



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

SEGMENTACIÓN Y CARACTERIZACIÓN DE CLIENTES LIBRES DEL SISTEMA
ELÉCTRICO NACIONAL PARA MODELAR DEMANDA FLEXIBLE

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL ELÉCTRICO

DIEGO ANDRÉS GUIRALDES DECK

PROFESOR GUÍA:
RODRIGO MORENO VIEYRA

MIEMBROS DE LA COMISIÓN:
JUAN ARAYA PÉREZ
ANDRÉS CABA RUTTE

Este trabajo ha sido parcialmente financiado por Colbún

SANTIAGO DE CHILE
2020

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE: INGENIERO CIVIL ELÉCTRICO
POR: DIEGO ANDRÉS GUIRALDES DECK
FECHA: 2020
PROFESOR GUÍA: Dr. RODRIGO MORENO VIEYRA

SEGMENTACIÓN Y CARACTERIZACIÓN DE CLIENTES LIBRES DEL SISTEMA ELÉCTRICO NACIONAL PARA MODELAR DEMANDA FLEXIBLE

Las mejoras tecnológicas en el sector eléctrico han permitido que se desarrollen nuevos negocios. Dentro de estos se encuentra la Respuesta en Demanda (DR por sus siglas en inglés), donde se llega a un acuerdo para que los clientes adapten su consumo en función de los precios del mercado. Para estudiar el comportamiento de un mercado que opere con DR, es necesario modelar correctamente la demanda del sistema y su flexibilidad.

El objetivo de este trabajo es segmentar a los clientes libres del Sistema Eléctrico Nacional (SEN) y caracterizar su flexibilidad, con el fin de representar la demanda en un modelo de operación con DR. Para esto se usa una base de datos que contiene el balance de inyecciones y retiros horarios del SEN desde septiembre de 2018 hasta agosto de 2019. Luego de adaptar y filtrar los datos, se segmentan los clientes según perfil de consumo y sector económico al que pertenecen. Finalmente se caracterizan las agrupaciones encontradas en función de su flexibilidad, mediante entrevistas a clientes de las agrupaciones.

Los perfiles de consumo resultan de agrupar las series de tiempo de los retiros de energía, en un promedio diario por semestre (invierno o verano) y por tipo de día (día lunes, día hábil distinto de lunes, sábado y domingo o feriado). Los sectores económicos son extraídos de una base de datos que relaciona los nombres con el rubro definido por Servicio de Impuestos Internos. La segmentación se lleva a cabo mediante un algoritmo de aprendizaje de máquinas denominado TSNE en conjunto con un algoritmo de clusterización denominado DBSCAN.

La metodología presentada para la segmentación demuestra tener excelentes resultados. Se logra identificar agrupaciones de clientes según perfil de consumo y sector económico, de las cuales destacan 8 que concentran la mayor parte de los retiros. Esto no sirve solamente para modelar la flexibilidad, sino que se puede usar para pensar en nuevos negocios específicos para cada segmento y mejorar los modelos de negocio actuales.

La principal conclusión de las entrevistas con clientes y de los análisis de resultados, es que los perfiles de consumo revelan información importante de la flexibilidad de un cliente, mientras que su sector económico no lo hace. La flexibilidad que puede ofrecer un cliente depende estrictamente de los procesos productivos que desempeña, los cuales se ven reflejados en sus perfiles de consumo. Por el contrario, muchas empresas poseen procesos que pueden asociarse a más de un sector económico, por lo que caracterizarlas por un solo sector puede llevar a una segmentación equivocada. Por ende, a la hora de caracterizar la flexibilidad de un retiro, el perfil de consumo revela mucha información y el sector económico no es relevante.

La continuación del trabajo es definir un modelo del SEN y simular el despacho con demanda flexible usando los resultados obtenidos. Sin embargo, antes de seguir con este paso, es preferible repetir el experimento con una serie de mejoras propuestas.

Tabla de Contenido

1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	2
1.3. Supuestos y alcances	3
1.4. Estructura del documento	3
2. Antecedentes generales	4
2.1. El mercado eléctrico chileno	4
2.1.1. Sistemas eléctricos	4
2.1.2. Segmentos del mercado	4
2.1.3. Operación del sistema	5
2.1.4. Clientes del mercado	5
2.1.5. Balances de inyecciones y retiros	6
2.2. Respuesta en demanda	6
2.2.1. Infraestructura necesaria	7
2.2.2. Programas de DR	7
2.2.3. Ejemplos internacionales	8
2.2.4. Gestión de demanda en Chile	10
2.3. Técnicas de data science y machine learning	11
2.3.1. Clusterización	11
2.3.2. Principal component analysis (PCA)	12
2.3.3. t-Distributed Stochastic Neighbor Embedding (t-SNE)	12
3. Metodología	14
3.1. Modelo de operación con demanda flexible	14
3.1.1. Descripción general	14
3.1.2. Descripción de generación	14
3.1.3. Descripción de transmisión	15
3.1.4. Descripción de demanda	16
3.1.5. Restricciones del modelo	17
3.1.6. Función objetivo	19
3.2. Datos de balance	19
3.3. Corrección de nombres	20
3.4. Asignación de sectores	20
3.5. Clasificación por perfil de consumo y sector	22
3.5.1. Manejo de outliers	23

3.5.2.	Información para clasificación	28
3.5.3.	Reducción de dimensiones y clusterización	29
3.6.	Entrevistas a clientes representativos	30
4.	Resultados	31
4.1.	Clusterización de clientes	31
4.1.1.	Corrección de nombres, confección de retiros y asignación de sectores	31
4.1.2.	Limpieza de datos	33
4.1.3.	Representación bidimensional	36
4.1.4.	Descomposición por sector	37
4.1.5.	Perfiles por cluster	38
4.2.	Entrevistas	41
4.2.1.	Cliente 1: Sector agrícola, perfil estacional	41
4.2.2.	Cliente 2: Sector industria manufacturera no metálica, perfil oficina industrial	42
4.2.3.	Cliente 3: Sector comercio al por mayor y menor, perfiles típico industrial y estacional	43
4.2.4.	Cliente 4: Sector industria manufacturera no metálica, perfiles típico industrial, escalón semanal y estacional	44
4.2.5.	Cliente 5: Sector comercio al por mayor y menor, perfiles escalón semanal y típico industrial	46
4.2.6.	Cliente 6: Sector comercio al por mayor y menor, perfil retail	47
5.	Análisis	49
5.1.	Discusión metodológica	49
5.1.1.	Base de clientes y confección de perfiles	49
5.1.2.	Corrección de nombres	50
5.1.3.	Asignación de sectores	50
5.1.4.	Confección de retiros	51
5.1.5.	Filtros de datos	54
5.1.6.	Métodos de clusterización	55
5.2.	Análisis de clusters	57
5.2.1.	Selección de clusters más importantes	57
5.2.2.	Caracterización de clusters importantes	60
5.3.	Entrevistas	65
5.3.1.	Deducciones generales	65
5.3.2.	Análisis por cluster	66
6.	Conclusiones y trabajo futuro	70
6.1.	Trabajo futuro	71
	Bibliografía	73
	Anexo A. Modelo propuesto	75
	Anexo B. Entrevista	77
B.1.	Introducción	77
B.2.	Entrevista	78

Anexo C. Tablas y figuras complementarias	79
C.1. Resultados	79
C.2. Análisis	82

Capítulo 1

Introducción

1.1. Motivación

La generación de electricidad en Chile es un mercado competitivo, cuya operación está a cargo de un organismo autónomo denominado Coordinador Eléctrico Nacional. Las empresas que generan electricidad lo hacen según las órdenes del Coordinador, recibiendo una remuneración por lo que brindan de energía y de potencia. El negocio de energía está dividido en dos: producción y comercialización.

En el negocio de producción interactúan los generadores con el Coordinador, donde los primeros producen energía según las órdenes del segundo, la cual se tranza a un precio competitivo. Este precio varía con frecuencia horaria y viene dado por el costo marginal del sistema, es decir el costo de inyectar un MW adicional. Por lo tanto, bajo la perspectiva del negocio de producción, a los generadores les conviene que el precio sea alto para vender la energía más cara.

En el negocio de comercialización los generadores interactúan con los clientes y el Coordinador. Para esto compran energía del sistema (valorizada a costo marginal) y luego se la venden a los consumidores a un precio que tranzan mediante un contrato. Los contratos con clientes regulados (entender como clientes pequeños) son acordados mediante un ente estatal (Comisión Nacional de Energía), mientras que para clientes libres (entender como clientes grandes) se hace mediante una negociación directa.

El contrato con los clientes fija un precio de energía, lo que hace que estos no vean señales de precio según el momento que consumen. De esta forma, los generadores perciben ganancias de comercialización cuando sus clientes retiran en horas en que el precio de la energía es menor que su precio de contrato y perciben pérdidas en las horas que el precio de energía es mayor que el precio de contrato. Por lo tanto, desde el punto de vista del negocio de comercialización, a los generadores les conviene que el costo marginal del sistema sea bajo en los horas que consumen sus clientes.

Si el generador logra que sus clientes trasladen sus consumos de las horas caras a las horas más baratas, podría tener ahorros considerables, que pueden repartirse con el cliente que

gestionó su consumo. Al modelo de negocios en que el cliente adapta su consumo en función de los precios del mercado se le conoce como Respuesta en Demanda (DR por sus siglas en inglés). La multinacional Enel X define DR como la acción que *permite a los consumidores comerciales e industriales responder a las señales del mercado aumentando o reduciendo su consumo energético, con vistas a responder a los picos de oferta o demanda de electricidad, permitiendo una mayor flexibilidad y estabilidad de la red y un uso más eficientes de las infraestructuras y los recursos energéticos.* [1]

Si son pocas las empresas que efectúan DR y su consumo no es significativo en el mercado, probablemente el único efecto que esto tendría es que el generador a cargo del retiro tendría un ahorro (que naturalmente debiese compartirse con el cliente que gestiona). Ahora, si son muchas las empresas que gestionan su demanda en función de las señales del mercado, o bien lo hacen grandes consumos que tienen una participación importante en el mercado, podrían llegar a cambiar el precio al que se tranza la energía, lo que en teoría significaría una disminución de costos del sistema. Este cambio de precio afectaría tanto en el negocio de producción como de comercialización.

En Chile, recién se está comenzando a desarrollar la infraestructura que permite realizar DR, por lo que aún no se ha visto alta penetración en el mercado. Tampoco se tiene claro cuales serían las empresas generadoras que más se beneficiarían ni los consumidores que podrían brindar el servicio. Por lo tanto, se requiere hacer estudios que permitan tener nociones de los clientes que podrían prestar servicios de DR, para luego simular comportamientos del sistema donde los generadores incorporen este modelo.

En este trabajo se propone un modelo para simular la operación del sistema con demanda flexible, a través de una metodología de segmentación y caracterización de los clientes libres. Las simulaciones y análisis quedan propuestas como trabajo futuro. El foco de esta investigación es segmentar correctamente la demanda de electricidad y caracterizar su potencial de flexibilidad.

1.2. Objetivos

- **Objetivo general:** Segmentar a los clientes libres del Sistema Eléctrico Nacional y caracterizar su flexibilidad, con el fin de representar la demanda en un modelo de operación con respuesta en demanda.
- **Objetivos específicos:**
 - i Proponer un modelo que simule la operación del sistema en que la demanda responda a las señales del mercado.
 - ii Armar una base de datos de retiros horarios del sistema por cliente y su sector económico, para identificar patrones de consumo.
 - iii Filtrar la base de datos, para eliminar anomalías que no corresponden a patrones de consumo.
 - iv Segmentar a los clientes del mercado según sector económico y perfil de consumo.
 - v Caracterizar el potencial de flexibilidad de agrupaciones principales, mediante entrevistas con clientes representativos.

1.3. Supuestos y alcances

- Se estudia la demanda del Sistema Eléctrico Nacional (SEN) de un año (Sep 2018 - Ago 2019), cuyos datos se obtienen de los balances de inyecciones y retiros mensuales que entrega el Coordinador.
- El modelo de Respuesta en Demanda asume que todos los agentes del mercado tienen la capacidad de actuar con respuesta en demanda. Bajo lo anterior, se asume que existe la infraestructura adecuada y que los clientes son capaces de variar su consumo en función de las señales de forma racional.
- La simulación del modelo de despacho con respuesta en demanda queda propuesta como trabajo futuro.
- Solo se consideran como potenciales clientes flexibles a los libres y libres en distribución. Es decir, los clientes regulados son tratados como demanda fija.
- El problema se aborda desde el punto de vista del mercado de energía y no de potencia ni servicios complementarios.
- A pesar de que se analizan posibles modelos de negocios acorde a la experiencia internacional, está fuera del alcance del trabajo definir alguno.

1.4. Estructura del documento

- **Capítulo 2: Antecedentes.** En este capítulo se definen los conceptos necesarios para entender el documento. Se explican los fundamentos básicos del sector eléctrico de Chile, se define el concepto de respuesta en demanda, se dan ejemplos internacionales y finalmente se explican las técnicas de data science utilizadas.
- **Capítulo 3: Metodología.** En el tercer capítulo se especifica el procedimiento que se lleva a cabo para realizar el estudio completo. Se propone el modelo matemático para simular la operación con DR, se detallan los procedimientos para extracción y análisis de datos, y la metodología para entrevistarse con clientes.
- **Capítulo 4: Resultados.** En este apartado se exponen todos los resultados obtenidos: el manejo de bases de datos, la agrupación de clientes, las características de las agrupaciones y los resultados de las entrevistas.
- **Capítulo 5: Análisis.** Se discuten las metodologías adoptadas y los resultados. Además se le da un sentido lógico a la segmentación, para luego relacionar los resultados con las entrevistas y caracterizar la flexibilidad de la demanda.
- **Capítulo 6: Conclusiones y trabajo futuro.** En este último capítulo se extraen las conclusiones del trabajo, determinando así el valor de la investigación. Luego se plantean los desafíos para seguir ahondando en el estudio.
- **Anexos.** Se incluye el modelo de optimización para simular el despacho con demanda flexible, la pauta de entrevista a clientes e información adicional de tablas y figuras que no son expuestas en el informe. Además, los códigos utilizados están disponibles en el siguiente repositorio: <https://github.com/dguiraldes/Memoria>

Capítulo 2

Antecedentes generales

2.1. El mercado eléctrico chileno

El mercado eléctrico chileno comienza a desarrollarse a fines del siglo XIX. Este pasa de estar compuesto por pequeñas centrales generadoras, empresas de alumbrado público y tranvías a un complejo sistema donde interactúan miles de empresas. En esta sección se explican los aspectos fundamentales del mercado.

2.1.1. Sistemas eléctricos

Actualmente, el sistema eléctrico chileno se divide en 3 sistemas independientes que abarcan la totalidad del territorio: Sistema Eléctrico Nacional (SEN), Sistema de Aysén (SEA) y Sistema de Magallanes (SEM). El SEN concentra el 99.27 % de la capacidad de generación instalada, mientras que SEA y SEM poseen el restante 0.27 % y 0.46 % respectivamente. En este estudio solo se considera el SEN, por lo que este se detalla a continuación.

El SEN está compuesto por los antiguos sistemas Interconectado Central (SIC) e Interconectado del Norte Grande (SING), abarcando desde la región de Arica y Parinacota hasta la región de Los Lagos. Para abril de 2019 cuenta con una capacidad instalada de 23.218 MW. El 47 % de la capacidad instalada corresponde a fuentes renovables (28 % hidráulica, 10 % solar, 7 % eólico y 2 % biomasa) mientras que el 53 % corresponde a fuentes térmicas (21 % carbón, 19 % gas natural y 13 % petróleo)[2]. En los últimos 5 años se ha visto un gran aumento en la generación mediante energía eólica y solar fotovoltaica, las cuales pasaron de representar 0.8 % y 0.01 % de la generación anual respectivamente en 2013, a 5.1 % y 6.7 % en 2018. [3]

2.1.2. Segmentos del mercado

El mercado eléctrico en Chile, desde el lado de la oferta de energía, está compuesto por tres sectores cuyas actividades hacen posible la generación de energía eléctrica y su disposición para todos los clientes finales. Las empresas que participan del mercado son privadas y se rigen por normas dependientes del sector. Dichos sectores se describen a continuación.[3]

- **Generación:** Tiene como función la producción y comercialización de energía eléctrica. Este sector se rige por las reglas de libre competencia, donde los inversionistas definen los proyectos y tecnología que quieran desarrollar. La generación se puede dividir en dos mercados: obligatorio y voluntario. El primero es dependiente de la operación (detallado en 2.1.3) y el segundo, de contratos con los clientes del mercado.
- **Transmisión:** Tiene como función la transmisión, en niveles altos de voltaje (mayor a 23 kV), de la energía producida a todos los puntos del sistema eléctrico. Este es un servicio público que está regulado, sin embargo, las inversiones son privadas.
- **Distribución:** Tiene como función el distribuir la energía, en niveles de voltaje menores que los de Transmisión (menor o igual a 23 kV), desde un cierto punto del sistema eléctrico a los consumidores finales. Al igual que la transmisión, se considera un servicio público, en que se regula la operación, calidad de servicio y tarificación. Las empresas distribuidoras se dividen en zonas de concesión definidas por las autoridades.

2.1.3. Operación del sistema

La operación del sistema se diseña con el fin de abastecer la demanda eléctrica en todo instante de forma segura y al menor costo. El encargado de coordinar la operación de la generación es una empresa centralizada que se denomina Coordinador Eléctrico Nacional (CEN). Esta es una corporación autónoma de derecho público, sin fines de lucro, con patrimonio propio y de duración indefinida. [4]

El mercado de generación sigue un esquema tipo pool centralizado de costos auditados. El precio de la energía en la operación (precio spot), se define como el costo marginal del sistema, el cual es el costo de inyectar un MW adicional con la unidad generadora más eficiente disponible. Es el CEN el encargado de auditar los costos de las centrales, calcular los precios de energía e informarlo a los coordinados.

2.1.4. Clientes del mercado

Los clientes del mercado de electricidad en Chile se dividen en dos tipos: libres y regulados. Los clientes libres son aquellos que negocian su compra de electricidad directamente con los generadores, mientras que los otros acceden a tarifas reguladas a cargo de las empresas distribuidoras. Cabe destacar que pueden participar en el mercado de inyecciones y retiros de energía, solamente empresas generadoras acorde a lo dispuesto en el Artículo 17 del DS 291 [5] [6]. Esto implica, que las empresas que no son generadoras, no pueden comprar energía del mercado spot, por lo que deben contratar su energía con aquellas empresas que están habilitadas para hacer retiros.

De acuerdo a la Ley General de Servicios Eléctricos (LGSE), todos los consumos que tengan una capacidad conectada mayor a 5MW deben ser clientes libres. Por otro lado, aquellos cuya capacidad sea menor a 500 kW deben ser regulados. Los clientes cuya potencia conectada es mayor a 500 kW, pero menor a 5 MW, pueden elegir si ser clientes libres o regulados. [7]

Los clientes libres se catalogan en dos grupos: libre y libre en distribución. Los primeros deben conectarse al sistema de transmisión bajo sus propios medios, cumpliendo con la

normativa vigente y pagando los peajes de transmisión que corresponde. Los segundos están conectados al sistema mediante redes cuyos propietarios son empresas distribuidoras, por lo que deben pagar a estas un peaje adicional por el uso de su infraestructura. Los generadores a cargo de retiros de sus clientes deben declararlos al Coordinador con precisión horaria,

Los clientes regulados están conectados al sistema mediante las redes de las empresas distribuidoras, quienes ofrecen distintos tipos de tarifa reguladas por las autoridades. En cuanto a la energía, las distribuidoras la compran a las empresas que participan en el mercado spot mediante licitaciones de suministro coordinadas por la Comisión Nacional de Energía (CNE).

2.1.5. Balances de inyecciones y retiros

El Coordinador publica de forma mensual una base de datos con el balance nodal horario del mes anterior. Es decir, se declaran todas las inyecciones de energía por central, todos los retiros por empresa, la transmisión por líneas y consumos del sistema. Además se incluyen los costos marginales de las barras. El proceso de cálculo de transferencias se detalla en una guía de aplicación confeccionada por el Coordinador [8]. A continuación se explica este proceso.

A más tardar el quinto día hábil de cada mes, las empresas participantes de los balances de transferencia, empresas distribuidoras y transmisoras deben enviar al Coordinador, las lecturas de inyecciones y retiros de energía u otros parámetros necesarios para la valorización de transferencias del mes anterior. El Coordinador procesa la información y la combina con otros datos como nuevas instalaciones eléctricas que ingresen a la red, cambios en la topología de red ,etc. Luego, se envían los cuadros de pago, los balances físicos y el de transferencias económicas a las empresas participantes. Además se publican los archivos con los antecedentes del cálculo, que confeccionan el balance preliminar.

A las empresas participantes del balance de transferencias se les otorga un plazo de dos días para emitir las observaciones al balance preliminar. En el caso de existir observaciones al cálculo enviado, el Coordinador evalúa la pertinencia y de ser válidas las corrige.

Una vez realizada las correcciones (debido a las observaciones), se obtiene el balance definitivo de transferencias económicas, y se emiten los reportes asociados. Después se envían a los Coordinados los cuadros de pago, los balances físicos y el de transferencias económicas. Finalmente se publican los archivos con los antecedentes del cálculo, que representan la topología del sistema, las inyecciones, los retiros y los precios utilizados en cada una de las barras del sistema. El balance definitivo debe ser publicado a más tardar el día veinticuatro (24) del mes, o el día hábil inmediatamente siguiente si aquél no lo fuese.

2.2. Respuesta en demanda

Como se menciona en la introducción, la respuesta a la demanda (en adelante DR) se refiere a los cambios en el patrón de consumo de los clientes del sistema eléctrico en función de las señales del mercado. En esta sección se describen las condiciones necesarias para la implementación de DR, las clasificaciones que pueden hacerse y los beneficios que trae, para luego revisar ejemplos internacionales y los desafíos para Chile. La revisión se basa

principalmente en el estudio de Paterakis et al. [9]

2.2.1. Infraestructura necesaria

Los programas de DR se pueden implementar siempre y cuando existan las señales y la infraestructura adecuada. En cuanto a las primeras, se necesita un mercado que refleje los costos del sistema en sus señales de precio, como por ejemplo, contar con un precio horario de electricidad definido por el costo marginal del sistema. En cuanto a la infraestructura, se necesitan las tecnologías de medición y comunicaciones adecuadas para que exista una correcta interacción entre proveedor de electricidad y cliente. Por ejemplo, que el proveedor pueda hacer un seguimiento en tiempo real al precio de energía del sistema y al consumo de los clientes, para luego enviar una señal para adaptar la demanda. A continuación se detallan los requerimientos de medición y comunicaciones.

Medición

Para efectuar DR es fundamental contar con sistemas de medición inteligente. Estos sistemas entregan datos de carga en tiempo real, con lo que los proveedores pueden monitorear el consumo. Además, a través de estos mismos medidores, los proveedores pueden enviar a los clientes señales de precio o potencia máxima permitida en un intervalo de tiempo, para que estos adapten su consumo. Esto hace que se requiera un sistema de comunicación bidireccional. Una red que cuenta con muchos medidores inteligentes, puede considerarse como “infraestructura de medición avanzada” o AMI (por sus siglas en inglés “Advanced Metering Infrastructure”).

Las estructuras más avanzadas, además de incluir medidores inteligentes, tienen “sistemas de gestión de energía” o EMS (por sus siglas en inglés “Energy Management System”). Estos reciben la información del consumo de distintas cargas (controlables o no controlables) y de la generación propia de los clientes (como paneles solares, baterías o generadores convencionales). Además, las señales de precio e instrucciones de DR enviadas por el proveedor, son traspasadas al EMS mediante la AMI. De esta forma, el EMS toma todas estas entradas y decide la estrategia de operación óptima del cliente, para cumplir con las instrucciones del proveedor sin afectar sus necesidades de consumo.

Comunicaciones

Otro requisito fundamental para efectuar DR, es la capacidad para soportar la transferencia bidireccional de una gran cantidad de datos. Es por esto que se necesita un sistema de comunicaciones entre proveedor y cliente con baja latencia y alto ancho de banda. La latencia corresponde al retraso percibido por el receptor de un mensaje en contraste con el momento que fue enviado, mientras que el ancho de banda corresponde a la tasa de transferencia de datos de los actores involucrados.

2.2.2. Programas de DR

Los programas de DR se pueden clasificar en dos tipos: basados en incentivos y basados en señales de precio. En los primeros se remunera a los clientes por bajar su carga, mientras

que en los segundos, los clientes lo hacen de forma voluntaria en respuesta a distintos precios de energía. A continuación se describen algunos programas de cada tipo.

En base a incentivos

- **Control de carga directo:** Se ocupa cuando se tiene un gran número de clientes pequeños con cargas altamente controlables (por ejemplo, aire acondicionado, luces, calentamiento de agua, bombas de agua, etc). El proveedor controla estas cargas sujeto a las condiciones de los clientes y les remunera por su flexibilidad.
- **Demanda recortable:** Principalmente destinado a clientes con grandes consumos; bajo este programa el proveedor recompensa a los clientes por reducir o cortar su consumo en los momentos que se le indique.
- **Subastas de reducción de consumo:** Este programa permite que la demanda participe activamente de la operación del sistema al ofertar sus reducciones. A este acceden grandes consumidores y agregadores de demanda.

En base a señales de precio

- **Tarifas diferenciadas por tiempo de consumo:** El objetivo de esta tarifa es que los clientes perciban los cambios en el precio de la energía según horario, tipo de día o estación. Para esto, se definen precios diferenciados para distintos momentos en el año.
- **Valoración en tiempo real:** Se define el precio de energía en tiempo real en función del precio de mercado (que generalmente es horario). Para las aplicaciones de esta tarifa, los proveedores informan a los clientes del pronóstico de costos marginales y luego se realiza la valoración según los precios del sistema real.
- **Tarifas críticas:** Este programa propone una tarifa diferenciada solo para los momentos críticos del sistema (como condiciones climáticas extremas o precios muy altos de energía). Los clientes se someten a tarifas planas más bajas que lo normal, pero en un número definido de eventos críticos el precio sería mucho mayor.

2.2.3. Ejemplos internacionales

Estados Unidos

Estados Unidos es el país que lidera el desarrollo de programas de DR a nivel mundial, dada su gran penetración en los Estados más importantes. Esto ha sido impulsado tanto por los proveedores de energía como por los operadores de los sistemas. A continuación se describen brevemente las situaciones de California y Texas, los dos Estados más grandes.

En California, empresas como *Pacific Gas&Electric Company* (PG&E), *San Diego Gas&Electric Company* (SDGE) y *Southern California Edison* (SCE) han desarrollado amplios programas de DR tanto para clientes industriales y comerciales como residenciales. Por ejemplo, el programa de “SmartAC” para controlar el aire acondicionado a clientes residenciales durante los días más calurosos del verano, asegurándoles un rango de temperatura adecuado y remuneración por participar del programa. También está el “Peak Day Pricing” que reduce el precio de electricidad en el verano a cambio de aumentarlo en los días más calurosos. Otro ejemplo es el “Base Interruptible Program” donde se solicita a los clientes que reduzcan su

consumo con 30 minutos de anticipación a cambio de primas mensuales por unidad reducida. Existen otros ejemplos como “Scheduled Load Reduction Program”, “Optional Binding Mandatory Curtailment”, “Agricultural and Pumping Interruptible Program”, que demuestran la variedad que ofrecen los proveedores en California.

En Texas, su operador ERCOT ofrece programas para que los usuarios finales ofrezcan reducciones de consumo en respuesta a precios de mercado. Recursos de cargas controlables están habilitados para participar en el mercado de reserva en frío¹. Asimismo, los proveedores de energía como *CPS Energy*, *American Electric Power*, *Austin Energy Company*, *Center-Point Energy Company*, *El Paso Electric Company*, *Oncor Company* y *TXU Energy*, ofrecen programas de DR similares a los de California. Se destaca el programa "Free Nights or Weekends", donde se permite el consumo totalmente gratuito en fines de semana o durante la noche, a cambio de tarifas considerablemente más altas durante los días de semana u horas de día.

Otros países

Otro exponente de los programas de DR a nivel mundial es el Reino Unido. Esto se debe principalmente a las políticas del RU para incorporar energías renovables a su matriz. La distribuidora *UK Power Networks Company* en el marco de su proyecto “Low Carbon London” está trabajando con las empresas *Flexitricity*, *EDF Energy* y *EnerNOC* para suscribir clientes en sus programas de DR con el fin de reducir el nivel de demanda en las horas punta.

En Australia, tanto proveedores de electricidad como autoridades han desarrollado estrategias para potenciar la adopción de DR. En 2012 la Comisión del Mercado de Energía de Australia (AMEC), anunció una reforma que permitía la participación de recursos de gestión de demanda en el mercado de energía. Esto incentivó a los proveedores de energía quienes rápidamente incrementaron sus programas de DR. De esta forma, empresas como *Ausgrid Company*, *AusNet Services Company*, *Endeavour Energy*, *Energex Company* y *SA Power Networks*, han estado desarrollando estrategias de DR.

En Asia, Singapur es uno de los exponentes en desarrollar estrategias de DR. La Autoridad del Mercado de Energía (EMA) ha introducido programas de DR para promover la competencia en su sistema. A este pueden acceder directamente consumidores que tengan capacidad de ofrecer una reducción de carga mayor a 0.1 MW durante por lo menos media hora, o bien pequeños consumidores mediante un agregador de demanda. A los consumidores que accedan a este programa se les ofrece un tercio de los ahorros por su recorte, con un tope de 4,500 USD/MWh. [10]

En Sudamérica el líder en incentivar programas de DR es Brasil. En Noviembre de 2017 la Agencia Nacional de Energía Eléctrica (ANEEL), publicó un programa para que los consumidores industriales pueden ofrecer reducciones de demanda. Esto se hace mediante un sistema de subasta, al igual que en la generación, en que los consumidores pueden ofrecer un programa de disponibilidad de reducción semanal, el cual se puede ajustar de forma diaria con un día de anticipación. [11]

¹Capacidad de reserva de generación que puede entrar en operación ante una contingencia en el sistema en unos pocos minutos.

2.2.4. Gestión de demanda en Chile

En Chile se realiza gestión de demanda mediante la definición de horas de punta. Sin embargo, la legislación no es clara respecto a los horarios de punta y coexisten dos definiciones en los reglamentos. El estudio de los efectos de horario de punta en Chile publicado en mayo de 2019 en el sitio web de la CNE [12], realiza un extenso análisis de este mecanismo y su implementación.

En el estudio se concluye que el concepto de “Horas de punta” u “Horario de punta” no se encuentra definido en forma explícita en la LGSE ni en sus reglamentos. Se deduce de la LGSE que por horario de punta debe entenderse aquellas horas donde se presenta la demanda máxima anual del sistema o demanda de punta. Sin embargo, en la normativa hay dos períodos horarios en que se considera que se presentan las demandas máximas anuales:

- Para efectos del Balance de Transferencias de Potencia entre generadores, que considera como período de punta todas las horas del año, y como Demanda de Punta el promedio de las mayores 52 demandas horarias anuales.
- Para efectos de fijar tarifas a la demanda regulada, que considera como horario de punta 5 horas diarias en los meses comprendidos entre abril y septiembre en el SEN CentroNorte y Sur, y todos los meses del año en el SEN Norte Grande, y como Demanda de Punta el promedio de las 52 demandas máximas leídas durante las horas de punta de los últimos 12 meses. La demanda máxima leída corresponde al más alto valor de las demandas integradas en períodos sucesivos de 15 minutos.

Así entonces, la normativa vigente mantiene señales de horario de punta diferentes para el mercado de la suficiencia del sistema y para el mercado de suministro a la demanda regulada.

Luego de estudiar la normativa, en el informe antes mencionado se analizan los efectos del horario punta en los clientes del sistema. De este cabe destacar las siguientes conclusiones:

- Los mecanismos de gestión de demanda en el período de control de horas de punta, aparentemente no son utilizados en el SEN-SING y se incrementan en la medida que los consumidores se localizan hacia el sur del Sistema, alcanzando los mayores valores porcentuales de gestión de demanda de punta en el sistema comprendido entre Ciruelos 220kV y Chiloé 220 kV (denominado S3).
- El aumento de gestión de demanda en el subsistema antes mencionado (S3), se explica fundamentalmente por la migración de la condición de cliente regulado a cliente libre.
- El cambio de horario (verano e invierno) solo afecta al perfil de consumo de los clientes regulados. Esto se debe a los cambios en los patrones de consumo de los clientes residenciales. Su efecto para el conjunto del sistema se diluye al considerar la totalidad de los clientes libres y regulados.
- El tamaño de un cliente libre no se relaciona con su porcentaje de reducción de consumo en horas de punta, sino más bien en el tipo de actividad desarrollada. El efecto se concentra en los clientes que tienen procesos industriales (agua, alimentos, cemento, papel y celulosa, manufactura, petroquímica, portuario, químicos y siderurgia), dependiendo principalmente de la decisión de optimizar su costo energético por parte del propietario del establecimiento y de las posibilidades que da la naturaleza del proceso productivo asociado

- A excepción del SEN-SING, la demanda máxima en el período de control de horas de punta es menor que fuera de ellas, principalmente en los subsistemas SIC-Norte y SIC-Sur, por la gestión de demanda que realizan los clientes industriales en ambos subsistemas.

Se demuestra que existen reducciones considerables de demanda en Chile debido a la gestión que hacen los consumidores. Esto lleva a la idea de que un mejor mecanismo que el de horario punta podría tener un efecto considerable en el funcionamiento del sistema.

2.3. Técnicas de data science y machine learning

En esta sección se describen brevemente las técnicas de data science y machine learning usadas para la elaboración del modelo. Si el lector necesita profundizar en los temas, se recomienda visitar las referencias asociadas.

2.3.1. Clusterización

El término clusterización es un anglicismo derivado de “clustering”, que se refiere a agrupar un conjunto de objetos según sus características. En palabras técnicas, clusterización podría definirse de la siguiente forma: dada una representación de n objetos, se encuentran K clusters (grupos) basados en una medida de similitud, tal que la similitud entre objetos del mismo cluster sea alta, y que entre objetos de distinto cluster sea baja.[13] Un ejemplo gráfico de clusterización se muestra en la figura 2.1, que fue extraída de [14].

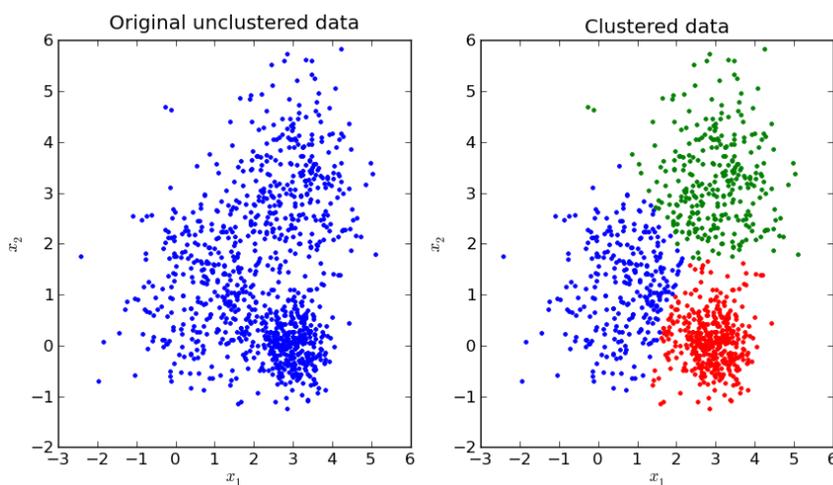


Figura 2.1: Ejemplo de clusterización de datos. En el gráfico de la izquierda se muestran los datos sin clasificar y a la derecha se colorean según el cluster asignado por un algoritmo de clusterización.

Existen muchos algoritmos de clusterización, cuya elección depende del problema y la solución que se quiere alcanzar [13]. Dentro de los más famosos se encuentra K-Means y DBSCAN. El primero recibe como entrada el número de clusters en que se quiere segmentar el conjunto inicial y luego este los agrupa según distancia. Por el contrario, el segundo no

recibe como parámetro el número de clusters deseado, si no que el tamaño de vecindad en términos de distancia y el mínimo número de puntos en una vecindad para su inclusión en un cluster, con lo que determina las distintas agrupaciones.

Algunos conjuntos de datos son demasiado grandes para clusterizar o bien su información no está separada de forma tan explícita, por lo que requiere hacerse un tratamiento de datos previo. Para esto existen técnicas como las que se ven en la siguiente sección.

2.3.2. Principal component analysis (PCA)

Este es un algoritmo de reducción de dimensionalidad mediante cálculo de covarianza y una transformación lineal. Es decir, se reduce la cantidad de datos, conservando aquellos que representan de mejor forma la variación dentro de la muestra.

A continuación, se resumen los pasos que realiza un algoritmo de pca. Para esto se contempla un conjunto de datos compuesto por N muestras (por ejemplo clientes del sistema) y M dimensiones (por ejemplo, su consumo horario en un año). Para mayor detalle se recomienda consultar [15]

- i **Estandarización:** Se realiza una estandarización de cada dimensión para que las variables puedan ser comparables (es distinto comparar una variable que va de 0 a 1, que una de 0 a 100).
- ii **Matriz de covarianza:** Se calcula la matriz de covarianza del conjunto de datos. Esto resulta en una matriz de $M \times M$.
- iii **Vectores y valores propios:** Se calculan los vectores y valores propios de la matriz de covarianza. Los valores propios indican cuánta información de varianza contiene su vector propio asociado respecto a las demás componentes. De esta forma se ordenan los valores propios desde los que contienen mayor información a menor. Luego se selecciona un número p de componentes a preservar, que indicará la nueva dimensionalidad del conjunto. Es decir, si antes tenía M dimensiones, ahora se tienen p .
- iv **Transformación:** Con los vectores propios asociados a las p componentes escogidas se genera una matriz de transformación. Luego se realiza la transformación lineal al dataset inicial estandarizado.

2.3.3. t-Distributed Stochastic Neighbor Embedding (t-SNE)

El método t-SNE es usado para reducir la dimensionalidad de un conjunto de datos. Este fue publicado por Van der Maaten et al. en el año 2008 [16]. El mismo autor da una cátedra acerca de la técnica en una “Google Tech Talk” el año 2013, la cual está disponible en Youtube [17]. Basado en la publicación original, la charla y los artículos [18] [19], se realiza a continuación una descripción del procedimiento.

Lo que hace t-SNE es proyectar los datos de una dimensión alta a una dimensión menor, de forma que se mantengan los cluster (subconjunto de datos similares) de la dimensión alta. Este procedimiento puede resumirse en los 3 pasos descritos a continuación.

i **Similitud entre puntos en el plano de alta dimensión:** Se toma un punto (x_i) del espacio de alta dimensión y se computa una distribución Gaussiana centrada en este. Luego se mide la densidad de cada punto bajo la Gaussiana (x_j) y se normaliza para todos los puntos. Esto se repite tomando como referencia (centro de Gaussiana) a todos los puntos en el conjunto, obteniendo para cada par i, j una probabilidad p_{ij} de ser similares. En la ecuación 2.1 se muestra la fórmula descrita en este paso. La cantidad de puntos que engloba la Gaussiana depende del parámetro de perplejidad (que es definido por el usuario) y que se ve reflejado en σ_i .

$$p_{ij} = \frac{\exp(-\|x_i - x_j\|^2 / 2\sigma_i^2)}{\sum_{k \neq l} \exp(-\|x_k - x_l\|^2 / 2\sigma_i^2)} \quad (2.1)$$

ii **Similitud entre puntos en el plano de baja dimensión:** Los puntos en el plano de baja dimensión son posicionados de forma aleatoria. A estos se les computa la probabilidad de manera muy similar. Se sigue el mismo procedimiento que el paso anterior, pero en lugar de ocupar una distribución Gaussiana se usa una distribución t-student con un grado de libertad. Luego la probabilidad de semejanza en el espacio de baja dimensión se calcula mediante la fórmula expuesta en 2.2.

$$q_{ij} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq l} (1 + \|y_k - y_l\|^2)^{-1}} \quad (2.2)$$

iii **Optimización:** Finalmente se busca que el conjunto de probabilidades del espacio de baja dimensión refleje aquellas de la alta dimensión de mejor manera. Para esto se reubican los puntos en el espacio de baja dimensión, con tal de minimizar la diferencia entre las probabilidades en ambas dimensiones. Esta diferencia se representa usando la divergencia de Kullback-Liebler, donde la función queda como se muestra en la ecuación 2.3.

$$KL(P||Q) = \sum_i \sum_{j \neq i} p_{ij} \log\left(\frac{p_{ij}}{q_{ij}}\right) \quad (2.3)$$

Capítulo 3

Metodología

3.1. Modelo de operación con demanda flexible

3.1.1. Descripción general

El modelo consiste en una minimización de costos de operación del sistema considerando una demanda variable. Este se considera lineal y se desprecian las restricciones de operación asociadas a variables enteras. Aún así, estas podrían ser incorporadas eventualmente. Cabe destacar que el modelo es genérico, no se acota a un sistema particular.

Para describir el modelo se considera una resolución horaria, que no existe almacenamiento (salvo embalses), que no hay pérdidas por transmisión, que la capacidad de generación cambia de forma mensual y es un flujo DC. Aún así esto es sencillo de editar, agregando restricciones y/o cambiando los conjuntos de valores.

Los conjuntos que representan el tiempo se exponen a continuación.

- $T := \{\text{Horas del año}\}$
- $Tm(m) := \{\text{Subconjunto de } T \text{ que diferencia las horas del mes } m\}$
- $Td(d) := \{\text{Subconjunto de } T \text{ que diferencia las horas del día } d\}$
- $M := \{\text{Meses del año}\}$
- $H := \{\text{Días del año}\}$

Es importante tomar en cuenta que el período de estudio puede ser pasado o futuro. El primero sirve para tener certeza de algunas variables y simular escenarios que se pudieron haber dado. Por ejemplo, tener la demanda real, generación real de fuentes renovables, etc. Por otro lado, si se quiere estudiar un período futuro, se tienen que usar proyecciones de demanda y de generación renovable. Ambos escenarios son compatibles con el modelo.

3.1.2. Descripción de generación

La generación se descompone en centrales despachables y no despachables. Así, las primeras se modelan según su costo variable, potencia máxima, potencia mínima y disponibilidad

mensual¹. Las segundas se consideran un parámetro, es decir, se debe ingresar la generación real de estas fuentes o las proyecciones de esta. Dependiendo del nivel de detalle que se quiera, se pueden modelar los generadores con más características.

Un caso particular es el comportamiento de los embalses. Estos se consideran centrales despachables cuyo costo variable es 0, pero que tienen disponibilidad mensual fija, por lo que la decisión es las horas del mes en que se despacha. En caso de que el modelo se ambiente en el pasado, la decisión es simplemente sumar la generación real de los embalses para cada mes. De ser futuro, habría que ingresar proyecciones de la energía disponible en cada embalse por mes.

Conjuntos

- $G := \{\text{Generadores despachables}\}$
- $GND := \{\text{Generadores no despachables}\}$

Variables de decisión

- $P_{i,t}$: Generación de central i en instante t

Parámetros

- $C_{i,t}$: Costo variable de la central i en el mes del instante t
- $Pmax_{i,m}$: Potencia máxima de generador i en mes m
- $Pmin_{i,m}$: Potencia mínima de generador i en mes m
- $E_{j,t}$: Energía inyectada por generador no despachable j en instante t
- $Emb_{k,m}$: Energía máxima disponible para embalse k en mes m

3.1.3. Descripción de transmisión

La topología de la red se describe mediante un conjunto de barras y líneas que las unen. Estas conforman un grafo direccionado, en que los arcos se componen de pares de nodos cuya primera componente es el nodo de partida y la segunda el de llegada. Por ejemplo, la línea que une a la barra $b1$ con la barra $b2$, se representa como $(b1,b2)$ e indica que la dirección del flujo se considera positiva si va desde $b1$ a $b2$.

La conexión de los generadores y los consumos a las barras se definen mediante subconjuntos de los generadores y consumos, respectivamente. Por ejemplo, se define el subconjunto $G(b)$, que contiene a todos los generadores despachables conectados a la barra b .

Conjuntos

- $B := \{\text{Barras del sistema}\}$
- $L := \{\text{Líneas de transmisión}\}$

¹Si se considera que una central no ha entrado en operación en un mes, se modela fijando su potencia máxima en cero

VARIABLES DE DECISIÓN

- $F_{t,l}$: Flujo por línea l en instante t
- $\theta_{t,b}$: Ángulo de voltaje en p.u. de barra b en instante t

PARÁMETROS

- $F_{max,l}$: Flujo máximo de línea l
- X_l : Impedancia de línea l

3.1.4. Descripción de demanda

La flexibilidad que tenga la demanda depende de la flexibilidad de cada uno de los clientes. Aún así, en el sistema existen claros patrones de consumo debido a procesos definidos, por lo que grupos de clientes de similares características podrían tener una flexibilidad parecida. Es por esto que el eje central de este trabajo es determinar estas agrupaciones (que tienen flexibilidad parecida) y dar luces de cómo cuantificar la flexibilidad para incorporarla en este modelo.

A diferencia de los modelos clásicos que consideran una única demanda por nodo, en este se propone una descomposición de la demanda en agrupaciones que comparten potencial de flexibilidad. A estas agrupaciones se les denomina clientes genéricos y se componen por la suma de la demanda de todos los clientes reales que la conforman. Por lo tanto, la demanda nodal se compone por la suma de los clientes genéricos de ese nodo.

En este modelo, la demanda considera una componente fija y una variable. La componente fija viene dada por el retiro real (o esperado) del cliente genérico durante el período de estudio. Esta podría variar dentro de los parámetros de flexibilidad de cada cliente genérico. De esta forma se representa la demanda fija del cliente genérico i en la hora t como $\bar{D}_{i,t}$.

La componente variable resulta de definir un rango de aumento y decremento que puede tener cada cliente genérico en cada instante. Este rango se determina como un porcentaje de la componente fija y se le asigna un costo que representa la compensación que el cliente aceptaría por unidad de energía disminuida. Por ejemplo, el cliente i estaría dispuesto a disminuir su consumo hasta en un 20% entre las 20:00 y 22:00, a cambio de C [\$/MWh]. Así, la componente variable decide cuánto aumenta o disminuye la demanda fija, dentro del rango permitido.

El costo de la flexibilidad de los clientes puede variar dentro del rango definido. Por ejemplo, el cliente dispuesto a disminuir su consumo en un 20% entre las 20:00 y 22:00, a cambio de C [\$/MWh], podría reducir un 10% adicional en el mismo horario si se le compensa con $3C$ [\$/MWh]. Así, podrían haber N rangos para cada cliente, determinados por la compensación que están dispuestos a recibir. El porcentaje de reducción y su costo asociado representan la curva de oferta de flexibilidad del cliente.

Cabe destacar que no todos los clientes pueden gestionar sus consumos. Dichos clientes se denominan “no gestionables” y se modelan mediante asignar que su rango de disminución o aumento es 0.

Un caso particular de la respuesta en demanda, son aquellos clientes que pueden desplazar sus consumos. Es decir, clientes que pueden reducir su consumo en ciertas horas del día, pero luego deben aumentarlo en la misma cantidad en otras. En el modelo se especifica aquellos clientes genéricos que desplazan sus consumos, mediante un subconjunto de todos.

Conjuntos

- $R := \{\text{Clientes genéricos}\}$
- $DS := \{\text{Clientes genéricos que desplazan demanda}\}$
- $N(j) := \{\text{Escalones de flexibilidad de cliente genérico } j\}$

VARIABLES DE DECISIÓN

- $D_{j,t}$: Demanda del cliente j en el instante t
- $\bar{\Delta}_{j,t,n}$: Porcentaje de aumento en escalón n de cliente j en instante t
- $\underline{\Delta}_{j,t,n}$: Porcentaje de reducción en escalón n de cliente j en instante t

Parámetros

- $\bar{\zeta}_{j,t,n}$: Costo de aumentar una unidad de energía del cliente j en instante t para escalón n
- $\underline{\zeta}_{j,t,n}$: Costo de reducir una unidad de energía del cliente j en instante t para escalón n
- $\bar{D}_{i,t}$: Demanda real (o esperada) del clientes genérico i en el instante t
- $\bar{\Delta}max_{j,n}$: Porcentaje máximo de aumento en escalón n de cliente j en instante t
- $\underline{\Delta}max_{j,n}$: Porcentaje mínimo de reducción en escalón n de cliente j en instante t

3.1.5. Restricciones del modelo

Balance nodal

La primera ecuación impone que se abastezca la demanda de todos los clientes en todo instante.

$$\sum_{i \in G(b)} P_{i,t} + \sum_{k \in GND(b)} E_{k,t} - V_{t,b} = \sum_{j \in R} \bar{D}_{j,t} + \sum_{l \in O(b)} F_{t,l} - \sum_{l \in I(b)} F_{t,l} \quad \forall (t \in T)(b \in B) \quad (3.1)$$

En caso que la generación no despachable sea mayor que la demanda del nodo y el flujo que sale de este, entonces habrá vertimiento. Para esto se define la variable de decisión $V_{t,b}$, que representa el vertimiento en el instante t en la barra b . Además se definen los conjuntos $I(b)$ y $O(b)$, que son un subconjunto de L y contienen las líneas que entran y salen de b respectivamente. Matemáticamente se puede representar de la siguiente manera: $I(b) = \{(x, b) \in L, x \in B\}$ y $O(b) = \{(b, x) \in L, x \in B\}$.

Demanda flexible

La demanda de cada cliente genérico se compone por una componente fija y su porcentaje de desviación. En esta ecuación se asume que los escalones hacia arriba y abajo son simétricos, pero basta con diferenciar el conjunto de escalones si se quiere tener diferencia entre aumento y reducción.

$$D_{j,t} = \bar{D}_{j,t} \left(1 + \sum_{n \in N} (\bar{\Delta}_{j,t,n} - \underline{\Delta}_{j,t,n}) \right) \quad \forall (j \in R)(t \in T) \quad (3.2)$$

Damanda desplazable

Esta restricción impone que los clientes que pertenecen al conjunto de DS, puedan desplazar sus consumos. Es decir, si reducen/aumentan su consumo en una hora del día, tienen que aumentar/reducir la misma cantidad en otro momento.

$$\sum_{t \in Td(d)} D_{j,t} = \sum_{t \in Td(d)} \bar{D}_{j,t} \quad \forall (j \in DS)(d \in H) \quad (3.3)$$

Generación de embalses

El despacho de los embalses está acotado a una producción máxima mensual.

$$\sum_{t \in Tm(m)} P_{i,t} \leq Emb_{i,m} \quad \forall (i \in G(Embalses))(m \in M) \quad (3.4)$$

Cotas de variables

Cada escalón de cada cliente está acotado por su máximo valor posible.

$$\begin{aligned} \bar{\Delta}_{j,t,n} &\leq \bar{\Delta}max_{j,t,n} \\ \underline{\Delta}_{j,t,n} &\leq \underline{\Delta}max_{j,t,n} \quad \forall (j \in R)(t \in T)(n \in N) \end{aligned} \quad (3.5)$$

Cada generador tiene restricciones de operación máxima y mínima.

$$Pmin_{i,t} \leq P_{i,t} \leq Pmax_{i,t} \quad \forall (i \in G)(t \in T) \quad (3.6)$$

El flujo de potencia por las líneas está acotado a la capacidad de la línea.

$$\begin{aligned} F_{t,l} &\leq Fmax_{t,l} \\ -F_{t,l} &\leq Fmax_{t,l} \quad \forall (l \in L)(t \in T) \end{aligned} \quad (3.7)$$

Naturaleza de variables

$$D_{j,t}, P_{i,t}, V_{t,b}, \bar{\Delta}_{j,t,n}, \underline{\Delta}_{j,t,n} \geq 0 \quad \forall (j \in R)(i \in G)(t \in T)(b \in B)(n \in N) \quad (3.8)$$

3.1.6. Función objetivo

Finalmente la función objetivo consiste en minimizar los costos de operación de las centrales y las compensaciones a los clientes por su flexibilidad.

$$\min \sum_{t \in T} \left(\sum_{i \in G} C_{i,t} \cdot P_{i,t} + \sum_{j \in R} \bar{D}_{j,t} \left(\sum_{n \in N} (\bar{\zeta}_{j,t,n} \cdot \bar{\Delta}_{j,t,n} + \underline{\zeta}_{j,t,n} \cdot \underline{\Delta}_{j,t,n}) \right) \right) \quad (3.9)$$

En el anexo se puede encontrar el modelo completo.

3.2. Datos de balance

Los datos usados en este trabajo provienen de la base del balance definitivo (detallado en 2.1.5). En particular, se ocupan los consumos horarios de los clientes libres, indicando la empresa generadora a cargo del retiro y la barra Troncal en que se conecta. A esta información se le agrega el tipo de día del retiro, es decir, si es día de semana, sábado o domingo/feriado. Los días lunes se diferencian de los días de semana, por lo que se define el día de “trabajo” de martes a viernes. Estos datos, desde septiembre de 2018 a agosto de 2019 conforman la base a utilizar.

Desde 2017 los clientes libres que aparecen en las bases han aumentado de forma exponencial. El gráfico de la figura 3.1, muestra el conteo de nombres distintos que aparece en cada base de balance mensual publicada desde enero de 2011 hasta agosto de 2019. Para hacer el estudio lo más robusto posible, es ideal contar con la mayor cantidad de clientes distintos, que dispongan de la mayor cantidad de datos. Es por esto que se decide definir el año de estudio lo más actualizado que se pueda.

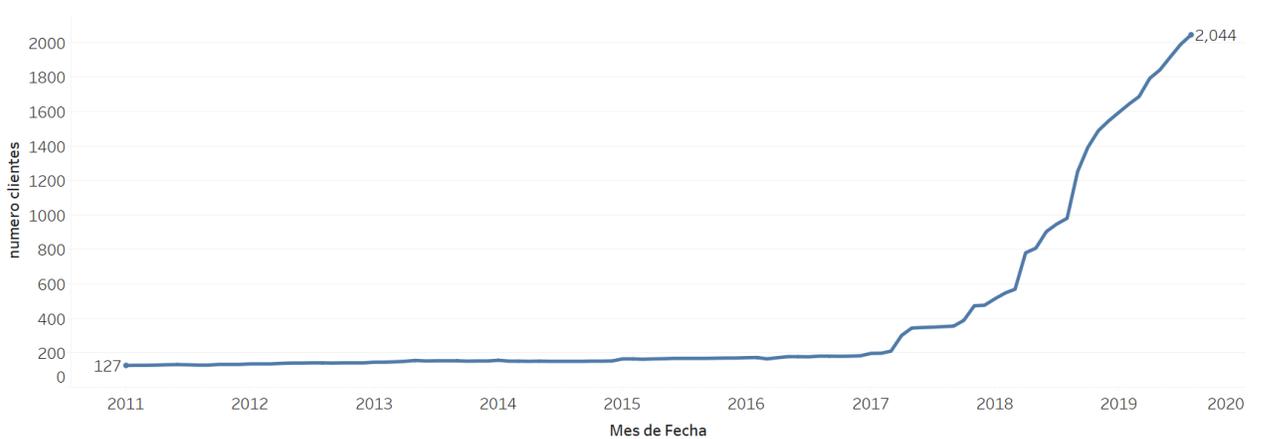


Figura 3.1: Evolución de clientes libres en bases de balance mensual de inyecciones y retiros, desde enero de 2011 hasta agosto de 2019.

3.3. Corrección de nombres

Uno de los problemas más comunes con las bases de datos es que se ingresa un mismo cliente con distinto nombre. Por ejemplo, se tiene el cliente ficticio “EMPRESA MINERA”; este puede ser ingresado un mes como “EMPRESA MINERA SA”, luego “EMPRESA MINERA S.A” y en otro como “EMPRESA_MINERA”. A simple vista, un humano puede saber que son el mismo cliente ingresado con distinto nombre, pero para un software estos son clientes distintos.

La primera intervención que se le hace a la base de datos es corregir los nombres de los clientes. Para esto, en una primera iteración se toma una base de clientes y se asigna de forma manual. Luego, con todos los datos, se programa un script en python que toma los nombres originales de los clientes que no han sido asignados y los compara con el listado completo de nombres originales mediante funciones de la librería “fuzzywuzzy”. De esta forma, se definen 3 niveles de puntaje de asignación: asignación automática, en que se asume que son el mismo dado que son muy parecidos; cliente nuevo, en que no hay ningún cliente que sea lo suficientemente parecido, por lo que se asume que es nuevo; asignación manual, en que hay clientes con nombres parecidos, pero necesitan una revisión. Luego de la revisión de estos clientes se crea un archivo que sirve como diccionario, en que a cada cliente que aparece en la base de datos, se le asigna un nombre corregido.

Habiendo corregido los nombres, es importante destacar que hay empresas que tienen retiros en distintas partes, identificados por el mismo nombre. Estos deben considerarse como retiros distintos, puesto que puede afectar la ubicación y/o el tipo de trabajo que se hace en cada uno. Por ejemplo, un retiro de Walmart asociado a un gran supermercado de la región de Antofagasta puede tener retiros muy distintos que un supermercado express de Chiloé. Por lo tanto, se usa la barra troncal de conexión para diferenciar los retiros de un mismo cliente.

3.4. Asignación de sectores

La primera característica de clasificación que se desea agregar a los clientes es el sector económico en que se desenvuelven. Para esto se ocupan los distintos sectores según define el Servicio de Impuestos Internos (SII). En la tabla 3.1 se muestran los sectores junto con un mnemotécnico asignado.

Sector	Mnemotécnico
Primario	
Agricultura, ganadería, caza y silvicultura	AGCS
Pesca	PESC
Explotación de minas y canteras	MINE
Secundario (Industrial)	
Industrias manufactureras no metálicas	IMNM
Industrias manufactureras metálicas	IMM
Suministro de electricidad, gas y agua	SEGA
Construcción	CONS
Terciario (Servicios)	
Comercio al por mayor y menor	COME
Hoteles y restaurantes	HORE
Transporte, almacenamiento y comunicaciones	TRAC
Intermediación financiera	FINA
Act. inmobiliarias, empresariales y de alquiler	AIEA
Adm. pública y defensa	ADPD
Enseñanza	ENSE
Servicios sociales y de salud	SOSA
Otras actividades de servicios	OTRO
Consejo de adm. de edificios y condominios	EDCO
Organizaciones y órganos extraterritoriales	ETER

Tabla 3.1: Tabla de sectores definidos por Servicio de Impuestos Internos y mnemotécnico asignado a sector.

La asignación de sector se hace mediante web-scraping ². Para esto se revisan las páginas web “genealog.cl” (que contiene el rubro de la empresa según su nombre y RUT) y “mercantil.cl” (que contiene nombre y rut de empresas). La automatización se realiza mediante un script de python que ocupa la librería “Selenium”. En el diagrama de la figura 3.2, se muestra la estructura de cómo funciona el script.

²Técnica que sirve para extraer información de páginas web de forma automatizada. La traducción del inglés sería “escarbar una web”.

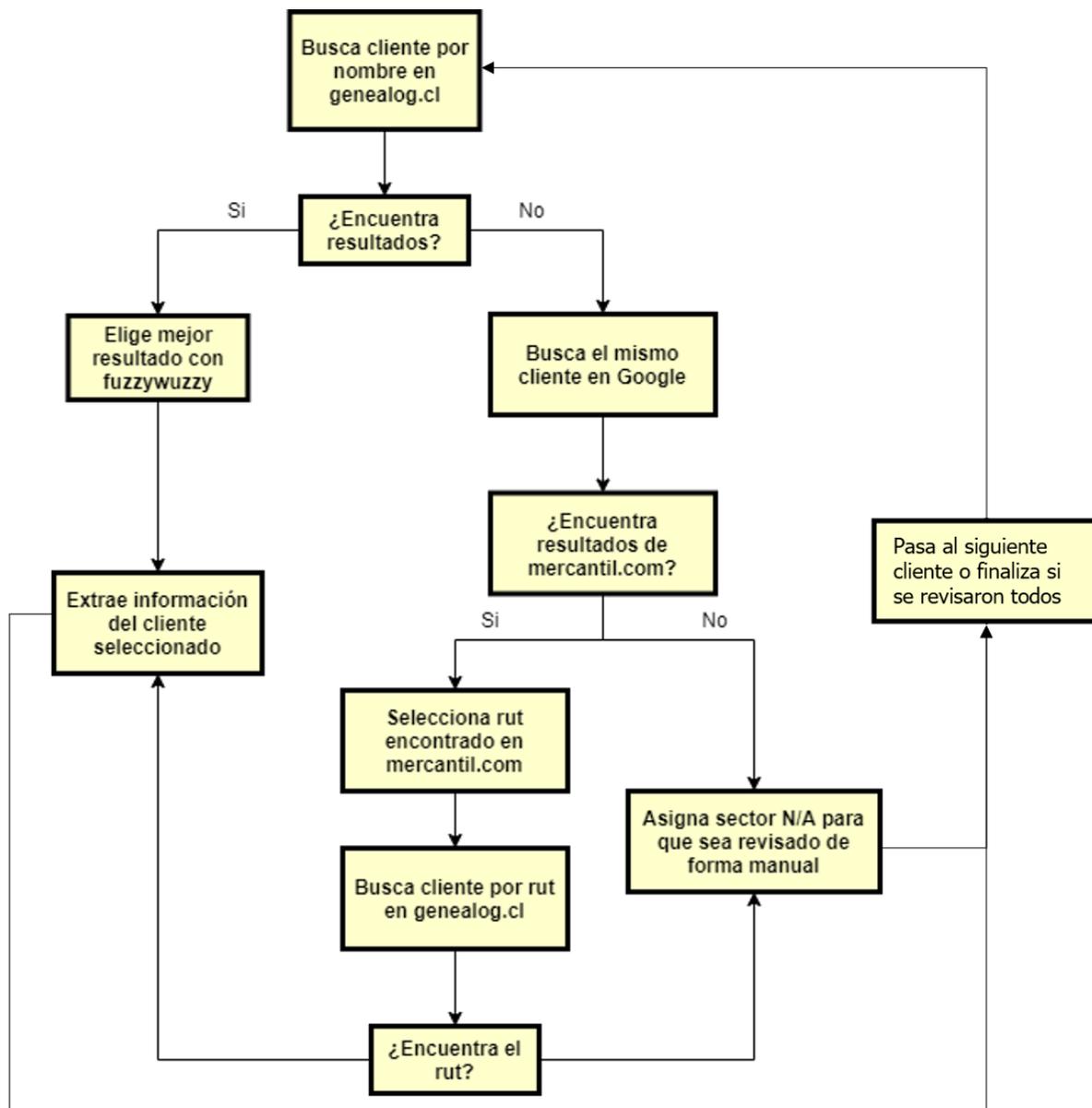


Figura 3.2: Diagrama de flujo que describe el script de python que realiza la asignación automática de sector económico.

3.5. Clasificación por perfil de consumo y sector

Los datos considerados para la clasificación, son el sector económico al que pertenecen los clientes y perfiles de consumo de los retiros según semestre y tipo día. Es decir, se considera un semestre de verano (definido como los meses en que no existe el cargo por horario de punta, de octubre a marzo) y semestre de invierno (abril a septiembre). También se consideran los distintos tipos de día que típicamente inciden en el consumo de un cliente, es decir, lunes hábil, trabajo (martes a viernes hábil), sábado y domingo (o feriado). Los datos de consumo horario para cada cliente se agrupan según el tipo día y tipo semestre mediante el promedio. Finalmente el consumo se normaliza dividiendo por el máximo valor alcanzado en la agrupación mencionada anteriormente.

En la figura 3.3, se muestra el perfil de consumo normalizado por tipo semestre y tipo día, al que se desea llegar. Esta información, en conjunto con el sector económico, permiten realizar la clasificación para cada cliente. A continuación se detalla la metodología para clasificar.

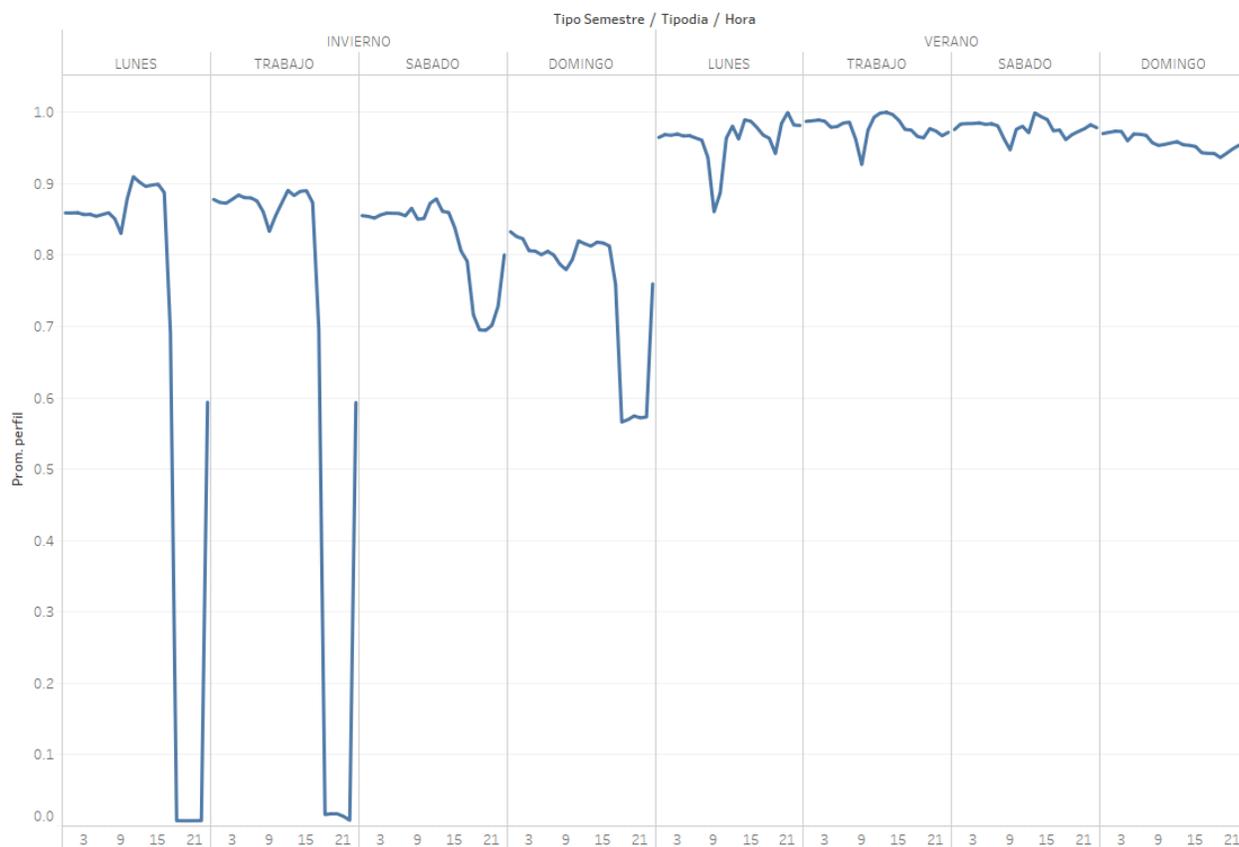


Figura 3.3: Perfil de consumo normalizado por tipo semestre y tipo día de un cliente del balance

3.5.1. Manejo de outliers

Los datos de energía en las bases no siempre reflejan el perfil de consumo de los clientes. Esto se debe a problemas con los medidores, cortes de energía o problemas humanos por parte de la empresa a cargo del retiro. Es por esto, que para tener una muestra fidedigna, el primer paso es limpiar los datos erróneos o outliers. A continuación se identifican los tipos de outliers y se detalla la forma en que se manejan.

Cantidad de información

Durante los meses de estudio, empiezan a aparecer muchos clientes libres nuevos de los cuales no se tiene información completa (por ejemplo, un cliente que aparece en noviembre de 2018, no tiene datos de septiembre y octubre del mismo año). Al tener menos meses completos se distorsiona el promedio de la agrupación por tipo día y tipo semestre, puesto que algunos clientes incorporan variables propias de meses que otros no tienen. Por ejemplo, que en días de septiembre existan tratos especiales debido a las fiestas patrias, lo cual hace

que los clientes varíen su perfil de consumo. Lo anterior es igual para todos, siempre y cuando todos tengan la información de septiembre. De esta forma este comportamiento incide a la hora de armar los perfiles promedio, lo cual es un error que debe ser minimizado.

El primer filtro se realiza con el fin de minimizar el error asociado a clientes con datos incompletos. Para esto, se eliminan todos los clientes que tengan algún dato vacío desde noviembre de 2018 a agosto de 2019. Es importante destacar, que es distinto que el dato sea cero a que esté vacío. Si es cero, puede atribuirse a que el cliente no consumió en tal hora, a que fue un corte de electricidad o que fue un error de medición.

Declaraciones de distribuidoras

Hasta mediados de 2018, las empresas distribuidoras agrupaban una gran cantidad de sus clientes libres en distribución como retiros a nombre de la distribuidora (por ejemplo Chilectra Libre). Gracias a una interpelación al Coordinador, se exige a las empresas que entreguen el detalle de estos clientes para conformar el balance. Por lo tanto de a poco, empiezan a disminuir los retiros asociados a empresas distribuidoras.

Si bien las grandes distribuidoras están entregando el detalle de los clientes, existen algunas más pequeñas que siguen ocultándolo. Debido a que estos retiros en realidad no son consumos de clientes particulares sino una agrupación de ellos, no se deben considerar en este estudio. De esta forma, el segundo filtro consiste en eliminar los retiros bajo el nombre de empresas de distribución.

Cantidad de energía

En clientes grandes, errores de medición pueden ser muy perjudiciales debido a que se puede tratar de una suma importante de dinero. Es por esto que las mediciones de estos son más rigurosas que en clientes pequeños. Esto se puede ver en las bases de balance, en que es más probable encontrar consumos pequeños con errores evidentes.

Como el fin de este trabajo es clasificar a los clientes según su perfil de consumo, no se considera la cantidad de energía que retira. Lo anterior implica que se compara de igual forma el retiro de una minera que consume del orden de un TWh al año, con un edificio que consume del orden de MWh anuales. Por lo tanto, tener clientes pequeños con alta probabilidad de ser un outlier, puede contaminar mucho la muestra, lo que haría obtener una clasificación no deseada.

Además existen retiros que claramente no representan consumos asociados a algún proceso productivo con un patrón específico. Esto se puede deber al sistema de medición, como por ejemplo, clientes que no tienen medidor horario y declaran su consumo en mediciones acumuladas. Esto se puede ver en los ejemplos de la figura 3.4. Si bien puede que se muestren consumos horarios de grandes magnitudes, la energía que consumen en el año es baja. Por lo tanto, pueden catalogarse como un cliente pequeño.

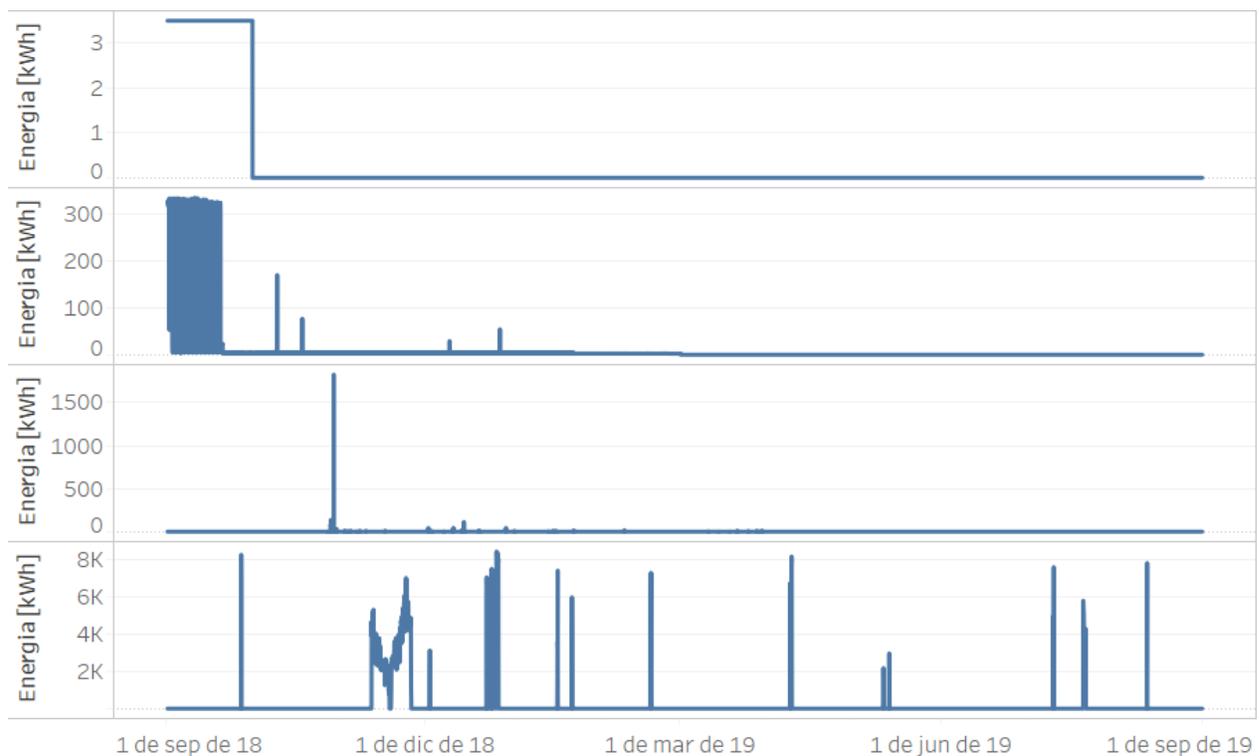


Figura 3.4: Consumo energético horario de 4 retiros que evidentemente no tienen un patrón de consumo asociado a un proceso productivo durante el año de estudio. Estas series son consideradas outliers evidentes.

Por lo anterior, se decide filtrar los retiros según tamaño, eliminando todos aquellos cuyo consumo anual sea menor a 200 MWh. Esta cantidad se decide ordenando los retiros de menor a mayor consumo anual y luego observando el detalle horario (como la figura 3.4). El valor de 200, resulta ser sensato para eliminar outliers, manteniendo la mayor cantidad de clientes completos.

Mediciones horarias erróneas

Como se describe al comienzo del apartado, existen errores horarios asociados a las lectura de los medidores o a la existencia de cortes de suministro. En el primer caso, algunos presentan fallas en que el medidor repite la medida de la hora anterior y acumula la medida real para la hora siguiente. Esto se refleja como un aumento intempestivo en el consumo. Por otro lado, en el caso de los cortes, se pueden detectar como bajadas abruptas que no son habituales. En la figura 3.5, se muestran ejemplos de clientes con datos horarios erróneos.

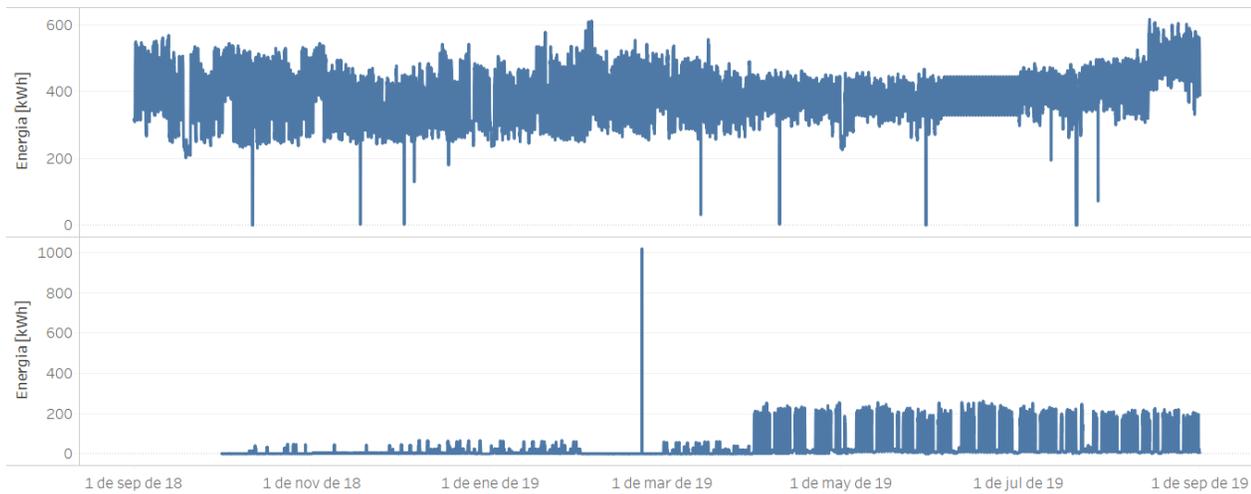


Figura 3.5: Consumo energético horario de año de estudio de 2 retiros que tienen datos puntuales cuyo comportamiento está fuera de lo normal.

Para eliminar este tipo de outlier, se propone hacer un análisis estadístico mensual a la señal de consumo de energía. Esto se realiza definiendo un rango mensual en que se considera aceptable que existan datos. El valor máximo permitido se define como el promedio + 6 desviaciones estándar, mientras que el mínimo como el promedio - 4 desviaciones estándar. Se da más holgura hacia arriba, ya que en las muestras se puede observar que en más casos los cambios intempestivos hacia arriba corresponden realmente a un hábito del cliente. En la figura 3.6, se muestran los mismos retiros que la figura 3.5, pero con sus bandas de referencia.

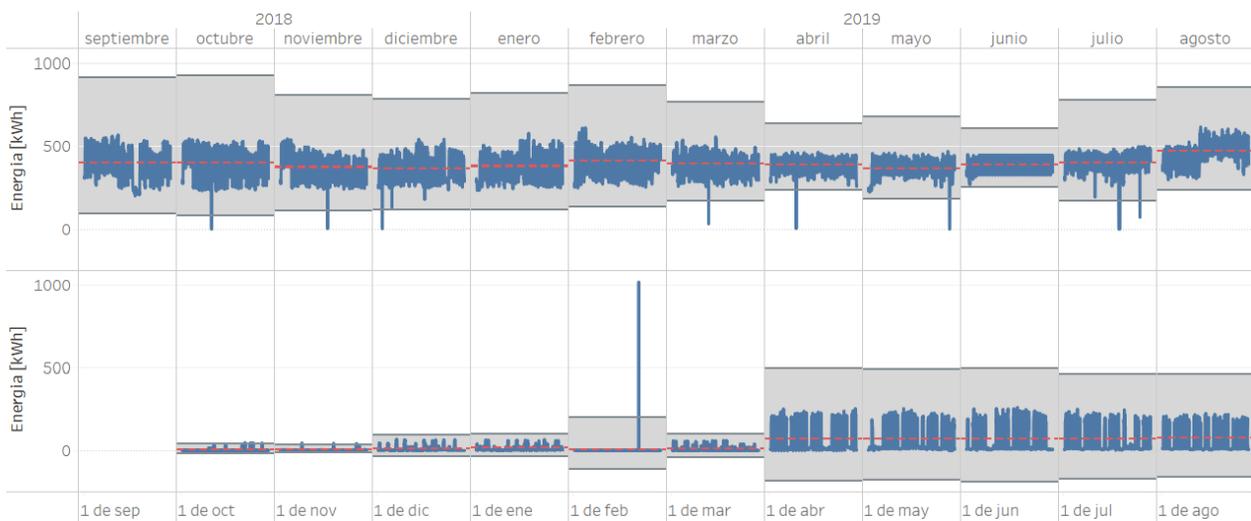


Figura 3.6: Consumo energético horario de año de estudio de los retiros de figura 3.5, con medidas estadísticas para detección de puntos erróneos

En la literatura, la detección de outliers en series de tiempo es un campo ampliamente estudiado, pero requiere de mayor conocimiento. Por temas de simplicidad, se opta por el método previamente expuesto. Queda propuesto mejorar la metodología de manejo de errores horarios mediante un sistema sofisticado para trabajo futuro.

Meses erróneos

La última fuente de errores que se detecta, es que existen meses en que toda la señal tiene un comportamiento muy distinto al resto de la muestra. Por ejemplo, hay meses en que el consumo es una constante, lo que se debe a que hubo problemas con la medición, por lo que se factura asumiendo un promedio del cliente. También, clientes que consumen con un patrón claro, pero inesperadamente un mes su consumo desciende a cero. A su vez, hay retiros que en un mes determinado cambian drásticamente su consumo promedio y/o el perfil. Esto se explica con que las distribuidoras hayan declarado bien su consumo agregado, pero mal la separación por cliente. Los errores expuestos, se pueden ver en la figura 3.7.

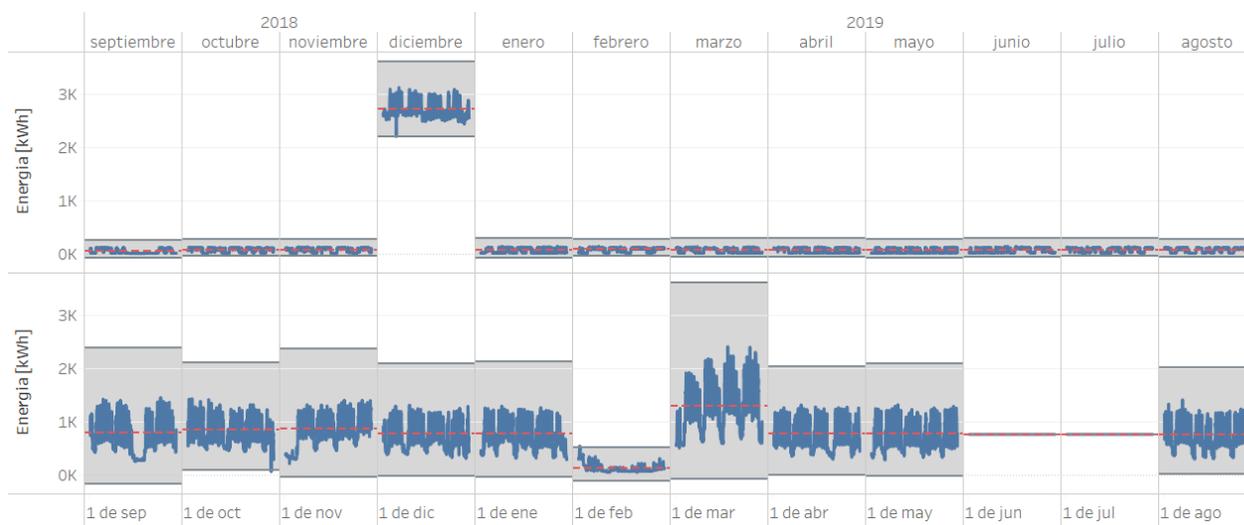


Figura 3.7: Consumo energético horario con filtros estadísticos mensuales de retiros con meses cuyos datos no corresponden al patrón de consumo que exhiben en los otros meses, sin explicación lógica. El gráfico de arriba muestra un cliente cuyo consumo en diciembre es muy distinto a su comportamiento normal y el de abajo un cliente al que le declararon junio y julio como su promedio típico.

Si bien en un inicio, se podría considerar como solución eliminar los meses erróneos con un criterio de promedio y desviación estándar, se descarta debido a la existencia de consumos estacionales. Es decir, un mes sería outlier si el promedio y desviación estándar del mes son muy distintos al resto de la muestra. Por otro lado, existen clientes (principalmente agrícolas) cuyos consumos son estacionales. Es decir, su producción y por tanto, consumo de energía, cambia mucho dependiendo de la estación del año. Si se ocupa el criterio de comparación de promedio y desviación estándar, se eliminarían meses en que los consumos son muy distintos, pero que sí corresponde que lo sean. En la figura 3.8, se muestra un ejemplo de cliente estacional, cuyo consumo en meses de invierno es 0.

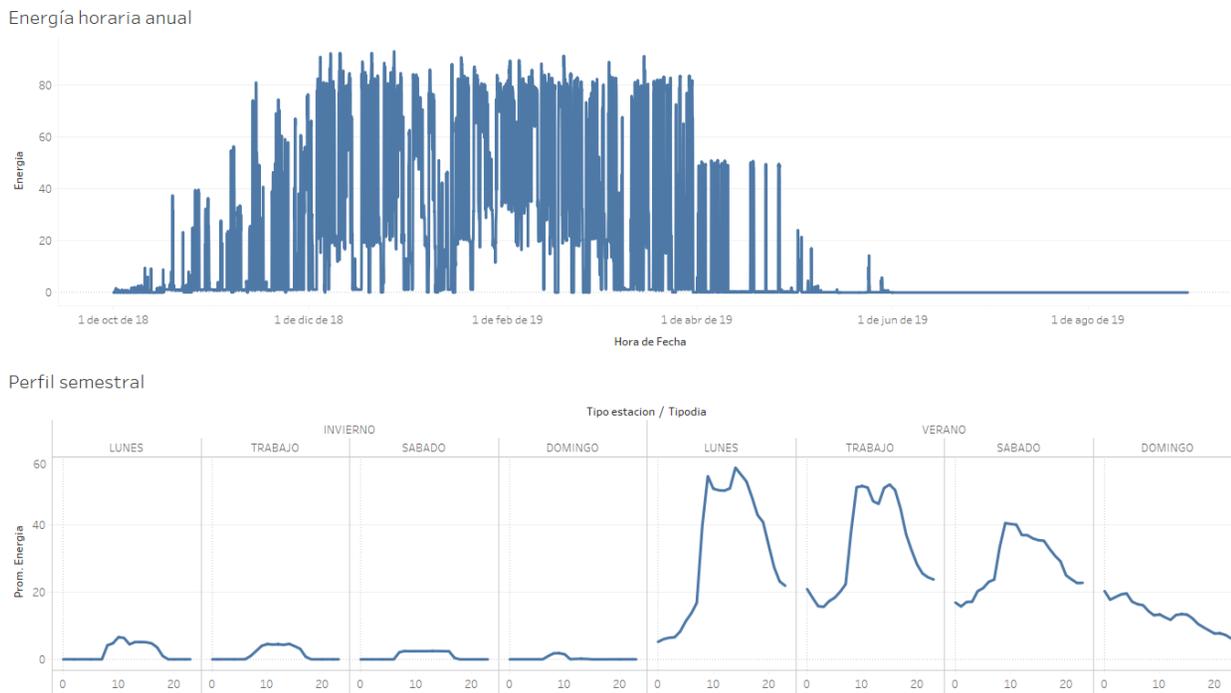


Figura 3.8: Gráficos de un retiro perteneciente a un cliente agrícola, cuyo consumo es estacional. El gráfico de arriba muestra el consumo horario en el año de estudio, mientras que el de abajo se muestran los perfiles típicos que se generan con los datos del gráfico de arriba.

Finalmente, en vista de que los retiros que ya pasaron los filtro anteriores son alrededor de 1500, se decide hacer la eliminación de meses de forma manual. Es decir, se observa la señal horaria de consumo y los perfiles mensuales que generan, además de analizar el rubro al que pertenece y se eliminan aquellos meses de los retiros que claramente son distintos a la muestra y no tienen una explicación intuitiva de por qué pudiesen serlo.

Una de las ideas que surge para automatizar el proceso, es usar el criterio de promedio y desviación estándar mencionado previamente, pero con restricciones que dependen del sector. Esto queda propuesto como trabajo futuro.

3.5.2. Información para clasificación

Una vez filtrada la base de datos, a cada retiro se le asignan 209 atributos correspondientes a los perfiles de consumo y el sector económico. Los primeros 192 son cada hora de los perfiles (24) agrupados por tipo día importante (4) y tipo semestre (2). Estos valores están normalizados, por lo que cada uno de los atributos asociados al perfil de consumo toma valores entre 0 y 1. Luego, se agrega la información del sector económico mediante el método “One-Hot”, es decir, se agregan como atributos los 17 sectores posibles y se indica con un 1 si el cliente pertenece al sector y 0 en caso contrario. Así, cada uno de los retiros que queda en la base es representado en estas 209 dimensiones.

3.5.3. Reducción de dimensiones y clusterización

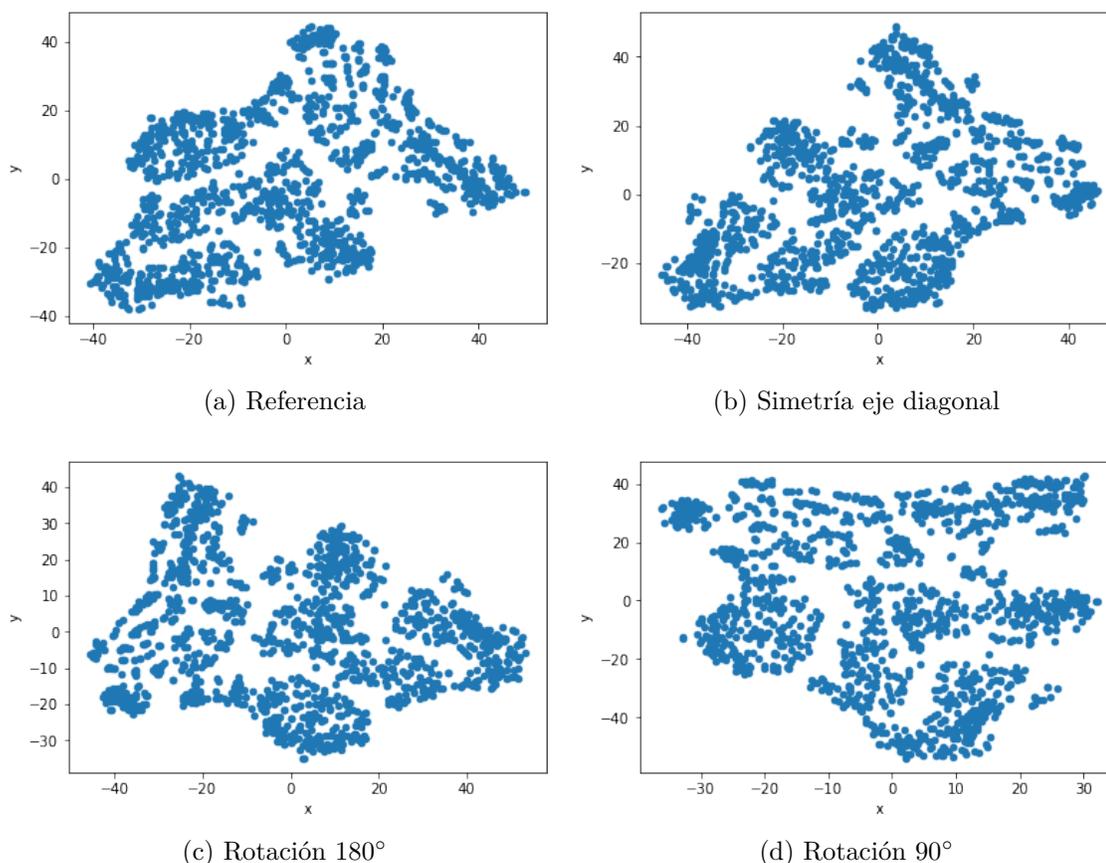


Figura 3.9: Resultados de 4 repeticiones de reducción de dimensionalidad con TSNE, para un mismo set de datos e iguales parámetros. Se aprecia que la estructura es igual, pero tienen distinta orientación espacial. Usando el gráfico (a) como referencia, se obtiene (b) con una simetría con eje diagonal, (c) con una rotación de 180° y a (d) con una rotación de 90° en sentido antihorario.

En este punto, se busca agrupar en clusters a los retiros que estén más cercanos en la representación de 209 dimensiones expuesta en el apartado anterior. Para esto, primero se transforma la representación mencionada a 2 dimensiones mediante TSNE. Esta técnica permite conservar la estructura del plano de alta dimensión y representarla en 2 dimensiones (ver apartado 2.3.3). Así, se pueden exhibir los puntos en un gráfico bidimensional, donde se puede visualizar la estructura de los datos.

Visualizar los datos permite ajustar adecuadamente los parámetros que se requieren para clusterizar y tener una noción de la posición de los retiros en la representación propuesta. Los métodos dbSCAN y kmeans requieren de parámetros de entrada que deben ser ajustados (ver apartado 2.3.1), lo cual resulta intuitivo en la representación bidimensional. Además, las nociones de proximidad espacial de los datos graficados en 2 dimensiones, facilita el análisis de los resultados.

La técnica de clusterización (dbSCAN o kmeans) y sus parámetros correspondientes se eligen

mediante prueba y error. En el caso de dbscan, se considera una mejor representación la que tenga más clusters y menos puntos de ruido. En el caso de kmeans, el número de clusters se ajusta observando las agrupaciones en la representación bidimensional. La comparación entre métodos se hace analizando los perfiles de los retiros agrupados en cada cluster. Es por esto que la etapa de clusterización requiere bastante tiempo, ya que hay que analizar muchos casos.

Cabe destacar que TSNE es una técnica probabilística, por lo tanto la representación bidimensional resultante es distinta en cada iteración. Sin embargo, la estructura es prácticamente la misma, pero con distinta orientación. Por ejemplo, en la figura 3.9 se muestran cuatro resultados de TSNE asociados al mismo conjunto de datos. Si bien todas las representaciones son distintas, se puede observar una estructura clara: tomando como referencia la de (a), se obtiene la de (b) con una simetría con eje diagonal, la de (c) con una rotación de 180° y la de (d) con una rotación de 90° en sentido antihorario. Aún cuando la forma del gráfico es muy distinta, al clusterizar se nota que las agrupaciones son muy parecidas.

3.6. Entrevistas a clientes representativos

Una vez armados y analizado los clusters se busca caracterizarlos según el nivel de flexibilidad que pueden ofrecer. Para esto hay que entender las razones de por qué consumen de la forma en que lo hacen. Esto se logra mediante entrevistas telefónicas a clientes del sistema.

La entrevista busca entender los procesos productivos de los clientes, los horarios en que consumen, la capacidad de flexibilidad que pueden otorgar y su costo. La pauta se muestra en el anexo B. En la práctica, esta pauta sirve como guía, dado que la entrevista consiste más en una conversación.

Los clientes a entrevistar son uno por cada cluster de interés. Este se selecciona según la composición por sectores del cluster: el cliente debe pertenecer al sector más representativo de este. Algunos de los clientes seleccionados tienen más de un retiro, por lo que se les pregunta por aquellos que tengan perfiles pertenecientes a clusters distintos.

Finalmente, mediante el análisis de las respuestas se proponen técnicas para determinar parámetros de flexibilidad, pero estos no se calculan explícitamente.

Capítulo 4

Resultados

4.1. Clusterización de clientes

En esta sección se exponen los resultados desde el manejo de datos hasta confección de los clusters.

4.1.1. Corrección de nombres, confección de retiros y asignación de sectores

El primer procedimiento corresponde a corregir los nombres de las empresas que han sido ingresados al balance de forma distinta, pero en realidad son la misma. Luego se separan las empresas por retiros, asociados a su barra troncal. En la tabla 4.1 se muestra la cantidad de nombres distintos en el balance, la cantidad de clientes luego de la reasignación y la cantidad de retiros. Se puede ver que los nombres recorregidos implican una reducción de 461 clientes. Luego el número de retiros asciende por aquellos clientes que tienen consumos conectados en más de una zona.

Clientes en balance	2142
Clientes post corrección de nombres	1681
Retiros	2100

Tabla 4.1: Cantidad de nombres distintos en el balance de clientes luego de la reasignación y de retiros

A estos 1681 retiros se les asigna el sector económico mediante web-scraping. El número de clientes pertenecientes a cada sector y la cantidad de energía que concentran todos sus retiros se exhiben en la tabla 4.2. En esta se muestran los sectores ordenados según energía, de forma descendente.

Sector	Clientes	Retiros	Energía (GWh/año)
MINE	106	120	24,725
IMNM	399	475	4,519
COME	159	291	2,795
AGCS	308	355	2,556
SEGA	72	104	1,468
IMM	80	86	1,350
TRAC	54	69	935
AIEA	80	104	804
OTRO	79	80	696
FINA	55	75	629
PESC	55	67	373
CONS	45	54	234
SOSA	33	40	215
EDCO	90	92	187
HORE	35	39	123
ENSE	25	42	90
ADPD	6	7	50
Total	1681	2100	41,748

Tabla 4.2: Número de clientes, retiros y energía anual de cada sector económico.

En la figura 4.1, se grafica el número de retiros y la energía anual por sector asociados a la tabla 4.2.

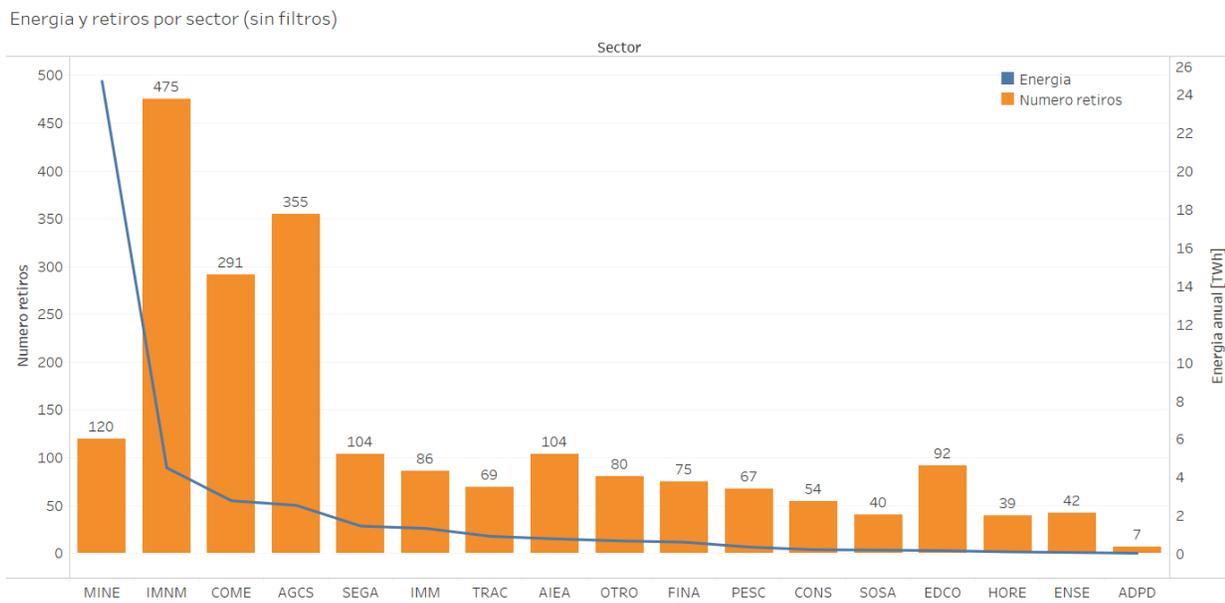


Figura 4.1: Gráfico de número de retiros y energía anual por sector asociado a datos de tabla 4.2.

4.1.2. Limpieza de datos

En este apartado se cuantifican los datos que son filtrados de la base original. Estos filtros eliminan retiros completos, meses particulares de ciertos retiros y horas particulares.

En la figura 4.2 se muestra la cantidad de retiros que fueron removidos según su criterio de eliminación. Los criterios de eliminación son aplicados de forma secuencial, por lo que se muestran en el orden de aplicación. Los clientes que cumplen con más de un criterio, se cuentan solamente en el primero con que cumplen de la secuencia.

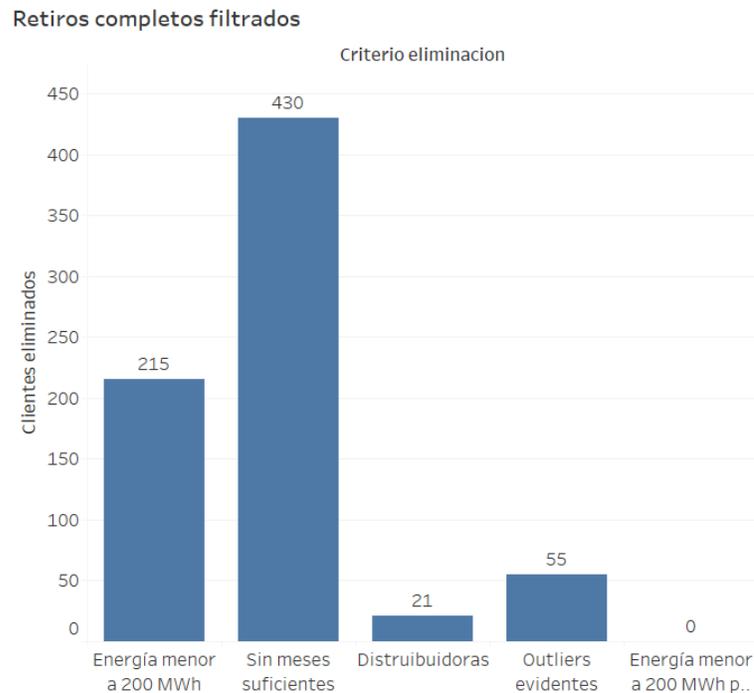


Figura 4.2: Cantidad de retiros removidos, según su criterio de eliminación.

El siguiente filtro aplicado es el que elimina los meses calificados como anómalos. En la figura 4.3 se muestra el número de retiros a los cuales se le extrajo un mes completo según mes. Se aprecia que el mes de diciembre es considerablemente el que más veces se quita de la serie de retiros. La mayoría de los retiros que tiene mal el mes de diciembre se ubican en la zona de concesión de Enel Distribución. Por lo tanto, se deduce que fue la distribuidora la que presentó problemas con sus medidas en ese período.

Meses completos filtrados

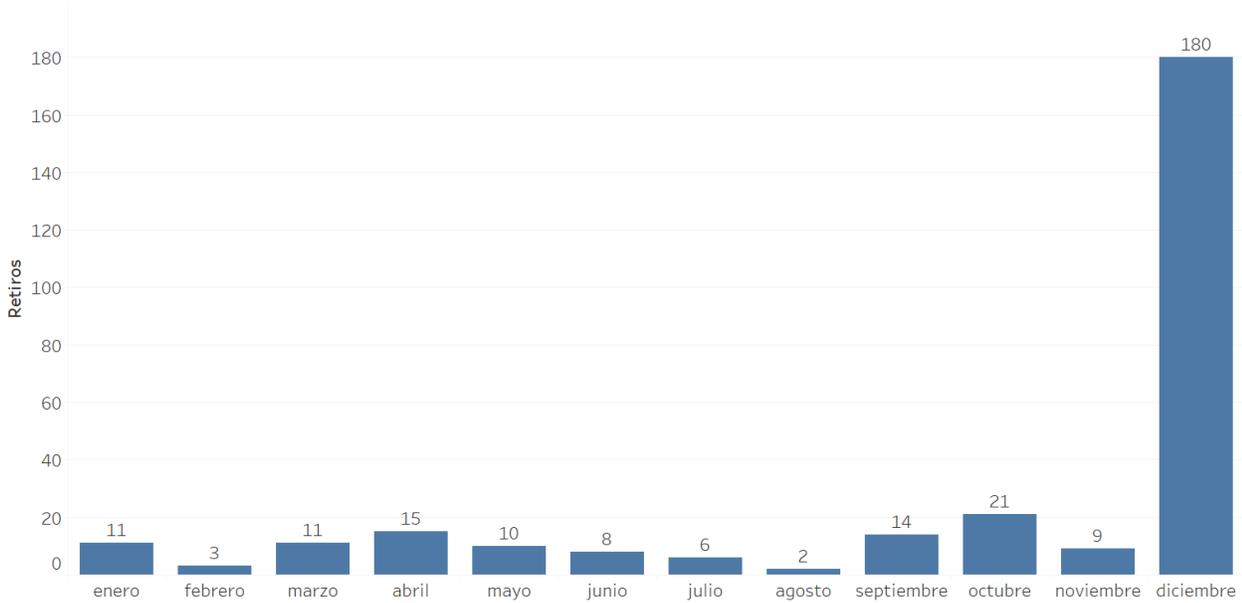


Figura 4.3: Número de retiros a los cuales se le extrae un mes completo según mes.

Finalmente, se filtran las horas anómalas de la serie anual. La cantidad de horas eliminadas en total se muestran en el gráfico de la figura 4.4.

Horas filtradas por mes

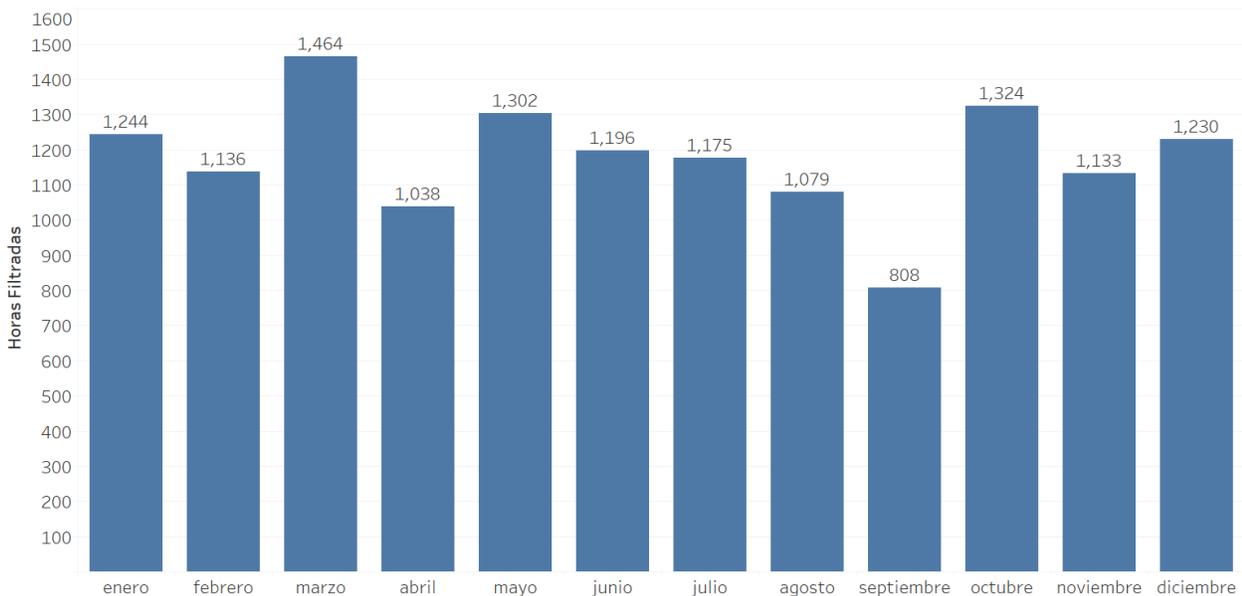


Figura 4.4: Número total de datos horarios eliminados por mes.

Una vez realizados todos los filtros, la tabla 4.3 y figura 4.5, muestran la distribución de energía y retiros por sector.

Sector	Clientes	Retiros	Energía (GWh/año)
MINE	81	88	23,492
IMNM	271	326	4,138
COME	110	199	2,636
AGCS	177	203	2,211
IMM	52	53	1,214
TRAC	33	39	890
AIEA	52	71	747
OTRO	40	40	651
FINA	41	50	597
SEGA	51	63	560
PESC	49	60	365
SOSA	28	34	211
CONS	25	31	210
EDCO	70	71	160
HORE	23	25	113
ENSE	13	23	59
ADPD	3	3	46
Total	1119	1379	38,301

Tabla 4.3: Número de clientes, retiros y energía anual de cada sector económico, luego de aplicar filtros.

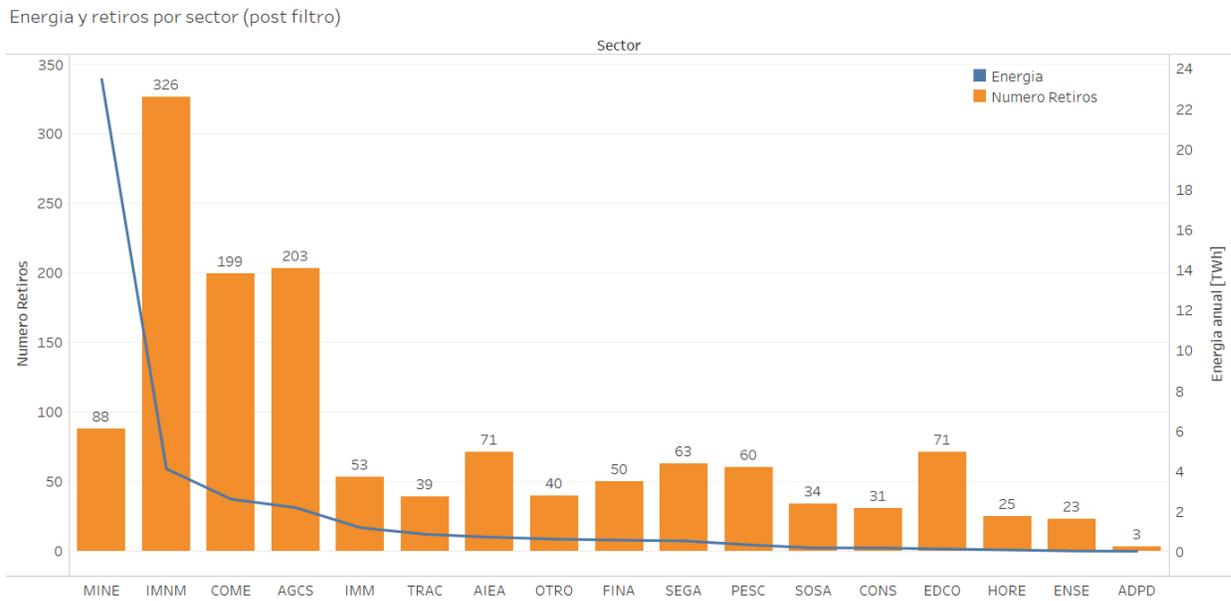


Figura 4.5: Gráfico de número de retiros y energía anual por sector asociado a datos de tabla 4.3.

4.1.3. Representación bidimensional

Luego de confeccionar los perfiles tipo para cada retiro con la información de la base filtrada, se procede a realizar la transformación a 2 dimensiones mediante TSNE. Para esto se usa una perplejidad de 35, se obtiene un σ medio de 0.9480 y diferencia de Kullback-Liebler de 0.8353 luego de 1000 iteraciones. La representación bidimensional se muestra en la figura 4.6.

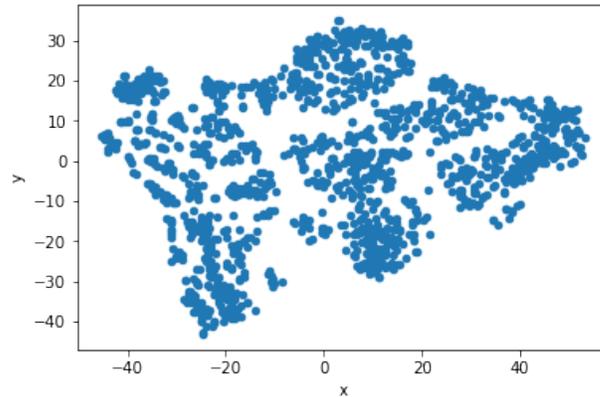


Figura 4.6: Representación bidimensional de retiros dimensionados por su perfil de retiro y sector económico.

Para clusterizar los retiros, se prueba con KMeans y DBSCAN, seleccionando la segunda por inspección. Además se hacen pruebas con distintas combinaciones de parámetros de entrada, de los cuales se escogen dos combinaciones para mostrar: $\varepsilon = 0,188$ con $min_samples = 15$ y $\varepsilon = 0,245$ con $min_samples = 30$. Los resultados de la clusterización se muestran en la figura 4.7. En el caso de la figura 4.7a se designan 13 clusters y 49 outliers, mientras que para la figura 4.7b se designan 11 clusters y 126 outliers.

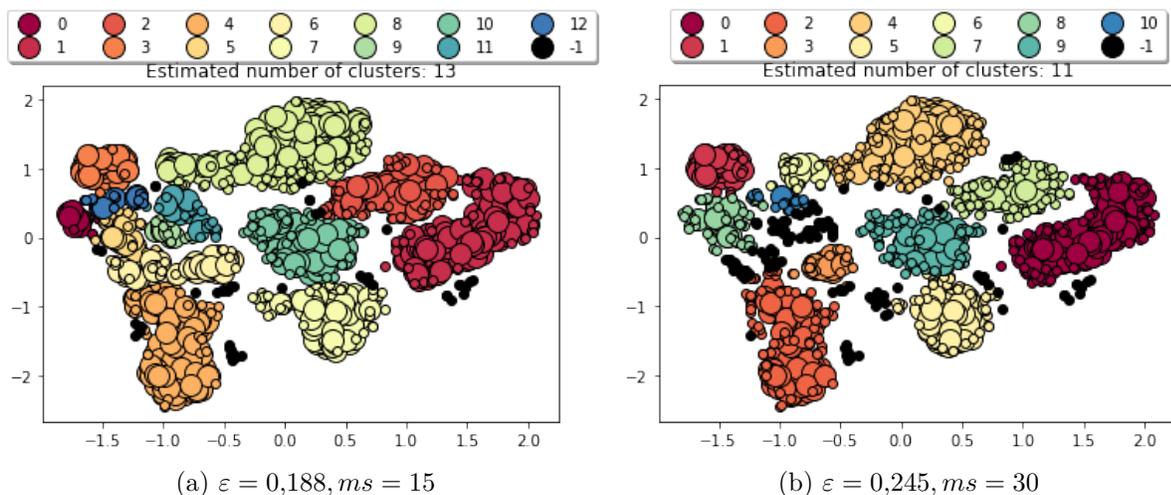


Figura 4.7: Clusterización de resultados con DBSCAN, para distintos parámetros de entrada. Los círculos grandes representan a los “core samples” de DBSCAN. Los colores indican el cluster. Aquellos puntos con color negro y cluster -1 son outliers.

De los resultados expuestos en la figura anterior, se ve que los clusters más evidentes se agrupan de casi igual forma. Estos son [1,2,3,4,7,8,10] y [0,1,2,4,5,7,9] para las figuras 4.7a y 4.7b, respectivamente. Inspeccionando los resultados, se decide usar los asociados a la figura 4.7b, por lo que en adelante, se asume que los clusters son los mostrados en esta. El detalle de la inspección se presenta en el siguiente capítulo.

4.1.4. Descomposición por sector

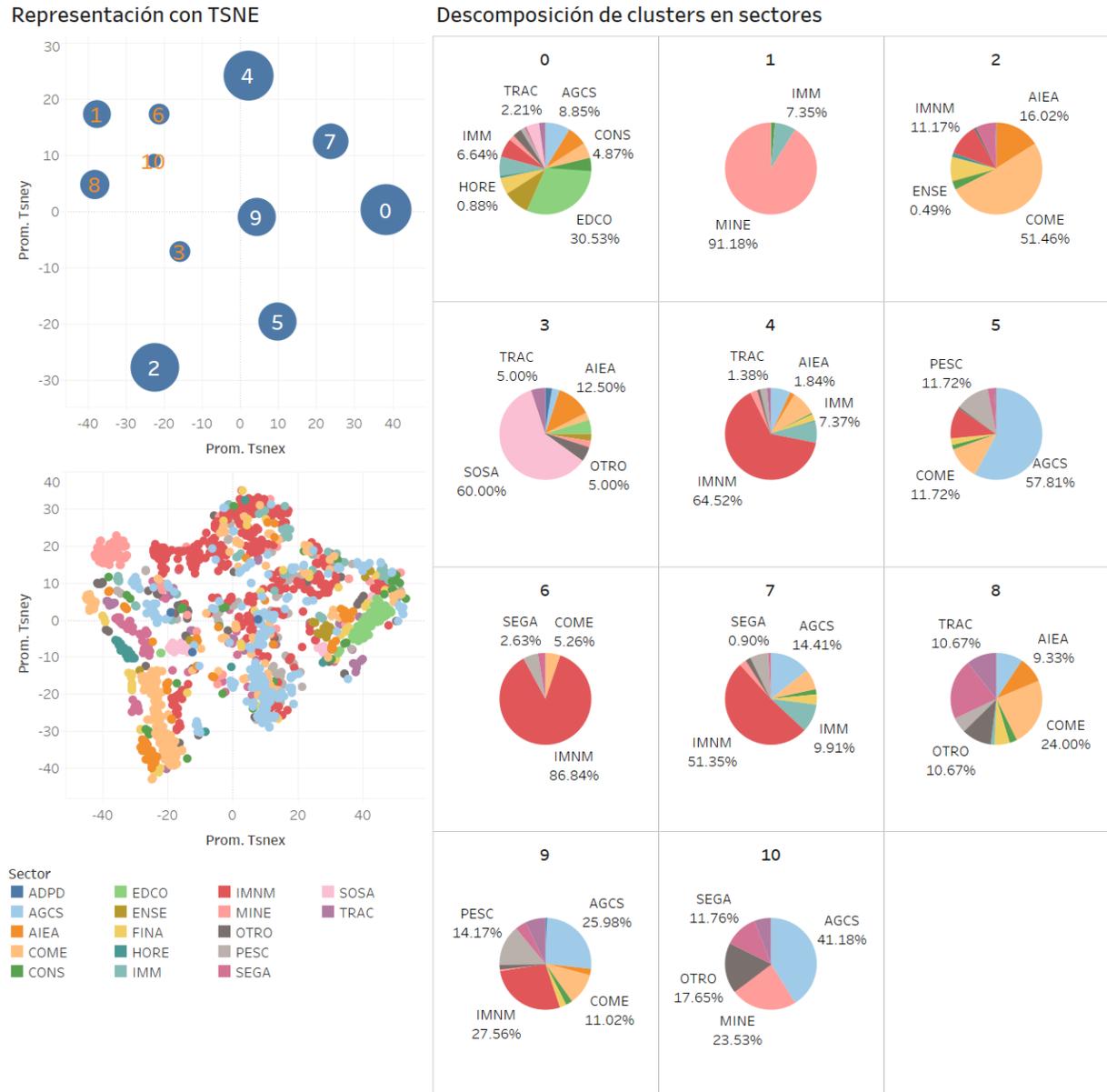


Figura 4.8: Se relacionan los clusters con los sectores económicos de los retiros que contienen. En el sector izquierdo de la figura se muestran dos representaciones del plano generado con tsne: la superior muestra la ubicación y tamaño (según cantidad de retiros) de los clusters y la inferior, muestra todos los retiros coloreados por sector. A la derecha se muestra un gráfico de pie asociado a cada cluster, donde se resalta la descomposición porcentual de los sectores predominantes.

Con el fin de relacionar los clusters con los sectores económicos a los que pertenecen los clientes, se confecciona la figura 4.8. En esta se muestran dos representaciones del plano generado con tsne: una con la ubicación y tamaño (según cantidad de retiros) de los clusters y otra mostrando todos los retiros diferenciados por sector (cada color representa un sector). Luego se muestra un gráfico de pie asociado a cada cluster, donde se indica el porcentaje de retiros de cada sector respecto al número total de retiros en el cluster.

Se aprecia que los retiros se reparten en el plano a excepción de los clientes mineros, que en su mayoría se agrupan en el cluster 1. Sin embargo, la mayoría de los clusters tiene un sector predominante. En la figura de tsne coloreado por sector, se aprecia que dentro de las agrupaciones, los retiros pertenecientes al mismo sector se tienden a juntar. Esto demuestra que la segmentación funciona como se espera.

4.1.5. Perfiles por cluster

Luego de analizar los contenidos de los clusters (detallado en apartado 5.2.2), se definen los 8 clusters principales para el estudio. A estos, se les asigna un nombre y son los que se busca caracterizar. A continuación se muestra el perfil promedio de cada uno de los clusters importantes y se contrastan con 10 perfiles reales de clientes pertenecientes a este. Los perfiles promedio de todos los clusters se incluyen en el anexo (figura C.3).

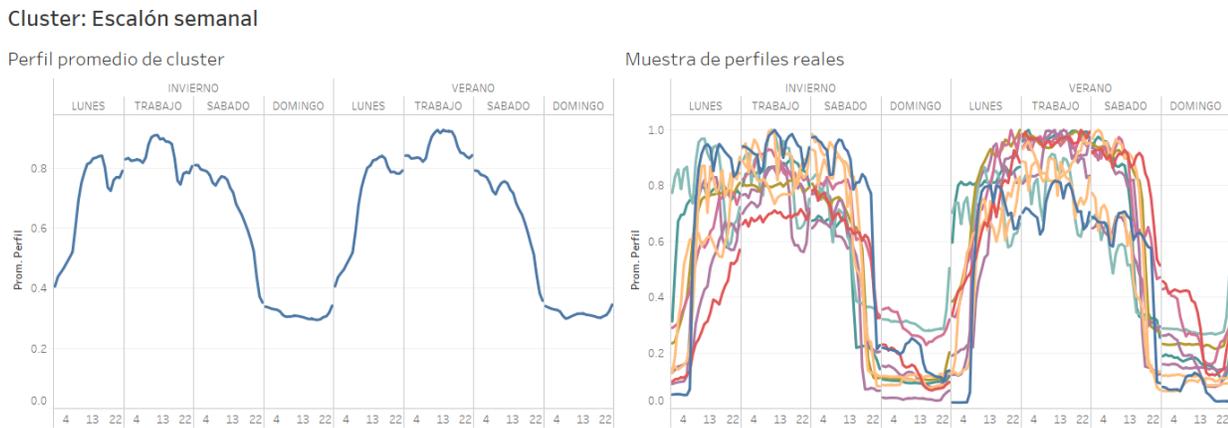


Figura 4.9: Gráficos de perfiles de cluster 4, apodado “Escalón semanal”. A la izquierda se muestra el perfil promedio del cluster y a la derecha los perfiles de 10 retiros reales.

Cluster: Estacional

Perfil promedio de cluster



Muestra de perfiles reales

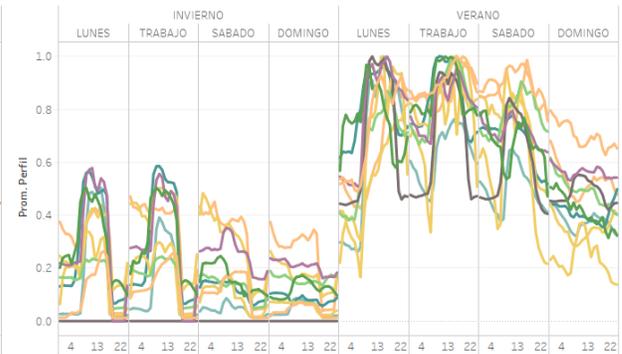
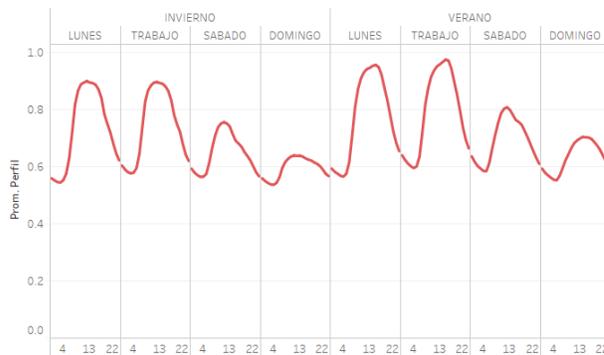


Figura 4.10: Gráficos de perfiles de cluster 5, apodado “Estacional”. A la izquierda se muestra el perfil promedio del cluster y a la derecha los perfiles de 10 retiros reales.

Cluster: Hospital

Perfil promedio de cluster



Muestra de perfiles reales

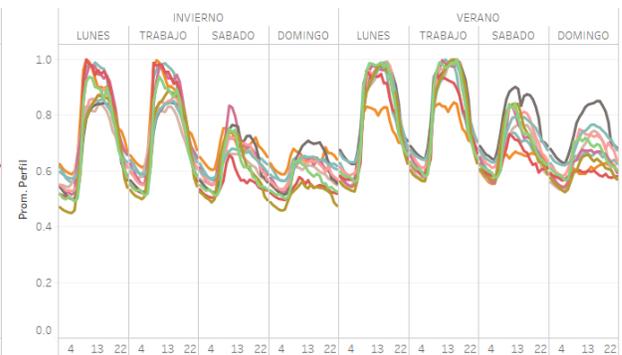


Figura 4.11: Gráficos de perfiles de cluster 3, apodado “Hospital”. A la izquierda se muestra el perfil promedio del cluster y a la derecha los perfiles de 10 retiros reales.

Cluster: Minero

Perfil promedio de cluster



Muestra de perfiles reales

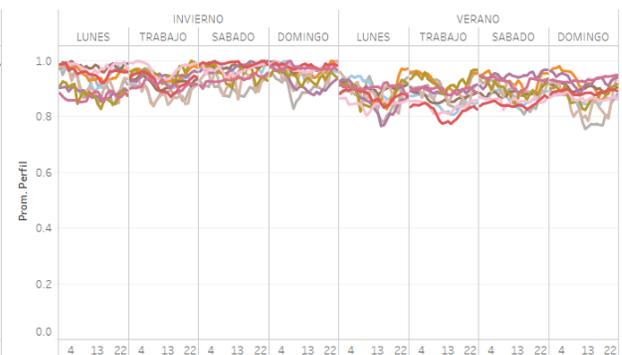


Figura 4.12: Gráficos de perfiles de cluster 1, apodado “Minero”. A la izquierda se muestra el perfil promedio del cluster y a la derecha los perfiles de 10 retiros reales.

Cluster: Oficina

Perfil promedio de cluster



Muestra de perfiles reales

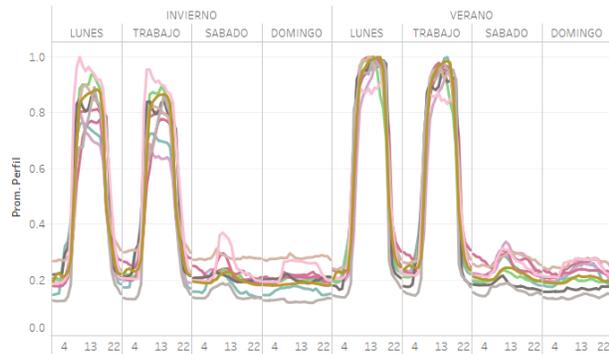


Figura 4.13: Gráficos de perfiles de cluster 0, apodado “Oficina”. A la izquierda se muestra el perfil promedio del cluster y a la derecha los perfiles de 10 retiros reales.

Cluster: Oficina industrial

Perfil promedio de cluster



Muestra de perfiles reales

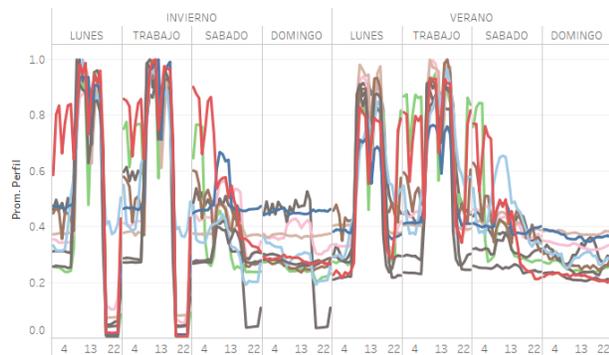
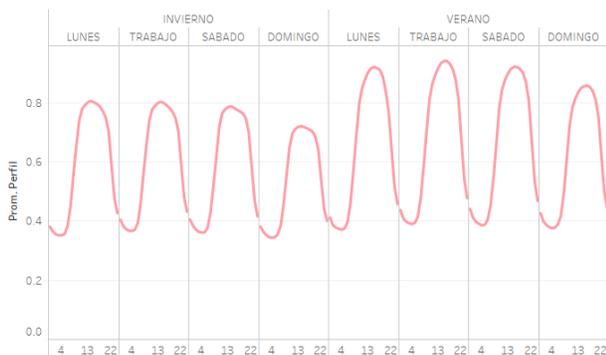


Figura 4.14: Gráficos de perfiles de cluster 7, apodado “Oficina industrial”. A la izquierda se muestra el perfil promedio del cluster y a la derecha los perfiles de 10 retiros reales.

Cluster: Retail

Perfil promedio de cluster



Muestra de perfiles reales

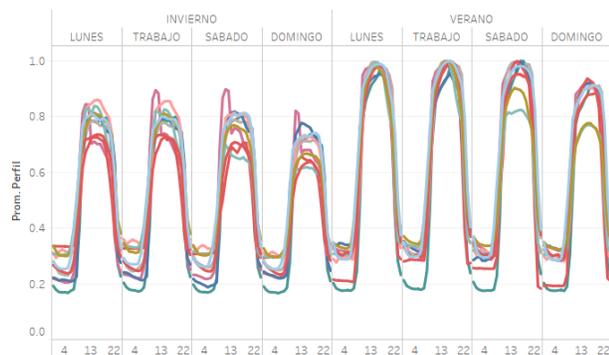


Figura 4.15: Gráficos de perfiles de cluster 2, apodado “Retail”. A la izquierda se muestra el perfil promedio del cluster y a la derecha los perfiles de 10 retiros reales.

Cluster: Típico industrial

Perfil promedio de cluster



Muestra de perfiles reales

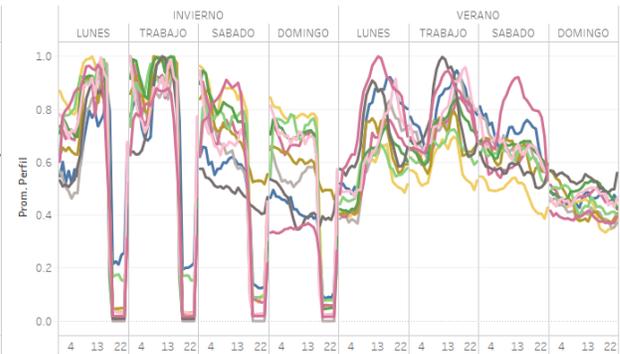


Figura 4.16: Gráficos de perfiles de cluster 9, apodado “Típico industrial”. A la izquierda se muestra el perfil promedio del cluster y a la derecha los perfiles de 10 retiros reales.

4.2. Entrevistas

Luego de analizar los clusters (expuesto en apartado 5.2.2), se realizan entrevistas a 6 clientes. Estos poseen retiros pertenecientes a los clusters más importantes que tienen potencial de flexibilidad. Los resultados de las entrevistas se muestran a continuación, donde se sintetiza lo conversado con el cliente y se muestran los perfiles de los retiros por los que se les pregunta, contrastados con el perfil medio del cluster al que pertenece. Los nombres de los clientes no son mostrados, pero se identifica su sector y los clusters por los que se le pregunta (los nombres de los clusters se explican en el apartado 5.2.2).

Cabe destacar que a medida se realizan las entrevistas, estas se guían de mejor manera. Es decir, se aprende a cómo tratar con los clientes para obtener más información. Es por esto que los resultados se muestran en el orden cronológico en que se realizan las entrevistas.

4.2.1. Cliente 1: Sector agrícola, perfil estacional

Este cliente basa su negocio en el cultivo de uva. Si bien también tiene plantas de vino, su principal consumo se asocia al cultivo.

Perfil estacional

Perfil de retiro, Cliente 1 - Estacional



Figura 4.17: Perfil real de cliente 1, catalogado como “Estacional”

- Declara que su estado normal de operación es consumir a máxima capacidad de sus procesos durante todas las horas del año, pero que los procesos se diferencian por estación. En invierno (abril - septiembre) bombean agua para llenar tranques, mientras que en verano además de bombear se riega.
- El consumo asociado al riego es mayor y trata de no tener interrupciones. Por su parte, las bombas se detienen solo en caso de que no haya agua o necesiten una mantención.
- El horario de punta es realmente una complicación para este cliente. Estiman que pierden mucho por no poder bombear durante el período de control, al nivel que evalúan no considerar el horario punta y aceptar la penalización. Para regularlo, antes cortaban el suministro y ahora están ocupando interruptores con relojes.
- En la planta de vinos (que funciona 24/7), tienen un grupo electrógeno diesel de respaldo, el cual encienden para el horario de punta. El consumo de la planta es constante, pero su volumen depende de la cantidad de uva que reciben.

4.2.2. Cliente 2: Sector industria manufacturera no metálica, perfil oficina industrial

Esta empresa se dedica al procesamiento de madera.

Perfil oficina industrial

Perfil de retiro, Cliente 2 - Oficina industrial



Figura 4.18: Perfil real de cliente 2, catalogado como “Oficina industrial”

- Declara que se preocupan del horario de consumo, debido al horario de punta, Para esto tienen generadores de respaldo que suplen el 100 % de los consumos en ese horario.
- Tienen turnos rotativos 24/7 de lunes a sábado. El entrevistado asegura que todos sus consumos son críticos y no gestionables: calderas (que no pueden aprovechar inercia térmica) y regadío continuo.

4.2.3. Cliente 3: Sector comercio al por mayor y menor, perfiles típico industrial y estacional

El rubro de esta empresa es el cultivo y venta frutas. Su perfil típico industrial se debe al almacenamiento de productos en condiciones óptimas y el estacional a sus cultivos.

Perfil típico industrial

Perfil de retiro, Cliente 3 - Típico industrial



Figura 4.19: Perfil real de cliente 3, catalogado como “Típico industrial”

- El principal consumo eléctrico asociado a su planta industrial es mantener la fruta en condiciones específicas de presión y temperatura. Este consumo representa aproximadamente un 70 % del total. El consumo es variable pues está en función de la cosecha.

- Cuando trabajaban con productos congelados no tenían problemas con la interrupción del suministro, pero con el método de ahora si lo hacen. Cuenta que tuvieron un corte de una hora y fue muy perjudicial. Los cortes hacen que exista un desbalance en las máquinas que mantienen presión y temperatura debido al cambio en esta variables.
- Sí se preocupan de cortar la electricidad en el horario de punta. Para eso tienen grupos electrógenos para cubrir la demanda que denomina fundamental. Por ejemplo, en la sala de máquina de una de sus plantas importantes tienen un consumo de 3 MVA el cual reemplazan en la hora punta por un generador diesel de 500 kVA. Este generador consumo alrededor de 70 L por hora y se opera a aprox 450 kW.
- La demanda fundamental depende de las condiciones externas de temperatura. Por ejemplo, en el periodo de control de hora punta pueden operar con mucho menos consumo debido a que son horas en que hay poca (o nula) radiación solar. Afirma que están dispuestos a ser flexibles de ser informados con antelación, durante horas en que la temperatura de ambiente sea baja.

Perfil estacional

Perfil de retiro, Cliente 3 - Estacional



Figura 4.20: Perfil real de cliente 3, catalogado como “Estacional”

- Respecto a su consumo estacional, mantiene la posición del primer cliente, que afirma que sus consumos se asocian a bombas y regadío, que se operan a máxima capacidad durante todo el año.

4.2.4. Cliente 4: Sector industria manufacturera no metálica, perfiles típico industrial, escalón semanal y estacional

Este cliente está ligado al negocio del vino, en que trabajan con el proceso productivo completo: Desde el cultivo de uvas, procesamiento y embotellamiento.

Perfil escalón semanal

Perfil de retiro, Cliente 4 - Escalón semanal



Figura 4.21: Perfil real de cliente 4, catalogado como “Escalón semanal”

- La planta cuyo perfil pertenece al cluster escalón semanal, es una embotelladora que funciona todo el año. Esta no se preocupa de los horarios punta.
- En meses de verano se hacen turnos rotativos para distintos procesos, entre los cuales algunos solo funcionan en horas de día. Esto explica el aumento de consumo durante el bloque laboral (aprox de 7 a 19).
- Los turnos son de lunes a sábado, debido a la ley laboral. Para poder tener turnos los domingos, se debe justificar que la naturaleza del trabajo implica que hayan turnos continuos. Cada turno depende del proceso que se lleve a cabo y esto depende de las ventas de la empresa.

Perfil típico industrial

Perfil de retiro, Cliente 4 - Típico industrial



Figura 4.22: Perfil real de cliente 4, catalogado como “Típico industrial”

- El perfil típico industrial se asocia a bodega que no tiene embotellado y es donde se hacen las mezclas. Los principales consumos se deben a bombas y compresores, los cuales operan en función de las ventas.
- El peak de consumo lo tienen en mayo y abril.

- Durante las horas de punta arriendan generadores. En los meses de peak estos cubren el 100% del consumo, mientras que durante los otros meses en que hay demanda en punta, los arriendan por menos capacidad en función de sus ventas.
- El entrevistado afirma estar abierto a la flexibilidad, especialmente si eso implica arrendar menos generadores que suplen el consumo en la hora de punta.

Perfil estacional

Perfil de retiro, Cliente 4 - Estacional



Figura 4.23: Perfil real de cliente 4, catalogado como “Estacional”

- Por último, en cuanto al perfil estacional, ocurre lo mismo que los clientes 2 y 3, en que se afirma que no poseen flexibilidad en las bombas y el regadío.

4.2.5. Cliente 5: Sector comercio al por mayor y menor, perfiles escalón semanal y típico industrial

Esta empresa se dedica a la elaboración de productos alimenticios.

Perfil escalón semanal

Perfil de retiro, Cliente 5 - Escalón semanal



Figura 4.24: Perfil real de cliente 5, catalogado como “Escalón semanal”

- La entrevista se enfoca en su perfil de escalón semanal. Este se asocia a 5 fábricas de la planta donde se hace el retiro:
 1. UHT - Ultra High Temperature (Proceso néctar y leches, es el que requiere de más consumo de electricidad)
 2. Refrigerados (Jugos, frutas, cremas)
 3. Concentrado de fruta (Diciembre y Julio) (Mix de frutas)
 4. Mermeladas
 5. Conservas
 6. Procesado de frutas para yoghurt
- Tienen 3 turnos rotativos (para llevar procesos las 24 hrs del día) que funcionan de lunes a sábado. (No pueden ser continuos por ley de trabajo). El pronunciamiento en horario laboral se debe a procesos que tienen menos turnos, por ejemplo 1 o 2.
- En cuanto a los recortes en hora punta, en su consumo de escalón semanal no pueden hacerlo debido a lo crítico del proceso. Este funciona manteniendo condiciones de alta temperatura. Reducen un poco el consumo, pero no completo. No pueden flexibilizar en el proceso lácteo.

Perfil típico industrial

Perfil de retiro, Cliente 5 - Típico industrial



Figura 4.25: Perfil real de cliente 5, catalogado como “Típico industrial”

- El retiro calificado como típico industrial se asocia a una planta donde tienen productos congelados. El entrevistado admite que recortan punta con facilidad. Pueden cortar el consumo completo durante todo el período de control sin problemas.
- Está abierto a flexibilidad dada una correcta remuneración.

4.2.6. Cliente 6: Sector comercio al por mayor y menor, perfil retail

Este es un gran cliente de retail, que posee tiendas a lo largo de todo Chile.

Perfil retail

Perfil de retiro, Cliente 6 - Retail

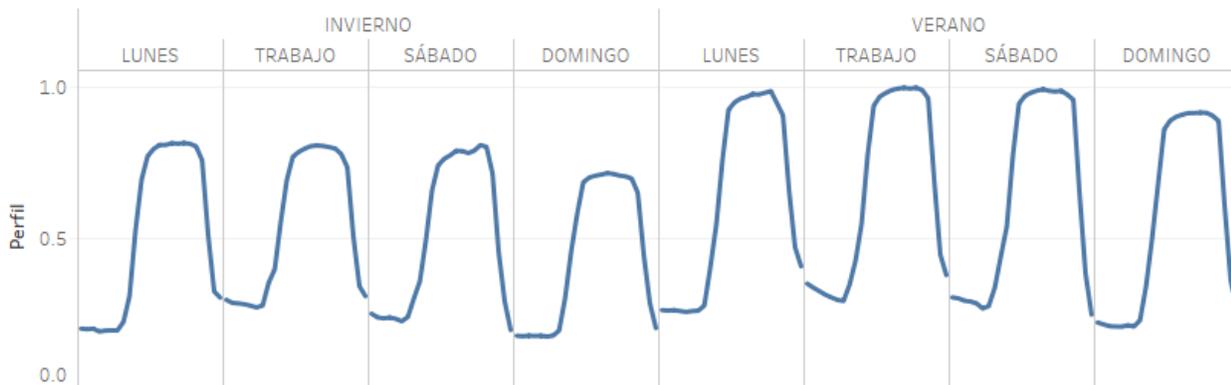


Figura 4.26: Perfil real de cliente 6, catalogado como “Retail”

- El mayor consumo se asocia a climatización, en particular ventilación y aire acondicionado. El volumen depende de la época del año y del horario.
- Tienen medidores para monitorear en línea el consumo de electricidad y se preocupan por el horario punta. Cuando empieza el período de control, se le avisa a las tiendas para que disminuyan el consumo. Cuando esto ocurre, bajan los compresores (aire acondicionado) y sobreventilan. Al proceso anterior se le denomina “Free Pooling”.
- Las tiendas no tienen calefacción, salvo pocas excepciones en el sur de Chile. Cuando tienen calefacción, esta es a gas. El problema es enfriar, debido al calor emanado por las personas, equipos encendidos y luces (estas siempre están encendidas en el horario de atención a clientes). El uso de luces LED los ha ayudado mucho para reducir el calor de la tienda, necesitando de menos acondicionamiento.
- Los controladores de los equipos de aclimatización son muy rápidos. Pueden brindar flexibilidad a una correcta remuneración, pero se tienen que estudiar bien los efectos en la temperatura de la tienda. Disminuir su consumo por un tiempo implica aumentarlo en el periodo inmediatamente a continuación.

Capítulo 5

Análisis

5.1. Discusión metodológica

5.1.1. Base de clientes y confección de perfiles

La calidad de los datos y la forma en que se usan para confeccionar los perfiles de clientes son críticos en la metodología. Todos los supuestos considerados en esta sección afectan a los resultados, por lo que se debe ocupar la información disponible de forma óptima. Para efectos de la confección de perfiles, la metodología óptima es aquella que permita comparar de mejor forma los patrones de consumo evitando de mejor manera las distorsiones por errores en las bases de datos.

Para que la metodología sea óptima, se busca tener la mayor cantidad de retiros posible. Mientras más información asociada a distintos clientes existe, más patrones se pueden determinar y es más probable encontrar perfiles de consumo similares. De esta forma, más retiros implica clusters más representativos del sistema y una segmentación más adecuada.

Mientras mayor es la historia de los retiros, mejor se puede representar el perfil de consumo de cada uno de ellos. Por ejemplo, mediante la serie horaria de los retiros desde septiembre de 2017 a agosto de 2019, se podrían haber confeccionado los perfiles promedio por tipo día para verano 2017, invierno 2018, verano 2018, invierno 2019. Así, al añadir la componente cronológica se puede corroborar que el patrón se repita año a año o bien analizar la evolución.

Sin embargo, los datos disponibles implican un tradeoff entre el número de clientes y el historial de la información. Los clientes que tienen mediciones horarias públicas son solo los clientes libres, los cuales comienzan a ser un número considerable recientemente (ver figura 3.1). Cada mes, más clientes del sistema se están pasando de regulado a libre, por lo cual su información comienza a estar disponible a partir del mes en que se cambian. Como se detalla en la sección 3.2, es importante que para la confección de los perfiles se ocupe el mismo período de estudio. Por lo tanto, mientras más largo el periodo de estudio, menos clientes tienen datos completos, lo que hace que haya que eliminarlos.

Si bien sería más preciso que la descomposición temporal que define los perfiles fuese por

mes en vez de por estación, se opta por la segunda debido a la calidad de los datos. Para emplear TSNE, a todos los clientes se les debe asignar las mismas dimensiones; por lo tanto, para la descomposición mensual se tendrían que incluir los meses que son considerados outlier. Además, descomponer por mes obliga a descartar a todos los clientes libres que comienzan a aparecer en los balances desde octubre o noviembre de 2018 (es decir, que no tienen la serie anual completa). La descomposición por estación permite filtrar información errónea e incluir más retiros para el análisis, por lo que se considera que es la descomposición óptima.

Se destaca que la metodología tiene un pequeño error asociado a la disponibilidad de datos en el momento que se realiza el estudio: el retiro del mes de septiembre no corresponde al invierno de 2019, sino al de 2018. Esto hace que se incorpore al promedio de invierno un mes que no coincide cronológicamente, por lo que podría no incorporar cambios en infraestructura ocurridos a partir de octubre de 2018. Este error no puede ser corregido debido a que en el momento de obtener los resultados oficiales no se encuentran disponibles los datos de septiembre de 2019, sin embargo, su impacto en los resultados es insignificante.

5.1.2. Corrección de nombres

La corrección de nombres es un proceso fundamental, y aunque se logran buenos resultados, no está libre de errores. No corregir los nombres implicaría que los clientes que son ingresados de forma distinta en algún mes, pierdan dicha información, dejando su serie temporal incompleta. En la metodología planteada en este trabajo, se eliminan aquellos clientes que tienen datos incompletos, entonces se eliminarían aquellos que tienen datos faltantes debido a su error de nombre.

Sin embargo, existen nombres asociados a un mismo cliente que deben ser diferenciados, pero el script de python interpreta que son el mismo. Por ejemplo, las plantas 1 y 2 de la empresa “Empresa Minera”, se conectan a una misma barra, pero están diferenciadas debido a que sus consumos son muy distintos. Es por esto que una entra al balance como “Empresa Minera 1” y la otra como “Empresa Minera 2”. El algoritmo programado, detecta que estas son muy parecidas, por lo que las junta en una sola. Lo anterior implica que en ocasiones se mezclen retiros que no debiesen hacerlo y pasan desapercibidos al ser revisadas por un humano.

5.1.3. Asignación de sectores

El sector económico que se le asigna a los clientes se extrae de una base de datos que contiene información del SII. Por lo tanto, los nombres de los clientes que aparecen en las bases de balance se comparan con las razones sociales de la base del SII. Lo anterior es una fuente de errores que serán detallados a continuación.

No siempre calza el nombre de un cliente en el balance con su razón social del SII. Por ejemplo, el nombre de algunos clientes del balance está escrito en siglas, mientras en su razón social aparece el nombre completo o viceversa. Si bien el algoritmo de asignación automática tiene una forma de solucionar el problema de no encontrar el nombre en la base del SII, no tiene como detectar cuando encuentra un nombre parecido al buscado, pero que realmente no lo es. Estos errores son revisados al final del proceso, pero no se pueden detectar todos.

En ocasiones, una misma empresa puede tener más de una razón social lo que hace que no siempre se asocie el nombre en el balance a la razón social que corresponde. El algoritmo de automatización tiene un método para detectar la razón social más adecuada para el nombre, pero en ocasiones este no corresponde al de la planta que efectivamente hizo el retiro. Esto hace que haya retiros que se asocian al sector económico equivocado

Por último, hay empresas que tienen una sola razón social, pero en realidad sus actividades pertenecen a más de un sector económico. Por ejemplo, una empresa cuyo negocio es la manufactura y venta de vino, pero se preocupan del proceso completo. Es decir, tienen viñas de donde extraen las uvas, por lo que pueden considerarse agrícolas, también tienen una planta donde procesan las uvas para hacer vino y embotellarlo, lo que las hace pertenecer a las industrias manufactureras no metálicas, y finalmente se dedican a vender sus productos, pudiendo considerarse del sector de comercio. Este error implica que haya clientes cuyos retiros tengan el mismo perfil de consumo (porque efectivamente se dedican a lo mismo), pero se les atribuya un sector económico distinto.

5.1.4. Confección de retiros

Al iniciar la investigación se tiene la hipótesis de juntar a todos los retiros de un mismo cliente según su barra troncal no afecta al estudio, debido a que sería el mismo cliente al que se le aplica la flexibilidad. Esta hipótesis resulta ser totalmente errónea, puesto que hay clientes que tienen plantas muy distintas, cuya barra de conexión se asocia a la misma troncal. Por lo tanto, el efecto de sumar las plantas distorsiona el perfil del retiro.

Una de las distorsiones que se provoca por añadir consumos, se produce cuando aparecen nuevas plantas en el período de estudio, es decir que tienen parte de su serie incompleta. Esto se ejemplifica con las figuras 5.1 y 5.2, en que se muestran los gráficos de dos retiros del mismo cliente (perteneciente al sector de comercio) conectados a distinta barra, pero ambas se asocian a la misma barra troncal. Por lo anterior, en el modelo ambos son considerados parte del mismo retiro aún cuando sus series de tiempo difieren en extensión.

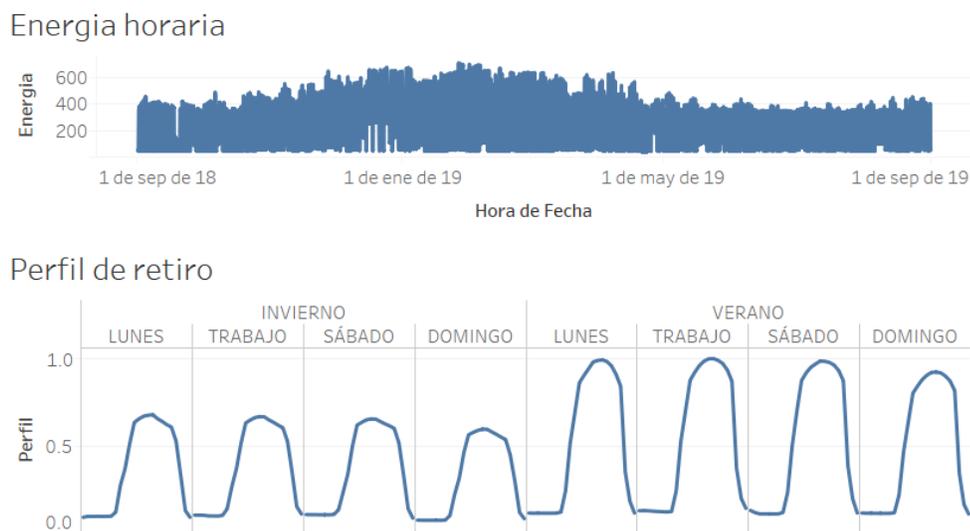


Figura 5.1: Gráficos de consumo horario y perfil generado del retiro de un cliente comercial

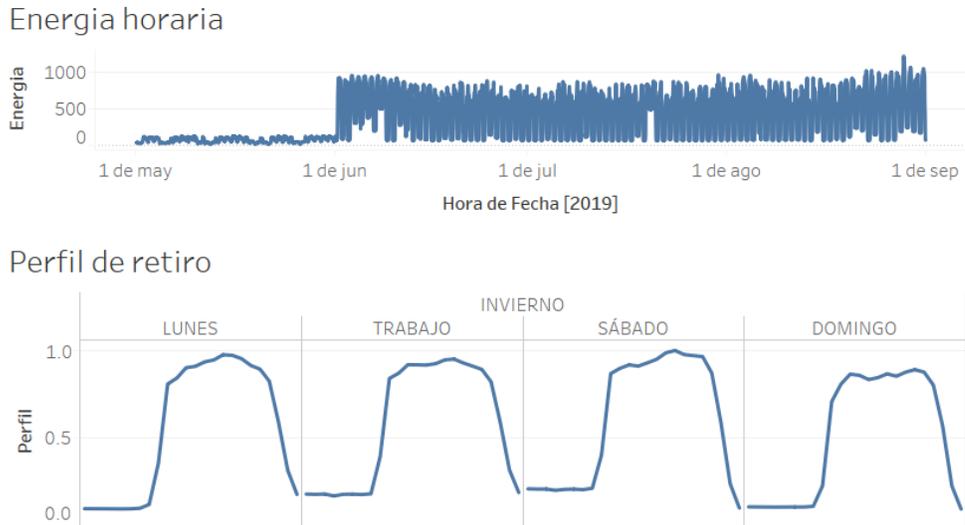


Figura 5.2: Gráficos de consumo horario y perfil generado de otro retiro del mismo cliente de la figura 5.1, que se encuentra conectado a la misma barra troncal.

La composición de los retiros mostrados en 5.1 y 5.2 se muestra en la figura 5.3. Como el retiro de la barra 2 comienza en mayo de 2019, este se agrega al perfil de invierno y su efecto no se muestra en verano. Por lo tanto, pareciera que es un cliente comercial que disminuye su consumo en invierno, pero en realidad esto es porque en medio del periodo de estudio aparece un nuevo retiro asociado a la misma barra troncal. Estas distorsiones pueden hacer que se extraigan conclusiones erradas respecto al comportamiento de los clientes.



Figura 5.3: Gráfico de consumo horario y perfil generado de la composición de los retiros mostrados en las figuras 5.1 y 5.2. Este es el retiro que aparece en el modelo.

Otro caso de distorsión debido a confeccionar los retiros mediante la barra troncal y no

la de retiro es que se mezclen perfiles que son muy distintos y podrían ser diferenciables. Un ejemplo de esto se muestra en la figura 5.4, donde se muestran dos retiros de distintas barras de un mismo cliente, pero la misma barra troncal y luego la composición resultante. Cada retiro por separado podría pertenecer a un cluster distinto, pero como se consideran uno solo, prima el más predominante.

Perfil de retiro



(a) Retiro barra 1

Perfil de retiro



(b) Retiro barra 2

Perfil de retiro



(c) Conjunto barra 1 y 2

Figura 5.4: Perfiles de 2 retiros de un mismo cliente y su composición.

Si bien confeccionar los retiros por barra es una mejora para diferenciar, no es suficiente para asegurar que los procesos se asocien a plantas distintas. Existen clientes que poseen distintas plantas conectadas a la misma barra, lo cual hace que el consumo que se muestra en las bases de balance sea una composición. Lo anterior implica que haya que buscar otro método para detectar estas composiciones y así poder analizar de mejor manera los procesos productivos de los clientes.

5.1.5. Filtros de datos

El objetivo de filtrar los datos es eliminar las anomalías sin perder representatividad de la realidad. El riesgo de eliminar datos es que se desprece información importante para la investigación. Sin embargo, de no eliminarse los datos anómalos, se producen distorsiones en los resultados que pueden influir en la segmentación.

Los retiros más críticos son aquellos cuya serie completa debe ser eliminada. Esto se debe a que se generan perfiles ficticios muy distintos a los reales que alteran la representación bidimensional. Ahora, eliminar clientes completos aumenta la probabilidad de que los datos pierdan representatividad de la realidad, por lo que se debe analizar su impacto en la muestra.

Si bien puede ser considerable la cantidad de retiros que se pierden, principalmente por falta de información, no se pierde mucha representatividad de la demanda. Lo anterior se ve reflejado en la figura 5.5, donde se aprecia que se elimina un gran porcentaje de los retiros de cada sector, pero la variación de la energía que concentran es poca. El único sector donde la disminución de energía puede ser considerable es en SEGA (Suministro de electricidad, gas y agua), lo cual es de esperar, ya que se eliminan los retiros asociados a agrupaciones de clientes libres de distribuidoras. Por lo tanto, se puede decir que la eliminación de retiros hace que se pierda información de algunos patrones de consumo, pero esto no implica una gran pérdida en la representatividad de la demanda total.

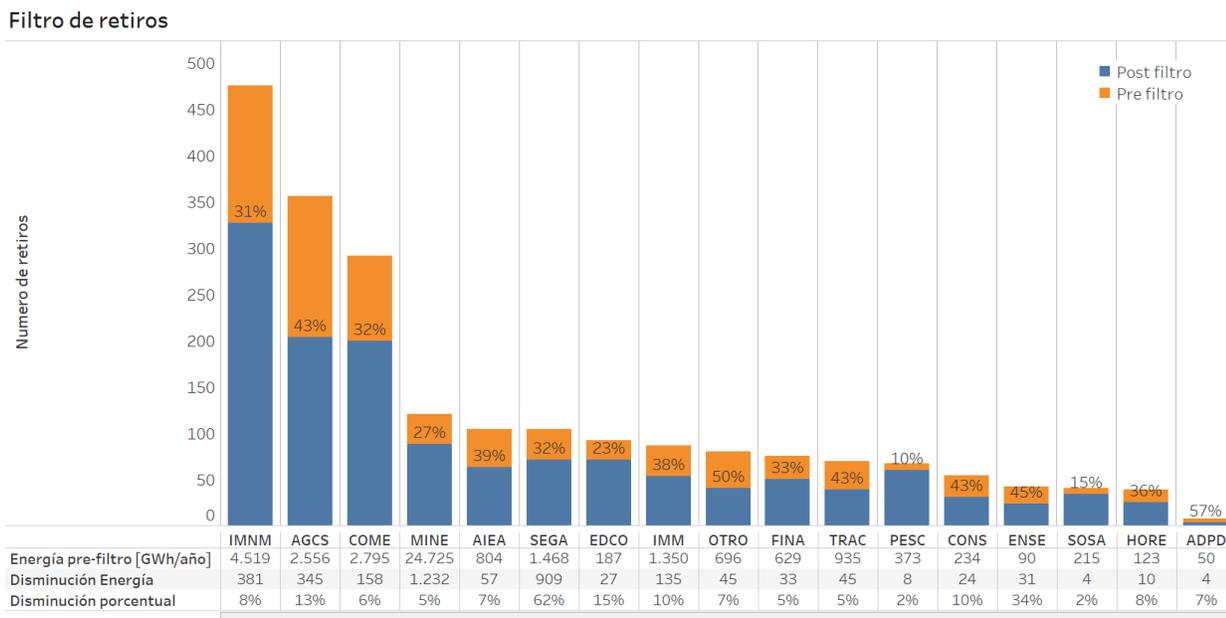
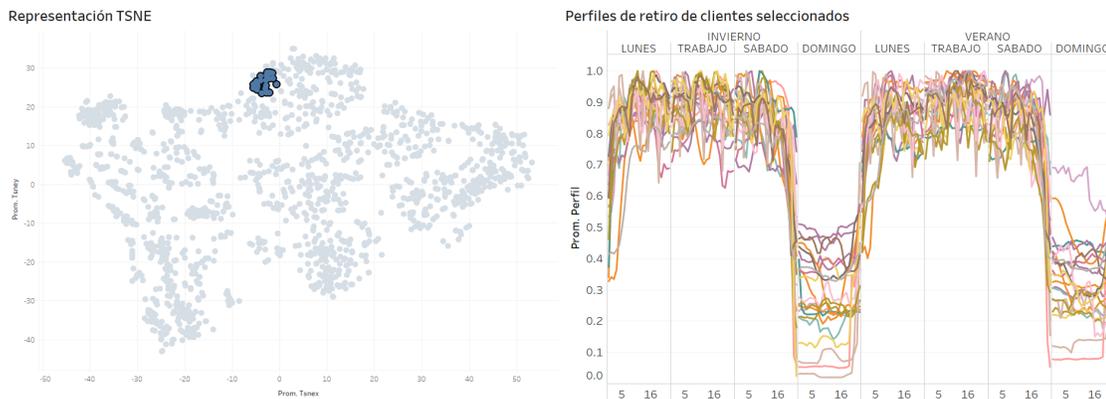
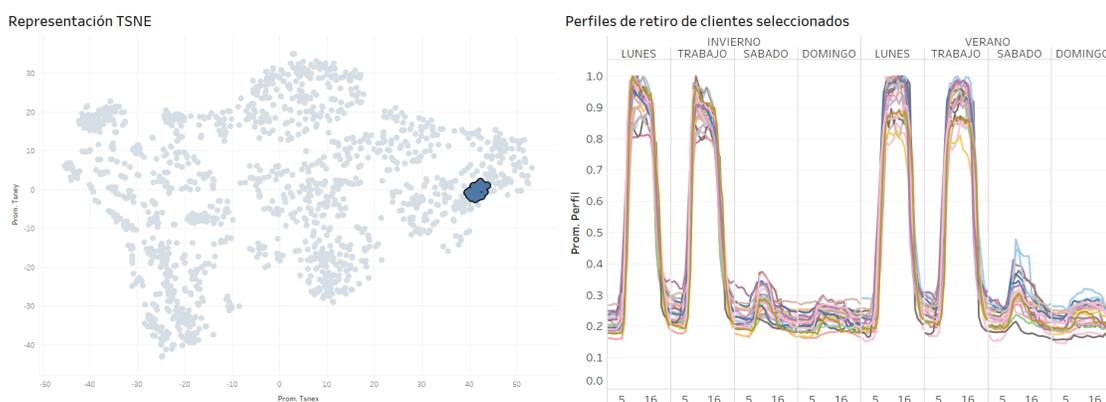


Figura 5.5: Comparación de la cantidad de retiros y energía anual antes y después del filtro de datos. En el gráfico de barras se muestra la cantidad de retiros y la variación porcentual. En la tabla se muestra la variación real y porcentual de energía por sector.

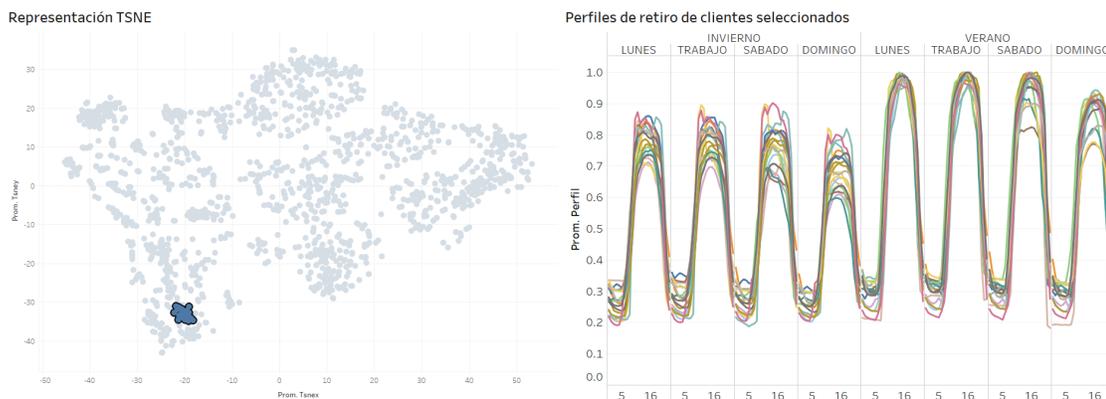
5.1.6. Métodos de clusterización



(a) Inspección 1



(b) Inspección 2



(c) Inspección 3

Figura 5.6: Inspección de perfiles de retiros vecinos en TSNE. En los gráficos de la izquierda se muestra la representación bidimensional obtenida con TSNE, resaltando a un grupo de retiros. En la derecha se muestra el perfil de consumo de cada retiro seleccionado, diferenciándolos por colores.

Una buena forma de corroborar que la metodología de reducción de dimensionalidad está funcionando como se espera, es analizar los perfiles de los retiros que se encuentran cerca en

el plano de TSNE. Lo anterior se muestra en el gráfico de la figura 5.6, donde se seleccionan grupos de puntos en la representación bidimensional y luego se grafican sus perfiles. En estas inspecciones se demuestra la efectividad del método, en que claramente se pueden ver patrones similares en los perfiles de los clientes que aparecen representados cerca.

Habiendo validado que los retiros cercanos en el plano de TSNE tienen perfiles similares, sigue agrupar la mayor cantidad de los que tengan perfil similar mediante técnicas de clusterización. La efectividad de las técnicas se mide comparando los perfiles agrupados entre ellos y teniendo como referencia el perfil medio del cluster. De esta forma, la técnica escogida es la que haga que se agrupen juntos la mayor cantidad de retiros semejantes.

En el apartado 4.1.3 se presentan dos resultados de clusterización mediante DBSCAN usando distintos parámetros (figuras 4.7a y 4.7b), de los cuales se escoge una (figura 4.7b). Lo anterior se decide puesto que la primera clusterización no discrimina correctamente a grupos relativamente cercanos a cluster grandes, pero que en realidad los perfiles de estas son muy distintos. Esto se explica mediante las figuras 5.7 y 5.8.

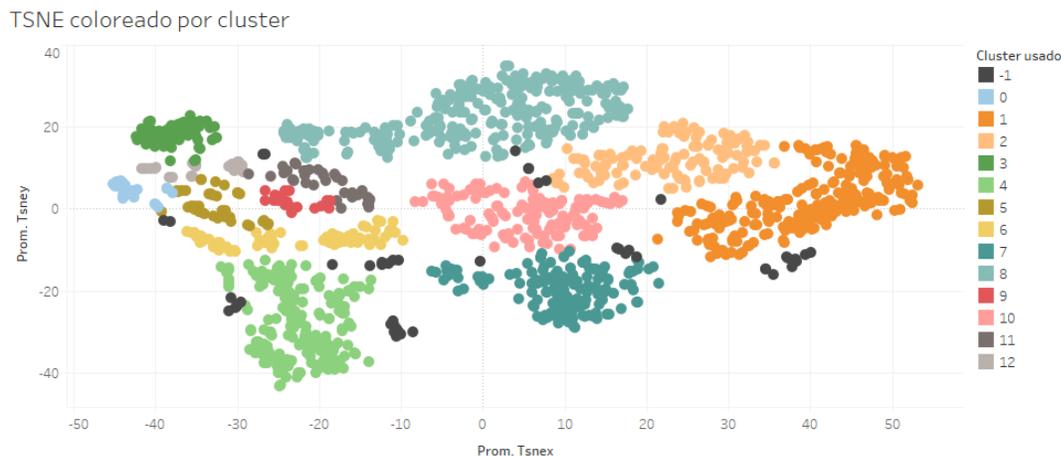


Figura 5.7: Representación bidimensional obtenida de TSNE coloreado para los resultados de la clusterización mediante DBSCAN y parámetros $\epsilon = 0,188$ y $min_samples = 15$

En la figura 5.7 se muestra el plano TSNE donde se colorean los retiros según el cluster al que pertenecen bajo los resultados asociados a los parámetros $\epsilon = 0,188$ y $min_samples = 15$. Llama la atención como se agrupan en un mismo cluster agrupaciones de retiros que están muy alejadas entre sí. Para corroborar la pertinencia de lo anterior, se inspeccionan los retiros alejados y se comparan con la media del cluster. En dicha inspección se determina que efectivamente los retiros alejados poseen perfiles muy distintos a la mayor parte del cluster, por lo que se decide optar por la otra alternativa de resultados, que o bien los excluye o los agrupa en otro cluster.

Un ejemplo de inspección se muestra en la figura 5.8, en que se selecciona un grupo alejado perteneciente al cluster 8. Se puede apreciar que los perfiles de los retiros seleccionados no tienen semejanza a la media del cluster, pero si tienen entre sí.

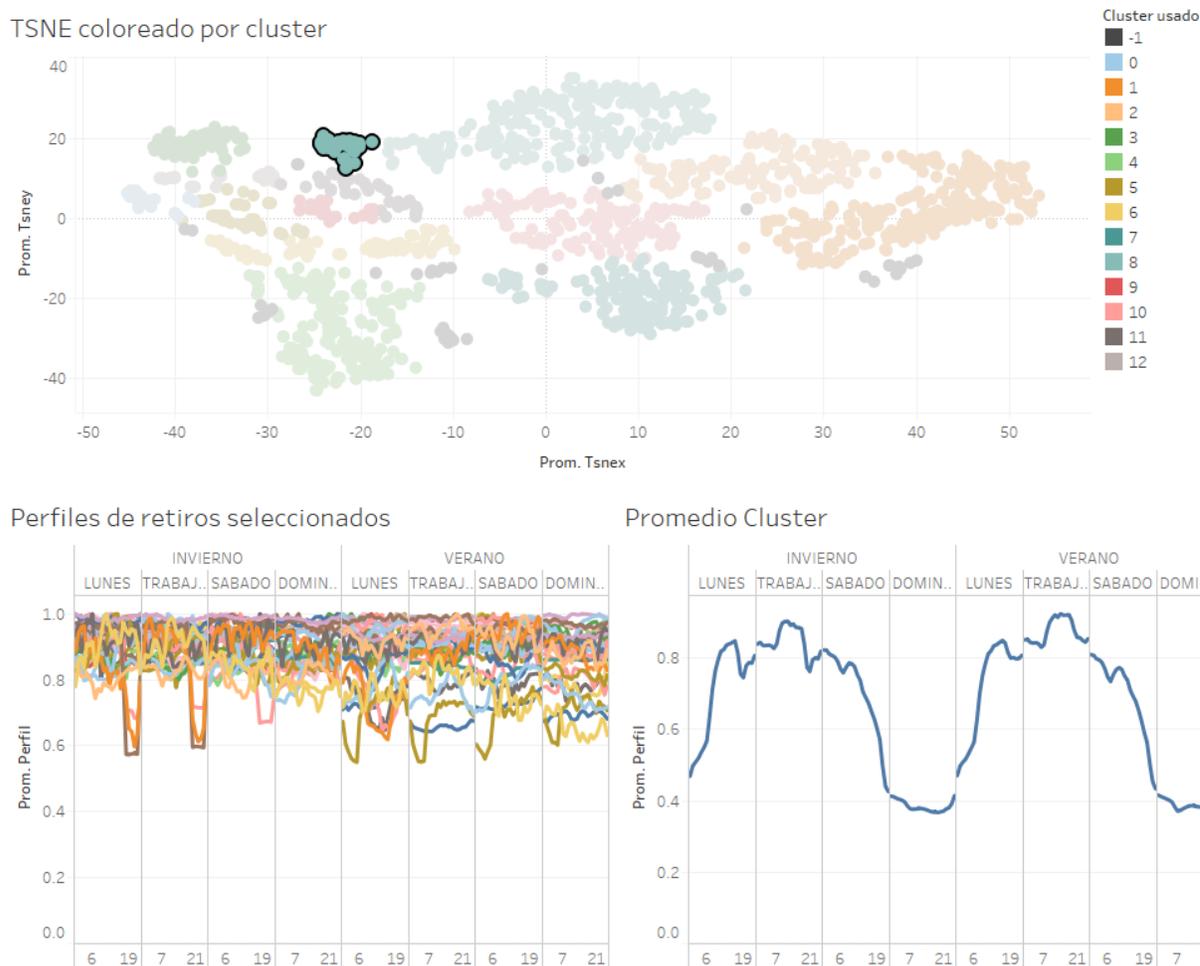


Figura 5.8: Selección particular de retiros en el plano TSNE mostrado en la figura 5.7. En el gráfico de la parte inferior izquierda de la figura, se muestran los perfiles reales de los retiros seleccionados mientras que a la derecha se muestra el promedio del cluster al que pertenecen (cluster 8).

5.2. Análisis de clusters

5.2.1. Selección de clusters más importantes

Los resultados definitivos de clusterización indican que hay 10 clusters de perfiles y 126 retiros que se catalogan como outliers. El plano TSNE con los retiros coloreados por cluster se muestra en la figura 5.9. En esta se detecta que hay pequeñas agrupaciones que podrían ser clusters, pero son tratadas como outliers, debido a que contienen muy pocos retiros.

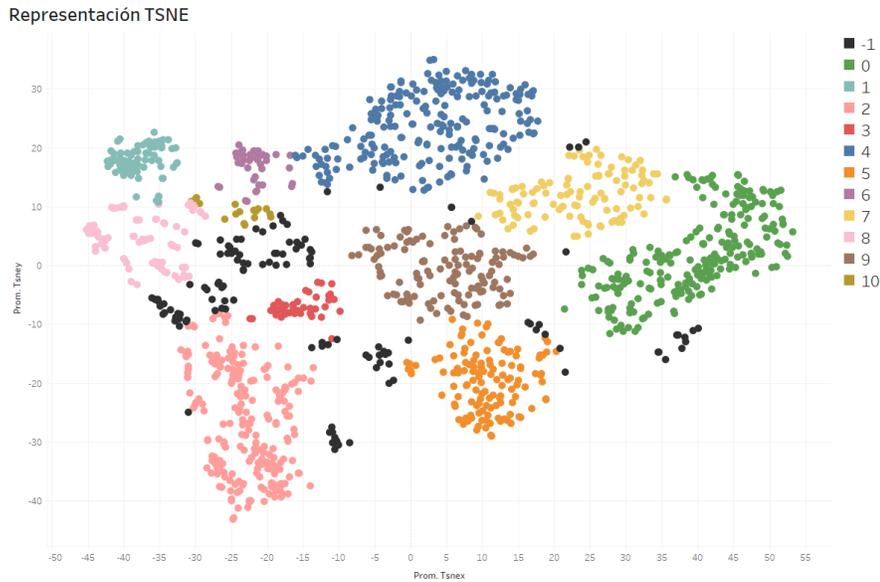


Figura 5.9: Representación bidimensional obtenida de TSNE coloreado para los resultados de la clusterización mediante DBSCAN y parámetros $\varepsilon = 0,245$ y $min_samples = 30$

Si se hace una inspección a los perfiles de las pequeñas agrupaciones consideradas como outliers, se puede ver que estos son claramente similares y distintos a los perfiles de los otros clusters. Esto se muestra en la figura 5.10, donde se exhiben 3 inspecciones de este tipo de agrupaciones pequeñas. Lo anterior lleva a cuestionarse si todas estas pequeñas agrupaciones debiesen estar clusterizadas.

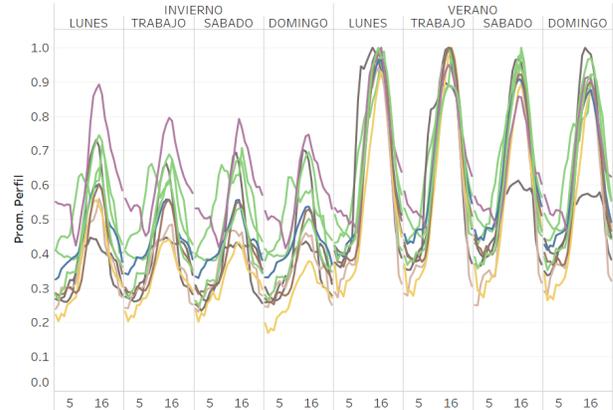
Existe un tradeoff entre el número de clusters y el análisis que se puede hacer sobre estos. Si se configura el DBSCAN con tal de identificar todas las agrupaciones pequeñas por separado, se pueden encontrar más patrones diferentes y que estos se parezcan más entre sí. Sin embargo, esto resulta ser un problema a la hora de analizar, puesto que habrían demasiados, lo que obligaría a hacer un trabajo mucho mas intenso, siendo que en realidad la mayor parte del sistema se puede describir con menos perfiles.

En la representación bidimensional de los datos, hay un sector en que los retiros están claramente separados en pequeños grupos. Este es el sector donde están los clusters 6, 8, 10, y los retiros que se resaltan en las figuras 5.10b y 5.10c. Si bien tienen cosas en común, estas agrupaciones no son tan similares entre sí como los otros clusters. Por lo tanto, se decide tratar los clusters de esa zona (6,8 y 10) como outliers y dedicarse a estudiar en profundidad a los demás.

Representación TSNE

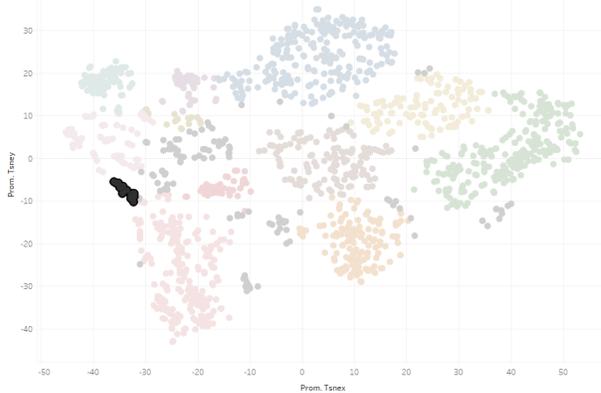


Perfiles de retiro de clientes seleccionados

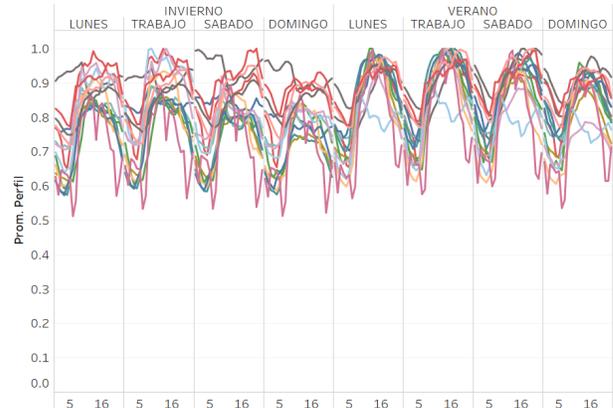


(a) Inspección 1

Representación TSNE

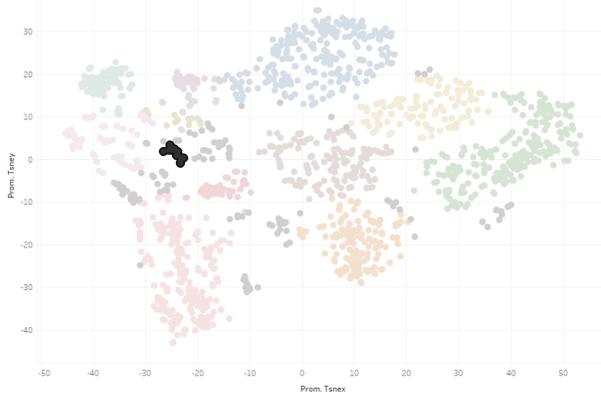


Perfiles de retiro de clientes seleccionados

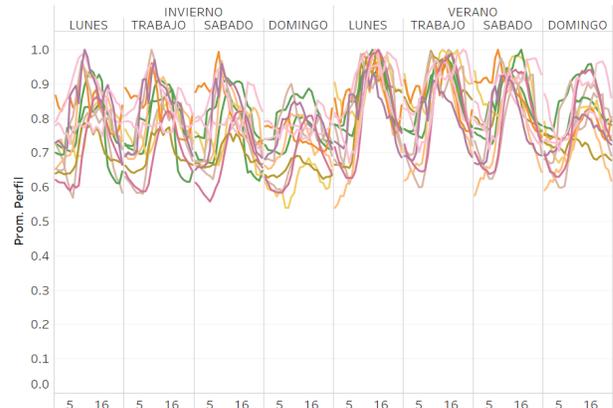


(b) Inspección 2

Representación TSNE



Perfiles de retiro de clientes seleccionados



(c) Inspección 3

Figura 5.10: Inspección de perfiles de pequeñas agrupaciones que se encuentran en el plano TSNE. En los gráficos de la izquierda se muestra la representación bidimensional, resaltando a un grupo de retiros. En la derecha se muestra el perfil de consumo de cada retiro seleccionado, diferenciándolos por colores.

Finalmente quedan los 8 clusters considerados más importantes, los cuales son nombrados según la forma de su perfil promedio. La representación de TSNE con los clusters más importantes se muestra en la figura 5.11 y los nombres asignados se muestran en la leyenda. Las características de estos clusters se describen a continuación.

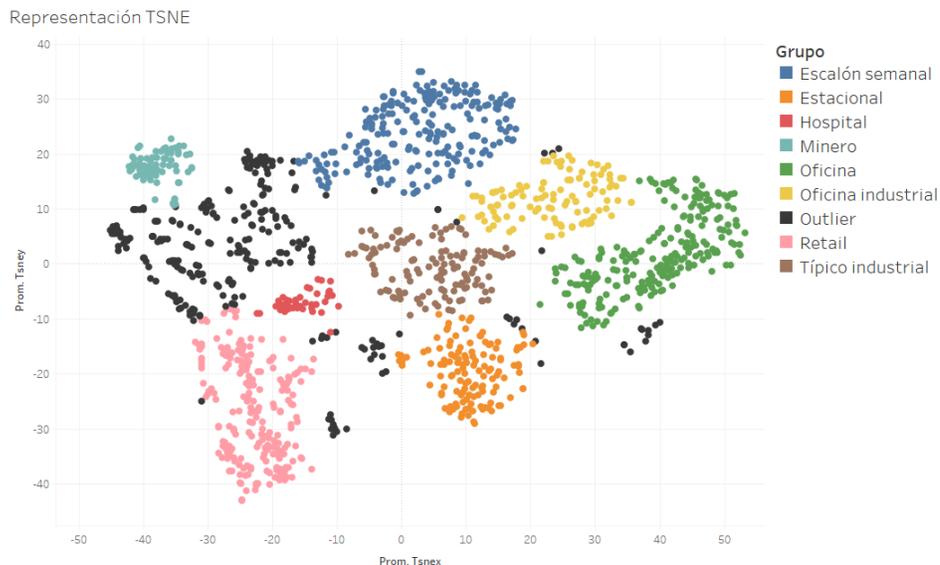


Figura 5.11: Representación final del plano TSNE, donde se colorean los puntos (retiros) por cluster y en la leyenda se exhiben los nombres que se les asigna.

5.2.2. Caracterización de clusters importantes

Los clusters se comparan según su perfil promedio, composición de sectores y ciertas características que poseen los retiros que contienen. Tanto los perfiles promedio como la composición de sectores son exhibidos en el capítulo anterior. En cuanto a las características para comparar los retiros, está el consumo en las madrugadas de los días hábiles, el consumo promedio en invierno y verano, los recortes en punta, la baja en la hora de almuerzo y el pronunciamiento en hora de trabajo. Estas se parametrizan y se evalúan para cada cliente, para luego elaborar una descripción por cluster según la tendencia de los retiros que contiene.

A cada cluster se le confecciona una ficha que incluye su perfil promedio, composición de sectores y descripción según las características de sus retiros. El detalle de la comparación por características se encuentra en las tablas del anexo C. A continuación se presentan las fichas por cluster.

Escalón semanal

Perfil medio cluster



Descripción

Se caracteriza por mantener un consumo muy alto durante los días hábiles y bajo los días domingo. Los lunes y sábado sirven como transición para pasar de consumo bajo a alto y viceversa. En general no existe mayor diferencia entre los perfiles de invierno y verano. Son pocos los retiros de este cluster que recortan en horario punta.

Composición

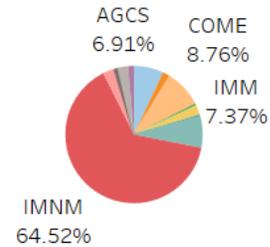


Figura 5.12: Ficha de caracterización de cluster denominado “Escalón Semanal”

Estacional

Perfil medio cluster



Descripción

Se caracteriza por tener un consumo en invierno es mucho menor al de verano. En invierno el consumo es bajo y en general se recorta en horario de punta. En cambio, en verano el consumo es alto y sigue perfiles similares al escalón semanal, pero las bajadas son menos pronunciadas.

Composición

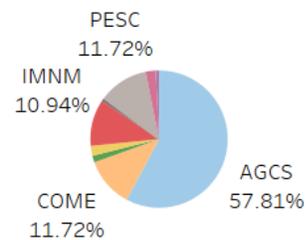


Figura 5.13: Ficha de caracterización de cluster denominado “Estacional”

Hospital

Perfil medio cluster



Descripción

Se caracteriza por mantener un consumo base alto durante todos los días y aumentarlo en horario laboral. El aumento es más pronunciado en los días de semana (lunes y trabajo). El domingo no existe un pronunciamiento claro en horario laboral y el sábado es más bien un promedio entre domingo y día de trabajo. El consumo no cambia en horario de almuerzo y no se recorta en horario punta.

Composición

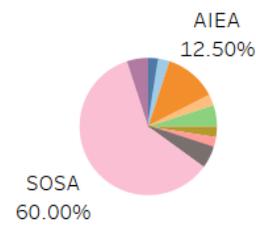


Figura 5.14: Ficha de caracterización de cluster denominado “Hospital”

Minero

Perfil medio cluster



Descripción

Se caracteriza por mantenerse prácticamente constantes a lo largo del día, para todos los tipos de día y en ambas estaciones. El consumo se mantiene en un nivel muy alto y en general no se presentan mayores variaciones. Al contrario de la mayoría de los otros perfiles, en general se ve una leve disminución del consumo en el horario laboral, lo cual puede tener relación con el encendido de luces artificiales.

Composición

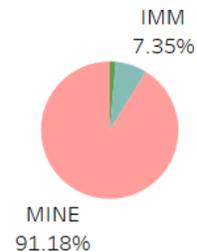


Figura 5.15: Ficha de caracterización de cluster denominado “Minero”

Oficina

Perfil medio cluster



Descripción

Se caracteriza tener un consumo base bajo, pero con un alto pronunciamiento durante el horario laboral en los días de semana. Los domingos se presenta un consumo prácticamente constante (consumo base), mientras que los sábados son un promedio entre trabajo y domingo. Los lunes no se diferencian de los otros días de trabajo. En general se puede distinguir una bajada en el horario de almuerzo y no se recorta en punta.

Composición

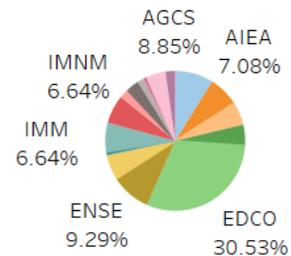


Figura 5.16: Ficha de caracterización de cluster denominado “Oficina”

Oficina industrial

Perfil medio cluster



Descripción

Se caracteriza por tener un consumo base medio en los días de semana con un alto pronunciamiento durante el horario laboral y consumo base bajo los fines de semana. Los días lunes son prácticamente iguales a los días de trabajo, mientras que los sábados son una transición hacia el domingo. Se distingue el recorte en punta de la mayoría de los retiros al igual que la disminución durante el horario de almuerzo.

Composición

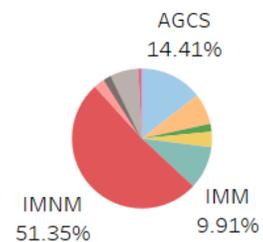
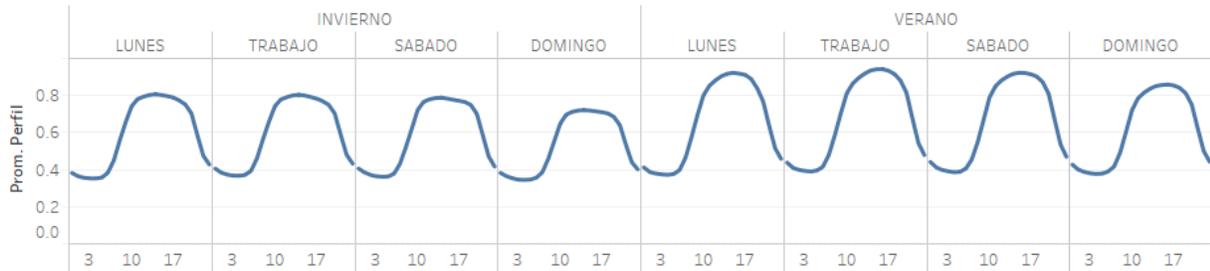


Figura 5.17: Ficha de caracterización de cluster denominado “Oficina Industrial”

Retail

Perfil medio cluster



Descripción

Se caracteriza por tener un consumo base bajo todos los días, pero con un alto pronunciamiento en el horario laboral. No presenta variaciones según tipo de día, pero en general existe un aumento de los peak de horario laboral durante el verano. No recortan en punta ni se distingue una disminución en el horario de almuerzo.

Composición

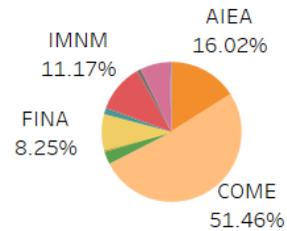


Figura 5.18: Ficha de caracterización de cluster denominado “Retail”

Típico industrial

Perfil medio cluster



Descripción

Se caracteriza por tener un consumo base alto durante la semana (muy alto en los días de trabajo), y se nota claramente el recorte en horas de punta. Las madrugadas de lunes y tardes de sábado demuestran una transición desde/hacia el domingo, donde el consumo es un poco más bajo. En general, el consumo en verano es casi igual al de invierno (salvo por los recortes en punta) y algunos demuestran una disminución en las horas de almuerzo.

Composición

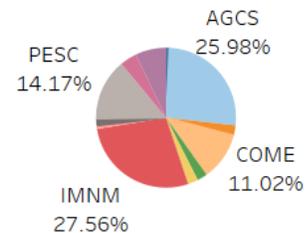


Figura 5.19: Ficha de caracterización de cluster denominado “Típico Industrial”

5.3. Entrevistas

5.3.1. Deducciones generales

Analizando las entrevistas se pueden deducir aspectos generales de los procesos productivos de los clientes de electricidad. Se observan patrones de consumo que se repiten en distintos clusters, que complementado con las entrevistas dan cuenta de ciertos índices de flexibilidad. Este análisis se detalla a continuación.

El pronunciamiento de consumo en el horario laboral implica que hay turnos desplazables. De lo conversado en las entrevistas se desprende que los procesos pueden ser operados en 1, 2 o 3 turnos de 8 horas cada uno. El pronunciamiento en horario laboral se debe a que hay un proceso que tiene solo 1 turno, es decir, no se opera de forma continua sino durante el horario laboral. Este turno podría ser adelantado o retrasado a cambio de una compensación a los trabajadores y a la empresa.

Las bajadas en horario de almuerzo demuestran que es un consumo no crítico operado por personas. Muchos de los perfiles cuyo consumo se pronuncia durante el horario laboral, denotan una disminución durante una hora que coincide con típicas horas de almuerzo. Como pueden hacer esa pausa, se deduce que el proceso no es crítico y depende de quien lo opera. Cuantificar la bajada de la hora de almuerzo permite deducir el volumen que podría dejar de consumir una empresa en el horario de pronunciamiento laboral, siempre que la remuneración ofrecida sea mayor que el costo de oportunidad de no producir en ese período.

Resulta muy útil saber la capacidad de los grupos electrógenos de respaldo que tienen los clientes que recortan en horario punta, debido a que evidencian su consumo crítico y el costo al que incurren para satisfacerlo. En general, los clientes que en horario punta prefieren autoabastecerse de electricidad mediante generadores diesel, no lo hacen por su consumo completo, sino por parte de este. En ese momento priorizan sus consumos críticos, demostrando además cuánto están dispuestos a pagar (precio diesel) por operar en esas horas. En los casos en que no ocupan generadores de respaldo, demuestran que ninguno de sus consumos son críticos y mediante una buena oferta estarían dispuestos a disminuirlos.

Los procesos que requieren operar en condiciones de bajas temperaturas pueden brindar flexibilidad, ya que debido a la inercia térmica, la temperatura se demora en subir. Esto se demuestra con que los entrevistados declaran que pueden cortar parcial o totalmente el suministro en las horas de punta sin sufrir mayores consecuencias. Por el contrario, los procesos que requieren operar en altas temperaturas demuestran ser poco flexibles debido a que son muy sensibles a las bajas de temperatura y que es muy difícil alcanzar las condiciones óptimas de trabajo. Reconocer la importancia de la temperatura en un proceso productivo, inmediatamente brinda nociones de su flexibilidad.

Los perfiles de consumo de cada retiro, reflejan el proceso productivo implícito. Al entrevistar a clientes que tienen retiros cuyo perfil pertenece al mismo cluster, se percibe coincidencia en la explicación que dan de su consumo, aún cuando son de sectores económicos diferentes. Por ejemplo, se entrevista a 3 clientes, de sectores económicos distintos, pero todos tienen retiros cuyo perfil es estacional. Al preguntarles por el consumo de esos retiros, todos coinciden en que se debe principalmente al bombeo de agua durante el año y riego en el verano. Con

esto se demuestra que resulta un buen método agrupar a los retiros según perfil de consumo, para evaluar sus procesos productivos y en consecuencia su flexibilidad.

Es necesario estudiar la demanda de los mercados donde los consumidores de electricidad son productores, para comprender mejor su flexibilidad. Tal y como declaran los clientes entrevistados, muchos de sus consumos dependen directamente de las ventas. Por ejemplo, en el caso de almacenar alimentos depende de la disponibilidad de estos y/o la época de cosecha, de la demanda de vino en el caso de embotelladora, o la cantidad de clientes en una tienda en el caso de retail. Predecir los comportamientos de los mercados permite negociar de mejor manera la flexibilidad que puedan ofrecer los clientes.

5.3.2. Análisis por cluster

Al asociar los resultados de las entrevistas con el análisis de perfiles exhibido en el apartado 5.2.2, se pueden extraer conclusiones de la flexibilidad de cada cluster estudiado. De los clusters importantes, existen dos que se asumen inflexibles sin entrevistarse con clientes y otro cuya flexibilidad no puede ser determinada mediante entrevistas. Estos son: minería, hospital y oficina.

A continuación se presentan las deducciones de flexibilidad de cada cluster.

Clusters inflexibles

La inflexibilidad del cluster "minería" se justifica mediante un estudio de Systep publicado en 2015 [20]. En este se analizan en profundidad los procesos mineros, determinando que no poseen flexibilidad, por lo que no podrían operar con respuesta en demanda. Puede ser que haya procesos específicos que no se hayan considerado en ese estudio, pero la entrevista que se plantea en este trabajo no tiene la capacidad de detectarlos. Por lo tanto, se considera que el cluster minero no puede ofrecer flexibilidad.

En el caso de "hospital", su mayor consumo proviene de servicios de salud, lo cual es difícilmente flexibilizable. Los procesos de hospitales son sensibles debido a que se atiende la salud de las personas, que muchas veces se encuentran en situaciones delicadas. Por lo tanto, los hospitales deben mantener sus condiciones para atender a personas y poder resolver emergencias. El resto de los retiros del cluster son la minoría y concentran muy poca energía, por lo que se consideran despreciables.

Oficina

En este se agrupan principalmente clientes cuyos consumos se deben a propiciar un ambiente de trabajo a ejecutivos. En general estos clientes son agrupaciones de empresas que tienen oficinas en el mismo condominio o edificio, lo que resulta difícil poder entrevistar a uno en particular. Sin embargo, se pueden deducir los consumos mediante la experiencia personal de trabajar en una oficina.

De la entrevista con el cliente 6, se deduce que los perfiles que siguen a la hora sol deben su consumo a la climatización. Esto resulta ser intuitivo para una oficina, ya que trabajan muchas personas y es necesario acondicionar el ambiente. Además, es claro que a medida

aumenta la radiación solar, es necesario más esfuerzo de climatización, lo que se ve reflejado en el perfil de consumo.

Por último, en muchos de los clientes pertenecientes a este cluster, se ve una disminución en el horario de almuerzo, lo cual se debe a que los ejecutivos dejan de trabajar en ese periodo. Estos trabajos probablemente requieren del uso de computador, por lo tanto este es el consumo que se ve reflejado.

La flexibilidad que pueden otorgar los clientes de este cluster va ligado al manejo de los dispositivos de climatización. Para coordinar la climatización de tantas oficinas se requiere una infraestructura de control más avanzada, como las exhibidas en el apartado 2.2.1. Por lo tanto, para determinar las restricciones y parámetros que modelan la flexibilidad de los clientes pertenecientes a este cluster, es necesario investigar los costos y la operación de los sistemas de control de energía centralizados.

Escalón semanal

El perfil de escalón semanal debe su forma a la ley laboral que impide que existan turnos productivos los días domingo mientras no se pueda justificar que la naturaleza del proceso obliga a que sea continuo. Esto hace que se vea en el consumo cómo encienden las máquinas los días lunes, mantienen los procesos mediante turnos rotativos continuos durante la semana y luego apagan los días sábado. Esto da la forma de escalón semanal.

Este cluster es uno de los que puede aportar menos flexibilidad. Esto queda en evidencia con la naturaleza del perfil, en que se ve que su mayoría no recorta en hora de punta, por lo que se deduce que están dispuestos a asumir el costo. Sin embargo, aquellos retiros que tienen pronunciamientos en horas del día, podrían ofrecer cierta flexibilidad al desplazar aquellos procesos que tienen menos de 3 turnos (no funcionan las 24 horas).

Estacional

Si bien el perfil estacional está compuesto por clientes provenientes de varios sectores, todos estos tienen alguna relación con el sector agrícola. Esto se deduce de analizar a los clientes, que gran parte son conocidos y se sabe que tienen procesos agrícolas, o bien su nombre dice explícitamente que es una sociedad agrícola, pero está catalogado como industria manufacturera no metálica o comercial. Por lo tanto, sus procesos de consumo debiesen ser muy parecidos.

Todos los entrevistados declaran ser inflexibles, pero evidentemente están dispuestos a disminuir el consumo debido a que hacen recorte en punta y no tienen grupo electrógeno de respaldo. Sus consumos se atribuyen a bombas que funcionan a máxima capacidad todo el año y riego en la temporada de verano. Para estimar realmente la flexibilidad hay que calcular el costo de oportunidad que les implica tener menos agua, o bien, que la remuneración permita aumentar la cantidad de bombas y así poder gestionar las horas en que operan en lugar de estar a máxima capacidad continuamente.

Oficina industrial

La entrevista con el cliente perteneciente a este sector sugiere que no hay flexibilidad, sin embargo el entrevistado no reconoce características importantes que demuestran lo contrario. Por ejemplo, que existe consumo asociado a un turno que no es continuo y que claramente su consumo no es crítico, ya que se detiene al almuerzo. Aún así queda claro que su mayor proceso es crítico y no gestionable puesto que en la hora punta suplen el 100 % del corte con un grupo electrógeno.

Se considera que la flexibilidad que pueden ofrecer los clientes de oficina industrial se atribuyen al desplazamiento de turnos que implican pronunciamiento del consumo base.

Retail

El cluster retail debe su perfil a que los retiros pertenecen a tiendas que atienden a personas. En su mayoría, los retiros de este cluster son de clientes pertenecientes al sector de “comercio al por mayor y menos”, donde se distinguen supermercados o tiendas conocidas. Sin embargo, también hay clientes pertenecientes a otros sectores, pero se puede deducir que sus consumos se atribuyen a tiendas. Esto pues se distinguen ferreterías que pertenecen a sector “industria manufacturera no metálica”, centros comerciales que pertenecen a “intermediación financiera” o “actividades inmobiliarias, empresariales y de alquiler”, entre otros. Por lo tanto, se puede decir que el cluster denominado retail, contiene a las tiendas y centros comerciales, pudiendo generalizarse sus condiciones de flexibilidad.

El perfil de consumo debe su forma a la climatización. Claramente se comporta de la misma forma que la radiación solar, es decir mientras más radiación hay más consumo. Esto se confirma mediante la entrevista, en que el cliente asegura que la mayor parte de su consumo se debe al uso de ventiladores y compresores para el aire acondicionado. Esto reafirma que los retiros de este cluster están asociados a tiendas comerciales.

Para que los clientes de este cluster accedan a flexibilizar su consumo se les tiene que asegurar que la temperatura siempre estará dentro de una banda determinada. Esto implica que las reducciones que ofrecen, tienen que ser seguidas de aumentos en el consumo, para no desestabilizar. Por ende, se modelan como retiros pertenecientes al conjunto de demanda desplazable.

La cantidad de energía que se puede disminuir y la prolongación de esta, está relacionada con la transferencia de calor desde afuera de la tienda y la de los consumidores que están adentro. Estas transferencias se pueden cuantificar según la temperatura externa y el número de personas dentro de la tienda. Por lo tanto, la flexibilidad está en función del clima y de la demanda del local.

Típico industrial

El cluster típico industrial es el que contiene más procesos diferentes, por lo que no es muy fidedigno asociarle a todos el mismo potencial de flexibilidad. A diferencia de los otros clusters, este se conforma por clientes pertenecientes a muchos sectores distintos, los cuales no necesariamente tienen características comunes en cuanto a sus procesos. Por esto, no se

puede afirmar que todos puedan dar una flexibilidad similar, por lo que habría que atribuir más características a estos retiros que permitan intuir sus procesos.

Los entrevistados de este cluster coinciden en que su procesos se deben a tratamiento de productos que deben mantenerse a bajas temperaturas. Estos retiros son muy gestionables: si se trata de congelados, pueden disminuir la totalidad de su consumo por un período prolongado y si se trata de temperaturas que bordean los 0 grados, pueden disminuir hasta un 85 % en períodos dependientes de la temperatura de ambiente. A raíz de los resultados de este trabajo se puede concluir que este es el cluster que puede brindar mayor flexibilidad, pero esto debiese ser comprobado mediante más entrevistas.

Capítulo 6

Conclusiones y trabajo futuro

En este trabajo se propone una segmentación de la demanda eléctrica de los clientes libres del SEN, según su perfil de consumo y sector económico, con el fin de modelar su flexibilidad. Para esto se propone un modelo de despacho con demanda flexible, cuyos parámetros son caracterizados mediante entrevistas a clientes pertenecientes a las distintas segmentaciones. Del proceso se extraen una serie de conclusiones que son expuestas a continuación.

La metodología presentada para la segmentación demuestra tener excelentes resultados. Se logran identificar agrupaciones de clientes según perfil de consumo y sector económico, de las cuales destacan 8 que concentran la mayor parte de los retiros. Además de poder usar esta información para el modelo de DR, la metodología permite entender mejor a los clientes y sirve para pensar en otros negocios y modelos de negocios específicos para cada segmento. Esto resulta ser información relevante para empresas comercializadoras de energía.

Los perfiles de consumo de los retiros dicen mucho acerca de la flexibilidad, no así el sector económico que se le designa al cliente. Mediante las entrevistas, se detecta que perfiles de consumo similares reflejan procesos productivos similares, por lo que su flexibilidad también debiese ser similar. No obstante, estos retiros podrían corresponder a sectores económicos diferentes, debido a que las empresas pueden pertenecer a distintos rubros y acá solo se les asigna uno. Por lo tanto, se concluye que para determinar la flexibilidad de un proceso es más importante considerar el perfil de consumo que el sector económico del cliente.

Para obtener mejores resultados es indispensable que los retiros estén separados por planta y no agrupados. Las plantas que realizan el mismo proceso productivo tienen un perfil de consumo muy parecido y por lo tanto su flexibilidad también debiese serlo. De esta forma, mientras más diferenciados estén los procesos, resultan mejores agrupaciones por perfil, con lo que se pueden extraer conclusiones de flexibilidad más precisas.

Con el fin de discernir mejor entre los procesos que se llevan a cabo en las plantas con perfil parecido, es necesario saber a qué se dedica la empresa que efectúa el retiro con más detalle que un sector económico del SII. Existen empresas cuyas actividades podrían atribuirse a más de un sector económico, por lo que asignarle uno solo, provoca que se asuma que todos los retiros de esta, pertenecen al mismo sector. Lo anterior provoca distorsiones en el análisis,

donde quedan clusters conformados por muchos sectores, pero en realidad todos los retiros se dedican a lo mismo.

La limpieza de datos es indispensable para obtener buenos resultados. Uno de los grandes problemas en la ciencia de datos, es que las bases siempre contienen muchos errores, los cuales introducen ruido a los modelos, obteniendo resultados erróneos. Lo propio ocurre con la base de datos ocupada, en que existen errores en los nombres de los clientes y mediciones. El proceso de limpieza de datos es parte importante del trabajo y sin esta los resultados serían menos fidedignos.

La información disponible en el momento en que se hace el estudio limita los resultados. La migración de clientes regulados a libre es un fenómeno reciente, que en el último tiempo ha experimentado un alza considerable. Esto implica que la base de datos podría tener muchos más retiros y con eso caracterizar de mejor forma los perfiles del sistema.

Las entrevistas a los clientes, no pueden ser guiadas con una pauta estricta, sino que debe ser acorde a cada perfil. De la experiencia que se adquiere entrevistando, los clientes son más receptivos cuando se les pregunta de procesos específicos que se pueden ver en su patrón de consumo. Mediante una entrevista más precisa, se obtienen mejores nociones de los procesos de los clientes.

6.1. Trabajo futuro

Naturalmente la continuación del trabajo es definir un modelo del SEN y simular el despacho con demanda flexible usando los resultados obtenidos. Para esto hay que confeccionar los clientes genéricos según la topología de la red del modelo y luego calcular sus parámetros de flexibilidad en función de las características aquí presentadas. Sin embargo, antes de seguir con este paso, sería mejor repetir el experimento con las mejoras propuestas, como diferenciar los retiros por planta y luego caracterizarlos por rubro, hacer un manejo de outliers más sofisticado, descomponer los perfiles por trimestres en lugar de semestres, entre otras. De esta forma se pueden obtener mejores resultados y así caracterizar mejor la demanda flexible para el modelo.

Con el fin de hacer el modelo más detallado, se pueden considerar todos los clusters que aparecen en lugar de seleccionar solo los más importantes. Si bien esto puede significar una pérdida de generalidad, permite segmentar los perfiles de forma más precisa, y de requerirse más generalidad, simplemente pueden juntarse los clusters más parecidos desde el punto de vista del perfil y la flexibilidad. Así, se reducen los clientes considerados como outliers y por lo tanto, la demanda agregada puede ofrecer más flexibilidad.

Para profundizar en el entendimiento de la flexibilidad de los perfiles sería necesario entrevistar a más clientes. Mediante más entrevistas se pueden validar de mejor manera las inferencias de los perfiles y se pueden encontrar nuevas explicaciones de los perfiles que reflejan un mayor potencial de flexibilidad. Más entrevistas permiten que se puedan obtener resultados más precisos del modelo operado con Respuesta en Demanda.

Luego de simular la operación del despacho con Respuesta en Demanda, sigue analizar los

resultados para tener nociones de qué empresas se beneficiarían. Incorporando las mejoras propuestas, se puede simular un año de operación pasada, pero flexibilizando la demanda. Posteriormente se comparan los resultados con el mismo caso, pero en un escenario con demanda fija. Esto permite analizar cuáles son las empresas que se beneficiarían de esto y cuáles son los mejores clientes para ofrecer este servicio.

Por último, podría repetirse la simulación, pero esta vez en un escenario futuro. Para esto se necesitan predicciones de la demanda y la generación, donde se debiesen incorporar componentes que se desprecian al día de hoy, como vehículos eléctricos, almacenamiento y una masiva incorporación de generación distribuida. Este escenario permite a los comercializadores de energía tener una noción de los clientes que debiesen ir a buscar.

Bibliografía

- [1] EnelX, “¿qué es la respuesta a la demanda y por qué es conveniente?” <https://www.enelx.com/es/faq/eindustry/la-respuesta-a-la-demanda>. Visitado: 04/06/2019.
- [2] C. N. de Energía, “Capacidad instalada.” <http://energiaabierta.cl/visualizaciones/capacidad-instalada>. Visitado: 06/06/2019.
- [3] G. de Chile, “Generación eléctrica en Chile.” <http://generadoras.cl/generacion-electrica-en-chile>. Visitado: 15/05/2019.
- [4] C. E. Nacional, “Nosotros.” <https://www.coordinador.cl/nosotros/>. Visitado: 06/06/2019.
- [5] F. Y. R. MINISTERIO DE ECONOMÍA, “Decreto supremo 291,” Ago 2008. Última Modificación: Ago 2013.
- [6] C. N. de Energía, “Resolución exenta n669: Cálculo y determinación de transferencias económicas de energía,” nov 2017.
- [7] F. Y. R. MINISTERIO DE ECONOMÍA, “Dfl 4: Ley general de servicios eléctricos,” feb 2007. Última Modificación: Nov 2018.
- [8] D. D. P. CDECSIC, “Guía de aplicación: Transferencias de energía,” Jun 2016.
- [9] N. G. Paterakis, O. Erdinç, and J. P. Catalão, “An overview of demand response: Key elements and international experience,” *Renewable and Sustainable Energy Reviews*, vol. 69, pp. 871 – 891, 2017.
- [10] “Introducing demand response to enhance competition in the national electricity market of Singapore,” Jun 2016.
- [11] A. D. IMPRENSA, “Programa de resposta da demanda para consumidores industriais é aprovado pela ANEEL,” *ANEEL*, Nov 2017.
- [12] Narvik, “Estudio análisis y determinación de horario de punta en el sistema eléctrico nacional.” web CNE, dec 2018. Publicado por CNE: May 2019.
- [13] A. K. Jain, “Data clustering: 50 years beyond k-means,” *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651 – 666, 2010. Award winning papers from the 19th International

Conference on Pattern Recognition (ICPR).

- [14] N. Viswarupan, “K-means data clustering,” Jul 2017.
- [15] Z. Jaadi, “A step by step explanation of principal component analysis.” <https://towardsdatascience.com/an-introduction-to-t-sne-with-python-example-5a3a293108d1>, 2019. Visitado: 10/05/2019.
- [16] L. van der Maaten and G. Hinton, “Visualizing high-dimensional data using t-sne,” *Journal of machine learning research*, 2008.
- [17] GoogleTechTalks, “Visualizing data using t-sne.” <https://www.youtube.com/watch?v=RJVL80Gg31A>, 2013. Visitado: 10/05/2019.
- [18] A. keitakurita, “Paper dissected: "visualizing data using t-sne.explained,"” Sep 2018.
- [19] A. Violante, “An introduction to t-sne with python example.” <https://towardsdatascience.com/an-introduction-to-t-sne-with-python-example-5a3a293108d1>, 2018. Visitado: 10/05/2019.
- [20] O. Morales, S. Mocarquer, H. Rudnick, and P. Miquel, “Benefits of industrial demand response in the chilean electricity market.” web Systep, 2015.

Anexo A

Modelo propuesto

Conjuntos

- $T := \{\text{Horas del año}\}$
- $Tm(m) := \{\text{Subconjunto de } T \text{ que diferencia las horas del mes } m\}$
- $Td(d) := \{\text{Subconjunto de } T \text{ que diferencia las horas del día } d\}$
- $M := \{\text{Meses del año}\}$
- $H := \{\text{Días del año}\}$
- $G := \{\text{Generadores despachables}\}$
- $GND := \{\text{Generadores no despachables}\}$
- $B := \{\text{Barras del sistema}\}$
- $L := \{\text{Líneas de transmisión}\}$
- $R := \{\text{Clientes genéricos}\}$
- $DS := \{\text{Clientes genéricos que desplazan demanda}\}$
- $N(j) := \{\text{Escalones de flexibilidad de cliente genérico } j\}$

Variables de decisión

- $P_{i,t}$: Generación de central i en instante t
- $V_{t,b}$, Vertimiento en el instante t en la barra b
- $F_{t,l}$: Flujo por línea l en instante t
- $\theta_{t,b}$: Ángulo de voltaje en p.u. de barra b en instante t
- $D_{j,t}$: Demanda del cliente j en el instante t
- $\bar{\Delta}_{j,t,n}$: Porcentaje de aumento en escalón n de cliente j en instante t
- $\underline{\Delta}_{j,t,n}$: Porcentaje de reducción en escalón n de cliente j en instante t

Parámetros

- $C_{i,t}$: Costo variable de la central i en el mes del instante t
- $Pmax_{i,m}$: Potencia máxima de generador i en mes m

- $Pmin_{i,m}$: Potencia mínima de generador i en mes m
- $E_{j,t}$: Energía inyectada por generador no despachable j en instante t
- $Eemb_{k,m}$: Energía máxima disponible para embalse k en mes m
- $Fmax_l$: Flujo máximo de línea l
- X_l : Impedancia de línea l
- $\bar{\zeta}_{j,t,n}$: Costo de aumentar una unidad de energía del cliente j en instante t para escalón n
- $\underline{\zeta}_{j,t,n}$: Costo de reducir una unidad de energía del cliente j en instante t para escalón n
- $\bar{D}_{i,t}$: Demanda real (o esperada) del clientes genérico i en el instante t
- $\bar{\Delta}max_{j,n}$: Porcentaje máximo de aumento en escalón n de cliente j en instante t
- $\underline{\Delta}max_{j,n}$: Porcentaje mínimo de reducción en escalón n de cliente j en instante t

Modelo

$$\min \sum_{t \in T} \left(\sum_{i \in G} C_{i,t} \cdot P_{i,t} + \sum_{j \in R} \bar{D}_{j,t} \left(\sum_{n \in N} (\bar{\zeta}_{j,t,n} \cdot \bar{\Delta}_{j,t,n} + \underline{\zeta}_{j,t,n} \cdot \underline{\Delta}_{j,t,n}) \right) \right) \quad (A.1)$$

s.a.

$$D_{j,t} = \bar{D}_{j,t} \left(1 + \sum_{n \in N} (\bar{\Delta}_{j,t,n} - \underline{\Delta}_{j,t,n}) \right) \quad \forall (j \in R)(t \in T) \quad (A.2)$$

$$\sum_{t \in Td(d)} D_{j,t} = \sum_{t \in Td(d)} \bar{D}_{j,t} \quad \forall (j \in DS)(d \in H) \quad (A.3)$$

$$\sum_{t \in Tm(m)} P_{i,t} \leq Eemb_{i,m} \quad \forall (i \in G(Embalses))(m \in M) \quad (A.4)$$

$$\begin{aligned} \bar{\Delta}_{j,t,n} &\leq \bar{\Delta}max_{j,t,n} \\ \underline{\Delta}_{j,t,n} &\leq \underline{\Delta}max_{j,t,n} \quad \forall (j \in R)(t \in T)(n \in N) \end{aligned} \quad (A.5)$$

$$Pmin_{i,t} \leq P_{i,t} \leq Pmax_{i,t} \quad \forall (i \in G)(t \in T) \quad (A.6)$$

$$\begin{aligned} F_{t,l} &\leq Fmax_{t,l} \\ -F_{t,l} &\leq Fmax_{t,l} \quad \forall (l \in L)(t \in T) \end{aligned} \quad (A.7)$$

$$D_{j,t}, P_{i,t}, V_{t,b}, \bar{\Delta}_{j,t,n}, \underline{\Delta}_{j,t,n} \geq 0 \quad \forall (j \in R)(i \in G)(t \in T)(b \in B)(n \in N) \quad (A.8)$$

Anexo B

Entrevista

B.1. Introducción

Mi nombre es Diego Guiraldes Deck, soy egresado de Ingeniería Eléctrica de la Universidad de Chile y me encuentra realizando mi memoria de título.

Estoy haciendo un estudio para incorporar una nueva técnica que permite ocupar electricidad de forma eficiente, lo que se traduce en ahorros tanto para los clientes como para los proveedores de energía. Esta se conoce internacionalmente como “Demand Response” (respuesta en demanda), la cual actualmente está siendo implementada en mercados eléctricos desarrollados y competitivos, como en California, Nueva York, Londres, etc.

En términos generales, este concepto se refiere a los cambios en el patrón de consumo de los clientes del sistema eléctrico en función de las señales del mercado. Un ejemplo de esto sería informarle a un cliente que existe mucha energía solar en el sistema, entonces este podría adelantar o atrasar la operación de una máquina a cambio de un pago o compensación que se ve reflejado en su cuenta final.

Para implementar esta innovadora técnica se requiere una coordinación inteligente entre los procesos de los consumidores de electricidad y la generación de esta. Es por esto, que mediante una serie de entrevistas anónimas se modelará la capacidad de flexibilidad de clientes de distintos sectores económicos para tener una noción más acabada de los efectos que podría tener “Demand Response” en el sistema chileno.

Este estudio se enmarca bajo mi trabajo de título, que está siendo apoyado por Colbún, cuyos resultados podrían ser enviados a quienes quieran participar de la entrevista. En ningún caso quedará registrado el nombre de la empresa, ni de la persona que dio la entrevista, solamente se indicará su sector económico y el tipo de perfil de consumo, como por ejemplo, “minería, tipo 1”, “agricultura tipo 2”, etc.

B.2. Entrevista

La entrevista consiste en 3 partes: capacidad de flexibilidad, de disminución neta y de desplazamiento.

1. En su empresa, ¿se preocupan de los horarios en que consumen?

Capacidad de flexibilidad

2. ¿Poseen consumos eléctricos no críticos que puedan gestionarse con flexibilidad? ¿Tiene rangos horarios en que le sea factible hacer una gestión de dichos consumos?

- Por ejemplo: una bomba de agua que pueda prenderse en distintas horas, una máquina que no haya problema de apagar en ciertos intervalos, algo que pueda esperar a encenderlo más tarde.

3. ¿Poseen consumos eléctricos críticos que puedan gestionarse considerando un aviso con suficiente anticipación? ¿Tienen rangos horarios en que le sea factible hacer una gestión de dichos consumos?

- Por ejemplo: una máquina que inicia un proceso a una determinada hora,

Capacidad de disminución

4. ¿Cuánto consumo toleraría disminuir en cada uno de los rangos horarios? ¿Cuánto esperaría como compensación/pago?

(a) Puede ser energía o porcentaje aproximado de disminución

(b) Pueden ser distintos precios a distintas horas

Capacidad de desplazamiento

5. ¿Le es posible desplazar sus consumos? ¿Cuándo y cuánto? ¿Cuánto esperaría como compensación/pago? Indicar volúmenes, rangos horarios y duración del desplazamiento

- Por desplazamiento me refiero a consumir una fracción de energía en otro momento
- Por ejemplo: puedo desplazar el 20% de mi consumo desde la 20h a las 22h
- Por ejemplo: consumos dependiendo de turnos.
- El pago puede ser el equivalente de comprar un estanque y bomba mas grande, contratar más personal, o puede no tener costo.

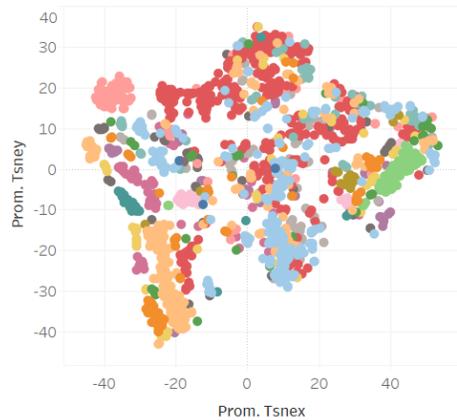
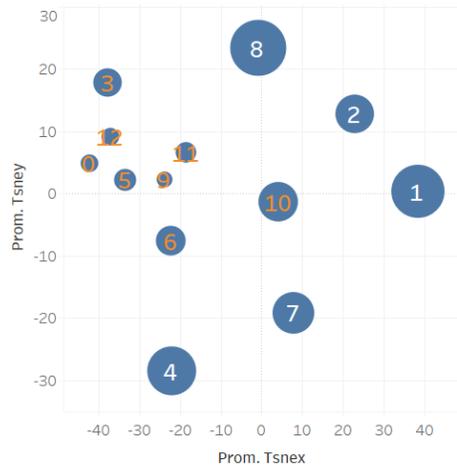
Anexo C

Tablas y figuras complementarias

C.1. Resultados

En esta sección se muestran gráficos de resultados de la clusterización que no se incluyen en el informe. Estos pueden servir de información adicional.

Representación con TSNE



Descomposición de clusters en sectores

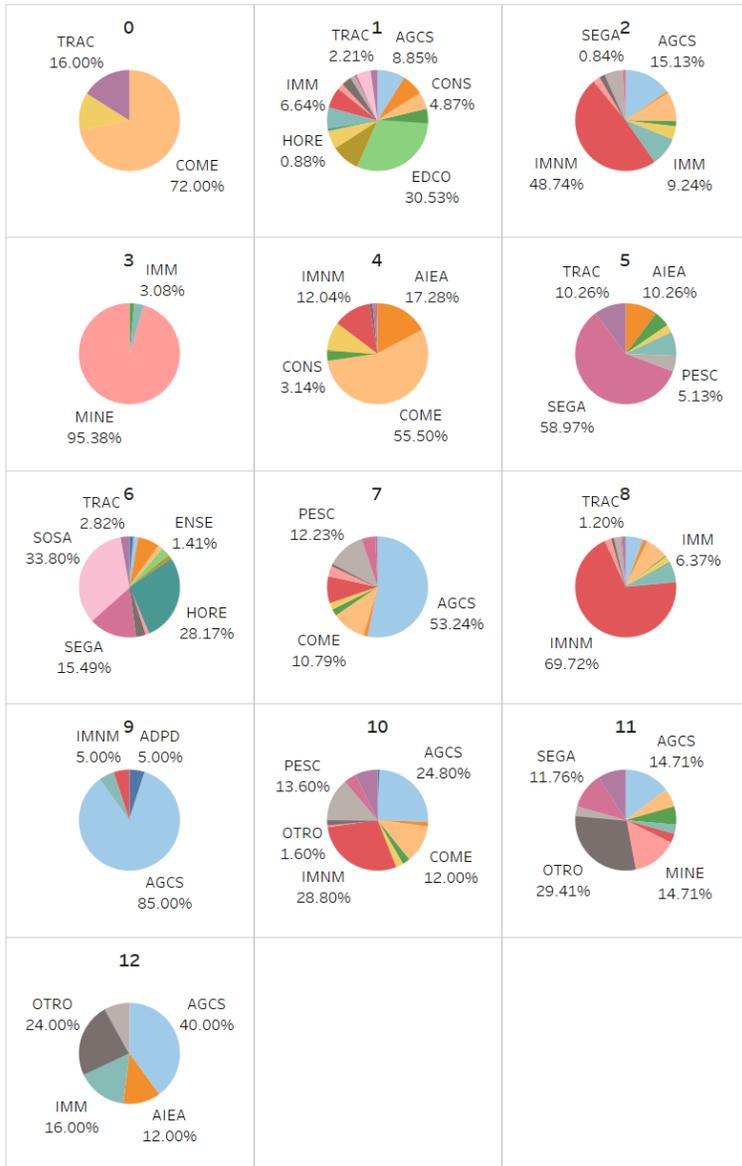


Figura C.1: Descomposición de clusters en sectores para clusterización de $\varepsilon = 0,188$ y $min_samples = 15$. Se relacionan los clusters con los sectores económicos de los retiros que contienen. En el sector izquierdo de la figura se muestran dos representaciones del plano generado con tsne: la superior muestra la ubicación y tamaño (según cantidad de retiros) de los clusters y la inferior, muestra todos los retiros coloreados por sector. A la derecha se muestra un gráfico de pie asociado a cada cluster, donde se resalta la descomposición porcentual de los sectores predominantes.

Perfiles promedio por cluster

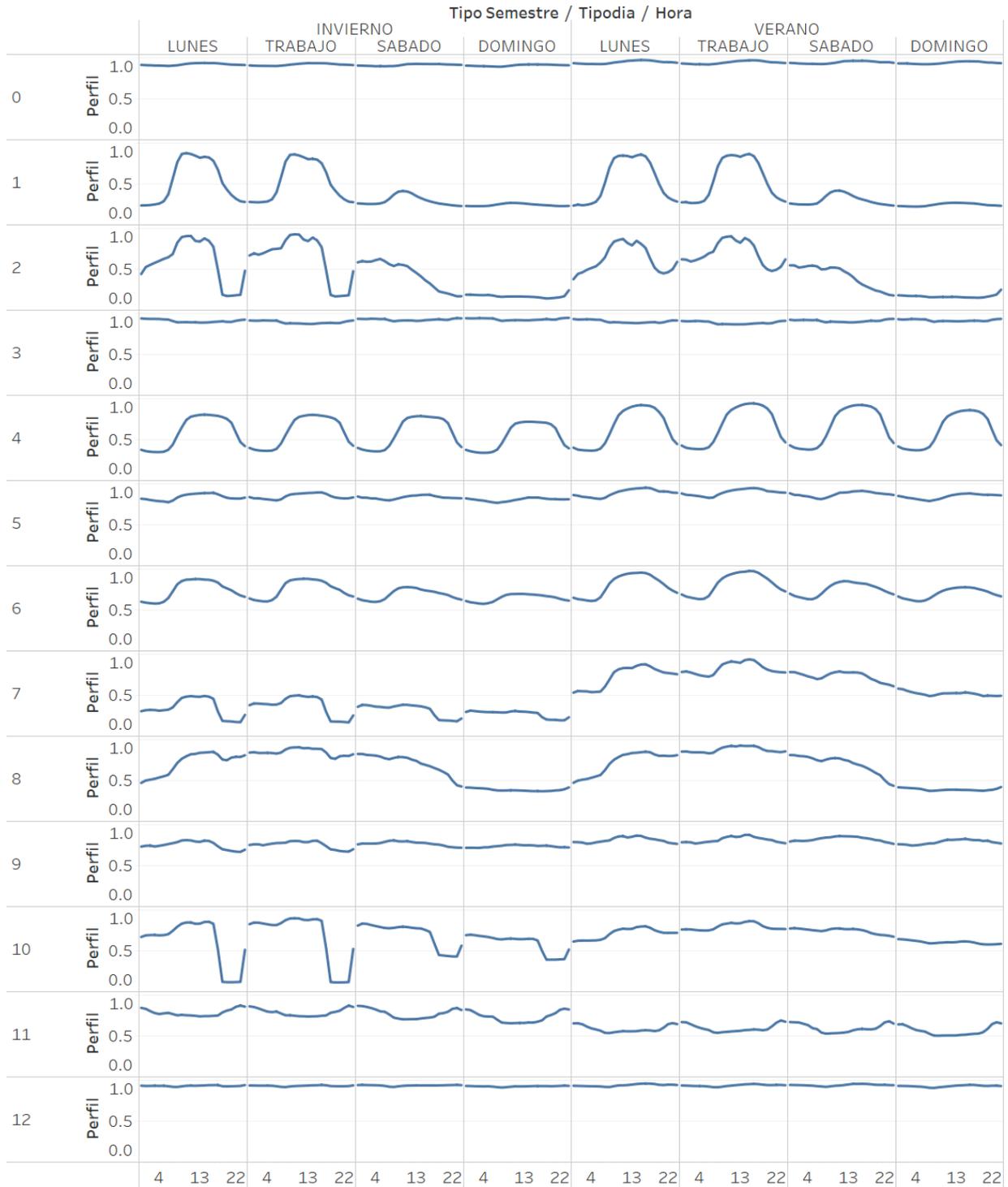


Figura C.2: Perfiles promedio de cada cluster para la descomposición temporal propuesta (tipo de semestre, tipo de día y hora del día), para la clusterización con parámetros $\varepsilon = 0,188$ y $min_samples = 15$.

Perfiles promedio por cluster

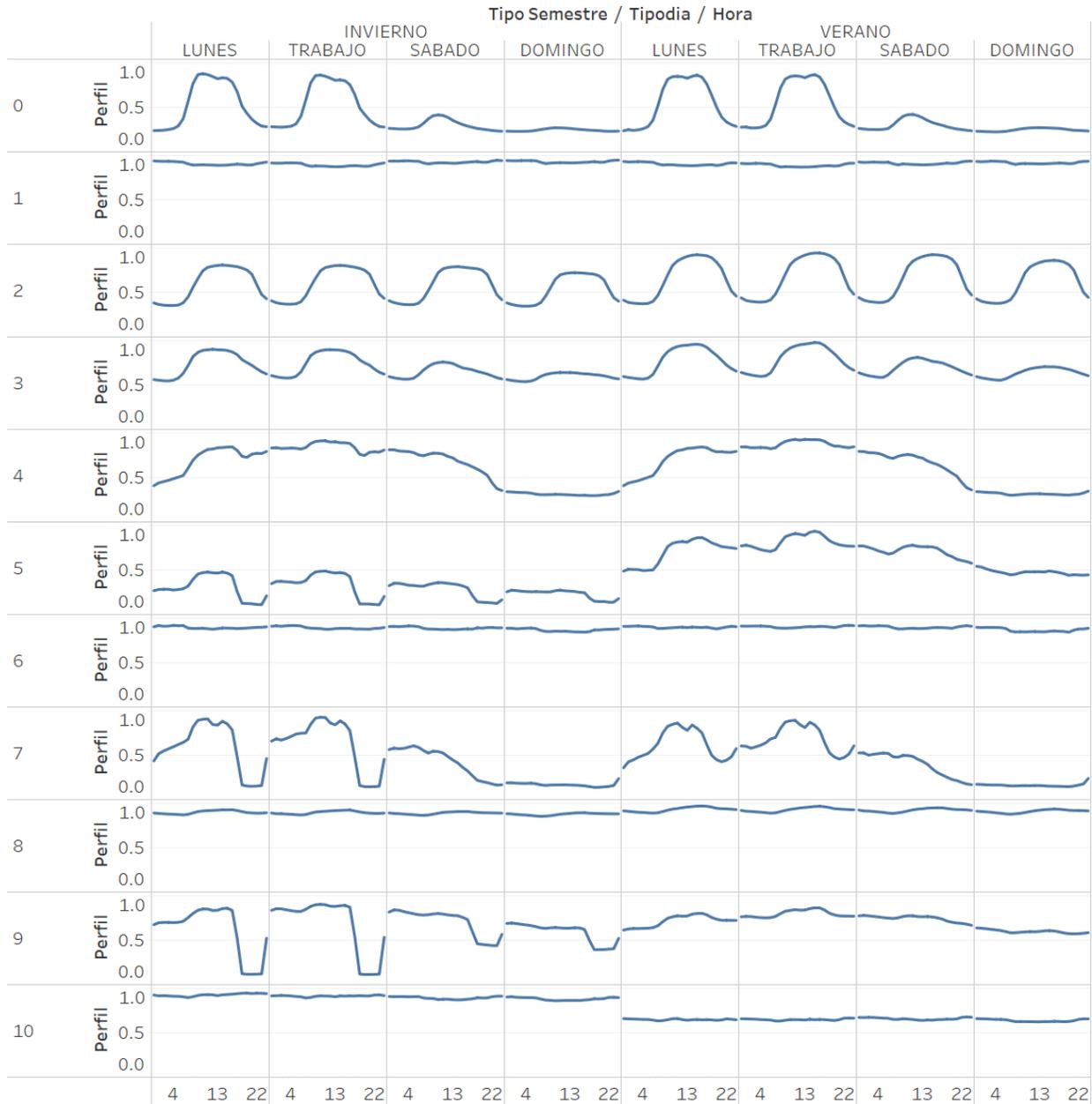


Figura C.3: Perfiles promedio de cada cluster para la descomposición temporal propuesta (tipo de semestre, tipo de día y hora del día), para la clusterización con parámetros $\varepsilon = 0,245$ y $min_samples = 30$.

C.2. Análisis

En esta sección se muestran las tablas que se usan para construir la descripciones que se presentan en las fichas de los clusters.

	Consumo de madrugada			
	Invierno		Verano	
	Lunes	Trabajo	Lunes	Trabajo
Escalón semanal	medio	muy alto	medio	muy alto
Estacional	bajo	bajo	medio	alto
Hospital	medio	medio	medio	alto
Minero	muy alto	muy alto	muy alto	muy alto
Oficina	bajo	bajo	bajo	bajo
Oficina industrial	medio	alto	medio	alto
Retail	bajo	bajo	bajo	medio
Típico industrial	alto	muy alto	alto	alto

Tabla C.1: Comparación de características de consumo de madrugada para los días de semana de invierno y verano.

	Promedio invierno	Promedio verano
Escalón semanal	alto	alto
Estacional	bajo	alto
Hospital	alto	alto
Minero	muy alto	muy alto
Oficina	medio	medio
Oficina industrial	medio	medio
Retail	medio	alto
Típico industrial	alto	alto

Tabla C.2: Comparación de consumo promedio de invierno y verano

	Recorte en punta
Escalón semanal	no
Estacional	si
Hospital	no
Minero	no
Oficina	no
Oficina industrial	si
Retail	no
Típico industrial	si

Tabla C.3: Identificación de retiros que recortan en punta.

	Tipo de consumo		
	Sábado	Lunes	Fin de semana
Escalón semanal	transición	paulatino	medio
Estacional	transición	paulatino	medio
Hospital	promedio	abrupto	alto
Minero	igual a trabajo	igual a trabajo	muy alto
Oficina	promedio	abrupto	bajo
Oficina industrial	transición	abrupto	bajo
Retail	igual a trabajo	igual a trabajo	medio

Tabla C.4: Caracterización de tipo de transición de sábado a domingo y domingo a lunes, en conjunto con consumo medio de fin de semana.

	Horaio de almuerzo
Escalón semanal	poco o nada
Estacional	claro
Hospital	poco o nada
Minero	poco o nada
Oficina	claro
Oficina industrial	claro
Retail	poco o nada
Típico industrial	algunos

Tabla C.5: Detección de disminución en horario de almuerzo.

	Pronunciamento en horario de trabajo
Escalón semanal	no
Estacional	medio
Hospital	medio
Minero	no
Oficina	pronunciado
Oficina industrial	pronunciado
Retail	pronunciado
Típico industrial	no

Tabla C.6: Detección de pronunciamento del consumo durante el horario laboral.