



**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FISICAS Y MATEMATICAS
DEPARTAMENTO DE INGENIERIA ELECTRICA**

**ANALISIS DE ESTACIONALIDAD DE LA CURVA DE DEMANDA
PARA CLIENTES BT-1**

MEMORIA PARA OPTAR AL TITULO DE INGENIERO CIVIL ELECTRICISTA

DANIEL IGNACIO GOMEZ SAGNER

**PROFESOR GUIA:
GUILLERMO PEREZ DEL RIO**

**MIEMBROS DE LA COMISION:
ARIEL VALDENEGRO ESPINOZA
RODRIGO PALMA BEHNKE**

**OCTUBRE 2007
SANTIAGO, CHILE**

RESUMEN DE LA MEMORIA
PARA OPTAR AL TITULO DE
INGENIERO CIVIL ELECTRICISTA
POR: DANIEL IGNACIO GOMEZ SAGNER
FECHA: 23/10/2007
PROF. GUIA: GUILLERMO PEREZ DEL RIO

“ANALISIS DE ESTACIONALIDAD DE LA CURVA DE DEMANDA PARA CLIENTES BT-1”

En el marco de la caracterización de la curva de demanda para los clientes sin medidor horario de demanda surge una serie de dificultades. La principal de ellas es la limitación económica, ya que instalar un medidor horario para cada cliente involucraría un monto de inversión muy superior a los beneficios que otorga el conocer el comportamiento de consumo de estos clientes. Es por ello, que las campañas de medición destinadas a este tipo de clientes y las técnicas estadísticas que las certifican toman una gran relevancia.

El objetivo de esta memoria de título es presentar una metodología que permita caracterizar la curva de demanda de clientes BT-1. Para estos fines se presenta un resumen con los campos de aplicación de *Load Research* y se presentan las principales técnicas estadísticas que permiten efectuar una campaña de medición representativa para un universo de clientes ya definido. Se detalla el proceso de desarrollar una minería de datos, destacando las técnicas de agrupamiento o *clustering* comúnmente utilizadas en investigaciones de esta índole.

Utilizando la base de datos de clientes BT-1 de Chilectra, se propone un muestreo estadístico representativo para efectuar una campaña de medición. Posteriormente, se realiza un análisis de la estacionalidad de la curva de demanda de los clientes que poseen tarifa BT-1 de la empresa Chilectra S.A. a través de perfiles típicos y de curvas agregadas por estrato de consumo, utilizando una base de 1.050 mediciones obtenidas en el año 1999. A partir de estos resultados, se efectúa un estudio sobre algunos supuestos del cobro por energía adicional de invierno planteados por la Comisión Nacional de Energía en 1981, que dice que los clientes de bajos consumos tienen un comportamiento estable en la potencia a lo largo del año y que son los clientes de mayores consumos los responsables de producir el “sobrepick” de invierno.

Como resultado a esta investigación, se obtiene una caracterización de los clientes BT-1 basada en la existencia de cuatro perfiles típicos de consumo. Tres de ellos poseen la forma característica de un consumo residencial, que se diferencian en el uso medio de la potencia, destacando una estacionalidad horaria de la máxima demanda y una variación estacional del factor de carga. El perfil de consumo restante corresponde a una curva con la forma de un consumo comercial, que no posee estacionalidad horaria. Finalmente, se demuestra que efectivamente los clientes de mayor consumo, de la muestra de mediciones de 1999, son los responsables de producir el “sobrepick” de invierno corroborando el supuesto de la Comisión.

Como conclusión se propone ampliar el espectro de alternativas de investigación incorporando al análisis otras variables, tales como el nivel de tensión o la potencia reactiva. Adicionalmente se puede considerar variables no eléctricas, como la temperatura y la luminosidad o variables a través de encuestas que reflejen la demanda de electricidad por el uso de equipos electrodomésticos y así encontrar nuevas fuentes para futuros desarrollos.

A mis padres

Agradecimientos

A través de estas líneas quiero entregar mis más sinceros agradecimientos a todas las personas que me han ayudado de alguna u otra forma a lograr este objetivo.

En primer lugar agradezco a toda mi familia, especialmente a mis padres, que a través de su esfuerzo y sus consejos han hecho posible que logre mi objetivo.

En segundo lugar quisiera agradecer a Don Guillermo Pérez del Río por su gran colaboración en el desarrollo de este trabajo, pero por sobretodo, por brindarme esta oportunidad y depositar toda su confianza en mí. Tampoco puedo dejar de agradecer a Gabriela, Rosa y Luigi, del área de Regulación de Tarifas de la empresa Chilectra por su constante apoyo y su ayuda entregada en la búsqueda de información.

También quisiera extender mis agradecimientos a los profesores Ariel Valdenegro y Rodrigo Palma del Departamento de Ingeniería Eléctrica de la Universidad de Chile, por sus valiosos aportes.

Agradezco a mis amigos de la universidad, Alfredo, Felipe, Oscar y Pablo por estos años de amistad, que seguramente seguirán para toda la vida.

De igual modo, agradezco mis tíos Jorge y Elizabeth que me entregaron su apoyo en los momentos en que lo necesité durante mi estadía en la Universidad.

Finalmente, agradezco a Fernando Flatow por su buena disposición y colaboración en el comienzo de este trabajo, sin su ayuda este desafío no habría sido posible.

Indice General

INDICE GENERAL	5
INDICE DE FIGURAS	7
INDICE DE TABLAS	8
1 INTRODUCCIÓN	9
1.1 MOTIVACIÓN.....	9
1.2 OBJETIVOS Y ALCANCES.....	9
1.3 ESTRUCTURA.....	10
2 CARACTERÍSTICAS DEL SECTOR ELÉCTRICO CHILENO	12
2.1 GENERACIÓN.....	12
2.2 TRANSMISIÓN.....	12
2.3 DISTRIBUCIÓN.....	13
2.3.1 Tarifas de distribución.....	13
2.4 MODELO ELÉCTRICO CHILENO.....	15
2.4.1 Mercado Spot.....	16
2.4.2 Mercado de clientes libres.....	16
2.4.3 Mercado regulado.....	16
2.5 CALIDAD DE SERVICIO ELÉCTRICO.....	16
3 CARACTERIZACIÓN DE LA CURVA DE CARGA	18
3.1 FUNDAMENTOS Y APLICACIONES.....	18
3.1.1 Diseño de tarifas.....	18
3.1.2 Planificación del sistema.....	18
3.1.3 Operación y mantenimiento del sistema.....	19
3.1.4 Marketing.....	19
3.1.5 Load management.....	19
3.2 CAMPAÑAS DE MEDICIÓN.....	20
3.2.1 Técnicas de muestreo.....	20
3.2.2 Implementación de la campaña.....	26
3.3 MINERÍA DE DATOS.....	27
3.3.1 Etapas de una minería de datos.....	27
3.3.2 Técnicas de agrupamiento.....	28
3.4 ESQUEMA GENERAL PROPUESTO.....	37
4 EXPERIENCIA PREVIA EN CAMPAÑAS DE MEDICIÓN	38
4.1 CASO TAIWÁN.....	38
4.1.1 Diseño de la muestra.....	38
4.1.2 Obtención de las curvas características.....	39
4.1.3 Aplicación.....	40
4.2 CASO MARYLAND.....	41
4.3 CASO PERÚ.....	42
4.3.1 Diseño de la muestra.....	42
4.3.2 Obtención de las curvas.....	44
4.3.3 Aplicación.....	45
4.4 CASO CHILECTRA.....	46
4.4.1 Diseño de la muestra.....	47
4.4.2 Aplicación.....	47
5 DISEÑO DE LA MUESTRA PARA UNA CAMPAÑA DE MEDICIÓN EN CHILECTRA	50
5.1 INFORMACIÓN DE LA POBLACIÓN OBJETIVO.....	50
5.2 DISEÑO DE LA MUESTRA.....	50
5.2.1 Estratificación.....	50
5.2.2 Muestreo.....	52

6	APLICACIÓN DE <i>LOAD RESEARCH</i>: ANÁLISIS DE ESTACIONALIDAD	55
6.1	SELECCIÓN DE DATOS OBJETIVO	55
6.2	PREPROCESO DE DATOS	56
6.3	TRANSFORMACIÓN DE DATOS	57
6.4	MINERÍA DE DATOS	58
6.4.1	<i>Clientes agrupados en clusters</i>	59
6.4.2	<i>Clientes en forma agregada</i>	65
6.5	DISEÑO DE UN MODELO MATEMÁTICO PARA SIMULAR CURVAS	68
7	ANÁLISIS DE SUPUESTOS DEL COBRO POR ENERGÍA ADICIONAL DE INVIERNO	70
7.1	ESTRUCTURA DE LA TARIFA BT-1A	70
7.2	SUPUESTOS DE LA TARIFA BT-1A	74
7.3	ANÁLISIS DE DATOS	74
7.4	ANÁLISIS CRUZADO ENTRE CLUSTERS Y ESTRATOS	81
8	CONCLUSIONES Y DESAFÍOS FUTUROS	84
9	BIBLIOGRAFÍA	87
	ANEXOS	89

Indice de Figuras

Figura 2.1: Modelo pool chileno	15
Figura 3.1: Diagrama de flujo <i>Load Research</i>	20
Figura 3.2: Etapas de un proceso <i>KDD</i>	28
Figura 3.3: Técnicas de análisis	29
Figura 3.4: Esquema de la estructura de una red <i>SOM</i>	33
Figura 3.5: Índice DBI	35
Figura 3.6: Índice MIA	36
Figura 3.7: Índice KC	36
Figura 3.8: Diagrama completo de una caracterización de la carga	37
Figura 4.1: Aporte a la demanda por estrato en el estado de Kaohsiung	40
Figura 4.2: Curvas típicas de las SED	44
Figura 4.3: Curva de carga agregada de la tarifa BT5B	46
Figura 4.4: Curva de carga resultante para los clientes BT-1	48
Figura 5.1: Variación del intervalo de confianza en función del tamaño de muestra	53
Figura 6.1: Curvas agregadas de demanda para el mes de julio	57
Figura 6.2: Demanda agregada de la muestra mensual	59
Figura 6.3: Índice Davies – Bouldin	60
Figura 6.4: Promedio Mensual de Índice DBI	61
Figura 6.5: Perfiles típicos mensuales	62
Figura 6.6: Movimiento estacional de cada cluster	63
Figura 6.7: Estimación de la curva de demanda agregada de julio	64
Figura 6.8: Error absoluto de la estimación	65
Figura 6.9: Demanda agregada mensual del estrato 1	66
Figura 6.10: Demanda agregada mensual del estrato 2	66
Figura 6.11: Demanda agregada mensual del estrato 3	67
Figura 6.12: Curvas reales y simuladas tipo comercial	68
Figura 7.1: Esquema planteado por Bernstein	74
Figura 7.2: Curvas de demanda mensuales del estrato 1	76
Figura 7.3: Estacionalidad estrato 1	76
Figura 7.4: Variación de la máxima demanda entre invierno y verano estrato 1	77
Figura 7.5: Curvas de demanda mensuales del estrato 2	77
Figura 7.6: Estacionalidad estrato 2	78
Figura 7.7: Variación de la máxima demanda entre invierno y verano estrato 2	78
Figura 7.8: Curvas de demanda mensuales del estrato 3	79
Figura 7.9: Estacionalidad estrato 3	79
Figura 7.10: Variación de la máxima demanda entre invierno y verano estrato 3	80
Figura 7.11: Incidencia de cada estrato sobre la curva agregada	80
Figura 7.12: Caracterización por perfiles típicos del Estrato 1	81
Figura 7.13: Caracterización por perfiles típicos del Estrato 2	81
Figura 7.14: Caracterización por perfiles típicos del Estrato 3	82

Indice de Tablas

Tabla 3.1: Resumen de fortalezas y debilidades de cada técnica de agrupamiento	34
Tabla 4.1: Muestreo propuesto para el estado de Kaohsiung	39
Tabla 4.2: Muestreo en estrato BT	43
Tabla 4.3: Muestreo en estrato MT	43
Tabla 4.4: Muestreo en estrato BT5B	43
Tabla 4.5: Distribución de cada SED en curvas típicas	44
Tabla 4.6: Resumen de resultados obtenidos	46
Tabla 4.7: Muestreo realizado	47
Tabla 5.1: Categorías de clientes	51
Tabla 5.2: Intervalos de consumo promedio de energía de clientes	51
Tabla 5.3: Estratos de consumo	51
Tabla 5.4: Información estadística de los estratos	52
Tabla 5.5: Ponderación de cada estrato en la muestra	53
Tabla 5.6: tamaño de la muestra para distintos valores de k	54
Tabla 5.7: Estratificación para distintos tamaños de muestra	54
Tabla 6.1: Formato de la base de mediciones	56
Tabla 6.2: Estratificación para análisis de curvas agregadas	65
Tabla 7.1: Total de Energía Consumida	75
Tabla 7.2: Estratos de consumo	75
Tabla 7.3: Porcentaje de participación de cada cluster en cada estrato	82

1 Introducción

En este capítulo se presentan las bases que motivan este trabajo titulado “Análisis de la estacionalidad de la curva de demanda para clientes BT-1”, tema que nace de la necesidad de disponer de mayor información de un segmento de clientes a los cuales, para efectos de facturación, sólo se les mide la energía. Como los temas que se vinculan a un estudio de este tipo son demasiado amplios, este trabajo tiene un planteamiento de investigación general sobre las bases de las campañas de medición y una aplicación específica que corresponde al tema de estacionalidad planteado en el título.

1.1 Motivación

Las compañías eléctricas de distribución requieren información sobre la curva de carga de sus clientes para un sinnúmero de propósitos, ya sea de carácter técnico, económico o comercial. Actualmente los grandes clientes, principalmente industriales y comerciales, cuentan con tarifas que utilizan un medidor horario de demanda. En el caso de los clientes de menor consumo, sólo algunos de ellos poseen medidores horarios. La gran mayoría de éstos, sólo posee un medidor de energía y limitadores de potencia, razón por la cual el comportamiento de estos clientes a lo largo del día y del año no es posible obtenerlo a partir de información comercial. Los clientes con tarifa residencial (BT-1) representan el caso más emblemático, ya que en conjunto se acercan al 90% de la totalidad de los clientes, mientras que su consumo energético agregado arroja cifras similares a un 30% del total.

En el marco de la caracterización de la curva de demanda para los clientes sin medidor horario de demanda surgen una serie de dificultades. La principal de ellas es la limitación económica, ya que instalar un medidor horario para cada cliente involucraría un monto de inversión muy superior a los beneficios que otorga el conocer el comportamiento de consumo de estos consumidores. También existen dificultades técnicas y logísticas, que se relacionan con el manejo de grandes volúmenes de información, ya que las medidas deben registrarse en intervalos pequeños de tiempo y luego ser almacenada para su posterior análisis. Es por ello, que las campañas de medición destinadas a estos clientes y las técnicas estadísticas que las certifican toman una gran relevancia.

1.2 Objetivos y Alcances

El objetivo de esta memoria de título es presentar una metodología que permita caracterizar la curva de demanda de clientes BT-1. Para estos fines se realizará un estudio de las principales técnicas estadísticas que permiten efectuar una campaña de medición representativa para un universo de clientes ya definido. También se detallan las técnicas de agrupamiento o *clustering* comúnmente utilizadas en investigaciones de esta índole. Luego, como aplicación se realizará un análisis de la estacionalidad de la curva de demanda de los clientes que poseen tarifa

BT-1 de la empresa Chilectra S.A., utilizando una base de datos de 1.050 mediciones obtenidos a través de una campaña de medición realizada en 1999.

Los objetivos específicos son:

- Presentar un resumen de las distintas aplicaciones que se pueden realizar con un estudio de la curva de carga de los clientes (*Load Research*) en un sistema de distribución.
- Adquirir una visión general sobre el estado del arte en campañas de medición realizadas por distintas empresas u organismos, que se hayan sido planteadas con fines tarifarios.
- Proponer una metodología para realizar una campaña de medición, que incluya técnicas de muestreo estadístico y técnicas de análisis de minería de datos.
- Realizar un análisis de estacionalidad de la curva de demanda para clientes BT-1, utilizando una base de mediciones de la empresa Chilectra S.A.
- Revisar algunos supuestos del cobro por energía adicional de invierno de la tarifa BT-1 utilizando los resultados obtenidos en el análisis de estacionalidad.

1.3 Estructura

La memoria consta de ocho capítulos, a través de los cuales se abarcan los objetivos planteados.

En el primer capítulo se presenta la introducción al tema, explicando los objetivos y alcances de la memoria y se define la estructura del trabajo.

En el segundo capítulo se explica, en líneas generales, el funcionamiento del sector eléctrico chileno, incluyendo los agentes participantes y los mercados existentes.

El tercer capítulo define las actividades involucradas en un proceso de caracterización de la curva de carga, mostrando un esquema basado en técnicas de muestreo, minería de datos y técnicas de agrupamiento. Además se resumen algunas de las aplicaciones que se pueden desarrollar.

El cuarto capítulo recoge la experiencia previa en campañas de medición realizadas por distintas empresas u organismos, que se hayan sido planteadas con fines tarifarios. Se presentan algunos casos internacionales y un estudio previo realizado por la empresa Chilectra S.A.

En el quinto capítulo se propone un muestreo para realizar una nueva campaña de medición en la empresa Chilectra S.A.

En el sexto capítulo se realiza un estudio de estacionalidad de la curva de demanda para clientes BT-1, como aplicación de *Load Research*. Se analiza la estacionalidad de perfiles típicos de consumo y de la máxima demanda a nivel agregado.

El capítulo siete corresponde a un análisis de algunos supuestos que posee el cobro por energía adicional de invierno de la tarifa BT-1a, utilizando resultados obtenidos en el capítulo anterior. Específicamente se analiza el comportamiento de la potencia de invierno de los clientes por estratos de consumo.

Finalmente, en el octavo capítulo se presentan las conclusiones y los desafíos futuros que deja este trabajo realizado.

2 Características del Sector Eléctrico chileno

El sector eléctrico chileno distingue tres segmentos independientes que definen los eslabones de la cadena productiva de electricidad: generación, transmisión y distribución [1].

2.1 Generación

Este segmento está conformado por las centrales generadoras de electricidad. Este segmento se caracteriza, a diferencia de los otros, por ser un mercado competitivo, donde los precios son el reflejo de los costos marginales de producción. Dentro de este segmento se distingue cuatro mercados claros:

- Mercado de clientes libres¹: La empresa generadora establece un contrato de abastecimiento bilateral con el cliente, en que el precio es libremente acordado entre las partes.
- Mercado de empresas distribuidoras: La empresa generadora establece un contrato de abastecimiento con empresas distribuidoras a precio de nudo de energía y de potencia.
- Mercado spot: En este mercado se transan los excedentes de generación en relación a los contratos de abastecimiento que posea cada empresa generadora, valorizados a costo marginal.
- Mercado de empresas generadoras: Las empresas generadoras pueden establecer contratos de suministro a un precio de comun acuerdo entre las partes.

2.2 Transmisión

El segmento de transmisión o transporte de electricidad está conformado por líneas y subestaciones que forman parte de un sistema eléctrico, en un nivel de tensión nominal superior a 23 kV. En este sector se distingue tres subsectores:

- Sistema de transmisión troncal: Conjunto de líneas y subestaciones eléctricas, económicamente eficientes, necesarias para posibilitar el abastecimiento de la totalidad de la demanda de sistema eléctrico respectivo.
- Sistemas de subtransmisión: Conjunto de líneas y subestaciones eléctricas que están dispuestas para el abastecimiento exclusivo de grupos de consumidores finales libres o regulados que se encuentren en zonas de concesión de empresas distribuidoras.
- Sistemas de transmisión adicionales: Instalaciones de transmisión destinadas al abastecimiento de clientes libres o inyecciones de centrales al sistema eléctrico.

¹ Clientes con una demanda mayor a 2.000 kW. Clientes con demanda entre 500 y 2.000 kW pueden optar a ser libres o regulados.

Debido a las fuertes economías de escala que se presentan en este segmento, se caracteriza por funcionar como un monopolio natural. Esta realidad hace necesaria su regulación. Además, este segmento debe garantizar el libre acceso, bajo condiciones técnicas y económicas, no discriminatorio a los usuarios.

2.3 Distribución

Los sistemas de distribución están constituidos por las líneas, subestaciones y equipos que entregan la electricidad a los consumidores finales en un nivel de tensión menor o igual a 23 kV, que se encuentran dentro de una zona geográfica explícitamente delimitada llamada zona de concesión. Las empresas de distribución operan bajo un régimen de empresas de servicio público, por lo que sus tarifas se encuentran reguladas y su servicio se encuentra sujeto a restricciones de calidad de suministro. Dentro de este segmento se distinguen los siguientes mercados:

- Mercado de clientes libres: Del mismo modo que las empresas generadoras, las distribuidoras pueden establecer contratos de abastecimientos a precio libremente acordado entre las partes.
- Mercado regulado: Este mercado abastece a los clientes regulados² con tarifas que son fijadas por la autoridad.

2.3.1 Tarifas de distribución

Para el mercado regulado se definen tarifas en alta y baja tensión. Los clientes pueden elegir libremente su opción tarifaria [2].

2.3.1.1 Tarifas en baja tensión

Son clientes en baja tensión aquellos que están conectados con su empalme a líneas cuyo voltaje es igual o inferior a 400 V. Las opciones tarifarias son las siguientes:

- Tarifa BT-1: Opción tarifaria simple en baja tensión. Para clientes con medidor simple de energía. Posee un límite de potencia conectada de 10 kW. Esta tarifa posee dos modalidades.
 - BT-1a: Aplicable a clientes abastecidos por empresas cuya demanda máxima se produce en los meses en que se ha definido horas de punta.
 - BT-1b: Aplicable a clientes abastecidos por empresas cuya demanda máxima se produce en los meses en que no se ha definido horas de punta.
- Tarifa BT-2: Opción de tarifa con potencia contratada. Para clientes con medidor simple de energía y potencia contratada. La potencia se puede contratar libremente y posee una vigencia de 12 meses.

² Clientes con una demanda menor a 500 kW. Clientes con demanda entre 500 y 2.000 kW pueden optar a ser libres o regulados.

- Tarifa BT-3: Opción de tarifa con demanda máxima leída. Para clientes con medidor simple de energía y demanda máxima leída. Se entenderá por demanda máxima leída al mes, al más alto valor de las demandas integradas en períodos sucesivos de 15 minutos.
- Tarifa BT-4: Opción de tarifa horaria. Para clientes con medidor simple de energía y demanda máxima contratada o leída, y demanda máxima contratada o leída en horas de punta del sistema eléctrico. En esta opción existen las siguientes opciones de medición:
 - i. BT-4.1: Medición de la energía mensual total consumida, y contratación de la demanda máxima de potencia en horas de punta y de la demanda máxima de potencia.
 - ii. BT-4.2: Medición de la energía mensual total consumida y de la demanda máxima de potencia en horas punta, y contratación de la demanda máxima de potencia.
 - iii. BT-4.3: Medición de la energía mensual total consumida, de la demanda máxima de potencia en horas de punta y de la demanda máxima de potencia suministrada.

2.3.1.2 Tarifas en alta tensión

Son clientes en alta tensión aquellos que están conectados con su empalme a líneas cuyo voltaje es superior a 400 V. Las opciones tarifarias son las siguientes:

- Tarifa AT-2: Opción de tarifa con potencia contratada. Para clientes con medidor simple de energía y potencia contratada. La potencia se puede contratar libremente y posee una vigencia de 12 meses.
- Tarifa AT-3: Opción de tarifa con demanda máxima leída. Para clientes con medidor simple de energía y demanda máxima leída. Se entenderá por demanda máxima leída al mes, al más alto valor de las demandas integradas en períodos sucesivos de 15 minutos.
- Tarifa AT-4: Opción de tarifa horaria. Para clientes con medidor simple de energía y demanda máxima contratada o leída, y demanda máxima contratada o leída en horas de punta del sistema eléctrico. En esta opción existen las siguientes opciones de medición:
 - i. AT-4.1: Medición de la energía mensual total consumida, y contratación de la demanda máxima de potencia en horas de punta y de la demanda máxima de potencia.

- ii. AT-4.2: Medición de la energía mensual total consumida y de la demanda máxima de potencia en horas punta, y contratación de la demanda máxima de potencia.
- iii. AT-4.3: Medición de la energía mensual total consumida, de la demanda máxima de potencia en horas de punta y de la demanda máxima de potencia suministrada.

2.4 Modelo Eléctrico Chileno

En resumen, los tres segmentos anteriormente señalados articulan un mercado eléctrico, en el cual existen dos maneras básicas para realizar contratos comerciales entre compradores y vendedores: trato directo (un comprador y un vendedor realizan contratos bilaterales físicos) o mediante un intermediario y/o *pool* (los oferentes venden con la coordinación centralizada de un intermediario). El modelo *pool* chileno tiene un organismo central de despacho denominado CDEC³. El CDEC coordina los despachos de energía realizando una programación por orden de mérito de las centrales que están conectadas a cada sistema interconectado, de modo de abastecer la demanda tanto de los consumidores con precios regulados como de los consumidores libres. En la figura 2.1 se muestra la relación entre los agentes participantes en el sector [3].

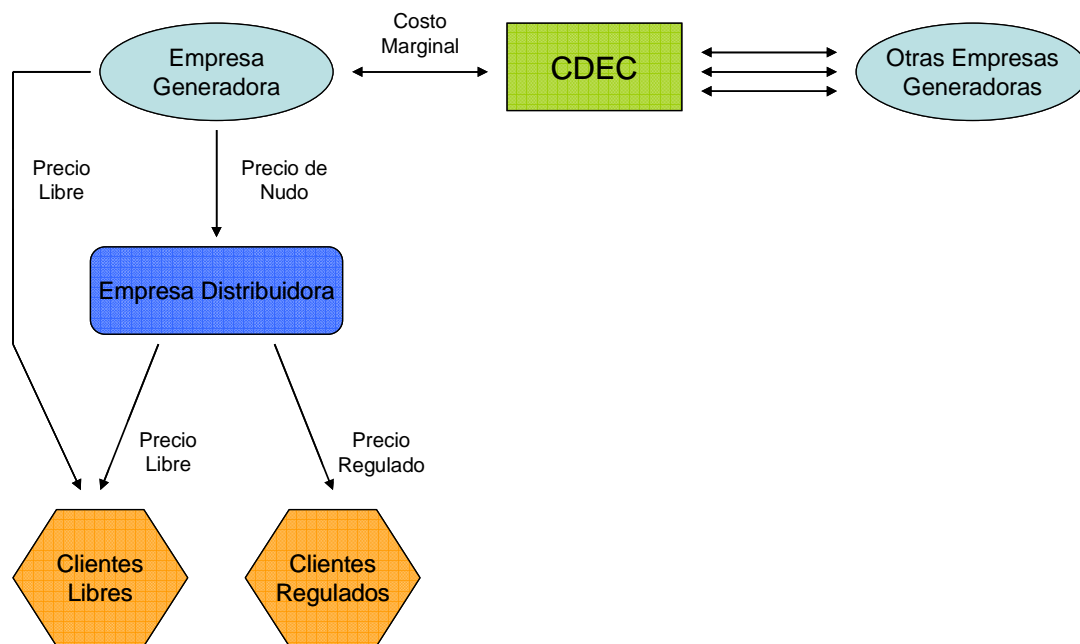


Figura 2.1: Modelo pool chileno.

Como se muestra en la figura 2.1 existen diversas relaciones y precios que actúan en el sector, dependiendo de quienes sean los agentes que participan en la transacción.

³ Centro de Despacho Económico de Carga

En este esquema se reconocen tres mercados distintos:

2.4.1 Mercado Spot

Al mercado Spot pertenecen todas las empresas generadoras que operan en sincronismo con el sistema. El CDEC es el organismo encargado de coordinar el despacho físico de energía. Las empresas generadoras tienen derecho a vender la energía que evacúan al sistema al costo marginal del sistema, así como sus excedentes de potencia a precio de nudo de la potencia a otras empresas generadoras, para que estas últimas puedan cumplir con sus contratos de abastecimiento.

2.4.2 Mercado de clientes libres

En este mercado las empresas generadoras y distribuidoras actúan como comercializadoras, pudiendo establecer contratos de abastecimiento a precios libremente acordados entre ellas y los clientes libres.

2.4.3 Mercado regulado

El mercado de distribución lo constituyen los clientes regulados y las empresas de distribución. Su operación está regulada y restringida a las áreas de concesión con tarifas fijadas por la autoridad cada cuatro años, en base a los costos medios de una empresa ideal o modelo. El precio de venta de electricidad a clientes regulados está estipulado como el precio de nudo de la barra de retiro respectiva más el VAD⁴. Los precios de nudo corresponden a los precios a nivel de generación-transporte y son definidos para todas las subestaciones desde las cuales se efectúa el suministro. Poseen una componente del precio de la energía y otra por el precio de la potencia de punta.

Bajo el esquema de este mercado, las empresas distribuidoras deben abastecer la demanda de los clientes regulados y deben permitir el abastecimiento no discriminatorio de la demanda de los clientes libres con contratos ubicados dentro de la zona de concesión de distribución. La legislación vigente obliga a las empresas distribuidoras a contar con contratos con empresas generadoras que les asegure el abastecimiento total de la demanda de los consumidores regulados, en un horizonte de al menos tres años.

2.5 Calidad de Servicio Eléctrico

Bajo el esquema planteado en los párrafos anteriores, a los consumidores finales se les debe asegurar un producto de calidad. El consumidor percibe, por lo general, el suministro de electricidad como un elemento compuesto solamente de potencia y energía.

⁴ Valor Agregado de Distribución

La calidad de suministro se define como el conjunto de parámetros físicos y técnicos que debe cumplir el producto final que reciben los consumidores. Sin embargo la calidad de servicio es un concepto más amplio que debe ser exigible por los clientes finales, que corresponde al conjunto de propiedades y estándares normales que, son inherentes a la actividad de distribución de electricidad concesionada, y constituyen las condiciones bajo las cuales debe desarrollarse esta actividad [4].

- Las normas y condiciones establecidas en los decretos de concesión.
- La seguridad de las instalaciones y de su operación, y el mantenimiento de las mismas.
- La satisfacción oportuna de las solicitudes de servicio.
- La correcta medición y facturación de los servicios prestados, y el oportuno envío a los usuarios y clientes.
- El cumplimiento de los plazos de reposición de suministro.
- La oportuna atención y corrección de situaciones de emergencia, interrupciones de suministro, accidentes y otros imprevistos.
- El cumplimiento de los plazos de reposición de suministro
- La utilización de adecuados sistemas de atención e información a los usuarios y clientes
- La continuidad del servicio.
- Los estándares de calidad del suministro.

3 Caracterización de la Curva de Carga

La demanda de un sistema eléctrico varía con el tiempo y el espacio y la generación debe responder a la demanda en cada instante. Es por este motivo que los agentes del sistema, descritos en el capítulo anterior, requieren conocer cual es el comportamiento temporal y espacial de la demanda, de forma de poder asegurar y optimizar su funcionamiento, tanto a nivel técnico, como económico. La información más requerida o de mayor utilidad es la curva de demanda del sistema.

3.1 Fundamentos y Aplicaciones

Por definición, *Load Research* es una actividad que abarca la medición y el estudio de las características de la carga eléctrica, entregando como resultado una proyección y un comportamiento general de forma minuciosa y confiable para la empresa de servicio público. Otra definición, menos estricta, pero más detallada corresponde a una actividad que permite a las empresas de servicios públicos estudiar la forma en que sus consumidores utilizan la energía, ya sea en forma agregada, por grupos o en forma individual [5].

Variadas disciplinas participan en el desarrollo de esta actividad, como ingeniería, estadística, computación y marketing, por mencionar algunas, siendo el objetivo principal entregar valor financiero y estratégico a la compañía que lo realiza. Dentro de las aplicaciones más comunes destacan: diseño de tarifas, planificación del sistema, operación y mantenimiento del sistema, marketing y *Load Management* (manejo de carga) [5, 6, 7, 8, 9]. A continuación se presenta una breve descripción de las aplicaciones antes mencionadas.

3.1.1 Diseño de tarifas

En el marco de apertura a la comercialización de energía que incluya a los clientes que posean medidores de energía simple, es donde las campañas de medición poseen las mayores aplicaciones, ya que las tarifas se diseñan, principalmente, para perfiles típicos de consumo. En el caso de mercados con tarifas reguladas las campañas de medición se realizan para calcular parámetros tarifarios, como son las horas de uso o los factores de coincidencia. En el capítulo siguiente se detallará algunos ejemplos.

3.1.2 Planificación del sistema

Desde el punto de vista técnico, la planificación de los sistemas es una de las líneas más importantes de abarcar dentro de las empresas distribuidoras. Para estos efectos se utilizan datos como el crecimiento de la demanda, las nuevas urbanizaciones de las comunas que pertenecen a las áreas de concesión y los requerimientos de calidad de servicio que exija la normativa vigente.

Es por ello que un estudio de mediciones mejora la disposición de información para efectos del desarrollo de diversos análisis, tales como:

- Estimación de demanda para el cálculo apropiado de nuevas redes y equipos.
- Estudio de la congestión de redes existentes.
- Cálculo de respaldo de alimentadores ante contingencias a nivel de distribución.

3.1.3 Operación y mantenimiento del sistema

Como operación y mantenimiento puede entenderse la actividad de explotación técnica del negocio. Bajo ese esquema las mediciones resultan imprescindibles para el monitoreo del funcionamiento de las redes, de tal forma que se asegure la calidad de servicio, principalmente los estándares de continuidad y calidad de suministro.

3.1.4 Marketing

Dentro de las aplicaciones que se mencionan, ésta destaca por ser la menos técnica. A partir del conocimiento del comportamiento eléctrico de los clientes o de un grupo de ellos se pueden ofrecer productos diseñados exclusivamente para ese tipo de comportamiento. Un ejemplo de ello podría ser ofrecer un sistema de calefacción eléctrica que se ajuste a la demanda del cliente.

3.1.5 Load management

Esta aplicación está muy relacionada con la operación del sistema. Cuando la demanda de los clientes es muy grande, o cuando el aumento de demanda es muy acelerado se utiliza el sistema de *Load Management*, que consiste en desconectar a ciertos clientes, previo acuerdo mediante un contrato, por el tiempo que sea necesario. Es por esta razón que el conocimiento del comportamiento eléctrico de los clientes se hace necesario, para poder ofrecer este tipo de contratos.

Realizar *Load Research* requiere de una planificación determinada. El primer paso debe ser definir los objetivos que se persiguen, ya que éstos condicionan los pasos siguientes. Luego hay que seleccionar los datos que se va a utilizar. Si estos datos no están disponibles, habrá que realizar una campaña de medición para obtenerlos. Un ejemplo de esto ocurre con los clientes residenciales, que sólo poseen un medidor de energía simple, por lo que hacer un estudio de la curva de carga requerirá la instalación de medidores especiales que sean capaces de registrar la curva de demanda. Posteriormente se debe realizar una minería de datos, que entregará el conocimiento o la conclusión necesaria para desarrollar la aplicación. En la figura 3.1 se muestra un diagrama de flujo básico de las actividades fundamentales.

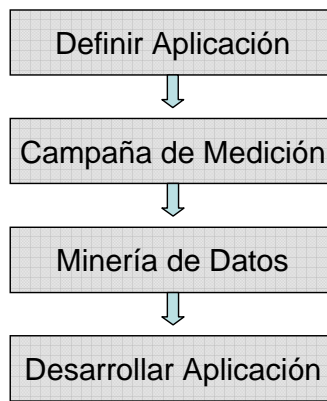


Figura 3.1: Diagrama de flujo *Load Research*.

En los párrafos siguientes se revisarán las actividades involucradas en el proceso de realizar un estudio de caracterización de la curva de demanda.

3.2 Campañas de Medición

Una campaña de medición es una actividad que se realiza para establecer las distintas modalidades de consumo con objeto de caracterizar la carga, sobre una muestra del universo de clientes de la empresa distribuidora. Las etapas que se deben considerar en una campaña de medición son:

- Diseño de la campaña: Esta etapa consiste en definir el muestreo que se utilizará para asegurar la representatividad de la caracterización, el tiempo de duración de la campaña y el intervalo de registro de cada medida.
- Implementación de la campaña: Esta etapa corresponde a la instalación de los medidores y a las pruebas necesarias para realizar la campaña.
- Recopilación de datos: Corresponde a la extracción y almacenaje de los datos que entregará cada medidor instalado.

3.2.1 Técnicas de muestreo

Previamente se mencionó que antes de realizar un estudio de caracterización de la carga es necesario seleccionar qué datos se utilizará en el estudio. Estos datos pueden pertenecer a una base histórica de datos o también puede darse el caso de que los datos requeridos no existan, por lo que será necesario obtener esa información mediante una campaña de medición.

El primer paso para realizar un muestreo es definir el universo objetivo, denominado población. Una muestra será un subconjunto de la población, que debe ser estadísticamente representativo, dado un cierto grado de confiabilidad. Si bien lo ideal sería poder realizar un

estudio sobre la población completa, esto no es posible fundamentalmente por razones económicas, técnicas y logísticas.

Existen varias formas de elegir una muestra de una población, sin embargo todas deben cumplir con la condición de que sean representativas, para que no se presenten errores imprevistos llamados sesgos. Es por ello que el principal objetivo de la teoría del muestreo sea determinar el tamaño adecuado que debe tener la muestra para asegurar la representatividad y validez del estudio que se propone.

Básicamente existen dos tipos de muestreos [10]:

- Muestreo no probabilístico: La muestra no se elige al azar, sino mediante determinados criterios subjetivos.
- Muestreo probabilístico: La muestra se elige al azar. En este caso se pueden distinguir varios tipos:
 - Muestreo aleatorio simple: Aquel en que cada individuo, o cliente en este caso, de la población tiene las mismas posibilidades de ser elegido en la muestra.
 - Muestreo sistemático: En el que se elige un cliente al azar, y a partir de él, a intervalos constantes de una numeración continua, se eligen los demás hasta completar la muestra.
 - Muestreo estratificado: Este muestreo divide a la población en estratos, que corresponden a subconjuntos de la población que presentan características homogéneas con respecto a una variable conocida. Luego se escoge aleatoriamente un número de individuos de cada estrato de forma proporcional al número de componentes del estrato o con algún criterio que además considere la varianza.
 - Muestreo por conglomerados: Si no se dispone de alguna información de los clientes de la población, o de los posibles estratos, se puede realizar este tipo de muestreo. Se divide a la población en conglomerados, que corresponden a subconjuntos de la población con características tan heterogéneas como la población. Luego se elige una cantidad de conglomerados aleatoriamente y posteriormente se analizarán sus individuos o se toma una muestra aleatoria simple del conglomerado seleccionado.

La elección del tipo de muestreo que se utilizará dependerá del conocimiento previo que se tenga de la población, ya que un muestro estratificado puede entregar mucha mas representatividad que un muestreo aleatorio simple si es que los estratos están bien conformados.

En la literatura técnica consultada [6, 7, 8, 11] principalmente se propone el uso del muestreo estratificado, ya que existe suficiente información que puede ser utilizada para definir estratos, ya sea el tipo de contrato o el consumo de energía promedio en kWh. Sin embargo, se

propone varios métodos distintos de muestreo estratificado, que dependen, entre otras cosas, de la información con la que se cuente inicialmente. A continuación se revisarán algunas propuestas.

3.2.1.1 Propuesta 1

Esta propuesta separa como estrato a las distintas opciones tarifarias existentes, ya que supone que dentro de esas clases existe una homogeneidad entre los clientes. Como dato conocido se considera el consumo facturado en kWh de cada cliente, lo que permite calcular el promedio de consumo anual y la desviación estándar para cada estrato. Luego se determina la muestra óptima utilizando un intervalo de confianza, típicamente de un 95% de significación para distribuciones normales. La expresión que determina el número óptimo de clientes que se deben muestrear en cada estrato, es la que se muestra en la ecuación (3.1) [6].

$$n_i = (1,96)^2 \cdot \frac{\sigma_i^2}{\mu_i^2} \quad (3.1)$$

Donde

σ_i : Desviación estándar del consumo de energía en el estrato i

μ_i : Promedio del consumo de energía en el estrato i

Luego, el tamaño total de la muestra corresponderá a la suma de cada muestreo por estrato, como se muestra en la ecuación (3.2).

$$n = \sum_{i=1}^N n_i \quad (3.2)$$

3.2.1.2 Propuesta 2

Esta propuesta, a diferencia de la anterior, calcula en primer lugar el número total de individuos que se incluirán en la muestra utilizando una ecuación clásica de teorías de muestreo [10]. Se utiliza como dato el consumo facturado por cada cliente en kWh.

Entonces, utilizando una ecuación del error de muestreo se determina el número de clientes que serán muestreados. En la ecuación (3.3) se muestra la expresión [7].

$$d = Z \cdot \sqrt{\frac{S}{n} \cdot \frac{N-n}{N}} \quad (3.3)$$

Donde

d : Error de muestreo

Z : Nivel de confianza

S : Varianza del consumo de energía

N : Total de clientes en la población

n : Total de clientes en la muestra

Luego, para determinar la proporción de cada estrato sobre la muestra total se utiliza la varianza del consumo de energía de cada estrato, así la cantidad de clientes que se debe seleccionar en cada estrato esta dado por la relación que muestra la ecuación (3.4).

$$n_i = \frac{N_i \cdot S_i}{\sum_{i=1}^N N_i \cdot S_i} \cdot n \quad (3.4)$$

Donde

n_i : Clientes a ser muestreados en el estrato i

S_i : Varianza del consumo en el estrato i

N_i : Total de clientes en el estrato i

N : Total de clientes en la población

n : Total de clientes en la muestra

3.2.1.3 Propuesta 3

Esta propuesta es la más completa de las revisadas en la bibliografía, ya que considera los costos de realizar una medición dentro de los parámetros que definen el número óptimo de clientes que deben ser muestreados. Es por esta razón que se analizará con más detalle que las anteriores.

Para implementar esta propuesta es necesario definir los siguientes parámetros y posteriormente los estadísticos que se muestran en el conjunto de expresiones (3.5) [9]:

H : Número de estratos

N : Total de clientes en la población

Luego, para un estrato genérico h se tiene:

N_k : Total de clientes en el estrato h

n_k : Total de muestras en el estrato h

y_{hi} : Valor de la variable y de la muestra i en el estrato h

$$\begin{aligned}
u_h &= \frac{1}{N_h} \cdot \sum_{i=1}^{N_h} y_{hi} \\
\sigma_h^2 &= \frac{1}{N_h - 1} \cdot \sum_{i=1}^{N_h} (y_{hi} - \mu_h)^2 \\
\hat{\mu}_h &= \frac{1}{n_h} \cdot \sum_{i=1}^{n_h} y_{hi} \\
\hat{\sigma}_h^2 &= \frac{1}{n_h - 1} \cdot \sum_{i=1}^{n_h} (y_{hi} - \hat{\mu}_h)^2
\end{aligned} \tag{3.5}$$

Donde,

u_h : Promedio real en el estrato h

σ_h^2 : Varianza real en el estrato h

$\hat{\mu}_k$: Promedio estimado en el estrato h

$\hat{\sigma}_h^2$: Varianza estimada en el estrato h

Una vez calculados los estadísticos anteriores, es necesario proyectar esos valores a la población utilizando los valores estimados, ya que a partir de esa proyección se encontrará una ecuación para determinar el número óptimo de individuos que deben pertenecer a la muestra. Las expresiones son las que se muestran en el grupo de ecuaciones (3.6).

$$\begin{aligned}
N &= \sum_{h=1}^H N_h \\
n &= \sum_{h=1}^H n_h
\end{aligned} \tag{3.6}$$

Luego, el promedio real de la población y su estimado se detallan en las expresiones (3.7).

$$\begin{aligned}
\mu &= \frac{1}{N} \cdot \sum_{h=1}^H N_h \mu_h \\
\hat{\mu} &= \frac{1}{N} \cdot \sum_{h=1}^H N_h \hat{\mu}_h
\end{aligned} \tag{3.7}$$

La varianza asociada a $\hat{\mu}$ es la que se muestra en la ecuación (3.8).

$$\hat{\sigma}^2 = \frac{1}{N} \cdot \sum_{h=1}^H \left[N_h^2 \cdot \left(\frac{N_h - n_h}{N_h - 1} \right) \cdot \frac{\sigma_h^2}{n_h} \right] \tag{3.8}$$

Luego, sin cometer mucho error, se puede estimar un intervalo de confianza $\mu \approx \hat{\mu} \pm k\hat{\sigma}$, donde k va a representar el grado de confiabilidad.

Finalmente, para seleccionar el número de muestras de cada estrato hay que resolver el problema de optimización planteado en la expresión (3.9).

$$\begin{aligned} & \min_{n_h} \hat{\sigma}^2 \\ & \text{sujeto a} \\ & C = c_0 + \sum_{h=1}^H c_h n_h \end{aligned} \quad (3.9)$$

Donde

C : Presupuesto total

c_0 : Costo fijo de realizar mediciones

c_h : Costo de realizar una medicion en el estrato h

Este problema de optimización se resuelve utilizando un lagrangeano. El resultado que se obtiene es el que se muestra en la expresión (3.10).

$$\begin{aligned} n_h &= n \cdot \frac{N_h \cdot \sigma_h \cdot \sqrt{\frac{N_h}{N_h - 1}}}{\sum_{m=1}^H N_m \cdot \sigma_m \cdot \sqrt{\frac{N_m}{N_m - 1}}} \\ n &= \frac{C - c_0}{c'} \end{aligned} \quad (3.10)$$

con $c' = c_h \forall h$

Sin embargo, la ecuación que determina el número total de clientes que se debe incluir en la muestra no es suficiente, ya que no existe una conclusión en base a estadísticas. Es por ello que hay incluir la ecuación que certifica que la muestra entrega una aproximación con cierta probabilidad de encontrarse dentro de un intervalo de confianza determinado.

Una representación aproximada del intervalo de confianza para el promedio de los consumos de la población se muestra en la ecuación (3.11).

$$\mu \approx \hat{\mu} \pm k\hat{\sigma} = \hat{\mu} \pm d \quad (3.11)$$

Donde d representa el intervalo.

A partir de la ecuación (3.11), es útil definir el intervalo de manera proporcional al promedio de consumo. De esta forma la ecuación de cálculo queda representada como se muestra en la ecuación (3.12).

$$\frac{d}{\mu} = k \cdot \frac{\hat{\sigma}}{\mu} \quad (3.12)$$

Entonces para calcular el número de clientes de la muestra se debe decidir que intervalo de confianza se requiere (por ejemplo 5%) y definir con que probabilidad de acierto se necesita pertenecer a ese intervalo (por ejemplo 95%, $k = 1,96$). Luego, reemplazando la ecuación de n_h (3.10) en la ecuación de $\hat{\sigma}^2$ (3.8), se obtiene una expresión que depende del valor de n que se muestra en la ecuación 3.13.

$$\hat{\sigma}^2 = \frac{\left(\sum_{h=1}^H N_h \cdot \sigma_h \cdot \sqrt{\frac{N_h}{N_h - 1}} \right)^2 - n \cdot \sum_{h=1}^H \frac{N_h^2 \cdot \sigma_h^2}{N_h - 1}}{N^2 \cdot n} \quad (3.13)$$

El resultado que entregue la expresión (3.13) debe chequearse con la restricción presupuestaria, o bien a partir de esta ecuación definir cual será el costo de realizar la campaña de medición.

3.2.2 Implementación de la campaña

La implementación en terreno de la campaña es un proceso técnico que normalmente lo realizará una empresa externa especializada. Se debe considerar principalmente los siguientes puntos:

- Documentar los datos del cliente que será medido, los datos del medidor instalado y la fecha y hora de la instalación.
- Verificar el correcto funcionamiento del medidor instalado.
- Verificar constantemente que el cliente posee servicio y que los medidores siguen funcionando correctamente.

Dentro de este proceso se enmarca también la extracción de los datos obtenidos por el medidor, lo que se debe hacer con una frecuencia que sea permitida por la capacidad de la memoria de masa del medidor.

3.3 Minería de Datos

Una vez obtenida la base de datos de la campaña, o bien teniendo disponible la base existente de mediciones que se quiere estudiar, será necesario realizar una minería de datos para encontrar los patrones comunes que caracterizarán la carga.

La minería de datos es una técnica que se divide en varias etapas. En los párrafos siguientes se detallará las etapas y las técnicas más utilizadas en estudios de *Load Research* en una empresa distribuidora.

3.3.1 Etapas de una minería de datos

La minería de datos es una técnica que busca patrones o reglas comunes dentro de una base de datos. El proceso general, que involucra todas las actividades previas y posteriores, para realizar una minería de datos se denomina *KDD*⁵ y posee cinco pasos fundamentales [9].

- Selección de datos: Se refiere a la elección de los datos que se utilizará desde una base existente que contenga los datos necesarios para resolver el problema planteado.
- Pre-procesamiento de datos: Esta etapa consiste en eliminar los datos que contengan errores o eliminar el ruido que pudiera existir.
- Transformación de datos: Significa convertir datos cualitativos en cuantitativos, o convertir los datos a unidades equivalentes de manera que puedan ser procesados. Un ejemplo podría ser convertir medidas eléctricas que estén en unidades reales a p.u.
- Minería de datos: Corresponde al proceso de identificar patrones o reglas comunes. Existen muchos métodos. Entre ellos los más populares son: Lógica Difusa, Redes Neuronales y Técnicas Estadísticas.
- Interpretación: Es el proceso final y corresponde a interpretar y concluir sobre los resultados obtenidos en el paso anterior.

En la figura 3.2 se muestra un esquema con las etapas y los resultados que se obtienen en cada una de ellas.

⁵ *Knowledge Discovery in Databases*

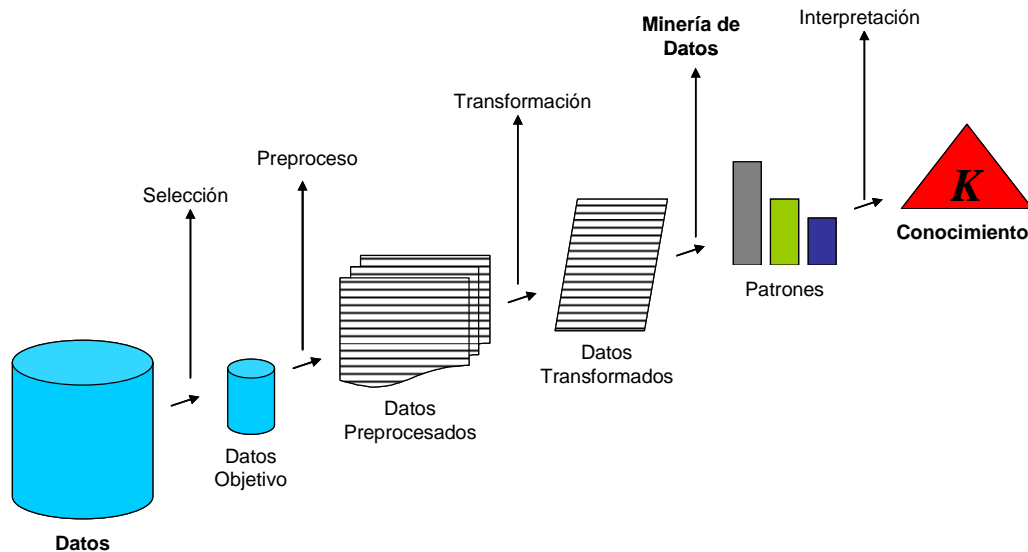


Figura 3.2: Etapas de un proceso *KDD*.

Posteriormente, y a modo de revisión, se debe comprobar que las reglas o patrones encontrados cumplan ciertas condiciones que entreguen un resultado claro al análisis realizado. Estas condiciones son:

- Validez: implica la importancia de las reglas obtenidas, ya que se espera que los resultados obtenidos puedan ser extrapolados a datos que no son conocidos.
- Novedad: significa que los resultados obtenidos no puedan ser obtenidos previamente a priori, sino que resulten de un descubrimiento.
- Utilidad: se refiere a que los resultados encontrados tengan alguna aplicación válida para el usuario.
- Simplicidad: las reglas obtenidas deben ser fáciles de interpretar, de tal manera que se pueda encontrar alguna aplicación con datos desconocidos.

La minería de datos también puede ser utilizada con otros fines como, análisis estadístico, inteligencia artificial o visualización de datos.

3.3.2 Técnicas de agrupamiento

Para realizar una minería de datos es necesario conocer cuales son las técnicas que se utilizan para clasificar o agrupar los datos, que en este caso son los consumos eléctricos, para formar los llamados *cluster* o grupos que poseen características similares y que juntos forman una curva típica llamada perfil de consumo (*Load Profile*).

3.3.2.1 Criterios de clasificación

En de la literatura se distingue los siguientes criterios para realizar una clasificación de consumos eléctricos [4, 12, 13, 14]:

- Estacionalidad: supone la existencia de ciclos anuales en el comportamiento de los consumos, los que se correlacionan en forma importante con las cuatro estaciones del año.
- Sectorial: se basa en que el comportamiento eléctrico de los consumos está determinado por actividades inherentes al sector al que pertenecen. Se distingue entre tipo residencial, comercial o industrial.
- Días característicos: supone la existencia de comportamientos típicos de acuerdo a los días de la semana. Normalmente se distingue entre día laboral y fin de semana o feriado.
- Demográfico: debido a la naturaleza de sus actividades, los consumos urbanos y rurales merecen tratamientos distintos.
- Horario: diferencia los consumos de acuerdo a la distribución horaria de su demanda de electricidad.

3.3.2.2 Algoritmos de clasificación

Existen variados métodos o algoritmos de clasificación para realizar *load profiling*. Principalmente se clasifican en tres grandes grupos: técnicas estadísticas, técnicas de lógica difusa y técnicas de redes neuronales [9, 12, 14, 15, 16, 17, 18, 19, 20]. En la figura 3.3 se muestra algunas de las técnicas revisadas para el desarrollo de esta memoria.

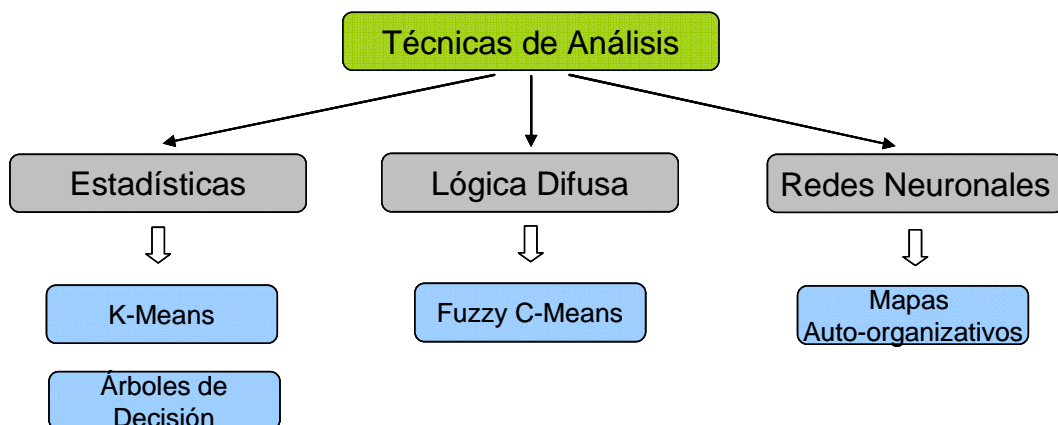


Figura 3.3: Técnicas de análisis.

3.3.2.2.1 Técnicas Estadísticas

Los métodos estadísticos son del tipo determinístico, en decir, cada medición pertenece a un único cluster. Además pueden ser jerárquicos, lo que significa que los consumos se clasifican de manera ascendente o descendente en pasos sucesivos, o pueden estar basadas en una función objetivo, donde la agrupación va a ser resultado de una función de optimización. A continuación se detallará las técnicas revisadas.

K-Means

Este método, también conocido como C-Means, está basado en una función objetivo de minimización de distancia entre mediciones. Consiste en agrupar un set de datos (vectores) $x^{(n)}$ ($n = 1, \dots, N$) en $k = 1, \dots, K$ clusters a través de un proceso iterativo. El primer paso es definir los centros $c^{(k)}$ de los k clusters de forma aleatoria. El segundo paso es calcular la distancia (usualmente Euclideana) $\|x^{(n)} - c^{(k)}\|$ entre cada dato y cada centro de cluster. El tercer paso es agrupar cada dato con el centro del cluster en que la distancia haya sido la mínima y se calcula el nuevo centro del cluster $c^{(k)}$ como el promedio de los datos que pertenecen a ese cluster. Luego se repite el segundo paso hasta que el centro de los clusters se estabilice y su variación sea inferior a un valor ε antes definido. La función objetivo corresponde a la expresión que se muestra en la ecuación (3.14).

$$J(U, C) = \sum_{k=1}^K \sum_{n=1}^N u_{k,n} \cdot \|x^{(n)} - c^{(k)}\|^2 \quad (3.14)$$

Donde

J : Función objetivo

U : Matriz de pertenencia a cada cluster (binaria)

C : Centro de cada cluster

Los coeficientes de la matriz de pertenencia se calculan con la expresión 3.15.

$$u_{k,n} = \begin{cases} 1 & , \text{ si } \|x^{(n)} - c^{(k)}\| = \min(\|x^{(n)} - c^{(j)}\|) \forall j \\ 0 & , \text{ si } \|x^{(n)} - c^{(k)}\| \neq \min(\|x^{(n)} - c^{(j)}\|) \forall j \end{cases} \quad (3.15)$$

Finalmente, los centros de cada cluster se calculan como se muestra en la ecuación 3.16.

$$c^{(k)} = \frac{\sum_{n=1}^N u_{k,n} \cdot x^{(n)}}{\sum_{n=1}^N u_{k,n}} \quad (3.16)$$

Árboles de Decisión

Este método es del tipo jerárquico. El primer paso consiste en agrupar el set de datos (vectores) $x^{(n)}$ ($n=1, \dots, N$) en N *clusters*, en que cada uno de esos clusters será un dato distinto del set y que representará el primer nivel del árbol. Luego se construye una matriz de $N \times N$ que representa la similitud (utilizando algún criterio de distancia) entre cada una de las mediciones. Utilizando esta matriz se construye el árbol de decisión, en que el segundo nivel será de $N-1$ *clusters*, ya que los *clusters* más cercanos se unirán para formar un nuevo *cluster*. De esta forma para llegar a K *clusters* el árbol deberá tener $N+1-K$ niveles. Existen dos criterios para decidir cuáles serán los *clusters* que se fusionaran al pasar al nivel siguiente:

Average distance: con este método la fusión de dos cluster s y t se decidirá de acuerdo al promedio de las distancias de todos los pares de datos de los dos clusters $d_A(s, t)$, según se muestra en la ecuación (3.17).

$$d_A(s, t) = \frac{1}{n^{(s)} \cdot n^{(t)}} \cdot \sum_{i=1}^{n^{(s)}} \sum_{j=1}^{n^{(t)}} d(x_i^{(s)}, x_j^{(t)}) \quad (3.17)$$

Donde

$n^{(s)}, n^{(t)}$: Número de objetos en grupos s y t

Ward: con este método la fusión de dos clusters s y t se decidirá de acuerdo a que el incremento de la suma de distancias al cuadrado de los centros de los cluster al ser unidos $d_W(s, t)$ sea mínima, según lo indica la ecuación (3.18).

$$d_W(s, t) = \frac{n^{(s)} \cdot n^{(t)}}{n^{(s)} + n^{(t)}} d^2(\bar{x}^{(s)}, \bar{x}^{(t)}) \quad (3.18)$$

3.3.2.2.2 Técnicas Difusas

Las técnicas difusas se caracterizan por ser de tipo probabilísticas, es decir, a cada dato se le asigna un grado o probabilidad de pertenecer a un *cluster* determinado. Dentro de esta familia se analizará el método Fuzzy C-Means (FCM).

Fuzzy C-Means

Este método se diferencia de K-Means en la naturaleza de la matriz de pertenencia U , que en este caso deja de ser binaria y se transforma en continua, tomando valores entre 0 y 1 según el grado de pertenencia de cada medición a cada *cluster*. Además, como es una técnica difusa se define el parámetro m llamado *amount of fuzzyness*⁶, que pertenece al rango $(1, \infty)$ y que será el exponente de la matriz de pertenencia U . La función objetivo corresponde a la expresión que se muestra en la ecuación 3.19.

$$J_m(U, C) = \sum_{k=1}^K \sum_{n=1}^N (u_{k,n})^m \cdot \|x^{(n)} - c^{(k)}\|^2 \quad (3.19)$$

Los coeficientes de pertenencia de la matriz U se calculan como se detalla en la ecuación (3.20).

$$u_{k,n} = \frac{1}{\sum_{i=1}^K \left(\frac{d^2(x^{(n)}, c^{(k)})}{d^2(x^{(n)}, c^{(i)})} \right)^{\frac{1}{m-1}}} \quad (3.20)$$

$$\sum_{k=1}^K u_{k,n} = 1, \forall n$$

Finalmente, los centros de cada cluster se calculan mediante la ecuación (3.21).

$$c^{(k)} = \frac{\sum_{n=1}^N u_{k,n} \cdot x^{(n)}}{\sum_{n=1}^N u_{k,n}} \quad (3.21)$$

3.3.2.2.3 Técnicas de Redes Neuronales

Las redes neuronales artificiales tratan de reproducir la forma de actuar del cerebro humano, es decir, un procesador de información altamente complejo, no-lineal y paralelo que es capaz de obtener resultados mucho más rápido que en un proceso binario o digital. Se componen de elementos de procesamiento interconectados (neuronas) que resuelven los problemas, en este caso, de agrupamiento. Existe una diversidad de métodos desarrollados para realizar *clustering*. En este documento se revisará el de Mapas Auto-Organizativos.

⁶ Grado difuso.

Mapas Auto-organizativos⁷

Este método pertenece al tipo de redes neuronales de aprendizaje no supervisado, que proyecta un set de datos $x^{(m)} (m=1, \dots, M)$ de n-dimensiones en un mapa unidimensional. El mapa auto-organizativo está compuesto por una grilla de neuronas n-dimensionales $c^{(i)}$, en que sólo una de ellas se activará ante la entrada de un dato $x^{(m)}$. La función de activación será inversamente proporcional a la función $\|x^{(m)} - c^{(i)}\|$, es decir se activará la neurona más cercana al valor $x^{(m)}$. Luego, el valor de la unidad ganadora se actualizará según lo expresado en la ecuación (3.22).

$$c^{(i)}(t+1) = c^{(i)}(t) + \eta \cdot (x^{(m)} - c^{(i)}(t)) \quad (3.22)$$

Donde

t : Cuenta el número de activaciones de cada neurona $c^{(i)}$

η : Grado de aprendizaje

En la figura 3.4 se muestra esquemáticamente la estructura de un mapa auto-organizativo unidimensional.

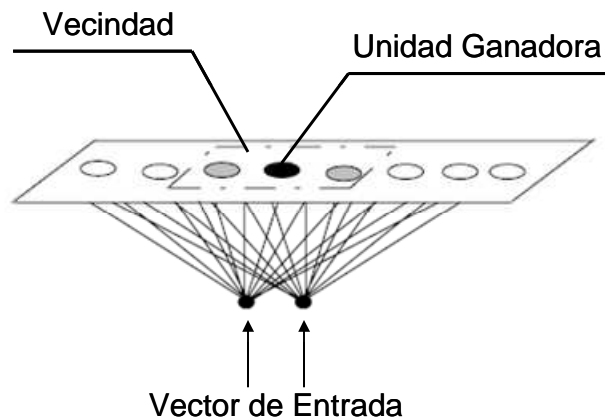


Figura 3.4: Esquema de la estructura de una red SOM.

Cada una de las técnicas presentadas en los párrafos anteriores posee fortalezas y debilidades frente a las demás. A modo informativo, en la tabla 3.1 se describe un resumen con las principales características de cada método.

⁷ Self-Organizing Maps, SOM

Método	Fortalezas	Debilidades
K-Means	Velocidad de procesamiento	Sensibilidad al punto de partida
	Habilidad para clasificar datos de entrada	Matriz de pertenencia binaria
Arboles de Decisión (Average)	Detección de curvas atípicas	Clusters demasiado grandes
	No es sensible al punto de partida	Velocidad de procesamiento
Arboles de Decisión (Ward)	Previene la formación de clusters grandes	Velocidad de procesamiento
	No es sensible al punto de partida	
Fuzzy C-Means	Matriz de pertenencia continua	Velocidad de procesamiento
	Habilidad para clasificar datos de entrada	
SOM	Habilidad para clasificar nuevos datos distintos a los de entrada	Programación compleja
	Velocidad de procesamiento	

Tabla 3.1: Resumen de fortalezas y debilidades de cada técnica de agrupamiento.

La decisión particular de utilizar alguna de las técnicas presentadas recientemente, dependerá de las aplicaciones que se quiera realizar, teniendo en consideración las características expuestas en la tabla 3.1

3.3.2.3 Índices de validez

Todas las técnicas descritas anteriormente tienen el inconveniente de que hay que ingresar el número de *clusters* que se desea tener, por ello un punto importante será poder definir cual es el número óptimo de *clusters* para cada set de datos. Desde el punto de vista de desarrollar alguna aplicación el número de *clusters* debería ser lo más pequeño posible, tal que se encuentren bien definidos, pero por otro lado un número pequeño arrojará una gran dispersión en los resultados y el error cometido será muy grande. Es por esta razón que existen los índices de validez de *clusters*, que determinan el número adecuado haciendo una comparación de los resultados obtenidos para distintos números. La literatura propone una diversidad de métodos para analizar la validez de cada agrupación [14, 16, 18], sin embargo no es un objetivo de esta memoria estudiar detalladamente cada uno de los métodos, por lo que sólo serán nombrados algunos de ellos. Los más utilizados son:

Índice DBI (Davies-Bouldin Index)

Este índice es el más popular dentro de la literatura consultada. Compara la distancia entre los datos al interior de cada cluster con la distancia entre los centros de cada cluster. La mejor elección será cuando el índice sea más pequeño. La ecuación (3.23) muestra el cálculo del índice.

$$DBI = \frac{1}{K} \cdot \sum_{\substack{i=1, \dots, K \\ i \neq j}} \max \left\{ \frac{\sigma_i + \sigma_j}{D_{i,j}} \right\}, \quad K = 2, \dots, N-1$$

$$\sigma_l = \left(\frac{1}{n^{(k)}} \cdot \sum_{v=1}^{n^{(k)}} |x^{(v)} - c^{(k)}|^2 \right)^{\frac{1}{2}}, \quad k = i, j \quad (3.23)$$

Donde

$D_{i,j}$: Distancia entre los centros del cluster i y j

$n^{(k)}$: Datos que pertenecen al cluster l

Un resultado gráfico de este índice para distintos números de cluster se muestra en la figura 3.5.

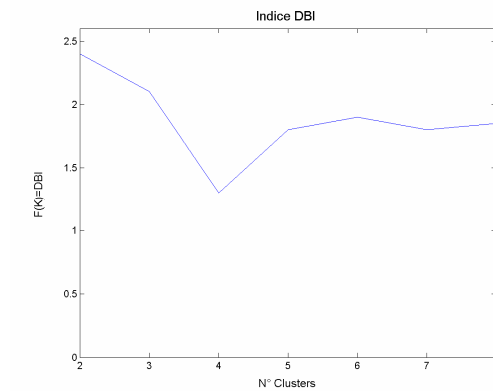


Figura 3.5: Índice DBI.

Índice MIA (Mean Index Adequacy)

Este índice está basado en la distancia Euclídeana que tiene cada medida con respecto al centro del cluster al que pertenece. Calcula el promedio de las distancias (o dispersiones) de cada centro de cluster frente a las medidas que lo integran. Al igual que el método anterior, la mejor elección es la que entrega el índice más pequeño. La ecuación que define este índice es la (3.24) y la figura 3.6 muestra el resultado para distintos números de *clusters*.

$$MIA = \left(\frac{1}{K} \cdot \sum_{k=1}^K \left(\frac{1}{n^{(k)}} \cdot \sum_{j=1}^{n^{(k)}} d^2(c^{(k)}, x^{(j)}) \right) \right) \quad (3.24)$$

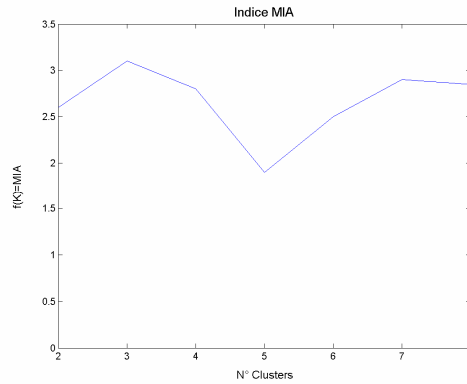


Figura 3.6: Indice MIA

Índice *Knee Citation*

Este índice difiere de los anteriores, ya que es una función decreciente a medida que se aumenta el número de clusters, por lo que su interpretación es más subjetiva. Se define el valor KC mediante la ecuación (3.25), que corresponde a la suma del error de cada medición asignada a cada *cluster* para un número k de *clusters*.

$$KC = \sum_{k=1}^K \sum_{j=1}^{n^{(k)}} \|x^{(j)} - c^{(k)}\| \quad (3.25)$$

La elección del número óptimo de clusters N_0 debe hacerse cuando la diferencia entre $KC(K) - KC(K+1) \approx cte \quad \forall K \geq N_0$.

Un resultado gráfico de este índice para distintos números de cluster se muestra en la figura 3.7.

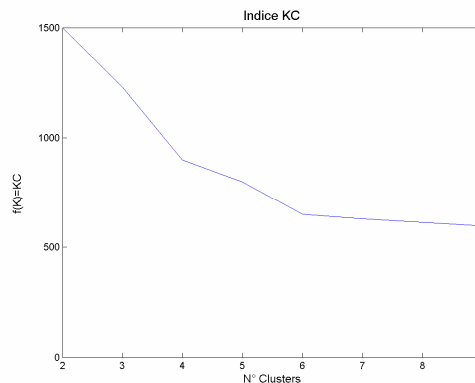


Figura 3.7: Indice KC

Es importante destacar que no siempre los resultados serán los mismos, ya que cada indicador realiza distintos cálculos, pero sirven para tener una certeza en que se está disminuyendo el error cometido en las aproximaciones de consumos por clusters.

3.4 Esquema General Propuesto

En los párrafos anteriores se presentaron las metodologías más destacadas de las actividades que deben realizarse para una aplicación de caracterización de la carga. Se expusieron algunas técnicas de muestreo utilizadas en la literatura técnica, se detalló el proceso de realizar una minería de datos y se explicaron algunas de las distintas técnicas de agrupamiento que se utilizan para estos fines.

A modo de resumen en la figura 3.8 se presenta un diagrama de flujo de todas las actividades que debe incluir un estudio de este tipo basado en una campaña de medición.

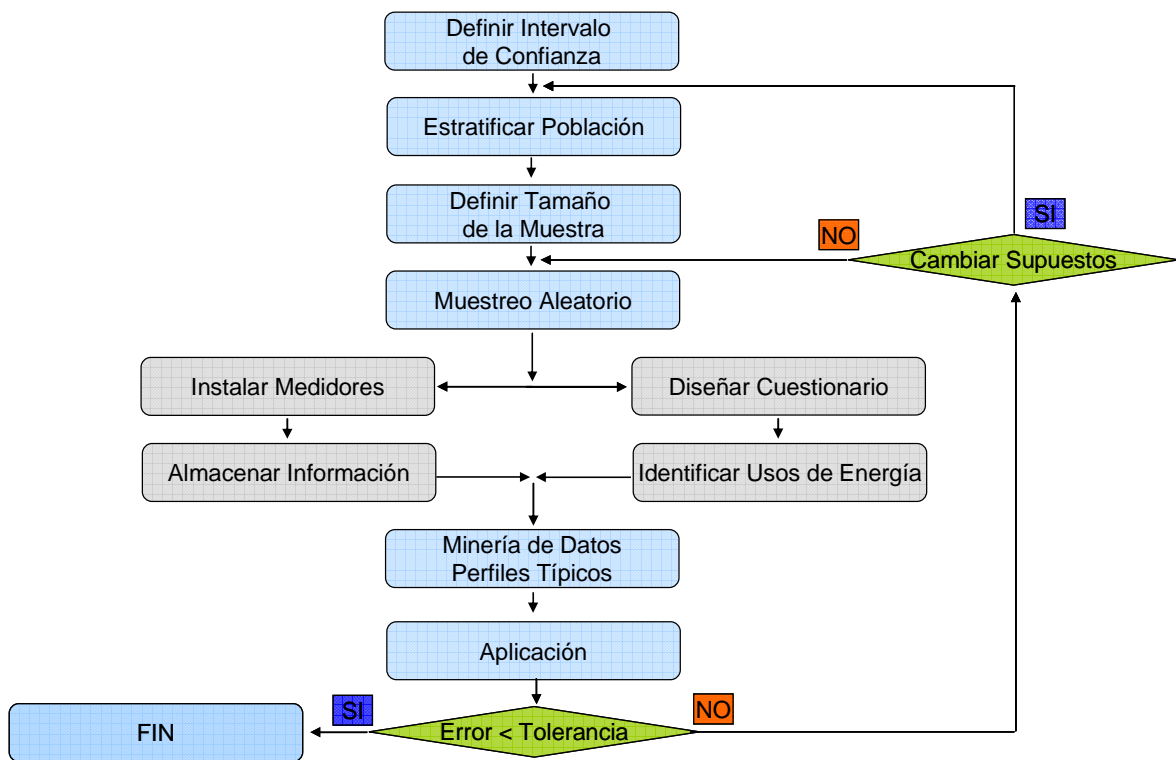


Figura 3.8: Diagrama completo de una caracterización de la carga.

En el diagrama destaca la propuesta de diseñar un cuestionario que se entregará a los clientes a los que se les instale medidores, con el fin de identificar los usos que se le da a la energía. Así, en caso de encontrar alguna correlación entre el perfil de consumo y el resultado del cuestionario, se podrá asignar perfiles de consumo aplicando sólo el cuestionario.

4 Experiencia Previa en Campañas de Medición

En el capítulo anterior se presentó el proceso completo que abarca un estudio de *Load Research*, específicamente un estudio de caracterización de la carga. A nivel mundial estos estudios se han realizado masivamente, sin embargo la información disponible es escasa, debido principalmente a la confidencialidad de los resultados. En este capítulo se mostrarán algunos casos internacionales de campañas de medición que han sido efectuadas con fines tarifarios. También se presentará el último estudio que se realizó en la empresa Chilectra. Se incluye un estudio realizado en Taiwán, que funciona como un mercado monopólico integrado verticalmente, un estudio realizado en Perú, que pertenece a un mercado regulado en distribución, y el caso del estado de Maryland (EEUU), donde se ha introducido la comercialización de energía a nivel de distribución.

4.1 Caso Taiwán

Taiwan Power Company (Taipower) es una empresas estatal que se encarga del suministro eléctrico del país. Actualmente posee una participación dominante en las redes de transmisión y distribución; no obstante los activos de generación están divididos en varias empresas, desde que en 1994 se permitió a los Productores de Independientes de Energía (PIE) ingresar al sector. A finales de 2002, Taipower controlaba una capacidad instalada superior a los 32.000 MW, de los cuales un 69% provenían de fuentes térmicas, un 16% de fuentes nucleares y un 14% de fuentes hidroeléctricas. Desde 2004 se ha estado estudiando en el parlamento una nueva Ley Eléctrica con objeto de lograr la privatización de Taipower, debido a la aparente inhabilidad de Taipower para construir la capacidad suficiente para satisfacer la demanda.

Durante el año 1992 la energía vendida alcanzó los 92 GWh y la capacidad instalada era de 19.562 MW. El aumento promedio anual de la demanda en los años anteriores alcanzaba un 8,6%, mientras que la reserva de capacidad era sólo de un 6,3%. Por esta razón, investigar la curva de demanda mediante una campaña de medición se volvió crítico. Así, a través del diseño de una nueva tarifa, se implementó un esquema de *load management*. [6, 7, 8].

Durante la década del 90 se han realizado varias campañas de medición en Taiwán, todas con características muy similares. Para fines de esta memoria se ha seleccionado la campaña realizada en el distrito de Kaohsiung en 1997.

4.1.1 Diseño de la muestra

Para determinar la curva característica de cada clase de consumidor la forma mas precisa sería medir a todos los clientes que pertenecen a la clase, sin embargo esto resulta impracticable por los costos asociados y por la gran cantidad de datos que habría que almacenar para realizar el

análisis. Por este motivo fue necesario aplicar teoría de muestreo para seleccionar el tamaño adecuado de la muestra. Con este método los perfiles de carga para cada clase de clientes se pueden determinar con un determinado nivel de confianza.

El método utilizado es el muestreo aleatorio estratificado de Neymum [6, 7, 8, 10, 11]. Los estratos utilizados para realizar el muestreo son 9 y corresponden a las distintas opciones tarifarias. La variable que se utilizó para determinar el número de clientes que serían medidos la potencia contratada. Se consideró además la varianza como valor de ponderación para cada estrato. El nivel de confiabilidad requerido corresponde a un 95%.

De una población total de 624.381 clientes los resultados que se muestran en la tabla 4.1.

Estado de Kaohsiung			
Estrato	Descripción	Muestra	Población
1	Residencial (110 V)	121	533.242
2	Comercial (110 V)	95	69.613
3	Comercial (220 V)	31	3.274
4	No Comercial (220 V)	42	1.072
5	Industrial (220 V)	39	15.964
6	Comercial (11 kV)	45	235
7	No Comercial (11 kV)	43	185
8	Industrial (11 kV)	44	756
9	Industrial (69 o 161 kV)	40	40
Total		500	624.381

Tabla 4.1: Muestreo propuesto para el estado de Kaohsiung.

El muestreo propuesto considera 500 mediciones que obedecen al nivel de confiabilidad propuesto y a la varianza del consumo de cada estrato. De esta forma se puede observar que a pesar de que el 85% de los clientes son residenciales, la muestra seleccionada es muy pequeña, ya que la desviación estándar del consumo de energía es muy baja. Por otro lado en el caso de los clientes industriales de alta tensión (estrato 9) la muestra seleccionada es la totalidad de la población, ya que la desviación estándar del consumo de energía es alta y la demanda de potencia representa más de la mitad del sistema.

4.1.2 Obtención de las curvas características

Luego de definir los clientes que se medirían se instalaron los 500 medidores inteligentes que almacenan información de consumo de energía activa y reactiva cada 15 minutos, con el fin de registrar el comportamiento estacional de la curva durante un periodo de dos años. Además se diseñó un cuestionario que fue entregado a cada cliente para efectos de conocer los usos típicos de energía, con el fin de poder proyectar los resultados obtenidos aplicando este cuestionario a los clientes que no serán medidos.

Terminado el proceso de medición se determina la curva típica para un día laboral representativo para cada estrato. Luego, con la información de facturación se puede hacer una proyección de la curva para la totalidad de la población.

4.1.3 Aplicación

La aplicación diseñada con esta campaña fue poder determinar el aporte de cada clase de cliente a la punta del sistema, con el fin de diseñar una estructura de tarifas que sea proporcional al aporte a la máxima demanda. Los resultados del aporte de cada estrato a la demanda agregada se pueden ver en la figura 4.1.

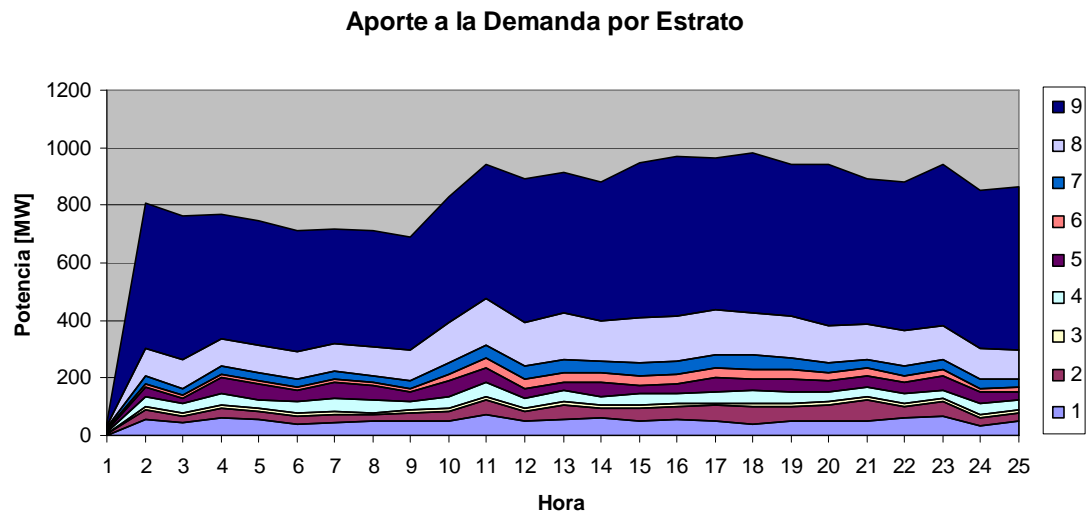


Figura 4.1: Aporte a la demanda por estrato en el estado de Kaohsiung.

Para verificar el correcto resultado de la aplicación se realizó una comparación de la curva de carga agregada estimada, con las mediciones instantáneas del sistema SCADA del centro de despacho de Taipower.

Acorde a los aportes de cada estrato de clientes a la demanda del sistema, se propuso una estructura tarifaria proporcional a los costos de capacidad de potencia llamada Método de Green, basado en el teoría de *Peak Load Pricing*.

La contribución al peak de potencia del sistema F_i de cada clase i se muestra en la ecuación (4.1).

$$F_i = D_i \cdot \frac{D_s - A_s}{\sum D_j - A_s} + A_i \cdot \frac{\sum D_j - D_s}{\sum D_j - A_s} \quad (4.1)$$

Donde,

D_i : peak de demanda de los clientes de la clase i

A_i : promedio de demanda de los clientes de la clase i

D_s : peak de demanda del sistema

A_s : promedio de carga del sistema

El costo de capacidad de potencia se calcula mediante la ecuación 4.2, que recoge los aportes a la demanda de cada estrato de clientes.

$$C_i = C_s \cdot \frac{F_i}{\sum_{k=1}^n F_k} \quad (4.2)$$

4.2 Caso Maryland

En el marco de la desregulación del sector eléctrico en el estado de Maryland, EEUU, se decidió implementar el sistema de comercialización de energía y potencia para los usuarios finales. Es por ello que se diseñó un plan de tarifas a través de perfiles típicos de consumo. Dado que la mayoría de los clientes residenciales posee un medidor simple de energía, fue necesario realizar mediciones para determinar cuales eran los distintos perfiles típicos de consumo para asignar a cada cliente. [21]

Se creó un organismo llamado *GTIWG*⁸, que sería el encargado de asesorar y recomendar a las empresas para la aplicación del sistema de mercado minorista. Este organismo define una serie de actividades que se debe realizar para determinar los perfiles típicos, entre ellas destacan:

- Diseñar un muestreo de datos acorde a los reglamentos y protocolos establecidos.
- Implementar *Load Research* para determinar los perfiles de consumo para todas las clases de clientes.
- Verificar la exactitud de los datos y de los perfiles obtenidos.

Se mencionan tres métodos que pueden ser utilizados para determinar los perfiles de consumo:

- Perfiles Estáticos: Se basan en datos históricos que determinan un perfil de consumo fijo.
- Perfiles Dinámicos Modelados: Se basan en datos históricos que determinan un perfil de consumo que se ajusta temporalmente mediante diversos parámetros, principalmente como función del estado del tiempo.

⁸ *Generic Technical Implementation Working Group*

- **Perfiles Dinámicos Medidos:** Se obtienen a través de campañas de medición diseñadas con teoría de muestreo, que se realizan temporalmente de tal manera de ajustar los perfiles en el tiempo.

Cada una de las empresas participantes en el estado de Maryland diseñó su propia campaña para obtener los perfiles típicos de sus clientes, utilizando alguno de los tres métodos explicados recientemente, que se ajustaban a sus universos de clientes.

4.3 Caso Perú

El estudio de caracterización de carga, tiene por objetivo la determinación de los factores que caracterizan el comportamiento de los diagramas de cargas de clientes, grupos de clientes y del sistema eléctrico en su conjunto. Esta información es el soporte para la regulación tarifaria, dentro del proceso de fijación de las tarifas de distribución eléctrica realizado por *OSINERG*⁹ para el periodo Noviembre 2005 – Octubre 2009. [22]

La metodología empleada en el estudio involucra las siguientes actividades: análisis de muestreo estadístico, desarrollo de una campaña de medición, procesos de control de calidad de la información y finalmente el cálculo de los factores de coincidencia y contribución a la máxima demanda de los grupos de clientes. Para el presente informe se detallará la campaña realizada en el sector típico 2, que corresponde al sistema eléctrico de la ciudad de Huancayo.

4.3.1 Diseño de la muestra

Para el diseño de la muestra se consideró como universo del estudio a clientes libres y regulados de todas las opciones tarifarias (Media Tensión y Baja Tensión). Dentro de las opciones tarifarias existen tarifas binomiales, que incluyen cobro por energía y potencia y la tarifa BT5B, que sería el caso análogo a la tarifa BT-1a de Chile. Para el elegir que clientes serían medidos se definieron tres estratos: Media Tensión, Baja Tensión y tarifa BT5B. Los dos primeros estratos estarían divididos en sub-estratos que serían las opciones tarifarias, mientras que el estrato BT5B estará dividido según la ubicación geográfica.

Se determinó que en Media Tensión se medirían 29 clientes y en Baja Tensión 34 clientes. Para la tarifa BT5B se realizarían 50 mediciones de forma indirecta, tomando una muestra sobre las Subestaciones de Distribución (SED) que tengan al menos un 85% del consumo de energía realizado por clientes BT5B.

⁹ Oficina Supervisora de Inversiones en Energía

Para decidir cuantos clientes se medirían en cada estrato se utilizó la relación que muestra la ecuación (4.3). Los resultados finales del muestro se muestran en las tablas 4.2, 4.3 y 4.4.

$$n_i = \left(\frac{0,65 \cdot \sum_{j=1}^{N_i} E_j}{\sum_{j=1}^N E_j} + \frac{0,35 \cdot N_i}{N} \right) \cdot n \quad (4.3)$$

Donde

n_i : Clientes medidos en el estrato i

n : Total de clientes medidos

N_i : Total de clientes en el estrato i

N : Total de clientes

E_j : Energia mensual consumida por el cliente j

Sector Típico 2: Huancayo		
Estrato BT	Muestra	Población
BT2	1	4
BT3FP	2	6
BT3P	1	2
BT4FP	7	35
BT4P	9	17
BT5A	14	14
Total	34	78

Tabla 4.2: Muestreo en estrato BT.

Sector Típico 2: Huancayo		
Estrato MT	Muestra	Población
MT2	3	9
MT3FP	10	19
MT3P	5	7
MT4FP	5	41
MT4P	6	6
Total	29	82

Tabla 4.3: Muestreo en estrato MT.

Sector Típico 2: Huancayo		
Estrato BT5B	Muestra	Población
Huancayo	22	38
Chilca	4	241
El Tambo	23	208
San Agustín	1	8
Total	50	495

Tabla 4.4: Muestreo en estrato BT5B.

4.3.2 Obtención de las curvas

Para la obtención de las curvas típicas se agruparon las muestras de distintos estratos considerando varias tarifas que tienen comportamientos similares y luego se realizó un análisis de clustering a través de árboles de decisión utilizando la distancia euclídeana. A modo de ejemplo se detallará el resultado obtenido para el estrato de tarifa BT5B.

El estrato BT5B se analizó en forma independiente y los resultados obtenidos son 4 curvas típicas. La distribución de cada tipo se muestra en la tabla 4.5.

Estrato BT5B	
Cluster	N° SED
Tipo 1	44
Tipo 2	1
Tipo 3	4
Tipo 4	1
Total	50

Tabla 4.5: Distribución de cada SED en curvas típicas.

En la figura 4.2 se muestran las curvas típicas encontradas para el día representativo seleccionado. [22]

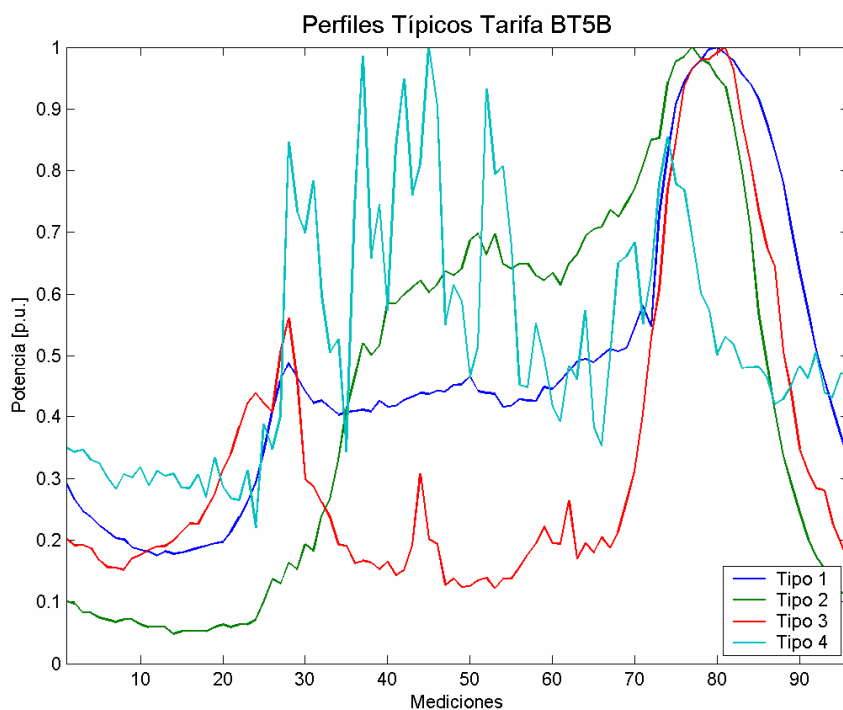


Figura 4.2: Curvas típicas de las SED.

4.3.3 Aplicación

La finalidad de esta campaña de medición era calcular los parámetros que se utilizan en la regulación de tarifas. Para ello se determinaron las curvas típicas de cada estrato (o grupo de estratos) y luego se proyectó la curva agregada para toda la muestra y se realizaron cálculos de factor de carga, factor de pérdidas y número de horas de uso según muestran las ecuaciones (4.4), (4.5) y (4.6).

Factor de Carga:

$$F_c = \frac{P_{med}}{P_{max}} \quad (4.4)$$

Factor de Perdidas:

$$F_p = \sum_{i=1}^{96} \frac{\left(\frac{P_i}{P_{max}}\right)^2}{96} \quad (4.5)$$

Número de Horas de Uso:

$$NHUBT = \frac{E_{mes}(\text{promedio})}{P_{max}(\text{coincidente con la del sistema})} \quad (4.6)$$

Donde

P_{med} : Potencia media de la curva típica

P_{max} : Potencia máxima de la curva típica

P_i : Potencia en instante i

E_{mes} : Energía mensual

En la figura 4.3 se muestra el resultado obtenido para la curva agregada de la tarifa BT5B en el sector de Huancayo.

En la tabla 4.6 se muestra un resumen con los resultados obtenidos para los parámetros factor de carga, factor de pérdidas y número de horas de uso de la tarifa BT5B.

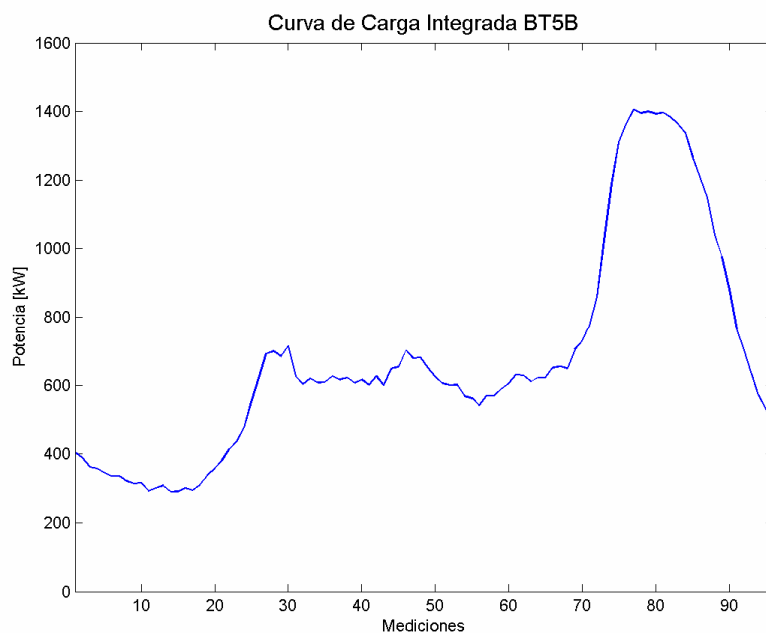


Figura 4.3: Curva de carga agregada de la tarifa BT5B.

Parámetros Registrados								
Máxima Demanda			Demanda Promedio			Energía Activa		
HP	1405,6	kW	HP	1198,1	kW	HP	5991	kWh
HFP	860,4	kW	HFP	540	kW	HFP	10259	kWh
DIA	1405,4	kW	DIA	677,1	kW	DIA	16250	kWh
Parámetros Calculados								
Factor de Carga			Factor de Pérdidas			Número de Horas de Uso		
HP	0,85		HP	0,75		NHUBT1	334	hrs/mes
HFP	0,63		HFP	0,16		NHUBT2	344	hrs/mes
DIA	0,48		DIA	0,28				

Tabla 4.6: Resumen de resultados obtenidos.

4.4 Caso Chilectra

Durante el año 1999 la empresa Chilectra realizó una campaña de medición con el fin de determinar el parámetro Horas de Uso de los clientes BT-1, el cual permite la recuperación de la potencia demandada por ellos, por intermedio del consumo de energía. Un aspecto fundamental para la determinación de las “Horas de Uso” consiste en determinar la contribución de la potencia demandada por el conjunto de los clientes, que realizan estos al sistema de distribución, como también a las compras de potencia.

Para lograr el objetivo, Chilectra realizó un completo plan de mediciones a una muestra estadísticamente representativa de 1.050 clientes pertenecientes a la tarifa BT-1. Estas mediciones se efectuaron continuamente durante el transcurso de un año (período abril 1999 a marzo de 2000). A cada cliente perteneciente a la muestra, se le instaló un medidor digital con

memoria de masa suficiente para 70 días, que registró variables eléctricas, relacionadas a la potencia, tanto activa como reactiva, al voltaje suministrado, a la energía consumida y a los cortes de suministro, segregando estos últimos en cortes mayores a tres minutos y menores que ese tiempo.

4.4.1 Diseño de la muestra

En el diseño de la muestra se definieron cuatro categorías de consumidores, denominadas residenciales, comerciales, industriales y otros. En la categoría otros se agruparon los clientes clasificados como municipales, fiscales, agrícolas y de utilidad pública. Por otra parte se definieron cuatro rangos de consumo para cada una de las categorías señaladas anteriormente. Los rangos escogidos fueron 0 a 110 kWh; 111 a 150 kWh; 151 a 300 kWh y mayor a 300 kWh. Con estos criterios de clasificación se generaron 16 estratos de clientes a los cuales se les estudió su demanda durante el período de muestreo. Dada esta estratificación se diseñó una muestra, en que el tamaño de cada estrato se hizo de carácter proporcional a la población. El número total de clientes muestreados se determinó arbitrariamente en 1.050. En la tabla 4.7 se muestra el resultado que se obtuvo [23].

Estrato	Proporción [%]	Tamaño de la Muestra
Com1	2,18	23
Com2	0,6	6
Com3	1,35	14
Com4	2,56	27
Ind1	0,07	1
Ind2	0,02	0
Ind3	0,06	1
Ind4	0,19	2
Otr1	0,07	1
Otr2	0,02	0
Otr3	0,07	1
Otr4	0,29	3
Res1	33,96	356
Res2	18,88	198
Res3	29,12	306
Res4	10,54	111
TOTAL	100	1050

Tabla 4.7: Muestreo realizado.

4.4.2 Aplicación

De las mediciones efectuadas a los clientes de la muestra, se estimó la potencia demandada por la totalidad de los clientes BT-1. En particular se determinó la potencia de los clientes BT-1 para el día 30 de agosto de 1999, fecha en la cual se registró la máxima demanda de potencia en horas de punta, tanto en el sistema de distribución de Chilectra como en el Sistema Interconectado Central (SIC).

La metodología empleada para efectuar esta estimación consistió en extrapolar los resultados de cada estrato de consumo definido para este estudio (Tabla 4.7). En cada uno de los 16 estratos se obtuvo un valor promedio de curva de demanda para el día elegido, valores que fueron multiplicados por el número total de clientes que pertenecen a cada estrato. Luego estos valores fueron sumados y se obtuvo la curva de carga estimada.

En la ecuación (4.7) se muestra la expresión que describe el proceso.

$$\hat{Y}_t = \sum_{i=1}^{16} N_i \cdot \bar{x}_{i,t} \quad (4.7)$$

Donde

\hat{Y}_t : Valor de la potencia estimada en el tiempo t

N_i : Número total de clientes que pertenecen al estrato i

$\bar{x}_{i,t}$: Promedio de la potencia por cliente en el estrato i y en el tiempo t

En la figura 4.4, se muestra la curva de carga resultante para los clientes BT-1, conjuntamente con la curva de carga medida a la entrada del sistema de distribución de la empresa, ambas para el día 30 de agosto de 1999 [24].

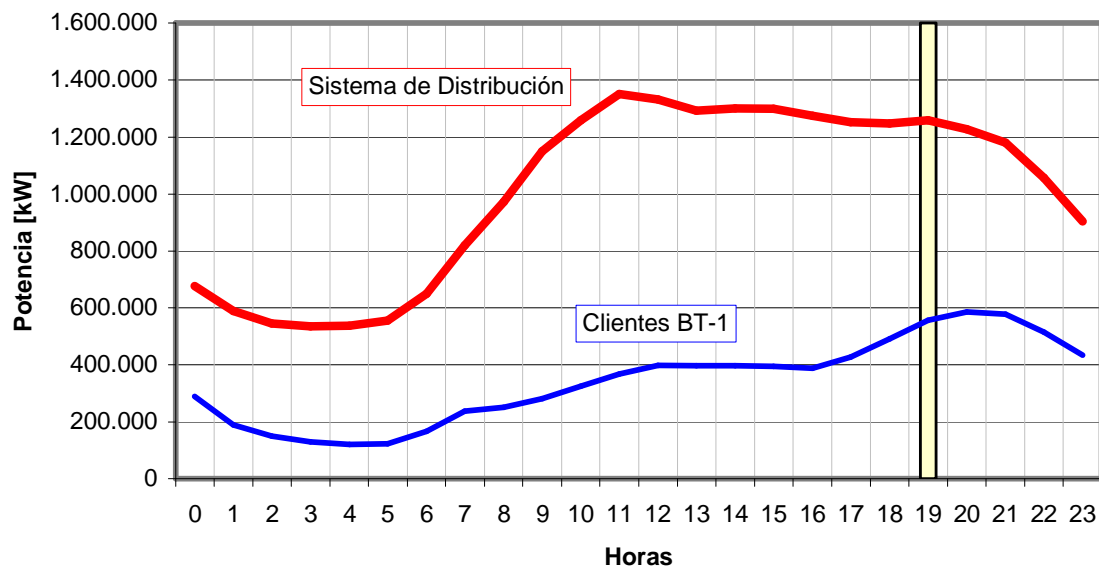


Figura 4.4: Curva de carga resultante para los clientes BT-1

Para el día 30 de agosto a las 19 horas la demanda máxima del sistema de distribución alcanzaba a 1.258,246 MW, en tanto que la contribución de demanda de los clientes BT-1 alcanzaba a esa misma hora a 574,301 MW.

A partir de estos resultados y de la base de consumo de energía se determinaron las horas de uso para esta tarifa, arrojando como resultado 380,5 horas.

5 Diseño de la muestra para una campaña de medición en Chilectra

En los capítulos anteriores se revisó la metodología necesaria para realizar un estudio de curvas de carga de una empresa distribuidora de energía eléctrica, se presentaron algunas de las aplicaciones que se pueden diseñar a partir del estudio y además se entregó una visión general, incluyendo algunos ejemplos, del desarrollo a nivel internacional en la materia, considerando como aplicación el diseño de tarifas o el cálculo de parámetros tarifarios.

En este capítulo se realizará un estudio estadístico utilizando la base de datos de consumos de energía de los clientes de la empresa Chilectra, específicamente el universo de clientes BT-1. El estudio considerará el diseño y la propuesta de una muestra de clientes representativa, que servirá como base para una futura campaña de medición que tenga como objetivo caracterizar la curva de demanda.

5.1 Información de la Población Objetivo

Chilectra distribuye energía a un total de 1.310.691¹⁰ clientes BT-1, que en conjunto consumen 3.522.356 MWh anualmente (periodo sep 05 – ago 06). Sin embargo, dentro de este grupo de clientes BT-1 existe una diversidad de clases de clientes, que se diferencian por su naturaleza. Estas clases son: Residenciales, que representan un 92% de la población, comerciales, que representan un 6,7%, industriales, agrícolas, fiscales, municipales, FF.CC y utilidad pública, que en conjunto representan al 1,3% del total.

5.2 Diseño de la Muestra

Como se ha explicado en el capítulo 3, el diseño de la muestra es un paso fundamental de la campaña de medición, es por esto que hay que decidir que estrategia resultará más conveniente para la aplicación que se quiere implementar. Dada la experiencia internacional, y los estudios previos realizados por la empresa se ha decidido utilizar un muestreo estratificado, con el objeto de obtener una muestra que asegure una representatividad a todas las clases involucradas, ya que al hacer un muestreo aleatorio simple se corre el riesgo de que los tipos de clientes menos numerosos no sean elegidos. [23, 24]

5.2.1 Estratificación

Para hacer una estratificación correcta es necesario que cada cliente pertenezca sólo a un estrato, es decir, que no exista superposición. Además es importante que cada estrato este formado por clientes con características homogéneas.

¹⁰ Fuente: evolución masiva 04 – 06 de Chilectra.

Por ello la primera hipótesis planteada para seleccionar la muestra será dividir la población BT-1 según el tipo de cliente, ya que los perfiles de consumo de cada tipo deberían tener un comportamiento similar. Sin embargo, como son muchos tipos, los menos frecuentes se agruparán en una sola categoría. Así, se definen las siguientes categorías caracterizadas por el índice i como se muestra en la tabla 5.1.

Indice i	Categoría
1	Residencial
2	Comercial
3	Industrial
4	Otros

Tabla 5.1: Categorías de clientes.

También es importante destacar que cada una de estas categorías definidas anteriormente, posee distintos niveles de consumo. Por esta razón se definirá el índice j , que separará el promedio de consumo de energía mensual de cada cliente según los intervalos que se muestran en la tabla 5.2.

Indice j	Rango [kWh/mes]
1	0-110
2	111-150
3	151-300
4	>301

Tabla 5.2: Intervalos de consumo promedio de energía de clientes.

Los niveles de energía se han definido con esos intervalos aprovechando la experiencia del estudio de las horas de uso desarrollado anteriormente en la empresa, aunque se trata de un criterio revisable a la luz de nueva información. Otro criterio podría ser utilizar los cuartiles de energía para efectuar la definición.

Utilizando las categorías descritas en la tabla 5.1 y en la tabla 5.2 se divide a la población en 16 estratos definidos por los índices $E_{(i,j)}$. Así, por ejemplo el estrato $E_{(2,3)}$ representa a los clientes comerciales con consumo promedio mensual entre 150 y 300 kWh. Para que la lectura resulte más fácil a cada estrato se le asignará el nombre que se muestra en la tabla 5.3.

		Categoría			
		Indice i			
Intervalo de Energía	Indice j	1	2	3	4
	1	RES1	COM1	IND1	OTR1
	2	RES2	COM2	IND2	OTR2
	3	RES3	COM3	IND3	OTR3
	4	RES4	COM4	IND4	OTR4

Tabla 5.3: Estratos de consumo.

5.2.2 Muestreo

Como se expuso en el capítulo 3, la literatura consultada entrega tres propuestas distintas para realizar el muestreo. Para escoger qué técnica es la más adecuada se debe chequear que la distribución que sigue la población con respecto a la variable consumo promedio mensual se asemeje a una distribución normal, ya que de esta forma se podrá aplicar la teoría de intervalos de confianza sin cometer un error significativo. Para estos fines se ha realizado un test de ajuste a la distribución normal, que efectivamente entregó como resultado esta distribución para valores cercanos al promedio de consumo.

Dado este resultado, el muestreo que se utilizará será el explicado en la propuesta 3, del capítulo 3, que considera la varianza de cada estrato y el costo de realizar la campaña. Se utilizará la estratificación definida en el punto 5.2.1, por lo que será necesario conocer los consumos facturados por cada cliente y su categoría para poder asignar cada cliente al estrato al cual pertenece. Los resultados se muestran en la tabla 5.4.

Estrato	Promedio Mensual [kWh]	Desviación Estandar [kWh]	Población
Res1	69,4	31,9	302.200
Res2	130,5	11,5	211.623
Res3	210,4	41,0	488.055
Res4	482,4	240,3	203.934
Com1	42,9	36,4	28.826
Com2	130,2	11,4	7.345
Com3	217,6	43,2	18.977
Com4	792,6	645,5	33.988
Ind1	33,7	36,7	932
Ind2	131,8	12,0	179
Ind3	224,1	42,8	684
Ind4	947,9	1.005,6	1.906
Otr1	32,4	37,9	1.956
Otr2	129,7	11,6	386
Otr3	224,8	42,7	1.480
Otr4	909,3	599,5	8.220
TOTAL	223,4	234,0	1.310.691

Tabla 5.4: Información estadística de los estratos.

El paso siguiente es encontrar la ponderación que debe tener cada estrato sobre la muestra, ya que el número total se debe determinar considerando un intervalo de confianza para el consumo de energía promedio y los costos asociados a realizar cada medición según lo planteado por las ecuaciones (3.10) del capítulo 3. Los resultados de la ponderación se muestran en la tabla 5.5, donde se compara el resultado de realizar un muestreo ponderado por varianza, frente al muestreo que sea proporcional a la población.

Estrato	Ponderación [%]	Población [%]
Res1	8,597	23,06
Res2	2,17	16,15
Res3	17,85	37,24
Res4	43,75	15,56
Com1	0,94	2,2
Com2	0,07	0,56
Com3	0,73	1,45
Com4	19,59	2,59
Ind1	0,03	0,07
Ind2	0	0,01
Ind3	0,03	0,05
Ind4	1,71	0,15
Otr1	0,07	0,15
Otr2	0	0,03
Otr3	0,06	0,11
Otr4	4,4	0,63
TOTAL	100	100

Tabla 5.5: Ponderación de cada estrato en la muestra.

A partir de los resultados expuestos en la tabla 5.5 se observa que al considerar la varianza, los estratos mas grandes resultaron ser los que tienen el consumo de energía mas elevado para cada categoría de cliente, es decir, los estratos Res4, Com4, Ind4 y Otr4. De cualquier modo, esto es una ventaja, ya que es conveniente tener una mayor precisión sobre los clientes que poseen una mayor variabilidad, considerando la hipótesis de que las curvas de cada cliente en cada estrato serán similares acorde al consumo de energía.

Posteriormente se debe determinar el número de clientes que se debe incluir en la muestra. Para esto se utilizará la expresión de intervalos de confianza que se propone en el método y que se muestra en las ecuaciones (3.11) a (3.13). Se escogen tres niveles de probabilidad para realizar la comparación de cómo varía el número de elementos necesarios en la muestra. En la figura 5.1 se exponen los resultados obtenidos y en la tabla 5.6 se detallan algunos valores representativos para tamaño de muestra.

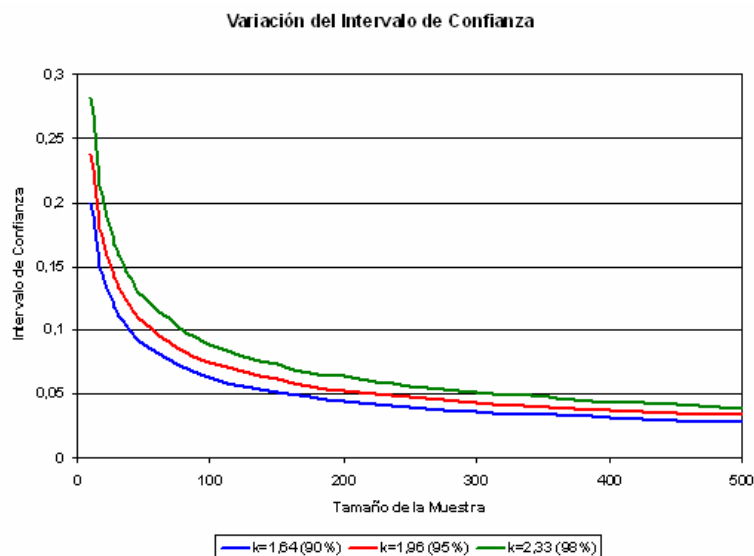


Figura 5.1: Variación del intervalo de confianza en función del tamaño de muestra.

	k=1,64	k=1,96	k=2,33
Intervalo 10%	40	57	80
Intervalo 5%	158	225	318

Tabla 5.6: tamaño de la muestra para distintos valores de k.

En la figura 5.1 es claro ver que a medida que se requiere un intervalo de confianza más pequeño el tamaño de la muestra debe ser mayor. Lo mismo sucede cuando se aumenta la probabilidad de pertenecer a ese intervalo. Por ello, si se quiere determinar el tamaño de la muestra para un intervalo de confianza de un 5% con una probabilidad de un 98% de pertenecer a él, se necesita una muestra de 318 clientes. Sin embargo, es conveniente escoger un tamaño mayor, debido a que el periodo de medición puede ser muy extenso y pueden ocurrir algunos imprevistos, como por ejemplo la falla de un equipo de medida o el cambio de domicilio de algún cliente. En la tabla 5.7 se muestra un resumen que indica cuantos clientes deben seleccionarse por estrato para distintos tamaños de muestra considerando el intervalo y la probabilidad antes mencionados. Finalmente resta comprobar si el presupuesto disponible es suficiente para medir la muestra resultante. Ese fin no se encuentra dentro de los objetivos de esta memoria.

Estrato	Ponderación [%]	Tamaño de la muestra		
Res1	8,597	27	30	34
Res2	2,17	7	8	9
Res3	17,85	57	63	72
Res4	43,75	140	153	175
Com1	0,94	3	3	4
Com2	0,07	0	0	0
Com3	0,73	2	3	3
Com4	19,59	63	69	78
Ind1	0,03	0	0	0
Ind2	0	0	0	0
Ind3	0,03	0	0	0
Ind4	1,71	5	6	7
Otr1	0,07	0	0	0
Otr2	0	0	0	0
Otr3	0,06	0	0	0
Otr4	4,4	14	15	18
TOTAL	100	318	350	400

Tabla 5.7: Estratificación para distintos tamaños de muestra.

Recapitulando lo planteado en las últimas páginas, se ha presentado una metodología para determinar el tamaño que debe tener una muestra, que cumpla con los requisitos estadísticos de intervalos de confianza y probabilidad, y logre el objetivo de reducir la población a un número tal que se pueda realizar un estudio de la curva de carga de los clientes BT-1.

6 Aplicación de *Load Research*: Análisis de Estacionalidad

Uno de los principales problemas para la empresa distribuidora, es el desconocimiento que se tiene acerca del comportamiento eléctrico de los clientes, principalmente los clientes que son masivos y que poseen un medidor de energía simple. En el caso chileno, estos clientes son los que poseen la tarifa BT-1, que representan más del 90% de los clientes y cerca del 30% del consumo de energía. Para las empresas sería de gran utilidad poder conocer cual es la curva de consumo de este gran grupo de clientes. Sin embargo, no se justifica económicamente medirlos a todos, toda vez que los costos actuales de inversión y de operación para ejecutar una actividad de esta envergadura, supera con largueza los potenciales beneficios económicos que se pueden obtener.

En el capítulo 5 se ha expuesto las bases sobre las cuales se recomienda desarrollar una campaña de medición de los clientes BT-1 de Chilectra. Dado que una campaña de esta naturaleza toma tiempos que exceden al desarrollo de esta memoria de titulación, para los efectos de análisis de datos, de resultados y conclusiones, se utilizarán las mediciones realizadas durante el año 1999,

En este estudio de estacionalidad se puede realizar dos análisis independientes que pueden entregar más información y así obtener mejores conclusiones. Un estudio de tipo cualitativo será analizar la estacionalidad de las curvas utilizando técnicas de agrupamiento, es decir, estudiar como cambia la forma de cada curva típica o *cluster* durante el año y a la vez analizar la pertenencia estacional de los clientes a cada cluster. El otro estudio posible está relacionado con la curva agregada de la muestra, lo que permitirá efectuar algunos cálculos que demuestren cual es el comportamiento de la máxima demanda a lo largo del año.

En los párrafos siguientes se desarrollarán los pasos necesarios para un estudio *KDD* basado en el esquema de la figura 3.2 explicado en el capítulo 3.

6.1 Selección de datos objetivo

La base de datos cuenta con un total de 1.050 mediciones de energía activa y reactiva de clientes BT-1, efectuadas en intervalos de 1 hora (en principio era cada 15 minutos, pero luego se integraron las medidas para que la base de datos no sea tan grande) entre los meses de abril de 1999 y marzo de 2000.

En la tabla 6.1 se muestra el formato que posee esta base de medidas.

N° Cliente	Estrato	Fecha	Hora	Potencia

Tabla 6.1: Formato de la base de mediciones.

Sin embargo, sólo algunos clientes están medidos por los 12 meses, ya que se presentaron algunos imprevistos en esa campaña y no fue posible tener la totalidad de las mediciones por todo el año. Como uno de los objetivos principales de esta memoria es estudiar la estacionalidad de las curvas, es conveniente que sean los mismos clientes los que se analicen durante un año completo.

Otro análisis previo consiste en buscar dentro de la base de datos algunos días que no tuvieran todas las mediciones o que las medidas fueran muy desproporcionadas en comparación a valores esperados (*outliers*). Realizar un filtraje de este tipo resulta muy complicado en el caso de mediciones individuales, ya que la curva de consumo es muy dependiente de las actividades que efectúe el cliente en cada uno de los días medidos. Distinto sería el caso de filtrar curvas erróneas para un transformador o un conjunto de clientes, ya que las curvas agregadas sí poseen un comportamiento típico definido. Sin embargo, se encontraron algunos errores de ese tipo, aunque fueron muy pocos principalmente porque estas mediciones ya se han utilizado para estudios previos y han sido reparadas en esas instancias. Así, se ha llegado a un resultado de 398 mediciones válidas para los objetivos de esta memoria.

Es por ello que se ha decidido tomar la base de 398 clientes, y aunque las ponderaciones no correspondan al muestreo propuesto, si se puede asegurar que son proporcionales a la población.

6.2 Preproceso de Datos

Las mediciones se presentan como una matriz de datos, con la identificación del cliente, la fecha, la hora y las mediciones de energía. Dada esta configuración, se debe seleccionar un día típico para realizar el estudio, ya que manejar demasiada información puede afectar el desarrollo de la minería de datos. Por ello, y porque el estudio se quiere realizar para cada mes, se calculará un día promedio mensual para cada cliente, ya que pueden existir días con consumos atípicos y la selección de un solo día puede arrojar resultados no representativos.

Como el objetivo es encontrar perfiles de consumo, los días promedio resultan ser una buena aproximación, ya que de esta forma un cliente que posee variaciones en sus consumos diarios tendrá una curva esperada que corresponderá a este valor.

Considerando lo expuesto en párrafo anterior, resta decidir que día será considerado como el día de consumo típico. Para efectos de este trabajo, la selección de día típico corresponde a los días miércoles, utilizando como criterio que para el mes de julio, que corresponde al de máxima

demanda de la muestra, la curva promedio de los días miércoles resulta ser la curva que posee la demanda mas elevada en el periodo de horas de punta. El resultado de ese análisis se muestra en la figura 6.1.

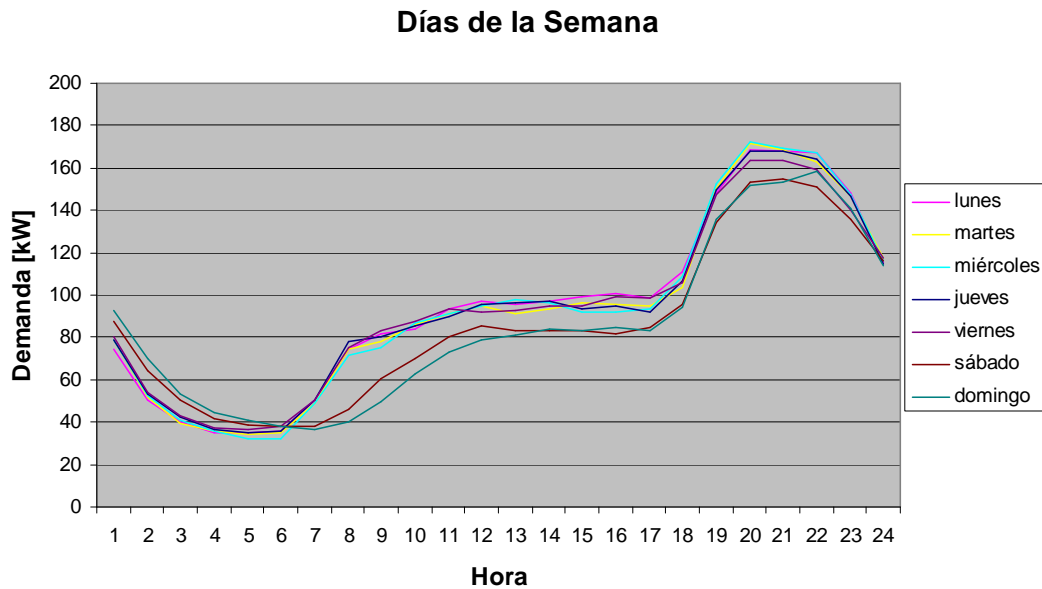


Figura 6.1: Curvas agregadas de demanda para el mes de julio.

Esta elección puede parecer a simple vista una decisión arbitraria, sin embargo no existe un estudio concluyente que diga cuál es el día típico más representativo de clientes sin medidor horario. Además, el día miércoles corresponde a un día laboral común y el comportamiento de la curva es muy similar a los demás días laborales, existiendo sólo pequeñas variaciones.

6.3 Transformación de Datos

Este punto es relevante para poder realizar una clasificación de consumos según el comportamiento horario de la curva. Como las curvas deben ser comparables con respecto a su forma, sin considerar la magnitud de cada una de ellas y como se está trabajando con medidas de distancia, se requiere que los valores se encuentren normalizados en una misma escala. Por ello, la transformación necesaria será normalizar las curvas para que cada una de ellas tome valores entre 0 y 1. Este procedimiento consiste en tomar los datos de una curva y dividirlos por su máximo valor. Una expresión genérica que representa este proceso es el que se indica en la ecuación (6.1).

$$E'_{ijk} = \frac{E_{ijk}}{\max_k \{E_{ijk}\}} \quad (6.1)$$

Donde

E_{ijk} : Medición de cliente i en el mes j en la hora k

De este modo se logra que todas las curvas puedan ser comparables por su forma en el tiempo y además la anti-transformación resulta simple, ya que sólo hay que multiplicar por el valor máximo de cada curva.

6.4 Minería de Datos

El proceso de minería de datos requiere realizar variados cálculos, eventualmente programación de algoritmos y análisis estadístico. Es por ello que se hace necesario encontrar alguna plataforma que permita realizar estas actividades, principalmente la programación de los distintos métodos de clasificación estudiados en el capítulo 3. Una opción es escoger un lenguaje de programación como JAVA o algún software de cálculo como Matlab. Cada uno de ellos presenta ventajas y desventajas sobre el otro, pero se ha decidido utilizar Matlab, ya que ha sido utilizado estudios previos de caracterización de clientes que se han revisado en la literatura técnica. Además, Matlab posee algunos *toolbox* con herramientas de *clustering* [25]. Así, los datos serán cargados a Matlab en forma matricial, con 398 filas, que representarán a los clientes y 24 columnas, que corresponderán a las mediciones horarias normalizadas.

Antes de resolver o realizar algún cálculo es importante conocer los datos. Una primera aproximación a ellos puede ser observar que ocurre con ellos en conjunto, es decir observar la forma de la curva agregada de las mediciones. Teóricamente debería tener la forma de una curva típica residencial, ya que en su gran mayoría los clientes son de este tipo. Otro análisis que debe hacerse es ver el comportamiento durante los meses, es decir, observar la diferencia existente entre el consumo (altura de la curva) para cada mes. En la figura 6.2 se presenta un gráfico que muestra el movimiento de la curva.

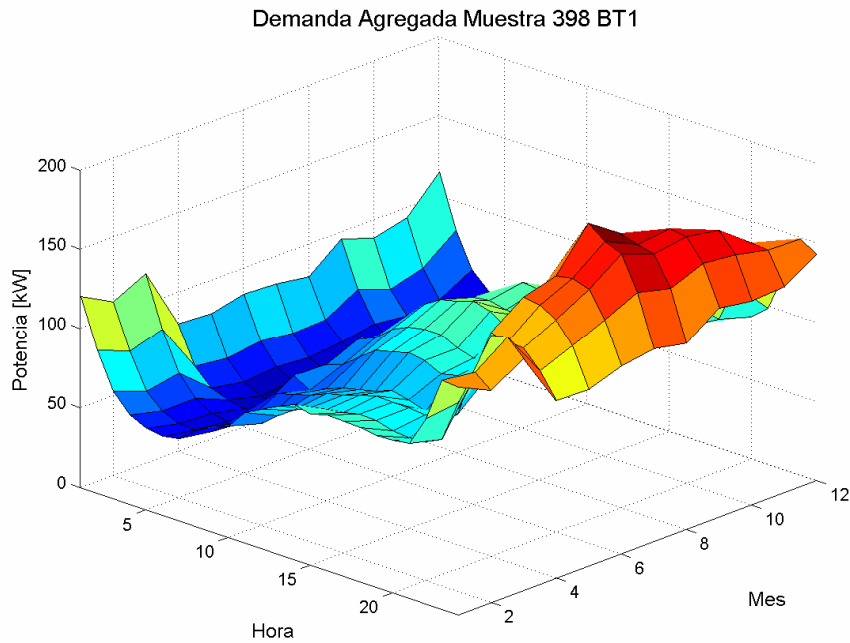


Figura 6.2: Demanda agregada de la muestra mensual.

En la figura 6.2 se muestra que la forma de la curva agregada se asemeja a una de tipo residencial, resultado que era esperado, ya que la composición de los clientes de la muestra es, en su mayoría, de este tipo. También hay que notar que el instante de máxima demanda ocurre alrededor de las 22 hrs. y como era lógico esperar, la máxima demanda se presenta en los meses de invierno.

Como se ha mencionado antes, son dos las formas de analizar la estacionalidad de la curva. En primer lugar se hará el análisis desde el punto de vista de grupos de clientes con curvas típicas. Posteriormente se analizará para las curvas agregadas.

6.4.1 Clientes agrupados en *clusters*

Para realizar un estudio de *clustering* hay que plantear ciertas hipótesis, que serán las que otorgarán validez a los resultados obtenidos. Estas hipótesis serán:

- Se considera día típico al promedio de las mediciones de los días miércoles de cada mes. De esta forma el año se representa por 12 días con comportamiento promedio.
- En caso de existir algún día miércoles festivo, éste no será considerado para el cálculo del día promedio
- No será relevante la magnitud de la demanda, sólo importará su forma en el tiempo.

- Cada mes deberá tener el mismo número de clusters, para que puedan ser comparables entre ellos.

Como se mencionó en el capítulo 3, son diversas las técnicas que existen para clasificar los consumos, sin embargo hay que decidir que técnica es la mas conveniente. Para fines exploratorios o primeros acercamientos se recomienda el uso de *k-means* [19, 20], debido a la velocidad de convergencia del método. Es por ello que se ha decidido utilizar esta técnica para determinar el número óptimo de clusters que debe tener el agrupamiento.

Dentro de las condiciones que debe cumplir una minería de datos se mencionó la simplicidad de las reglas obtenidas. Por este motivo se ha decidido que el máximo número de cluster debe ser 6, para no obtener tantas curvas que puedan resultar complejas de analizar. Además, el número de clusters debe ser al menos 2, ya que tener un solo cluster representaría el promedio de todas las curvas, lo que no entrega ninguna información adicional a la curva agregada. Para escoger el número óptimo de clusters hay que utilizar alguno de los índices de validez mencionados en el capítulo 3. El más utilizado corresponde al índice de Davies-Bouldin, y además posee la ventaja de estar implementado en Matlab [25].

Como la función *k-means* es sensible al punto de partida, el método que se utiliza realiza varios agrupamientos distintos para cada número de clusters y calcula el índice DBI para el agrupamiento que tenga la función objetivo más pequeña, es decir, la que tenga el menor error asociado. En la figura 6.3 se observa el resultado para cada uno de los meses.

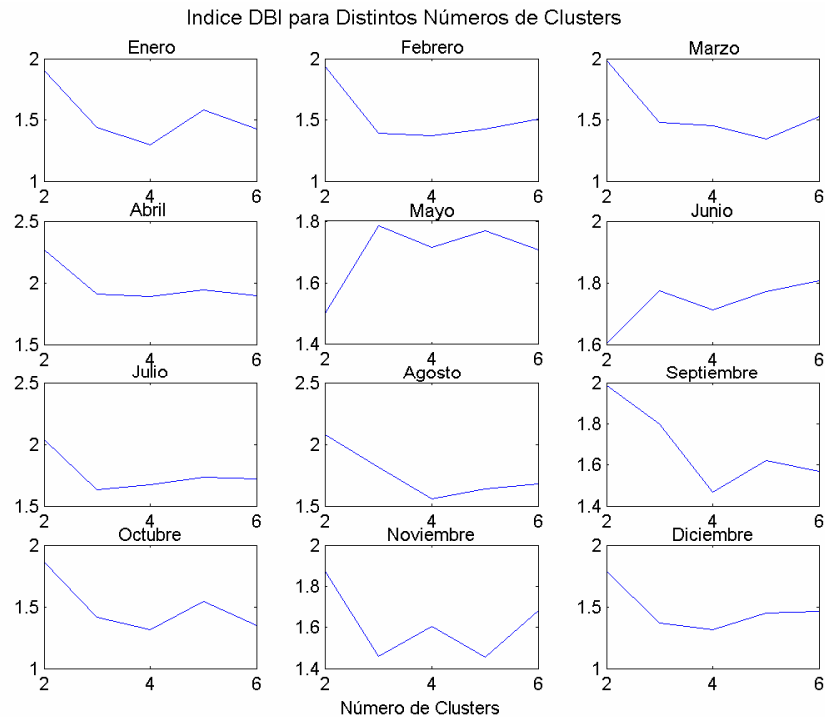


Figura 6.3: Índice Davies – Bouldin.

De la figura 6.3 se concluye que el número óptimo de clusters más frecuente es 4. Sin embargo para algunos meses el resultado es otro. Por ello en la figura 6.4 se muestra que el promedio del índice DBI es más pequeño para 4 clusters, así, se dirá que 4 clusters es el número óptimo.

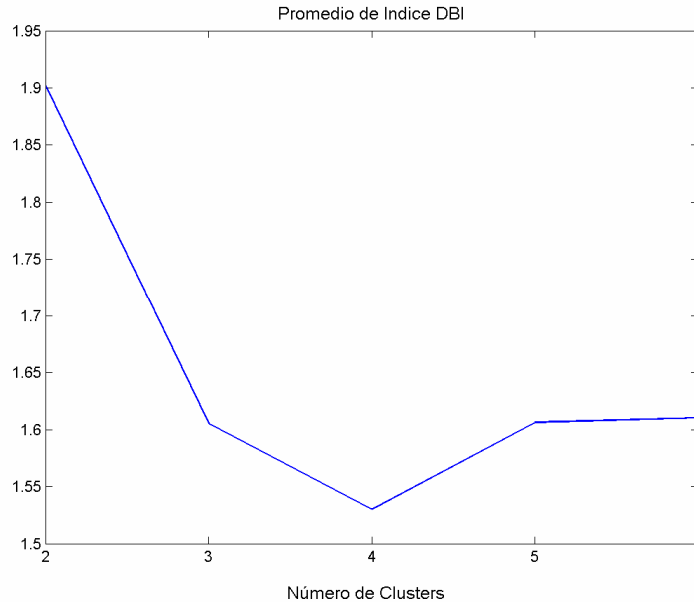


Figura 6.4: Promedio Mensual de Índice DBI.

Para realizar la clasificación por cluster que entregará los resultados definitivos hay que escoger la técnica que resulte más adecuada para los fines del estudio. Es por esto que la elección debe realizarse según lo expuesto en la tabla 3.1. Dada la experiencia internacional y la fortaleza de poseer una matriz de pertenencia con coeficientes continuos, es que se ha escogido la técnica *Fuzzy C-Means*.

A pesar que esta técnica también presenta una sensibilidad al punto de partida, esta se ve disminuida a medida que el factor m aumenta su valor. Además, para disminuir la dependencia del punto de partida y aumentar la confiabilidad será necesario realizar varios agrupamientos y escoger como definitivo el que tenga la función objetivo más pequeña, de la misma forma en que se hizo para determinar el número óptimo de clusters.

Como se explicó anteriormente, hay que definir el valor del factor m , ya que valores distintos entregan distintas soluciones. En la literatura se utiliza el factor $m=1,2$ [15] de forma arbitraria, ya que no se presenta ningún método para escoger un valor idóneo. Para este caso se hicieron pruebas para distintos valores, un valor que tiende a 1 da como resultado una técnica que pierde su categoría de difusa, mientras que valores mayores a 1,8 entregan como resultado curvas idénticas para cada cluster. Por ello, se ha elegido el valor $m=1,2$, siguiendo la experiencia previa presentada en la literatura técnica.

Para simplificar la lectura, en la figura 6.5 se presenta un resumen con los resultados obtenidos para cada mes. Los valores numéricos serán incluidos en el anexo A en forma de matrices.

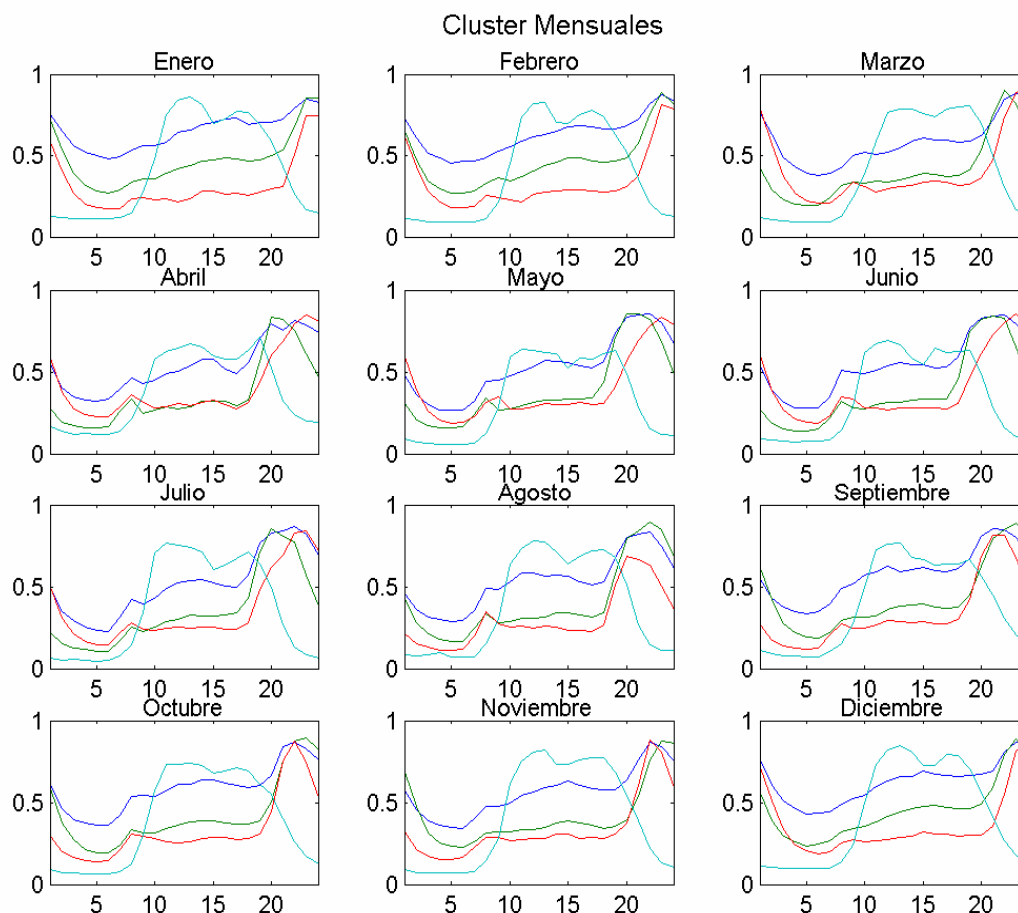


Figura 6.5: Perfiles típicos mensuales.

En la figura 6.6 se observa la existencia de cuatro perfiles de consumo bien definidos. Tres de ellos tienen la forma característica de un consumo residencial, identificadas por los colores azul, verde y rojo, que se diferencian en el uso medio de la potencia (o factor de carga) y en el instante en que ocurre la máxima demanda. En estas curvas se observa un desplazamiento horario estacional que presumiblemente está correlacionado con las horas de luz de cada día, ya que en los meses de invierno el aumento de la potencia ocurre más temprano que en los meses de verano. Otro resultado visible corresponde a la estacionalidad del uso medio de la potencia, ya que se observa que en los meses de verano las curvas tienden a ser más planas. Este fenómeno puede deberse en gran medida a que en los meses de invierno se utiliza más iluminación eléctrica y artefactos de calefacción.

El cuarto consumo típico, identificado con el color celeste, que también se repite a lo largo de los meses del año, corresponde a una curva característica comercial. Se observa que esta curva, a diferencia de las otras, no posee una estacionalidad horaria. Sin embargo, el fenómeno de la estacionalidad del uso medio de la potencia es inverso al que ocurre con las curvas residenciales, ya que en los meses de invierno la curva tiende a ser más plana que en los meses de verano. Este resultado puede ser provocado porque este tipo de clientes efectúa un mayor uso de equipos de refrigeración y/o aire acondicionado.

Sin embargo, para tener certeza de las aseveraciones planteadas en los párrafos anteriores, sería necesario conocer efectivamente cuál es el uso que cada cliente le da a la electricidad, información que podría obtenerse a través de una encuesta o cuestionario como la que se plantea en la figura 3.8.

En la figura 6.6 se han agrupado las curvas obtenidas para cada *cluster* en cuatro gráficos distintos, para observar cuál es el comportamiento estacional de cada una de ellas.

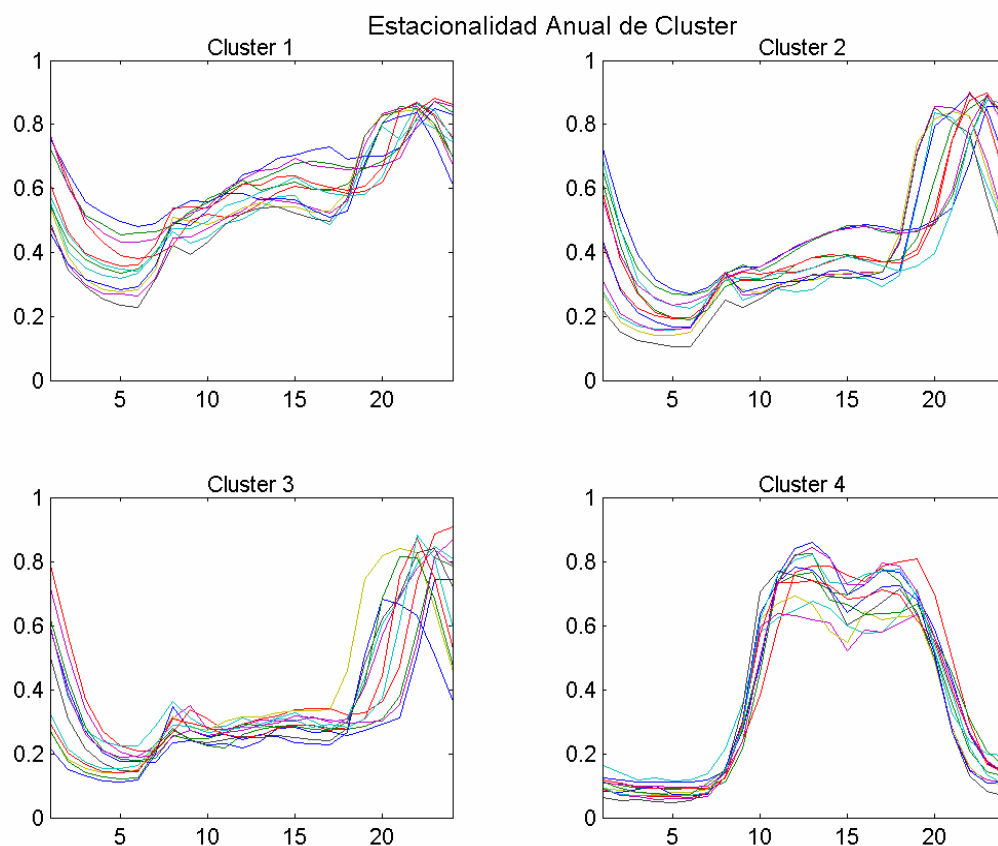


Figura 6.6: Movimiento estacional de cada cluster.

En los gráficos expuestos en la figura 6.6 se observa el movimiento estacional de cada *cluster*. Las curvas se mueven en una banda bastante acotada, aunque entre los *cluster* 2 y 3 puede observarse a simple vista que en algunos meses, principalmente en invierno, existe un traslape entre ellos, lo que provoca que las bandas antes mencionadas sean coincidentes en algunos valores.

A partir de esta definición de *clusters*, se pueden realizar variadas aplicaciones. Una de ellas, corresponde a la estimación de la curva agregada conociendo sólo la energía consumida por los clientes en el mes. Como el *cluster* representa una curva típica diaria, bastará ajustar la magnitud de esta curva de tal forma que la energía consumida corresponda al valor promedio diario del mes en estudio. Así, se obtendrán curvas agregadas para cada *cluster* y la suma de éstas entregará como resultado una curva general. En la figura 6.7 se muestra como ejemplo la estimación que se realiza para el mes de Julio, que corresponde al mes de máxima demanda.

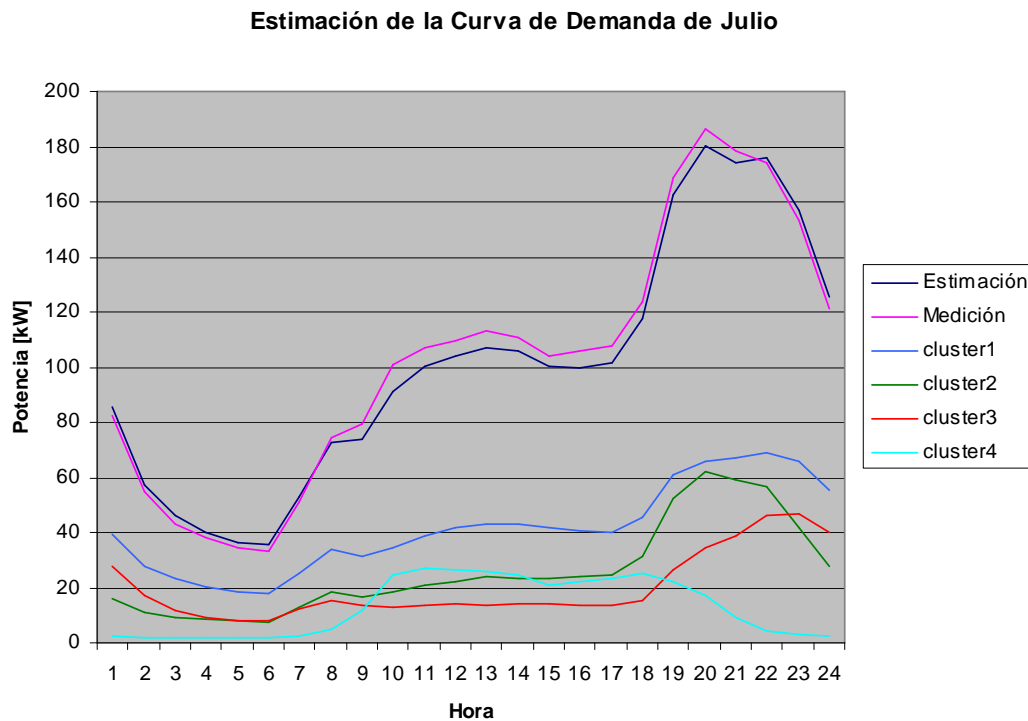


Figura 6.7: Estimación de la curva de demanda agregada de julio.

En la figura 6.7 se observa la participación de cada *cluster* en la curva agregada, de donde se concluye que ninguno de ellos es despreciable con respecto a otro. Sobre todo el *cluster* 4 que representa al consumo del tipo comercial, que a pesar de que son pocos clientes, su participación es bastante significativa a nivel agregado.

En la figura 6.8 se muestra el error de la estimación.

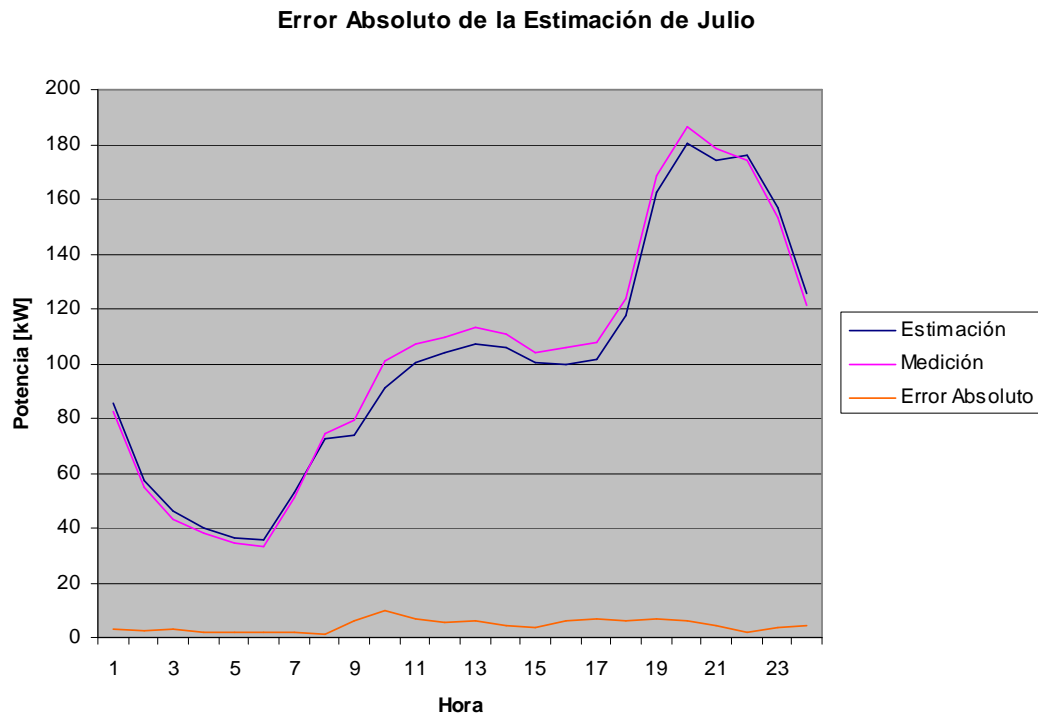


Figura 6.8: Error absoluto de la estimación.

En la figura 6.8 se muestra el error de la estimación, que en ningún punto alcanza a ser mayor a un 10%.

6.4.2 Clientes en forma agregada

El segundo análisis importante corresponde a analizar la estacionalidad de la curva agregada. La metodología incluye la separación original de la muestra en estratos, aunque para hacerla mas representativa se ha decidido unir los estratos 1 y 2 en un gran estrato. Así la estratificación utilizada es la que se indica en la tabla 6.2.

Indice j	Rango [kWh/mes]
1	0-150
2	151-300
3	>300

Tabla 6.2: Estratificación para análisis de curvas agregadas.

Para los tres estratos mencionados anteriormente se representarán las curvas agregadas de cada mes. En las figuras 6.9, 6.10 y 6.11 se muestran los resultados obtenidos.

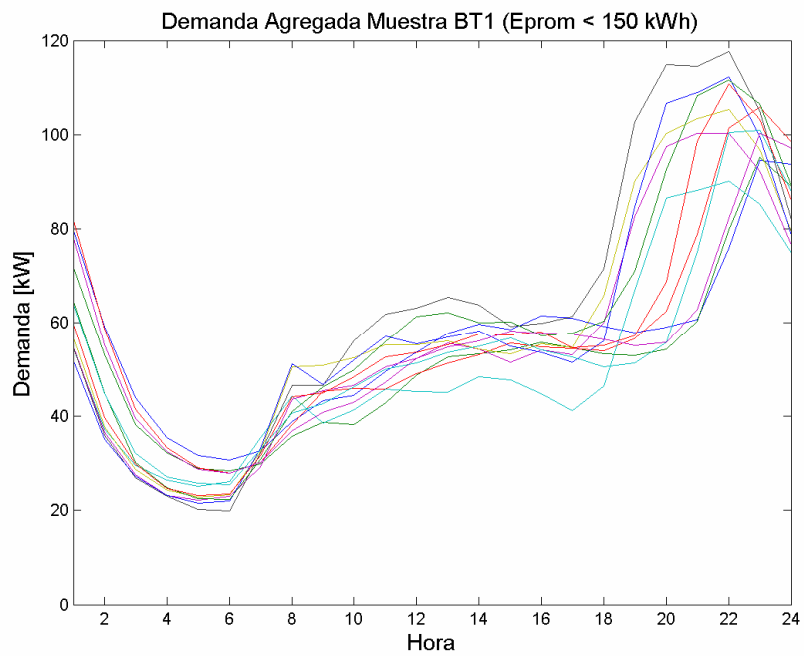


Figura 6.9: Demanda agregada mensual del estrato 1.

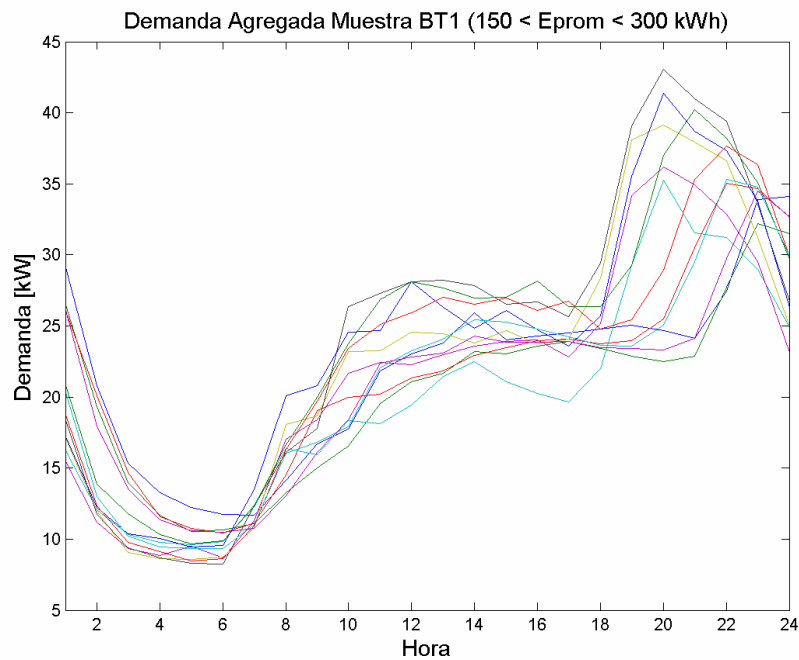


Figura 6.10: Demanda agregada mensual del estrato 2.

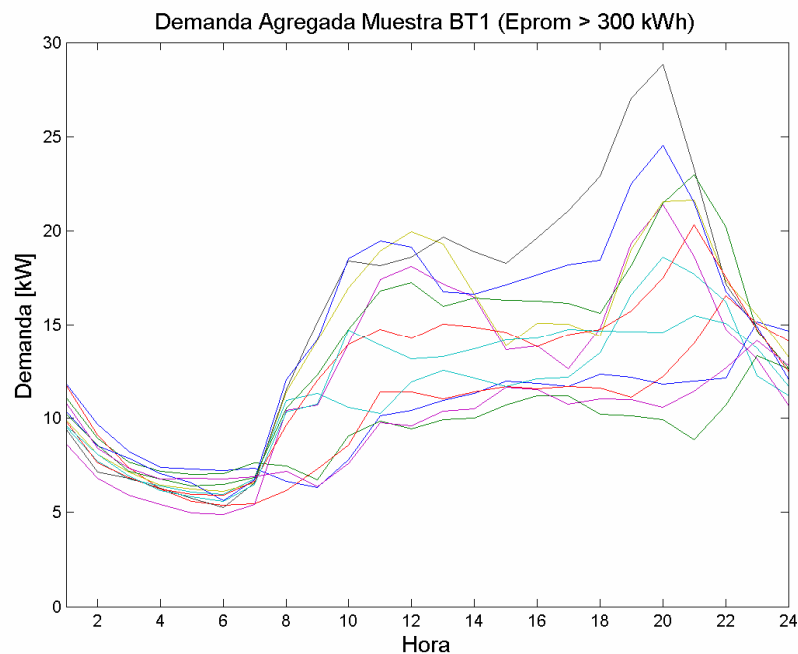


Figura 6.11: Demanda agregada mensual del estrato 3.

A partir de las figuras se observan dos fenómenos de manera inmediata. El primero es que el aumento de consumo en los meses de invierno es mucho más marcado para el estrato de mayor consumo en comparación con la variación que experimentan los estratos de consumo medio y pequeño, aunque para afirmarlo con mayor certeza será necesario realizar un análisis de curvas de demanda promedio por cliente de cada estrato, de tal forma de identificar el aumento porcentual de la demanda. El segundo fenómeno que se aprecia es que la máxima demanda para el estrato 1 ocurre, en promedio, dos horas más tarde que en los otros dos estratos. También se observa que en invierno la máxima demanda ocurre más temprano que en verano, lo que muestra la correlación existente entre este valor y la luminosidad de cada periodo del año.

Del mismo modo que en el análisis de *clustering*, se puede realizar una estimación de demanda utilizando estas curvas. Para este caso se puede estimar la demanda agregada BT-1, ya que podemos definir una curva promedio para cada estrato y luego proyectarla a la totalidad de clientes con esta opción tarifaria, como se realizó en el estudio previo realizado por Chilectra que se expuso en el capítulo 4.

Los análisis hechos en esta parte son del tipo cualitativo, por lo que en el capítulo siguiente se harán algunos análisis cuantitativos que servirán para darle más formalidad al estudio.

6.5 Diseño de un modelo matemático para simular curvas

Una aplicación que se puede desarrollar con los resultados de la parte 6.4.1 es un modelo matemático que permita realizar simulaciones de curvas de consumo que se ajusten a un perfil típico de consumo. El diseño se plantea como un proceso estocástico que genera puntos horarios para cada curva típica. La ecuación (6.1) muestra la expresión que rige el proceso es [23]:

$$X(t)_{i,m} = \mu(t)_{i,m} + \sigma(t)_{i,m} W(t) \quad (6.1)$$

Donde,

$\mu(t)_{i,m}$: Centro del cluster i del mes m en el periodo t

$\sigma(t)_{i,m}$: Desviación estándar del cluster i del mes m en el periodo t

$W(t)$: Movimiento Browniano

A partir de esta ecuación se obtiene una serie de curvas simuladas, que pueden utilizarse para desarrollar distintas aplicaciones, principalmente realizar estimaciones de carga cuando sólo se conoce la energía consumida. En la figura 6.12 se muestra a modo de ejemplo una comparación entre las curvas reales de un cluster comercial frente a las simulaciones obtenidas con el modelo.

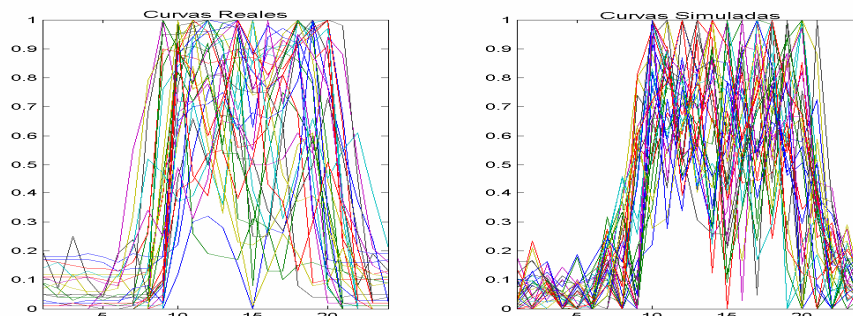


Figura 6.12: Curvas reales y simuladas tipo comercial.

A pesar que a simple vista las curvas se parecen es necesario validar el modelo mediante algún método estadístico válido. Es por ello que se utiliza el método de Kolmogorov-Smirnov, que compara dos muestras y decide si siguen la misma distribución. Un método para utilizar este test para muestras multivariantes se indica en la referencia [26], ya que el desarrollo matemático escapa de los objetivos de la memoria. A modo de resumen, lo que hace el test es realizar múltiples simulaciones (al menos 5.000) y comparar cada una de ellas con las curvas reales. Para cada comparación se decidirá si la distribución es la misma. Luego se calcula el porcentaje de rechazo frente al total de simulaciones realizadas. En la referencia citada se recomienda elegir aceptable un porcentaje de rechazo de un 5%, aunque se puede dar una mayor libertad. Para este

caso se decidió realizar 10.000 simulaciones y aceptar el modelo si es que el grado de rechazo es menor a un 10%.

Los resultados que entrega el test son cercanos a un rechazo de un 8% como valor promedio, salvo para unos pocos casos que se obtienen valores de rechazo elevados, por lo que el modelo es aceptado como válido. Una tabla resumen del test para cada cluster se presenta en el Anexo B.

7 Análisis de supuestos del cobro por energía adicional de invierno

La tarifa BT-1 corresponde a la opción tarifaria utilizada por la mayor cantidad de clientes del país, destinada principalmente a clientes residenciales. Esta tarifa posee dos modalidades: La tarifa BT-1a, que se aplica a los clientes abastecidos por empresas cuya demanda máxima ocurre en los meses en que se han definido horas punta, y la tarifa BT-1b, que se aplica a los clientes abastecidos por empresas cuya demanda máxima ocurre en los meses en que no se han definido horas de punta. Esta diferenciación se hace debido a que los mayores consumos ocurren en distintos periodos para cada caso, en invierno para el caso BT-1a y en verano para el caso BT-1b. Actualmente la única empresa que ofrece la modalidad BT-1b es la empresa Litoral S.A.

El análisis que se realizará en este capítulo está orientado a la tarifa BT-1a, que corresponde a la modalidad que se aplica en el caso de Chilectra S.A.

7.1 Estructura de la tarifa BT-1a

El sistema tarifario chileno considera la aplicación de precios para la energía y precios para la potencia, de suerte que la factura mensual es calculada a través de la expresión genérica representada en la ecuación (7.1).

$$\text{Facturacion} = Pe \cdot E + Pp \cdot P + \text{cargo fijo} \quad (7.1)$$

Donde,

Pe : Precio de la Energía \$/kWh

E : Energía kWh

Pp : Precio de la Potencia \$/kW

P : Potencia kW

cargo fijo : Considera los costos administración

En todas las tarifas existentes en el sistema chileno la energía es medida. En todas las tarifas, con la excepción de la BT-1, la potencia de facturación se determina a partir de una estimación o de una medición. No obstante, en la tarifa BT-1a se emplea el mismo esquema de facturación señalado en la expresión 7.1 y en donde la potencia se deriva a partir de la energía medida, haciendo uso de las relaciones que la ingeniería eléctrica ha establecido entre la energía y la potencia. Esta relación se muestra en la ecuación (7.2).

$$E = D_{\max} \cdot fc \cdot T \quad (7.2)$$

Donde,

E : Energía consumida

D_{\max} : Máxima Demanda

fc : Factor de carga

T : Periodo

Luego, despejando el valor D_{\max} de la ecuación se obtiene la expresión (7.3).

$$D_{\max} = \frac{E}{T \cdot fc} = \frac{E}{H} \quad (7.3)$$

Donde H son las Horas de Uso, que corresponde al parámetro que define la relación existente entre la Energía y la Potencia.

Así, en términos genéricos, la facturación de la tarifa BT-1a se reduce a la expresión (7.4).

$$Facturacion = \left(Pe + \frac{Pp}{H} \right) \cdot E + \text{cargo fijo} \quad (7.4)$$

Donde,

$\left(Pe + \frac{Pp}{H} \right)$ es un precio expresado en [\$/kWh], y que por el hecho de involucrar tanto a la energía como potencia, se le conoce también como un precio monómico.

La estructura de la tarifa BT-1a incorpora además un cargo por energía adicional de invierno. Este cargo surge de considerar que hay un grupo de clientes que tienen una demanda de potencia estable a lo largo del año y otros que suben la demanda de potencia en horas de punta. La tarificación persigue entonces que sólo los que contribuyen a este “sobrepeak” en el invierno paguen dicha contribución. Dado que en esta tarifa no se conoce el comportamiento de la potencia de cada cliente, la estimación de quienes contribuyen a este “sobrepeak” se hace a través de la energía.

El consumo de energía adicional de invierno se determina utilizando un límite de consumo mensual para los meses de invierno, calculado a partir del promedio de los consumos en los meses de verano. De esta forma, todo consumo que exceda ese límite, y que además sea

superior a 250 kWh, calificará como energía adicional de invierno. Así, la expresión (7.3) debe ser adaptada a la expresión (7.5).

$$D_{\max} = \frac{\sum_{i=1}^{12} E_{basei}}{NHUB} + \frac{12}{5} \frac{\sum_{i=5}^9 E_{adicionali}}{NHUI} \quad (7.5)$$

Donde,

D_{\max} : Maxima Demanda

E_{basei} : Energía Base total consumida en el mes i

$E_{adicionali}$: Energía Adicional de Invierno consumida en el mes i

$NHUB$: Numero de Horas de Uso para el cálculo de la Potencia Base

$NHUI$: Numero de Horas de Uso para el cálculo de la potencia de Invierno

Explicada la estructura general de la tarifa BT-1a, a continuación se presentan los cargos detallados que considera la tarifa.

La tarifa BT-1a se aplica a los clientes alimentados en baja tensión cuya potencia conectada sea inferior a 10 kW y a aquellos clientes que instalen un limitador de potencia para cumplir esta condición. Esta tarifa comprende los siguientes cargos [2]:

- a) Cargo fijo mensual
Es independiente del consumo y se aplicará incluso si este es nulo.
- b) Cargo único por uso del sistema troncal
Se determina en proporción al consumo de energía.
- c) Cargo por energía base
Se determina multiplicando los kWh de consumo base por su precio unitario calculado a través de la expresión (7.6).

$$PEBT \cdot PEAT \cdot Pe + \frac{PPBT \cdot PPAT \cdot Pp}{NHUNI} + \frac{CDBT}{NHUDI} \quad (7.6)$$

- d) Cargo por energía adicional de invierno
Este cargo se aplicará en cada mes del período 1° de mayo – 30 de septiembre, en que el consumo del cliente exceda 250 kWh/mes, a cada kWh consumido al mes en exceso del límite de invierno del cliente.

El límite de invierno será igual al mayor valor que resulte de comparar 200 kWh, con un séptimo de la totalidad de la energía consumida en el período 1° de octubre – 30 de abril inmediatamente anterior, incrementada en un 20%.

Se determina multiplicando los kWh de consumo adicional de invierno por su precio unitario calculado a través de la expresión (7.7).

$$PEBT \cdot PEAT \cdot Pe + 2,4 \cdot \frac{PPBT \cdot PPAT \cdot Pp}{NHUNI} + 2,4 \cdot \frac{CDBT}{NHUDI} \quad (7.7)$$

Donde,

Pe : Precio de nudo de energía a nivel de distribución

Pp : Precio de nudo de potencia a nivel de distribución

CDBT : Costo de distribución en baja tensión

PEBT : Factor de expansión de pérdidas de energía en baja tensión

PPBT : Factor de expansión de pérdidas de potencia en baja tensión

PEAT : Factor de expansión de pérdidas de energía en alta tensión

PPAT : Factor de expansión de pérdidas de potencia en alta tensión

NHUNB : Número de horas de uso para el cálculo de la potencia base coincidente con la punta del sistema

NHUDB : Número de horas de uso para el cálculo de la potencia base coincidente con la punta del sistema de distribución

NHUNI : Número de horas de uso para el cálculo de la potencia adicional de invierno coincidente con la punta del sistema

NHUDI : Número de horas de uso para el cálculo de la potencia adicional de invierno coincidente con la punta del sistema de distribución

La ecuación (7.7) que define el precio de la energía adicional de invierno, difiere por una parte en el valor de las horas de uso y por otra en un factor de ponderación que aplica a la componente de la potencia y de los costos de distribución frente al precio de la energía base representado en la ecuación (7.6). Este factor que es igual a 2,4 corresponde a la razón que existe entre los 12 meses del año, frente a los 5 meses en que se definen horas de punta donde se realiza este cobro adicional.

7.2 Supuestos de la tarifa BT-1a

De acuerdo a lo señalado por Bernstein en [27], el “sobrepeak” de invierno debe ser pagado por los clientes de mayores consumos. Actualmente, y para efectos de la aplicación de esta tarifa, se entiende por un consumo mayor aquel que supera los 250 kWh mensuales. El supuesto es entonces que los clientes de bajos consumos tienen un comportamiento estable en la potencia a lo largo del año y que son los clientes de mayores consumos los responsables de producir el “sobrepeak” de invierno.

En la figura 7.1 se muestra un esquema con la idea señalada por Bernstein para los clientes de mayor consumo cuando se desarrolló esta tarifa.

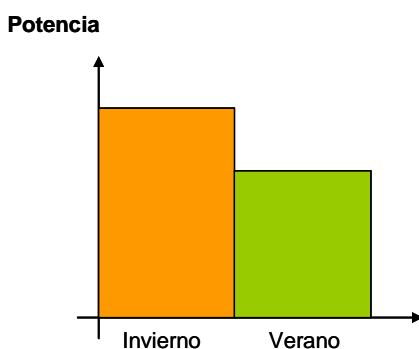


Figura 7.1: Esquema planteado por Bernstein.

Este criterio fue establecido en 1983, no siendo hasta la fecha ni corroborado ni refutado. No obstante, hay especialistas que ponen en duda este supuesto, ya que el comportamiento de los clientes pudiera ser exactamente lo contrario. Actualmente, la información disponible es la estacionalidad de la energía. Pero no debe olvidarse que el fenómeno que se pretende tarifificar es el de la potencia.

Con la información de mediciones disponibles, en el punto siguiente se analizará por estrato de consumos la estacionalidad de la demanda por potencia para corroborar o refutar la idea recién planteada.

7.3 Análisis de datos

Para realizar este análisis se ha decidido tomar como base sólo a los clientes que están clasificados como residenciales dentro del grupo válido, que corresponden a 374 clientes del total de 398 mediciones. El motivo de esto es poder reconocer algunos patrones comunes de los clientes con consumo energético similar y eliminar posibles curvas atípicas, por ejemplo un cliente industrial de gran consumo, que pudieran afectar la conclusión del análisis.

Una hipótesis fuerte que se considerará para el desarrollo de este capítulo es que no existe un crecimiento en el consumo anual de energía de los clientes de la muestra. Esto con el fin de poder comparar los valores máximos de potencia mensual. Del mismo modo, se utilizarán los valores de consumo de los meses de verano posteriores al periodo de invierno para calcular el la energía adicional de invierno, ya que no se cuenta con más información. En la tabla 7.1 se muestra el detalle con la totalidad de energía consumida por los clientes seleccionados de la muestra.

Mes	Energía Total [kWh]	Energía Base [kWh]	Energía Adicional de Invierno [kWh]
enero	57.899	57.899	0
febrero	51.586	51.586	0
marzo	60.154	60.154	0
abril	54.265	54.265	0
mayo	59.841	58.228	1.613
junio	61.439	57.947	3.492
julio	67.973	60.333	7.640
agosto	66.897	60.729	6.168
septiembre	63.380	59.593	3.787
octubre	60.954	60.954	0
noviembre	56.398	56.398	0
diciembre	58.007	58.007	0
Total	718.793	696.093	22.700

Tabla 7.1: Total de Energía Consumida.

En la tabla 7.1 se observa que los consumos mas elevados corresponden a los meses de julio y agosto. A partir de estos resultados y utilizando la ecuación 7.5 que define la máxima demanda se obtiene como resultado una potencia de 164,6 kW.

El paso siguiente corresponde a preguntarse si son los clientes de mayores consumos, es decir, los consumos mayores a 250 kWh/mes los que poseen la mayor estacionalidad y por ende deben pagar por la potencia de invierno, sujeto a la restricción de que el consumo de invierno también sea mayor al límite definido por los consumos en los meses de verano. Para este análisis se definieron tres estratos por intervalo de consumo promedio mensual según el criterio que se muestra en la tabla 7.2, ya que los intervalos definidos en la muestra de 1999 no son los óptimos para esta aplicación.

Estrato	Energía Min [kWh]	Energía Max [kWh]
1	0	100
2	101	250
3	251	Mayores

Tabla 7.2: Estratos de consumo.

La metodología considera utilizar valores de potencia y energía en pu¹¹ en base propia, que será el resultado del promedio de los meses de verano para cada estrato, ya que así se pueden hacer comparaciones cuantitativas del aumento de demanda en los meses de invierno.

El estrato 1 corresponde a los clientes que poseen un consumo promedio menor a 100 kWh mensual. La curva de demanda para cada uno de los meses se muestra en la figura 7.2. Estos clientes poseen un comportamiento estacional, de consumo de energía y de máxima demanda en horario punta, como el que se indica en la figura 7.3, en que el valor 1,00 del eje de la ordenada representa el valor promedio de las variables en el periodo de verano (1° de octubre a 30 de abril).

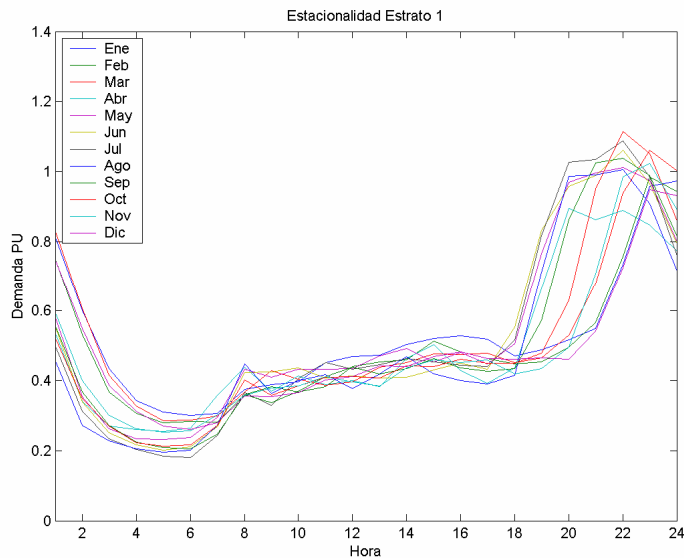


Figura 7.2: Curvas de demanda mensuales del estrato 1.

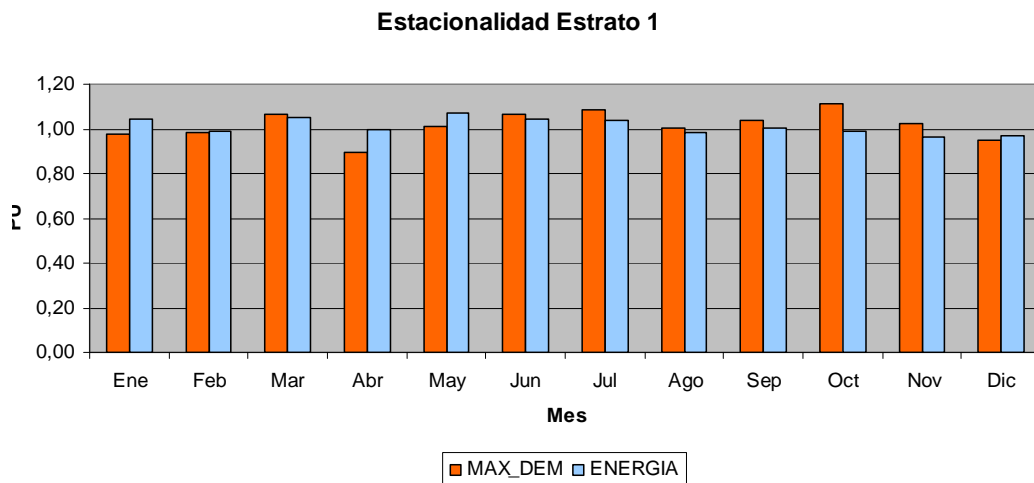


Figura 7.3: Estacionalidad estrato 1.

¹¹ Por unidad

De la figura 7.2 se concluye que las curvas de demanda tienen poca movilidad mes a mes. Además en la figura 7.3 se observa que el consumo de energía presenta baja estacionalidad y la máxima demanda, a pesar de que su estacionalidad es un poco mas marcada, también se puede catalogar como baja, ya que se mueve en una banda de 0,89 y 1,11 pu. Esta variación en promedio según los periodos invierno y verano se representa en la figura 7.4, en que el aumento promedio de invierno es de un 4%.

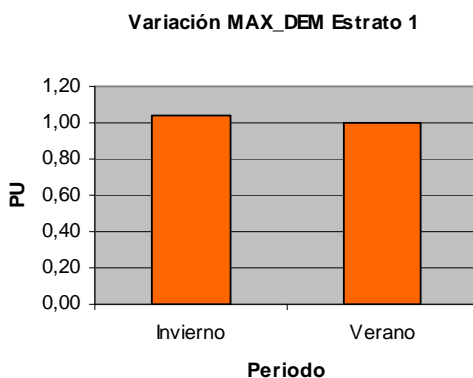


Figura 7.4: Variación de la máxima demanda entre invierno y verano estrato 1.

El estrato 2 corresponde a los clientes que poseen consumos promedio entre 101 y 250 kWh. En la figura 7.5 se muestran las curvas de demanda para este estrato. La estacionalidad que presentan estos clientes se muestra en la figura 7.6.

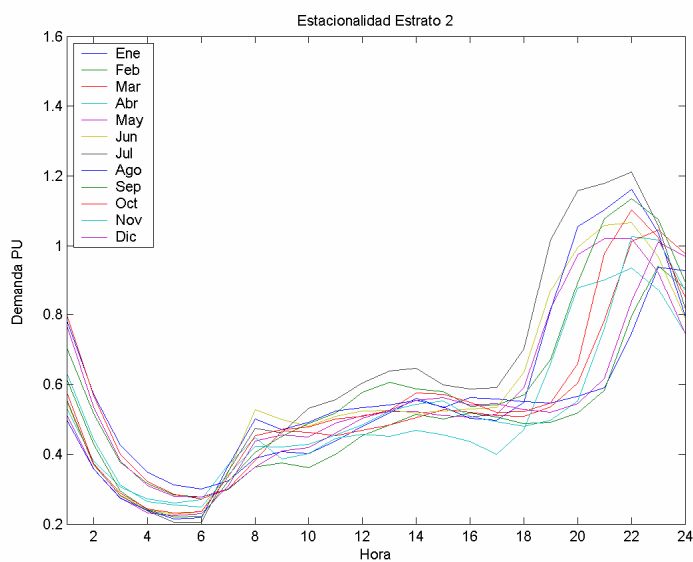


Figura 7.5: Curvas de demanda mensuales del estrato 2.

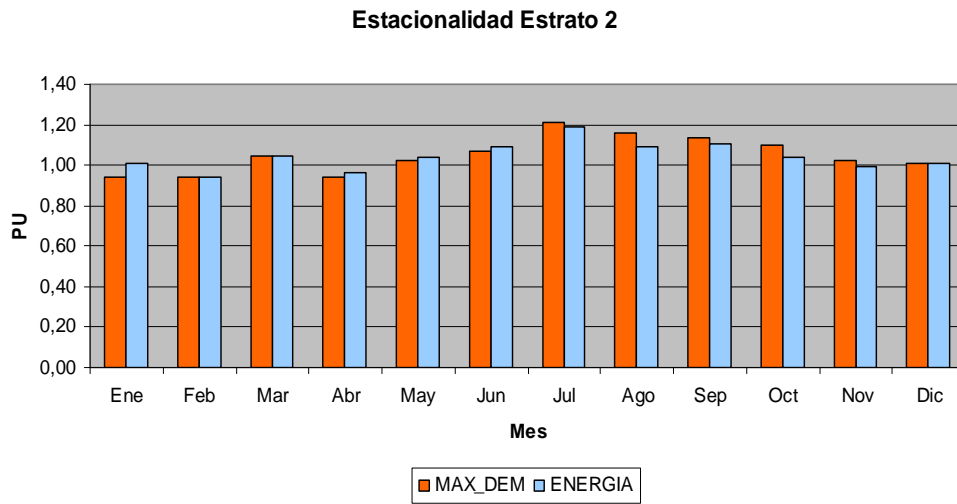


Figura 7.6: Estacionalidad estrato 2.

En la figura 7.6 se puede apreciar que las curvas de demanda del estrato 2 poseen movimiento mas marcado que el estrato 1, mientras que la estacionalidad de la potencia máxima está acotada por los valores 0,94 y 1,21 pu, alcanzando su valor máximo en el mes de máxima demanda agregada que corresponde a Julio. Esta variación en promedio según los periodos invierno y verano se representa en la figura 7.7, en que el aumento promedio de invierno es de un 12%.

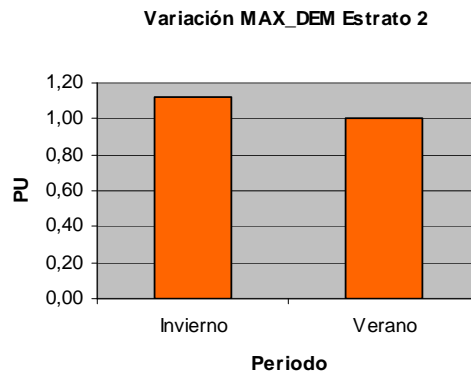


Figura 7.7: Variación de la máxima demanda entre invierno y verano estrato 2.

El estrato 3 corresponde a los clientes que consumen en promedio más de 250 kWh al mes. En la figura 7.8 se muestran las curvas de demanda para este estrato. La estacionalidad que presentan estos clientes se muestra en la figura 7.9.

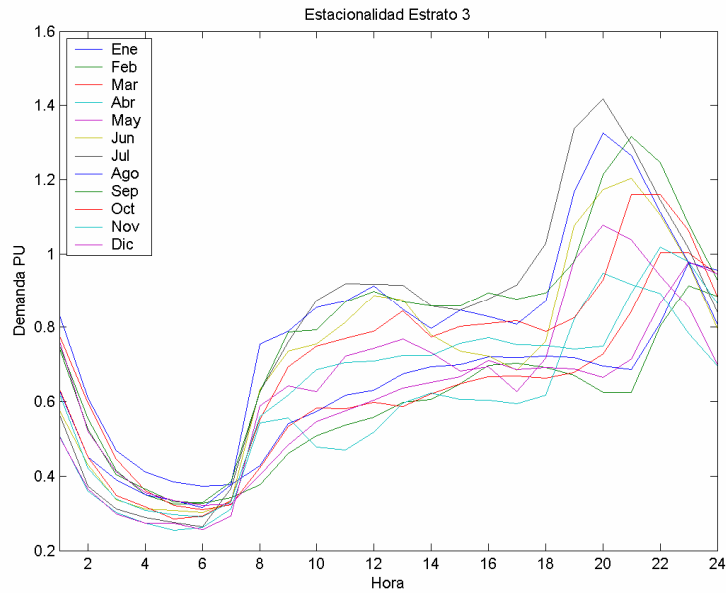


Figura 7.8: Curvas de demanda mensuales del estrato 3.

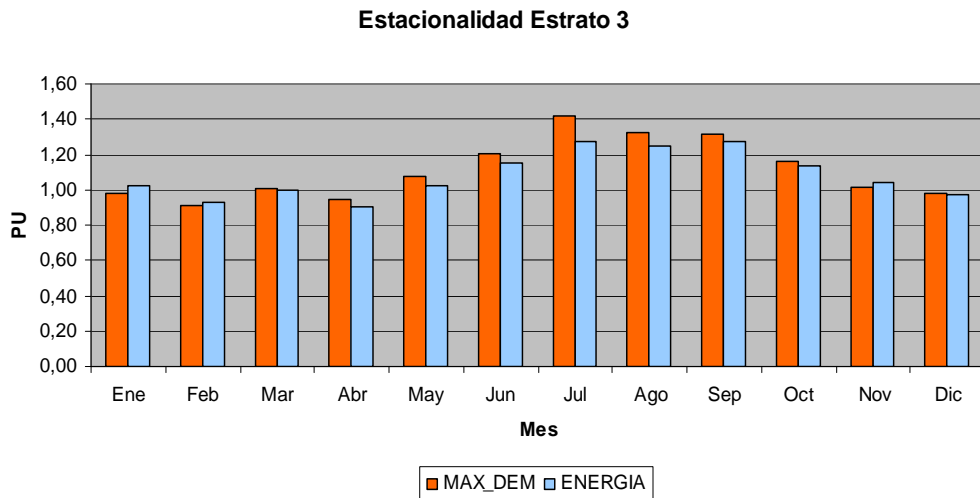


Figura 7.9: Estacionalidad estrato 3.

En la figura 7.9 se muestra que los clientes de este estrato poseen una estacionalidad de la potencia mayor que la de los otros estratos, ya que los clientes aumentan considerablemente su demanda máxima en los meses de invierno, principalmente en el mes de julio. Incluso en el mes de octubre se observa un crecimiento cercano al 20% con respecto al promedio de los meses de verano. La banda en la que se mueve la máxima demanda tiene como cota a los valores 0,91 y 1,42 pu. La variación promedio en invierno con respecto al verano se observa en la figura 7.10, en que el aumento en invierno es de un 27%.

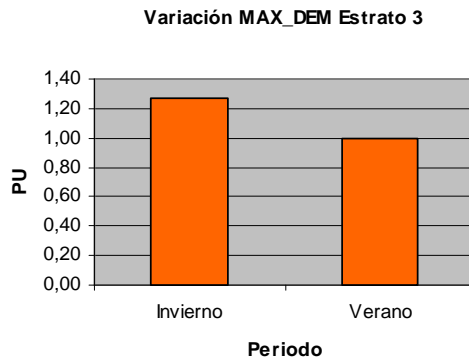


Figura 7.10: Variación de la máxima demanda entre invierno y verano estrato 3.

A partir de los análisis realizados anteriormente se puede sostener que Bernstein estaba en lo correcto al señalar que son los clientes de mayor consumo los que poseen la mayor estacionalidad, y que por ende deben pagar por el “sobrepico”, siendo julio, agosto y septiembre los meses que presentan la más alta demanda con respecto a los meses de verano. Sin embargo, debe tenerse muy presente que este estudio está realizado en base a unas mediciones del año 1999, por tanto no se puede aseverar que este resultado siga siendo efectivo a la fecha, principalmente porque los hábitos de consumo pueden haberse modificado.

Como resultado, y a modo de resumen, en la figura 7.11 se presenta una estimación sobre la incidencia de cada estrato en la curva agregada de demanda. Los estratos 1 y 2 son los responsables, principalmente, del aumento de la demanda en horas punta frente al promedio del día, ya que poseen un factor de carga muy bajo. Sin embargo, son los clientes del estrato 3 los responsables de las variaciones estacionales de la curva agregada, específicamente de los aumentos de potencia en invierno.

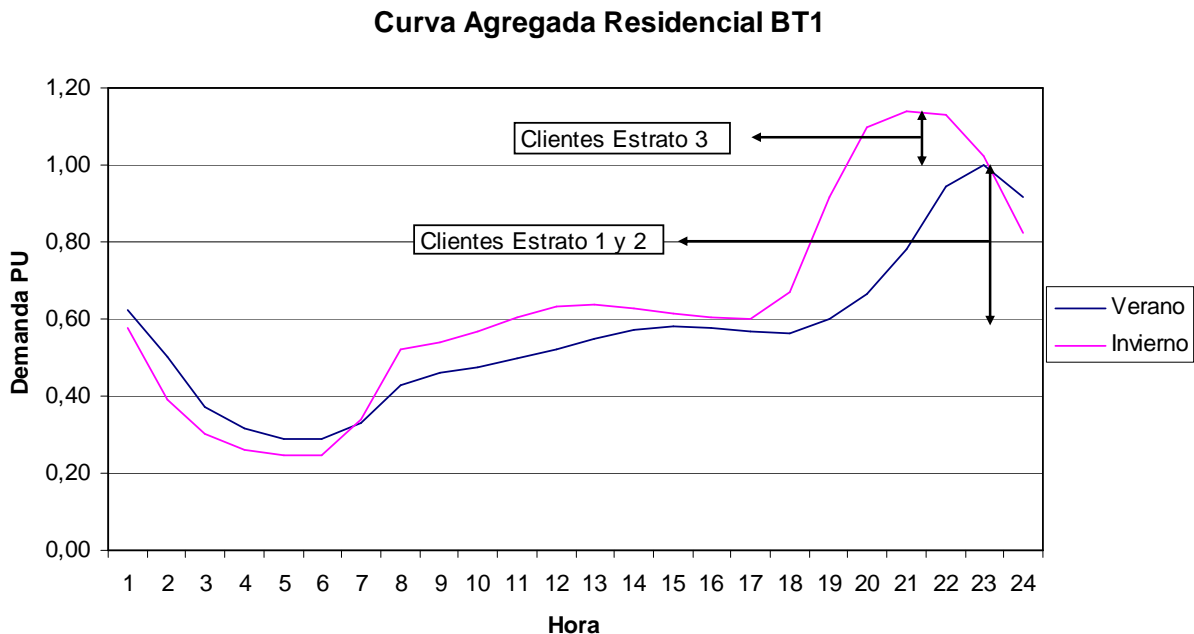


Figura 7.11: Incidencia de cada estrato sobre la curva agregada.

7.4 Análisis cruzado entre clusters y estratos

Otro análisis que se puede realizar con la información y los resultados obtenidos en este estudio, corresponde a la relación que existe entre los estratos de consumo y el comportamiento por curvas típicas de los clientes. A partir de este resultado se pueden encontrar variadas aplicaciones, entre las cuales destaca la tarificación a través de perfiles típicos. En las figuras 7.12, 7.13 y 7.14 se muestra la caracterización temporal de perfiles para cada estrato definido al comienzo de este capítulo.

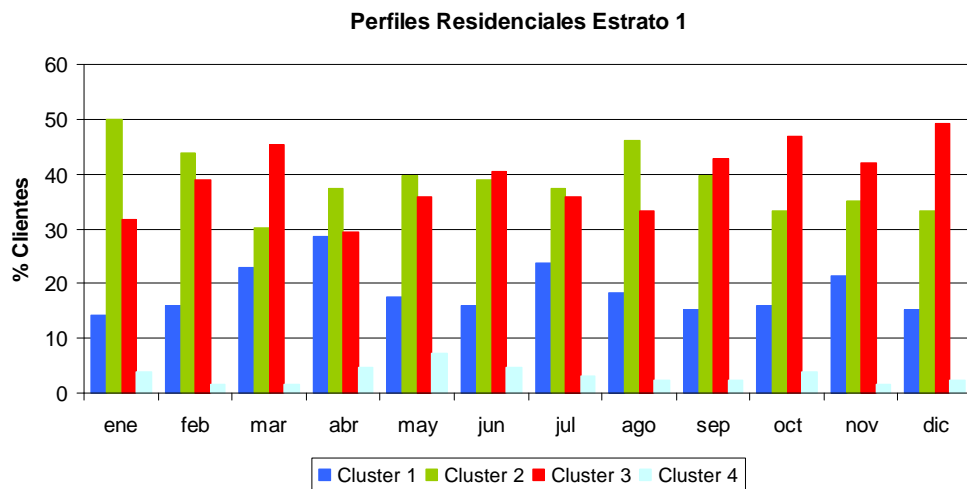


Figura 7.12: Caracterización por perfiles típicos del Estrato 1

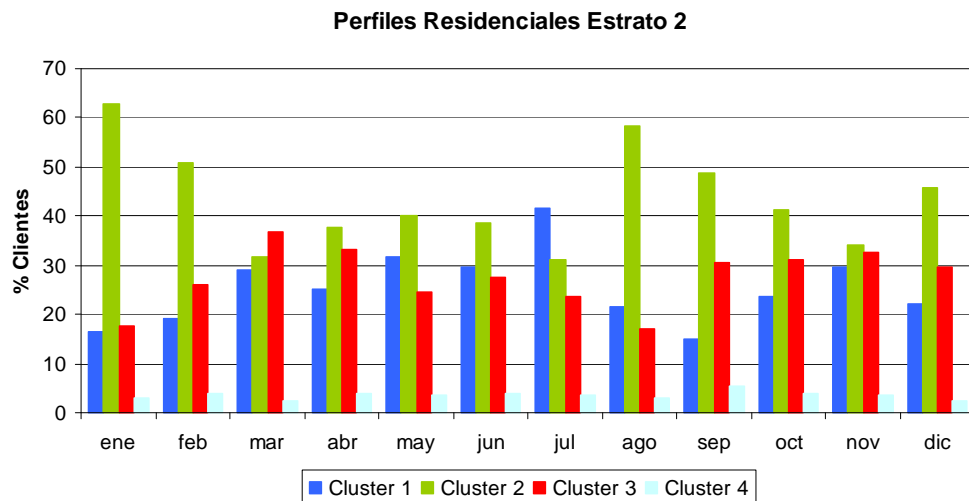


Figura 7.13: Caracterización por perfiles típicos del Estrato 2

Perfiles Residenciales Estrato 3

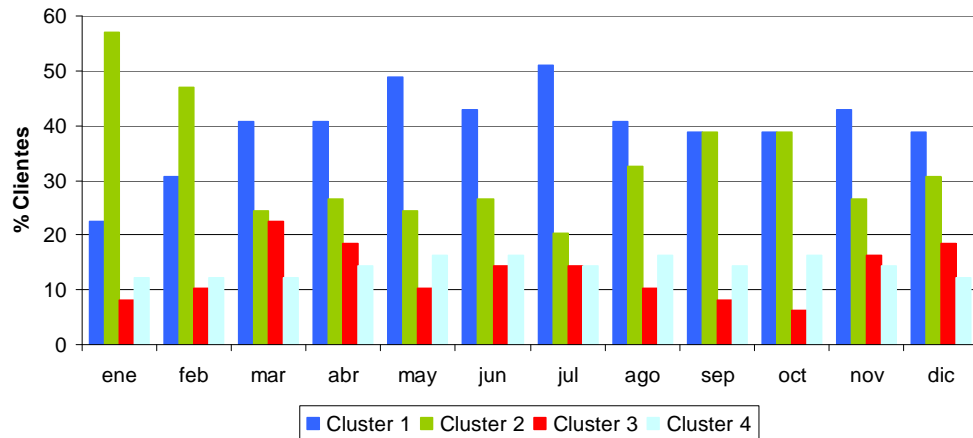


Figura 7.14: Caracterización por perfiles típicos del Estrato 3

De las figuras anteriores se puede observar que para cada estrato existe una curva distinta que predomina por sobre las otras. Así, en el estrato 1 los *clusters* predominantes son el 2 y 3 (curva verde y roja respectivamente). Para simplificar la lectura de los gráficos, en la tabla 7.3 se presenta un resumen con los resultados obtenidos como participación porcentual de cada curva en cada estrato.

Mes	Estrato 1				Estrato 2				Estrato 3			
	1	2	3	4	1	2	3	4	1	2	3	4
ene	14,3	50,0	31,7	4,0	16,6	62,8	17,6	3,0	22,4	57,1	8,2	12,2
feb	15,9	43,7	38,9	1,6	19,1	50,8	26,1	4,0	30,6	46,9	10,2	12,2
mar	23,0	30,2	45,2	1,6	29,1	31,7	36,7	2,5	40,8	24,5	22,4	12,2
abr	28,6	37,3	29,4	4,8	25,1	37,7	33,2	4,0	40,8	26,5	18,4	14,3
may	17,5	39,7	35,7	7,1	31,7	40,2	24,6	3,5	49,0	24,5	10,2	16,3
jun	15,9	38,9	40,5	4,8	29,6	38,7	27,6	4,0	42,9	26,5	14,3	16,3
jul	23,8	37,3	35,7	3,2	41,7	31,2	23,6	3,5	51,0	20,4	14,3	14,3
ago	18,3	46,0	33,3	2,4	21,6	58,3	17,1	3,0	40,8	32,7	10,2	16,3
sep	15,1	39,7	42,9	2,4	15,1	48,7	30,7	5,5	38,8	38,8	8,2	14,3
oct	15,9	33,3	46,8	4,0	23,6	41,2	31,2	4,0	38,8	38,8	6,1	16,3
nov	21,4	34,9	42,1	1,6	29,6	34,2	32,7	3,5	42,9	26,5	16,3	14,3
dic	15,1	33,3	49,2	2,4	22,1	45,7	29,6	2,5	38,8	30,6	18,4	12,2
Promedio	18,7	38,7	39,3	3,3	25,4	43,4	27,6	3,6	39,8	32,8	13,1	14,3

Tabla 7.3: Porcentaje de participación de cada cluster en cada estrato.

En un estudio anterior [12] se asignó cada perfil típico a clases económicas según su forma. De este modo se dedujo que el perfil 1 correspondía a la clase alta, el perfil 2 a la clase media y el perfil 3 a la clase baja, mientras que el perfil 4 correspondía a clientes comerciales. De la tabla 7.3 se deduce que los clientes de bajo consumo tienen un comportamiento, mayoritariamente, como los perfiles 2 y 3. Para el segundo estrato de consumo el perfil dominante es el 2. En el estrato de consumo mas alto el perfil dominante es el 1, lo que resume que la tesis planteada de las clases económicas puede ser certera, sin embargo, para poder

asegurarle sería necesario conocer la exacta correlación que existe entre la clase socioeconómica, el consumo energético mensual y las curvas típicas, información que podría obtenerse realizando un estudio de mediciones con alguna estratificación acorde a este requerimiento.

8 Conclusiones y desafíos futuros

El análisis y la estimación de la curva de carga de los clientes representan nuevas alternativas de desarrollo de las compañías eléctricas. Con la entrada de nuevas tecnologías de información y de equipos digitales sofisticados se abre un sinnúmero de posibilidades para realizar estudios que entreguen las conclusiones necesarias que generen valor agregado. Las aplicaciones de *load research* expuestas en el capítulo 3 entregan una visión sobre las distintas alternativas que se pueden implementar a partir de mediciones de curvas de carga de los distintos grupos de clientes. Destacan aplicaciones que pueden tener utilidad en diversas áreas, desde las áreas más técnicas hasta las comerciales, como:

- Diseño de tarifas e introducción de flexibilidad tarifaria.
- Planificación
- Operación y mantenimiento
- Marketing
- Manejo de carga (*load management*)

En el marco de las campañas de medición, el presente trabajo muestra la importancia que adquieren las técnicas de muestreo estadístico para el desarrollo de aplicaciones en grupos masivos de clientes que no poseen medidores horarios de demanda, como es el caso de los clientes BT-1, y propone tres alternativas distintas, ya que es necesario tener la certeza de que se obtiene una representatividad dentro de los rangos aceptables definidos con anterioridad. Posteriormente se presenta una metodología para realizar una minería de datos a un conjunto de mediciones horarias de potencia eléctrica y luego se detalla una serie de alternativas de *clustering* para agrupar datos con características similares.

La aplicación desarrollada en Chilectra considera una base de muestras de 1.050 mediciones horarias que se obtuvieron en una campaña de medición efectuada entre abril de 1999 y marzo de 2000. El primer análisis realizado consiste en la agrupación de consumos mediante técnicas de *clustering*, que arrojó como resultado la existencia de cuatro consumos típicos en la tarifa BT-1, donde destaca la existencia de una curva del tipo comercial dentro de una población mayoritariamente residencial. Sin embargo, estas curvas poseen una estacionalidad en cuanto a su forma observándose una tendencia a disminuir el factor de carga (o elevar la demanda en horario punta) en los meses de invierno. Además se detectó un traslape en las bandas de movimiento estacional de los *cluster* 2 y 3. Este análisis arroja las siguientes conclusiones:

- Junto a la campaña de medición es conveniente realizar una encuesta para identificar el uso que se le da a la energía, por ejemplo a través del equipamiento electrodoméstico. Así, será posible a través de un análisis estadístico encontrar la

correlación existente entre la encuesta y los perfiles típicos, que servirá para la asignación de los clientes no medidos a un perfil de consumo.

- La clasificación de consumos debe hacerse en forma dinámica, ya que de esta forma se podría implementar de manera mas apropiada un sistema tarifario basado en perfiles típicos de consumo.

El segundo análisis consistió en estudiar el comportamiento estacional de la demanda a nivel agregado. Para ello se utilizó una estratificación de la muestra, utilizando como variable el consumo promedio mensual de cada cliente. Este análisis tenía la motivación de verificar el grado de cumplimiento de los supuestos que fueron planteados hace casi 25 años y que se relacionan con el cobro por energía adicional de invierno. Para el desarrollo de este estudio se hizo un filtrado de datos y se tomaron sólo curvas residenciales, ya que se desea conocer el comportamiento para este tipo de clientes. Así se detectó una estacionalidad tanto a nivel de magnitud de la máxima demanda como de horario de ocurrencia, en que los clientes de menor consumo presentan un comportamiento mucho más estático que los de consumo elevado. Por ende se concluye que:

- Los clientes de menor consumo poseen un aumento en su demanda en invierno muy pequeña al ser comparada con el aumento que presentan los clientes de mayor consumo, esto indica que efectivamente son los clientes de mayor consumo los mayores responsables del aumento de la potencia de invierno.
- Las bases de la estructura tarifaria BT-1 están en lo correcto al dar una señal de aumento de precio en los meses en que sea han definido horarios punta, ya que se cobra a los clientes que efectivamente aumentan la demanda. Sin embargo, un análisis futuro podría ser revisar que este objetivo se cumpla a cabalidad, ya que se ha definido un límite de consumo para los meses de invierno que aun no se ha demostrado si es el adecuado.

Este trabajo tenía como finalidad mostrar una aplicación de *load research* dentro de los campos de investigación que se han desarrollado a nivel mundial. Para ello, se ha empleado una base de datos disponible en Chilectra, la que fue levantada por esa empresa entre los años 1999 y 2000.

Existen muchas otras aplicaciones que se basan en mediciones de la curva de demanda que sería interesante implementar. Como propuesta se presentan las siguientes:

- Profundizar los estudios de un simulador de curvas e implementar un sistema que permita, a través de simulaciones, planificar expansiones de las redes y los respectivos equipos que se ajusten a los requerimientos.
- Realizar nuevas campañas de medición para tener bases de datos actualizadas. En particular, se propone como alternativa la implementación de un plan de mediciones con lectura de datos *online* que funcione en forma periódica, por ejemplo cada 15 minutos, de tal manera que pueda tener aplicaciones en tiempo real.

- Incorporar al análisis otras variables, como el nivel de tensión o la potencia reactiva, para ampliar el espectro de alternativas de investigación. Adicionalmente se pueden considerar variables no eléctricas, como la temperatura o la luminosidad, y así encontrar nuevas fuentes de desarrollo.
- Recoger variables a través de encuestas que reflejen el uso que se le da a la electricidad con el equipamiento electrodoméstico y variables proxy como por ejemplo la superficie construida o la orientación geográfica de la vivienda, entre otras.

En resumen, el trabajo presentado en esta memoria, entrega los elementos necesarios para formular nuevos estudios en el tema de mediciones de curva de demanda.

9 Bibliografía

- [1] Comisión Nacional de Energía CNE, “Decreto con Fuerza de Ley N°4” 2006, Chile.
- [2] Ministerio Economía, Fomento y Reconstrucción. “Decreto N° 276”, 11 de Febrero de 2005, Chile.
- [3] Lemus, C., “El Comercializador como Agente de Competencia en el Mercado Eléctrico Chileno”, Tesis para optar al grado de Magíster en Ciencias de la Ingeniería, Pontificia Universidad Católica de Chile, Junio 2006
- [4] Comisión Nacional de Energía CNE, “Decreto Supremo N° 327” 10 de Septiembre de 1998, Chile.
- [5] The Association of Edison Illuminating Companies, Load Research Committee. [en línea] http://www.aeic.org/load_research/index.html
- [6] C.S. Chen, J.C. Hwang, Y.M. Tzeng, C.W. Huang, M.Y. Cho, “Determination of Customer Load Characteristics by Load Survey System at Taipower” IEEE Transactions on Power Delivery, Vol. 11, No. 3, Julio 1996.
- [7] C.S. Chen, J.C. Hwang, C.W. Huang, “Application of Load Survey Systems to Proper Tariff Design”, IEEE Transactions on Power Systems, Vol. 12, No. 4, Noviembre 1997.
- [8] C.S. Chen, M.S. Kang, J.C. Hwang, C.W. Huang, “Implementation of the Load Survey System in Taipower” IEEE Transmission and Distribution Conference, Vol. 1, pp. 300 – 304, 11 – 16 Abril 1999.
- [9] Hiroyuki Mori, “State-of-the-Art Overview on Data Mining in Power Systems”, IEEE Power Engineering Society General Meeting, 18 – 22 Junio 2006.
- [10] W.G. Cochran, “Técnicas de Muestro”, Primera Edición, John Wiley & Sons, 1980.
- [11] E. Bompard, E. Carpaneto, G.Chicco, R. Napoli, F. Piglione, P. Postolache, M. Scutariu, “Stratified Sampling of the Electricity Customers for Setting Up a Load Profile Survey”, Proc. RIMAPS 2000, Funchal, Madeira, Portugal, Sept. 25 – 28, 2000, paper RUR-017.
- [12] F. Flatow, “Alternativas de Diversificación del Producto Electricidad”, Memoria de Título para optar a Ingeniero Civil Electricista, Depto. De Ingeniería Eléctrica, Universidad de Chile, 2000
- [13] R. Flores, “Propuesta de Implementación de un Mercado Minorista de Energía en el Sector Eléctrico Chileno”, Memoria de Título para optar a Ingeniero Civil Electricista, Depto. De Ingeniería Eléctrica, Universidad de Chile, 2003

- [14] G. Chicco, R. Napoli, F. Piglione, P. Postolache, M. Scutariu, C. Toader, “A Review of Concepts and Techniques for Emergent Customer Categorisation” Telmark, Londres, 2002.
- [15] D. Gerbec, S. Gasperic, F. Gubina, I. Simon, “Determining the Load Profiles of Consumers based on Fuzzy Logic and Probability Neural Networks” IEEE Proceedings Generation Transmission and Distribution, Vol. 151, No. 3, Mayo 2004.
- [16] D. Gerbec, S. Gasperic, F. Gubina, I. Simon, “Allocation of the Load Profiles to Consumers Using Probabilistic Neural Networks” IEEE Transactions on Power Systems, Vol. 20, No. 2, Mayo 2005.
- [17] D. Gerbec, S. Gasperic, F. Gubina, “Determination of the Consumers’ Load Profiles”, Faculty of Electrical Engineering, University of Ljubljana, Eslovenia, 2001.
- [18] G. Chicco, R. Napoli, F. Piglione, “Application of Clustering Algorithms and Self Organising Maps to Classify Electricity Customers” IEEE Power Tech Conference Proceedings, Vol. 1, 23 – 26 Junio 2003.
- [19] S.V. Verdú, M.O. García, C. Senabre, A. Gabaldón Marín, F.J.G. Franco, “Classification, Filtering, and Identification of Electrical Customer Load Patterns Through the Use of Self-Organizing Maps” IEEE Transactions on Power Systems, Vol. 21, No. 4, Noviembre 2006.
- [20] Medel, F. “Agrupamiento k-means”. [presentación] Curso IN720, Departamento de Ingeniería Industrial, Universidad de Chile.
- [21] Generic Technical Implementation Working Group, GTIWG, Supplier Coordination and Delivery Sub-Team, “Load Profiling for Implementing Retail Choice in Maryland”, 30 Junio 1999.
- [22] OSINERG-GART, CENERGIA, “Estudio de la Caracterización de la Carga de los Sistemas de Distribución Eléctrica – Sectores Típicos 1 2 3 4 5 y Especial”, Perú, Junio 2005.
- [23] R. Rebolledo, “Procedimiento de Toma de Muestras para el Estudio de Curvas de Consumo de Chilectra” Informe Interno Chilectra S.A., Santiago, 29 Diciembre 1998.
- [24] Chilectra S.A., “Determinación de Horas de Uso Clientes BT-1”, Informe Interno, Agosto 2000
- [25] Helsinki University of Technology, “SOM Toolbox for Matlab 5.0”. [en línea] <<http://www.cis.hut.fi/projects/somtoolbox/>>.
- [26] Cuesta-Albertos, J.A., Fraiman, R. and Ransford, T. , “Random projection sand goodness of tests in infinite dimensional spaces”, Bull. Braz. Math. Soc. (N.S.), Vol 37, No 4, pp. 477-501. 2006.
- [27] Bernstein, S., “Tarificación a Costo Marginal en las Empresas Distribuidoras de Electricidad del Sistema Interconectado Central de Chile”, Comisión Nacional de Energía, Santiago, Mayo 1981.

Anexos

Anexo A

En las tablas A.1, A.2, A.3 y A.4 se presentan los valores obtenidos para cada cluster en cada hora bajo el método *fcm* de Matlab.

Enero				Febrero				Marzo			
Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 1	Cluster 2	Cluster 3	Cluster 4
0,7300	0,7544	0,1231	0,5917	0,6562	0,1132	0,6237	0,7282	0,4260	0,7649	0,7962	0,1190
0,5311	0,6482	0,1176	0,4042	0,4705	0,1022	0,4320	0,6063	0,2868	0,6264	0,5663	0,1068
0,3942	0,5596	0,1120	0,2701	0,3467	0,0931	0,2813	0,5150	0,2255	0,4894	0,3689	0,0963
0,3157	0,5229	0,1105	0,2015	0,2939	0,0922	0,2111	0,4869	0,2013	0,4320	0,2685	0,0926
0,2821	0,4978	0,1115	0,1771	0,2710	0,0921	0,1817	0,4546	0,1933	0,3908	0,2244	0,0886
0,2713	0,4797	0,1104	0,1751	0,2681	0,0904	0,1785	0,4626	0,1947	0,3793	0,2085	0,0914
0,2892	0,4907	0,1164	0,1740	0,2815	0,0888	0,1861	0,4648	0,2397	0,3892	0,2092	0,0931
0,3345	0,5317	0,1455	0,2327	0,3319	0,1109	0,2553	0,4835	0,3196	0,4225	0,2651	0,1251
0,3546	0,5614	0,2734	0,2392	0,3606	0,2132	0,2444	0,5295	0,3395	0,4986	0,3362	0,2415
0,3549	0,5589	0,4798	0,2262	0,3420	0,4530	0,2252	0,5543	0,3317	0,5187	0,3095	0,3826
0,3880	0,5817	0,7550	0,2323	0,3702	0,7371	0,2167	0,5864	0,3403	0,5101	0,2734	0,5909
0,4158	0,6409	0,8407	0,2175	0,4046	0,8213	0,2604	0,6138	0,3371	0,5185	0,2934	0,7668
0,4414	0,6575	0,8608	0,2362	0,4378	0,8256	0,2749	0,6304	0,3517	0,5508	0,3090	0,7841
0,4656	0,6928	0,8103	0,2799	0,4604	0,7052	0,2845	0,6518	0,3720	0,5867	0,3164	0,7866
0,4758	0,7051	0,6949	0,2808	0,4835	0,6944	0,2880	0,6788	0,3900	0,6084	0,3386	0,7564
0,4856	0,7216	0,7265	0,2646	0,4850	0,7496	0,2898	0,6853	0,3875	0,5979	0,3409	0,7387
0,4798	0,7293	0,7718	0,2705	0,4643	0,7784	0,2799	0,6769	0,3719	0,5959	0,3390	0,7851
0,4673	0,6896	0,7666	0,2575	0,4566	0,7380	0,2756	0,6645	0,3786	0,5828	0,3196	0,7989
0,4744	0,7010	0,6922	0,2726	0,4677	0,6295	0,2843	0,6604	0,4098	0,5913	0,3259	0,8089
0,5001	0,7010	0,5884	0,2929	0,4874	0,5225	0,3094	0,6856	0,5308	0,6209	0,3613	0,6971
0,5373	0,7272	0,4311	0,3103	0,5756	0,3560	0,3796	0,7265	0,7487	0,7200	0,4721	0,4988
0,6764	0,7901	0,2602	0,5093	0,7571	0,2068	0,5845	0,8217	0,8999	0,8438	0,7382	0,2999
0,8551	0,8503	0,1685	0,7426	0,8867	0,1428	0,8118	0,8731	0,8227	0,8828	0,8852	0,1675
0,8568	0,8297	0,1450	0,7432	0,8158	0,1273	0,7857	0,8356	0,6297	0,8618	0,9067	0,1366

Tabla A.1: Valores de cada cluster por mes.

Abril				Mayo				Junio			
Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 1	Cluster 2	Cluster 3	Cluster 4
0,5584	0,2806	0,5971	0,1643	0,5984	0,4841	0,3119	0,0914	0,2683	0,0916	0,6070	0,5418
0,4034	0,1957	0,3773	0,1395	0,3858	0,3609	0,2078	0,0705	0,1828	0,0850	0,3807	0,3860
0,3506	0,1711	0,2730	0,1165	0,2584	0,2994	0,1750	0,0653	0,1529	0,0783	0,2684	0,3153
0,3294	0,1591	0,2380	0,1241	0,2075	0,2690	0,1575	0,0578	0,1411	0,0727	0,2148	0,2834
0,3181	0,1563	0,2248	0,1158	0,1890	0,2691	0,1609	0,0600	0,1392	0,0794	0,1914	0,2779
0,3336	0,1670	0,2242	0,1169	0,1904	0,2648	0,1648	0,0591	0,1503	0,0803	0,1848	0,2833
0,3950	0,2579	0,2853	0,1377	0,2258	0,3189	0,2387	0,0651	0,2228	0,0872	0,2343	0,3455
0,4659	0,3312	0,3623	0,2133	0,3117	0,4454	0,3387	0,1268	0,3178	0,1480	0,3500	0,5086
0,4295	0,2497	0,3125	0,3550	0,3499	0,4475	0,2668	0,2895	0,2822	0,2842	0,3318	0,4965
0,4472	0,2694	0,2778	0,5772	0,2749	0,4737	0,2710	0,5945	0,2723	0,6177	0,2774	0,4865
0,4914	0,2870	0,2852	0,6257	0,2749	0,5058	0,2947	0,6372	0,3005	0,6704	0,2780	0,5106
0,5038	0,2767	0,3106	0,6477	0,2846	0,5287	0,3154	0,6322	0,3132	0,6933	0,2651	0,5383
0,5356	0,2844	0,2926	0,6748	0,3041	0,5704	0,3271	0,6187	0,3139	0,6674	0,2774	0,5560
0,5752	0,3229	0,3109	0,6525	0,3003	0,5619	0,3306	0,6096	0,3286	0,5827	0,2818	0,5424
0,5772	0,3215	0,3274	0,5969	0,3024	0,5548	0,3326	0,5226	0,3336	0,5459	0,2797	0,5412
0,5142	0,3186	0,3013	0,5771	0,3130	0,5392	0,3369	0,5841	0,3348	0,6421	0,2807	0,5268
0,4885	0,2933	0,2758	0,5795	0,2973	0,5235	0,3393	0,5799	0,3377	0,6161	0,2765	0,5302
0,5547	0,3272	0,3125	0,6351	0,3068	0,5618	0,4369	0,6094	0,4584	0,6259	0,3086	0,5920
0,7043	0,5647	0,4359	0,7116	0,4066	0,7365	0,7161	0,6342	0,7436	0,6294	0,4670	0,7665
0,7959	0,8360	0,6046	0,5203	0,5685	0,8349	0,8574	0,4872	0,8176	0,4896	0,6139	0,8306
0,7530	0,8216	0,6841	0,3240	0,6934	0,8500	0,8532	0,2809	0,8396	0,2785	0,7263	0,8384
0,8147	0,7518	0,7936	0,2409	0,7776	0,8554	0,8227	0,1506	0,8275	0,1594	0,8038	0,8453
0,7871	0,6107	0,8471	0,1992	0,8360	0,7995	0,6842	0,1167	0,6541	0,1079	0,8550	0,7925
0,7415	0,4711	0,8081	0,1942	0,7887	0,6712	0,4833	0,1092	0,4569	0,1037	0,7734	0,6980

Tabla A.2: Valores de cada cluster por mes.

Julio				Agosto				Septiembre			
Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 1	Cluster 2	Cluster 3	Cluster 4
0,2205	0,4951	0,5030	0,0636	0,4614	0,2169	0,0824	0,4395	0,1120	0,6202	0,2720	0,5501
0,1507	0,3462	0,3065	0,0523	0,3637	0,1488	0,0799	0,2815	0,0911	0,4207	0,1742	0,4320
0,1259	0,2916	0,2138	0,0570	0,3156	0,1292	0,0875	0,2122	0,0779	0,2748	0,1397	0,3757
0,1153	0,2556	0,1686	0,0501	0,2987	0,1148	0,0977	0,1818	0,0745	0,2172	0,1268	0,3495
0,1054	0,2332	0,1466	0,0472	0,2845	0,1113	0,0741	0,1668	0,0689	0,1950	0,1198	0,3352
0,1052	0,2270	0,1433	0,0528	0,2933	0,1175	0,0708	0,1673	0,0714	0,1892	0,1247	0,3478
0,1748	0,3162	0,2185	0,0781	0,3570	0,1991	0,0737	0,2451	0,1079	0,2212	0,2087	0,3965
0,2517	0,4238	0,2786	0,1428	0,4928	0,3478	0,1530	0,3349	0,1506	0,2924	0,2721	0,4890
0,2291	0,3945	0,2406	0,3421	0,4843	0,2725	0,3031	0,2773	0,2751	0,3164	0,2460	0,5136
0,2526	0,4323	0,2354	0,7061	0,5313	0,2538	0,6359	0,2890	0,4978	0,3121	0,2482	0,5695
0,2885	0,4871	0,2441	0,7702	0,5831	0,2590	0,7354	0,3047	0,7291	0,3197	0,2664	0,5895
0,3004	0,5223	0,2522	0,7553	0,5830	0,2447	0,7823	0,3091	0,7577	0,3612	0,2923	0,6258
0,3305	0,5402	0,2490	0,7393	0,5646	0,2583	0,7715	0,3144	0,7676	0,3823	0,2894	0,5927
0,3237	0,5423	0,2553	0,7106	0,5693	0,2531	0,7141	0,3405	0,6775	0,3870	0,2816	0,6080
0,3189	0,5221	0,2504	0,6029	0,5657	0,2345	0,6393	0,3443	0,6661	0,3939	0,2856	0,6194
0,3255	0,5053	0,2422	0,6342	0,5324	0,2313	0,6858	0,3276	0,6336	0,3774	0,2757	0,5960
0,3381	0,4977	0,2406	0,6728	0,5084	0,2269	0,7196	0,3141	0,6384	0,3694	0,2727	0,5930
0,4304	0,5705	0,2806	0,7145	0,5293	0,2671	0,7247	0,3406	0,6405	0,3783	0,2985	0,6146
0,7114	0,7654	0,4826	0,6402	0,6793	0,5068	0,6786	0,5748	0,6659	0,4460	0,4217	0,6699
0,8539	0,8267	0,6213	0,4988	0,8033	0,6839	0,5110	0,7942	0,5619	0,6258	0,6846	0,8045
0,8070	0,8423	0,6948	0,2678	0,8237	0,6663	0,2684	0,8397	0,4503	0,7973	0,8142	0,8569
0,7708	0,8656	0,8280	0,1291	0,8351	0,6311	0,1468	0,8943	0,3086	0,8486	0,8109	0,8485
0,5711	0,8243	0,8396	0,0823	0,7440	0,5065	0,1093	0,8450	0,2094	0,8852	0,6819	0,8011
0,3791	0,6944	0,7186	0,0666	0,6085	0,3628	0,1119	0,6831	0,1569	0,8188	0,4736	0,6963

Tabla A.3: Valores de cada cluster por mes.

Octubre				Noviembre				Diciembre			
Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 1	Cluster 2	Cluster 3	Cluster 4
0,0886	0,5916	0,2963	0,6143	0,5794	0,6928	0,3227	0,0911	0,5668	0,7215	0,7694	0,1085
0,0731	0,3779	0,2026	0,4646	0,4549	0,4741	0,2143	0,0732	0,3908	0,5075	0,6051	0,1051
0,0706	0,2710	0,1674	0,3958	0,3916	0,3121	0,1711	0,0687	0,2953	0,3363	0,5096	0,0984
0,0657	0,2136	0,1432	0,3736	0,3607	0,2536	0,1544	0,0699	0,2572	0,2430	0,4617	0,0999
0,0658	0,1946	0,1409	0,3592	0,3475	0,2360	0,1541	0,0682	0,2357	0,2036	0,4323	0,0968
0,0634	0,1944	0,1479	0,3603	0,3416	0,2245	0,1626	0,0686	0,2434	0,1873	0,4336	0,0947
0,0757	0,2411	0,2125	0,4227	0,4036	0,2574	0,2295	0,0796	0,2659	0,1966	0,4426	0,0992
0,1274	0,3353	0,3068	0,5378	0,4755	0,3129	0,2896	0,1479	0,3206	0,2547	0,4919	0,1407
0,2985	0,3118	0,2945	0,5428	0,4756	0,3208	0,2854	0,2824	0,3426	0,2712	0,5238	0,2372
0,5719	0,3144	0,2821	0,5393	0,4967	0,3179	0,2676	0,6201	0,3542	0,2578	0,5422	0,4989
0,7348	0,3431	0,2610	0,5714	0,5450	0,3364	0,2712	0,7552	0,3824	0,2671	0,5978	0,7304
0,7337	0,3610	0,2511	0,6127	0,5631	0,3338	0,2831	0,8058	0,4176	0,2725	0,6305	0,8182
0,7396	0,3841	0,2579	0,6111	0,5881	0,3498	0,2827	0,8198	0,4405	0,2896	0,6555	0,8450
0,7253	0,3919	0,2778	0,6353	0,6065	0,3734	0,3067	0,7349	0,4652	0,2960	0,6607	0,8084
0,6809	0,3880	0,2888	0,6401	0,6344	0,3853	0,3043	0,7303	0,4742	0,3192	0,6955	0,7266
0,6893	0,3781	0,2897	0,6153	0,6040	0,3738	0,2829	0,7599	0,4808	0,3104	0,6728	0,7286
0,7127	0,3703	0,2735	0,6043	0,5842	0,3585	0,2905	0,7708	0,4718	0,3050	0,6652	0,7943
0,6945	0,3674	0,2820	0,5943	0,5777	0,3430	0,2824	0,7746	0,4612	0,2941	0,6594	0,7845
0,6129	0,3916	0,3094	0,6074	0,5803	0,3571	0,3076	0,6879	0,4629	0,2974	0,6645	0,7064
0,5537	0,5041	0,4451	0,6640	0,6369	0,3954	0,3731	0,5416	0,4925	0,3025	0,6705	0,5616
0,3994	0,7505	0,7540	0,8432	0,7624	0,5359	0,6070	0,3899	0,5901	0,3544	0,6937	0,4162
0,2616	0,8758	0,8731	0,8696	0,8650	0,7590	0,8809	0,2203	0,7935	0,5479	0,8057	0,2676
0,1743	0,8974	0,7480	0,8295	0,8429	0,8759	0,8068	0,1342	0,8913	0,8140	0,8713	0,1789
0,1262	0,8204	0,5273	0,7596	0,7540	0,8642	0,5912	0,1021	0,7882	0,8698	0,8547	0,1324

Tabla A.4: Valores de cada cluster por mes.

Anexo B

La tabla B.1 muestra los valores de rechazo del test ks aplicado al simulador de curvas para cada caso.

MES	Cluster			
	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Enero	0,132	0,042	0,013	0,105
Febrero	0,465	0,021	0,116	0,075
Marzo	0,091	0,098	0,106	0,039
Abril	0,053	0,049	0,061	0,039
Mayo	0,077	0,034	0,091	0,015
Junio	0,183	0,067	0,075	0,046
Julio	0,053	0,062	0,066	0,023
Agosto	0,094	0,072	0,013	0,104
Septiembre	0,079	0,145	0,108	0,058
Octubre	0,095	0,099	0,097	0,086
Noviembre	0,092	0,089	0,078	0,023
Diciembre	0,178	0,131	0,096	0,009

Tabla B.1: Valores de rechazo del test ks.

El programa implementado en Matlab para efectuar este cálculo se compone de los métodos presentados a continuación:

Método rechazo: calcula el rechazo que se obtiene al realizar el test ks para comparar dos muestras de datos. Una de ellas corresponde a la muestra real de mediciones y la otra corresponde a las muestras simuladas.

```
function [sumh,res]=rechazo(real,prom,std,k)
res=zeros(k,1);
n=length(real);
p3=simul(prom,std,n);
for i=1:k
    h=randn(24,1);
    p1=real*h;
    p2=p3*h;
    res(i,1)=kstest2(p1,p2);
end
sumh=sum(res)/k;
```

Método simul: crea una matriz de datos simulados utilizando la metodología planteada en el capítulo 6.

```
function [ma_]= simul(prom,desv,t)
prom=prom';
desv=desv';
ma = zeros(24,t);
dv=diag(desv);
for i=1:t
    ma(:,i)=prom+dv*(randn(24,1));
end
j=1;
k=1;
for j=1:t
    for k=1:24
        if ma(k,j)>1
            ma(k,j)=1;
        elseif ma(k,j)<0
            ma(k,j)=0;
        else
            end
        end
    end
end
ma_=ma';
```

Anexo C

Este anexo se incluirá en el CD bajo el nombre de “anexo_c.mat” e incluirá las curvas de demanda agregada por estrato de consumo para clientes residenciales y los perfiles típicos de consumo obtenidos en el desarrollo de la memoria.