be incorporated. In the existing grid, there exists a mix of hydroelectric, diesel and gas power plants, whose generation capacity mainly contributes to satisfying the power demand of the metropolitan area.



(**a**) Satellite view (**b**) Schematic view

Figure 1. Graphical representation of the studied instance. We use squares to represent to substations; circles to represent generation zones (V_1); green color indicates possible substations (V_2); and red color for existing substations (O).

In Figure 1b, we display a schematic representation of the geographical data of the considered case study. In the considered zone, we have identified 20 areas where the generation plant could be located; these 20 areas, which comprise set V_1, are represented by yellow circles in Figure 1a,b. The location of potential substations, which yields set V_2, are represented by 15 green squares. The existing substations, where energy is to be injected (i.e., sold) to the power grid, are represented by 11 red squares, comprising set O. Existing, as well as potential transmission lines are not shown for ease of exposition. Besides this topological data, we have curated all other problem parameters; this will be described in detail in the following subsection.

4.2. Estimation and Model Parameters

We have divided the problem parameters into five sets: (i) energy resource parameters, associated with generation zones; (ii) technological parameters, associated with power plants and transmission lines; (iii) generation and transmission construction costs; (iv) energy storage parameters; and (v) spot prices. The process for curating all of these parameters is presented below.

4.2.1. Solar Radiation Potential

The solar radiation potential of the 20 generation zones was calculated using one year of data (with a 30-min resolution) using the tool Explorador Solar. This tool, which was developed by an initiative of the Chilean Ministry of Energy [25], allows on to simulate, using sophisticated meteorological models, daily solar radiation in any location of the Chilean territory for a whole (typical) year.

In Figure 2, we report the radiation potential curve (along the day) of the 20 locations, for four different days; 4 February (summer), 8 May (autumn), 15 August (winter) and 18 November (spring). In these examples, we can see that while the generation potential of the different areas is similar in summer and somewhat in spring, in autumn and, in particular, in winter, this is not the case. This latter situation, which does not occur in northern regions, reinforces the need for an optimized policy design in order to maximize the potential mix along the year.

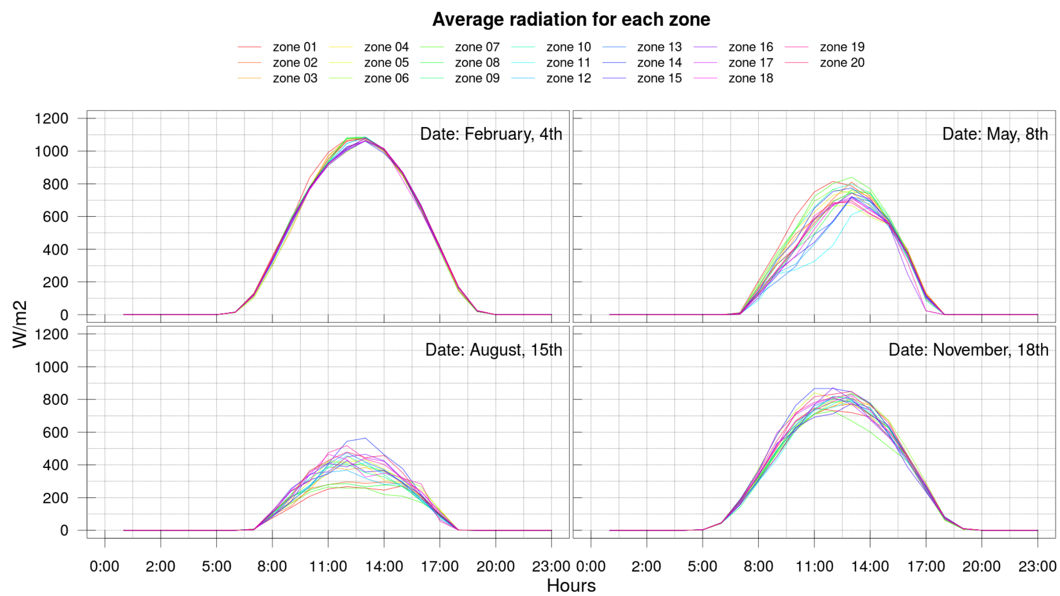


Figure 2. Representation of available average radiation for each studied zone.

4.2.2. Generation and Transmission Technological Parameters

To determinate the appropriate instance setup, we needed to find some technological parameters related to power plants and transmission lines.

The necessary parameters related to power plants have been determined using information from Vivest Energías Renovables S.A. (Vivest) [26]. Then, we obtain P_{stc} = 1000 (W/m^2), η_{inv} = 98%, ϕ_{dust} = 5%, η_{DC} = 97% and η_{AC} = 97%. On the other hand, we also receive the electric transmission loss performance parameter from Vivest. This corresponds to η_{trans} = 1.5% [26].

Furthermore, in our model, we have P^{max} = 200 (MWp). This parameter comes from the latest renewable energy projects that have been installed and that actually operate in Chile [27]. Finally,

D_{max} = 40 (km), which is the distance threshold defined by the Chilean power operator for ensuring distribution quality of the Chilean grid [28].

4.2.3. Generation and Transmission Construction Costs

The total construction cost associated with PV power plants (measured by U.S. dollars per MW) was curated according to the information provided by [26]. We obtained construction costs associated with different plant sizes, and we used this information to construct, via regression, a linear function associating generation power with construction cost. This is shown in Figure 3; the points in the plot correspond to the information at hand, while the line (along with the information incorporated in the plot) corresponds to the linear regression. As expected, we can see that the construction cost lineally depends ($r^2 = 0.99$) on the generation potential of the PV power plant. Therefore, $A_{IG} = 1{,}337{,}963$ (USD) and $B_{IG} = 1106$ (USD/MW).

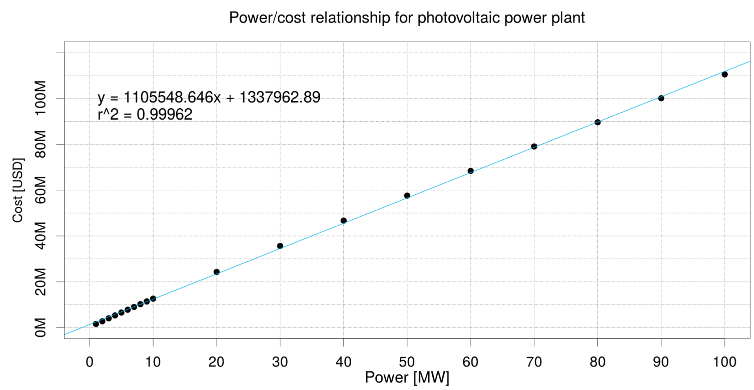


Figure 3. Relationship between power and construction cost for PV power plants.

For defining the construction cost of transmission lines, we used the, relatively recent, information published in [29–32], where the construction cost (in USD per km) associated with seven different lines' transmission capacity (in MW) is reported. In Figure 4, we display these values by colored points (colors depending on the source). As for the plant construction cost, we also compute a linear regression for approximating both, a fixed and a variable line construction cost. From the obtained regression, we have that $A_{IL} = 239,754$ (USD/km) and $B_{IL} = 0.396$ (USD/MW/km). Although the obtained regression is not that accurate ($r^2 = 0.60$), we believe that it is a quite fair approximation; a better estimation of this cost is out of the scope of this work, and it can be incorporated later on when the tool is used in practice.

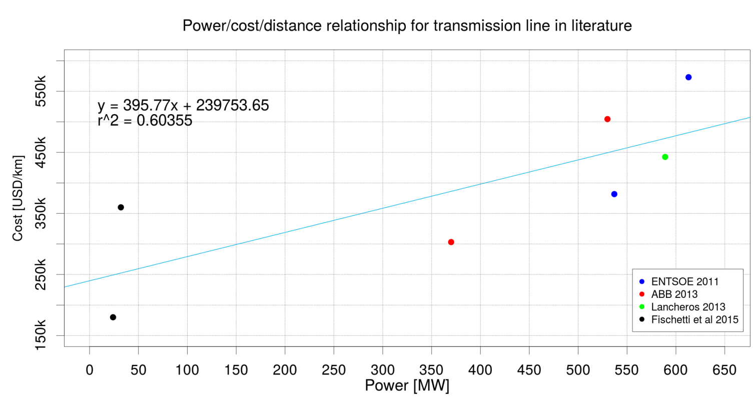


Figure 4. Relationship between power and construction cost for a 1 (km) transmission line.

The installation cost of substations was approximated by a similar scheme. We considered the installation cost data (in USD/MW) published in [29–31,33]. In Figure 5, we display the published construction costs (colored points) and the linear regression associated with these costs. From the obtained regression ($r^2 = 0.99$), we have that $C_{IS/E} = 125.6$ (USD/MW) (fixed installation costs were negligible).

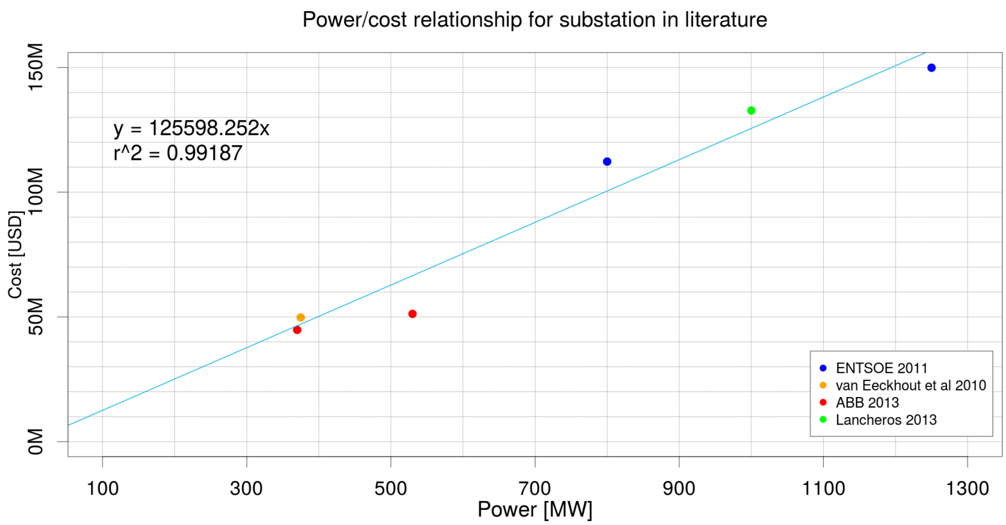


Figure 5. Relationship between power and construction cost for substations.

4.2.4. Energy Storage Parameters

As said before, one of the key contributions of our work consists of assessing the economic impact of different ESSs as part of a solar power generation system embedded into a large existing power grid. For such a purpose, we evaluated four storage technologies: the considered technologies are a sodium-nickel chloride cell technology from the so-called Zeolite Battery Research Africa battery protect (ZEBRA), Pb-acid batteries (PBAC), pumped hydroelectric storage (PHS) and compressed air energy storage (CAES). These four technologies are based on different technological principles and, therefore, operate according to different values of their parameters (see Section 3.1). In Table 2, we present the values of the parameters associated with the different technologies and the literature reference from where these values were obtained.

Table 2. Different technologies for ESS that we consider in this work. ZEBRA, Zeolite Battery Research Africa battery of sodium-nickel chloride cell; PBAC, Pb-acid batteries; PHS, pumped hydroelectric storage; CAES, compressed air energy storage.

	S_{max} (MWh)	Λ^c, Λ^d (%)	Φ^c, Φ^d (MW)	SOC_{min} (%)	Cost (USD/MWh)	Lifetime (Years)	Reference
ZEBRA	0.3	94.9	0.075	25	220,000	12	[34]
PBAC	20	86.6	5	40	450,000	10	[26,34,35]
PHS	200	92.0	50	10	200,000	30	[34]
CAES	200	89.0	25	10	50,000	30	[34]

These technologies are in different stages of development. While ZEBRA is a quite new technology, still in the R&D stage, PBAC, PHS and CAES are well-known solutions. Since PHS is the most common technology in our decision-making setting [34], we will consider it as our default technology.

Nonetheless, in the following subsection, we will also compare the economic impact of using each of these ESS technologies. As a final remark, in all of our experiments, we assume that the installed ESSs are renewed, within the evaluation horizon (30 years in this case), according to the lifetime reported in Table 2.

4.2.5. Modeling and Estimation of Spot Prices

The energy generated in the different power plants (either the existing or the potential ones) is injected into the grid at the substations. At these substations, the system operator pays for the injected energy at the so-called spot price, which depends on the demand and the generation mix.

For estimating the spot prices in the existing substations, we used historical data from the Economic Dispatch Load Center of the Central Interconnected System (CDEC-SIC) between the years 2015 and 2016 [36]. The resolution of this database is 1 h for each substation. These data were processed by an ad hoc adaptation of a bootstrapping technique [37,38], to produce, for each substation, 1000 possible series (each of them associating one value per hour). Afterwards, these series were used to approximate a so-called empirical density distribution price (EDDP).

In Figures 6 and 7, we show for the Itahue154 and Rapel220 substations, respectively, the EDDPs estimated for four different days in eight different hours. In the presented examples, we can clearly see that the series follow bi-modal distributions; the first mode (associated with lower prices) corresponds to the efficient supply of energy that satisfies most of the demand, while the second mode (associated with higher prices) is likely to be related with contingencies in the power supply or peaks of the demand. The described behavior is quite clear for the case of Itahue154, which represents the spot price behavior of most of the considered substations. As a matter of fact, the different percentiles (25, 50 and 75) are homogeneously located along the series. However, this is not the case for Rapel220, where spot prices present a higher variability among different days and hours, and the chances of having higher spot prices (>250 (USD/MWh)) are unusually high. It is important to point out that a bi-modal distribution of spot prices shall be common in efficiently-operated power systems, where spot prices are highly correlated with mixed power generation technologies and efficient demand response.

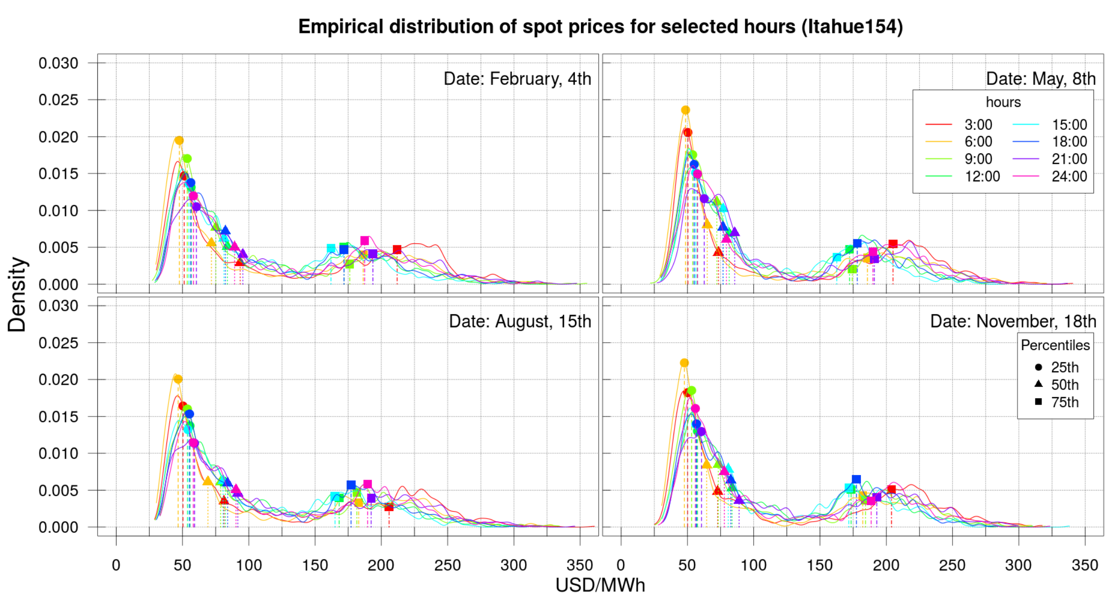


Figure 6. Empirical density distribution spot prices for Itahue154 in four representative days of the year.

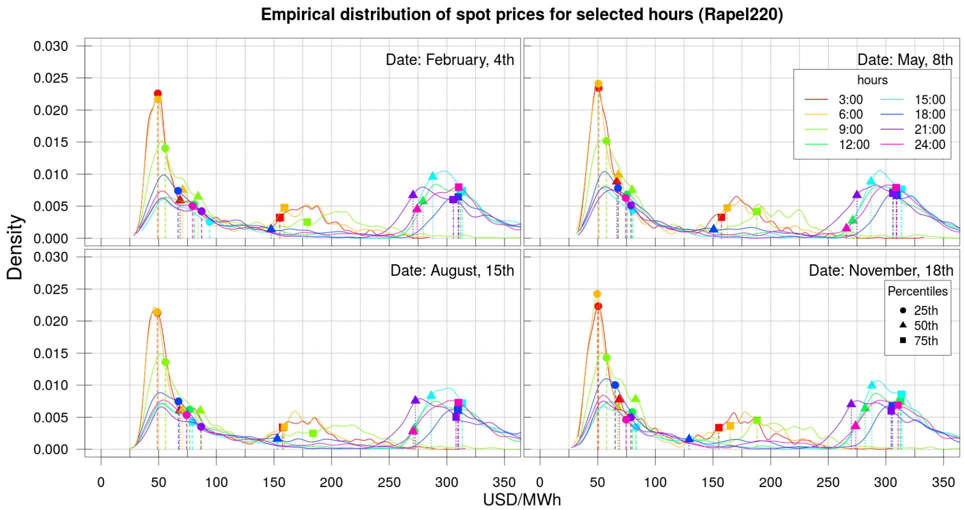
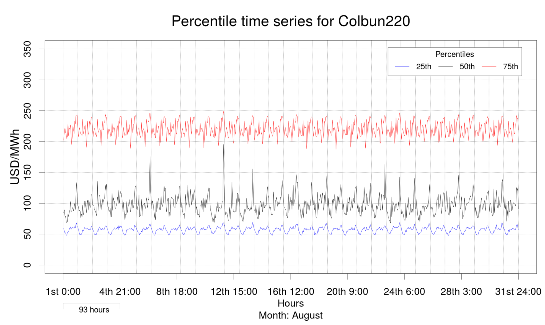
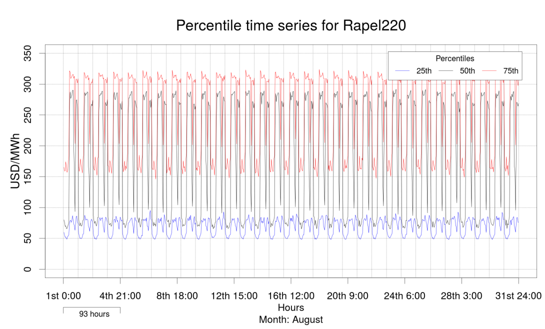


Figure 7. Empirical density distribution spot prices for Rapel220 in four representative days of the year.

Once the spot price distributions were estimated, we followed the typical approach used for the economical-financial evaluation of projects and generated three scenarios: a pessimistic one (associated with the 25th percentile, P25, of the corresponding EDDP), a possible one (associated with the 50th percentile, P50, of the corresponding EDDP) and an optimistic one (associated with the 75th percentile, P75, of the corresponding EDDP). In Figure 8, we display the series associated with these three scenarios for two substations (Colbun220 and Rapel220) during the month of August. For Colbun220, there is a clear differentiation among the three scenarios, while for Rapel220, it holds that P50 and P75 tend to overlap.



(**a**) Colbun220



(**b**) Rapel220

Figure 8. Percentile time series for (**a**) Colbun220 and (**b**) Rapel220 for the month of August.

For complementing the results shown in Figure 8, we show in Figure 9 the time series associated with the 11 substations for P50 during the month of August. From this plot, it is clear that injecting energy in different substations leads to very different outcomes in terms of the collected revenue. While most of the prices fluctuate in the band 75–100 (USD/MWh), there are some cases (Rapel220 and Colbun220) where the price can be as high as 280 (USD/MWh).

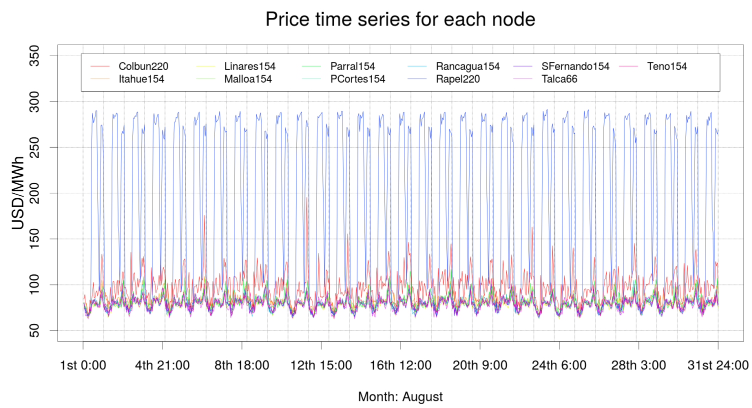


Figure 9. Spot price time series for each existing substation on the studied region for P50 and for the whole month of August.

The complex way that spot prices behave at different substations and at different times of the day reinforces the need for an accurate design tool that properly incorporates the nature of spot prices into an ad hoc mathematical optimization model.

4.3. Discussion and Results Analysis

In Section 4.2, we have described from where we have obtained or how we have estimated the parameters of the GTSELSP. In the remainder of this section, we present and discuss the results obtained when solving the GTSELSP and performing a sensitivity analysis.

Since the GTSELSP seeks for an optimal investment plan, in terms of the attained profit, the obtained solutions are not only measured by the underlying objective function value (annualized profit, measured by millions of USD per year), but also by its components (annualized investment costs and annualized incomes, measured by millions of USD per year) and by the corresponding internal rate of return (IRR, in percentage), which corresponds to the minimum discount rate that makes the project economically feasible (in this case, in a 30-year planning horizon).

4.3.1. Experimental Setting

We run our experiments on an Intel® Core™ i7-4702MQ 2.20-GHz machine with 16 GB RAM and Ubuntu 16.04 LTS. The underlying GTSELSP instances were solved using ILOG® CPLEX® 12.6.3; we used default setting, except for the primal-dual gap, which was set to 0.2%.

4.3.2. Time Resolution

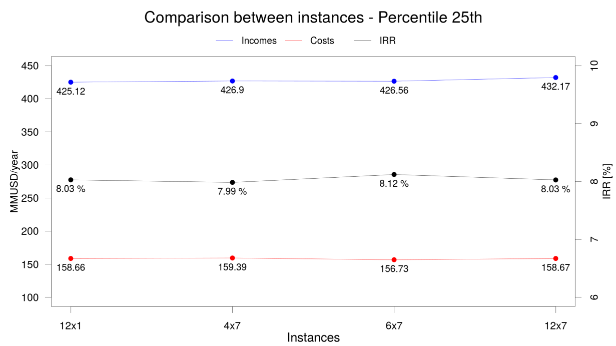
The first part of our computational analysis consists of determining which is the best time resolution for running the whole set of experiments. For such a purpose, we considered four time resolution structures:

- '12×1': One year is characterized by 12 days, so that each day encodes one month. Since every day is comprised of 24 h, this setting yields $T = \{1, 2, \ldots, 24 \times 12 = 288\}$.
- '4×7': One year is characterized by four weeks, so that each week encodes three months. In this case, we have $T = \{1, 2, \ldots, 4 \times 7 \times 24 = 672\}$.

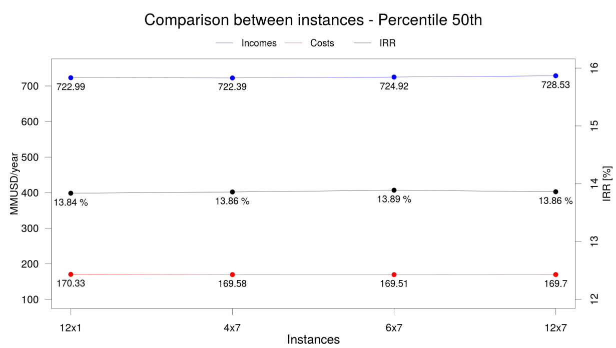
- ‘6×7’: One year is characterized by six weeks, so that each week encodes two months. In this case, we have $T = \{1, 2, \ldots, 6 \times 7 \times 24 = 1008\}$.
- ‘12×7’: One year is characterized by 12 weeks, so that each week encodes one month. In this case, we have $T = \{1, 2, \ldots, 12 \times 7 \times 24 = 2016\}$.

Since these four time resolution alternatives yield different sizes of set T, one would expect that the resulting MILP behaves numerically and computationally different.

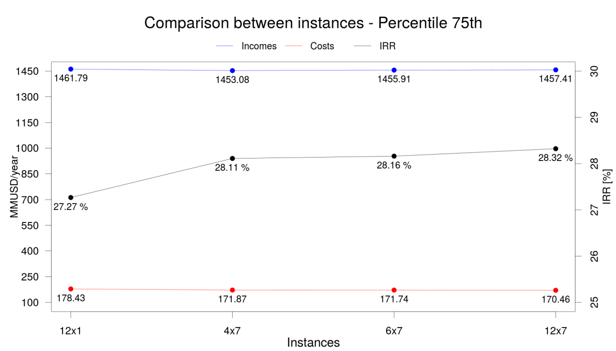
In Figure 10a–c, we report the economical behavior (measured by the investment costs, incomes and IRR) of the solutions obtained when solving the GTSELSP for different scenarios of spot prices (i.e., P25, P50 and P75, respectively) and the different time resolutions. Surprisingly, we can see that regardless of the scenario, the solutions attained when using different time resolutions are quite similar in all economic dimensions (although, evidently, they differ when compared among scenarios).



(**a**) Percentile 25: pessimistic scenario



(**b**) Percentile 50: possible scenario



(**c**) Percentile 75: optimistic scenario

Figure 10. Comparison between economical indicators attained for different time resolution structures and scenarios. (**a**) Percentile 25; (**b**) Percentile 50; (**c**) Percentile 75.

From the computational difficulty point of view, in Figure 11, we compare the computing time (vertical axis, in seconds) required for solving the MILP instances induced by the different time resolution structures (horizontal axis) and by the different price scenarios (colored lines). In this case, the impact of the different resolution alternatives is clear: the resolution '12 × 7' induces resolutions times that are at least two orders of magnitude larger than those induced by '12 × 1'. Since the resolution times attained by '6 × 7' are reasonable for the considered decision-making context, we use this time resolution as part of our default setting.

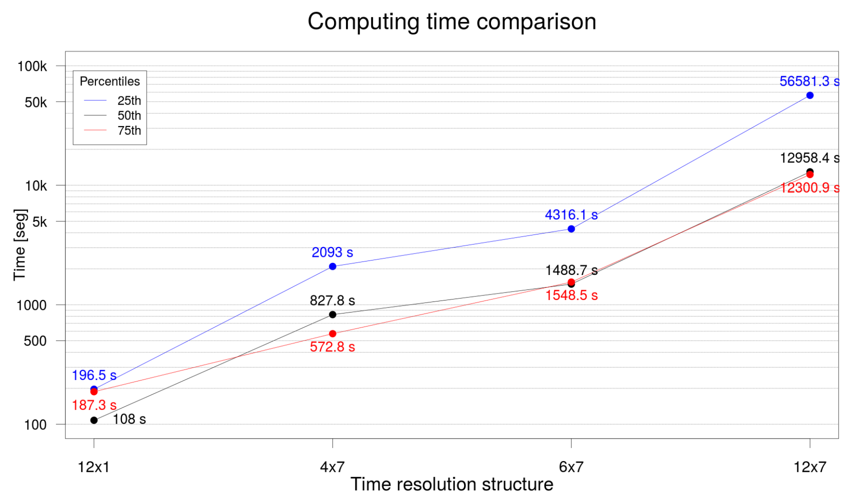


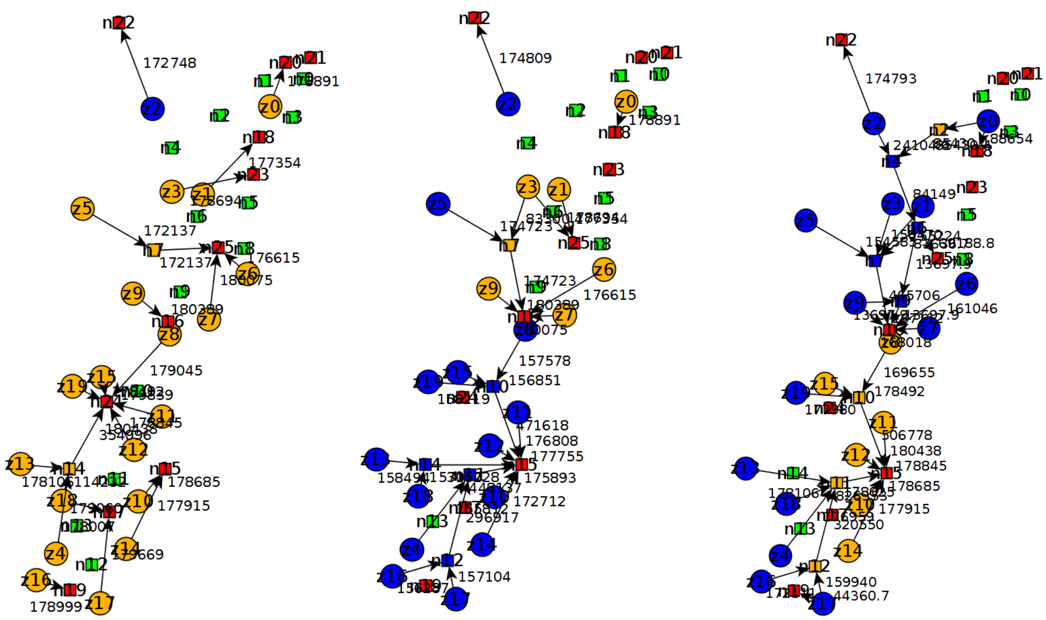
Figure 11. Computing time comparison for different time resolution structures.

4.3.3. Spot Price Scenarios

As shown in Figure 10a–c, different price scenarios induce different economical performances. While the IRR attained when considering P25 (pessimistic scenario) is near 8.0%, the one attained for P50 (possible scenario) is almost 14.0%, and the one attained for P75 (optimistic scenario) can be above 28.0%. Interestingly, different scenarios of spot prices are not associated with noticeable differences in the investment cost (whose annualized value is shown by red lines), and the differences in the economical performance mainly depend on the different levels of collected revenues (blue lines).

Although the investments do not fluctuate considerably according to the scenario, the infrastructure induced by these investments can be different, especially when comparing the one corresponding to P25 with respect to the one corresponding to P75 (i.e., the strategic design component of the GTSELSP solution differs when having different spot price realizations). In Figure 12, we report the solutions computed for these three scenarios (in the caption of each figure, we report total investment in MMUSD), the description of which is given as follows:

- The solution obtained for P25 (with a total investment of 4702 MMUSD) is comprised by 20 generation points (orange and blue circles), where an ESS facility is installed in only one of them (blue circle); these generation points inject the generated power through nine existing substations (red squares) and require one additional intermediate substation (orange square).
- The solution obtained for P50 (with a total investment of 5085 MMUSD) is comprised by 20 generation points, where ESS facilities are installed in 14 of them; these generation points inject the generated power through six existing substations and require five additional intermediate substations, and ESS facilities are installed in four of them (blue squares).
- The solution obtained for P75 (with a total investment of 5152 MMUSD) is comprised by 20 generation points, where ESS facilities are installed in 14 of them; these generation points inject the generated power through six existing substations and require eight additional intermediate substations, and ESS facilities are installed in four of them.



(**a**) pessimistic (4702 MMUSD) (**b**) possible (5085 MMUSD) (**c**) optimistic (5152 MMUSD)

Figure 12. Generation, Transmission, Storage Energy Location and Sizing Problem (GTSELSP) solutions when using PHS technology and considering three spot price scenarios (squares = substations ($V_2 \cup O$); circles = generation zones (V_1); green = unselected; orange = selected; red = sales node (O); blue = storage system). (**a**) pessimistic (4702 MMUSD; (**b**) possible (5085 MMUSD); (**c**) optimistic (5152 MMUSD).

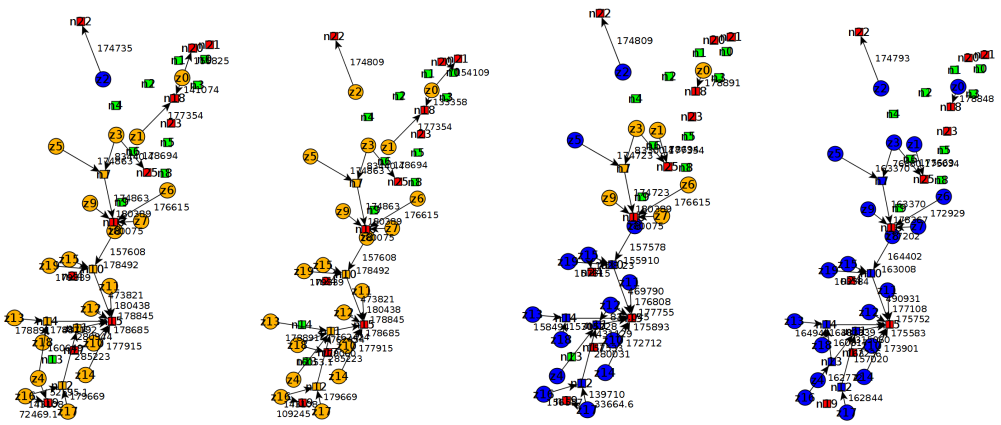
In light of the descriptions of the attained solutions and considering the values obtained for the IRR, the total incomes and the total investment costs, we consider that the scenario associated with P50 is a conservative way of capturing the spot price variability; hence, the possible scenario will be set as our default spot price realization.

Besides describing the effect of the spot price scenarios, one of the main outcomes of the analysis carried out is the important role that ESSs play within the designed solar power system. The solution shown in Figure 12b considers 1067 MWh of energy storage potential, while the generation potential is 4000 MWp. This solution shows that ESSs are an economically effective means of exploiting the generation from renewable sources, especially if they are incorporated within the system during its design stage.

4.3.4. Storage Technology

So far, we have considered that the ESS technology corresponds to PHS (which operates according to the parameters given in Table 2). However, decision-makers could be interested in knowing how the design and performance of the system changes when considering different storage technologies. For such a purpose, we solved the GTSELSP considering the four aforementioned storage systems; the produced solutions are shown in Figure 13. In the caption of each figure, we report in parenthesis the attained objective function value and the corresponding IRR. While the solutions obtained when considering ZEBRA and PBAC basically do not incorporate storage systems, the solutions associated with PHS and CAES incorporate ESS modules in many of the generation and intermediate points (totaling a storage capacity of 1067 and 733 MWh, respectively). Although from the objective function point of view, the PHS technology is associated with the best value (552 MMUSD), the results reported in Figure 13 make clear that recent technologies, such as ZEBRA, might boost the performance of power grids. The fact that no PBAC storage is installed should not be surprising, since this technology does not fit with the operational scheme of our design setting [34,35].

Additionally, the results reported in Figure 13, suggest that Φ^c and Φ^d are the most relevant parameters in the design strategy; both PHS and CAES allow larger charge and discharge flows (50 MW and 25 MW, respectively) than the ZEBRA and PBAC technologies (0.075 MW and 5 MW, respectively). This is explained by the fact that it is pointless to invest in a large capacity, i.e., a large value of S_{max}, when it is impossible to charge or discharge such capacity. As a matter of fact, when looking at Figures 13c,d, we see many small CAES systems and fewer, but larger, PHS systems.



(**a**) ZEBRA (546 MMUSD, 14.18%)
(**b**) PBAC (546 MMUSD, 14.18%)
(**c**) PHS (552 MMUSD, 13.88%)
(**d**) CAES (551 MMUSD, 14.21%)

Figure 13. Results for technological ESS comparison in P50 (squares = substations ($V_2 \cup O$); circles = generation zones (V_1); green = unselected; orange = selected; red = sales node (O); blue = storage system).

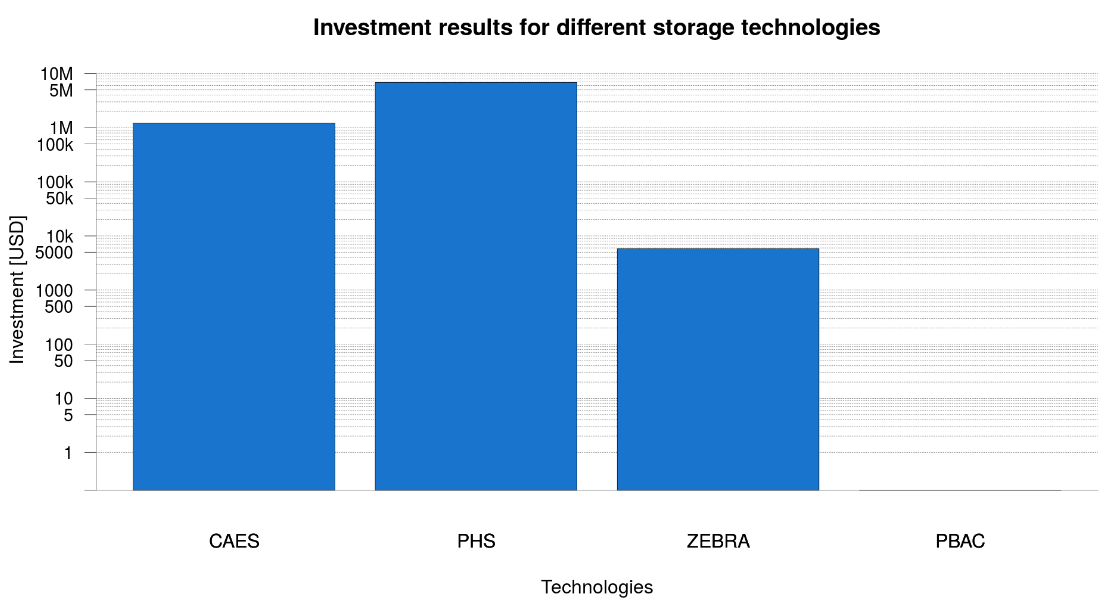


Figure 14. Investment results for the studied ESS technologies in P50.

The considered technologies do not only differ in their operation, but also from the investment point of view (see Table 2). As can be seen from Figure 14 (ESS installation cost associated with the solutions reported in Figure 13), the investment in PHS is five-times higher than the one for CAES, and about 1000-times higher than the one for ZEBRA. This reveals that although PHS is an expensive technology, it pays back due to its efficiency in storing energy.

4.3.5. Sensitivity Analysis on Storage Parameters

In light of the results presented above, we wanted to answer the following question: how sensitive are the ZEBRA storage installation decisions with respect to different values in the technology's operative parameters?. By answering such a question, we would have an idea of how the ZEBRA technology should evolve in order to be economically feasible. Hence, we performed a sensitivity analysis on four parameters: the lifetime (years), the installation cost (USD/kWh), the maximum storage capacity S_{max} (kWh) and the charge (Φ^c) and discharge rates (Φ^d) (kW). Using a ceteris paribus approach, we solved the GTSELSP for different values of the mentioned parameters and analyze how the installed storage (measured by capacity and total cost) changes as a consequence.

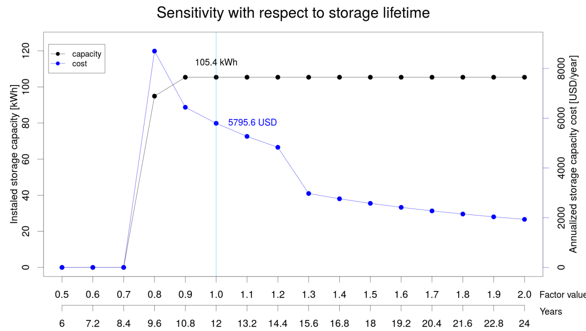
In Figure 15a, how the storage decisions would change if the ZEBRA lifetime goes from six (factor value = 0.5, i.e., half of the current lifetime of 12 years) to 24 years (factor value = 2, i.e., double the current lifetime of 12 years). Besides, the values associated with the current lifetime of 12 years are highlighted by a vertical line. From this plot, we can see that a lifetime of eight years makes the ZEBRA technology completely unpractical (installed capacity is 0 kWh). However, being able to extend the lifetime to double does not lead to any increase in the amount of installed capacity when compared to the original lifetime; only the present value of the total installation costs reduced significantly.

The sensitivity with respect to the installation cost is shown in Figure 15b, which is similar to Figure 15a. From this plot, we can see that an installation cost equal to or larger than 352 USD/KWh (1.6-times the original value of 220 USD/kWh) makes it economically impractical to use this technology. Nonetheless, even if the installation costs reduce to only half of the original value, i.e., falls to 110 USD/kWh, we would not install more capacity than that installed with the original value (highlighted by a vertical line).

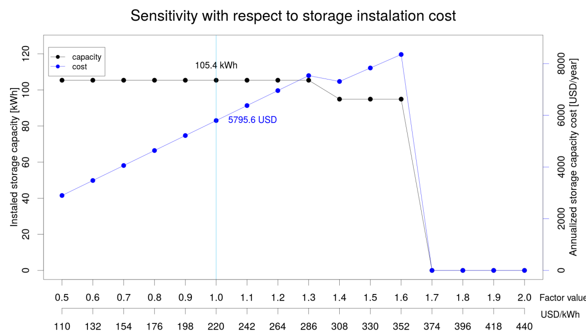
The impact of modifying the maximum storage capacity S_{max} can be analyzed from Figure 15d. In this case, it seems that the storage decisions are not sensitive to different values of S_{max}; even if it takes half or triple the original value, the storage capacity (and therefore, the cost) remains at 105.4 kWh. This result makes sense, since the original value of 0.3 MWh is already small.

Finally, we analyze how sensitive the storage decisions are with respect to the maximum charge and discharge flow. The plot shown in Figure 15d makes evident that these are the most influential parameters. There is a clear linear relation between the value of these parameters and the installed storage capacity. However, the increase in the installed capacity is bounded by $\Phi^c = \Phi^d = 300$ kW, which coincides with the nominal maximum storage capacity $S_{max} = 300$ kWh; therefore, having a larger charge and discharge flow capacity will not bring any change in the storage decision.

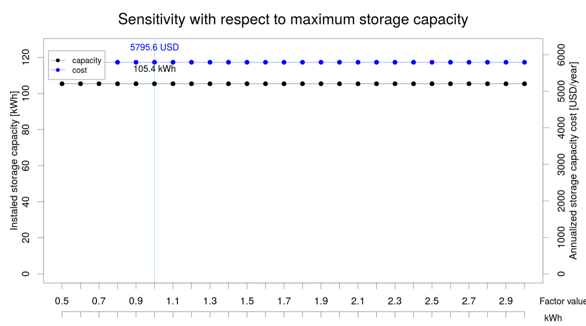
As a consequence of this sensitivity analysis, we can conclude that the key characteristics that comprise the viability of using ESSs within a large power system basically depend on allowing large amounts of stored energy (S_{max}) and the technological capacity of charging or discharging proportionally large amounts of energy (Φ^c and Φ^d, respectively). According to this conclusion, we could affirm that the ZEBRA technology is unlikely to respond, at its current stage of development, to the needs of the power system as those considered within the scope of this paper.



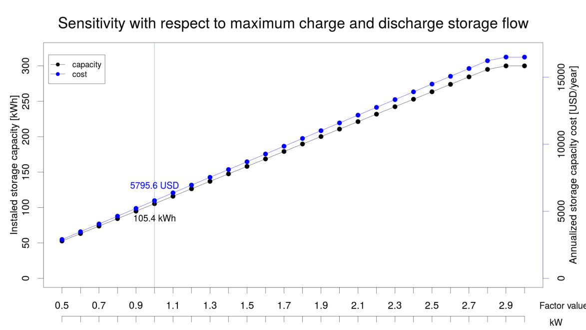
(**a**) Lifetime



(**b**) Cost



(**c**) S_{max}



(**d**) Φ^c, Φ^d

Figure 15. Sensitivity analysis results for four ESS parameters: (**a**) lifetime, (**b**) cost, (**c**) max. storage capacity and (**d**) charge/discharge storage flows. Lines represent the aggregated installed ESS cost and capacity for ZEBRA technology.

5. Conclusions and Future Work

One of the most crucial challenges that society faces today corresponds to evolving our current power systems towards a scenario where renewable energies play a predominant role. This challenge brings several technological and strategic issues that have attracted the attention of researchers from different areas. In particular, several works have focused on the development of decision support tools for designing and sizing power systems based (completely or partially) on renewable sources. In this sense, the incorporation of energy storage systems as part of the decision-making scope is fundamental, and it leads to an additional level of modeling difficulty due to the complex interplay among the generation, transmission and storage decisions.

In this paper, we have presented a methodological framework for a strategic decision-making setting involving the design of a solar power system embedded within an existing large power grid. The developed tool relies on an optimization problem coined as the generation, transmission, storage energy location and sizing problem (GTSELSP), which allows on to compute economically-attractive investments given by the location and sizing of the generation and storage energy systems, along with the corresponding layout of transmission lines. By means of a real case study (located in the central region of Chile), characterized by carefully-curated data, we showed the potential of the developed tool for aiding long-term investment planning. The obtained solutions show that decision-makers could exploit ESSs if their location and sizing are defined, simultaneously, with the generation and transmission planning. Additionally, we showed how the proposed tool could be used to have an accurate measure of the impact of different ESS technologies. Furthermore, we also demonstrated the possible use of the developed tool for investigating how sensitive the storage decisions are with respect to the different operative parameters of a given technology; this latter characteristic is particularly important for those researchers working on improving existing storage technologies.

An interesting path for future work would be to devise a more complex setting in which the design power system is evaluated by a proxy of its actual functioning regime; for instance, by solving a unit commitment problem that incorporates the designed power system. Such an approach is likely to lead to an iterative scheme that shall provide a much more accurate evaluation of the designed power system.

Acknowledgments: David Olave-Rojas and Eduardo Álvarez-Miranda acknowledge the support of the Chilean Council of Scientific and Technological Research (FONDECYT Grant No. 11140060 and Complex Engineering Systems Institute ICM:P-05-004-F/CONICYT:FB0816). The authors would like to thank the support of C. Wander and R. Fuentes-Ryks from Vivest Energías Renovables S.A.

Author Contributions: David Olave-Rojas contributed with the problem definition and formulation, performed most of data processing, computational implementation, experimental design, results analysis, and wrote most of the paper. Eduardo Álvarez-Miranda contributed with the problem definition and formulation, and partially contributed with the computational implementation, experimental design and results analysis. Alejandro Rodríguez performed the bootstrapping analysis of spot prices, and partially contributed with results analysis. Claudio Tenreiro partially contributed with the problem definition, the experimental design and results analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. International Energy Agency (IEA). *Key World Energy Statistics 2015;* Technical Report; International Energy Agency: Vienna, Austria, 2015.
2. Carlyle, R. How big are the currently known oil reserves and what are the chances of finding new ones?—Forbes. Available online: http://www.forbes.com/sites/quora/2013/03/27/how-big-are-the-currently-known-oil-reserves-and-what-are-the-chances-of-finding-new-ones/ (accessed on 21 July 2016).
3. Tully, A. How long will world's oil reserves last? 53 Years, Says BP. Available online: http://www.csmonitor.com/Environment/Energy-Voices/2014/0714/How-long-will-world-s-oil-reserves-last-53-years-says-BP (accessed on 21 July 2016).

4. Petroleum, B.B. Oil Reserves | About BP | BP Global. Available online: http://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy/oil/oil-reserves.html (accessed on 21 July 2016).
5. Roboam, X. (Ed.) *Integrated Design by Optimization of Electrical Energy Systems*; Wiley-ISTE: Hoboken, NJ, USA, 2012.
6. Suazo-Martínez, C.; Pereira-Bonvallet, E.; Palma-Behnke, R. A simulation framework for optimal energy storage sizing. *Energies* **2014**, *7*, 3033–3055.
7. Kim, J.; Powell, W. Optimal energy commitments with storage and intermittent supply. *Oper. Res.* **2011**, *59*, 1347–1360.
8. Koltsaklis, N.; Georgiadis, M. A multi-period, multi-regional generation expansion planning model incorporating unit commitment constraints. *Appl. Energy* **2015**, *158*, 310–331.
9. Jiang, R.; Wang, J.; Guan, Y. Robust unit commitment with wind power and pumped storage hydro. *IEEE Trans. Power Syst.* **2012**, *27*, 800–810.
10. Dehghan, S.; Amjady, N. Robust transmission and energy storage expansion planning in wind farm-integrated power systems considering transmission switching. *IEEE Trans. Sustain. Energy* **2016**, *7*, 765–774.
11. MacRae, C.; Ozlen, M.; Ernst, A.; Behrens, S. Locating and sizing energy storage systems for distribution feeder expansion planning. In Proceedings of the TENCON 2015—2015 IEEE Region 10 Conference, Macao, China, 1–4 November 2015; pp. 1–6.
12. Bradbury, K.; Pratson, L.; Patiño-Echeverri, D. Economic viability of energy storage systems based on price arbitrage potential in real-time US electricity markets. *Appl. Energy* **2014**, *114*, 512–519.
13. Berrada, A.; Loudiyi, K. Operation, sizing, and economic evaluation of storage for solar and wind power plants. *Renew. Sustain. Energy Rev.* **2016**, *59*, 1117–1129.
14. Xiong, P.; Singh, C. Optimal planning of storage in power systems integrated with wind power generation. *IEEE Trans. Sustain. Energy* **2016**, *7*, 232–240.
15. Rahmann, C.; Palma-Behnke, R. Optimal allocation of wind turbines by considering transmission security constraints and power system stability. *Energies* **2013**, *6*, 294–311.
16. Jakus, D.; Krstulovic, J.; Vasilj, J. Algorithm for optimal wind power plant capacity allocation in areas with limited transmission capacity. *Int. Trans. Electr. Energy Syst.* **2014**, *24*, 1505–1520.
17. Qi, W.; Liang, Y.; Shen, Z. Joint planning of energy storage and transmission for wind energy generation. *Oper. Res.* **2015**, *63*, 1280–1293.
18. Ghofrani, M.; Arabali, A.; Etezadi-Amoli, M.; Fadali, M. A framework for optimal placement of energy storage units within a power system with high wind penetration. *IEEE Trans. Sustain. Energy* **2013**, *4*, 434–442.
19. Oh, H. Optimal planning to include storage devices in power systems. *IEEE Trans. Power Syst.* **2011**, *26*, 1118–1128.
20. Jabr, R.; Džafić, I.; Pal, B. Robust optimization of storage investment on transmission networks. *IEEE Trans. Power Syst.* **2015**, *30*, 531–539.
21. Wang, H.; Huang, J. Hybrid renewable energy investment in microgrid. In Proceedings of the 2014 IEEE International Conference on Smart Grid Communications, Venice, Italy, 3–6 November 2014; pp. 602–607.
22. Wang, H.; Huang, J. Joint investment and operation of microgrid. *IEEE Trans. Smart Grid* **2017**, *8*, 833–845.
23. Wang, H.; Huang, J. Cooperative planning of renewable generations for interconnected microgrids. *IEEE Trans. Smart Grid* **2016**, *7*, 2486–2496.
24. Ciabattoni, L.; Ferracuti, F.; Grisostomi, M.; Ippoliti, G.; Longhi, S. Fuzzy logic based economical analysis of photovoltaic energy management. *Neurocomputing* **2015**, *170*, 296–305.
25. Ministerio de Energía de Chile. Explorador de Energía Solar. 2011. Available online: http://walker.dgf.uchile.cl/Explorador/Solar3/ (accessed on 03 May 2016). (In Spanish)
26. Fuentes-Ryks, R.; Wander, C. Technical Information. Personal communication. Vivest Energías Renovables S.A., 2016.
27. Coordinador Eléctrico Nacional. Información Técnica Centrales SING. 2016. Available online: http://cdec2.cdec-sing.cl/pls/portal/cdec.pck_web_coord_elec.sp_pagina?p_id=5187 (accessed on 20 May 2016). (In Spanish)

28. Dirección de Planificación y Desarrollo CDEC-SIC. Determinación de Puntos de Conexión STT; Technical Report; Coordinador Eléctrico Nacional: Santiago, Chile, 2015. Available online: https://sic.coordinadorelectrico.cl/wp-content/uploads/2015/07/Determinación-de-puntos-de-conexión-STT.pdf (accessed on 4 May 2016). (In Spanish)
29. ABB AB. *It's Time to Connect-HVDC Light©*, 7th ed.; Grid Systems-NVDC: Ludvika, Sweden, 2013.
30. European Network of Transmission System Operators for Electricity. *Offshore Transmission Technology*; Technical Report Prepared by the Regional Group North Sea for the NSCOGI; ENTSOE: Brussels, Belgium, 2011.
31. Lancheros, C. Transmission Systems for Offshore Wind Farms: A Technical, Enviromental and Economic Assessment. Master's Thesis, Hamburg University of Technology, Hamburg, Germany, 2013.
32. Fischetti, M.; Leth, J.; Borchersen, A. A Mixed-Integer Linear Programming approach to wind farm layout and inter-array cable routing. In Proceedings of the 2015 American Control Conference (ACC), Chicago, IL, USA, 1–3 July 2015; pp. 5907–5912.
33. Van Eeckhout, B.; Van Hertem, D.; Reza, M.; Srivastava, K.; Belmans, R. Economic comparison of VSC, HVDC and HVAC as transmission system for a 300 MW offshore wind farm. *Eur. Trans. Electr. Power* **2010**, *20*, 661–671.
34. Chen, H.; Cong, T.; Yang, W.; Tan, C.; Li, Y.; Ding, Y. Progress in electrical energy storage system: A critical review. *Prog. Nat. Sci.* **2009**, *19*, 291–312.
35. Cho, J.; Jeong, S.; Kim, Y. Commercial and research battery technologies for electrical energy storage applications. *Prog. Energy Combust. Sci.* **2015**, *48*, 84–101.
36. Coordinador Eléctrico Nacional. Programación Semanal CDEC SIC. 2016. Available online: https://sic.coordinadorelectrico.cl/informes-y-documentos/fichas/operacion-programada-2/ (accessed on 3 May 2016).
37. Pascual, L.; Romo, J.; Ruiz, E. Bootstrap predictive inference for ARIMA processes. *J. Time Ser. Anal.* **2004**, *25*, 449–465.
38. Pascual, L.; Romo, J.; Ruiz, E. Bootstrap prediction intervals for power-transformed time series. *Int. J. Forecast.* **2005**, *21*, 219–235.

Apéndice C

Towards a complex investment and evaluation framework for renewable energy systems: A 2-stage-heuristic approach

Contiene:

- Confirmación de envío a la revista
- Artículo enviado

Gmail

David Olave Rojas <d.olave.rojas@gmail.com>

Thank you for your submission to Applied Energy

1 message

Applied Energy <eesserver@eesmail.elsevier.com> Thu, Apr 4, 2019 at 12:52 PM
Reply-To: Applied Energy <appliedenergy@elsevier.com>
To: ealvarez@utalca.cl
Cc: d.olave.rojas@gmail.com

*** Automated email sent by the system ***

Dear Dr. Álvarez-Miranda,

Thank you for sending your manuscript Towards a complex investment evaluation framework for renewable energy systems: A 2-stage-heuristic approach for consideration to Applied Energy. Please accept this message as confirmation of your submission.

When should I expect to receive the Editor's decision?
We publicly share the average editorial times for Applied Energy to give you an indication of when you can expect to receive the Editor's decision. These can viewed here: http://journalinsights.elsevier.com/journals/0306-2619
review_speed

What happens next?
Here are the steps that you can expect as your manuscript progresses through the editorial process in the Elsevier Editorial System (EES).

1. First, your manuscript will be assigned to an Editor and you will be sent a unique reference number that you can use to track it throughout the process. During this stage, the status in EES will be "With Editor".

2. If your manuscript matches the scope and satisfies the criteria of Applied Energy, the Editor will identify and contact reviewers who are acknowledged experts in the field. Since peer-review is a voluntary service, it can take some time but please be assured that the Editor will regularly remind reviewers if they do not reply in a timely manner. During this stage, the status will appear as "Under Review".

Once the Editor has received the minimum number of expert reviews, the status will change to "Required Reviews Complete".

3. It is also possible that the Editor may decide that your manuscript does not meet the journal criteria or scope and that it should not be considered further. In this case, the Editor will immediately notify you that the manuscript has been rejected and may recommend a more suitable journal.

For a more detailed description of the editorial process, please see Paper Lifecycle from Submission to Publication:
http://help.elsevier.com/app/answers/detail/a_id/160/p/8045/

How can I track the progress of my submission?
You can track the status of your submission at any time at http://ees.elsevier.com/APEN

Once there, simply:
1. Enter your username: Your username is: ealvarez@utalca.cl

If you need to retrieve password details, please go to: http://ees.elsevier.com/APEN/automail_query.asp

2. Click on [Author Login]. This will take you to the Author Main Menu
3. Click on [Submissions Being Processed]

Many thanks again for your interest in Applied Energy.

Kind regards,

Professor Jinyue Yan

If you require further assistance, you are welcome to contact our Researcher Support team 24/7 by live chat and email or 24/5 by phone: http://support.elsevier.com

Gmail

David Olave Rojas <d.olave.rojas@gmail.com>

Track your recent Co-Authored submission to APEN

1 message

Applied Energy <eesserver@eesmail.elsevier.com> Thu, Apr 4, 2019 at 12:52 PM
Reply-To: Applied Energy <appliedenergy@elsevier.com>
To: d.olave.rojas@gmail.com

*** Automated email sent by the system ***

Dear Dr. David Olave-Rojas,

You have been listed as a Co-Author of the following submission:

Journal: Applied Energy
Title: Towards a complex investment evaluation framework for renewable energy systems: A 2-stage-heuristic approach
Corresponding Author: Eduardo Álvarez-Miranda
Co-Authors: David Olave-Rojas;

To be kept informed of the status of your submission, register or log in (if you already have an Elsevier profile).

Register here: https://ees.elsevier.com/apen/default.asp?acw=&pg=preRegistration.asp&user=coauthor&fname=David&lname=Olave-Rojas&email=d.olave.rojas@gmail.com

Or log in: https://ees.elsevier.com/apen/default.asp?acw=&pg=login.asp&email=d.olave.rojas@gmail.com

If you did not co-author this submission, please do not follow the above link but instead contact the Corresponding Author of this submission at ealvarez@utalca.cl.

Thank you,

Applied Energy

Towards a complex investment evaluation framework for renewable energy systems: A 2-stage-heuristic approach

David Olave-Rojas[a], Eduardo Alvarez-Miranda[b,*]

[a] *DSc Program on Complex Engineering Systems, Universidad de Talca, Curicó, Chile*
[b] *Department of Industrial Engineering, Universidad de Talca, Curicó, Chile*

Abstract

The main challenge for humanity is developing an economically, socially and environmentally sustainable lifestyle. One of the main actions to achieve this is the generation and use of energy in a sustainable way. In this context, non-conventional renewable energies seem to be the only possible solution; but they still present technical and operational challenges, and their financial competitiveness is still an important issue. These challenges can be faced through energy storage systems, but their incorporation requires the development of decision support tools in the design and planning stages of power systems. In this work, we introduce a decision aid framework based on an optimization problem that supports the investment decision making process in energy projects. This framework consists of two problems that are solved iteratively in a 2-stage scheme. The results, based on a Chilean case study characterized by curated data, show an improvement in the economic evaluation corresponding to the annualized return on investment of 32.76% and a 31.47% on the internal rate of return (IRR) compared to the result without the presented framework. We also present the application of the methodology in other scenarios where budget and subsidy conditions are applied. In the first case, we discussed the effects of splitting a project into stages according to a certain budget. While in the second case, we determined 15% as an ideal subsidy for improving the economical results of our case study.

Keywords: Renewable energies, Energy storage systems, Power system planning, Mathematical optimization, Decision aid tools

1. Introduction and Motivation

The greatest challenge that humankind faces today is the design and implementation of effective policies and radical cultural changes towards an environmentally and economically sustainable lifestyle (see United Nations, 2016, where the United Nations (UN) *Sustainable Development Goals* are presented). Climate change is the main driver of this challenge (see UN Sustainable Development Goals, 2016a), and the sustainable generation and consumption of energy is one of the fundamental tasks for addressing it (see, e.g., Armaroli & Balzani, 2007; Dorian et al., 2006; UN Sustainable Development Goals, 2016b). Because of this critical situation, over the last decades there has been a rapid growth in the development of technologies associated with (non conventional) renewable energy sources (see, e.g., Dincer, 2000; Owusu & Asumadu-Sarkodie, 2016; World Future Council, 2016, for reviewes dealing with this issue over the last 20 years). This prominent development has lead, and has been benefit from, an accelerated reduction of investment costs (due to technological advances), and environmentally-driven regulations forcing the reduction of fossil fuels consumption, who are responsible for most of greenhouse gas emissions (Carley, 2009).

Although non conventional renewable sources seem to be the only alternative for a sustainable generation of energy, their use results in economical, technological and managerial issues, which take place when designing, planning and operating power systems with penetration of renewable sources. Evidently, in a given power system, the higher the penetration of these sources, the stronger the impact of these issues. Nowadays, massive (or "grid") energy storage systems (ESS) are emerging as a viable alternative

*Corresponding author
Email addresses: d.olave.rojas@gmail.com (David Olave-Rojas), ealvarez@utalca.cl (Eduardo Alvarez-Miranda)

to cope with them. As a matter of fact, ESS technologies can address most of these concerns and facilitate power balancing as they (i) decouple the time of generation and consumption, and (ii) can be used as part of frequency and shortage control systems (see, e.g., Kim & Powell, 2011; Harsha & Dahleh, 2015). Nonetheless, ESS are still costly solutions, with quite strict operative patterns, so their incorporation within large-scale power grids does not only associate technological challenges but also calls for decision-making tools that ensure appropriate design and sizing of the corresponding power systems (see, e.g., Chen et al., 2009; Luo et al., 2015; Zakeri & Syri, 2015, as examples of area analysis and reviews).

Power systems typically encompass a physical network of electrical devices, such as power plants (eventually, some of them based on non-conventional renewable sources), transformers, and transmission and distribution lines, deployed within a given territory. Such systems allow generating, distributing and supplying electricity to final (domiciliary and industrial) customers; in the classical framework, power systems were designed with the goal of meeting, typically from a centralized perspective, the electricity demand over a long-term time horizon. Due to their critical importance, the managerial aspects of the design and expansion of such systems have attracted the attention of researchers for decades; as a matter of fact, optimization approaches to power system planning have been proposed since the 70s (see, e.g., Oatman & Hamant, 1973). Currently, power systems (i) are not necessary centralized, as distributed generation is becoming more common with the increasing penetration of renewables (see, e.g., Ackermann et al., 2001; Mehigan et al., 2018); (ii) their design and operation is constrained by market projections, legal regulations, climatological and orography conditions, and technological advances and challenges, among other factors; (iii) and their performance is measured by economical, sustainability and quality-of-service criteria. Hence, the design and expansion of power systems associates an intricate strategical decision making process (we refer the reader to Momoh & Mili, 2009; Seifi & Sepasian, 2011; Koltsaklis & Georgiadis, 2015; Guerra et al., 2016; Chen et al., 2016; Caunhye & Cardin, 2018, for textbooks and papers presenting an overview of power systems planning methodologies and discussing the impact of renewable energies).

Complementary, the (daily) operation of these systems results from another decision making process, which decides on how and where energy will be produced and distributed to meet the demand. These decisions are commonly made by solving the so-called unit commitment problem (UCP). Briefly speaking, the UCP defines, usually on a 24 or 48 hours basis, the production scheduling of the generators and other devises, i.e., the *units*, taking into account technical limitations of these units, the spot market conditions, and the customers short-term demand. Due to its importance, the UCP has been exhaustively studied from a practical and theoretical point of view; a countless number of optimization and simulation model variants as well as algorithmic strategies have been published over 50 years since the mid sixties (see, e.g., Kennedy & Mabuce, 1965; Kerr et al., 1966; Lowery, 1966, as examples of the earliest works dealing with the UCP).

Nowadays, the high penetration of energy renewable sources, have lead to electrical and electronic engineers, applied mathematicians and computer scientists, among others, to devote efforts to provide decision aid tools for supporting decisions at strategical, tactical and operative levels. In the following section, an overview of some these tools will be presented; as it will be shown, most of these works typically focus on a single level of the decision process. Hence, they do not take into account the fact that power grid components work collectively and simultaneously, failing to respond to sustainability, quality-of-service and financial expectations of policy makers, customers and investors, respectively.

Contribution and paper outline. The main contribution of our paper is the design and implementation of a novel optimization framework, based on mixed integer programming, for supporting strategical decision-making in (renewable-based) complex energy systems planning. The proposed framework relies on solving an optimization problem, coined as the *Generation, transmission and storage location and sizing of operation-aware sustainable power system design problem* (GTSOPSD). The GTSOPSD encodes two nested problems that are solved, iteratively, through a two-stage-like scheme; one corresponds to a strategic power system design problem, while the other corresponds to an embedded UCP instance.

Using a case study obtained from the Chilean power system, we show that the proposed framework is capable of exploiting the nature of the electricity market as well as the potential of territorial differences, when deciding (i) location and size of renewable energy plants (photo-voltaic plants in this particular case), (ii) location and size of ESSs, and (iii) the layout of the power grid expansion.

Furthermore, the GTSOPSD can be used to investigate how high (public) subsidies on renewable technologies should be in order to have an economically sustainable high penetration of these technologies. This type of analysis is crucial for supporting public policies targeting to increase the commercial competitiviness of more sustainable projects. As it will be shown, we used the proposed tool to find recomendable levels of subsidy in order to balance the financial performance of the project and the public investment.

The paper is organized as follows. A literature review on related applications is presented in Section 2. The optimization model that embodies our modeling framework is presented in Section 3. In Section 4, we present the results obtained when applying the designed optimization model to our case study. Finally, conclusions and paths for future work are drawn in Section 5.

2. Literature Review

As explained in the previous section, incorporating non-conventional renewable energy sources in power systems is both, mandatory and challenging. And, therefore, power systems planning demands for sophisticated optimization tools for effectively and efficiently coping with such decision-making challenge; specially, at a strategical level (see Roboam, 2012, for a general reference on optimization for electrical energy systems design). In this section, we review some of the most recent and relevant examples associated to our planning setting.

Our review focused on articles proposing optimization tools encompassing decisions, in different levels, on (i) ESSs, (ii) renewable-based generation, (iii) transmission infrastructure, and/or (iv) grid operation (typically embodied by the UCP). Note that, in most cases, grid operation decisions are embedded into planning optimization models in order to capture the operational effects of the design decisions. The review process led to 23 articles, most of them published over the last 6 years, satisfying these four criteria, which are presented in Table 1. With respect to ESSs, we recognize decisions with respect to "capacity", "location", and others such as evaluation of operational profiles ("Op.") or technological assessment ("Tech."); an equivalent classification is considered for generation decisions. In the case of transmission infrastructure decisions, papers are classified according to three decision levels; layout design ("layout"), lines' capacity ("capacity"), and power flow ("flow"). Finally, in the case of grid operation decisions, we identified whether these decisions are or not considered as part of the design process.

Reference	Storage			Generation			Transmission			Grid
	capacity	location	other	capacity	location	other	layout	capacity	flow	operation
(Oudalov et al., 2007)	✓		Op.							
(Kim & Powell, 2011)			Op.			Op.				✓
(Oh, 2011)		✓	Op.						✓	
(Dominguez et al., 2012)			Op.			Op.				
(Jiang et al., 2012)			Op.							✓
(Ghofrani et al., 2013)		✓							✓	
(Rahmann & Palma-Behnke, 2013)					✓	Tech.				
(Bradbury et al., 2014)	✓		Tech.							
(Muche, 2014)			Op.							✓
(Suazo-Martínez et al., 2014)	✓									✓
(Chang & Lin, 2015)	✓		Op.	✓		Op.			✓	
(Jabr et al., 2015)	✓	✓							✓	
(Koltsaklis et al., 2015)	✓		Tech.	✓		Tech.		✓		
(MacRae et al., 2015)	✓	✓	Op.				✓			
(Qi et al., 2015)	✓	✓					✓	✓		
(Xia et al., 2015)	✓	✓	Op.							
(Berrada & Loudiyi, 2016)	✓		Op. & Tech.							
(Dehghan & Amjady, 2016)	✓	✓	Op.				✓			
(Guerra et al., 2016)				✓	✓	Op.	✓	✓	✓	✓
(Xiong & Singh, 2016)	✓	✓	Op.							
(Lan et al., 2017)	✓	✓	Op.			Op.				✓
(Alsaidan et al., 2018)	✓	✓	Op.							✓
(Haas et al., 2019)	✓	✓		✓	✓		✓	✓		
our work (GTSOPSD)	✓	✓	Op. & Tech.	✓	✓	Connect.	✓	✓	✓	✓

Table 1: Summary of recent literature providing optimization tools for energy systems planning incorporating: ESSs, renewable and conventional generation, transmission, and/or grid operation.

The information displayed in Table 1 shows that, over the last years, there has been a systematic effort to integrate ESSs optimization decisions along with long-term planning decisions, such as a generation capacity and location and transmission layout. Furthermore, operational decisions, such as transmission flow and grid operation, have been also incorporated in order to better measure how the planning decisions behave. For example, in (MacRae et al., 2015), the authors provide an optimization model that seeks for an optimal location and sizing of a storage system along with an optimal operation regime, and the optimal design of the transmission lines for coupling it with the existing grid. However, other papers, such as (Ghofrani et al., 2013), only consider ESSs location and transmission decisions within an existing power system, without taking into account any further design element, As well as these two cases, each of the articles summarized in Table 1 address, partially, the features of the decision making setting presented in the introduction; however, none of them offers a general optimization framework integrating all of the decisions in a consolidated manner. The present work, characterized in the last line of the table, offers an integrated decision-making framework that closes this methodological gap. Furthermore, by considering the investors perspective, the role of public subsidies, and the territorial diversity offered by the different location alternatives, our framework seeks for economically attractive plans for designing or expanding energy systems by optimizing the economical value of the sought investment strategies.

3. An optimization framework for the GTSOPSD

The core of our energy system planning tool corresponds to an optimization framework for solving the previously introduced GTSOPSD. As explained in the introduction, the GTSOPSD exploits the relationship between two (integer programming) problems: the generation, transmission, energy storage location and sizing problem (GTSELSP), presented in Olave-Rojas et al. (2017), and the well-known UCP. While the GTSELSP involves technological and economical considerations related to strategical decisions, the UCP involves economical considerations associated to operational decisions. As we will show in this section, by iteratively solving these two problems in a master and slave fashion, it is possible to converge to a design strategy that effectively responds to economical, technological, operational and sustainability criteria.

3.1. Modeling the GTSELSP and the UCP

The GTSELSP. As explained above, the GTSELSP is an optimization problem whose aim is to find an investment strategy for a renewable-based energy generation project, that could be incorporated into an existing power system. The strategy involves energy storage systems, renewable energy power plants and transmission infrastructure (lines and substations); the goal is to find a maximum profit strategy for a given planning horizon. Following Olave-Rojas et al. (2017), the profit can be expressed as:

$$profit = I - C_{PV} - C_{trans} - C_{ESS} - C_{S/E}, \tag{1}$$

where I represents present value of the revenues obtained from selling the produced energy over a planning horizon (according to the market spot prices), C_{PV} is the generation installation cost, C_{trans} corresponds to the transmission layout construction cost, C_{ESS} is the cost associated to the infrastructure of energy storage systems, and $C_{S/E}$ represents the total substations installation cost. Conceptually, the GTSELSP is encoded by

$$\begin{aligned}
\text{(GTSELSP)} \qquad & profit^* = \max profit \\
\text{s.t} \quad & - \text{generation constraints} \\
& - \text{transmission topology constraints} \\
& - \text{energy storage systems constraints} \\
& - \text{power balance constraints} \\
& - \text{transmission and substation capacities constraints} \\
& - \text{losses and boundary conditions constraints.}
\end{aligned}$$

In Olave-Rojas et al. (2017) the GTSELSP formulated as a MIP model; nonetheless, for ease of exposition we do not provide the details of the formulation, and only present its conceptual structure. The GTSELSP can be regarded as a two-stage optimization problem; investment decisions (power plants and transmission infrastructure) are made in a first stage, while generation and power flow decisions, whose modeling purpose is to provide an estimation of the economical performance of the investment decisions, are made in a second stage. Therefore, the solution of the GTSELSP must satisfy technological and operational constraints, respectively.

The UCP. Conceptually, a solution of the UCP corresponds to a one-day ahead generation scheduling of the power plants (units) of a given power system; therefore, the UCP defines the units that will operate (the *commitment*) and how much energy they will generate (the *dispatching*). The goal is to find a minimum cost scheduling that satisfies the corresponding demand, while respecting a set of technological limitations and operational requirements. In its more common form, the cost of a given UCP solution is given by

$$cost = C_{start} + C_{shutdown} + C_{gen}, \tag{2}$$

where, C_{start} corresponds to the total cost of turning the power plants on, $C_{shutdown}$ corresponds to the total cost of shutting the power plants off, and C_{gen} is the total cost of producing energy.

The UCP is usually structured as

$$\begin{aligned} \text{(UCP)} \qquad & cost^* = \min costs \\ & \text{s.t} \quad - \text{commitment constraints} \\ & \qquad - \text{ramping constraints} \\ & \qquad - \text{primary and production reserves constraints} \\ & \qquad - \text{demand constraints} \\ & \qquad - \text{boundary conditions constraints,} \end{aligned}$$

where the commitment and ramping constraints respond to technological conditions, the reserves and demand constraints correspond to operational requirements, and boundary conditions might model further requirements (e.g., CO_2 emissions limits, further security constraints, among other). As well as the GTSELSP, the UCP also falls within the class of two-stage optimization problems; in this case, commitment decisions are made in a first stage, and dispatching decisions are made in a second stage. The UCP and its variants have been usually formulated as a MIP model (we refer the reader to Bhardwaj et al., 2012a,b; Saravanan et al., 2013; Zheng et al., 2015, for comprehensive reviews on modeling and algorithmic aspects of the UCP).

3.2. Algorithmic scheme for the GTSOPSD

As explained above, the GTSOPSD combines the GTSELSP with the UCP in a two-stage framework; the GTSELSP corresponds to the (master) first stage problem, while the UCP corresponds to the (slave) second stage problem. The goal is to design a renewable-based power system expansion that performs well when evaluated, on a given planning horizon, using an operational-based metric (as the UCP allows to).

Roughly speaking, a solution of GTSELSP corresponds to an expansion infrastructure of a power system (generation, storage and transmission capacity). The profit achieved by this infrastructure is defined by the present value of the total revenues associated to is operation (into an existing power system) over a planning horizon. This operation can be estimated by scaling UCP solutions to an adequate planning horizon. Furthermore, these UCP solutions allow to map-back an estimation of energy sale spot prices, that can be then used to solve the GTSELSP again. This GTSELSP-UCP-GTSELSP dynamic repeats until a convergence criterion is satisfied. Convergence is an important issue because while the GTSELSP solution (investment-oriented) seeks for the higher spot prices, the UCP solution (operator-oriented) seeks for the lowest possible spot prices.

In Figure 1 we present a diagram representing the inter-dependence between the GTSELSP and UCP and how a heuristic can be designed exploiting this relation. The resulting algorithm is characterized by the following steps;

Step 1 Solve the GTSELSP using an initial series of spot prices;

Step 2 Incorporate the infrastructure induced by the current GTSELSP solution into the UCP input data (i.e., update the existing power system) and solve the resulting UCP instance;

Step 3 Update the energy sales spot prices using the attained UCP solution;
Step 4 Solve the GTSELSP model using the current energy sale spot price series.
Step 5 If the stopping criterion is verified, then resume, otherwise go to Step 2.

The computational performance of the previously outlined algorithm strongly relies on the technique used for solving the GTSELSP and UCP instances at each iteration; as it will be shown in Results and Discussion section, these two problems were solved using a tunned version of an off-the-shelf MIP solver.

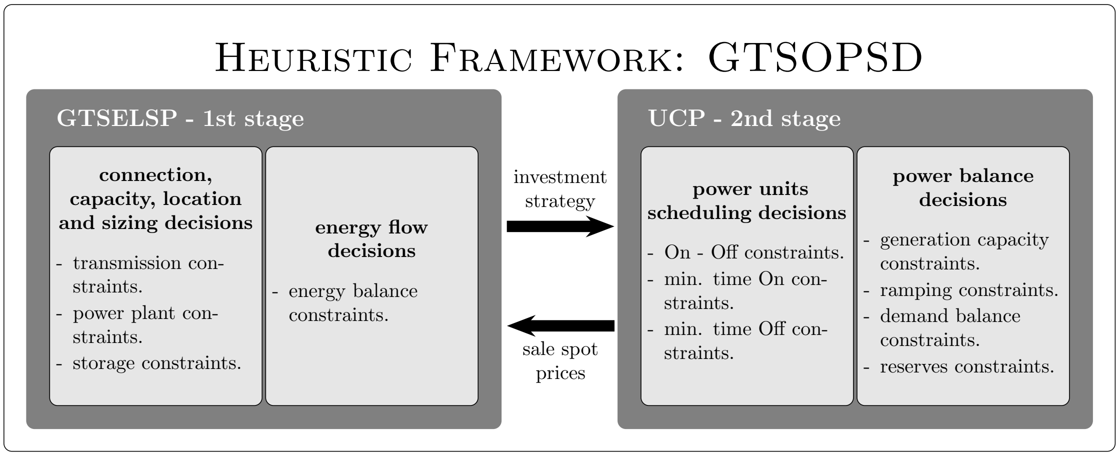


Figure 1: Diagram representing the presented heuristic framework.

Further considerations. Since the locations of the potential power plants could be scattered over a large territory, it might occur that the infrastructure to be installed (i.e., the solution associated to the GTSELSP) is comprised not by single but by several *sub*-grids (each of them encompassing power plants, ESSs, transmission lines) connected (through a sub-station) to the existing power system. Since the energy price at each of these connecting sub-stations might be different, the operation of the GTSELSP solution cannot be approximated, directly, by the standard UCP. Instead, we must transform the input data in order to take into account the territorial deployment and constraints associated to the existing power system. In our case study (described in Section 4), we first analyzed the historical data of the selling prices within the studied area; the records corresponded to hourly prices over a couple years. Afterwards, we processed the data and constructed probability density functions at each possible connection point (see Figure 4); using these profiles were merged into one using a normalization based on demand and price.

4. Results and Discussion

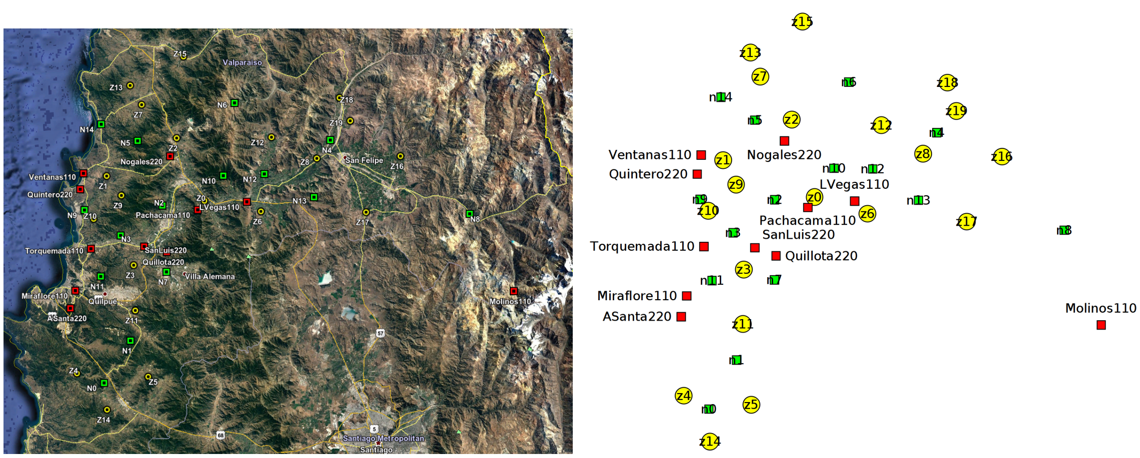
In this section we will present and discuss the results obtained when applying the proposed approach to a case study. The goal is finding an investment strategy for a large-scale (multi-site) PV-based power infrastructure. The section is organized as follows; we first describe the case study, later we outline the data processing and analysis, and finally we present and discuss the obtained results.

4.1. Case Study: Chilean North Central Region

Chile is one of the most attractive countries for solar energy projects due its great solar potential; specifically, in the northern region, the average radiation is as high as 2055 [kWh/m^2], the average annual sunshine time is 4000 hours, and there exists vast areas of flat land. However, the main energy demand are located in the central area, where the radiation potential has fluctuations and the orography is considerably more complex. Moreover, since affordable and clean energy is one of the United Nations' sustainable development goals (UN Sustainable Development Goals, 2016b), local solutions, renewable energy-based projects and energy efficiency approaches are gaining more importance. Therefore, it becomes a need the exploration of potential new solar projects, in such a way that these projects should be located near to demand areas, in order to enhance the territorial sustainability and improve life quality.

Some characteristics of the north central region play an important role in our decision setting. These are the following: (i) the average solar radiation is above 1478 [kWh/m^2], which is still higher than the radiation of countries as Germany (1261 [kWh/m^2]), with high penetration of solar generation; (ii) the portion of Chilean population in this area is around 56.00%, implying two of the three largest urban areas and representing a significant demand, but also less space for large energy projects; (iii) in addition to the domestic demand, there is also a massive need of electricity from industrial economical areas such as agriculture, retail, and activities related to the main Chilean port; (iv) the orography in this sector is more complex than in the northern region, which incorporates another level of complexity to the design process; (v) there is a significant quantity of thermal power plants, which use diesel, coal or gas, emitting greenhouse gases.

In this context, the studied territory corresponds to the 5th region (see Figure 2(a)). This zone satisfies two important requirements; the area contains the main port of the country, as it is located alongside the metropolitan area (where most of the demand is concentrated) as well, and there exists a quite complex power grid, where our designed power system could be connected to.



(a) Satellite view

(b) Schematic view

Figure 2: Graphical representation of studied territorial configuration. We use squares related to substations; circles to represent generation zones; green colour indicating possible substations and red colour for existing substations.

In Figure 2(b), we display a representation of the considered case study. Both Figures 2(a) and 2(b) contain the 20 areas that we have identified as possible locations of a generation plant; these areas are represented by yellow circles. Green squares represent locations for potential substations. In these 15 areas is also possible to locate an ESS if a substation is installed. The substations of the current power grid correspond to the 11 red squares; the generated energy of our designed power system is injected into these substations.

4.2. Estimation and Model parameters

The input data of the optimization problems encoding the GTSOPSD can be divided into generation and transmission technological parameters, generation and transmission construction costs, ESSs parameters, renewable energy source potential (solar radiation in this case), and market behavior (demand and energy prices). We used data curated by Kim et al. (2018); Olave-Rojas et al. (2017) to obtain the values corresponding to the first three groups of parameters, while for the later two sets we implemented statistical analysis methods which are outlined below.

Solar radiation potential. The solar radiation potential of the 20 generation zones was estimated using the data published by the Chilean Ministry of Energy Ministerio de Energía de Chile (2011), which was gathered and processed according to the methodology proposed in Molina et al. (2017). In Figure 3, we show the radiation curves (W/m^2) over a day for the 20 zones and for four days (each of them from a different season). Although the 20 zones belong to a relatively small region, there are notorious differences in the solar radiation potential profile during the day, specially in the case of autumn and spring (May and November, respectively). These behavior demonstrates the complexity of this type of decisions contexts, and the importance of taking into account the territorial dimension when planning a power system expansion based on renewables.

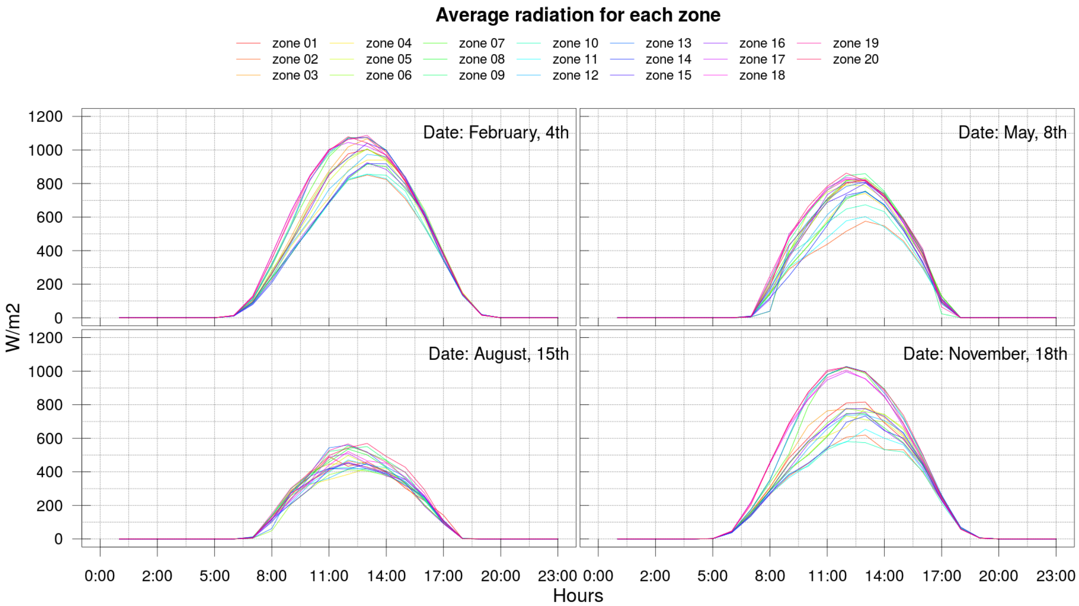


Figure 3: Representation of available average radiation for each studied zones. Is possible to see 4 selected days corresponding to the 4 year seasons.

Modeling and estimation of spot prices and demand. Spot prices and demand are obtained by processing the raw data from the Economic Dispatch Load Center of Central Interconnected System (CDEC-SIC) Coordinador Eléctrico Nacional (2016), following the methodology designed in Kim et al. (2018); Olave-Rojas et al. (2017).

In Figure 4, we present the behavior of spot prices series in ASanta220 (one of the 11 points where the generation plants can be connected to the system) in four different days and at four different hours of the day. These series were obtained using an ad-hoc adaptation of a bootstrapping technique (Pascual et al., 2004, 2005), according to the corresponding test, these spot prices follow, empirically, a bi-modal probability distribution. For a given day and a given hour, we mark three points in the corresponding curve; the 25-th percentile, the 50-th percentile and the 75-th percentile. Considering this, we can see that spot prices are typically between 25 and 75 USD per MWh (first mode), although prices between 125 and 175 USD per MWh are also relatively frequent (second mode); the first mode is associated with normal operation while the second mode is associated with disruptions and rare phenomena. This behavior is verified by the other 10 selling points.

Based on these distributions, we generated different series of spot prices which were then mapped into aggregated series according to a given time resolution. These estimated prices are used in the first iteration of the proposed heuristic, i.e., when solving the GTSELSP. Using spot prices based on historical data ensures to capture and propagate the real conditions of the energy market along the optimization process.

Complementary, in Figure 5 we display 6 curves characterizing the hourly (total) energy demand for each of the selected four days; these curves correspond to the minimum demand (*min*), the 25th-percentile demand (*25th pctl*), the 50th-percentile demand (*50th pctl*), the 75-th percentile demand (*75th pctl*), the maximum demand (*max*) and the expected demand (*E(X)*). These curves demonstrate that energy demand presents a similar pattern regardless of the season, specially in the interval 00:00 to 17:00;

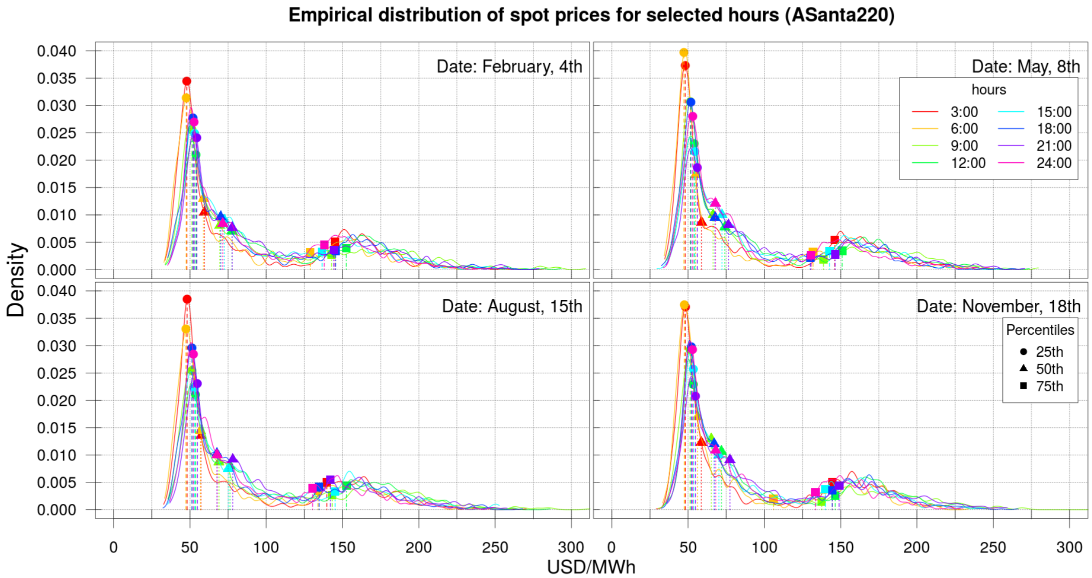


Figure 4: Empirical density distribution spot prices for ASanta220.

notwithstanding, in the interval 17:00 to 24:00, demand becomes much more volatile, as can be seen when comparing, in this interval, the minimum and maximum demand.

4.3. Discussion and Results Analysis

In Section 4.2, we have described where the necessary information to configure our model comes from in order to solve the problem raised in the case study. In addition, we have commented on some important characteristics of the data that model the area to be studied. In the remainder of this section, we present and analyze the obtained results after the application of our tool.

Recall that the objective value of the GTSELSP is the annualized profit, measured in millions of USD (MMUSD) per year. Additionally, we also present the annualized investment costs, measured in millions of USD per year, along with the corresponding internal rate of return (IRR) expressed as a percentage, which corresponds to the project's profit according to the evaluation planning horizon (in this case, 30 years). Likewise, the objective value of UCP is the equivalent annual operating cost, measured in millions of USD per year.

Experimental setting. We run our experiments in a Intel® Core™ i7-4702MQ 2.20GHz machine with 16GB RAM, and Ubuntu 16.04 LTS. The underlying GTSELSP instances were solved using ILOG® CPLEX® 12.7.0; we used default setting except for the primal-dual gap for GTSELSP which was set to 0.3% and 2.0% for UCP. Finally, for our heuristic we used a gap between iterations of 1.3 MMUSD. Since some early results shown a investment between 400 and 200 MMUSD for different configurations, the gap between iterations represent 0.3 – 0.6 %.

Time resolution. The first part of our computational analysis consists of determining which is the best time resolution for running the whole set of experiments. Also determining if a different area means a different behavior from a computational difficulty point of view. For such purpose, we considered three time resolution structures presented also in Olave-Rojas et al. (2017):

- '12x1': 1 year is characterized by 12 days, so that each day encodes 1 month. Since every day is comprised by 24 hours, this setting yields $T = \{1, 2, \ldots, 24 \times 12 = 288\}$.
- '4x7': 1 year is characterized 4 weeks, so that each week encodes three months. In this case, we have $T = \{1, 2, \ldots, 4 \times 7 \times 24 = 672\}$.

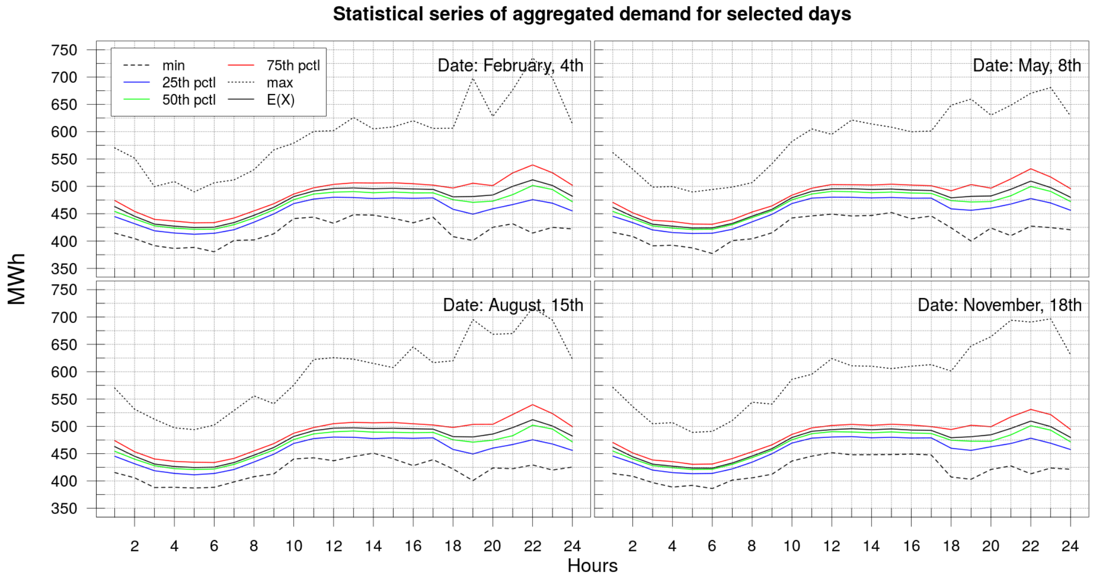


Figure 5: Statistical series of aggregated demand for four selected days.

- '6x7': 1 year is characterized 6 weeks, so that each week encodes two months. In this case, we have $T = \{1, 2, \ldots, 6 \times 7 \times 24 = 1008\}$.

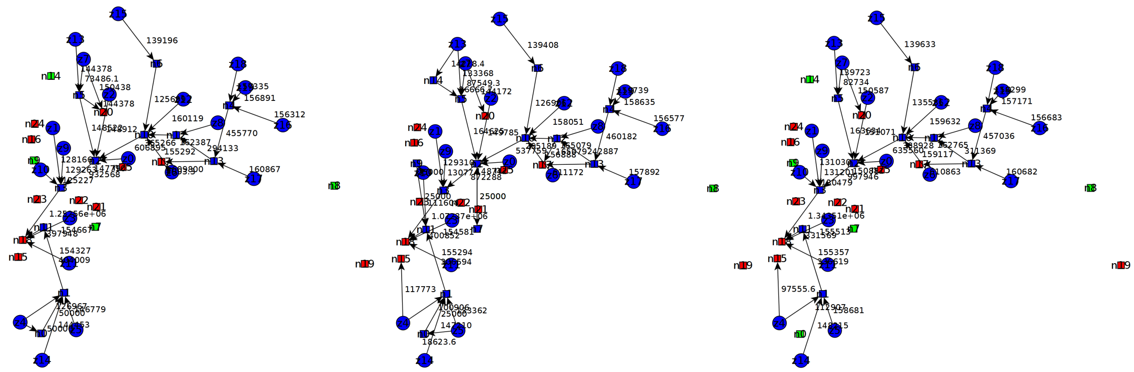
Instance	Obj. Value		Iterations	Gap (avg.)		Computing Time (avg.)	
	GTSELSP	UCP		GTSELSP	UCP	GTSELSP	UCP
12x1	276.252 MMUSD	116.881 MMUSD	6	0.26%	0.30%	276.94 s	83.05 s
4x7	276.396 MMUSD	114.504 MMUSD	31	0.26%	0.24%	1893.48 s	330.45 s
6x7	277.161 MMUSD	116.085 MMUSD	7	0.25%	0.35%	5870.49 s	607.74 s

Table 2: Computing results for each considered instance.

In Table 2, we present the results for the 3 instances from a numerical point of view. Also, in Figure 6, we present the results from a topological point of view. Surprisingly, in Table 2, we can see that the values of the objective functions in each instance are very similar to each other. This happens for both the GTSELSP and the UCP. Primal-dual gaps for each formulation are also stable. For the GTSELSP, the gap is around 0.26% and for the UCP around 0.30%. As expected, there is a direct relationship between the average computation time and the size of the instance. Nonetheless, we can not establish a relationship between the number of iterations and the size of the instance. However, given the strategic nature of the decisions to be made, the number of iterations seems to be acceptable.

From a topological point of view, we observe that, although there are some small differences, the solutions are similar, as can be seen from Figures 6(a), 6(b) and 6(c). For purposes of a better analysis, we will use the results of the '4x7' instance since it has the longest series of iterations. This allows us to visualize in a better way the behavior of the iteration process.

Investment strategy. A relevant result can be retrieved from Figure 7. We can observe the difference between of the value of the GTSELSP objective function when comparing iteration 1 with respect to iteration 31, where the first GTSELSP solution does not consider the UCP results. The resulting difference corresponds to a value 32.76% lower for the objective function and 31.47% lower for the IRR. This means approximately one third less than the annual earnings calculated by GTSELSP. this difference reveals that evaluating an investment project using solely a model such as the GTSELP, without encompassing a second-stage component, such as the UCP, is likely to lead to a wrong result.



(a) 12x1 (276.252 MMUSD, 7.15%) (b) 4x7 (276.396 MMUSD, 7.11%) (c) 6x7 (277.161 MMUSD, 7.07%)

Figure 6: Comparison of instance results for different time resolutions.

Please note that investment costs (generation, grid, storage and substations) are similar across the iteration process. As a matter of fact, the investment on generation power is almost the same in the 31 iterations and, as can be seen in the below plot of Figure 7, only the substation investment (called as SE cost) varies among the iterations. This can be explain by the fact that a minimum infrastructure is necessary to be able to distribute energy. In addition, there is a great difference between an almost zero investment in ESS in the first iteration and over 8 times more from the 2nd iteration. For this reason, ESSs could be essential in order to improve the competitiveness of projects of this nature.

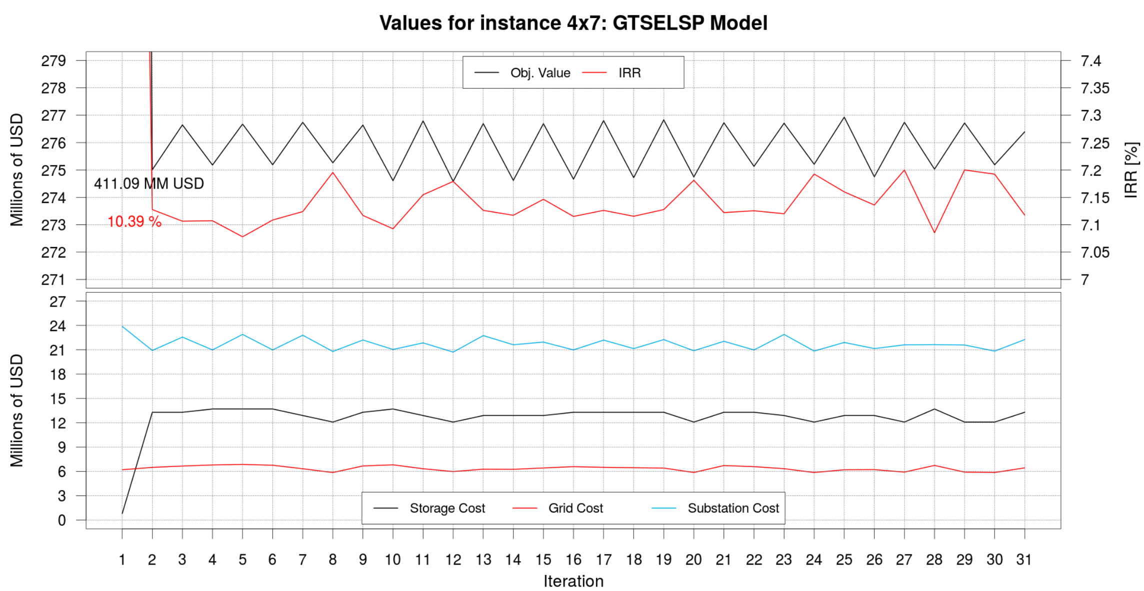


Figure 7: Iterations results from the GTSELSP model. The first value of each serie corresponds to the GTSELSP solution without consider the UCP result.

Energy flow. Unlike the work presented by Olave-Rojas et al. (2017), our decision context also considers the existing power plant in the area and the territorial deployment of the demand. Because of this difference, we also need to analyze the energy flows in the system. In Figure 8, we present total energy flow, obtained from the UCP solution for 4 selected days of the year. Although the generation from renewable energy sources is greater than the demand, the generation from conventional energy sources is necessary to ensure the stability ob the grid. Additionally, between 21:00 and 23:00 hours, we have a peak of renewable generation. Since the curve 'Renewable energy power' includes the renewable power plants and ESSs, the generation peak corresponds to the energy from ESSs because at the time of day there is no sunlight. Moreover, ESSs have a faster response than conventional power plants to sudden changes in demand. Due to the demand presented in Figure 5, where there is high variability between 21:00 and 23:00 hours, a quick response through ESSs seems to be the most convenient. This also implies an increase in the response flexibility of the entire power grid, thus improving its service levels to satisfy the demand.

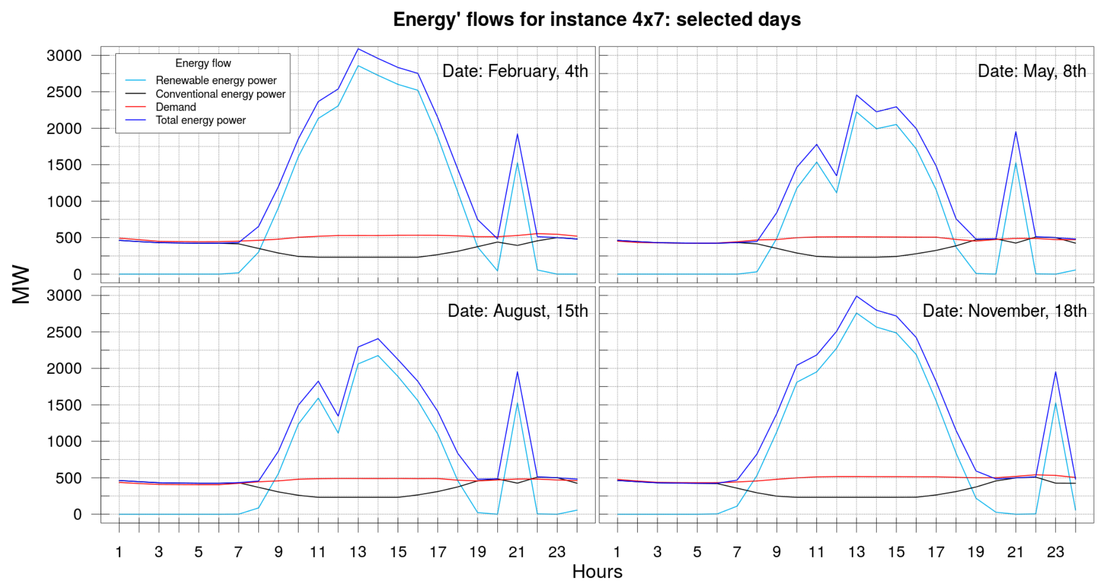
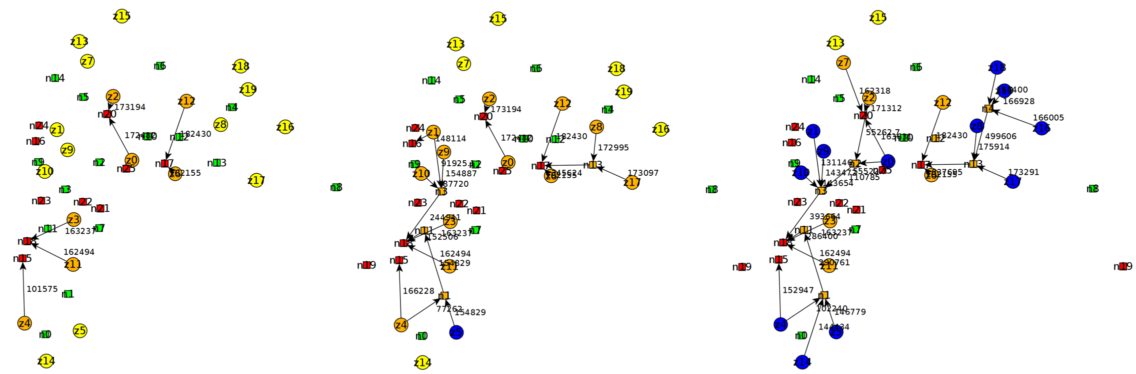


Figure 8: Daily energy flows after results for selected days representing the four year seasons.

Besides the presented results, the proposed methodology can be used for other purposes. Therefore, we present 2 different applications to determine long-term strategies from the point of views of an investor and policymaker.

Budget. Our methodology is also useful for deciding on an investment strategy when we have limited budget. For example, as a matter of fact, we defined three investment levels, associated with different proportions with respect to the investment obtained from the results presented above: 25%, 50% and 75% of the total investment. The results are presented in Figure 9, where there is a clear progression of the project between the different levels. This means that the zones and transmission lines chosen for the 25% budget (see Figure 9(a)) are also maintained in the strategy solution for the 50% budget (Figure 9(b)) in addition to those required for this budget solution in particular. This progression is also present between the Figures 9(b) and 9(c) (75% budget), as well as between the Figures 9(c) and 6(b), which represents a 100% investment solution. Furthermore, there is a logical prioritization: power plants are installed when there is a low budget, and ESSs are installed when the budget increases. The reader can also notice that ESSs (blue color) are built where the power plants are placed.

Knowing the results associated to different budgets allows reducing the risk on the side of the investor since it enables risking a smaller amount of money in the beginning. Additionally, this can be used for turning large projects into accessible ones for investors with a limited budget. This is positive as it increases the number of players with a real option to invest in these projects. Moreover, it can be used to decrease the amount of credit requested by an investor since the initial investment is lower. Finally, the total project becomes more flexible if the total budget is used in different stages (one stage associated to each level). Complementary, from the point of view of a policymaker, this can be useful to divide a large project into stages in order to attract investors.



(a) budget 25% (b) budget 50% (c) budget 75%

Figure 9: Budget comparison results considering three levels of total cost.

Subsidy. Although we include the investor's point of view in this formulation, the policymaker's point of view is equally important in order to support the decisions in complex evaluation scenarios. An example of the utilization of the proposed methodology is deciding whether an area needs subsidies or bonuses with the aim of attracting investment projects for improving its sustainability. In Figure 10, we present the results of IRR obtained for the case study for different horizons of economic evaluation, which are 15, 20, 25 and 30 years. The last one is the planning horizon that we used in all our previous analyses. Taking into account that the project without a subsidy has an IRR of 7.11% (graphically shown in Figure 7), we can compare the results to determine what subsidy is necessary to obtain the same IRR in less time. In this case, for a 25-year horizon, a subsidy of 6% is needed (point C); for a 20-year horizon, a subsidy of 15% should be applied (point B); and for 15 years, one of 27% (point A). Therefore, we can conclude that we should double the subsidy to reduce the return period of the expected investment in 5 years. In this context, we believe the best option for our case study is applying a subsidy of 15% since (i) it decreases in 10 years the payback, that is, in a third the time to obtain the expected return on investment; (ii) the IRR increases by 1.7% (for 30 years, a 15% subsidy means an IRR of 8.80%, that is 23.9% better); (iii) 15% is also a value that can be implemented by means of exemption from taxes in, for example, the place where the case study is developed; and (iv) 20 years is a horizon that has been commonly used to evaluate the inclusion of renewable technologies. In Figure 11, we present the obtained results for our case study in a detailed way for a project evaluation period of 20 years. In addition, for obtaining an IRR = 7.1%, a 15% subsidy is enough for reaching a constant investment in ESS. This is relevant when thinking about boosting investment in these technologies. At 30%, the investment in Substations increases. We can explain this by observing that a greater capacity or quantity in substations, is associated with a more flexible grid for meeting demands, as the system is able to distribute more energy from one point to another. This analysis raises a pattern of investment: first, the investment strategy is focused on generation, then on storage and finally, on improving the transmission power.

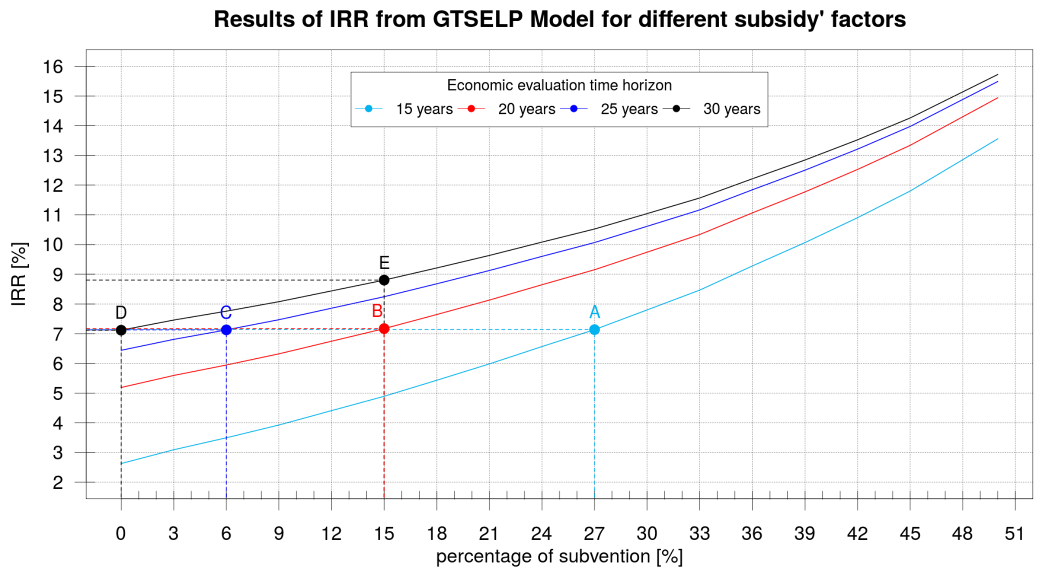


Figure 10: IRR results after different subsidy' factors. Points A, B, C and D represent the same IRR. Points E and B represent the same subsidy for different evaluation time horizons. Difference between points E und B means an IRR improvement of 0.9%.

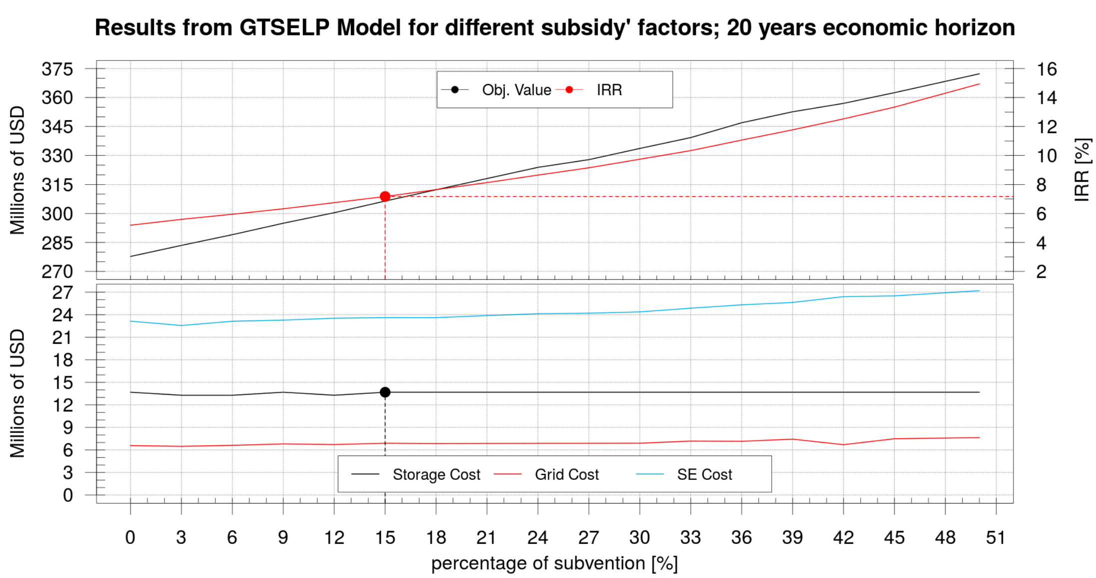


Figure 11: Investment results after different subsidy' factors. We can clearly see the the storage investment is indifferent after 15% subvention. This subsidy is also the indicate for an IRR of 7.1%.

5. Conclusions and Future Work

One of the greatest challenges for humanity is to make a sustainable social model. This has led to renewable energies having a predominant role now and in the future. However, this generates technological, managerial and economic challenges, which must be addressed taking into account every element involved.

In this paper, we present an optimization framework to support strategic decision making for complex energy systems planning, with special emphasis on renewable energies.

The developed methodology relies on an optimization problem coined as "Generation, transmission and storage location and sizing of operation-aware sustainable power system design" (GTSOPSD). This framework takes into account the operational decisions of the system in order to accurately assess strategical decisions associated with investments in energy systems.

By means of a real case study (in this case, located in a central region in Chile) characterized by curated data, we show an improvement with respect to existing tools (see section 2) in the literature. This improvement consists of a more realistic economic evaluation of an investment strategy in power systems since it takes into account the influence of the evaluated project on the grid operation strategy where it will be installed. In this context, for the annualized profit, there is a difference on the value of 32.76% lower and 31.47% lower for the internal rate of return when we evaluate our case study under the presented framework. This represents an improvement of around one third with respect to what exists in the literature.

Furthermore, we demonstrated the use of the framework to define investment stages according to a certain budget. This allows not only reducing the risk and the uncertainty present in these kinds of projects but also reassessing after each stage in order to contemplate updated parameters. Moreover, we showed the use of the framework to develop incentive policies for these projects through subsidies. The framework boosts the investment according to the definition of different objectives such as shortening the return time of the investment, motivating investments in Energy Storage Systems (ESS), determining a subsidy level to achieve a minimum internal rate of return, determining a subsidy level that makes a zone more attractive for investment, among others.

Using the tool for this purposes (associated with budget and subsidies), allows improving the competitiveness of a certain area in order to generate the necessary conditions for sustainable development. Additionally, this increases the possibilities of executing renewable energy projects since it permits more investors to be included.

An interesting path for future work would be developing a model based on a simulation-optimization approach in order to contemplate more parameters and their corresponding stochastic component. An effort in this line would allow including more of the complexities of real world systems, making an even more precise evaluation possible.

Acknowledgments

The authors acknowledge the support of the Chilean Council of Scientific and Technological Research, CONICYT, through the grant FONDECYT N.1180670 and through the Complex Engineering Systems Institute (ICM-FIC:P-05-004-F, CONICYT:FB0816). D. O.-R. acknowledges the support of the graduate school of the U. de Talca.

References

Ackermann, T., Andersson, G., & Söder, L. (2001). Distributed generation: a definition. *Electric power systems research*, *57*, 195–204.

Alsaidan, I., Gao, W., & Khodaei, A. (2018). Distribution network expansion through optimally sized and placed distributed energy storage. In *2018 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)* (pp. 1–5). IEEE.

Armaroli, N., & Balzani, V. (2007). The future of energy supply: Challenges and opportunities. *Angewandte Chemie International Edition*, *46*, 52–66.

Berrada, A., & Loudiyi, K. (2016). Operation, sizing, and economic evaluation of storage for solar and wind power plants. *Renewable and Sustainable Energy Reviews*, *59*, 1117–1129.

Bhardwaj, A., Kamboj, V., Shukla, V., Singh, B., & Khurana, P. (2012a). Unit commitment in electrical power system-a literature review. In *IEEE International Power Engineering and Optimization Conference* (pp. 275–280). IEEE.

Bhardwaj, A., Tung, N., & Kamboj, V. (2012b). Unit commitment in power system: A review. *International Journal of Electrical and Power Engineering*, *6*, 51–57.

Bradbury, K., Pratson, L., & Patiño-Echeverri, D. (2014). Economic viability of energy storage systems based on price arbitrage potential in real-time us electricity markets. *Applied Energy*, *114*, 512–519.

Carley, S. (2009). State renewable energy electricity policies: An empirical evaluation of effectiveness. *Energy Policy*, *37*, 3071–3081.

Caunhye, A. M., & Cardin, M.-A. (2018). Towards more resilient integrated power grid capacity expansion: A robust optimization approach with operational flexibility. *Energy Economics*, *72*, 20–34.

Chang, K., & Lin, G. (2015). Optimal design of hybrid renewable energy systems using simulation optimization. *Simulation Modelling Practice and Theory*, *52*, 40–51.

Chen, H., Cong, T., Yang, W., Tan, C., Li, Y., & Ding, Y. (2009). Progress in electrical energy storage system: A critical review. *Progress in Natural Science*, *19*, 291–312.

Chen, H., Zhang, Y., & Ngan, H. (2016). Power system planning. In H. Chen, Y. Zhang, H. Ngan, X. Wang, X. Wang, T. Chung, F. Zhang, Z. Hu, & M. Xie (Eds.), *Power System Optimization* chapter 4. (pp. 76–130). John Wiley & Sons.

Coordinador Eléctrico Nacional (2016). Programación Semanal CDEC SIC. `https://sic.coordinadorelectrico.cl/informes-y-documentos/fichas/operacion-programada-2/`. [Online; accessed 03-May-2016].

Dehghan, S., & Amjady, N. (2016). Robust transmission and energy storage expansion planning in wind farm-integrated power systems considering transmission switching. *IEEE Transactions on Sustainable Energy*, *7*, 765–774.

Dincer, I. (2000). Renewable energy and sustainable development: A crucial review. *Renewable and Sustainable Energy Reviews*, *4*, 157–175.

Dominguez, R., Baringo, L., & Conejo, A. (2012). Optimal offering strategy for a concentrating solar power plant. *Applied Energy*, *98*, 316–325.

Dorian, J., Franssen, H., & Simbeck, D. (2006). Global challenges in energy. *Energy Policy*, *34*, 1984–1991.

Ghofrani, M., Arabali, A., Etezadi-Amoli, M., & Fadali, M. (2013). A framework for optimal placement of energy storage units within a power system with high wind penetration. *IEEE Transactions on Sustainable Energy*, *4*, 434–442.

Guerra, O., Tejada, D., & Reklaitis, G. (2016). An optimization framework for the integrated planning of generation and transmission expansion in interconnected power systems. *Applied Energy*, *170*, 1–21.

Haas, J., Nowak, W., & Palma-Behnke, R. (2019). Multi-objective planning of energy storage technologies for a fully renewable system: Implications for the main stakeholders in chile. *Energy Policy*, *126*, 494–506.

Harsha, P., & Dahleh, M. (2015). Optimal management and sizing of energy storage under dynamic pricing for the efficient integration of renewable energy. *IEEE Transactions on Power Systems*, *30*, 1164–1181.

Jabr, R., Džafić, I., & Pal, B. (2015). Robust optimization of storage investment on transmission networks. *IEEE Transactions on Power Systems*, *30*, 531–539.

Jiang, R., Wang, J., & Guan, Y. (2012). Robust unit commitment with wind power and pumped storage hydro. *IEEE Transactions on Power Systems*, *27*, 800–810.

Kennedy, T., & Mabuce, E. (1965). Dispatch of pumped hydro storage on an interconnected hydrothermal system. *IEEE Transactions on Power Apparatus and Systems*, *84*, 446–457.

Kerr, R., Scheidt, J., Fontanna, A., & Wiley, J. (1966). Unit commitment. *IEEE Transactions on Power Apparatus and Systems*, (pp. 417–421).

Kim, H., Olave-Rojas, D., Álvarez-Miranda, E., & Son, S. (2018). In-depth data on the network structure and hourly activity of the central chilean power grid. *Scientific data*, *5*, 180209.

Kim, J., & Powell, W. (2011). Optimal energy commitments with storage and intermittent supply. *Operations research*, *59*, 1347–1360.

Koltsaklis, N., & Georgiadis, M. (2015). A multi-period, multi-regional generation expansion planning model incorporating unit commitment constraints. *Applied Energy*, *158*, 310–331.

Koltsaklis, N., Liu, P., & Georgiadis, M. (2015). An integrated stochastic multi-regional long-term energy planning model incorporating autonomous power systems and demand response. *Energy*, *82*, 865–888.

Lan, H., Yin, H., Wen, S., Hong, Y., Yu, D., & Zhang, L. (2017). Electrical energy forecasting and optimal allocation of ess in a hybrid wind-diesel power system. *Applied Sciences*, *7*, 155.

Lowery, R. (1966). Generating unit commitment by dynamic programming. *IEEE Transactions on Power Apparatus and Systems*, (pp. 422–426).

Luo, X., Wang, J., Dooner, M., & Clarke, J. (2015). Overview of current development in electrical energy storage technologies and the application potential in power system operation. *Applied Energy*, *137*, 511–536.

MacRae, C., Ozlen, M., Ernst, A., & Behrens, S. (2015). Locating and sizing energy storage systems for distribution feeder expansion planning. In *TENCON 2015-2015 IEEE Region 10 Conference* (pp. 1–6). IEEE.

Mehigan, L., Deane, J., Gallachóir, B., & Bertsch, V. (2018). A review of the role of distributed generation (DG) in future electricity systems. *Energy*, *163*, 822–836.

Ministerio de Energía de Chile (2011). Explorador de Energía Solar. `http://walker.dgf.uchile.cl/Explorador/Solar3/`. [Online; accessed 03-May-2016].

Molina, A., Falvey, M., & Rondanelli, R. (2017). A solar radiation database for Chile. *Scientific reports*, *7*, 14823.

Momoh, J., & Mili, L. (Eds.) (2009). *Economic Market Design and Planning for Electric Power Systems*. Wiley IEEE Press.

Muche, T. (2014). Optimal operation and forecasting policy for pump storage plants in day-ahead markets. *Applied Energy*, *113*, 1089–1099.

Oatman, E., & Hamant, L. (1973). A dynamic approach to generation expansion planning. *IEEE Transactions on Power Apparatus and Systems*, *PAS-92*, 1888–1897.

Oh, H. (2011). Optimal planning to include storage devices in power systems. *IEEE Transactions on Power Systems*, *26*, 1118–1128.

Olave-Rojas, D., Álvarez-Miranda, E., Rodríguez, A., & Tenreiro, C. (2017). An optimization framework for investment evaluation of complex renewable energy systems. *Energies*, *10*, 1062.

Oudalov, A., Cherkaoui, R., & Beguin, A. (2007). Sizing and optimal operation of battery energy storage system for peak shaving application. In *Power Tech, 2007 IEEE Lausanne* (pp. 621–625). IEEE.

Owusu, P., & Asumadu-Sarkodie, S. (2016). A review of renewable energy sources, sustainability issues and climate change mitigation. *Cogent Engineering*, *3*, 1–14.

Pascual, L., Romo, J., & Ruiz, E. (2004). Bootstrap predictive inference for arima processes. *Journal of Time Series Analysis*, *25*, 449–465.

Pascual, L., Romo, J., & Ruiz, E. (2005). Bootstrap prediction intervals for power-transformed time series. *International Journal of Forecasting*, *21*, 219–235.

Qi, W., Liang, Y., & Shen, Z. (2015). Joint planning of energy storage and transmission for wind energy generation. *Operations Research*, *63*, 1280–1293.

Rahmann, C., & Palma-Behnke, R. (2013). Optimal allocation of wind turbines by considering transmission security constraints and power system stability. *Energies*, *6*, 294–311.

Roboam, X. (Ed.) (2012). *Integrated Design by Optimization of Electrical Energy Systems*. Wiley-ISTE.

Saravanan, B., Das, S., Sikri, S., & Kothari, D. (2013). A solution to the unit commitment problem - a review. *Frontiers in Energy*, *7*, 223–236.

Seifi, H., & Sepasian, M. (Eds.) (2011). *Electric Power System Planning: Issues, Algorithms and Solutions*. Power Systems. Springer.

Suazo-Martínez, C., Pereira-Bonvallet, E., & Palma-Behnke, R. (2014). A simulation framework for optimal energy storage sizing. *Energies*, *7*, 3033–3055.

UN Sustainable Development Goals (2016a). Goal 13: Take urgent action to combat climate change and its impacts. https://www.un.org/sustainabledevelopment/climate-change-2/. Visited 2019-02-18.

UN Sustainable Development Goals (2016b). Goal 7: Affordable and Clean Energy. https://www.un.org/sustainabledevelopment/energy/. Visited 2019-02-18.

United Nations (2016). Sustainable Development Goals. https://www.un.org/sustainabledevelopment/sustainable-development-goals/. Visited 2019-02-18.

World Future Council (2016). Renewable Energy and Sustainable Development. https://www.worldfuturecouncil.org/renewable-energy-sustainable-development/. Visited 2019-02-18.

Xia, H., Chen, H., Yang, Z., Lin, F., & Wang, B. (2015). Optimal energy management, location and size for stationary energy storage system in a metro line based on genetic algorithm. *Energies*, *8*, 11618–11640.

Xiong, P., & Singh, C. (2016). Optimal planning of storage in power systems integrated with wind power generation. *IEEE Transactions on Sustainable Energy*, *7*, 232–240.

Zakeri, B., & Syri, S. (2015). Electrical energy storage systems: A comparative life cycle cost analysis. *Renewable and Sustainable Energy Reviews*, *42*, 569–596.

Zheng, Q., Wang, J., & Liu, A. (2015). Stochastic optimization for unit commitment - a review. *IEEE Transactions on Power Systems*, *30*, 1913–1924.