

Sähkönkulutuksen tuntimittauksien hyödyntäminen tyyppikäyrien muodostamiseksi

Henri Selenius

Sähkötekniikan korkeakoulu

Diplomityö
Espoo 23.04.2020

Työn valvoja

Prof. Matti Lehtonen

Työn ohjaaja

DI Juha Haikonen



Tekijä	Henri Selenius				
Työn nimi	Sähkönkulutuksen tuntimittauksien hyödyntäminen tyyppikäyrien muodostamiseksi				
Koulutusohjelma	Automation and Electrical Engineering				
Pääaine	Sähköenergiatekniikka	Koodi	ELEC3024		
Työn valvoja	Prof. Matti Lehtonen				
Työn ohjaaja	DI Juha Haikonen				
Päiväys	23.04.2020	Sivumäärä	8+93	Kieli	suomi

Tiivistelmä

Sähkönkulutuksen pitkän aikavälin ennustaminen toimii sähköverkon strategisen suunnittelun lähtökohtana. Perinteisten kuormitusmallien maantieteelliseen sijaintiin sidottu luonne, liittymäsopimusten tekemisen ajoilta peräisin olevat asiakastyypitiedot ja ainoastaan tammikuun lämpötilariippuvuuskertoimien käyttö virheellisesti koko vuodelle ovat perinteisille kuormitusmalleille ominaisia piirteitä. Sähkönkäytön ennustamiseen liittyvien epävarmuustekijöiden minimoimiseksi on syytä luoda uusia verkkoyhtiökohtaisesti määritettyjä alueellisia kuormitusmalleja.

Työssä muodostettiin uudet sähkönkulutuksen indeksisarjat eli tyyppikäyrät asiakkaiden pitkän aikavälin sähkönkäytön mallintamiseksi. Tyyppikäyrät luotiin käyttämällä sähkönjakeluyhtiö Kajave Oy:n toimittamia tuntitasolla mitattuja käyttöpaikkakohtaisia kahden vuoden tuntimittauksia perinteisissä malleissa käytetyn arvioidun vuosienenergiaan sijaan. Asiakkaat ryhmiteltiin tuntimittausten perusteella kuluttajaryhmiin. Indeksisarjat muodostettiin erikseen jokaiselle ryhmittelyn tuloksena saadulle kuluttajaryhmälle. Näin muodostettuja tyyppikäyriä verrattiin arviotarkkuuden osalta Kajavella nykyisin käytössä oleviin kuormituskäyriin eli tyyppikäyriin ja todelliseen vuoden 2019 tuntimittausten vertailudataan, jota ei käytetty indeksisarjojen luonnissa. Tämän perusteella tehtiin johtopäätökset siitä, kumpi indeksisarjoista mallintaa todellista tuntimittausten vertailudataa paremmin.

Työn tavoitteena oli parantaa ja tarkentaa Trimblen järjestelmien tehonjakolaskennan ja keskeytyksestä aiheutuneen haitan laskennan (KAH) antamia tuloksia. Työn tuloksien perusteella voidaan sanoa, että uusista indeksisarjoista lasketut tuntitehot mallintavat Kajavella nykyisin käytössä olevia indeksisarjoja paremmin todellista tuntimitattua kulutusta ja ovat potentiaaliset tarkentamaan järjestelmien tehonjako- ja KAH-laskennasta saatavia tuloksia. Uusien indeksisarjojen käyttöönotto vaatii kuitenkin lisätutkimusta, koska ryhmittelyn ulkopuolelle jätettyjen käyttöpaikkojen lukumäärä oli suuri. Tekoälyn avulla voitaisiin määrittää ryhmittelyn ulkopuolelle jätetyille käyttöpaikoille sopivin neljästä muodostuneesta kuluttajaryhmästä ja laskea indeksisarjoista tällaisille käyttöpaikoille tuntitehot, joita verrattaisiin todelliseen tuntimittausten vertailudataan. Mikäli mallinnus toimisi paremmin myös tässä tapauksessa verrattuna käytössä oleviin indeksisarjoihin, uudet indeksisarjat voitaisiin ottaa käyttöön.

Avainsanat AMR, klusterointi, KAH, tehonjakolaskenta, sähkönkulutus, ennustaminen, PCA, K-means, tuntimittaus



Author	Henri Selenius				
Title	Utilizing hourly electricity consumption data to generate load curves				
Degree Programme	Automation and Electrical Engineering				
Major	Electrical Power and Energy Engineering	Code	ELEC3024		
Supervisor	Prof. Matti Lehtonen				
Advisor	MSc Juha Haikonen				
Date	23.04.2020	Number of pages	8+93	Language	Finnish

Abstract

Long-term electricity consumption forecasting serves as a starting point for strategic grid planning. The dependence from geographic location, old customer type data and temperature dependence coefficients of January incorrectly used throughout the year alone are all characteristics of traditional load models. In order to minimize the uncertainties associated with electricity use forecasting, new regional load models should be created for each network operator.

New electricity consumption index series were created in the thesis to model long-term electricity consumption of customers. The load curves were generated using hourly measured two-year data received from the electricity distribution company Kajave Oy instead of using estimated annual energy of customers. Customers were grouped into consumer groups based on hourly measurements. Index series were created separately for each consumer group resulting from the grouping. The load curves generated in this way were compared to the current load curves in use in Kajave and to the actual hourly measurement data of year 2019.

The goal of this thesis was to improve and refine the results of the load flow calculation and customer interruption cost calculation (CIC) of the Trimble systems. Based on the results of the thesis, it can be said that the hourly powers calculated from the new index series better model the actual hourly consumption than the index series currently in use in Kajave. The new index series have the potential to improve accuracy of the load flow calculation and CIC calculation of the systems. However, the introduction of the new index series will require further study due to the large number of excluded customer points. Artificial intelligence could determine the most appropriate of the four consumer groups formed for non-grouped customers. After that, the calculated hourly powers from the index series would be compared to the actual hourly measured data. If modeling worked better in this case as well, the new index series could be introduced.

Keywords AMR, clustering, CIC, load flow calculation, electricity consumption, forecasting, PCA, K-means, hourly measurement

Esipuhe

Ensinnäkin haluan kiittää Trimble Solutions Oy:tä mahdollisuudesta tehdä diplomityö ja käyttää tähän osa viikon työpäivistä. Työn aihe oli hyvin ajankohtainen ja monipuolinen. Haluan kiittää kaikkia, jotka auttoivat minua diplomityön teon eri vaiheissa. Kiitän lämpimästi ohjaajaani DI Juha Haikosta asiantuntevasta ohjauksesta ja neuvoista. Isot kiitokset työn valvojalle professori Matti Lehtoselle työn ohjauksesta ja vastauksista kysymyksiin, jotka askarruttivat työn teon aikana. Suuret kiitokset myös TkL Janne Mörskylle asiantuntevasta avusta työn eri vaiheissa ja insinööri Arto Matsiselle aiheen suuntaviivojen näyttämisestä.

Otaniemi tarjosi loistavat puitteet ja mahtavat tapahtumat näiden kaikkien kuluneiden viiden vuoden aikana. Kiitokset kaikille opiskelukavereilleni, että teitte näistä vuosista ikimuistoiset. Kiitos kuuluu myös perheelleni ja ystävilleni kaikesta tuesta, jota sain läpi opiskeluaikojen.

Espoossa, 23.04.2020

Henri Selenius

Sisällysluettelo

Tiivistelmä	ii
Abstract	iii
Esipuhe	iv
Sisällysluettelo	v
Symbolit ja lyhenteet	vii
1. Johdanto	1
2. Mahdolliset lähestymisnäkökulmat aiheeseen	4
3. Kuormituksen mallintaminen	6
3.1 Perinteiset kuormitusmallit.....	6
3.2 Kuormituksen summaaminen eli yhdistetyn kuorman laskeminen	10
3.3 Kuormituksen laskemiseen vaikuttavat merkittävimmät tekijät	14
3.4 Maksimikuormituksen arvioiminen	16
4. Sähkönkulutuskäyrien perinteisen muodostuksen pääperiaatteet	19
4.1 Kuluttajaryhmien muodostaminen	20
4.2 SENERin indeksisarjat	22
4.3 Indeksisarjojen muodostus	26
4.3.1 Sähkönkulutuksen indeksisarjat	26
4.3.2 Indeksisarjojen lämpötilakorjaus (mitatun aikasarjan normalisointi vai mitatun energian normalisointi)	28
4.4 Perinteisten sähkönkulutuskäyrien virheet ja ongelmat	32
4.5 Ajantasaisempien sähkönkulutuskäyrien muodostamisen edut	34
5. Tuntimitatun datan hyödyntäminen uusien kuormituskäyrien luomiseksi 35	
5.1 Verkkoyhtiöltä saatu aineisto.....	36
5.2 Tuntimitatun sähkönkulutusdatan esiprosessointi	37
5.3 Aineiston dimensioiden redusointimenetelmät.....	38
5.3.1 Aineiston jakaminen vuosijaksoittain	38
5.3.2 Pääkomponenttianalyysi.....	39
5.4 Valittu klusterointimenetelmä.....	44
5.5 Klusteroinnin tulokset	49
5.6 Lämpötilariippuvuuden huomiointi indeksisarjoissa	55
5.6.1 Lämpötilariippuvuuskertoimien laskenta	55
5.6.2 Mitatun aikasarjan normalisointi	59
5.7 Uusien indeksisarjojen luominen	61
5.7.1 Sisäisten indeksien muodostus.....	61
5.7.2 Ulkoisten indeksien muodostus	64
5.8 Korrelaation huomioiminen indeksisarjoissa	66
5.9 Tehokertoimet kuormituskäyrissä	66
5.10 Asiakaskohtainen/Verkkoyhtiökohtainen mallintaminen.....	68
6. Tulokset ja johtopäätökset	69
6.1 Käytetty tulosten vertailutapa	69

6.2	Tulosten vertailu	72
6.3	Uusien sähkönkulutuskäyrien mahdolliset epätarkkuustekijät	83
7.	Sähkönkulutuskäyrien tulevaisuuden haasteet	84
7.1	Hajautettu tuotanto	84
7.2	Sähköautojen lataus	85
7.3	Aurinkovoima	86
7.4	Kysyntäjousto	86
7.5	Sähkön varastointi	87
8.	Yhteenveto	89
	Lähteet	92

Symbolit ja lyhenteet

Symbolit

A	ampeeri
C_i	klusteri, johon datapiste i on sijoitettu klusteroinnin yhteydessä
c	vuodenajasta riippuva lämpötilariippuvuuskerroin
c_p	maksimikuormituksen arvioinnissa käytettävä korjauskerroin
$\cos(\varphi)$	tehokerroin
d	datapisteen ja klusterin keskipisteen välinen etäisyys
E_k	kuluttajaryhmästä k valittujen käyttöpaikkojen vuosienergioiden summa
k_1, k_2	Velanderin kertoimet
kW	kilowatti
$M_{k,i}$	kuluttajaryhmän k kaksiviikkojakson i ulkoinen indeksi
m	asiakkaan tietyn ajan odotettavissa oleva kulutus
m_{ij}	kuluttajaryhmään i kuuluvan asiakkaan j arvioitu kulutus
$m_{k,i}$	kuluttajaryhmän k kaksiviikkojakson i sisäinen indeksi
n_i	kuluttajaryhmään kuuluvien asiakkaiden lukumäärä
P	tietyn ajanhetken teho
P_{lk}	tietyn ajanhetken teho, josta on poistettu lämpötilan vaikutus
P_{max}	odotettavissa oleva huippukuormitus
$P_{m,ki}$	indeksisarjoista laskettu tietyn ajanhetken teho
Q	loisteho
Q_a	tilastollisesta jakaumasta riippuva kerroin
$s(i)$	datapisteen i silhouetten kerroin
T_d	vuorokauden keskilämpötila
T_n	kaksiviikkojakson pitkän aikavälin normaalilämpötila
u_1	vuoden 2017 tarkasteltavan kaksiviikkojakson ulkoinen indeksi
u_2	vuoden 2018 tarkasteltavan kaksiviikkojakson ulkoinen indeksi
W	vuosienergia
W_k	kuluttajaryhmään k kuuluvan yksittäisen asiakkaan vuosienergia
X	kaksiviikkojakson indeksien avulla mallinnettu energia
x_i	tarkasteltava datapiste
y_i	klusterin keskipiste
ΔT	vuorokauden keskilämpötilan ja kaksiviikkojakson normaalilämpötilan erotus
ρ_{ii}	samaan kuluttajaryhmään kuuluvien kahden kuorman välinen keskimääräinen korrelaatio
ρ_{ij}	yksittäisten kuormien välinen tilastollinen korrelaatio
σ	keskihajonta
σ^2	variassi

Lyhenteet

AMR	Automatic meter reading, Automaattinen mittarinluenta
CHP	Combined Heat and Power, Yhdistetty sähkön- ja lämmöntuotanto
CIC	Customer Interruption Cost, Keskeytyksestä aiheutunut haitta
DMS	Distribution Management System, Sähkönjakelun hallintajärjestelmä
GGK062003	Sähköverkkoyhtiö Kajavella nykyisin laskennassa käytössä olevat indeksisarjat
KAH	Keskeytyksestä aiheutunut haitta
KAJAVE20	Tässä työssä luodut Kainuun alueen asiakkaiden sähkönkulutusta ennustavat uudet indeksisarjat
KAJAVE OY	Kainuun sähköverkko
K-MEANS	Clustering algorithm, Klusterointialgoritmi
MATLAB	Multi-Paradigm numerical computing environment and proprietary programming language developed by MathWorks, MathWorks:n kehittämä matematiikkaohjelmisto
NAN	Not a number, Puuttuva mittaus
NIS	Network Information System, Verkkotietojärjestelmä
PCA	Principal component analysis, Pääkomponenttianalyysi
R8	Kahdeksasta suomalaisesta maaseutus sähköverkkoyhtiöstä koostuva ryhmä
SENER	Sähköenergialiitto ry (entinen Suomen Sähkölaitosyhdistys ry)
SLY	Suomen Sähkölaitosyhdistys ry
VTT	Teknologian tutkimuskeskus VTT (entinen Valtion teknillinen tutkimuskeskus)

1. Johdanto

Sähkönkulutus on kasvanut Suomessa jatkuvasti. Yksi sähköverkkoyhtiön tärkeimpiä tehtäviä on taata luotettava sähkönsaanti kaikille verkkoalueen asiakkaille. Kaupunkien kasvaessa ja väkimäärän lisääntyessä sähkökuormitus lisääntyy kasvavalla vauhdilla. Toisaalta tavoitteet ilmastonmuutoksen estämiseksi voivat johtaa poliittisiin määräyksiin, joilla pyritään vähentämään sähkönkulutusta. Uusiutuvan energian lisääntyminen sekä puhtaampaan energiantuotantoon siirtyminen vaikuttavat osaltaan sähkönkäytön tottumuksiin. Sähköverkkoyhtiöiden on tärkeää pystyä ennalta arvioimaan asiakkaiden sähkönkäytön muutokset sekä vastaamaan yhteiskunnan muutoksen seurauksena aiheutuneisiin sähkön kulutustottumusten muutoksiin.

Uusien sähkölinjojen, voimalaitosten sekä sähköasemien suunnittelu- ja rakentamisprojektit ovat yleensä hitaita ja kestävät useita vuosia. Käytössä olevien laitteistojen pitoajat ovat kestoiltaan vuosikymmeniä. [1] Edellä mainittujen seikkojen takia pitkän aikavälin sähkönkulutuksen ennustaminen on tärkeää. Monella sähköverkkoyhtiöllä on vielä tänäkin päivänä käytössä sähkönkäytön ennustamiseen luodut Sähköenergialiiton (SENER) vanhat indeksisarjat eli tyyppikäyrät. Yksi näiden suurimmista puutteista on se, että alueen maantieteellistä sijaintia ei ole huomioitu. Lisäksi sähkönkäyttö muuttuu vuosien kuluessa, minkä vuoksi kuormitusmalleja tulisi aika ajoin päivittää.

Asiakkaiden eli käyttöpaikkojen tunnitista sähkönkulutusta mittaavien etäluettavien älykkäiden sähkömittarien (Automatic Meter Reading) määrän huomattava lisääntyminen viime vuosina mahdollistaa tuntimittauksiin perustuvan uusien aluekohtaisten kuormitusmallien luomisen. Kuormitusmallit voidaan siis tuntimitatun datan yleistymisen vuoksi luoda alueellisia lämpötiloja ja mittauksia käyttämällä. Näin saadaan paremmin kyseisen alueen asiakkaiden kuormitusta ennustavat mallit.

Tässä työssä luodaan uudet sähkönkulutuksen tyyppikäyrät asiakkaiden pitkän aikavälin sähkönkäytön mallintamiseksi. Tyyppikäyrät muodostetaan käyttämällä sähkönjakeluyhtiö Kajave Oy:ltä saatavilla olevia tuntitasolla mitattuja asiakaskohtaisia kahden vuoden tuntimittauksia. Kajavelta saatua aineistoa käsitellään luottamuksellisesti ja ainoastaan tämän työn tutkimustarkoitusta varten. Työn

tavoitteena on parantaa ja tarkentaa Trimblen järjestelmien tehonjakolaskennan ja keskeytyksestä aiheutuneen haitan laskennan (KAH) antamia tuloksia. Tämä voidaan saavuttaa hyödyntämällä tehokkaammin olemassa olevia tuntimitattuja sähkönkulutusmittauksia. Trimblen järjestelmät tukevat suorien tuntimittausten käyttöä ainoastaan verkostolaskennassa, jossa mittauksia voidaan käyttää yhtäjaksoisesti viimeiset 12 kuukautta. Tämä tapa on mahdollistanut sen, että on voitu tarkastella esimerkiksi sitä, mitä viime vuoden tammikuussa verkossa on tapahtunut. Ongelmana puolestaan on, että verkon tehonjakolaskentaan menetelmä ei anna oikeaa kuvaa siitä syystä, että viime vuoden tammikuu ja koko talvi saattoi olla poikkeuksellisen lämmin. Tämän vuoksi kulutus saattoi olla joissakin käyttöpaikoissa normaalia alhaisempi. Tavoitteena on myös paketoita saatu ratkaisu verkkoyhtiöille sopivaksi tuotteeksi.

Asiakkaat ryhmitellään tuntimittaussarjojen perusteella kuluttajaryhmiin. Indeksisarjat muodostetaan jokaiselle ryhmittelyn tuloksena saadulle kuluttajaryhmälle erikseen. Näin muodostettuja tyyppikäyriä verrataan arviotarkkuuden osalta Kajavella nykyisin käytössä oleviin kuormituskäyriin eli tyyppikäyriin ja todelliseen tuntimittausten vertailudataan. Tämän perusteella voidaan luoda johtopäätökset siitä, kummat indeksisarjoista mallintavat todellista tuntimittausten vertailudataa paremmin. Käyttöpaikkakohtaisen tarkastelun sijaan toinen vaihtoehto olisi ollut tarkastella kulutuksia liittymätasolla. Summaamalla käyttöpaikkojen mittaukset yhteen jokaiselle tunnille olisi ollut mahdollista luoda tyyppikäyrät myös liittymille. Summausta olisi voitu jatkaa myös jakelumuuntajille asti ja muodostaa myös näille omat tyyppikäyränsä. Useissa aikaisemmissa tutkimuksissa on tarkasteltu kulutuksia liittymätasolla kulutuksen tasaisemman käyttäytymisen vuoksi.

Työssä on teoriaosuus, jonka luvussa 2 käsitellään työn rajausta eli mahdollisia lähestymisnäkökulmia aiheeseen, luvussa 3 kuormituksen perinteisiä mallintamismenetelmiä ja luvussa 4 sähkönkulutuskäyrien perinteisen muodostuksen pääperiaatteita, perinteisten menetelmien haasteita ja ongelmia sekä ajantasaisempien tyyppikäyrien muodostamisen etuja. Tutkimusosuuden muodostaa matemaattinen ongelma eli se, miten tyyppikäyriä muodostetaan tuntimitatusta datasta ja mihin kriteereihin perustuen. Tutkimusosuutta käsitellään luvuissa 5 ja 6. Työn lopussa luvussa 7 tarkastellaan indeksisarjoihin perustuvien kuormitusmallien tulevaisuuden

haasteita. Merkittävimmät haasteita aiheuttavat tekijät ovat hajautettu tuotanto, sähköautojen lataus, kysyntäjousto, aurinkovoima ja sähkön varastointi.

2. Mahdolliset lähestymisnäkökulmat aiheeseen

Työssä tarkasteltavaa aihetta eli sähkönkulutuksen tuntimittauksien hyödyntämistä tyyppikäyrien muodostamiseksi voi lähestyä monesta eri näkökulmasta. Tuntimitattua dataa on saatavilla verkkoyhtiöiltä useamman vuoden ajalta, mutta sitä onko tuntimittauksia hyödynnetty, riippuu verkkoyhtiöstä. Tuntimittauksien hyödyntämistä voidaan tutkia monella eri tapaa.

Aiheeseen liittyviä ajankohtaisia tutkimuskysymyksiä on useita. Verkkoyhtiötä haastatteleamalla voitaisiin selvittää, kuvaavatko nykyisin heillä käytössä olevat indeksisarjat käyttöpaikkojen sähkönkulutusta riittävällä tarkkuudella ja kuinka usein näitä indeksisarjoja päivitetään verkkoyhtiön omasta toimesta vai päivitetäänkö niitä ylipäätään ollenkaan. Toinen oleellinen kysymys on se, että onko kuluttajaryhmiä päivitetty eli vastaavatko ne nykyisiä sähkön kulutustottumuksia. Kolmas aiheeseen liittyvä kysymys on, että mitä SENERin indeksisarjojen kuluttajaryhmiä verkkoyhtiöllä on käytössä ja miten suuria virheitä nämä aiheuttavat käytössä oleviin indeksisarjoihin perustuvassa laskennassa. Vertailemalla tarkasteltavan käyttöpaikan kulutuskäyttäytymisen poikkeamaa olemassa olevan kuluttajaryhmän indeksisarjaan, nähdään mallintaako kyseinen sarja enää kulutuspistettä riittävän tarkasti. Tähän vaikuttaa asetettava raja-arvo poikkeaman suuruudelle. Tilanne voi olla esimerkiksi se, että ulkoinen indeksisarja vastaa hyvin asiakkaan kulutusta, mutta sisäinen indeksisarja ei vastaa ollenkaan.

Alla on listattu vielä erikseen aiheen kannalta oleellisimpia ajankohtaisia tutkimuskysymyksiä:

1. Miten verkkoyhtiöt hyödyntävät nykyisin käytössä olevia indeksisarjoja?
2. Miten verkkoyhtiöt päivittävät indeksisarjoihin vaikuttavia tietoja?
3. Mitä SENERin indeksisarjojen kuluttajaryhmiä verkkoyhtiöt käyttävät?
4. Miten tuntimittauksista voidaan muodostaa uudet ja ajankohtaisemmat indeksisarjat?

Tässä työssä tutkimuskysymykseksi valittiin se, että miten tuntimittauksista voidaan muodostaa uudet ja ajankohtaisemmat indeksisarjat. Indeksisarjat toteutetaan asiakaskohtaisesti käyttämällä verkkoyhtiöltä saatavaa tuntimitattua dataa arvioidun vuosienergian sijaan.

Nykyisiin kuormitusmalleihin ja niiden toimivuuteen liittyy myös muita ajankohtaisia kysymyksiä. Näitä ovat esimerkiksi seuraavat:

1. Ovatko kaksiviikkojaksoihin perustuvat indeksisarjat enää järkeviä?
2. Onko arkipäivä-lomapäivä-pyhäpäivä jaottelu enää ajankohtainen?
3. Korreloiko kaksi peräkkäistä viikkoa toisiaan kulutukseltaan?
4. Onko tarpeen siirtyä tuntimittauksien sijaan mallintamaan kulutusta 15-minuutin mittauksilla tulevaisuudessa?

Tässä työssä kuitenkin keskitytään tuntimitattuun dataan, koska 15-minuutin tase tulee näillä näkymin vasta vuonna 2023. Huomiota ei kiinnitetä myöskään yllä oleviin muihin kysymyksiin, vaan pääpaino on uusien, verkkoyhtiökohtaisten tyyppikäyrien luomisessa.

3. Kuormituksen mallintaminen

Jakeluverkon uusia osia suunnitellessa sekä nykyisen järjestelmäkapasiteetin riittävyyden mittaamisessa on tärkeää arvioida odotettua kuormituskäyttäytymistä eri jännitetasoilla. Tämä perustuu eri kuormitusmalleihin, suoriin mittauksiin tai näiden kahden yhdistelmän tuottamaan tietoon. Tällainen lähestymistapa antaa kohtuullisen hyvät tulokset korkeammille järjestelmätasojille, joissa yhdistetty kuorma koostuu useista tuhansista asiakkaista. Kuormituksen arvioiminen on sitä hankalampaa ja ennalta-arvaamattomampaa, mitä vähemmän asiakkaita on tarkastelussa mukana. Kuormituksen tunteminen on siis yksi oleellisimmista tekijöistä sähköverkon mitoituksen ja suunnittelun kannalta tarkasteltuna. Parhaassa mahdollisessa tapauksessa tiedetään kuormitus kuluttajakohtaisesti. [2]

3.1 Perinteiset kuormitusmallit

Suurin ongelma kuormituksen arvioimisessa ja mallintamisessa on pitkään ollut tuntitehomittausten saatavuuden vähyys. Tällaisia mittareita, jotka mittaavat tuntitehoja on löytynyt pelkästään isoimmilta kuluttajilta. Sähkömarkkinalain edellyttämänä nykyään tehotietoja on saatavilla myös pienemmiltä kuluttajilta, mikä luo mahdollisuuden entistä tarkempien kuormitusmallien muodostamiseksi. Tähän palataan myöhemmin.

Perinteisissä kuormitusmalleissa suurimpia kuluttajia lukuun ottamatta kuormituksen arvioiminen perustuu arvioituun vuosienergiaan. Kuormitusmallien kehittämisessä on hyödynnetty asiakaskohtaisia vuosienergioita. Tällainen vuosienergiaan perustuva perinteinen malli on Velerin kaava. Menetelmällä pystytään määrittämään huipputehot kuluttajaryhmälle sekä yksittäiselle kuluttajalle. Alla on määritelty Velerin kaava. [3]

$$P = k_1 W + k_2 \sqrt{W} \quad (1)$$

missä

W on kuluttajakohtainen vuosienergia

k_1 ja k_2 tietyille kuluttajaryhmälle ominaiset Velerin kertoimet

Kertoimet saadaan määritettyä keskijännitelähtöjen maksimitehoja hyödyntämällä [3]. Veleranderin kaavalla laskettujen huipputehojen avulla voidaan määrittää arvioitu alueen huipputeho. Tämä tapahtuu kertomalla tietyille kuluttajaryhmälle lasketut huipputehot niin kutsutun osallistumiskertoimen kanssa ja summaamalla saadut tulokset yhteen. Osallistumiskerroin kuvaa sitä, miten suuri osa kuluttajaryhmän huipputehosta ajoittuu alueen huipputehon kanssa samaan aikaan. Menetelmän huono puoli on se, ettei Veleranderin kaava huomioi kuormituksen ajasta riippuvaa vaihtelevuutta eli sitä, että kuluttajaryhmien huipputehot ilmenevät eri aikoina. Tämän seurauksena kuormituksen hetkellinen teho on paljon pienempi todellisuudessa. [2]

Edellä mainitun ongelman ratkaisemiseksi on toteutettu eri kuormitusmalleja. Näiden avulla on mahdollista määrittää asiakkaan kuluttama keskimääräinen teho jokaisena ajanhetkenä. Kuormitusmallit ovat aina kuluttajaryhmäkohtaisia eli ne on erikseen toteutettu jokaiselle ryhmälle. Nämä mallit perustuvat pitkän aikavälin tuntimittauksiin kuluttajaryhmän sisällä. Otannan mittaukset keräämällä yhteen on mahdollista laskea tiettyä ajanhetkeä vastaava kulutuksen keskiarvo ja keskihajonta. Laskentatuloksia voidaan käyttää kaikkien kuluttajaryhmän jäsenten kulutuksen mallintamiseen. [2]

Useimmissa Suomen sähköverkkoyhtiöissä on käytetty yksinkertaista kuormitusmallia jakeluverkon kuormituksen määrittämisessä. Mallit perustuvat kuluttajaryhmäkohtaisesti määritettyihin kuormituskäyriin eli indeksisarjoihin. Odotettavissa oleva kulutus ja sen keskihajonta voidaan laskea, kun kuluttajaryhmään i kuuluvan asiakkaan j vuosienergia W_j tunnetaan. Nämä voidaan laskea seuraavalla kaavalla. [4]

$$\begin{aligned} m_{ij}(t) &= \bar{\mu}_i(w(t), d(t), h(t)) \times W_j \\ \sigma_{ij}(t) &= \bar{v}_i(w(t), d(t), h(t)) \times W_j \end{aligned} \quad (2)$$

missä

m_{ij} on kuluttajaryhmään i kuuluvan asiakkaan j arvioitu kulutus

σ_{ij} on arvioitun kulutuksen keskihajonta

$\bar{\mu}_i$ on kuluttajaryhmän i normalisoitu keskimääräinen kulutus

\bar{v}_i on kuluttajaryhmän i normalisoitu keskihajonta

Normalisoitu keskimääräinen kulutus ja keskihajonta on määritelty tarkasteltavan kuukauden, viikonpäivän ja vuorokauden tunnin funktioina [5]. Yksinkertainen kuormitusmalli perustuu tuntikeskitehon ja keskihajonnan esittämiseen vuosienergian lineaarisena funktiona. Normalisoitu keskimääräinen kulutus $\bar{\mu}_i$ ja normalisoitu keskihajonta \bar{v}_i voidaan määrittää käytössä olevan mittausaineiston perusteella. Vuosienergioita ja tuntimittauksia hyödyntämällä saadaan laskettua $\bar{\mu}_i$ ja \bar{v}_i . [2]

$$\begin{aligned}\bar{\mu}_i(w, d, h) &= E \left\{ \frac{W_{h,i}(w, d, h)}{W_{a,i}} \right\} \\ \bar{v}_i(w, d, h) &= \sigma \left\{ \frac{W_{h,i}(w, d, h)}{W_{a,i}} \right\}\end{aligned}\tag{3}$$

missä

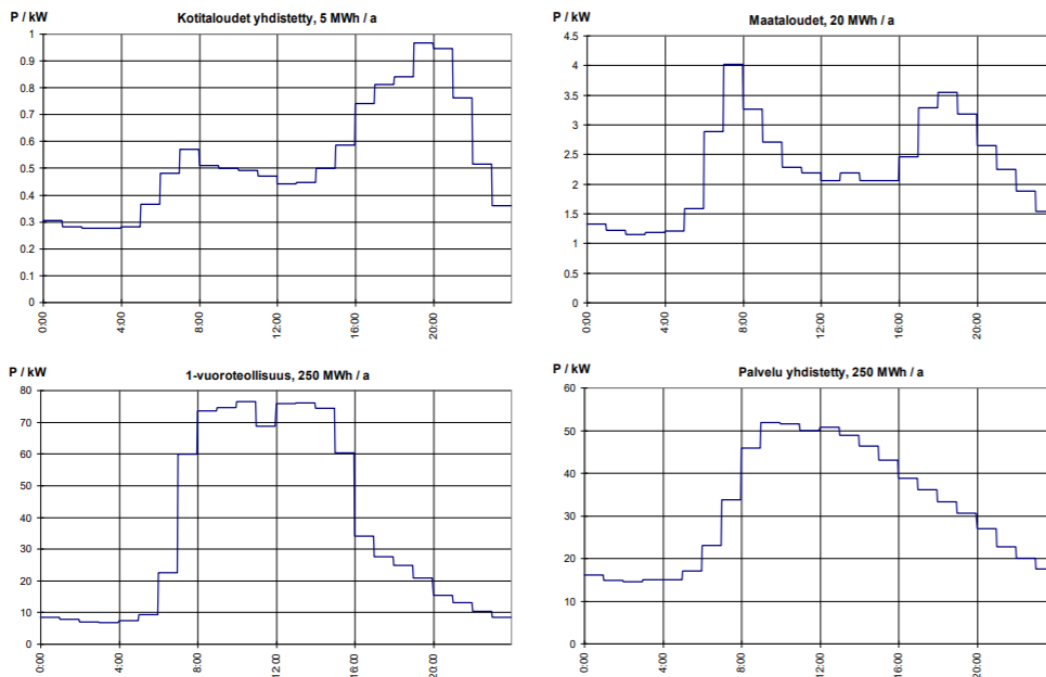
$W_{h,i}$ on kuluttajaryhmän i tiettyä ajanhetkeä vastaava tuntienenergia

$W_{a,i}$ on kuluttajaryhmään i kuuluvan asiakkaan vuosienergia

$E\{X\}$ on odotusarvo muuttujalle X

$\sigma\{X\}$ on keskihajonta muuttujalle X

Edellä kuvatun mallin funktion kuormituksen muuttujina on käytetty pelkästään kuukautta, viikonpäivää ja vuorokauden tuntia. On mahdollista muodostaa myös paljon monimutkaisempia malleja, joissa kuormituksen kannalta merkittäviä muuttujia olisi huomattavasti enemmän. Nykyisin kuluttajaryhmäkohtaiset mallit ovat eniten käytettyjä tarkasteltaessa jakeluverkon ja kuluttajien kuormitusta. Kuormitusmallien soveltaminen ja käyttö laskennassa toimii monien laskentaohjelmistojen perustana. Esimerkkinä tällaisesta ohjelmistosta on sähköverkon tilaseurantaohjelmistot. Kuva 1 havainnollistaa eri kuluttajaryhmien tyyppikuormitusmalleja. [2]



Kuva 1: Kuluttajaryhmien tyyppikuormitusmalleja [5]

Kuormitusmallit perustuvat suureen määrään mitattua dataa. Yhteensä 46 erilaista mallia on tällä hetkellä esitetty. Näistä malleista jokainen kuvaa tiettyä kuluttajaryhmää ja kulutuskäyttäytymistä kyseisessä ryhmässä. Nykyisin käytössä olevat Sähköenergiailiiton (SENER) indeksisarjat on luotu 46 eri kuluttajaryhmälle. Nämä mallit ovat saatavilla Sähköenergiailiitosta (entinen SLY). Alla oleva taulu 1 havainnollistaa joitakin nykyisin käytössä olevia kuluttajaryhmiä. [2]

Class	description
110	residential, direct electric heating, water boiler < 300 l
120	residential, direct electric heating, water boiler ≥ 300 l
220	residential, electric heating with heat storage
602	residential, no electric heating, electric sauna
712	agriculture, milk production, housing
810	1-shift industry
820	2-shift industry
910	public service
920	commercial service

Taulu 1: Nykyisin käytössä olevia kuluttajaryhmiä [4]

3.2 Kuormituksen summaaminen eli yhdistetyn kuorman laskeminen

Oletetaan, että yksittäisten asiakkaiden kulutuksen eli kuormien i keskiarvo m_i ja keskihajonta σ_i tunnetaan. Yhdistetylle kuormalle saadaan tällöin keskiarvoksi ja varianssiksi. [4]

$$\begin{aligned} m &= \sum_{i=1}^n m_i \\ \sigma^2 &= \sum_{i=1}^n \sum_{j=1}^n \sigma_i \sigma_j \rho_{ij} \end{aligned} \quad (4)$$

missä

ρ_{ij} on yksittäisten kuormien välinen tilastollinen korrelaatio

Yleensä korrelaatiokomponenttien todellisia arvoja ei ole tunnettu ja tämän vuoksi niitä ei ole huomioitu nykyisissä kuormitusmalleissa. Käytännön tapauksissa yksittäisten kuormien ominaispiirteitä ei tunneta, mutta niitä täytyy mallintaa käyttämällä keskimääräisiä arvoja, jotka on määritetty kuormaa mallintavalle kuluttajaryhmälle. Käyttämällä tyyppikuormitusmalleja edellinen varianssin kaava saa muodon. [4]

$$\sigma^2 = \sum_{i=1}^k n_i \bar{\sigma}_i^2 + \sum_{i=1}^k n_i (n_i - 1) \bar{\sigma}_i \bar{\rho}_{ii} + \sum_{i=1}^k n_i \bar{\sigma}_i \sum_{j=1}^k n_j \bar{\sigma}_j \bar{\rho}_{ij} \quad (5)$$

missä

$\bar{\sigma}_i$ on kuluttajaryhmään i kuuluvan kuorman keskihajonta

n_i on kuluttajaryhmään kuuluvien kuormien lukumäärä

$\bar{\rho}_{ii}$ on kahden kuluttajaryhmään i kuuluvan kuorman välinen keskimääräinen korrelaatio

$\bar{\rho}_{ij}$ on kuluttajaryhmien i ja j kuormien välinen keskimääräinen korrelaatio

Normalisoimalla hajonnat seuraavalla tavalla:

$$(\bar{\sigma}_1 = \bar{\sigma}_2 = \dots = \bar{\sigma}_n = 1) \quad (6)$$

Kaava 5 yksinkertaistuu muotoon:

$$\sigma^2 = n + \sum_{i=1}^k n_i (n_i - 1) \bar{\rho}_{ii} + \sum_{i=1}^k n_i \sum_{j=1}^k n_j \rho_{ij} \quad (7)$$

Jos kaikki asiakkaat kuuluvat samaan ryhmään, $k=1$, yhdistetyn kuorman varianssi saa muodon:

$$\sigma^2 = n + n(n-1) \bar{\rho} \quad (8)$$

Tämä yllä oleva hyvin yksinkertaistettu tapaus mallintaa korrelaation tärkeyttä. Mitä suurempi on kuluttajaryhmään kuuluvien asiakkaiden lukumäärä, sitä suurempi vaikutus korrelaatiolla on. Esimerkiksi 100 yksittäisen kuorman tapauksessa 0,01 korrelaatio on tarpeeksi kaksinkertaistamaan yhdistetyn kuorman varianssin. Vastaavasti 1000 yksittäisen kuorman tapauksessa vastaavaan ilmiöön riittää 0,001 korrelaatio. [4]

Tärkeää on tutkia, kuinka suuri vaikutus korrelaatiolla on käytännössä ja miten luotettavasti se voidaan määrittää. Tätä on tutkittu usean tuhannen yksittäisen kuorman tapauksessa. Korrelaatiokertoimet annetaan kuluttajaryhmittäin. Korrelaation estimaatteja arvioitaessa talvijaksoa eli vuoden päiviä 300–100 käytettiin tarkastelun pohjana. Tämä ajanjakso vastaa Pohjoismaisissa sähköjärjestelmissä huippukulutuksen aikaa. Ainoastaan tavalliset viikonpäivät sisällytettiin tarkasteluun. Taulut 2 ja 3 havainnollistavat sekä ryhmien sisäistä että ryhmien välistä korrelaatiota. Ensimmäinen sarake kertoo kuluttajaryhmän ja loput havainnollistavat korrelaation vaikutusta eri vuorokauden tunteina. Tauluissa 2 ja 3 esitetyt kuluttajaryhmät on selitetty taulussa 1. [4]

$\bar{\rho}$	7-8	9-10	11-12	13-14	15-16	17-18	19-20	21-22	23-24	1-4	5-6
110	326	279	256	248	239	177	128	152	253	337	363
120	342	342	330	351	340	256	151	132	148	234	404
220	76	50	63	75	54	42	26	23	111	92	160
602	48	13	27	36	24	12	37	35	93	59	53
712	19	10	35	40	9	33	4	44	76	92	77
810	15	38	39	57	44	56	57	16	12	72	42
820	19	62	58	53	81	38	40	56	71	69	88
910	3	42	83	58	50	27	43	61	45	55	119
920	5	47	14	34	35	86	51	11	-2	17	25

Taulu 2: Keskimääräinen korrelaatio samaan ryhmään kuuluvien kuormien välillä.

Ryhmien luokittelu on sama kuin taulussa 1 [4]

$\bar{\rho}$	7-8	9-10	11-12	13-14	15-16	17-18	19-20	21-22	23-24	1-4	5-6
110	326	279	256	248	239	177	128	152	253	337	363
120	348	320	327	320	316	281	171	184	195	221	342
220	194	173	143	136	130	119	94	86	34	159	268
602	122	52	34	24	15	36	26	34	137	140	170
712	99	79	70	49	47	-1	16	69	93	157	192
810	32	36	5	31	-22	-31	31	39	106	126	108
820	71	37	14	-18	1	-35	-12	15	19	70	122
910	90	33	26	21	36	13	-6	62	112	132	200
920	39	-8	-35	-19	0	-2	-10	12	-17	52	44

Taulu 3: Keskimääräinen korrelaatio ryhmän 110 kuormien ja muiden kuluttajaryhmien kuormien välillä. Ryhmien luokittelu on sama kuin taulussa 1 [4]

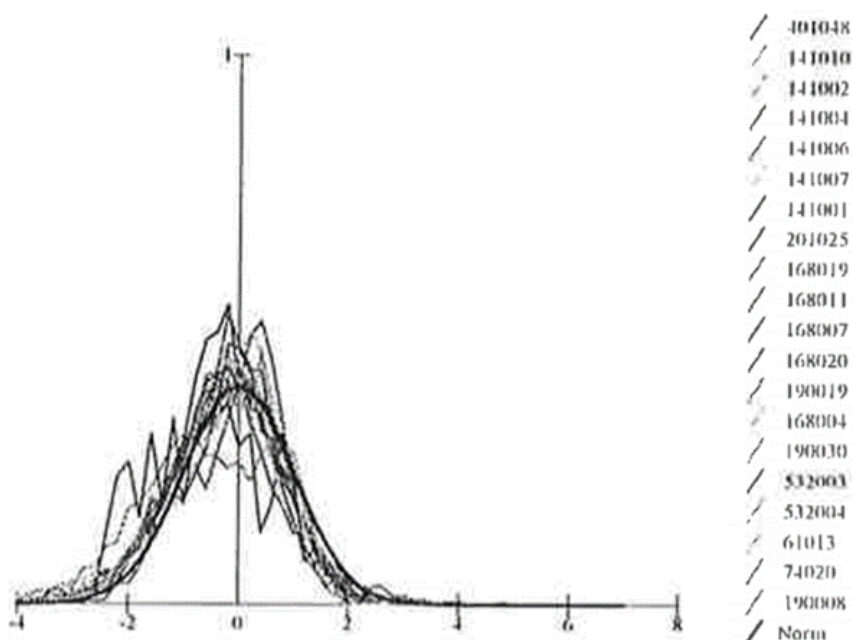
Tauluista 2 ja 3 voidaan havaita, että korrelaation merkitys on erityisen suuri varsinkin eri sähkölämmitysryhmien sisällä ja välillä. Pääasiallinen syy tähän on lämpötilariippuvuus, jota ei tutkimuksessa huomioitu kulutuksen keskiarvoja laskettaessa. Muiden kuin sähkölämmitysryhmien yhteydessä korrelaatio on pääasiassa positiivinen ja noin 0,05 -luokkaa. Olennaista on huomata, että 1000 asiakkaan esimerkkitapauksessa hajonta voi olla jopa 7-kertainen verrattuna tilanteeseen, jossa korrelaatiota ei huomioida. [4]

Korrelaation huomioimisen haasteena on se, että korrelaatio vaihtelee ajan sekä kuormitusasteen mukaan. Lisäksi on havaittu, että korrelaatioissa itsessään on myös huomattava vaihtelevuus. Tämä tarkoittaa sitä, että samaan ryhmään kuuluvat kuormat eivät ole käytökseltään kovinkaan homogeenisiä. Ongelma on haasteellisempi mitä matalammalla järjestelmätasolla eli mitä matalamman jännitetason verkossa

yhdistettyä kuormaa tarkastellaan. Tämä johtuu siitä, että yksittäisten kuormien lukumäärä on pieni verrattuna korkeampien jännitetasojen sähköjärjestelmiin. Korrelaation vaikutuksen arviointi on siis hankalaa, kun tarkasteltavien kuormien lukumäärä on pieni. [4]

Yleisin muoto tilastollista jakaumaa arvioitaessa on normaalijakauma. Valinta perustuu useimmiten keskeiseen raja-arvo lauseeseen. Tämän perusteella toisistaan riippumattomien muuttujien summajakauma lähestyy normaalijakaumaa muuttujien lukumäärän kasvaessa, riippumatta itsenäisten muuttujien alkuperäisistä tilastollisista jakaumista. Oletus pätee paremmin korkeampien tehotasojen sähköjärjestelmillä, kun taas jakeluverkon tasolla herää epäily siitä, onko yksittäisten kuormien lukumäärä tarpeeksi suuri tyydyttääkseen keskeisen raja-arvolauseen vaatimukset. [4]

Tutkimalla yksittäisten sähkökuormien vaihtelevuutta on havaittu, että ne noudattavat lognormaalista jakaumaa paremmin kuin normaalijakaumaa [5]. Alla oleva kuva 2 havainnollistaa tilannetta, jossa kuluttajaryhmään 110 kuuluvien yksittäisten asiakkaiden kuormituksen vaihtelevuuden logaritmeja verrataan normaalijakaumaan [4].



Kuva 2: Kuluttajaryhmään 110 kuuluvien asiakkaiden kuormituksen vaihtelevuuden logaritmit verrattuna normaalijakaumaan [4]

3.3 Kuormituksen laskemiseen vaikuttavat merkittävimmät tekijät

Kuormitukseen ja sen arviointiin vaikuttaa monia ulkoisia tekijöitä. Näitä ovat esimerkiksi lämpötila, kosteus, valo, ilmanpaine ja tuulen nopeus. Lämpötilalla on näistä suurin vaikutus ja nykyisissä kuormitusmalleissa ainoastaan se on huomioitu. Lämpötilan ja kuormituksen välinen suhde on hyvin monimutkainen. Tärkeimmät tekijät, jotka tulisi lämpötilariippuvuutta mallintaessa ottaa huomioon ovat lämpötilariippuvuuden epälineaarisuus, lämpötilan aikavaihtelut, lämpötilan vaikutuksen viive ja kuormituksen ohjattavuus. Nykyiset mallintamismenetelmät ovat pitkälti lineaarisia, käytännön sovelluksiin hyvin sopivia ja tämän vuoksi monet edellä mainitut seikat on jätetty huomioimatta. Lämpötilan aiheuttama tehonmuutos saadaan alla olevalla lämpötilariippuvuusmallilla. [6]

$$\Delta P(t) = c(T(t) - E[T(t)])E[P(t)] \quad (9)$$

missä

$T(t)$ on vallitseva lämpötila, joka on yleensä määritetty edellisten 24 tunnin lämpötilan keskiarvona eli vuorokauden keskilämpötilana

$E[T(t)]$ on vallitsevan lämpötilan odotusarvo eli pitkän aikavälin normaalilämpötila

c on vuodenajasta riippuva lämpötilariippuvuuskerroin

$E[P(t)]$ on kuormituksen odotusarvo, joka saadaan kuormituskäyristä

Nykyisissä malleissa lämpötilariippuvuuden on oletettu olevan lineaarinen. Lämpötilariippuvuuskerroimet on määritetty erikseen jokaiselle kuluttajaryhmälle sekä jokaiselle vuodenajalle. Kertoimien tyypillinen arvo vaihtelee 0...-6 %/°C välillä. [4]

SENERin tekemässä mittausaineiston tarkastelussa on päädytty siihen ratkaisuun, että kuormitusmallit ovat riippumattomia mitattavan kohteen sijainnista, kyseisessä sijainnissa vallitsevasta lämpötilasta, mittausvuodesta ja kyseisen vuoden lämpötilaolosuhteista. Tarkasteltavan mitatun kohteen alkuperäistä vuosienenergiaa ei myöskään ole huomioitu. Lämpötilariippuvuus on huomioitu siten, että jokaiselle

ajanjaksolle on määritetty erikseen referenssilämpötila, joka perustuu pitkällä aikavälillä tehtyihin lämpötilamittauksiin. Perusideana on se, että jos tarkasteltavalla ajanhetkellä mitattu lämpötila poikkeaa ajanjakson referenssilämpötilasta, niin tällöin laskettu teho korjataan vastaamaan ajanjakson referenssilämpötilaa. Mallissa jokaiselle kuukaudelle on määritetty erikseen ryhmäkohtaisesti lämpötilariippuvuuskerroin. Kuva 3 havainnollistaa lämpötilariippuvuusmallin normaalilämpötiloja, jotka vastaavat pitkän aikavälin lämpötilakeskiarvoja. [2]

	Kuukausi	Lämpötila °C
1	Tammikuu	-8,7
2	Helmikuu	-8,9
3	Maaliskuu	-5,4
4	Huhtikuu	1,3
5	Toukokuu	8,1
6	Kesäkuu	13,5
7	Heinäkuu	16,8
8	Elokuu	14,8
9	Syyskuu	9,6
10	Lokakuu	3,8
11	Marraskuu	-0,8
12	Joulukuu	-4,8

Kuva 3: Lämpötilariippuvuusmallissa käytettävät normaalilämpötilat [6]

Lämpötilariippuvuusmallissa käytössä olevia vuosijaksoja, jotka on määritetty vuoden alusta vuorokausina, on mallinnettu alla olevassa kuvassa 4.

1	Talvi	321 - 90
2	Kevät	91 - 180
3	Kesä	181 - 240
4	Syksy	241 - 320

Kuva 4: Lämpötilariippuvuusmallissa käytössä olevat vuosijaksot [6]

Kuten huomataan, menetelmässä ei tarvita muita lähtötietoja kuin edellisen vuorokauden keskilämpötila ja lämpötilariippuvuuskerroin, joka riippuu kuluttajaryhmän kuormituskäyrästä sekä laskenta-ajankohdasta. Näiden tietojen avulla nykymallissa saadaan kuormituskäyrällä laskettu teho korjattua vastaamaan vallitsevaa lämpötilaa. [6]

Toinen kuormituksen laskemiseen vaikuttava merkittävä tekijä on korrelaatio, jota nykyisin käytössä olevissa kuormitusmalleissa ei ole huomioitu. Yksittäisten asiakkaiden välisten kuormien korrelaatiolla on keskeinen merkitys kuormituksen huippuarvojen luotettavuusrajojen määrittämisessä. Korrelaation astetta on tarkasteltu sekä saman kuluttajaryhmän kuormien että eri kuluttajaryhmien kuormien välillä. Korrelaatiota tarkasteltiin jo edellisessä alikappaleessa 3.2. [4]

3.4 Maksimikuormituksen arvioiminen

Jakeluverkon kuormitusasteita arvioitaessa tärkein tehtävä on määrittää odotettavissa oleva huippukuormitus. Tämä saadaan seuraavasti:

$$P_{\max} = m + \sigma Q_{\alpha} \quad (10)$$

missä

m on arvioitu kulutus

σ on kuormituksen keskihajonta

Q_{α} tilastollisesta jakaumasta riippuva kerroin

Tässä tapauksessa oletetaan, että tilastollinen jakauma on normaalijakauma ja Q_{α} valitaan halutun luotettavuusrajan perusteella. Esimerkiksi tapauksissa, joissa haluttu luotettavuusraja on 95 %:a, 99 %:a tai 99,5 %:a kerroin Q_{α} on vastaavassa järjestyksessä 1,645, 2,326 tai 2,576. [4]

Kaavan 10 jokaisessa termissä on kuitenkin epätarkkuustekijöitä. Epätarkkuudet arvioidun kulutuksen keskiarvossa m johtuvat pääasiassa kolmesta syystä:

1. Virheistä tilastollisessa aineistossa, kuten vuosienergian arvioissa
2. Tyyppikuluttajaryhmiin kuulumattomat kuluttajat, joiden kulutusta ei voida mallintaa millään olemassa olevalla kuluttajaryhmällä
3. Virheet lämpötilakorjauksissa ja mittauksissa

Matalamman jännitetasojen järjestelmissä näiden virheiden merkitys voi nousta suureksi varsinkin tapauksissa, joissa tarkasteltavien yksittäisten kuormien lukumäärä on pieni. Vaikutus voidaan ottaa huomioon lisäämällä arvioidulle keskimääräiselle kulutukselle sopiva marginaali Δm . [4]

Yllä mainittujen virheiden lisäksi virheet keskihajonnassa johtuvat tilastollisista arviointivirheistä sekä tilastollisen hajonnan ja korrelaation luonnollisesta vaihtelevuudesta. Korrelaation vaihtelevuudella on suurin merkitys. Vaihtelevuuden merkitystä voidaan mallintaa lisäyksellä keskihajontaan seuraavasti. [4]

$$P_{\max} = m + \sqrt{\sigma^2 + \Delta\sigma^2} Q_{\alpha} \quad (11)$$

Tämä voidaan kirjoittaa vielä yksinkertaisemmin muodossa:

$$P_{\max} = m + c_p \sigma Q_{\alpha} \quad (12)$$

Yksinkertaisessa tapauksessa, jossa oletetaan, että kaikki asiakkaat kuuluvat samaan kuluttajaryhmään, saadaan korjauskertoimen c_p arvoksi:

$$\frac{\sqrt{\sigma^2 + \Delta\sigma^2}}{\sigma} = \sqrt{\frac{n + \sqrt{n(n-1)}\sigma_{\rho}}{n}} \approx \sqrt{1 + \sigma_{\rho}} \quad (13)$$

missä

σ_{ρ} on korrelaation keskihajonta

Edellä esitetty maksimikuormituksen arviointimenetelmä perustuu siis tyyppikuormitusmalleihin (mallista saadaan kuluttajaryhmän keskimääräinen teho) yhdessä kuormituksen keskihajonnan sekä kuluttajaryhmän sisäisten ja ryhmien välisten korrelaatiokertoimien kanssa. Kuormitusmallien epätarkkuudet otetaan huomioon lisäämällä sopivat marginaalit keskiarvoille ja korjauskertoimet hajonnoille, kuten edellä mainittiin. [4]

4. Sähkönkulutusikäyrien perinteisen muodostuksen pääperiaatteet

Kuten aiemmin jo todettiin, suurin ongelma kuormituksen arvioimisessa ja mallintamisessa on pitkään ollut tuntitehomittausten saatavuuden vähyys. Tuntitehoja mittaavia mittareita on löytynyt ainoastaan suurimmilta kuluttajilta. Tuntimittausten tuleminen pakolliseksi vuonna 2013 on mahdollistanut sen, että myös pienemmiltä kuluttajilta on alkanut kertymään tuntimitattua dataa. Perinteisissä malleissa kuormituksen arvioiminen on perustunut tuntitehomittausten sijaan kuluttajakohtaiseen arvioituun vuosienergiaan, jota on hyödynnetty kuormitusmallien kehittämisessä. Asiakaskohtainen vuosienergia on laskettu kaksiviikkojaksojen energioiden summana eli 26 kaksiviikkojakson energiat summaamalla on saatu vuosienergia määritettyä. [2]

Trimble NIS verkostolaskenta käyttää SENERin tuottamia indeksisarjoja kuluttajakohtaisen tehon määrittämisessä perinteisten Velanderin kertoimien sijaan. Näin saadaan kuluttajaryhmäkohtaiset sekä kuluttajakohtaiset kuorman ominaispiirteet ja lämpötilavaihtelevuudet huomioitua. SENERin indeksisarjoja voidaan käyttää laskennassa sellaisenaan. Verkkoyhtiöt voivat myös määrittää omat tyyppikäyränsä esimerkiksi asiakkaille, joita ei voida mallintaa riittävän tarkasti millään olemassa olevalla kuormituskäyrällä ja yhdistää nämä SENERin indeksisarjojen kanssa. [7]

Trimblen järjestelmissä tehonjakolaskennan perustana käytetään tunneittain mitattuja asiakaskohtaisia pätö- ja loistehomittauksia yhdessä indeksisarjojen kanssa. Jos käyttöpaikalta ei ole saatavilla tuntitehomittauksia, niin tällöin indeksisarjoja eli kuormituskäyriä käytetään mallintamaan käyttöpaikan kulutusta. Jos puolestaan käyttöpaikalta löytyy tuntimittauksia, mutta ei kaikille tunneille, niin niille tunneille, joilta ei löydy tuntimitattua dataa, käytetään indeksisarjoja kompensoimaan puuttuvaa mittausta. Sellaisille käyttöpaikoille, joilta löytyy ainoastaan tuntitehomittauksia pätöteholle, käytetään loistehon laskemisessa vakiota tehokerrointa. Indeksisarjoista lasketut kuluttajakohtaiset tuntitehot ja suorat tuntimittaukset yhdistetään joko tehonjakolaskennan aikana tai laskentaa ennen. Pienjänniteverkon säteisverkkolaskennan aikana yhdistetyt tehot lasketaan yhteen jakelumuuntajille.

Jakelumuuntajille summattuja tehoja käytetään puolestaan keskijänniteverkon säteisverkkolaskennan kulutuspisteinä ja laskennan perustana. [7]

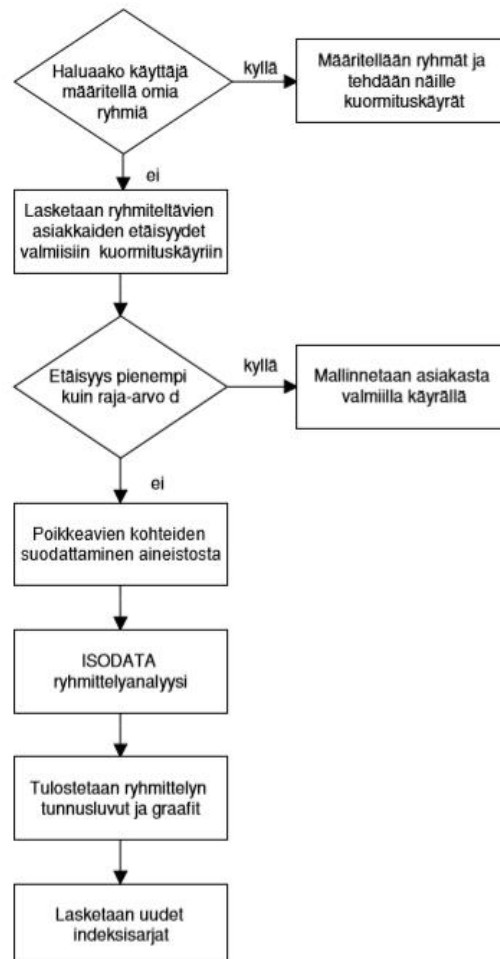
Laskennan periaatteena on Trimblen järjestelmissä siis, että ensisijaisesti tehonjakolaskennassa käytetään suoraa tuntimittauksia. Jos tarkasteltavalle tunnille ei ole saatavilla suoraa mittausta niin tällöin hyödynnetään indeksisarjoja (tietyille kuluttajaryhmälle ominaista kuormituskäyrää) yhdessä käyttöpaikan vuosienenergian kanssa. Tällöin saadaan käyttöpaikan arvioitu tietyn ajanhetken kulutus laskettua. [7]

4.1 Kuluttajaryhmien muodostaminen

Ennen indeksisarjojen muodostamista kuluttajat tulee jakaa ryhmiin. Ryhmittelyn periaatteena on luokitella sähkönkulutukseltaan samanlaiset asiakkaat samaan ryhmään. Kaikille kuluttajille ei ole tarpeellista luoda omia kuormituskäyriä, koska käyrien lukumäärästä tulisi valtava ja käsittely muodostuisi vaikeaksi niin tilastollisesta kuin tietoteknisestä näkökulmasta tarkasteltuna. Tuntitehon hajonnalle ei myöskään saada tarpeeksi tarkkoja arvioita yhden kuluttajan sähkönkulutuksen mittauksista, vaan mittauksia täytyy olla huomattavasti enemmän. [6]

Edellä mainittujen seikkojen takia on ryhmittely tarpeen. Ryhmittelyn perustana käytetään jotakin ryhmittelyalgoritmia, joka suorittaa varsinaisen ryhmiin luokittelun. Ryhmittelyanalyysi voi perustua esimerkiksi datapisteen ja klusterin keskipisteen väliseen euklidiseen etäisyyteen. Ennen kuin ryhmittely aloitetaan, on oleellista tarkastella ryhmiteltävien asiakkaiden sähkönkulutusmittausten ja olemassa olevien kuormituskäyrien yhtäläisyyttä. Perinteisissä malleissa on toimittu siten, että jos jokin näistä olemassa olevista kuormituskäyristä mallintaa asiakkaan sähkönkulutusta riittävän suurella tarkkuudella, niin tarkasteltavaa asiakasta ei ryhmitellä, vaan kyseisellä kuormituskäyrällä mallinnetaan asiakkaan kulutusta. Jos puolestaan mikään olemassa olevista kuormituskäyristä ei sovellu asiakkaan sähkönkulutuksen mallintamiseen, niin ryhmittely on tarpeen. Seuraavaksi mittausaineistosta suodatetaan poikkeavat kohteet. Poikkeavien kohteiden suodattamisen jälkeen valittua ryhmittelyalgoritmia käyttämällä suoritetaan kuluttajaryhmiin luokittelu. Ryhmittelyn tuloksena on käytetystä algoritmista riippuen tietty määrä kuluttajaryhmiä sekä näiden

tunnuslukuja. Loistehoryhmä määritetään myös jokaiselle asiakkaalle. Uudet indeksisarjat on mahdollista muodostaa tämän jälkeen. Ryhmittelyn perusidea on mallinnettu alla olevassa kuvassa 5. [6]



Kuva 5: Ryhmittelyn perusidean mallintaminen [6]

Ryhmittelyalgoritmille tulee asettaa parametrit, joihin ryhmittely perustuu. Alkuarvoina eli parametreina ryhmittelyalgoritmille voidaan algoritmista riippuen antaa esimerkiksi suurin sallittu etäisyys olemassa olevaan tyyppikäyrään, ryhmien haluttu lukumäärä tai pätötehon kuukausivaihtelu. Annetut parametrit toimivat ryhmittelyanalyysin perustana. [6]

Suurimmassa osassa nykyisin käytössä olevissa malleissa asiakastietojärjestelmästä tuleva tieto määrittää pitkälti, mitä kuormituskäyrää käytetään mallintamaan asiakkaan

kulutusta. Tätä ennen tulee määrittää kuitenkin kriteerit, joihin ryhmittely perustuu. Usein ryhmittelyn perustana käytettyjä kriteereitä ovat tariffit, vuosienergian määrä ja yö- ja päiväsiähkön suhteellisten osuuksien määrät. Tässä työssä käytetty menetelmä esitellään seuraavassa luvussa 5. Eri sektorien ryhmiin luokittelu on osoittautunut haasteelliseksi johtuen muun muassa liiallisesta yksinkertaistamisesta. Esimerkiksi vain muutamassa tapauksessa on käytetty perustana huipputehon käyttöaika. Maatalouden erilaisuutta on myös hankala mallintaa, koska SENERin kuormituskäyrät on luotu varsin vähäisestä mittausmateriaalista. Laskennan tarkentumisen ja onnistumisen kannalta ryhmittelyllä on erittäin suuri vaikutus. Nykyisissä malleissa ryhmittelyn osittainen epäonnistuminen johtaa suoraan epätarkkoihin laskennan tuloksiin. [7]

Ryhmittelyn perusteella muodostuneet kuluttajaryhmät sisältävät eri kuluttajaluokkia. Esimerkiksi asuminen on kuluttajaryhmä, joka sisältää seuraavat kuluttajaluokat: kerrostaloasunnot, rivitaloasunnot, pientaloasunnot, yhteiset tilat, peltoviljely ja vapaa-ajan asunnot. Toinen kuluttajaryhmä on esimerkiksi julkiset palvelut, johon kuuluu kuluttajaluokat julkinen hallinto, koulutus, järjestö ja uskonnollinen toiminta, ulkomaan edustustot, terveydenhuolto- ja sosiaalipalvelut sekä rahoitus- ja vakuutustoiminta. [2]

4.2 SENERin indeksisarjat

Perinteiset jakeluverkon kuormituksen arviointiin käytettävät kuormituskäyrät ovat peräisin vuosilta 1989 ja 1992. Nämä ovat Sähköenergioliiton useiden vuosien kehitystyön tuloksena toteutuneet kuluttajien sähkökuormitusta arvioivat tyypikäyrät, joita käytetään muun muassa tehonjakolaskennassa sekä KAH-laskennassa. Vuoden 1989 indeksisarjoissa on käytetty taulun 4 mukaista ryhmiin luokittelua mallintamaan kokonaissähkönkulutusta. [7]

1. Yksivuoroteollisuus	2. Kaksivuoroteollisuus
3. Yhdistetty teollisuus	4. Julkinen palvelu
5. Yksityinen palvelu	6. Yhdistetty palvelu
7. Pientalo- ja rivitaloasunnot	8. Maatalous
9. Kerrostaloasunnot	10. Yhdistetyt asunnot
11. Huonekohtainen sähkölämmitys ja asuminen	12. Osittain varaava sähkölämmitys ja asuminen
13. Varaava sähkölämmitys ja asuminen	14. Yhdistetty sähkölämmitys ja asuminen
15. Huonekohtainen sähkölämmitys	16. Osittain varaava sähkölämmitys
17. Varaava sähkölämmitys	18. Yhdistetty sähkölämmitys

Taulu 4: Vuoden 1989 kuluttajaryhmiin luokittelu

Viimeisimmät indeksisarjat julkaistiin vuonna 1992 ja ne sisältävät sekä yllä mainitut kokonaissähkökäytön ryhmittelyt että nykyisin käytössä olevat kaikki 46 eri tyyppikuluttajaryhmää. Taulussa 5 on esitetty kaikki 46 tyyppikuluttajaryhmää. [7]

<ul style="list-style-type: none"> ▪ Pientaloasunnot, huonekohtainen sähkölämmitys <ol style="list-style-type: none"> 1. 000110 käyttövesivaraaja < 300 l 2. 000120 käyttövesivaraaja 300 l 3. 000130 lattialämmitys > 2 kW 	<ul style="list-style-type: none"> ▪ Pientaloasunnot, osittain varaava sähkölämmitys <ol style="list-style-type: none"> 4. 000210 lyhyet sulkemisajat 5. 000220 pitkät sulkemisajat
--	---

<ul style="list-style-type: none"> ▪ Pientaloasunnot 6. 000300 varaava sähkölämmitys 	<ul style="list-style-type: none"> ▪ Pientaloasunnot 7. 000400 lämpöpumppulämmitys
<ul style="list-style-type: none"> ▪ Pientaloasunnot, tuplalämmitys 8. 000510 kertaluontainen korko 9. 000520 kaksinkertainen korko 10. 000530 kausiluontainen korko 	<ul style="list-style-type: none"> ▪ Pientaloasunnot, ei sähkölämmitystä 11. 000601 ei sähkökiuasta 12. 000602 sähkökiuas
<ul style="list-style-type: none"> ▪ Rivitalo- ja kerrostaloasunnot, ei sähkölämmitystä 13. 000611 ei sähkökiuasta 14. 000612 sähkökiuas 	<ul style="list-style-type: none"> ▪ Kerrostalo (kokonaan) 15. 001010 ilman asuntojen osuutta 16. 001020 asunnot mukaan lukien
<ul style="list-style-type: none"> ▪ Rivitalo (kokonaan) 17. 001030 huonekohtainen sähkölämmitys 18. 001120 loma-asuntoalue, jakeluverkon alue 	<ul style="list-style-type: none"> ▪ Maatalous, kasvituotanto 19. 000732 asuminen sisälletynä 20. 000733 asuminen, sähkökiuas ▪ Maatalous, eläinmaatalous ja maitotalous 21. 000711 asuminen pois luettuna 22. 000712 asuminen sisälletynä 23. 000713 asuminen, sähkökiuas 24. 000714 asuminen, sähkökiuas ja sähkölämmitys
<ul style="list-style-type: none"> ▪ Maatalous, lihatuotanto 25. 000721 asuminen pois lukien 26. 000722 asuminen sisälletynä 	<ul style="list-style-type: none"> ▪ Teollisuus 27. 810430 yksivuoroteollisuus, tekstiilien, vaatteiden ja nahkatuotteiden valmistus 28. 810440 yksivuoroteollisuus, puutavaran valmistus 29. 810452 yksivuoroteollisuus, paperituotteiden valmistus, graafinen tuotanto

	<p>30. 810460 yksivuoroteollisuus, kemikaalien, maaöljyn, kumin ja muovituotteiden valmistus</p> <p>31. 810480 yksivuoroteollisuus, metalli- ja konetuotteiden valmistus</p> <p>32. 820420 kaksivuoroteollisuus, ruuan, juomien ja tupakan valmistus</p> <p>33. 820430 kaksivuoroteollisuus, tekstiilien, vaatteiden ja nahkatuotteiden valmistus</p> <p>34. 820460 kaksivuoroteollisuus, kemikaalien, maaöljyn, kumin ja muovituotteiden valmistus</p>
<ul style="list-style-type: none"> ▪ Tuotteet <p>35. 820480 kaksivuoroteollisuus, metalli- ja konetuotteiden valmistus</p>	<ul style="list-style-type: none"> ▪ Palvelut, julkinen <p>36. 910810 yleinen ja muu hallinto</p> <p>37. 910820 julkiset ohjeet ja harjoitteet</p> <p>38. 910830 sairaalat ja terveydenhuolto</p>
<ul style="list-style-type: none"> ▪ Palvelut, yksityinen <p>39. 920610 tukkukauppa</p> <p>40. 920622 tavaratalot ja supermarketit</p> <p>41. 920623 muut vähittäiskaupat</p> <p>42. 920660 rahoituslaitokset ja vakuutustoiminta</p>	<p>43. 920630 automyynti ja palveluaseman toiminnot</p> <p>44. 920640 hotellit ja majoitustoiminta</p> <p>45. 920650 ravintolat ja kahvilat</p> <p>46. 920670 virkistys- ja kulttuuripalvelutoiminta</p>

Taulu 5: SENERin indeksisarjojen tyyppikuluttajaryhmät

4.3 Indeksisarjojen muodostus

Ryhmittelyanalyysin tulosten ja kuluttajien ryhmittelyn jälkeen voidaan indeksisarjat muodostaa. Indeksisarjojen muodostamisessa on perinteisissä menetelmissä käytetty pitkältä aikaväliltä kerättyjä mittauksia ja arvioita kuluttajien vuosienergioista.

4.3.1 Sähkönkulutuksen indeksisarjat

Indeksisarjat, kuten on jo todettu, on yleinen tapa mallintaa asiakkaiden sähkönkulutusta. Tunnittainen tietyn kuluttajaryhmän k tietyllä vuoden ajanhetkellä i tarvitsema keskimääräinen teho $P_{m,ki}$ ennustetaan käyttämällä seuraavaa kaavaa. [7]

$$P_{m,ki} = \frac{W_k}{8760} \times \frac{M_{k,i}}{100} \times \frac{m_{k,i}}{100} \quad (14)$$

missä

W_k on kuluttajaryhmään k kuuluvan yksittäisen asiakkaan vuosienergia

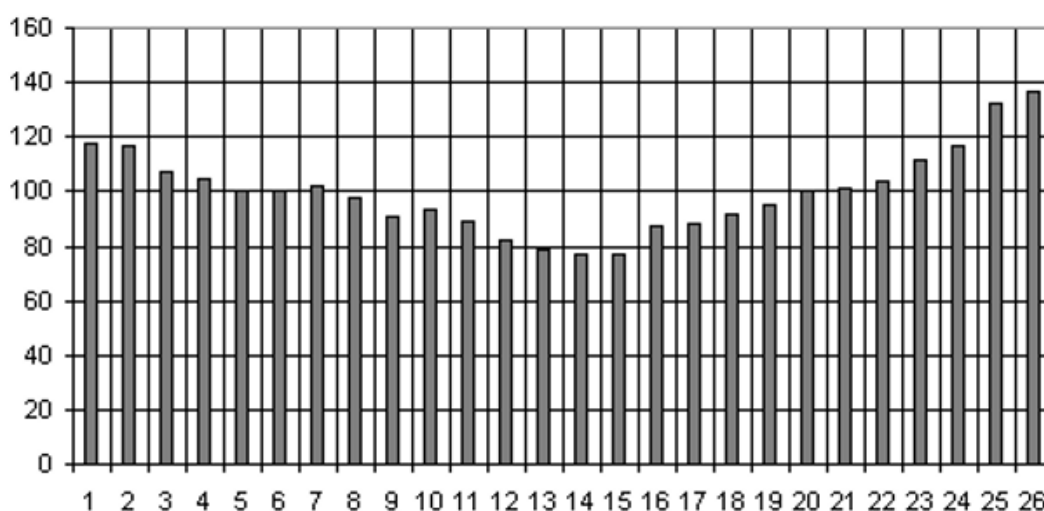
$M_{k,i}$ on kuluttajaryhmän k kaksiviikkojakson i ulkoinen indeksi (keskiarvo 100)

$m_{k,i}$ on kuluttajaryhmän k kaksiviikkojakson i sisäinen indeksi (keskiarvo 100)

Ulkoiset indeksit määritetään kuluttajaryhmäkohtaisesti. Ne kuvaavat kulutuksen kausiluontaista vaihtelua vuoden aikana. Jokaista kuluttajaryhmää kohden on 26 ulkoista indeksiä, joista jokainen kuvaa yhtä kaksiviikkojaksoa. Vuoden ensimmäinen kaksiviikkojakso alkaa tammikuun ensimmäinen päivä kello 7.00 ja loppuu tammikuun 15. päivä kello 7.00. Sen kaksiviikkojakson, joka sisältää karkauspäivän, pituus on 15 päivää eli karkausvuoden viidennen kaksiviikkojakson pituus on 15 päivää. Myös vuoden viimeisen kaksiviikkojakson pituus on 15 päivää, mutta kaikkien muiden kaksiviikkojaksojen pituus on 14 päivää. Ulkoisista indekseistä muodostunut kuvaaja

ilmaisee kaksiviikkojakson keskimääräisen tehon suhteessa koko vuoden keskimääräiseen tehoon. [7]

Kuluttajaryhmän ulkoiset indeksit saadaan määritettyä laskemalla kullekin kaksiviikkojaksolle energia eli laskemalla tuntitehot yhteen kahden viikon ajalta. Trimblen järjestelmissä ulkoiset indeksit skaalataan vielä siten, että niiden keskiarvo on 100 ja summa $26 \times 100 = 2600$. Jos oletetaan, että ryhmän vuosenergia on X (kaikkien 26 kaksiviikkojakson energioiden summa) ja tämän vuosieron haluttaisiin olevan 2600, niin tällöin kaikkien kaksiviikkojaksojen energiat kerrotaan kertoimella $2600/X$. Alla oleva kuva 6 havainnollistaa kuluttajaryhmän 000611 (rivitalot ja kerrostalot, joissa ei ole sähkölämmitystä eikä sähkökuuasta) kaksiviikkojaksojen ulkoisia indeksejä. [7]

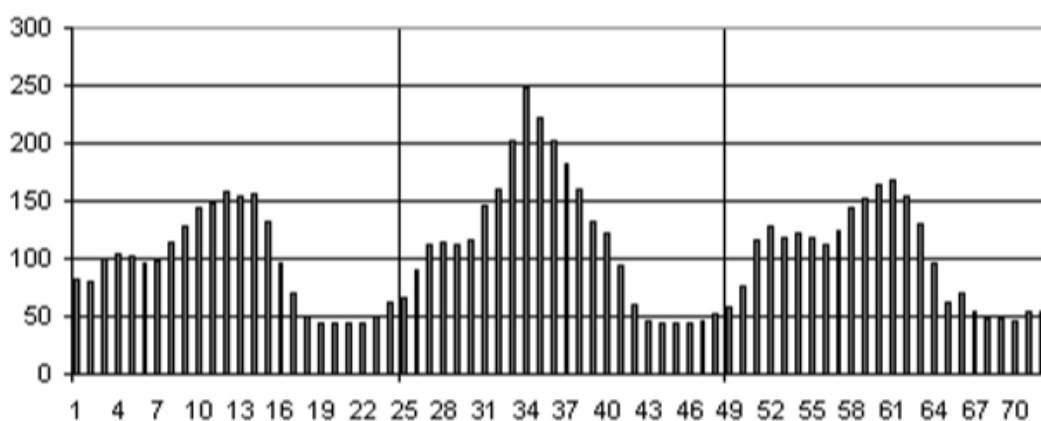


Kuva 6: Kuluttajaryhmän 000611 ulkoiset indeksit [7]

Sisäiset indeksit kuvaavat puolestaan päivän tunnittaisten tehojen vaihtelevuutta suhteessa ajan kausiluontaiseen vaihtelevuuteen. Tunnittaista vaihtelua kuvataan 24 tunti-indeksillä erikseen arkipäiville, aatoille ja pyhillä. Yleisesti aatoilla tarkoitetaan lauantapäiviä ja pyhillä sunnuntapäiviä. Lauantaita ja sunnuntaita mallinnetaan erikseen, koska näinä päivinä asiakkaiden sähkönkulutus poikkeaa merkittävästi arkipäivien kulutuksesta. Viikoppäivien mukaan jaoteltuina arki-indeksit kuvaavat tehon keskimääräistä tunnittaista vaihtelua yön ja päivän aikana maanantai kello

7.00:stä lauantai kello 7.00:ään asti, aattoindeksit lauantai kello 7.00:stä sunnuntai kello 7.00:ään asti ja pyhäindeksit sunnuntai kello 7.00:stä maanantai kello 7.00:ään asti. Trimblen järjestelmissä verkkoyhtiön pääkäyttäjä määrittää mitkä päivät lasketaan pyhäpäiviksi sunnuntaiden lisäksi. [7]

Sisäiset indeksit saadaan laskemalla kullekin tunnille keskimääräinen teho tietyllä kaksiviikkojaksolla. Tuntitehot skaalataan Trimblen järjestelmissä sellaisiksi, että niiden keskiarvo on 100 ja summa $14 \times 24 \times 100 = 33600$ tai $15 \times 24 \times 100 = 36000$ riippuen kaksiviikkojakson päivien lukumäärästä. Skaalauksen jälkeen arkipäiville, aattopäiville ja pyhäpäiville lasketaan omat keskiarvonsa. Jos kaksiviikkojaksolla ei ole yhtään erikoispyhää eli arkipyhää (pituus 14 päivää), niin tällöin arkipäiviä on 10, aattoja 2 ja pyhiä 2. Alla oleva kuva 7 mallintaa kuluttajaryhmän 000611 vuoden viimeisen kaksiviikkojakson sisäisiä indeksejä. [7]



Kuva 7: Kuluttajaryhmän 000611 kaksiviikkojakson 26 sisäiset indeksit [7]

4.3.2 Indeksisarjojen lämpötilakorjaus (mitatun aikasarjan normalisointi vai mitatun energian normalisointi)

Kuormituskäyrien perinteisissä muodostamismenetelmissä lämpötilakorjaus on huomioitu siten, että kuormituskäyrien parametrit on määritetty vastaamaan sähkönkulutusta kuukaudesta riippuvassa normaalilämpötilassa. Normaalilämpötilat on määritetty käyttämällä tarkasteltavan ajanhetken pitkän aikavälin keskimääräisiä lämpötiloja. Indeksisarjoilla kuormitusta arvioitaessa vuosienergiat on normalisoitava vastaamaan pitkän aikavälin lämpötiloja eli ne tulee lämpötilakorjata. Vuosienergiat

toimivat pitkälti perinteisissä menetelmissä kuormituksen arvioinnin lähtökohtana. Lämpötilakorjatulla vuosienenergialla voidaan laskea tiettyä ajanhetkeä vastaava lämpötilakorjattu teho. Näin toimitaan virheellisten laskentatuloksien välttämiseksi, joita syntyy laskettaessa tehoja suoraan mitatuilla vuosienenergioilla. Virhemarginaalin suuruus riippuu tarkasteltavan vuoden lämpötilaolosuhteista. Varsinkin mittausvuosina, jotka ovat olleet harvinaisen kylmiä tai lämpimiä laskentavuoden lämpötilaolosuhteisiin nähden, voi virhemarginaali olla hyvinkin suuri. Tämän vuoksi normalisoimattomia vuosienenergioita ei tule käyttää. Tarkasteltavan kulutuspuheen lämpötilariippuvuus tulee tuntea, jotta vuosienenergia voidaan normalisoida. [6]

Asiakkaan kulutuksen lämpötilariippuvuuden huomioonmissa on käytetty yksinkertaista lineaarista mallia. Lämpötilariippuvuusmalli kehitettiin niin kutsutun Ryhmä 8 toimesta, joka koostui kahdeksasta suomalaisesta sähköverkkoyhtiöstä. Tässä mallissa, kuten jo aliluvussa 3.3 todettiin, vuosi on jaettu neljään jaksoon, joista jokaisella on omat parametrinsa lämpötilariippuvuudelle. Lämpötilariippuvuuden vuorokauden sisäisiä vaihteluita ei mallissa ole huomioitu. Lämpötilariippuvuuden mallintaminen riippuu olennaisesti alueesta. Suomi on jaettu 10 alueeseen. Alla oleva kuva 8 havainnollistaa alueellisia säähavaintoasemia. [7]

- 0 Average
- 1 Helsinki-Vantaa
- 2 Turku
- 3 Tampere
- 4 Lappeenranta
- 5 Jyväskylä
- 6 Vaasa
- 7 Joensuu
- 8 Kajaani
- 9 Oulu
- 10 Sodankylä

Kuva 8: Alueelliset säähavaintoasemat [7]

Lämpötilan muuttuminen vaikuttaa kuormitukseen tietyllä viiveellä. Tämä huomioidaan lämpötilariippuvuusmallissa käyttämällä aina edellisen vuorokauden keskilämpötilaa tarkasteltavan vuorokauden keskilämpötilan sijaan. Kuormien

saturaation takia kyseistä mallia käytetään ainoastaan lämpötiloille $-15\dots+15\text{ °C}$ välillä. Jos edellisen vuorokauden keskilämpötila poikkeaa enemmän kuukauden keskilämpötilasta, käytetään lämpötilana -15 °C tai $+15\text{ °C}$ tilanteesta riippuen. Kaava 15 kuvaa jo aikaisemmin aliluvussa 3.3 esitettyä lämpötilariippuvuusmallia. [6] [7]

$$\Delta P = c(T - E[T])E[P] \quad (15)$$

missä

ΔP on lämpötilan aiheuttama tehonmuutos

T on vallitseva lämpötila, joka on yleensä määritetty edellisten 24 tunnin lämpötilan keskiarvona

$E[T]$ on vallitsevan lämpötilan odotusarvo

c on vuodenajasta riippuva lämpötilariippuvuuskerroin

$E[P]$ on kuormituksen odotusarvo, joka saadaan kuormituskäyristä

Tietyn tunnin t teho saadaan alla olevasta kaavasta:

$$P = (1 + c\Delta T)E[P] \quad (16)$$

missä

$$\Delta T = T - E[T]$$

Edellä esitetty lineaarinen malli ei tarvitse lähtöarvoikseen kuin edellisen vuorokauden keskilämpötilan ja lämpötilariippuvuuskerroimen. Näiden avulla lämpötilakorjaus tarkasteltavan hetken referenssilämpötilaan saadaan tehtyä. [6]

Suurimmalle osalle asiakkaista on käytetty perinteisissä malleissa mitatun energian normalisointia siitä syystä, että tuntimittauksia ei ole ollut saatavilla kuin suurimmille kuluttajille. Tämän vuoksi ei ole voitu määrittää yksilöllisiä lämpötilariippuvuuskertoimia. Vuosienergian normalisointi perustuu olemassa olevaan kuormituskäyrään ja tämän lämpötilariippuvuuskertoimiin, kuten edellä on jo mainittu.

Vuosienergian normalisoinnissa lasketaan ensin kuormituskäyrällä tarkasteltavan aikavälin jokaiselle tunnille tehot normalisoidun vuosienergian avulla. Kun oletetaan vuoden keskimääräiseksi tehoksi yksi, saadaan normalisoitu vuosienergia määritettyä. Normalisoitujen tehojen määrittämisen jälkeen voidaan laskea energia tietylle aikavälille. Seuraavaksi normalisoidut tehot täytyy lämpötilakorjata vastaamaan tarkasteltavan ajanhetken toteutunutta lämpötilaa. Tämän jälkeen lämpötilakorjattuja tehoja käyttäen lasketaan edellistä vaihetta vastaavan saman aikavälin energia. Edellisten vaiheiden jälkeen saadaan mitatun energian ja kuormituskäyrästä lasketun lämpötilakorjatun energian energioiden suhde tarkasteltavalla aikavälillä. Lämpötilakorjausvirheitä eikä mallinnusvirheitä huomioida. Energioiden suhteen määrittää ainoastaan energiaero. Kun sekä normalisoitujen tehojen avulla laskettu aikavälin energia että lämpötilakorjattu aikavälin energia kerrotaan energioiden suhteella, saadaan tuloksena normalisoidut tehot sekä todelliset tehot aikavälin jokaiselle tunnille. Normalisoidun aikasarjan perusteella pystytään määrittämään tiettyä aikaväliä vastaava normalisoitu energia. Myös aikavälin lämpötilakorjatun energian suhde normalisoituun energiaan nähden saadaan määritettyä. Normalisoinnin perustuminen kuormituskäyriin mahdollistaa normalisoidun energian laskemisen mitattua aikaväliä pidemmälle ajalle. [6]

Edellä mainittua vuosienergian normalisointia tarkempi tapa mallintaa lämpötilariippuvuutta on mitatun aikasarjan normalisointi. Menetelmän edellytyksenä on, että kulutuspuheen tuntitehot ja lämpötilat tiedetään koko vuodelle. Tällöin voidaan määrittää yksilölliset lämpötilariippuvuuskertoimet ja normalisoida aikasarjat näiden avulla. Johtuen siitä, että jokaisen tunnin todellinen teho tiedetään, saadaan mitatun energian normalisointia tarkemmin lämpötilariippuvat tehon osat selville. [6]

Mitatun aikasarjan normalisoinnissa lähdetään liikkeelle aikasarjan lämpötilariippuvuuskertoimien laskennasta. Koska jokaisen tunnin teho tiedetään ja tätä tuntitehomittausta vastaava lämpötila, niin saadaan yksilölliset lämpötilariippuvuuskertoimet laskettua. Lämpötilariippuvuuskerroin on muotoa $\%/^{\circ}\text{C}$. Lämpötilariippuvuuskertoimien määrittämisessä voidaan käyttää myös vuorokauden keskilämpötiloja tai valmiiksi määritettyjä kertoimia. Edellisen vaiheen jälkeen jokaisen tunnin teho normalisoidaan laskettuja lämpötilariippuvuuskertoimia käyttäen. Tämä tehdään edellä esitetyn kaavan 16 avulla, mutta käänteisesti eli poistamalla tuntitehosta lämpötilan vaikutus. Normalisoitu energia voidaan määrittää tarkasteltavalle aikavälille normalisoidun aikasarjan avulla. Kun lasketaan yhteen energioita, tulee kiinnittää huomiota siihen, että kesäajan lopussa ilmenevä ylimääräinen tunti huomioidaan. Tämä täytyy ottaa summaukseen mukaan. Normalisoidun ja normalisoimattoman aikasarjan energioiden suhde saadaan myös määritettyä laskettujen energioiden avulla. [6]

4.4 Perinteisten sähkönkulutuskäyrien virheet ja ongelmat

Nykyisin vielä pitkälti kuormituksen arvioinnissa ja laskemisessa käytettävät kuormituskäyrät ovat lähtöisin vuodelta 1992, jolloin Sähköenergialiitto julkaisi kuormitustutkimuksensa [8]. Kuluttajien sähkön käyttötottumukset ovat muuttuneet hyvin paljon näistä ajoista, mikä on luonut haasteen perinteisten kuormituskäyrien käytölle. Myös käytössä olevat sähkölaitteet sekä mittausten tekemisessä käytettävät välineet ja analysointimahdollisuudet ovat muuttuneet. Oleellista on huomata, että jo 1980-luvulla tehtyjä mittauksia on käytetty alkuperäisten kuormituskäyrien muodostamisessa. Mittaukset ovat siis vanhoja ja niiden perusteella tehdyt kuormituskäyrät eivät enää kuvaa kuluttajien sähkönkäyttöä yhtä tarkasti kuin ennen. [6]

Perinteisten kuormitusmallien virheet johtuvat pääasiassa kolmesta tekijästä, jotka ovat taustatietojen virheellisyys, tyyppikuluttajaryhmiin kuulumattomat kuluttajat sekä lämpötilamittausten ja lämpötilariippuvuuksien aiheuttamat virheet. Yleisin syy taustatietojen virheellisyyteen on vanha vuosenergia tai se, että kulutuspisteeltä ei löydy vuosenergiaa. Myöskään muuntajan syöttämän liittymän takana olevat

kuluttajat eivät ole aina selvillä. Toinen ongelma aiheutuu kuluttajista, joita on hankala mallintaa millään olemassa olevan kuluttajaryhmän tyyppikäyrällä. Nämä kuluttajat ovat useimmiten suuria kuluttajia, kuten palvelulaitoksia tai teollisuuden yrityksiä. Näiltä suurilta kuluttajilta löytyy sähkömarkkinoiden vapautumisen johdosta omat mittauksensa, jotka auttavat kyseisten asiakkaiden kulutuksen mallintamisessa merkittävästi. Lämpötilariippuvuutta on mallinnettu perinteisissä kuormituksen arvioinnin menetelmissä lineaarisella mallilla, vaikka lämpötila on epälineaarinen suure. Lämpötilamittausten ja lämpötilariippuvuuksien virheellisyydestä johtuvat epätarkkuudet vaikuttavat suurimmin sähkölämmitys asiakkaiden kulutuksen mallintamiseen. Esimerkiksi varaavien lämmitysjärjestelmien yhteydessä perinteinen lämpötilariippuvuusmalli ei anna tarkkoja tuloksia. Tämä johtuu pääosin siitä, että varaajat käyttävät energiaa suurimmilta osin ainoastaan yöaikaan asunnon lämmittämiseen. Varsinkin kun käytetään kuormitusmallia suurella lämpötilavälillä, tulee huolellisesti tarkastella lämpötilariippuvuuden epälineaarista käytöstä. Syytä olisi tällöin mallintaa lämpötilariippuvuutta esimerkiksi sopivalla epälineaarilla polynomilla. [2] [6]

Oleellista on myös huomata, että nykyisin vielä yleisesti käytössä olevat SENERin kuormituskäyrät käyttävät lämpötilariippuvuuden mallintamiseen ainoastaan tammikuun lämpötilariippuvuuskertoimia. Näitä kertoimia käytetään virheellisesti koko vuodelle. Muuten lämpötilariippuvuuskertoimia ei ole huomioitu kuormituksen arvioinnissa. [9]

Sähköenergialiiton luomat kuormituskäyrät ovat valtakunnalliset. Sähkönkulutuksessa ja käyttötottumuksissa on kuitenkin huomattavia alueellisia eroja. Tämän vuoksi kuormitusmalleja on oleellista muokata paikallisten olosuhteiden mukaan. Perinteisiä kuormitusmalleja määritettäessä mittauksia tehtiin ympäri Suomen. Varsinkin alueilla, jotka poikkeavat merkittävästi olosuhteiltaan, on tarve lisämittauksille suuri. Alueelliseen tutkimukseen siirtyminen valtakunnallisen tutkimuksen sijaan olisi paras keino kuormituskäyrien tarkemman käytettävyyden kannalta tarkasteltuna [8]. Tämän takia muun muassa tässä työssä keskitytään verkkoyhtiökohtaiseen alueelliseen mallintamiseen. [2]

4.5 Ajantasaisempien sähkönkulutuskäyrien muodostamisen edut

Kuten aiemmin jo todettiin, uusien kuormituskäyrien luominen auttaisi mallintamaan tarkemmin nykyisiä sähkön käyttötottumuksia sekä huomioimaan paremmin nykyaikana käytössä olevat sähkölaitteet ja mittausmenetelmät. Teknologian kehittymisen myötä vanhat mallit tulevat yhä epätarkemmiksi.

Ajantasaisilla ja uudemmilla sähkönkulutusta mallintavilla tyypikohtaisilla kuormituskäyrillä saavutettaisiin useita hyötyjä. Ensinnäkin kuormituskäyriin perustuvan tehonjakolaskennan antamat tulokset tarkentuisivat ja laskennan virhemarginaalin suuruus pienenesi. Tällä hetkellä Trimblen järjestelmissä tehonjakolaskennan antamissa tuloksissa on epätarkkuuksia vanhojen kuormituskäyrien käytön takia. Uusien kuormituskäyrien luominen mahdollistaisi myös eri skenaariotarkastelut. Voitaisiin luoda kuormituskäyrät erikseen maksimiarvoja, minimiarvoja ja keskimääräisiä arvoja käyttäen. Näin saataisiin laajennettua laskennan vaihtoehtoja. Nykyiset kuormituskäyrät perustuvat ainoastaan keskimääräisiin arvoihin. Eri skenaariotarkastelut loisivat Trimblen järjestelmiä käyttäville verkkoyhtiöille uusia mahdollisuuksia tutkia verkon kuormitusta ja laskea verkon osien mitoitusta eri vaihtoehtoja käyttäen.

KAH-laskenta on toinen kuormituskäyriin perustuva laskennan muoto. Ajantasaisemmat kuormituskäyrät mahdollistaisivat, että KAH voitaisiin laskea tarkemmilla ennustetuilla kuormilla. Tällöin Trimble DMS:n käytönaikainen KAH-laskenta vastaisi paremmin todellisuutta. Lisäksi ajantasaisemmat kuormitusennusteet tukisivat myös muita Trimble NIS, Optimizer ja Trimble DMS toimintoja.

5. Tuntimitatun datan hyödyntäminen uusien kuormituskäyrien luomiseksi

Tämän työn tutkimusosuuden tarkoituksena on selvittää, miten pelkkään tuntimitattuun aineistoon perustuva asiakkaiden klusterointi eli ryhmittely kuluttajaryhmiin ja indeksisarjojen muodostus saaduille ryhmille onnistuu. Tavoitteena on saada toisistaan selvästi sähkönkulutukseltaan eroavia ryhmiä. Tutkimus tehdään käyttöpaikkatasolla liittymätarkastelun sijaan. Käyttöpaikka- ja liittymätarkastelua on tutkinut muun muassa Matti Koivisto vuonna 2013. Liittymien kulutuksen tasaisemman käyttäytymisen vuoksi Koivisto päätyi käyttämään klusteroinnissa liittymiä käyttöpaikkojen sijaan. Tällöin saadut tulokset olivat helpommin tulkittavissa käyttöpaikkatarkasteluun verrattuna. [10]

Poiketen useimpiin aikaisempiin kuormitustutkimuksiin, pelkkään tuntimitattuun dataan perustuva indeksisarjojen muodostus ei käytä mitään muuta kuin sähkönkulutuksen tuntimittauksia lähtötietonaan. Useimmissa kuormitustutkimuksissa on lähtötietona käytetty olemassa olevia asiakastyypitietoja, jolloin karkeampi ryhmittely on jo ollut tiedossa etukäteen. Tällöin on välttytty siltä, että klusteroinnin tuloksena muodostuneeseen kuluttajaryhmään tulisi lämmitysmuodoltaan paljon erilaisia asuntoja. Tiedot rakennusten lämmitysmuodoista on haettu useimmiten rakennuskannasta koordinaattien perusteella.

Tässä työssä kiinnostuksen kohteena ovat kuitenkin yksittäisten asiakkaiden tuntimittausarjat ja kulutustottumukset eikä se mikä lämmitysmuoto asunnossa on käytössä. Olemassa oleviin asiakastyypitietoihin ei siis kiinnitetä huomiota. Tämän seurauksena, kuten aiemmin jo todettiin, samaan ryhmään voi tulla asiakastyypitiedoiltaan erilaisia asiakkaita. Asunnon lämmitysmuoto ja asiakkaan kulutustottumukset voivat olla myös ajan saatossa muuttuneet. Tämän vuoksi tuntimitattuun aineistoon pohjautuva tarkastelu voi tuottaa uutta tietoa vanhemman tiedon tilalle ja päivittää asiakkaan profiilin kuulumaan paremmin tälle sopivaan kuluttajaryhmään.

Tässä työssä on käytetty MATLAB -ohjelmistoa datan esiprosessointiin, klusterointiin ja indeksisarjojen muodostamiseen.

5.1 Verkkoyhtiöltä saatu aineisto

Verkkoyhtiö Kajave Oy:ltä saatiin sähkönkulutuksen 3,5 vuoden tuntimitattu aineisto käyttöön tutkimusta varten. Saatua aineistoa käsiteltiin luottamuksellisesti ja ainoastaan tämän työn tutkimustarkoitusta varten. Kajaven sähköverkko sijaitsee Kainuun alueella ja yhteensä verkon pituus on noin 13 380 km. Kajaven verkkovastuu alue käsittää koko Suomen pinta-alasta noin 7,7 %:a ja käyttöpaikkojen lukumäärä on noin 50 000. [11]

Tuntimitattu aineisto haettiin tietokannasta ja huomattiin, että osalta käyttöpaikoista puuttuivat vuoden 2018 lokakuun mittaukset kokonaan. Vuoden 2019 mittauksia ei otettu tarkasteluun mukaan sen vuoksi, että kaikilta käyttöpaikoilta puuttuivat vähintään huhtikuun ja toukokuun mittaukset. Virheet mittausaineistossa johtuvat luultavasti tietokanta-ajon aikana tapahtuneista virheistä, joiden vuoksi osa datasta ei ole koskaan päässyt kantaan asti. Tutkimusta varten käytettäväksi aineistoksi valittiin vuosien 2017 ja 2018 tuntimittaukset, koska näille vuosille löytyi suurimmilta osin kaikkien kuukausien mittaukset. Tutkimuksen pohjatietona toimivia mittauksia oli siis aikaväliltä 1.1.2017–31.12.2018.

Tarkasteltavien käyttöpaikkojen lukumääräksi ei valittu kaikkia alueen käyttöpaikkoja, koska kuluttajaryhmien erottamiseksi toisistaan riittää pienempi otos koko aineistosta eikä datan koko määrän käyttäminen merkittävästi vaikuta klusteroinnin tuloksiin. Käyttöpaikkojen lukumääräksi valikoitui aluksi 4938 asiakasta. Klusteroinnin tuloksista ei kuitenkaan voitu tunnistaa selvästi kaikkia saatuja asiakasryhmiä, joten lopullisena käyttöpaikkojen lukumääränä käytettiin pienempää 800 käyttöpaikan otosta. Tämän määrän oletettiin riittävän sähkönkulutukseltaan erilaisten ryhmien löytämiseen. Keskihajonnalle saadaan riittävän tarkat estimaatit, kun käyttöpaikkojen lukumäärä kuluttajaryhmässä on vähintään 10–20 välillä. Seuraavissa luvuissa tapauksella 1 tarkoitetaan tapausta, jossa tarkasteltavien käyttöpaikkojen lukumäärä oli 4938. Tapauksella 2 tarkoitetaan puolestaan tilannetta, jossa käyttöpaikkojen lukumäärä oli 800.

5.2 Tuntimitatun sähkönkulutusdatan esiprosessointi

Ensin kahden vuoden tuntimitattu aineisto ajettiin MATLAB:iin. Tuntimitattu data järjestettiin siten, että yhdellä rivillä oli yhden käyttöpaikan kaikki mittaukset kahden vuoden ajalta. Rivit kuvasivat siis käyttöpaikkoja ja sarakkeet muuttujina toimivia tuntimittauksia. Tietokannassa puuttuvia mittauservoja merkittiin luvulla -999999999. Jotta näiden lukujen käsittely onnistui MATLAB:ssa, täytyi ne muuttaa NaN arvoiksi. Näin ne pystyttiin jatkossa ohittamaan laskennassa. Kuten aiemmin jo mainittiin, vuoden 2018 mittauksista puuttuivat osalta käyttöpaikoista lokakuun mittaukset. Virheiden välttämiseksi ja tarkastelun pitämiseksi helposti hallittavana tällaiset käyttöpaikat suodatettiin pois ja käytettiin ainoastaan käyttöpaikkoja, joilta löytyi kaikkien kuukausien mittaukset. Käyttöpaikkojen lukumääräksi jäi ensimmäisessä tapauksessa suodatusten jälkeen 2272 ja toisessa tapauksessa 358. Suodatusten seurauksena poistuivat myös kaikki käyttöpaikat, joiden mittauksista 5 %:a oli nollia. Suodatusta kokeiltiin tehdä myös siten, että ensin poistettiin käyttöpaikat, joilta puuttuivat lokakuun 2018 mittaukset. Tämän jälkeen jäljelle jäävistä poistettiin ne, joiden mittauksista 10 %:a oli nollia. Mitään merkittävää eroa valitulla tavalla ei ollut, ja klusterointitulokset olivat samankaltaiset.

Seuraavaksi jäljelle jäänyt aineisto skaalattiin siten, että yksittäisten mittaussarjojen eli vuoden mittaisen aikasarjojen keskiarvoksi tuli 1. Skaalattiin siis jokaisen käyttöpaikan molempien vuosien mittaukset erikseen, jolloin saatiin skaalatut vuosien 2017 ja 2018 kulutukset jokaiselle käyttöpaikalle. Tuntimitatun aineiston rivien skaalaus vaikuttaa merkittävästi pääkomponenttianalyysin (Principal component analysis) antamiin tuloksiin. Tästä enemmän aliluvussa 5.3.2. Yksittäisten mittaussarjojen skaalaaminen on oleellista, koska käyttöpaikoilla on eri suuruiset vuosienergiat. Koon eli vuosienergian suuruuden vaikutus poistettiin siis skaalaamalla jokaisen käyttöpaikan tuntimittausten keskiarvoksi 1 erikseen molemmille tarkasteltaville vuosille.

Aineiston rivien skaalauksen lisäksi skaalattiin myös sarakkeet, jotta muuttujien eli tuntimittausten painokertoimet olisivat samat. Aineiston dimensioiden redusointi ja klusterointi suoritettiin sekä sarakkeet skaalaamalla että ilman skaalausta. Tuloksissa ei ollut merkittäviä eroja, mutta sarakkeet silti skaalattiin, jotta talvipäivien vaikutus ei

olisi niin hallitseva. Muuttujien skaalauksen seurauksena aineiston sarakkeiden keskiarvoksi tuli 0 ja keskihajonnaksi 1. [10]

5.3 Aineiston dimensioiden redusointimenetelmät

Aineiston dimensioiden suuren lukumäärän vuoksi (17520 tuntia = 2 vuotta), dimensioiden redusoiminen on tärkeää, jotta klusterointi on mahdollista suorittaa järkevässä ajassa ja jotta saadut tulokset ovat selkeitä. Dimensioiden vähennyksen periaate on korvata alkuperäiset muuttujat pienemmällä määrällä muuttujia samalla säilyttäen mahdollisimman paljon alkuperäisestä informaatiosta. Dimensioiden redusointiin kokeiltiin kahta erilaista tekniikkaa. Ensimmäisenä kokeiltu menetelmä oli aineiston jakaminen vuosijaksoihin ja summakulutuksen laskeminen jokaiselle vuosijaksolle käyttöpaikkakohtaisesti. Toisena kokeiltu menetelmä ja lopulta valittu menetelmä oli pääkomponenttianalyysi.

5.3.1 Aineiston jakaminen vuosijaksoittain

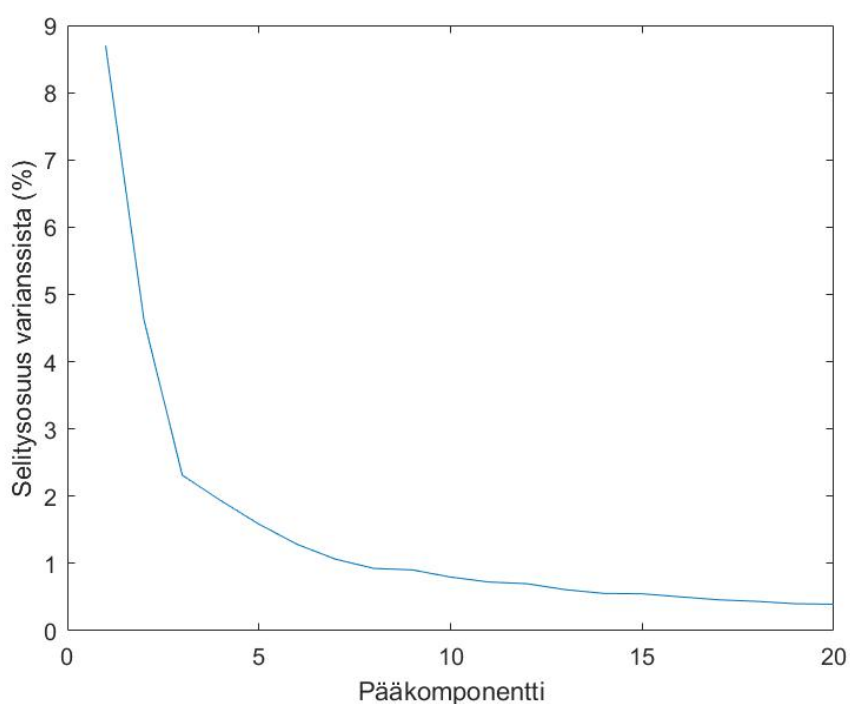
Normalisoitu tuntimitattu data sisälsi 17520 muuttujaa eli kahden vuoden tuntimittaukset. Ryhmittelyn suorittaminen näin suurella määrällä muuttujia on hankalaa ja hidasta. Tämän vuoksi ensimmäisenä kokeiltiin aineiston dimensioiden redusointia tuntimittausten jakamisella vuosijaksoihin. Talvijaksoon kuuluivat vuodesta päivät 321–90, kevätjaksoon päivät 91–180, kesäjaksoon päivät 181–240 ja syksyjaksoon päivät 241–320. Jokaiselle käyttöpaikalle laskettiin summakulutus neljälle eri vuosijaksolle erikseen ohittaen puuttuvat mittaukset. Näin saatiin dimensioiden lukumäärä vähennettyä neljään. Menetelmää käyttäen ei kuitenkaan saatu klustereita, jotka olisivat olleet helposti ryhmien keskiarvosarjoista tunnistettavissa, joten tätä menetelmää ei käytetty varsinaisessa ryhmittelyssä.

5.3.2 Pääkomponenttianalyysi

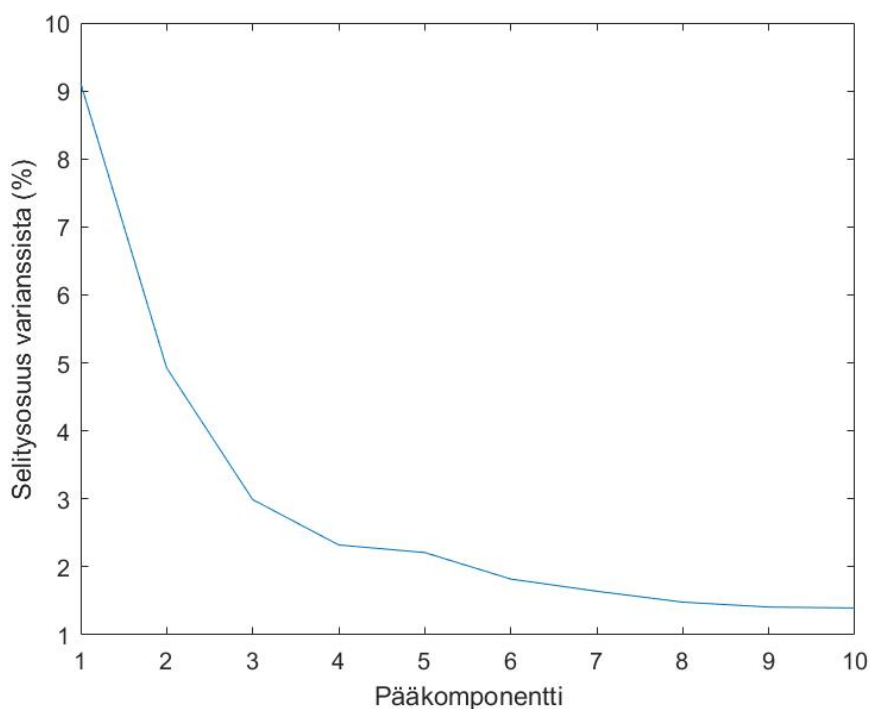
Toisena kokeiltu aineiston dimensioiden redusointimenetelmä oli pääkomponenttianalyysi (PCA). Se on tilastollinen menetelmä, jossa suuri määrä muuttujia pyritään korvaamaan vähemmällä määrällä muuttujia säilyttäen mahdollisimman suuri osa alkuperäisestä informaatiosta. Mitä vähemmän muuttujia jää jäljelle sitä parempi. [12] [13] Pääkomponenttianalyysi perustuu ortogonaaliseen muunnokseen, jossa alkuperäisen aineiston arvot pyritään muokkaamaan pienemmäksi määräksi muuttujia, jotka korreloivat keskenään mahdollisimman vähän. Tarkoituksena on saada siis toisistaan riippumattomia muuttujia. Näitä keskenään mahdollisimman vähän korreloivia muuttujia kutsutaan pääkomponenteiksi. Pääkomponenttien lukumäärä voi olla korkeintaan yhtä suuri kuin alkuperäisten muuttujien. [12] Normaalisti pääkomponenttianalyysissä havaintojen eli tässä yhteydessä käyttöpaikkojen lukumäärä on suurempi kuin tarkasteltavien alkuperäisten muuttujien, jotka ovat tässä tapauksessa tuntimittauksia. Kuitenkin usein tuntimittauksia tarkasteltaessa havaintojen lukumäärä on pienempi kuin muuttujien. Tällöin pääkomponenttianalyysin tuloksena saadaan enintään $n-1$ pääkomponenttia, missä n kuvaa käyttöpaikkojen lukumäärää. [13] Pääkomponenteista ensimmäinen muodostaa havaintoarvojen välille isoimman mahdollisen varianssin. Seuraavat pääkomponentit määrittävät havaintoarvojen välille myös mahdollisimman suuren varianssin ottaen huomioon pääkomponenttien korreloimattomuuden sekä ortogonaalisuuden. Ensimmäisen pääkomponentin jälkeiset pääkomponentit noudattavat siis kahta ehtoa. [14]

Kuten aiemmin jo mainittiin, ennen pääkomponenttianalyysia tuntimitattu data normalisoitiin siten, että jokaisen käyttöpaikan vuosienenergiaksi tuli 1. Myös sarakkeet normalisoitiin, jolloin jokaisen sarakkeen keskiarvoksi tuli 0 ja keskihajonnaksi 1. Tällä tavoin saatiin tuntimittausten painokertoimet tasoitettua. Pääkomponenttianalyysiin sijoitettiin matriisi, jonka jokainen rivi kuvasi yhtä käyttöpaikkaa ja sarakkeet kyseisen käyttöpaikan normalisoituja 17520 tuntimittausta. Eri mittausvuodet normalisoitiin erikseen. Jokainen muodostuva pääkomponentti on alkuperäisten muuttujien lineaarikombinaatio. Tällä tavoitellaan sitä, että kulutukseltaan erilaiset sähkökäyttäjät saataisiin erotettua toisistaan riittävällä tarkkuudella.

Pääkomponenttianalyysin tuloksena saatiin ensimmäisessä tapauksessa 2160 pääkomponenttia ja toisessa tapauksessa 341. Pääkomponenttien lukumäärä ei ollut sama kuin alkuperäisten muuttujien, koska käyttöpaikkojen lukumäärä aineistossa oli pienempi kuin muuttujien. Kuvat 9 ja 10 havainnollistavat 20 ensimmäisen pääkomponentin selityssasteita ensimmäisessä tarkastellussa tapauksessa ja 10 ensimmäisen pääkomponentin toisessa. Kuvista nähdään, että pääkomponenttien selityssasteet laskevat hyvin vahvasti.



Kuva 9: Pääkomponenttien selityssasteet tapauksessa 1

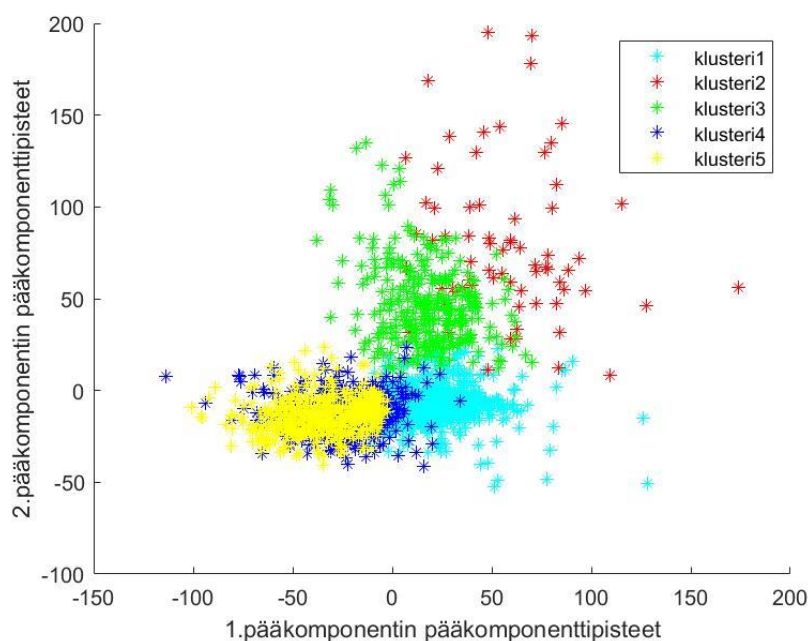


Kuva 10: Pääkomponenttien selitysasteet tapauksessa 2

Kuvista nähdään, että selvä taitepiste on kolmannen pääkomponentin kohdalla molemmassa tapauksissa. Ensimmäisessä tapauksessa valittiin kuitenkin pääkomponenttien lukumääräksi neljä, jotta viiden eri kuluttajaryhmän erottaminen olisi mahdollista. Kolmella pääkomponentilla kokeiltiin myös ryhmittelyä, mutta neljällä saatiin selkeämmin erotettavat kuluttajaryhmät, vaikka merkittävää eroa valinnalla ei ollut. Toisessa tapauksessa pääkomponenttien lukumääränä käytettiin myös neljää, mutta ainoastaan hieman tarkempien tulosten saamiseksi.

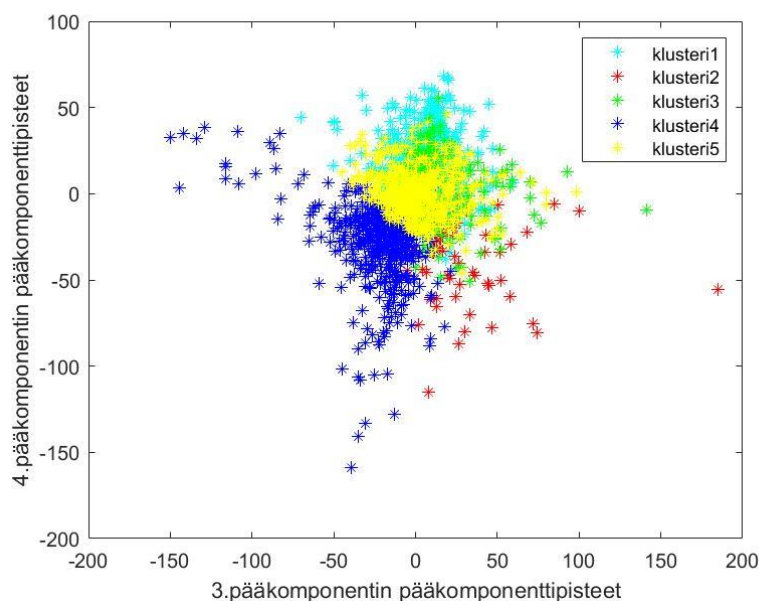
Kun pääkomponenttianalyysi ajetaan, niin jokaiselle tunnille määritetään omat pääkomponenttikertoimensa. Pääkomponenttikerroin muodostaa suurimman mahdollisen varianssin havaintoarvojen välille. Näiden pääkomponenttikertoimien avulla saadaan laskettua käyttöpaikkojen yksilölliset pääkomponenttipisteet käyttäen hyväksi sähkönkulutuksen tuntimittaussarjoja. [10] Varsinainen klusterointi suoritetaan käyttämällä asiakkaiden yksilöllisiä pääkomponenttipisteitä. Ennen klusterointia tehtiin ortogonaalinen maksimivarianssin kierto. Tämä ei vaikuta klusteroinnin tuloksiin mitenkään, vaan ainoastaan muokkaa tuloksia helpommin

havaittavaan muotoon. Alla oleva kuva 11 havainnollistaa ensimmäisen ja toisen pääkomponentin pääkomponenttipisteitä ensimmäisessä tapauksessa, jossa käyttöpaikkojen lukumäärä ennen suodatusta oli 4938. Eri väriset tähdet edustavat eri kuluttajaryhmiin kuuluvien käyttöpaikkojen pääkomponenttipisteitä. Kuluttajaryhmät saatiin erotettua toisistaan K-means -algoritmia käyttämällä.



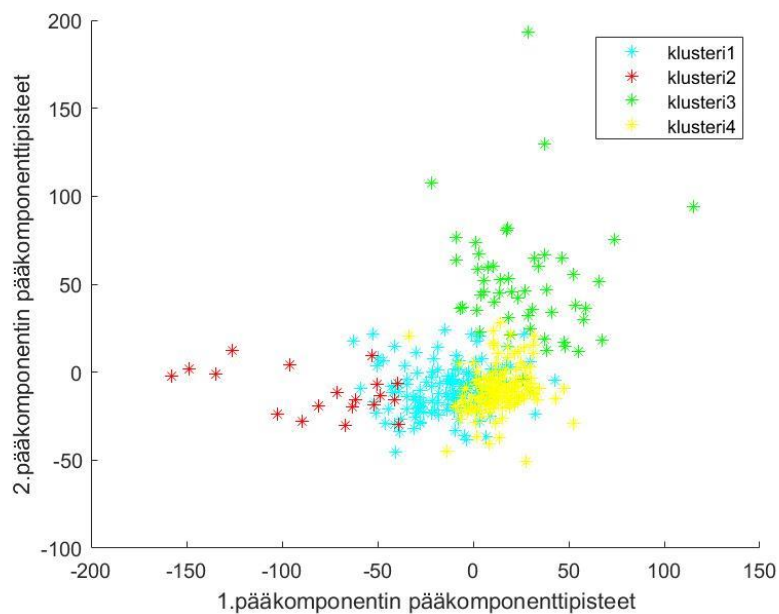
Kuva 11: Kainuun alueen käyttöpaikkojen pääkomponenttien 1 ja 2 pääkomponenttipisteet tapauksessa 1

Seuraava kuva 12 puolestaan kuvaa kolmannen ja neljännen pääkomponentin pääkomponenttipisteitä ensimmäisessä tapauksessa. Nämä saatiin muodostettua samalla tavalla kuin ensimmäisen ja toisen pääkomponentin pääkomponenttipisteet.

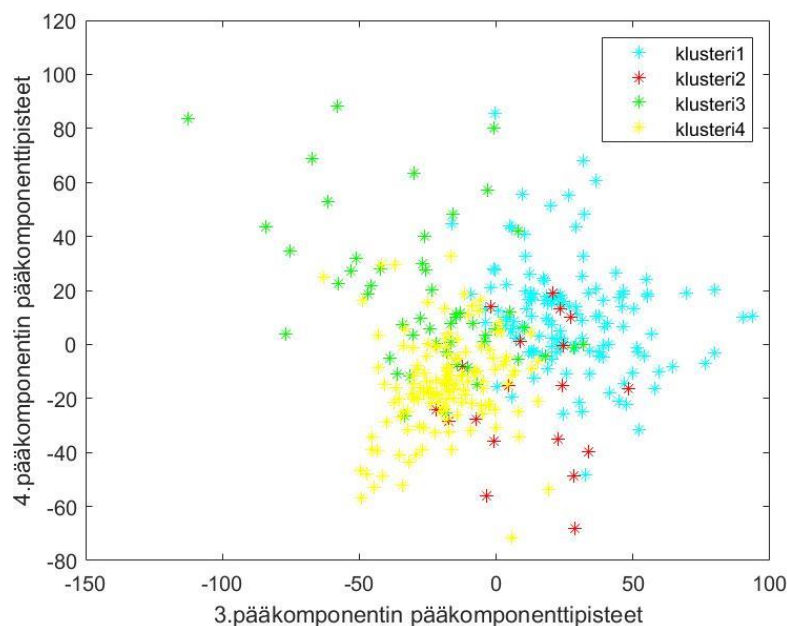


Kuva 12: Kainuun alueen käyttöpaikkojen pääkomponenttien 3 ja 4 pääkomponenttipisteet tapauksessa 1

Kuvissa 13 ja 14 on esitetty vastaavassa järjestyksessä tapauksen 2, jossa tarkasteltavien käyttöpaikkojen alkuperäinen lukumäärä oli 800, pääkomponenttien pääkomponenttipisteet.



Kuva 13: Kainuun alueen käyttöpaikkojen pääkomponenttien 1 ja 2 pääkomponenttipisteet tapauksessa 2



Kuva 14: Kainuun alueen käyttöpaikkojen pääkomponenttien 3 ja 4 pääkomponenttipisteet tapauksessa 2

Kuvista voidaan päätellä, että kuluttajaryhmiä erottelevat parhaiten pääkomponenttien lineaarikombinaatiot eivätkä mitkään yksittäiset pääkomponentit. Tämä havaitaan siitä, että ryhmät eivät muodostu akselien suuntaisesti. [1] Pääkomponenttianalyysin avulla saatiin aineisto siihen muotoon, että se on helppo jakaa kuluttajaryhmiin oikeastaan mitä tahansa klusterointialgoritmia käyttämällä.

5.4 Valittu klusterointimenetelmä

Kuten aikaisemmin on jo mainittu, klusterointimenetelmäksi valittiin K-means -menetelmä. Se on klusterointialgoritmi, jota hyödyntämällä voidaan jakaa aineisto pienempiin ryhmiin. Peruseriaate on se, että pyritään etsimään joukkoja, joilla on jotain yhteisiä ominaisuuksia ja ne ovat samankaltaisia keskenään. K-means -algoritmin toiminta perustuu vain neljään eri vaiheeseen. [15]

1. Algoritmi valitsee ensin sattumanvaraisesti N kappaletta osajoukkojen keskipisteitä.
2. Tarkasteltavan aineiston jokaisen datapisteen etäisyys lasketaan näihin osajoukkojen keskipisteisiin. Datapiste sijoitetaan siihen klusteriin, jonka keskipisteeseen nähden etäisyys on pienin.
3. Seuraavaksi muodostuneelle joukolle lasketaan uusi keskipiste ja edellinen keskipiste siirretään tähän kohtaan.
4. Kohtien 2 ja 3 toistamista jatketaan niin kauan, kunnes joukkojen keskipisteet eivät enää siirry.

K-means -algoritmin etäisyyden mittana käytettiin tässä työssä toiseen potenssiin korotettua euklidista etäisyyttä. Tämä eroaa euklidisesta etäisyydestä ainoastaan siten, että se ei ota neliöjuurta pisteiden välisen etäisyyden toisesta potenssista. Kaava 17 kuvaa datapisteen ja klusterin keskipisteen välisen etäisyyden laskentatapaa. [16]

$$d = \sum_{i=1}^n (x_i - y_i)^2 \quad (17)$$

missä

d on datapisteen ja klusterin keskipisteen välinen etäisyys

n on datapisteiden lukumäärä

x_i on tarkasteltava datapiste

y_i on klusterin keskipiste

K-means -algoritmin lähtötietona käytettiin pääkomponenttianalyysin tuloksena saatuja neljän ensimmäisen pääkomponentin pääkomponenttipisteitä. Klusterien lukumääräksi valittiin ensimmäisessä tapauksessa viisi, koska haluttiin erottaa sähkönkulutukseltaan viisi erilaista ryhmää. Toisessa tapauksessa valittiin klusterien lukumääräksi neljä, koska viiden eri klusterin tapauksessa suuremmalla määrällä dataa ei saatu viittä eri asiakasryhmää tunnistettua selkeästi. Kuten aikaisemmin jo mainittiin, ennen klusterointia tehtiin ortogonaalinen maksimivarianssin kierto, jotta klusteroinnin tulokset olivat selkeämmin havaittavassa muodossa. Huomattiin, että vaikka lähtötiedot olivat samat joka K-means -algoritmin ajossa, klusterointitulokset olivat hieman erilaiset. Tämä johtuu siitä, että K-means -algoritmi pyrkii etsimään uniikkeja tuloksia. Tämän vuoksi ennen algoritmin ajoa lisättiin tulosten toistettavuuden vuoksi parametri, joka antaa samoilla lähtötiedoilla aina saman tuloksen.

K-means -klusteroinnissa haasteena on se, että lähtötietona tulee antaa klusterien lukumäärä, vaikka ennalta on hankala tietää järkevää klusterien lukumäärää. Tämän vuoksi klusterointia kokeiltiin eri klusterien lukumäärillä ja klusterien hyvyttä arvioitiin silhouetten menetelmällä. Silhouetten menetelmä on yksi monista menetelmistä, joiden avulla voidaan yrittää määrittää klusterien optimaalinen lukumäärä. Silhouetten algoritmi olettaa, että data on jo klusteroitu k kappaleeseen klustereita. Tämän jälkeen jokaiselle datapisteelle määritetään kerroin a , joka kuvaa datapisteen i ja kaikkien muiden samaan ryhmään kuuluvien pisteiden keskimääräistä etäisyyttä saman klusterin sisällä. [17]

$$a(i) = \frac{1}{|C(i)|-1} \sum_{j \in C_i, i \neq j} d(i, j) \quad (18)$$

missä

$a(i)$ on kerroin, joka kuvaa kuinka hyvin datapiste i kuuluu klusteriin, johon se on sijoitettu

C_i on klusteri, johon datapiste i on sijoitettu klusteroinnin yhteydessä

$d(i, j)$ on datapisteiden i ja j välinen etäisyys klusterissa C_i

Seuraavaksi määritetään kerroin b , joka kuvaa datapisteen i pienintä keskimääräistä etäisyyttä sen naapuriklusteriin C . Tämä määritetään datapisteen i keskimääräisenä etäisyytenä kaikkiin naapuriklusterin C pisteisiin. [17]

$$b(i) = \min_{k \neq i} \frac{1}{|C_k|} \sum_{j \in C_k} d(i, j) \quad (19)$$

missä

$b(i)$ on kerroin, joka kuvaa datapisteen i pienintä keskimääräistä etäisyyttä lähimpään naapuriklusteriin C

Edellä esitettyjen kertoimien avulla saadaan määritettyä silhouetten kerroin jokaiselle datapisteelle erikseen. Silhouetten kerroin $s(i)$ kuvaa kuinka hyvin datapiste kuuluu klusteriin, johon se on K-means -algoritmin ajon aikana sijoitettu. Jos $|C_i| > 1$, niin alla olevalla kaavalla 20 saadaan yksittäiselle datapisteelle määritettyä silhouetten kerroin. [17]

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (20)$$

missä

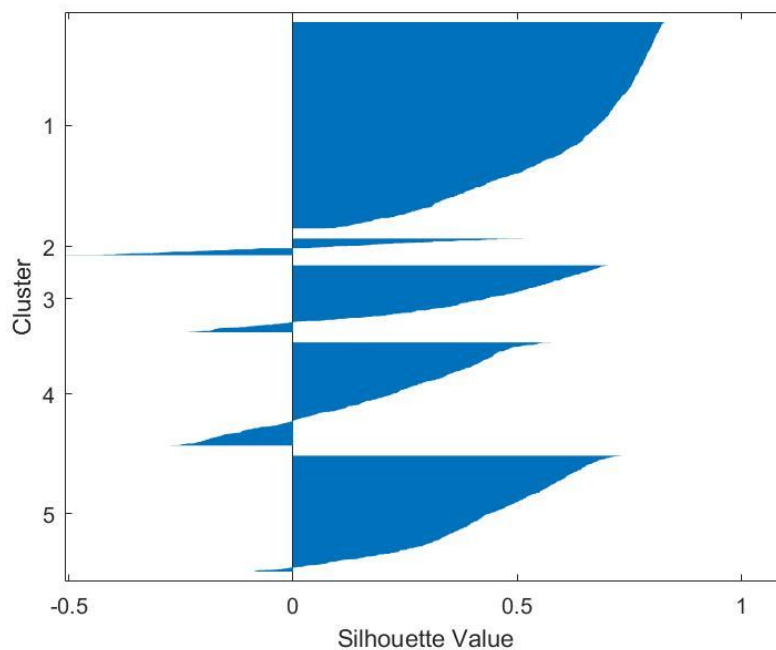
$s(i)$ on datapisteen i silhouetten kerroin

Jos puolestaan $|C_i| = 1$, niin tällöin edellä oleva kaava saa muodon:

$$s(i) = 0 \quad (21)$$

Yllä olevien määrittelyjen perusteella on selvää, että $-1 \leq s(i) \leq 1$. Mitä lähempänä datapisteen i silhouetten kerroin $s(i)$ on arvoa 1, niin sitä paremmin datapiste i kuuluu klusteriin, johon se on sijoitettu. Jos puolestaan datapisteen i silhouetten kerroin $s(i)$ on

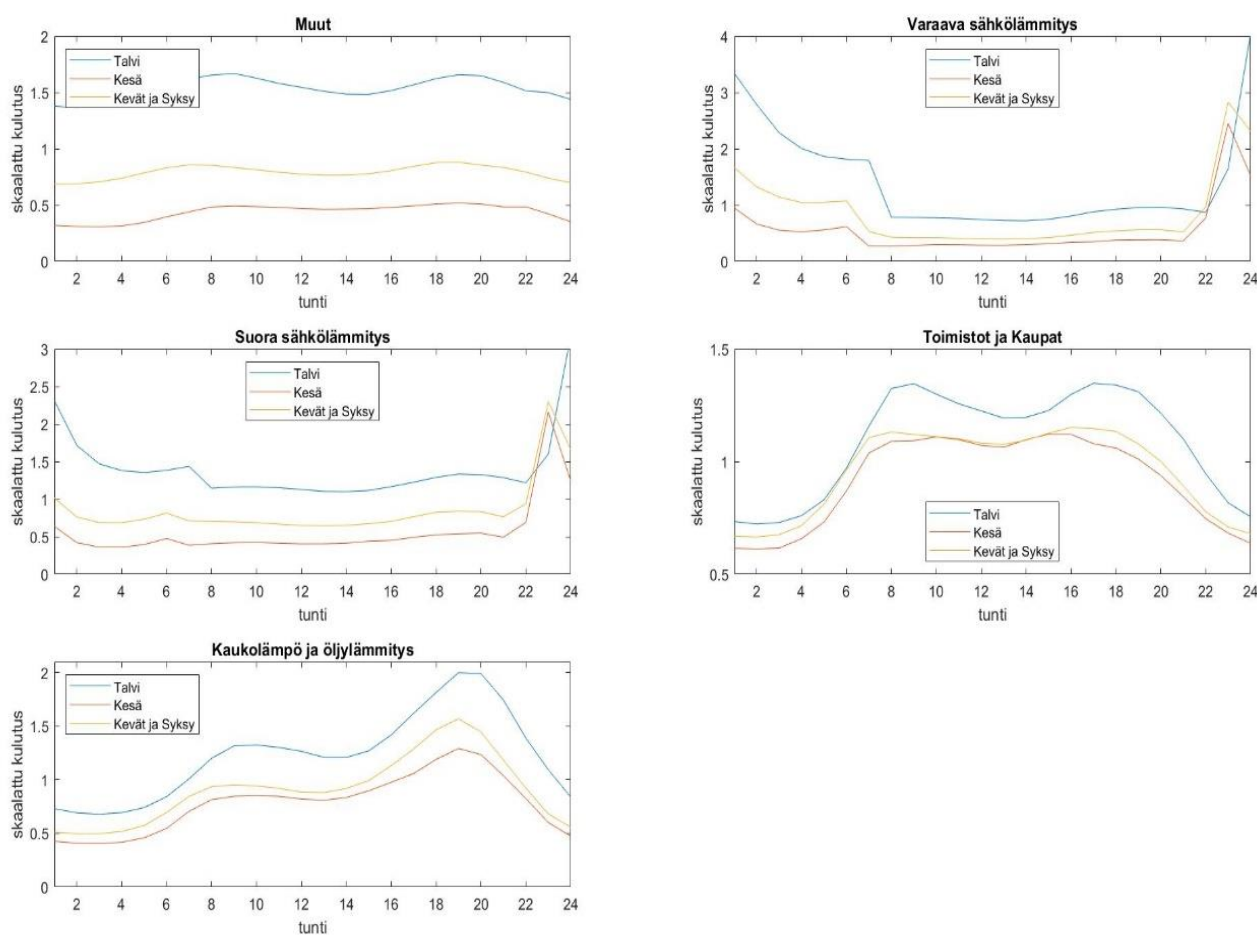
lähellä arvoa -1, niin tällöin datapiste i sopisi paremmin sen naapuriklusteriin. Tässä työssä silhouetten menetelmän tarkastelun yhteydessä datapisteillä tarkoitetaan käyttöpaikkoja ja niiden sopivuutta klustereihin, joihin ne on K-means -algoritmin ajon aikana sijoitettu. Alla oleva kuva 15 havainnollistaa käyttöpaikkojen silhouetten kertoimia. Mitä lähempänä arvoa 1 ollaan, sitä paremmin käyttöpaikka kuuluu klusteriin, johon se on sijoitettu. [17]



Kuva 15: Käyttöpaikkojen silhouetten kertoimet

5.5 Klusteroinnin tulokset

K-means -menetelmän tuloksena saatiin ensimmäisessä tapauksessa viisi kuluttajaryhmää. Jotta saatujen asiakasryhmien vastineet oli mahdollista selvittää, tuli jokaiselle ryhmälle laskea kuluttajaryhmäkohtaiset keskiarvosarjat. Keskiarvosarjat muodostettiin jokaiselle ryhmälle erikseen laskemalla ryhmän käyttöpaikkojen normalisoidut tuntitehot yhteen tunneittain ja ottamalla keskiarvo tästä summasarjasta. Saatujen keskiarvosarjojen profiileista voitiin päätellä, minkälaisesta lämmitysmuodosta on kyse. Alla oleva kuva 16 havainnollistaa tunnistettuja kuluttajaryhmiä ja näiden vuorokausittaisia tuntisarjoja ensimmäisessä tapauksessa, jossa käyttöpaikkojen alkuperäinen lukumäärä oli 4938.



Kuva 16: Eri kuluttajaryhmien vuorokausittaiset tuntisarjat

Kaukolämpö- ja öljylämmitysryhmä tunnistettiin vertaamalla ryhmän vuorokausittaisia tuntisarjoja suoran sähkölämmitysryhmän sarjoihin. Kuvasta 16 nähdään, että kaukolämpö- ja öljylämmitysryhmän kulutus on aina pienempää kuin suoran sähkölämmitysryhmän. Lisäksi ryhmän kulutus on pientä ympäri vuoden, joten ryhmä edustaa selvästi muuta lämmitysmuotoa kuin sähkölämmitystä. Kaukolämpöä ja öljylämmitystä on hankala erottaa toisistaan, joten ne nimettiin samaan ryhmään. Suora sähkölämmitysryhmä määritettiin puolestaan vastaavasti vertaamalla sitä kaukolämpö- ja öljylämmitysryhmän profiiliin. Sen vuoksi, että kulutus on suurempaa kuin edellä mainitun ryhmän tapauksessa ja talvella kulutustaso on selvästi suurempi kuin muina vuodenaikoina, saatiin tämä ryhmä erotettua. Kuvasta nähdään, että varaavan sähkölämmityksen ja suoran sähkölämmityksen vuorokausittaiset tuntisarjat muistuttavat hämmästyttävästi toisiaan, mikä puolestaan kertoo siitä, että tässä suoraksi sähkölämmitykseksi nimetty ryhmä voi edustaa myös varaavaa sähkölämmitystä. Toisaalta molemmille ryhmille on ominaista iltaan ajoittuva sähkön kulutuspiikki.

Varaava sähkölämmitysryhmä puolestaan tunnistettiin sen perusteella, että sähkönkulutuksessa on noin kello 22 suuri kulutuspiikki, joka kuvaa varaavien järjestelmien päälle kytkemistä. Päiväsaikaan kulutus on myös pientä, mikä on ominaista varaaville lämmittimille. Varaavien järjestelmien päälle kytkemisen ajankohdasta voidaan päätellä, että aika-siirto-tariffi on selvästi käytössä kyseisellä ryhmällä.

Toimistot ja kaupat tunnistettiin illan ja yön matalan kulutustason sekä päivän huomattavasti korkeamman kulutuksen perusteella. Ryhmää muut oli hankala tunnistaa vuorokausittaisten tuntisarjojen perusteella miksikään selkeästi tunnistettavaksi ryhmäksi. Tämän vuoksi se nimettiin muut nimellä. Skaalaamattomien tuntisarjojen avulla piirretyn kuvan perusteella nähtiin (klusteroinnissa käytettiin kuitenkin normalisoituja tuntitehoja), että kulutus muut ryhmässä on hieman matalampaa kuin suoran sähkölämmitysryhmän tapauksessa, mutta suurempaa kuin kaukolämpö- ja öljylämmitysryhmän tapauksessa. Tämä viittaa siihen, että kyseessä voisi olla jotain muuta kuin sähkölämmitystä käyttävä ryhmä. Varmuutta tähän ei kuitenkaan saatu, joten päädyttiin muut nimeen.

Alla olevassa taulussa 6 on kerrottu käyttöpaikkojen lukumäärät ryhmittäin ensimmäisessä tapauksessa ja ryhmän käyttöpaikkojen osuus prosentteina analyysissa käytettyjen kaikkien 2161 käyttöpaikan joukosta. K-means -algoritmin ajon aikana osaa käyttöpaikoista ei luokiteltu mihinkään ryhmään, vaan näiden klusterointiryhmän numeron tilalle tuli NaN eli puuttuva mittausta. Tämän vuoksi käyttöpaikkojen lukumäärä väheni 2272:sta 2161:een.

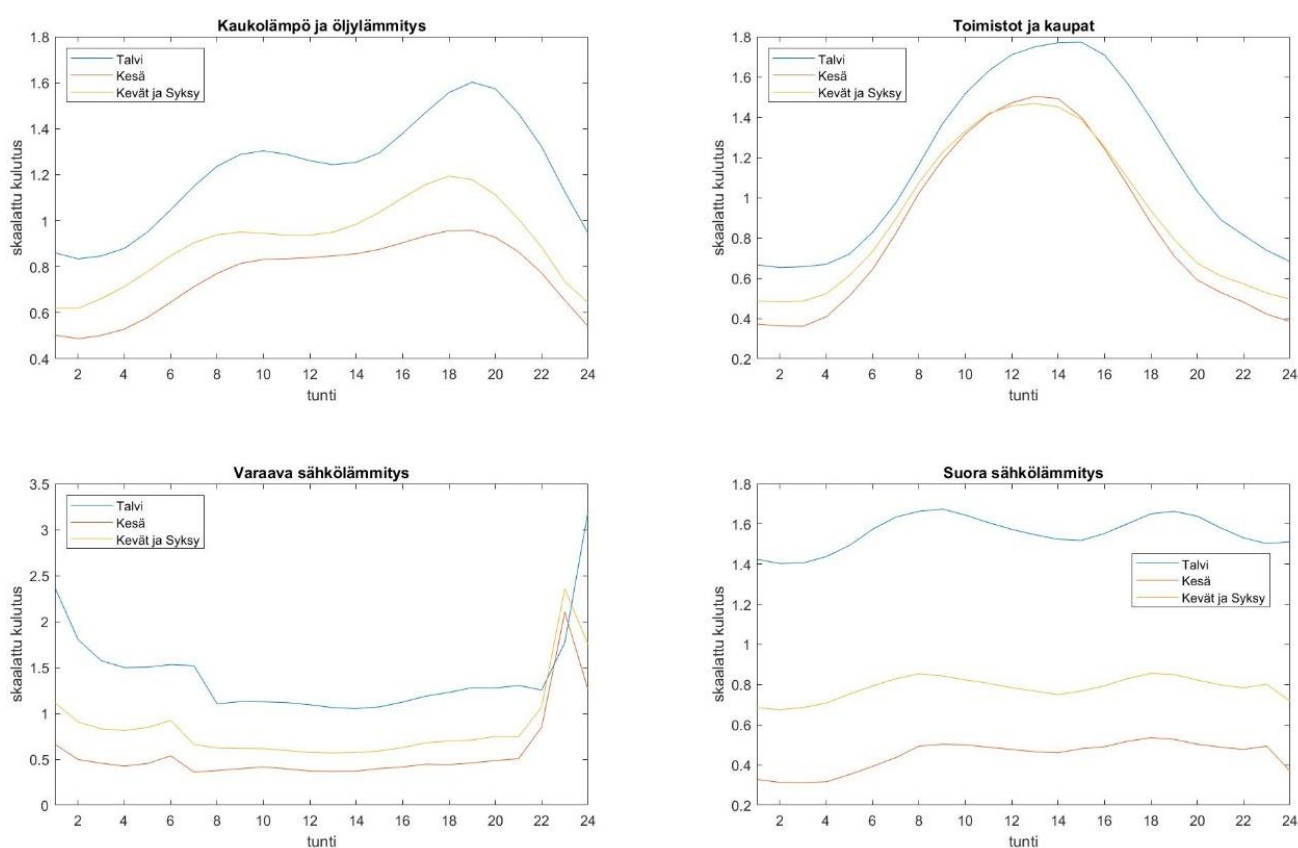
<u>Kuluttajaryhmän nimi</u>	<u>Käyttöpaikkojen lukumäärä</u>	<u>Osuus prosentteina</u>
Kaukolämpö ja öljylämmitys	492	22,8 %
Suora sähkölämmitys	284	13,1 %
Varaava sähkölämmitys	71	3,3 %
Toimistot ja kaupat	438	20,3 %
Muut	876	40,5 %

Taulu 6: Käyttöpaikkojen lukumäärät ryhmittäin ja osuus kaikkien käyttöpaikkojen joukosta tapauksessa 1

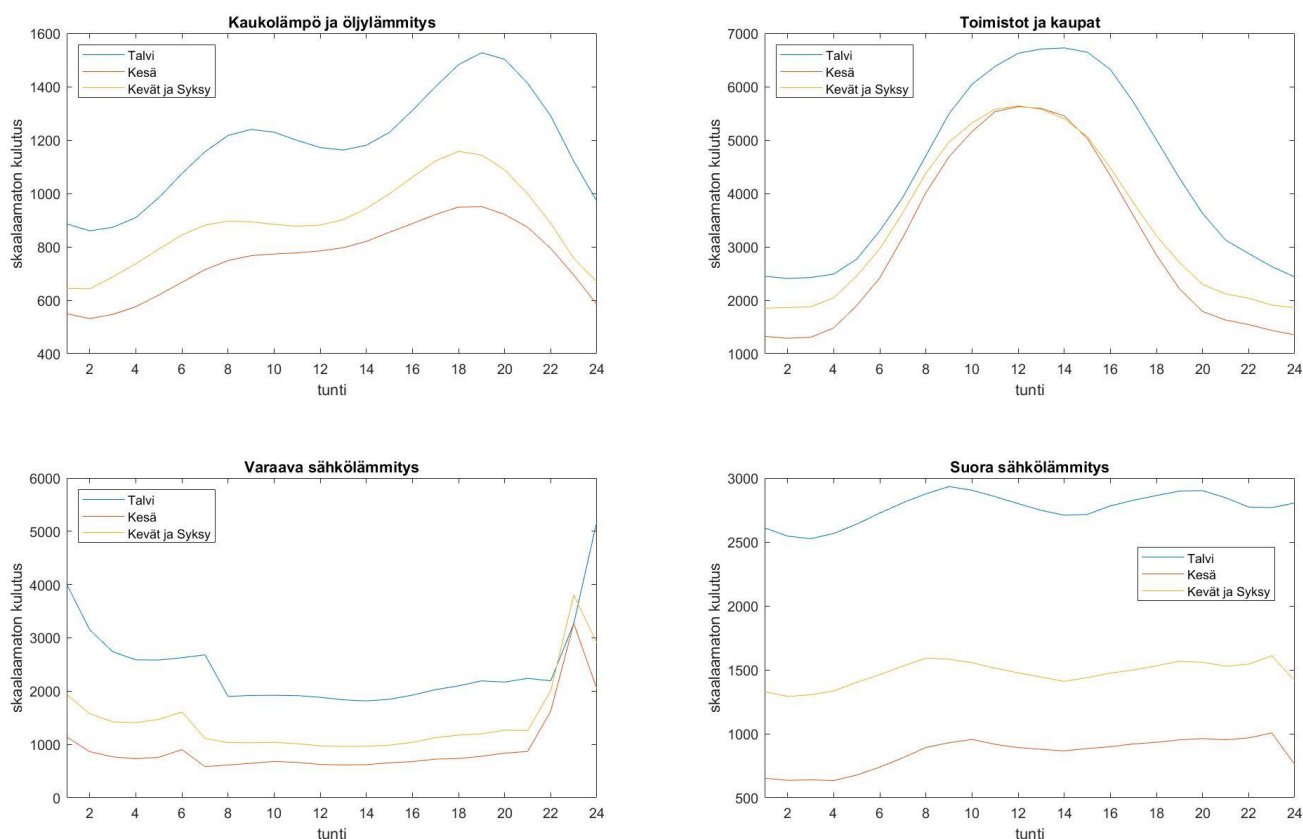
Taulun 6 havainnollistavien tuloksien mukaan otannan suurin ryhmä on ryhmä muut, johon kuuluu 876 käyttöpaikkaa. Tämän vuoksi klusteroinnin tulokset eivät tyydyttäneet, koska haluttiin saada selkeästi tunnistettavat ryhmät erotettua. Taulusta voi nähdä myös, että toimistojen ja kauppajen osuus on todella suuri koko otannasta, mikä herättää epäilyksen siitä, että ryhmittelyn aikana toimistoiksi ja kauppoiksi nimetty ryhmä ei vastaa kovin hyvin nimensä mukaista tarkoitusta. Myös kuvan 16 toimistot ja kaupat ryhmän vuorokausittaiset tuntisarjat näyttävät osittain oudolta

verrattuna ryhmän odotettavissa olevaan profiiliin. Aamulla noin yhdeksältä kulutus alkaa laskea ja nousta takaisin kello neljäntoista jälkeen, mikä ei ollut odotettua.

Edellä mainittujen seikkojen vuoksi, kuten aiemmin on jo mainittu, klusterointi suoritettiin myös 800 käyttöpaikan otannalla, josta suodatusten jälkeen jäi jäljelle 358 käyttöpaikkaa. Tässäkin tapauksessa klusteroinnin jälkeen muodostettiin kuluttajaryhmäkohtaiset keskiarvosarjat normalisoituja tuntitehoja käyttäen, jotta asiakasryhmien vastineet voitiin selvittää. Keskiarvosarjat muodostettiin myös tapauksessa, jossa käytettiin normalisoimattomia tuntitehoja (klusterointi tehty käyttäen normalisoituja tehoja). Tämä tehtiin sen vuoksi, jotta kulutuksen suuruuserot olivat helpommin nähtävissä. Kuvat 17 ja 18 havainnollistavat kuluttajaryhmäkohtaisia vuorokausittaisia tuntisarjoja molemmissa edellä mainituissa tapauksissa 800 käyttöpaikan otannassa.



Kuva 17: Eri kuluttajaryhmien vuorokausittaiset tuntisarjat (laskettu normalisoituja tuntitehoja käyttämällä)



Kuva 18: Eri kuluttajaryhmien vuorokausittaiset tuntisarjat (laskettu normalisoimattomia tuntitehoja käyttämällä)

Kuvista huomataan, että neljän klusterointiryhmän valitseminen ja pienempi aineiston otanta antoivat helpommin tunnistettavat profiilit. Kaukolämpö- ja öljylämmitysryhmä voitiin tunnistaa samalla tavalla kuin tapauksessa 1, koska ryhmän sähkönkulutus on matalaa ympäri vuoden ja pienempää kuin suoran sähkölämmitysryhmän kulutus. Suoran sähkölämmitysryhmän kulutus puolestaan on suurempaa kuin edellä mainitun ryhmän tapauksessa ja talvella kulutustaso on huomattavasti korkeampi kuin muina vuodenaikoina. Öisin kulutuksen taso on myös matalampi.

Toimistojen ja kauppojen vuorokausittaiset tuntisarjat muistuttavat myös huomattavasti enemmän ryhmän odotettua kulutusprofiilia verrattuna tapauksen 1 vastaavaan profiiliin. Varaava sähkölämmitysryhmä tunnistettiin vastaavalla tavalla kuin tapauksessa 1. Taulusta 7 nähdään, että käyttöpaikkojen lukumäärät ryhmittäin

vaikuttavat järkeviltä. Toimistot ja kaupat ryhmän osuus kaikkien käyttöpaikkojen joukosta on huomattavasti pienempi kuin tapauksessa 1. Suurin ryhmä tapauksessa 2 on suora sähkölämmitysryhmä, johon kuuluu 139 käyttöpaikkaa.

<u>Kuluttajaryhmän nimi</u>	<u>Käyttöpaikkojen lukumäärä</u>	<u>Osuus prosentteina</u>
Kaukolämpö ja öljylämmitys	133	38,9 %
Suora sähkölämmitys	139	40,6 %
Varaava sähkölämmitys	51	14,9 %
Toimistot ja kaupat	19	5,6 %

Taulu 7: Käyttöpaikkojen lukumäärät ryhmittäin ja osuus kaikkien käyttöpaikkojen joukosta tapauksessa 2

Tapauksessa 2 saatiin siis selkeämmin tunnistettavat kuluttajaryhmät ja järkevämmät käyttöpaikkojen lukumäärät ryhmittäin. Tämän vuoksi päätettiin indeksisarjojen muodostus suorittaa käyttämällä neljää kuluttajaryhmää ja tapauksen 2 mukaista luokittelua. Pienemmän määrän dataa oletettiin antavan keskihajonnalle riittävän tarkat estimaatit, koska lukumäärältään pienimmässäkin ryhmässä on 19 käyttöpaikkaa.

5.6 Lämpötilariippuvuuden huomiointi indeksisarjoissa

Klusteroinnin ja kuluttajaryhmien muodostamisen jälkeen voidaan ryhmäkohtaiset indeksisarjat muodostaa. Tätä ennen on syytä kuitenkin huomioida lämpötilan vaikutus tuntimittauksiin. Jotta eri vuosien tuntimittaukset olisivat vertailukelpoisia keskenään, tulee lämpötilan vaikutus poistaa tuntitehoista. Tämä toteutettiin mitatun aikasarjan normalisoinnilla [6]. Tuloksena saadaan pitkän aikavälin lämpötiloja eli normaalilämpötiloja vastaavat tuntitehot.

Lämpötilariippuvuuskertoimien laskentaa ennen normalisoidut tuntitehot järjestettiin siten, että aineiston jokaisen rivin ensimmäinen tuntimittaus kuvasi kello 7–8 välisen tunnin tehoa ja viimeinen puolestaan kello 6–7 välisen tunnin tehoa. Tähän asti tuntimittaukset ovat olleet muodossa 0–24, mutta Trimblen tietokannan taulujen rakenteen vuoksi nämä järjestettiin edellä kuvatulla tavalla. Tuntimittaukset olivat järjestämisen jälkeen siis muodossa 7ap.–7ap. ja kyseistä aikaväliä käytettiin vuorokauden uutena aikavälinä. Molempien vuosien 2017 ja 2018 tuntimittauksista poistettiin siis 7 ensimmäistä ja 17 viimeistä tuntimittauksista. Tällä tavoin saatiin edellä mainittu haluttu muoto. Sama toistettiin jokaisen kuluttajaryhmän tuntimittauksille erikseen.

5.6.1 Lämpötilariippuvuuskertoimien laskenta

Ennen kuin lämpötilan vaikutus voitiin poistaa, tuli laskea kuluttajaryhmäkohtaiset lämpötilariippuvuuskertoimet. Jokaiselle kuluttajaryhmälle määritettiin omat lämpötilariippuvuuskertoimensa vuosijaksoittain. Talvijaksona käytettiin päiviä 321–90, kevätjaksona päiviä 91–180, kesäjaksona päiviä 181–240 ja syysjaksona päiviä 241–320. Kuluttajaryhmään kuuluvien käyttöpaikkojen mittauksista laskettiin jokaiselle tunnille tuntitehon keskiarvo eli saatiin niin sanottu ryhmän keskiarvokäyrä. Näitä ryhmäkohtaisia keskiarvotehoja käyttämällä määritettiin vuorokausikohtaiset energiat jokaiselle vuorokaudelle kahden vuoden ajalta summaamalla vuorokauden tuntitehot yhteen. Laskettua keskiarvokäyrää vertaamalla vastaavaan lämpötilakäyrään voitiin tarkastella, miten lämpötila ja tuntimitatut tehot korreloivat keskenään eli miten

lämpötila vaikuttaa tuntitehoihin. Vastaavalla tavalla nähtiin, miten vuorokauden energiat muuttuivat vuorokauden keskilämpötilojen muuttuessa.

Lämpötilariippuvuuskertoimien laskennassa toimittiin niin, että määritettiin aina kahden peräkkäisen viikon vastaavien viikonpäivien vuorokauden energioiden energiaero. Edellisen vuorokauden keskilämpötilalla selitettiin tarkasteltavan vuorokauden energiaa, jotta lämpötilan vaikutuksen ajallinen viive tuli huomioitua. Alla oleva kaava 22 havainnollistaa yksittäisen lämpötilariippuvuuskertoimen laskentatapaa:

$$c = \frac{P_{E,d} - P_{E,d-1}}{(T_d - T_{d-1}) \times \left(\frac{P_{E,d} + P_{E,d-1}}{2} \right)} \quad (22)$$

missä

$P_{E,d}$ on kyseisen viikon tarkasteltavan viikonpäivän vuorokauden energia

$P_{E,d-1}$ on edellisen viikon vastaavan viikonpäivän vuorokauden energia

T_d on tarkasteltavaa viikonpäivää edeltävän vuorokauden keskilämpötila

T_{d-1} on edellisen viikon tarkasteltavaa viikonpäivää edeltävän vuorokauden keskilämpötila

$\left(\frac{P_{E,d} + P_{E,d-1}}{2} \right)$ on kyseisen viikon viikonpäivän ja edellisen viikon vastaavan viikonpäivän keskiteho

Edellä esitetystä lineaarisesta lämpötilariippuvuusmallista voidaan päätellä, että prosentuaalinen lämpötilariippuvuuskerroin on sama vuorokauden jokaisen tunnin teholle. Vuorokausien keskitehoa käytettiin lämpötilariippuvuuskerroimien laskemisessa osana jakajaa sen vuoksi, että saatiin yksittäisten tehopiikkien vaikutukset minimoitua ja jotta tehon lämpötilariippumatonta osaa voitiin mallintaa. Tämä mahdollisti sen, että eri suuruisten tehojen erotukset olivat paremmin vertailukelpoisia. Esimerkiksi tehoero $200 - 180 = 20$ vastaa 10 %:n eroa ja samaa vastaa myös tehojen $100 - 90 = 10$ ero. Keskiteholla jakaminen tekee tehoeroista siis paremmin vertailukelpoisia.

Kuten edellä jo mainittiin, lämpötilariippuvuuskerroimet määritettiin kullekin viikonpäivälle erikseen vuosijakson sisällä. Kuluttajaryhmäkohtainen lopullinen vuosijaksolla käytettävä lämpötilariippuvuuskerroin saatiin ottamalla viikonpäivien kertoimista keskiarvo. Laskettiin siis esimerkiksi tarkasteltavan viikon tiistain energian ja edellisen viikon tiistain energian energiaeron suhde vastaavassa järjestyksessä lasketuilla maanantaiden keskilämpötilojen lämpötilaeron ja tiistaiden keskitehon tulolla. Sama toistettiin vuosijakson sisällä kaikille viikonpäiville ja lopulta otettiin yksittäisten lämpötilariippuvuuskerroimien keskiarvo, jolloin saatiin siis lopullinen vuosijaksolla käytettävä kuluttajaryhmäkohtainen lämpötilariippuvuuskerroin. Tämä toistettiin jokaiselle vuosijaksolle, jolloin saatiin neljä kerrointa kuluttajaryhmää kohden.

Koska tehot voivat vaihdella suuresti tiettyinä jaksoina johtuen muista kuin lämpötilasta riippuvista syistä, asetettiin viikonpäiville määritettyjen lämpötilariippuvuuskerroimien ylärajaksi $\pm 10\% / ^\circ C$. Tällä voitiin huolehtia siitä, ettei laskennan tuloksena saatu yksittäisille lämpötilariippuvuuskerroimille järjettömiä arvoja, jotka olisivat vaikuttaneet merkittävästi vuosijakson lopullisen lämpötilariippuvuuskerroimen arvoon. VTT:n kuormitustutkimuksessa vuonna 2003 toimittiin samalla tavalla [6].

Vuorokauden keskilämpötiloina käytettiin Ilmatieteenlaitoksen arkistosta haettuja Kajaanin lentoaseman vuorokauden keskilämpötiloja. Kajaanista löytyi Ilmatieteenlaitoksen arkistosta yhteensä kolme mittauspaiikkaa, jotka ovat Kajaani Paltaniemi, Kajaani Petäisenniska ja Kajaani lentoasema, joista valittiin edellä mainittu

Kajaanin lentoasema. Alla oleva taulu 8 havainnollistaa kaavaa 22 käyttäen neljälle kuluttajaryhmälle laskettuja ryhmäkohtaisia lämpötilariippuvuuskertoimia (%/°C), joita käytettiin lämpötilan vaikutuksen poistamisessa tuntitehoista.

<u>Kuluttajaryhmän nimi</u>	<u>Talvi</u>	<u>Kevät</u>	<u>Kesä</u>	<u>Syksy</u>
Kaukolämpö ja öljylämmitys	-1,11	-1,04	-0,40	-1,69
Suora sähkölämmitys	-1,94	-3,46	-1,63	-3,55
Varaava sähkölämmitys	-1,41	-2,42	-1,35	-2,39
Toimistot ja kaupat	-0,29	-0,39	1,51	-0,61

Taulu 8: Lasketut kuluttajaryhmäkohtaiset lämpötilariippuvuuskertoimet (%/°C)

5.6.2 Mitatun aikasarjan normalisointi

Kuten aikaisemmin jo mainittiin, jotta eri vuosien tuntimittaukset olisivat vertailukelpoisia keskenään, tulee lämpötilan vaikutus poistaa tuntitehoista ennen indeksisarjojen muodostamista. Tämä tehtiin mitatun aikasarjan normalisoinnilla. Lämpötilan vaikutus poistettiin tuntimittauksista alla olevaa kaavaa 23 käyttäen:

$$P_{lk} = \frac{P}{1 + c \times \Delta T} \quad (23)$$

missä

P_{lk} on tietyn ajanhetken teho, josta on poistettu lämpötilan vaikutus

P on tietyn ajanhetken teho

c on vuodenajasta riippuva lämpötilariippuvuuskerroin

$\Delta T = T_d - T_n$ on vuorokauden keskilämpötilan ja normaalilämpötilan lämpötilaero

Lämpötilan vaikutuksen ajallinen viive huomioitiin mitatun aikasarjan normalisoinnissa siten, että laskettaessa tarkasteltavan vuorokauden korjattuja tuntitehoja käytettiin aina edellisen vuorokauden keskilämpötilaa tarkasteltavan vuorokauden keskilämpötilan sijaan. Kuukauden normaalilämpötilan sijaan käytettiin normaalilämpötilana kaksiviikkojakson normaalilämpötilaa eli lämpötilan odotusarvoa. Tämä määritettiin jokaiselle kaksiviikkojaksolle ottamalla keskiarvo kaksiviikkojakson sisältämien vuorokausien keskilämpötiloista vuosilta 2014–2019 eli kuudelta viimeisimmältä vuodelta. Yleensä normaalilämpötiloja laskettaessa on käytetty lämpötilamittauksia paljon pidemmältä noin 30 viimeisen vuoden ajalta. Edellä mainittuun ratkaisuun päädyttiin sen vuoksi, että 1990-luvun lämpötilamittaukset ja tätä vanhemmat eivät ole enää vertailukelpoisia lämpötilan nykymittausten kanssa. Tämä johtuu pääasiassa jatkuvasta ilmaston lämpenemisestä ja

ilmastonmuutoksesta. Kuuden viimeisen vuoden lämpötilamittausten todettiin riittävän mallintamaan kaksiviikkojaksojen normaalilämpötiloja riittävän tarkasti. Kaksiviikkojaksojen normaalilämpötiloja käytettiin kuukauden normaalilämpötilojen sijaan myös Trimblen tietokannan taulujen rakenteen takia. Alla olevassa taulussa 9 on havainnollistettu kaksiviikkojaksojen normaalilämpötiloja.

Kaksiviikkojakso	Aikaväli	Lämpötila (°C)
1	1.1-14.1	-8,91071
2	15.1-28.1	-12,3393
3	29.1-11.2	-7,45119
4	12.2-25.2	-5,08452
5	26.2-11.3	-4,78571
6	12.3-25.3	-3,20952
7	26.3-8.4	-0,50119
8	9.4-22.4	1,594048
9	23.4-6.5	4,229762
10	7.5-20.5	8,508333
11	21.5-3.6	11,35238
12	4.6-17.6	12,51548
13	18.6-1.7	13,24524
14	2.7-15.7	15,32619
15	16.7-29.7	17,87619
16	30.7-12.8	16,23214
17	13.8-26.8	14,01786
18	27.8-9.9	12,02976
19	10.9-23.9	9,39881
20	24.9-7.10	5,669048
21	8.10-21.10	3,278571
22	22.10-4.11	0,344048
23	5.11-18.11	-1,46548
24	19.11-2.12	-1,74643
25	3.12-16.12	-3,20476
26	17.12-31.12	-5,0119

Taulu 9: Kaksiviikkojaksojen normaalilämpötilat

Mitatun aikasarjan normalisointi edellyttää, että asiakkaiden mitatut tuntitehot ovat tiedossa koko vuodelle. Tällöin voidaan määrittää yksilölliset ryhmäkohtaiset/asiakaskohtaiset lämpötilariippuvuuskertoimet ja käyttää näitä normalisoinnissa. Mitatun aikasarjan normalisoinnissa todelliset tuntitehot siis

tiedetään, jolloin ei ole tarvetta arvioida tuntitehoja, kuten mitatun energian normalisoinnin tapauksessa. Tämän vuoksi menetelmän avulla saadaan tuntitehoista erotettua tarkemmin lämpötilariippuva tehon osa. [6]

5.7 Uusien indeksisarjojen luominen

Lämpötilan vaikutuksen poistamisen jälkeen voitiin muodostaa kuluttajaryhmäkohtaiset indeksisarjat. Ensin määritettiin sisäiset indeksit eli tunti-indeksit, jotka kuvaavat tehon tunnittaista vaihtelua kaksiviikkojakson sisällä. Tämän jälkeen laskettiin ulkoiset indeksit, jotka kuvaavat puolestaan kaksiviikkojaksojen energioiden kausittaista vaihtelua vuoden sisällä.

5.7.1 Sisäisten indeksien muodostus

Ennen sisäisten indeksien muodostamista jaettiin molemmat vuodet 2017 ja 2018 kaksiviikkojaksoihin. Saatiin siis 26 kaksiviikkojaksoa molempaa vuotta kohden. Jokaisesta kaksiviikkojaksosta eroteltiin arkipäivät, lauantait (aatot) ja sunnuntait (pyhät), koska näille muodostettiin sisäiset indeksit erikseen. Tavallisella kaksiviikkojaksolla, joka ei sisällä arkipyhäitä, on yhteensä 10 arkipäivää, 2 aattoja ja 2 pyhää. Arkipäivien osalta on yhdelle vuorokauden tunnille yhteensä 5 arkipäivää x 2 viikkoa x 2 vuotta x ryhmän käyttöpaikkojen lukumäärä = 20 mittauksia x käyttöpaikkojen lukumäärä. Esimerkiksi toimistot ja kaupat ryhmän tapauksessa yhdelle arkipäivän tunnille on $5 \times 2 \times 2 \times 19 = 380$ tuntitehomittauksia. Ottamalla näiden 380 mittauksen keskiarvo saadaan kaksiviikkojakson sisäinen indeksi tarkasteltavalle tunnille. Lauantaille on puolestaan yhtä tuntia kohden $4 \times$ käyttöpaikkojen lukumäärä mittauksia ja sunnuntaille pätee sama. Yksi kaksiviikkojakso sisältää siis 24 sisäistä indeksiä kullekin päivätyypille erikseen eli 24 arki-indeksiä, 24 aattoindeksiä ja 24 pyhäindeksiä.

Ennen kuin kaksiviikkojaksojen sisäiset indeksit muodostettiin, kaksiviikkojaksoon ajoittuvien arkipyhien tuntimittaukset tuli siirtää pyhäryhmään eli sunnuntain mittauksiin, koska kulutuksen on todettu arkipyhinä olevan sunnuntain kaltaista. Kun

kaksiviikkojaksojen arkipyhät oli siirretty sunnuntairyhmään, kerättiin molempien tarkasteltavien vuosien vastaavien kaksiviikkojaksojen arkipäivien mittaukset yhteen, aattojen mittaukset yhteen ja sunnuntain mittaukset yhteen kuluttajaryhmään kuuluvilta kaikilta käyttöpaikoilta. Kaksiviikkojakson sisäiset indeksit kullekin päivätyypille erikseen saatiin ottamalla jokaisen vuorokauden tunnin mittauksista keskiarvo. Yhteensä saatiin siis kaksiviikkojaksoa kohden 72 sisäistä indeksiä (24 arkipäiville, 24 aatoille, 24 pyhille). Sisäiset indeksit muodostettiin kuluttajaryhmäkohtaisesti sekä keskiarvoille että keskihajonnoille samalla tavalla.

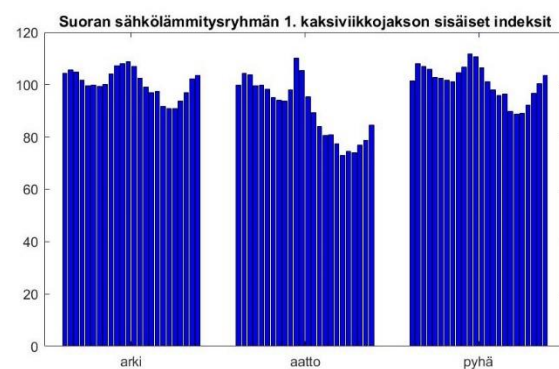
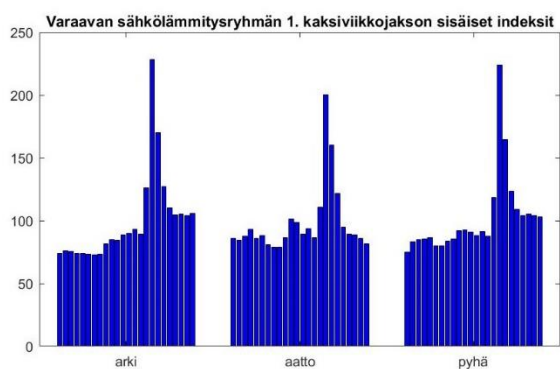
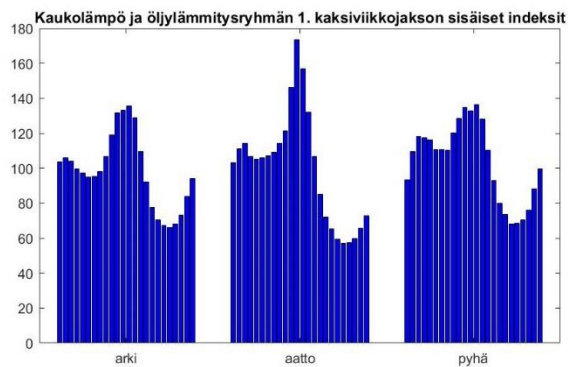
Keskiarvojen sisäiset indeksit skaalattiin vielä Trimblen tietokannan taulujen rakenteen vuoksi siten, että kaksiviikkojakson 72 sisäisen indeksin keskiarvoksi tuli 100 ja summaksi 33600 ($14 \times 24 \times 100$). Niillä kaksiviikkojaksoilla, jotka eivät sisältäneet arkipyhiä, käytettiin skaalauksessa painokertoimia 10 arkipäiville, painokertoimia 2 aatoille ja painokertoimia 2 pyhille. Laskettiin siis kaksiviikkojakson indeksien summa seuraavalla tavalla:

$$X = 10 \times \text{arki-indeksit} + 2 \times \text{aattoindeksit} + 2 \times \text{pyhäindeksit} \quad (24)$$

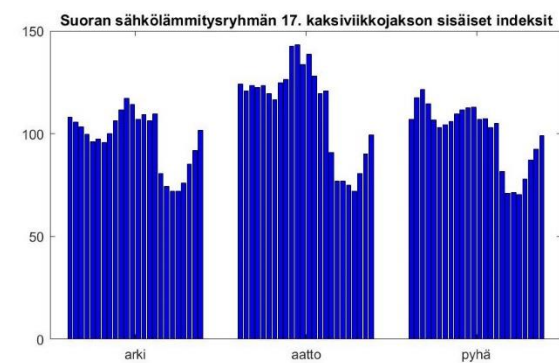
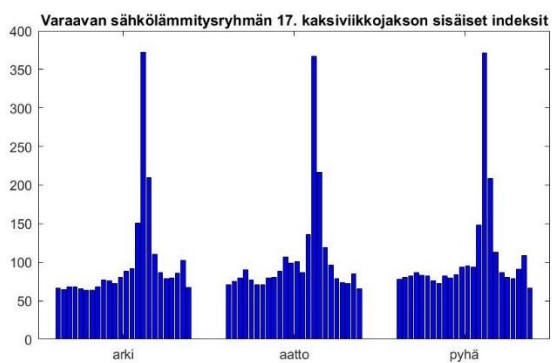
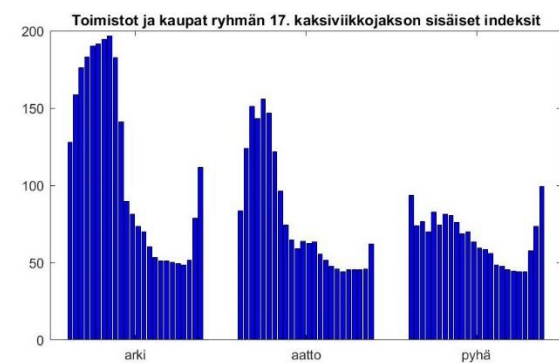
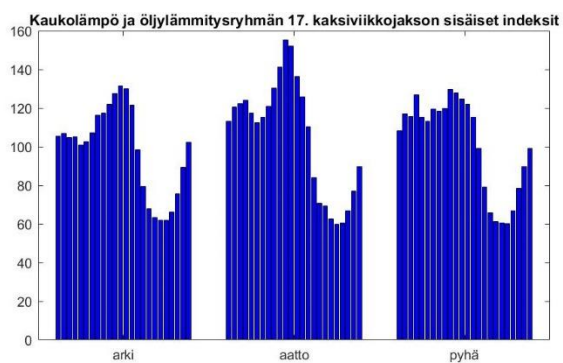
missä

X on kaksiviikkojakson indeksien avulla mallinnettu energia

Puolestaan niiden kaksiviikkojaksojen, jotka sisälsivät arkipyhiä, painokertoimet määräytyivät kaksiviikkojaksolle ajoittuvien arkipyhien lukumäärän mukaan. Painokertoimia muutettiin vain, jos arkipyhä ajoittui molempina vuosina samaan aikaan. Esimerkiksi jos kaksiviikkojaksolla oli yksi arkipyhä, niin tällöin painokertoimet olivat 9 arkipäiville, 2 aatoille ja 3 pyhille. Seuraavaksi kerrottiin kaksiviikkojakson sisäiset indeksit luvulla $33600/X$, jolloin saatiin indeksit haluttuun muotoon. Keskihajontojen skaalaus puolestaan tehtiin kertomalla keskihajontojen sisäiset indeksit luvulla 100. Kuvat 19 ja 20 havainnollistavat kuluttajaryhmien kaksiviikkojaksojen 1 ja 17 skaalattuja sisäisiä indeksejä keskiarvoille. Ensimmäinen arvo kuvaa tunnin 7 sisäistä indeksiä eli sisäiset indeksit ovat järjestyksessä $7ap.-7ap.$



Kuva 19: Kuluttajaryhmien 1. kaksiviikkojakson sisäiset indeksit keskiarvoille



Kuva 20: Kuluttajaryhmien 17. kaksiviikkojakson sisäiset indeksit keskiarvoille

Kuvista nähdään, että sisäisten indeksien profiilit muistuttavat huomattavasti klusteroinnin tuloksena saatujen kuluttajaryhmien keskiarvosarjojen kuvaajia. Tämä oli myös oletettavissa.

5.7.2 Ulkoisten indeksien muodostus

Lämpötilan vaikutuksen poistamisen jälkeen tuntitehoista, voitiin ulkoisten indeksien muodostaminen aloittaa. Laskettiin molempien tarkasteltavien vuosien 2017 ja 2018 kaksiviikkojaksojen energiat kuluttajaryhmäkohtaisesti erikseen. Tämä toteutettiin laskemalla ryhmän kaikkien käyttöpaikkojen tuntitehot yhteen kaksiviikkojakson ajalta. Näin saatiin kaksiviikkojaksojen energiat kaikille vuoden 26 kaksiviikkojaksolle. Vuosien 2017 ja 2018 vuosienergiat määritettiin laskemalla kaksiviikkojaksojen energiat yhteen vuoden ajalta. Tämän jälkeen laskettiin erikseen vuosien 2017 ja 2018 ulkoiset indeksit jokaiselle kaksiviikkojaksolle määrittämällä tarkasteltavan kaksiviikkojakson energian ja kyseisen vuoden vuosienergian suhde. Lopullinen kaksiviikkojakson ulkoinen indeksi saatiin ottamalla vuosien 2017 ja 2018 ulkoisten indeksien vuosienergioilla painotettu keskiarvo. Alla oleva kaava 25 havainnollistaa vuosienergioilla painotetun keskiarvon laskemista:

$$\frac{W_1}{W_1 + W_2} \times u_1 + \frac{W_2}{W_1 + W_2} \times u_2 \quad (25)$$

missä

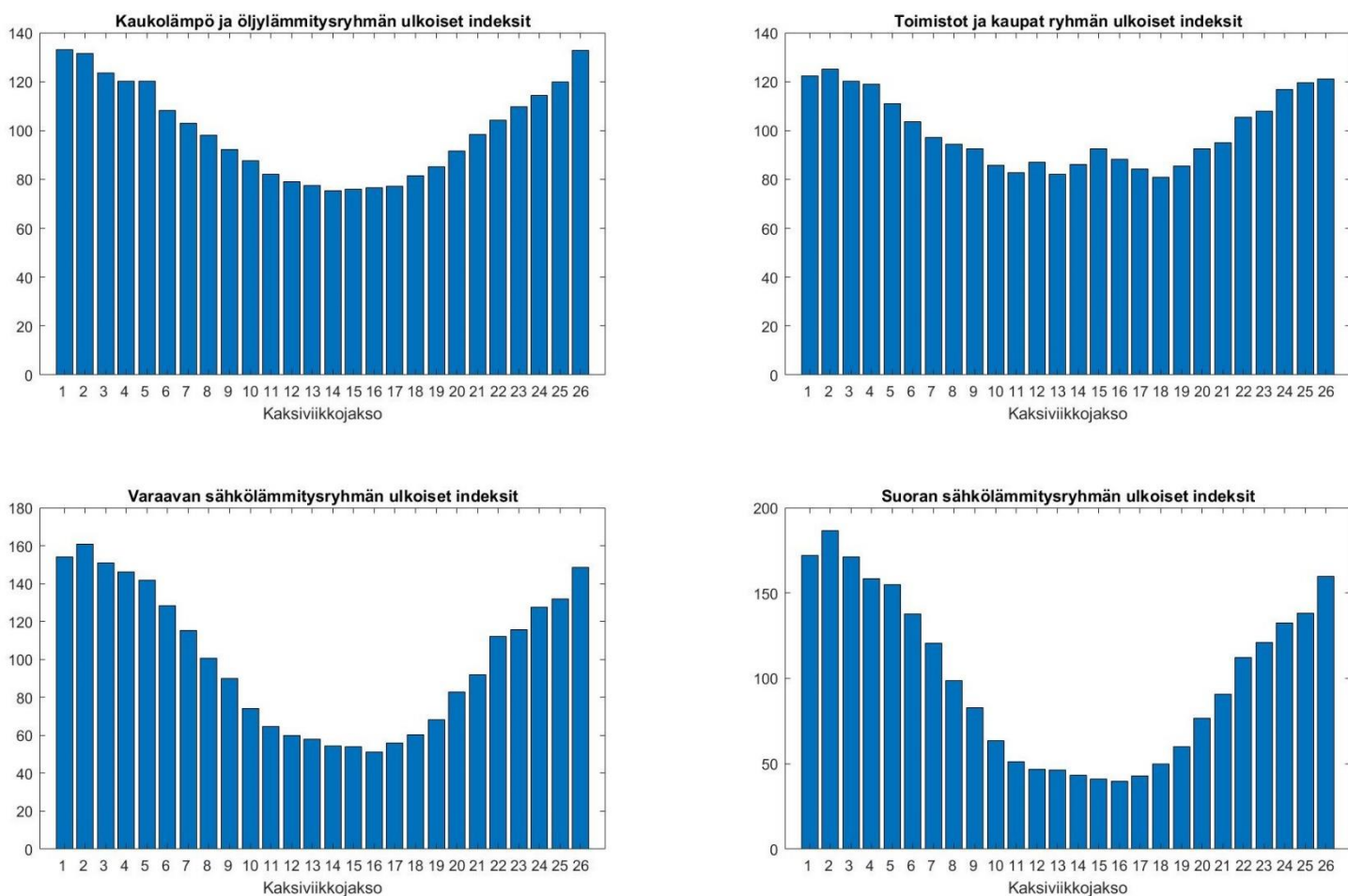
W_1 on vuoden 2017 vuosienergia

W_2 on vuoden 2018 vuosienergia

u_1 on vuoden 2017 tarkasteltavan kaksiviikkojakson ulkoinen indeksi

u_2 on vuoden 2018 tarkasteltavan kaksiviikkojakson ulkoinen indeksi

Kaavaa 25 käyttämällä saatiin siis kuluttajaryhmälle laskettua kaikki 26 lopullista ulkoista indeksiä. Ulkoiset indeksit skaalattiin vielä siten, että niiden keskiarvoksi tuli 100 ja summaksi 2600. Tämä tehtiin kertomalla lopulliset ulkoiset indeksit luvulla $26 \times 100 = 2600$. Skaalaus tehtiin Trimblen tietokantataulujen rakenteen vuoksi. Alla olevassa kuvassa 21 on havainnollistettu kaikkien kuluttajaryhmien ulkoisia indeksejä.



Kuva 21: Kuluttajaryhmien ulkoiset indeksit

Ulkoisista indekseistä nähdään, että talven kaksiviikkokajaksoilla indeksien arvot ovat korkeimmat ja kesän kaksiviikkokajaksoilla matalimmat. Tämä vaikuttaa järkevältä, koska talvella lämmitykseen kuluu luonnollisesti enemmän energiaa. Toimistot ja kaupat ryhmän ulkoiset indeksit kesän kaksiviikkokajaksoilla eivät ole niin matalat luultavasti siitä syystä, että jäähdytykseen kuluu paljon energiaa.

5.8 Korrelaation huomioiminen indeksisarjoissa

Kuten jo aliluvussa 3.2 todettiin, korrelaatiolla on keskeinen merkitys sekä kuluttajaryhmien sisällä että eri kuluttajaryhmien välillä. Tässä työssä korrelaation vaikutukseen ei kiinnitetty erityisemmin huomiota. Korrelaatio huomioitiin summaamalla tuntimitattu data yhteen, laskemalla keskiarvo jokaiselle vuorokauden tunnille päivätyypeittäin kaksiviikkojakson sisällä erikseen sekä tarkastelemalla lämpötilan aiheuttamaa tehonmuutosta. Oletettiin siis suuren määrän tuntimitatun raakadatan summaamisen riittävän korrelaation huomioimiseen. Lisäksi poistamalla lämpötilan vaikutus tuntitehoista ennen indeksisarjojen muodostamista voitiin varmistua siitä, ettei lämpötilan ja tuntimittauksen välinen korrelaatio vaikuta saatuihin tuloksiin. Samalla eri vuosien tuntimittauksista saatiin keskenään vertailukelpoisia, kuten jo aikaisemmin on mainittu.

Lämpötilariippuvuuskertoimien laskennassa tutkittiin vuorokauden keskilämpötilan ja tuntimittausten välistä korrelaatiota. Vertaamalla kahden peräkkäisen viikon samojen viikonpäivien tuntimittauksia toisiinsa ja tutkimalla miten näitä viikonpäiviä vastaavien edellisten vuorokausien keskilämpötilat poikkesivat toisistaan, voitiin seurata lämpötilan muutoksen vaikutusta tuntimittauksiin. Määritettyjen lämpötilariippuvuuskertoimien avulla lasketuilla tuntitehoilla saatiin korrelaatio myös huomioitua.

5.9 Tehokertoimet kuormituskäyrissä

Työn alussa oletettiin, että loistehosta olisi löytynyt myös tuntimittauksia, jolloin näiden tapauksessa olisi voitu toimia samalla tavalla kuin pätötehomittaustenkin kanssa. Klusteroinnissa käytetyiltä käyttöpaikoilta loistehomittauksia oli saatavilla kuitenkin niin vähän, ettei ollut järkevää lähteä mallintamaan loistehokulutusta samalla tavalla kuin pätötehoa. Tarkoitus oli laskea käyttämällä tuntimitattuja loisteho- ja pätötehoarvoja tehokertoimet jokaiselle tarkasteltavalle kaksiviikkojakson tunnille erikseen arkipäiville, aattopäiville ja pyhäpäiville. Seuraavaksi olisi määritetty sisäiset indeksit jokaiselle päivätyypille erikseen ottamalla keskiarvo samaa vuorokauden tuntia kuvaavista lasketuista tehokertoimista. Tällöin olisi saatu samaan tapaan kuin

pätötehonkin yhteydessä kaksiviikkojaksojen sisäiset indeksit päivätyypeittäin kaikille 26 kaksiviikkojaksolle. Nämä olisi syötetty sisäisten indeksien tauluun tietokantaan. Loistehomittauksia oli saatavilla ylipäättään vain harvoille käyttöpaikoille. Tehokertoimien laskemisessa olisi käytetty peruskaavaa:

$$\cos(\varphi) = \frac{P}{\sqrt{P^2 + Q^2}} \quad (26)$$

missä

$\cos(\varphi)$ on tietyn vuorokauden tunnin tehokerroin

P on tiettyä ajanhetkeä vastaava pätötehon tuntimittaus

Q on tiettyä ajanhetkeä vastaava loistehon tuntimittaus

Edellä mainittujen syiden takia tehokertoimia ei laskettu yllä olevaa kaavaa käyttäen. Loistehoa päätettiin mallintaa vanhalla tavalla eli kiinteitä tehokertoimia käyttäen. Jokaisen neljän kuluttajaryhmän tapauksessa käytettiin tehokertoimena 0,95:tä. Kertoimet syötettiin tietokantaan järjestelmien toimintaa varten. Myöskään suoraan lineaarisella regressiolla mallintamisella ei katsottu olevan työn kannalta niin merkittävää vaikutusta, koska alkuperäisenä tarkoituksena työssä oli mallintaa loistehoa pätötehon tapaan puhdasta tuntimitattua dataa käyttäen. Kiinteiden tehokertoimien käytössä tulee kuitenkin tiedostaa, että niissä ei huomioida loistehon ajasta riippuvaa vaihtelevuutta eikä asiakkaan loistehokulutus ja pätötehokulutus ole suorassa suhteessa toisiinsa nähden [2].

5.10 Asiakaskohtainen/Verkkoyhtiökohtainen mallintaminen

Kuten jo aikaisemmin on mainittu, verkkoyhtiökohtaisella alueellisella mallintamisella saadaan luotua huomattavasti tarkempia indeksisarjoja verrattuna valtakunnalliseen mallintamiseen. Valtakunnallisissa SENERin malleissa lämpötiloina on käytetty usean eri alueen mittauksista laskettuja keskilämpötiloja, jotka mallintavat sijainnista riippuen paremmin tai huonommin toisten alueiden keskilämpötiloja. Asiakaskohtaisessa alueellisessa mallintamisessa käytetään puolestaan tarkasteltavan alueen eli tässä työssä Kajaanin lentoaseman keskilämpötiloja. Tällöin saadaan laskettua huomattavasti tarkemmat kuluttajaryhmäkohtaiset lämpötilariippuvuuskertoimet. Mitatun aikasarjan normalisoinnilla eli poistamalla lämpötilan vaikutus tuntimittauksista saadaan myös tarkemmin todellisuutta vastaavaa luokkaa olevat tuntitehot.

Sähkönkulutuksen tuntimittausten, aluekohtaisten vuorokauden keskilämpötilojen ja normaalilämpötilojen perusteella muodostetut indeksisarjat ennustavat yleisiä valtakunnallisia malleja tarkemmin verkkoyhtiön asiakkaiden sähkönkulutusta. Pitkän aikavälin ennustamisesta tulee tämän seurauksena tarkempaa ja luotettavampaa. Verkkoyhtiökohtaisen mallintamisen on mahdollistanut, kuten jo aikaisemmin mainittiin, tuntimittausten yleistyminen.

6. Tulokset ja johtopäätökset

6.1 Käytetty tulosten vertailutapa

Muodostettujen indeksisarjojen, jotka nimettiin verkkoyhtiö Kajaven mukaan Kajave20 -sarjoiksi, hyvyyden selvittämiseksi niitä tuli verrata verkkoyhtiö Kajavella nykyisin käytössä olevien indeksisarjojen kanssa. Kajavella tällä hetkellä käytössä olevat GGK062003 -indeksisarjat perustuvat SENERin vuoden 1995 indeksisarjoihin. Näihin on kuitenkin tehty parannuksia lisäämällä sellaisille kulutuksille tyyppikäyriä, joita SENERin sarjat eivät pystyneet mallintamaan. Näistä esimerkkinä on hiihtokeskuksille sovellettu oma käyrä, johon ei soveltunut lumetuksen ja majoituskulutuksen ajallisen sijoittumisen poikkeuksellisuuden takia mikään oletuskäyristä. Kajaven omien indeksisarjojen taustalla vaikuttaa myös VTT:n 2000-luvun alussa tekemä kuormitustutkimus sekä maaseutuverkkoyhtiöiden R8-ryhmä. [18]

Muodostettujen Kajave20 -indeksisarjojen ja GGK062003 -sarjojen paremmuuden vertailemiseksi valittiin jokaisesta Kajave20 -indeksisarjojen neljästä muodostuneesta kuluttajaryhmästä sellaiset käyttöpaikat, jotka edustavat GGK062003 -indeksisarjoissa myös samaa kuluttajaryhmää. Tällä tavoin vertailu oli mahdollista suorittaa. Kajave20 -sarjojen kuluttajaryhmän kaukolämpö ja öljylämmitys joukosta valittiin käyttöpaikat, jotka kuuluvat GGK062003 -sarjoissa kuluttajaryhmään omakotitalot, ei sähkölämmitystä, ei sähkökuuasta. Käyttöpaikkoja valittiin yhteensä 31. Kajave20 -sarjojen kuluttajaryhmän toimistot ja kaupat joukosta valittiin puolestaan käyttöpaikat, jotka kuuluvat osaksi GGK062003 -sarjojen kuluttajaryhmää yksityinen palvelu, päivä h-aika 1699h. Näitä valittiin 5 kappaletta. Molempien Kajave20 -sarjojen kuluttajaryhmien varaava sähkölämmitys sekä suora sähkölämmitys tapauksessa valittiin käyttöpaikat, jotka kuuluvat GGK062003 -sarjoissa kuluttajaryhmään omakotitalot, sähkölämmitys, käyttövesivaraaja 300l. Ryhmästä varaava sähkölämmitys tällaisia käyttöpaikkoja valittiin 32 kappaletta ja suoran sähkölämmitysryhmän tapauksessa puolestaan 35 kappaletta.

Kuluttajaryhmästä valittujen käyttöpaikkojen vuoden 2019 vuosienenergiat laskettiin yhteen, jolloin saatiin valittujen käyttöpaikkojen yhteinen vuosienenergia jokaisen neljän

kuluttajaryhmän otannalle erikseen. Näin laskettua vuosienergiaa käytettiin sekä Kajave20 -indeksisarjojen että GGK062003 -indeksisarjojen vuosienergiana tuntitehojen laskemisessa indeksisarjoista. Vertailua ei tehty yksittäisille käyttöpaikoille siitä syystä, että yksittäisen kuluttajan käytös on hyvin vaihtelevaa eikä mikään malli pysty kovinkaan tarkasti arvioimaan yksittäisen kuluttajan sähkönkulutusta. Puolestaan yli kymmenen kuorman summakuorma muistuttaa jo huomattavasti paremmin mallien antamaa keskiarvoa. Yli 10 käyttöpaikan summakuorma mahdollisti myös sen, että kaupunkiverkossa voitiin mallintaa liittymätason käyttäytymistä ja maaseutuverkossa puolestaan runkojohtotason käytöstä.

Vertailun suorittamiseksi valittiin vuoden ajalta neljä päivää, jotka kiinnitettiin vuoden 2019 kalenteriin. Tämä tehtiin siitä syystä, että vuosienergia laskettiin vuoden 2019 tehoista. Indeksisarjojen muodostamisessa ei käytetty vuoden 2019 tuntimitattuja tehoja, vaan kyseisen vuoden tehoja käytettiin vertailun kohteena eli referenssinä tulosten tarkastelussa. Vuoden 2019 päivästä valittiin lämpötilan perusteella vuoden kylmin päivä (24.1), vuoden kuumin päivä (27.7), leudompi talvipäivä (11.1) ja viileämpi kesäpäivä (15.7). Tarkastelun kohteeksi valittiin vuoden kylmin ja kuumin päivä siitä syystä, että tällöin tarkastellaan pahimpia mahdollisia tilanteita tehonjakolaskennan kannalta, koska näinä päivinä laskennan antama tarkkuus voi poiketa todellisista tuloksista eniten. Tämä johtuu siitä, että vuoden kylmimmän ja kuumimman päivän lämpötilojen poikkeama ajanjakson referenssilämpötilasta on suurin. Lämpötilan perusteella valittujen talvipäivien (24.1 ja 11.1) vastaavina tuntimittauksina käytettiin seuraavan vuorokauden tuntimittauksia lämpötilan vaikutuksen ajallisen viiveen huomioimiseksi. Kesäpäivien tapauksessa käytettiin samojen päivien tuntimittauksia, koska jäähdytyksen vaikutus ilmenee jo saman päivän aikana toisin kuin lämmityksen. Tuntimittaukset olivat siis päiviltä 25.1 (arkipäivä), 12.1 (aattopäivä), 27.7 (aattopäivä) sekä 15.7 (arkipäivä). Valitut päivät kuuluivat kaksiviikkojaksoihin 1, 2, 14 ja 15.

Sekä Kajave20 -indeksisarjoista että GGK062003 -indeksisarjoista laskettiin edellä mainittujen vertailussa käytettävien päivien tehot tunneittain. Tietyn ajanhetken teho laskettiin seuraavalla kaavalla:

$$P_{m,ki} = \frac{E_k}{8760} \times \frac{M_{k,i}}{100} \times \frac{m_{k,i}}{100} \quad (27)$$

missä

E_k on kuluttajaryhmästä k valittujen käyttöpaikkojen vuoden 2019 vuosienenergioiden summa

$M_{k,i}$ on kuluttajaryhmän k kaksiviikkojakson i ulkoinen indeksi (keskiarvo 100)

$m_{k,i}$ on kuluttajaryhmän k kaksiviikkojakson i sisäinen indeksi (keskiarvo 100)

Molemmista indeksisarjoista laskettuja tunnittaisia tehoja verrattiin vuodelta 2019 valittujen päivien puhtaaseen tuntimittausten vertailudataan, jota ei käytetty Kajave20 -indeksisarjojen muodostamisessa. Tämän vuoksi vertailu oli mahdollista suorittaa kyseisellä tavalla. Kuten jo edellä sivuttiin, vertailussa ei käytetty vuoden 2019 todellista yksittäisten käyttöpaikkojen tuntimittaustietoa. Kuluttajaryhmästä valittujen käyttöpaikkojen todellinen tuntimitattu data laskettiin yhteen, jolloin saatiin tarkasteltaville päville summakuorma. Indeksisarjoista laskettuja vastaavien päivien tuntitehoja vertaamalla todelliseen tuntimittausten vertailudataan oli mahdollista tarkastella kuluttajaryhmäkohtaisesti, mallintaako Kajave20 -indeksisarjat vai GGK062003 -indeksisarjat paremmin tarkasteltavien päivien kulutusta. Vertailuun mukaan otettiin vielä kolmas vaihtoehto eli Kajave20 -indeksisarjoista lasketut tuntitehot, joiden ulkoinen indeksi lämpötilakorjattiin kaavan 28 osoittamalla tavalla:

$$M_{k,i,\text{lämpötilakorjattu}} = M_{k,i} \times (1 + c \times \Delta T) \quad (28)$$

missä

$M_{k,i}$ on kuluttajaryhmän k kaksiviikkojakson i ulkoinen indeksi

c on vuosijaksolla käytettävä lämpötilariippuvuuskerroin

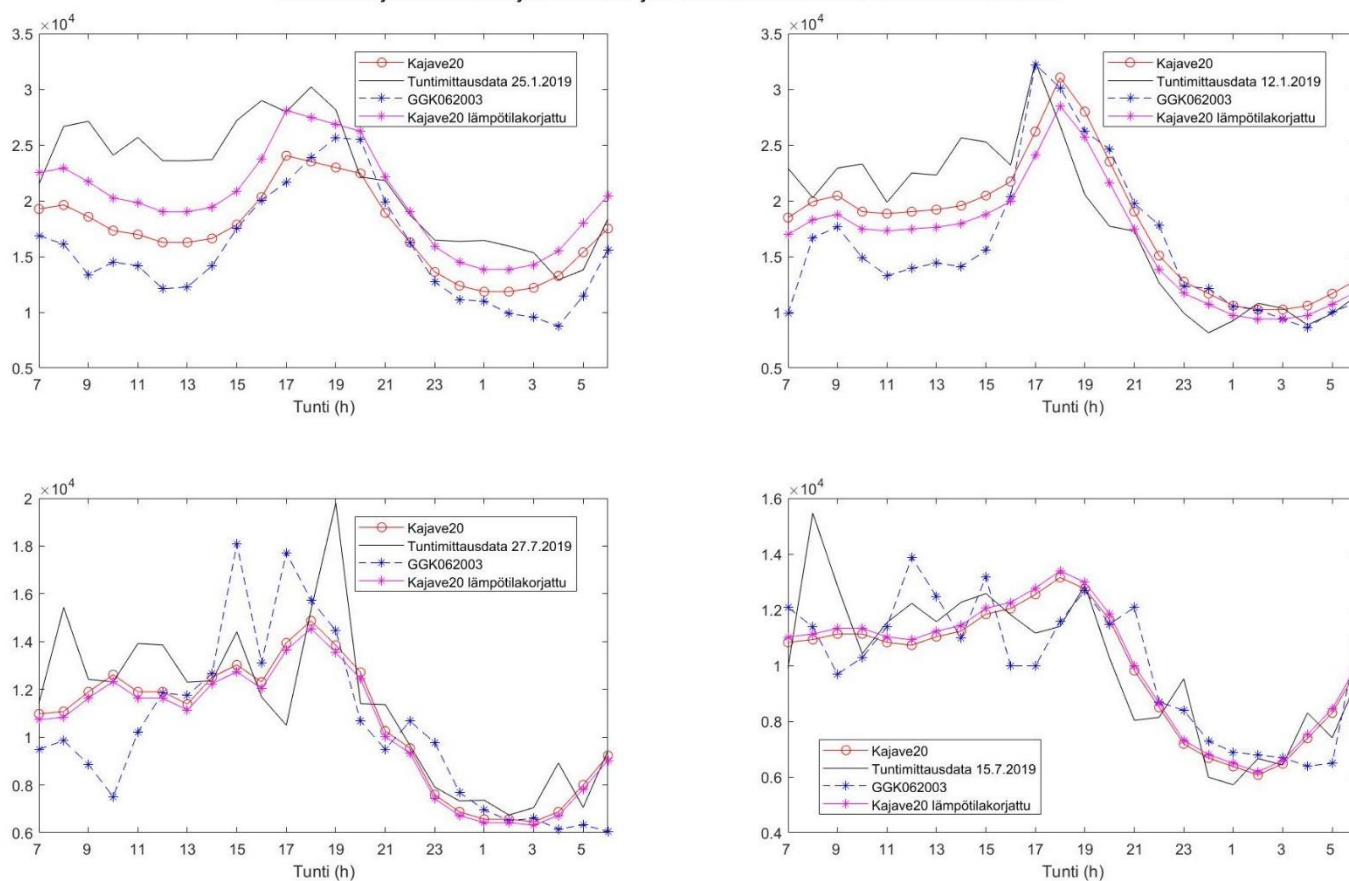
ΔT on vuorokauden keskilämpötilan ja kaksiviikkojakson normaalilämpötilan lämpötilaero

GGK062003 -sarjojen ulkoisia indeksejä ei lämpötilakorjattu, koska lämpötilariippuvuuskerroin ei ollut saatavilla. Verrattiin siis näitä kolmea edellä mainittua vaihtoehtoa kuluttajaryhmäkohtaisesti vuoden 2019 todelliseen tuntimittausten vertailudataan. Näin saatiin luotua johtopäätökset indeksisarjojen paremmuudesta. Tuloksia tarkastellaan seuraavassa alikappaleessa 6.2.

6.2 Tulosten vertailu

Kuten aiemmin jo mainittiin, muodostettujen Kajave20 -sarjojen ja nykyisin Kajavella käytössä olevien GGK062003 -sarjojen vertailemiseksi laskettiin molemmista indeksisarjoista tuntitehot tarkasteltaville päiville. Kajave20 -sarjojen kaukolämpö- ja öljylämmitysryhmästä valittiin 31 käyttöpaikkaa, jotka kuuluvat GGK062003 -sarjoissa kuluttajaryhmään omakotitalot, ei sähkölämmitystä, ei sähkökiuasta. Valittujen käyttöpaikkojen todellisesta tuntimitatusta datasta laskettiin summakuorma. Tällöin yksittäisten käyttöpaikkojen tarkastelun sijaan saatiin luotettavampia arvioita vertailtavien sarjojen paremmuudesta Kajave20 -sarjojen jokaisen neljän kuluttajaryhmän tapauksessa. Kuva 22 havainnollistaa, kuinka hyvin Kajave20 -sarjoista sekä GGK062003 -sarjoista lasketut tuntitehot mallintavat kaukolämpö- ja öljylämmitysryhmästä valittujen käyttöpaikkojen todellisten tuntimittausten summakuormaa tarkasteltavien päivien tapauksessa.

Indeksisarjoista laskettujen tuntitehojen vertaus todelliseen tuntimittausdataan



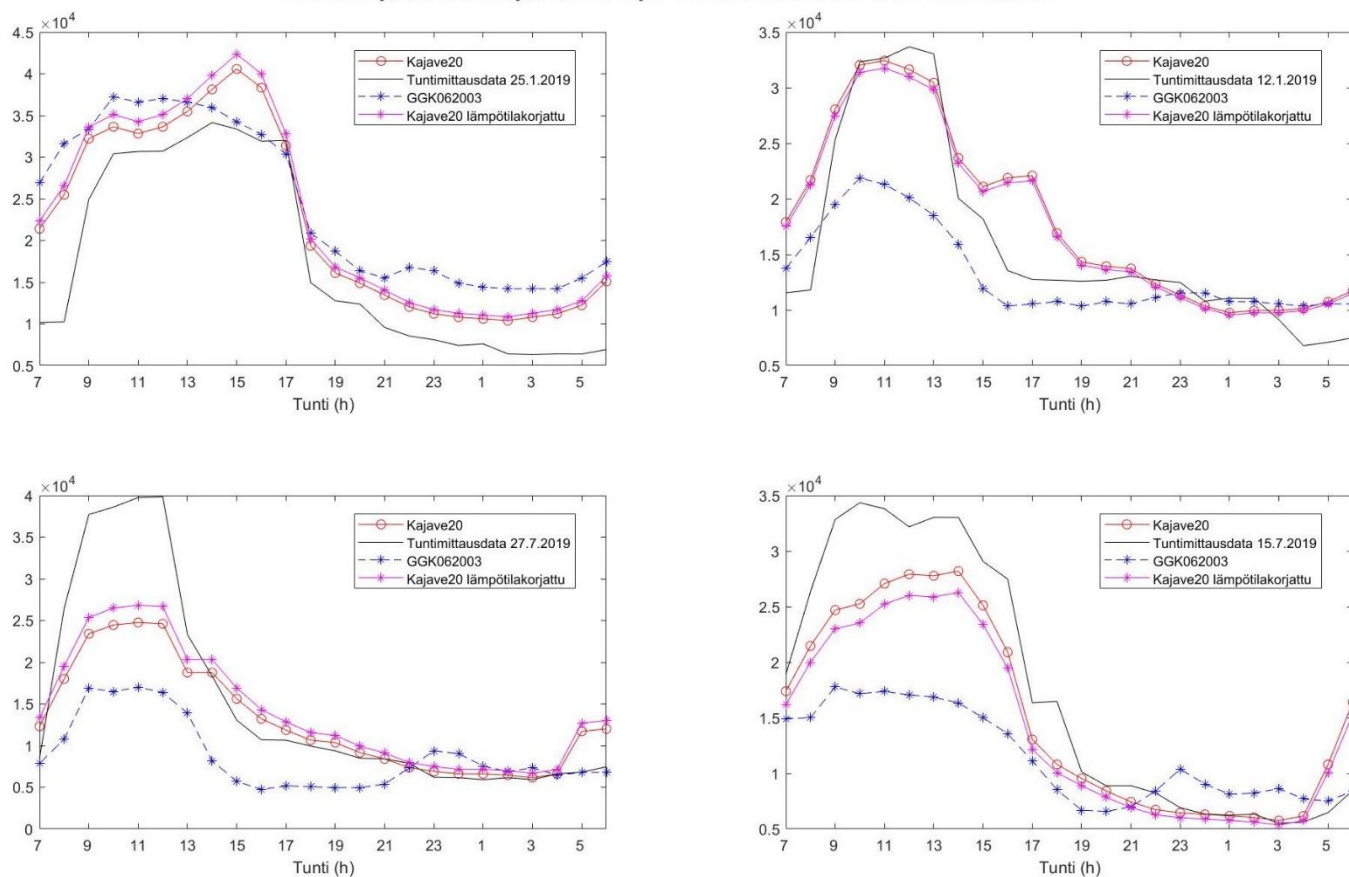
Kuva 22: Indeksisarjoista laskettujen tuntitehojen vertaus kaukolämpö- ja öljylämmitysryhmästä valittujen käyttöpaikkojen todellisten tuntimittausten summakuormaan

Kuvasta 22 nähdään, että keskimäärin Kajave20 -sarjoista lasketut tuntitehot mallintavat todellisen tuntimittausdatan summakuormaa paremmin kuin GGK062003 -sarjoista lasketut tuntitehot. Vuoden kylmimpänä päivänä Kajave20 -sarjoista lasketut lämpötilakorjatut tehot mallintavat parhaiten todellista kulutusta. Kesäpäivinä pienestä lämpötilakorjauksesta johtuen Kajave20 tehojen ja Kajave20 lämpötilakorjattujen tehojen profiilit ovat lähes samat. Asiakaskohtainen mallintaminen on varmasti yksi tekijä Kajave20 -sarjoista laskettujen tuntitehojen ja lämpötilakorjattujen tuntitehojen paremmuuteen GGK062003 -sarjoista laskettuihin tehoihin nähden. Toisena mahdollisena tekijänä voidaan pitää Kajaanin lentoaseman lämpötilojen käyttöä

valtakunnallisten lämpötilojen sijaan. Kesäpäivinä Kajave20 -sarjoista molemmilla tavoilla lasketut tuntitehot eivät vastaa vuorokauden kulutushuipputuntiin kovin hyvin, mutta myöskään GGK062003 -sarjoista lasketut tehot eivät mallinna huipputuntia sen paremmin.

Kajave20 -sarjojen toimistot ja kaupat ryhmästä valittiin vain 5 käyttöpaikkaa tulosten tarkasteluun. Syynä tähän oli klusteroinnin tuloksena kyseiseen ryhmään sijoittuneiden asiakkaiden vähäinen määrä. Valitut käyttöpaikat kuuluvat GGK062003 -sarjoissa kuluttajaryhmään yksityinen palvelu, päivä h-aika 1699h. Alla olevasta kuvasta 23 nähdään, miten Kajave20 -sarjoista sekä GGK062003 -sarjoista lasketut tuntitehot mallintavat toimistot ja kaupat ryhmästä valittujen käyttöpaikkojen kulutuksen summakuormaa.

Indeksisarjoista laskettujen tuntitehojen vertaus todelliseen tuntimittausdataan

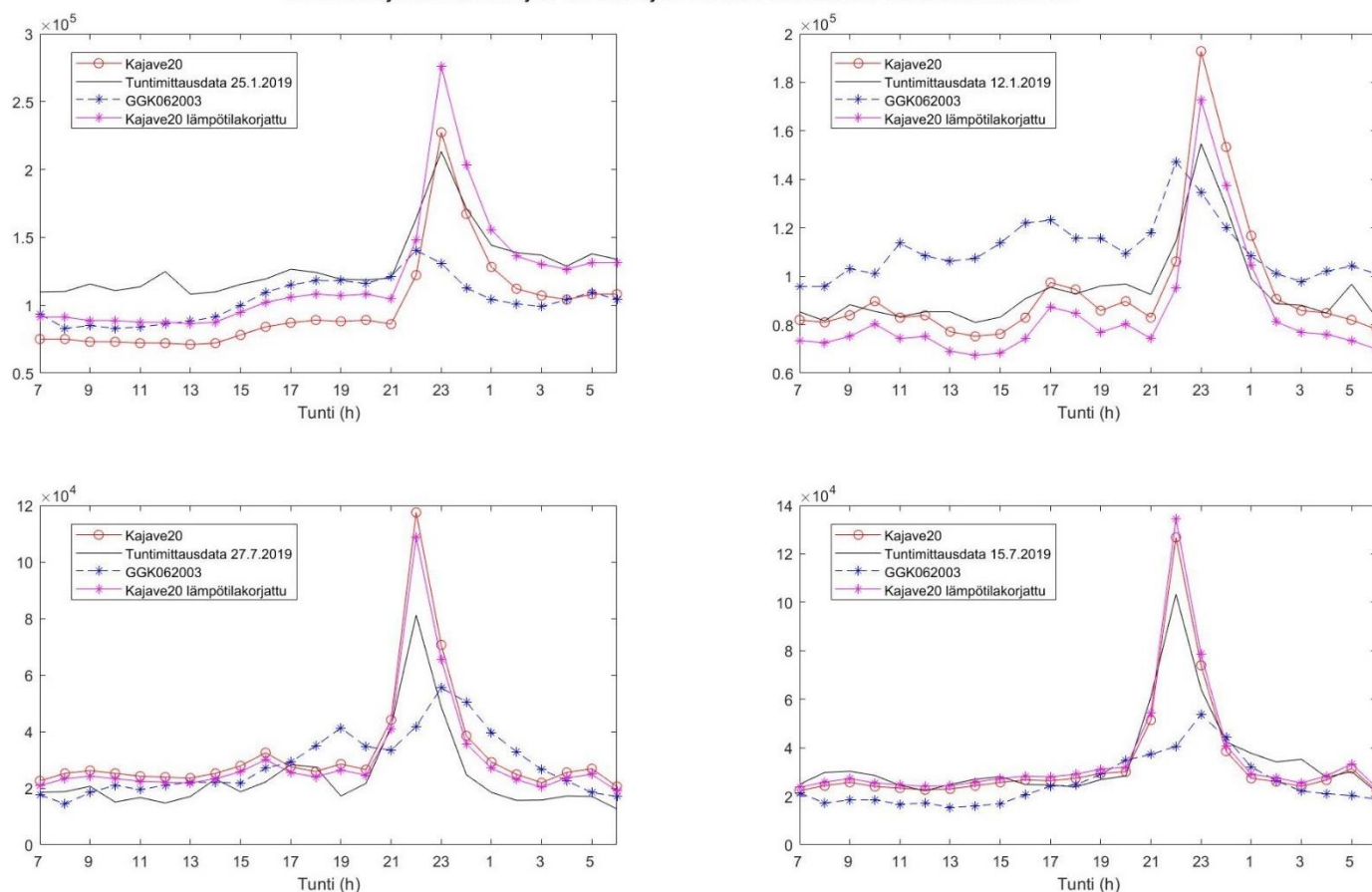


Kuva 23: Indeksisarjoista laskettujen tuntitehojen vertaus toimistot ja kaupat ryhmästä valittujen käyttöpaikkojen todellisten tuntimittausten summakuormaan

Kuvasta 23 voidaan havaita, että tarkasteltavina päivinä Kajave20 -sarjoista lasketut tuntitehot ja lämpötilakorjatut tuntitehot mallintavat todellista tuntimittausten vertailudataa lähes identtisesti. Kesäpäivinä ja talvipäivänä 12.1 GGK062003 -sarjoista laskettujen tuntitehojen huippuarvot ovat huomattavasti pienemmät verrattuna Kajave20 -sarjoista molemmilla tavoilla laskettuihin vastaaviin arvoihin. Kajave20 -sarjoista lasketut tehot myös mallintavat huippukulutusta selkeästi paremmin. Ainoastaan vuoden kylmimpänä päivänä maksimiarvon mallinnus on lähes samalla tasolla. Aattopäivinä sekä todellisen tuntimittaustiedon että indeksisarjoista laskettujen tuntitehojen päiväaikaa mallintavat profiilit ovat huomattavasti kapeammat kuin tarkasteltavien arkipäivien, mikä oli odotettua.

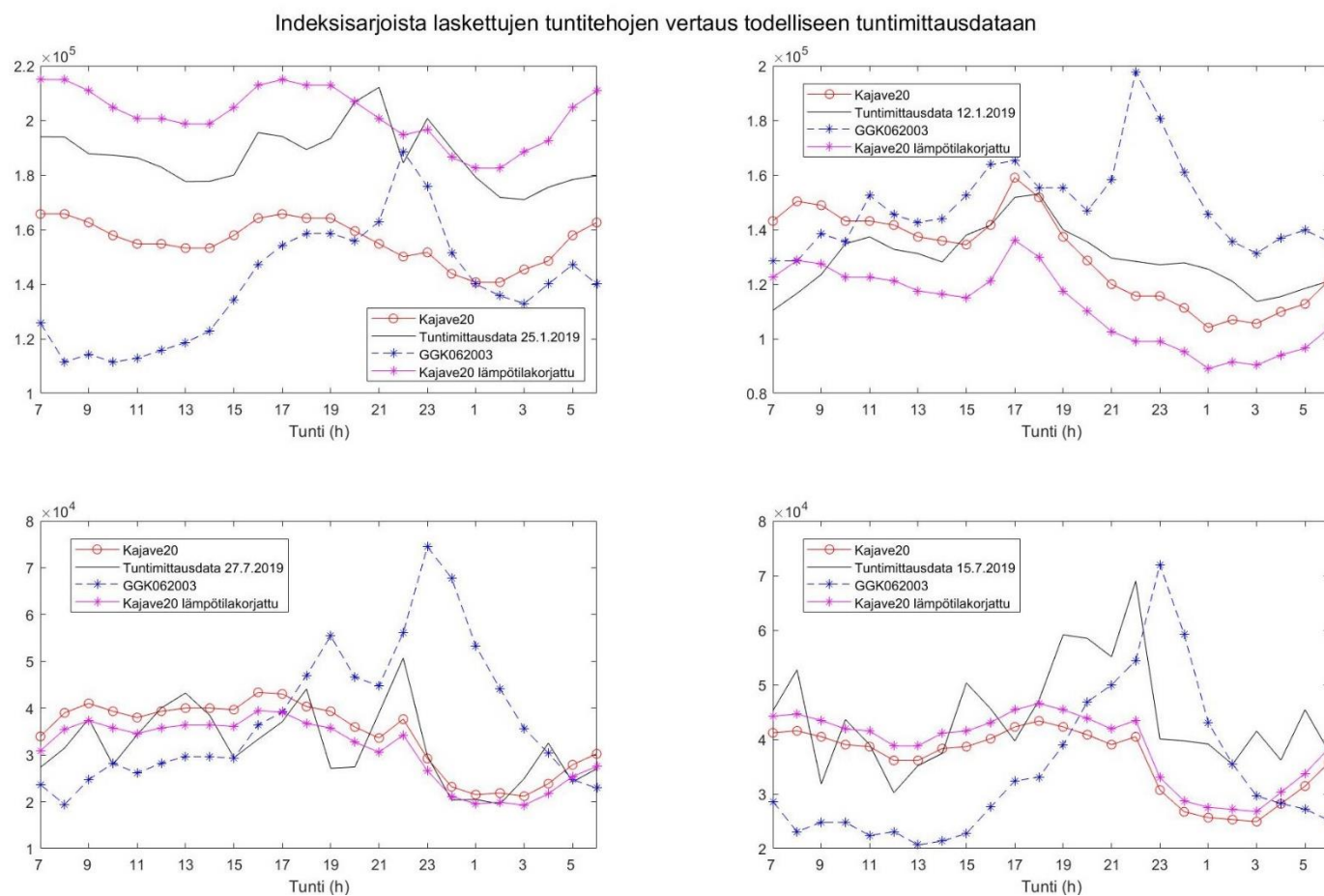
Kajave20 -sarjojen ryhmästä varaava sähkölämmitys puolestaan valittiin 32 käyttöpaikkaa, jotka kuuluvat GGK062003 -sarjoissa kuluttajaryhmään omakotitalot, sähkölämmitys, käyttövesivaraaja 300l. Suurin osa varaavan sähkölämmitysryhmän käyttöpaikoista kuuluvat edellä mainittuun ryhmään GGK062003 -sarjoissa. Kuva 24 havainnollistaa, kuinka tarkasti Kajave20 -sarjoista ja GGK062003 -sarjoista lasketut tuntitehot mallintavat varaavasta sähkölämmitysryhmästä valittujen käyttöpaikkojen todellisen kulutuksen summakuormaa. Kuvasta 24 voidaan havaita, että Kajave20 -sarjoista molemmilla tavoilla laskettujen tuntitehojen huippukulutustunti ajoittuu todellisen tuntimittaustiedon huipputunnin kanssa samaan aikaan. Indeksisarjoista laskettu tuntitehojen kulutushuippu eli varaavien sähkölämmittimien päälle kytkemistä vastaava aika, nousee kuitenkin liian suureksi. GGK062003 -sarjoista lasketut tuntitehot eivät mallinna yhtä tarkasti valittujen käyttöpaikkojen tuntimitatun kulutuksen summakuormaa. Varaavan sähkölämmityksen tapauksessa lämpötila ei vaikuta niinkään maksimitheoon, vaan enemmänkin siihen, kuinka monta tuntia maksimiteho on päällä aika-siirto-tariffin vaihtumisen jälkeen. Kajave20 -sarjoista laskettujen tuntitehojen ja lämpötilakorjattujen tuntitehojen profiilit ovat tarkasteltavina kesäpäivinä lähes identtiset, mutta talvella niissä on joitakin eroja havaittavissa.

Indeksisarjoista laskettujen tuntitehojen vertaus todelliseen tuntimittausdataan



Kuva 24: Indeksisarjoista laskettujen tuntitehojen vertaus varaavasta sähkölämmitysryhmästä valittujen käyttöpaikkojen todellisten tuntimittausten summakuormaan

Neljännestä Kajave20 -sarjojen kuluttajaryhmästä suora sähkölämmitys valittiin 35 käyttöpaikkaa, joista jokainen kuuluu GGK062003 -sarjoissa myös kuluttajaryhmään omakotitalot, sähkölämmitys, käyttövesivaraaja 300l. Kuvasta 25 nähdään, miten sekä Kajave20 -sarjoista että GGK062003 -sarjoista lasketut tuntitehot mallintavat suorasta sähkölämmitysryhmästä valittujen käyttöpaikkojen kulutuksen summakuormaa.



Kuva 25: Indeksisarjoista laskettujen tuntitehojen vertaus suorasta sähkölämmitysryhmästä valittujen käyttöpaikkojen todellisten tuntimittausten summakuormaan

Kuvasta 25 havaitaan, että Kajave20 -sarjoista laskettujen lämpötilakorjattujen tuntitehojen profiili vastaa parhaiten vuoden kylmimmän päivän kulutusta. Mallin arvioima kulutus on kuitenkin keskimäärin jonkin verran suurempaa kuin todellinen kulutus. Samana päivänä myös Kajave20 -sarjoista laskettujen tehojen ja lämpötilakorjattujen tehojen profiilit poikkeavat merkittävästi toisistaan verrattuna muihin tarkasteltaviin päiviin. Syynä tähän on tuntimittauksia edeltävän vuorokauden keskilämpötilan merkittävä poikkeama kaksiviikkojakson normaalilämpötilasta. Poikkeama on noin 15 astetta eli yli kaksinkertainen. Lämpötilakorjaus on kyseisessä tapauksessa siis suuri. Kuvasta voidaan nähdä, että GGK062003 -sarjoista lasketut

tuntitehot eivät mallinna hyvin minkään tarkasteltavan päivän kulutusta. Kulutushuippujen arviointi on hyvin epätarkkaa. Kesäpäivinä Kajave20 -sarjoista molemmilla tavoilla lasketut tuntitehot eivät kovin hyvin mallinna todellisen tuntimitatun datan huippuarvoja, mutta keskimäärin mallinnus on melko tarkkaa.

Seuraavat taulut 10, 11, 12 ja 13 havainnollistavat indeksisarjoista laskettujen tuntitehojen maksimivirheitä ja virheen itseisarvon keskiarvoja jokaisen kuluttajaryhmän jokaisen tarkasteltavan päivän tapauksessa. Maksimivirheellä tarkoitetaan indeksisarjoista laskettujen tuntitehojen suurinta poikkeamaa todellisista tuntimittauksista. Virheen itseisarvon keskiarvo puolestaan tarkoittaa indeksisarjoista laskettujen tehojen keskimääräistä poikkeamaa todellisista tuntimittauksista. Poikkeamat on ilmaistu kilowatteina.

<u>Kaukolämpö ja öljylämmitys</u>	Kajave20	GGK062003	Kajave20 lämpötilakorjattu
Virheen itseisarvon keskiarvo 25.1	4,835	6,642	2,949
Virheen itseisarvon keskiarvo 12.1	3,034	4,650	3,227
Virheen itseisarvon keskiarvo 27.7	1,276	2,261	1,386
Virheen itseisarvon keskiarvo 15.7	1,081	1,318	1,101
Virheen maksimiarvo 25.1	9,355	13,75	6,343
Virheen maksimiarvo 12.1	7,467	13,03	8,383
Virheen maksimiarvo 27.7	5,954	7,220	6,260
Virheen maksimiarvo 15.7	4,539	4,086	4,341

Taulu 10: Kaukolämpö- ja öljylämmitysryhmän indeksisarjoista laskettujen tuntitehojen maksimivirheet ja virheen itseisarvon keskiarvot yksikössä kW tarkasteltavien päivien tapauksessa

Taulusta 10 nähdään, että virheen itseisarvon keskiarvo on suurin GGK062003 -sarjoista laskettujen tuntitehojen tapauksessa kaikkina tarkasteltavina päivinä. Sama pätee virheen maksimiarvolle lukuun ottamatta päivää 15.7, jolloin Kajave20 -sarjoista laskettujen tuntitehojen virheen maksimiarvo on suurin.

<u>Toimistot ja kaupat</u>	Kajave20	GGK062003	Kajave20 lämpötilakorjattu
Virheen itseisarvon keskiarvo 25.1	4,914	7,120	5,827
Virheen itseisarvon keskiarvo 12.1	3,030	4,262	2,961
Virheen itseisarvon keskiarvo 27.7	4,040	7,123	4,158
Virheen itseisarvon keskiarvo 15.7	3,423	7,386	4,255
Virheen maksimiarvo 25.1	15,25	21,42	16,36
Virheen maksimiarvo 12.1	9,892	14,53	9,434
Virheen maksimiarvo 27.7	15,23	23,48	13,18
Virheen maksimiarvo 15.7	9,097	17,18	10,82

Taulu 11: Toimistot ja kaupat ryhmän indeksisarjoista laskettujen tuntitehojen maksimivirheet ja virheen itseisarvon keskiarvot yksikössä kW tarkasteltavien päivien tapauksessa

Yllä olevasta taulusta 11 havaitaan, että virheen itseisarvon keskiarvo on suurin GGK062003 -sarjoista laskettujen tuntitehojen tapauksessa jokaisena tarkasteltavana päivänä. Sama pätee virheen maksimiarvoille.

Taulusta 12 puolestaan huomataan, että vuoden kylmimpänä päivänä Kajave20 -sarjoista laskettujen tehojen virheen itseisarvon keskiarvo on suurin. Toisina tarkasteltavina päivinä keskimääräinen virhe on suurin GGK062003 -sarjoista laskettujen tuntitehojen tapauksessa. Virheen maksimiarvo on suurin GGK062003 -sarjoista lasketuista tuntitehoista lukuun ottamatta päivää 12.1.

<u>Varaava sähkölämmitys</u>	Kajave20	GGK062003	Kajave20 lämpötilakorjattu
Virheen itseisarvon keskiarvo 25.1	32,19	25,75	18,66
Virheen itseisarvon keskiarvo 12.1	7,744	19,24	12,68
Virheen itseisarvon keskiarvo 27.7	8,999	9,286	6,971
Virheen itseisarvon keskiarvo 15.7	4,864	10,07	5,007
Virheen maksimiarvo 25.1	52,75	82,47	62,82
Virheen maksimiarvo 12.1	38,20	32,83	23,29
Virheen maksimiarvo 27.7	36,30	39,36	27,55
Virheen maksimiarvo 15.7	23,45	62,78	31,18

Taulu 12: Varaavan sähkölämmitysryhmän indeksisarjoista laskettujen tuntitehojen maksimivirheet ja virheen itseisarvon keskiarvot yksikössä kW tarkasteltavien päivien tapauksessa

Seuraavan sivun taulusta 13 nähdään, että virheen itseisarvon keskiarvo on suurin Kajave20 -indeksisarjoista laskettujen lämpötilakorjattujen tuntitehojen tapauksessa päivänä 12.1, jolloin keskimääräinen virhe on hieman suurempi kuin GGK062003 -sarjoista laskettujen tehojen tapauksessa. GGK062003 -sarjoista laskettujen tuntitehojen tapauksessa virheen itseisarvon keskiarvo on kuitenkin suurin muina tarkasteltavina päivinä. Virheen maksimiarvo puolestaan on aina suurin GGK062003 -sarjoista lasketuista tuntitehoista.

<u>Suora</u> sähkölämmitys	Kajave20	GGK062003	Kajave20 lämpötilakorjattu
Virheen itseisarvon keskiarvo 25.1	31,20	47,52	16,58
Virheen itseisarvon keskiarvo 12.1	10,65	19,76	20,34
Virheen itseisarvon keskiarvo 27.7	5,548	12,71	4,731
Virheen itseisarvon keskiarvo 15.7	9,396	14,62	8,190
Virheen maksimiarvo 25.1	57,34	82,47	31,11
Virheen maksimiarvo 12.1	33,71	69,36	36,41
Virheen maksimiarvo 27.7	13,04	47,40	16,43
Virheen maksimiarvo 15.7	28,50	31,95	25,52

Taulu 13: Suoran sähkölämmitysryhmän indeksisarjoista laskettujen tuntitehojen maksimivirheet ja virheen itseisarvon keskiarvot yksikössä kW tarkasteltavien päivien tapauksessa

Tulosten tarkastelujen perusteella voidaan sanoa, että Kajave20 -indeksisarjoista molemmilla tavoilla lasketut tuntitehot mallintavat Kajavella nykyisin käytössä olevia GGK062003 -indeksisarjoja paremmin todellista tuntimitattua kulutusta valittuina neljänä päivänä. Oletus voidaan yleistää koskemaan myös muita päiviä, koska tutkimuksessa tarkasteltiin vuoden kylmintä ja kuuminta päivää. Näinä päivinä tulosten tarkkuus todennäköisimmin poikkeaa todellisista mittauksista. Täytyy kuitenkin huomioida, että tulosten tarkastelua varten valitut käyttöpaikat olivat osa klusteroinnin lähtöaineistoa. Indeksisarjojen paremmuuden vertailu olisi voitu tehdä myös valitsemalla käyttöpaikat, jotka eivät olleet klusteroinnissa mukana. Tätä ei kuitenkaan voitu tässä tapauksessa tehdä, koska kuluttajaryhmän päättelemisen klusteroinnin ulkopuolelle jätetyille käyttöpaikoille vaatisi jonkin yhteisen tekijän määrittämistä. Yhteisen tekijän perusteella käyttöpaikka liitettäisiin tiettyyn kuluttajaryhmään. Puhtaasti tuntimitattuun dataan perustuvassa kuluttajaryhmien muodostamisessa voi

samaan ryhmään tulla lämmitysmuodoiltaan hyvinkin monta erilaista käyttöpaikkaa. Tekoälyn avulla on kuitenkin mahdollista esimerkiksi käyttöpaikkakohtaiseen ulkoisten indeksien laskentaan perustuen määrittää kaikille klusteroinnin ulkopuolelle jätetyille käyttöpaikoille sopivin muodostuneista kuluttajaryhmistä. Tämä vaatii kuitenkin vielä lisäkehitystä eikä sen vuoksi menetelmää voitu suorittaa tähän työhön.

Saadut tulokset vaikuttavat kuitenkin lupaavilta ja uudet indeksisarjat ovat potentiaaliset parantamaan tehonjakolaskennan ja KAH-laskennan tarkkuutta. Uudet indeksisarjat näyttävät mallintavan vanhoja paremmin todellista käyttöpaikoilta mitattujen tuntimittausten summakuormaa, mutta uusien sarjojen käyttöönotto vaatii vielä lisätutkimusta. Kuten edellä jo mainittiin, indeksisarjojen paremmuutta olisi syytä vertailla vielä sellaisista käyttöpaikoista laskettujen todellisten tuntimittausten summakuormien kanssa, joita ei käytetty klusteroinnissa. Jos näistä saataisiin vastaavia tuloksia, voitaisiin uudet indeksisarjat ottaa käyttöön.

Muodostuneiden kuluttajaryhmien lukumäärä (4) on pieni verrattuna käytössä olevien indeksisarjojen ryhmien lukumäärään, mutta tärkeimmät kuluttajaryhmistä löydettiin klusteroinnin tuloksena. Lisäksi vanhoissa sarjoissa on monta nykyään tarpeetonta ryhmää. Monen asiakkaan kulutustottumukset ovat myös merkittävästi muuttuneet vanhojen indeksisarjojen perustamisajoista. Kuluttajaryhmien päivitys on siis tarpeen. Saadut indeksisarjat ja tulokset toimivat hyvänä pohjana lisäkehitystä varten.

Tulee kuitenkin muistaa, että sähkön käyttötottumukset tulevat tulevaisuudessa muuttumaan merkittävästi hajautetun tuotannon, aurinkovoiman, sähköautojen latauksen, kysyntäjouston ja sähkön varastointikapasiteetin kasvamisen seurauksena. Näitä indeksisarjojen käytettävyyteen ja tarkkuuteen vaikuttavia haasteita käsitellään luvussa 7.

6.3 Uusien sähkönkulutuskäyrien mahdolliset epätarkkuustekijät

Kuluttajaryhmien määrittämisessä ja indeksisarjojen muodostamisessa klusteroinnilla on lopputulosten kannalta keskeinen rooli. Jos klusteroinnin tulokset eivät ole tarpeeksi tarkat, niin indeksisarjat ja näistä lasketut tuntitehot estimoivat huonosti todellista tuntimitattua dataa sekä ennustavat pitkän aikavälin sähkönkulutusta epätarkasti.

Tässä työssä klusteroinnin tulokset olivat siinä suhteessa onnistuneet, että ryhmien keskiarvokäyrien perusteella voitiin tunnistaa selvästi neljä kuluttajaryhmää. Tosin silhouetten kuvaajan perusteella havaittiin, että pieni osa käyttöpaikoista olisi mahdollisesti kuulunut paremmin toiseen kuluttajaryhmään. Näitä tapauksia tosin tulee aina klusteroinnin yhteydessä. Epätarkkuuksia voi aiheutua myös K-means -algoritmin käytöstä. Algoritmille tulee aloitusparametrina antaa ryhmien haluttu lukumäärä, jota on usein hankala tietää etukäteen. Menetelmää joutuu tästä syystä toistamaan useita kertoja, jotta löytäisi optimaalisen ryhmien lukumäärän. Aineiston dimensioiden redusointiin käytetyn pääkomponenttianalyysin ansiosta ryhmien järkevä lukumäärä oli suhteellisen helppo arvata.

Myös klusteroinnin suorittamiseen käytetyillä eri käyttöpaikkojen lukumäärillä saadaan erilaisia tuloksia. Ongelmaksi lähtöaineiston pienen käyttöpaikkojen lukumäärän (358 suodatusten jälkeen) valinnassa osoittautui kuluttajaryhmän määrittäminen klusteroinnin ulkopuolelle jätetyille käyttöpaikoille, kuten aikaisemmin mainittiin. Tämä johtuu siitä, että pelkkään tuntimitattuun dataan perustuvassa indeksisarjojen muodostamisessa ei hyödynnetä lähtökohtana vanhoja olemassa olevia asiakastyypitietoja.

Lämpötilakorjauksella on myös oleellinen vaikutus saatujen lopputuloksien kannalta. Epätarkkuuksia voi syntyä johtuen virheistä lämpötilakorjauskertoimien määrittämisessä tai mitatun aikasarjan normalisoinnin suorittamisessa. Varsinkin huolella lasketut kuluttajaryhmäkohtaiset lämpötilariippuvuuskertoimet vähentävät epätarkkuuden suuruutta. Tässä työssä saatiin suhteellisen järkevän oloiset kuluttajaryhmäkohtaiset lämpötilariippuvuuskertoimet määritettyä.

7. Sähkönkulutusikäyrien tulevaisuuden haasteet

Nykyisille sähkönkulutusikäyrien muodostamismenetelmille on osoittautumassa haasteeksi useita tekijöitä. Näistä merkittävimpiä ovat hajautettu tuotanto, sähköautojen lataus, aurinkovoima, kysyntäjousto ja sähkön varastointi. Edellä mainittujen tekijöiden yleistyessä ongelmaksi muodostuu ainakin: Miten saadaan näiden tekijöiden aiheuttamat kulutukset erotettua asumisesta johtuvasta kulutuksesta? Miten omakotitalon tapauksessa saadaan yhtiösähkö ja asumissähkö erotettua toisistaan (näitä mittaavat eri sähkönkulutusmittarit)? Miten sähkön markkinahintaa voidaan seurata ja peilata tilanteeseen nähden? Miten kehitykseen sopeudutaan? Uusiutuvan energian jatkuva lisääntyminen ja tulevaisuuden älykäs sähköverkko (Smart Grid) tulevat muuttamaan kulutuskäyttäytymistä huomattavasti. Tämä tekee samalla nykyisistä kuormituskäyristä entistä epätarkemmat ja vanhentuneet mallintamaan sähkönkulutusta.

7.1 Hajautettu tuotanto

Hajautetulla tuotannolla tarkoitetaan pienimittakaavaista sähköntuotantoa, jossa tuotannon lähteenä toimivat uusiutuvat tai uusiutumattomat energialähteet sijaitsevat kuormakeskuksissa tai lähellä kulutusta. Yleisimpiä pienimittakaavaisia sähköntuotannon lähteitä ovat tuulivoima, aurinkokennot, biokaasu, mikroturbiinit ja biomassan yhdistetty sähkön- ja lämmöntuotanto (CHP). [9]

Hajautetun tuotannon lisääntyminen ja tuotannon siirtyminen lähemmäksi kulutusta vaikuttavat kuormituskäyriin kahdella eri tavalla. Säätoivoiman tarve tulee korostumaan muun muassa tuulivoiman tuotannon vaihtelevan ja huonon ennustettavuuden vuoksi. Vaihtoehtoisesti järjestelmän oman joustavuuden lisääminen voi toimia myös ratkaisuna säätoivoiman tarpeen tilalla. Tämän vuoksi tulevaisuuden älykkäälle sähköjärjestelmälle asetetaan suuria toiveita järjestelmän joustavuuden lisäämisessä. Hyvien ja toimivien kuormitusikäyrien tulisi huomioida joustavuuden ohjaus. Näiden avulla voitaisiin tutkia ja arvioida mahdollisia ohjauksia ja tämän jälkeen todentaa ohjausvasteita. [9]

Hajautetun tuotannon muodostama toinen merkittävä kuormituksen mallintamiseen tarvittava muutos on tuotantokäyrien tarve. Nykyiset kuormituskäyrät ovat mallintaneet ainoastaan kulutusta, mutta hajautetun tuotannon lisääntyessä korostuu selkeää tarve myös tuotantoa mallintaville käyrille. Varsinkin jakeluverkon mallintaminen voi nousta ongelmalliseksi ilman hajautetun tuotannon tarkastelua kuormituskäyrissä. Oleellista on huomata, ettei mittaustiedot välttämättä toimi edellytyksenä kaikkien kuormitusmallien muodostamiselle. Tämä johtuu siitä, että kuormituskäyrien muodostaminen voi perustua myös simulointimalleihin ja fysikaalisiin malleihin. Tuulivoiman tuotantoa voitaisiin mallintaa tällaisella menetelmällä jopa tarkemmin hyvin suuresti vaihtelevan tuotantotason takia. Ennustevirheen suuruuden välttämiseksi olisi kuitenkin viisasta käyttää keskituotantoa. [9]

7.2 Sähköautojen lataus

Sähköautojen määrä lisääntyy jatkuvasti muuttaen samalla kulutustottumuksia huomattavasti. Isoin lisääntymisen kannalta rajoittava tekijä on ollut varastointikyvyltään pienet ja kalliit akut. Akkujen hinnat tulevat kuitenkin alaspäin ja tulevaisuudessa sähköautojen osuus kaikista ajoneuvoista tulee olemaan merkittävä. Vuonna 2030 sähköautojen lukumäärä tulee olemaan karkeasti arvioiden 20–50 %:a kaikista autoista. [9]

Sähköautojen töpselin välityksellä tapahtuva lataaminen kestää arviolta noin 3–4 tuntia. Lataaminen aiheuttaa suuren piikin sähkönkulutukseen, joka on esimerkiksi 16 A:n sulakkeella arviolta 3,5 kW:a. Vaikutusta ei voida mallintaa nykyisillä SENERin kuormituskäyrillä mitenkään. Tätä ei voida korjata lisäämällä vuosienergian muutosta, vaan virhemarginaali kasvaa entisestään näin toimittaessa. Täysin uudet kuormituskäyrät tarvitaan mallintamaan edellä kuvattua kulutusta. Mallien tulisi kytkeä jokaisen omakotitalo kulutustyyppin yhteyteen rinnalle vastaava sähköautollinen kulutustyyppi. Älykkään latauksen mallintaminen vaatisi rinnalle vielä kolmannen omakotisarjan. [9] Siirrettävän energian määrä ei siis aiheuta suurta ongelmaa, vaan ongelmaksi muodostuu sähköautojen latauksesta johtuvat tehohuiput. Parhaassa

tapauksessa sähköautojen lataus tapahtuisi verkon kuormituksen ollessa pienimmillään, mutta silti uusia kuorman ohjaamismalleja tarvittaisiin. [19]

7.3 Aurinkovoima

Aurinkoenergian rajoittamattoman potentiaalin takia sekä aurinkosähköjärjestelmien jatkuvan teknisen ja taloudellisen kehittymisen ansiosta, aurinkosähkön määrä on lisääntynyt merkittävästi viime aikoina. Tulevaisuuden kannalta tarkasteltuna aurinkosähkö on merkittävimpiä uusiutuvan energian tuotantomuotoja. Globaalisti aurinkosähkö on nopeinten kasvava tapa tuottaa sähköenergiaa.

Huolimatta siitä, että pienimittakaavaisten aurinkosähköjärjestelmien lukumäärä on kasvanut Suomessa huomattavasti, verkon mitoituksen merkittävästi vaikuttaviin huipputehoihin tällä ei tule olemaan isompaa merkitystä. Syynä tähän on se, että aurinkosähkön tuotanto on suurimmillaan kesäaikana, kun taas kulutuksen huiput saavutetaan talviaikana. Kuitenkin esimerkiksi Saksassa kulutukseltaan vähäisinä ajankohtina pienimittakaavainen aurinkosähkön tuotanto on johtanut kesäaikana tilanteisiin, joissa kuluttajilta verkkoon siirrettävän sähkön teho on ollut huomattavan suuri. Suomessa vastaava ongelma ei tällä hetkellä eikä lähiaikoina ole ajankohtainen taloudellisten ja teknisten syiden vuoksi. Suomessa ei ole taloudellisesti kannattavaa myydä tuottamaansa sähköä verkkoon ja aurinkosähköjärjestelmät mitoitetaan tarkasteltavan asennuskohteen kulutuksen perusteella. Kuormituslaskennassa kuitenkin tulevaisuudessa tulee ottaa pientuotanto huomioon. Nykyiset vuosienenergiaan ja keskitehoihin perustuvat mallit tulevat aurinkosähkön pientuotannon lisääntyessä johtamaan epätarkkoihin ja virheellisiin tuloksiin. Aurinkosähkön vaikutus on siis tärkeä huomioida kuormituslaskennassa tulevaisuudessa. [19]

7.4 Kysyntäjousto

Kysyntäjoustolla tarkoitetaan yleisesti sähkönkulutuksen siirtämistä niille tunneille, joina kulutus on pienempi ja sähkön hinta matalampi. Samalla sähköntuotannon

päästöjä voidaan vähentää. Kuluttajien kuormien ohjaus vaikuttaa myös jakeluverkon kuormitukseen. [19]

Tuntimitattuun kulutukseen perustuva laskutus tulee lähiaikoina käyttöön suurimmalle osalle kuorman ohjauksen kannalta merkittävillä kuluttajilla. Tariffeja, jotka perustuvat tuntimitattuihin hintoihin on jo olemassa. Kuorman jouston tarpeellisuus lisääntyy jatkuvasti ja tätä varten tarvitaan jokin ohjausmalli arvioimaan kulutuskäyttäytymisen muutosta. Kuormitusmallien uudistuminen on edellytyksenä muun muassa spot- ja tasesäätöhintoihin perustuvassa ohjauksessa. Hintaohjaus toimisi hyvin esimerkiksi sähköautojen älykkään latauksen yhteydessä, koska se hyödyntää älykkäitä sähköverkkoja ja kulutuksen ohjausta. Luultavasti myös varaavat sähkölämmittimet tulevat siirtymään sähkömarkkinoiden nopeasti muuttuvien hintojen perusteella tapahtuviin dynaamisiin ohjauksiin aikaohjauksien sijaan. Tämä edellyttää kysyntäjoustoa. Lämpöpumputkin saattavat tulevaisuudessa siirtyä käyttämään kysyntäjoustoa. [9]

Nykyisillä kuormitusmalleilla on ollut ongelmia varaavien sähkölämmitysten ja erilaisten kaksiaikatariffien mallintamisessa. Näitä on ollut hankala mallintaa käyttämällä ainoastaan yhtä kuormitusmallia. Verkkoyhtiöissä päivatariffien vaihtuminen yötariffeiksi ei ole ollut yksiselitteistä. Nykyään myös osa varaavista sähkölämmittimistä pystyy jakamaan lämpövarauksen yön eri ajoille entistä tasaisemmin eikä tehoa tarvitse syöttää täysillä, kunnes varaaja on täyttynyt. [9] Kysynnän jouston merkitys tulee tulevaisuudessa siis lisääntymään ja kuormituskäyrissä se tulee huomioida oleellisena kulutustottumuksiin vaikuttavana tekijänä. Kysynnän jouston vaikutukset tulee huomioida kuormituksen alueellisissa ennusteissa.

7.5 Sähkön varastointi

Sähkön varastoinnin yleistyminen ja varastojen kapasiteetin kasvaminen aiheuttavat selviä haasteita kuormitusmallien muodostamiselle. Varastojen yhdistäminen kuormaan ja tuotantoon muodostaa merkittävän haasteen nykyisille kuormituskäyrille. Sähkön varastoinnin ansiosta kuluttajan olisi mahdollista siirtää sähkökäyttöään

edullisimpien tuntien kohdalle, jolloin kysyntä on pientä. Kaikkien näin toimiessa kulutushuiput ajoittuisivat nykyisten hintapohjien kohdalle. Tämän seurauksena kuormituksen mallintamismenetelmät muuttuisivat merkittävästi. Muutokset kulutushuippujen ajoittumisessa aiheuttaisivat myös hintahuippujen siirtymisen. Tällöin varastointi pitäisi ajoittaa siirtyneiden hintahuippujen perusteella uudestaan. Tämä malli johtaisi hinta- ja kulutushuippujen jatkuvaan siirtymiseen, jolloin kuormituksen mallintamisesta tulisi erittäin vaikeaa. Asiakkaiden yhteenlasketun kulutuksen perusteella määräytyvien tuntihintojen ajoittamisen seurauksena, sähkömarkkinoiden hallinnasta ja ennustettavuudesta tulisi hankalaa. [20]

Edellistä kuitenkin todennäköisempi tilanne on se, että yksittäiset kuluttajat siirtävät kulutustaan varastojen avulla halvempien tuntien eli hintakuoppien ajankohdille. Tällöin kokonaiskulutus ei muuttuisi niin merkittävästi, että se vaikuttaisi oleellisesti tuntihintoihin. Tilanteessa, jossa kulutuksen huippupiikit ja hintakuopat kohtaavat samoilla ajanhetkillä, yksittäisten asiakkaiden kuormituksen huippuja on mahdollista madaltaa hintaan perustuvalla varaston ohjauksella. Menetelmän seurauksena yksittäisen asiakkaan kulutus voisi olla tasaisempi. [20] Näin ei kuitenkaan välttämättä tapahdu. Esimerkiksi jos sähköauton lataus ja varaava sähkölämmitys keskittyvät halvoille tunneille, niin tämän seurauksena kokonaissähkökulutuksen maksimit hot kasvavat. Jos asiakkaalla on käytössä myös oma sähkövarasto, voi tämäkin tulla varatuksi samojen tuntien aikana.

Mikäli asiakkaalla on omia pienimittakaavaisia sähköntuotannon lähteitä, voidaan varastoinnin ansiosta vähentää sähköverkkoon syötettävän osuuden määrää ja mahdollisesti käyttää kaikki tuotanto itsenäisesti. Itse tuotetun sähkön oman käytön lisääntyminen mahdollistaa pienemmän verkkoon syötettävän ja verkosta otettavan tehon määrän. Näiden seurauksena kuormitus verkossa vähenee ja verkon suojauksen suunnittelu muuttuu verkkoon syötön eston seurauksena. Nykyiset kuormitusmallit eivät pysty mallintamaan edellä mainituista tekijöistä johtuvaa käytöstä. Varastoinnin rooli tulee myös olemaan merkittävässä osassa tulevaisuudessa kuormituksen ja uusien kuormitusmallien kannalta tarkasteltuna. [20]

8. Yhteenveto

Kasvavan sähkönkulutuksen ennustamiseksi ja kysynnän täyttämiseksi, sähköverkkoyhtiöiden tulee olla jatkuvasti valmiita muuttuviin olosuhteisiin. Luotettavan sähkönsaannin turvaaminen verkkoalueen asiakkaille on yksi sähköverkkoyhtiöiden tärkeimmistä tehtävistä. Verkkoon syötettävän uusiutuvan energian määrän kasvaminen ja ilmastonmuutosta ehkäisevien poliittisten sähkönkulutusta rajoittavien keinojen lisääntyminen tekevät kulutuksen ennustamisesta yhä hankalampaa. Sähköjärjestelmän pitää pystyä täyttämään yhteiskunnan sille asettamat vaatimukset. Pitkän aikavälin sähkönkulutuksen ennustaminen on siis tärkeää. Tämän vuoksi vanhoja SENERin indeksisarjoja tulee korvata uudemmilla verkkoyhtiökohtaisilla indeksisarjoilla asiakkaiden sähkönkulutuksen paremman ennustettavuuden vuoksi. Vanhat indeksisarjat aiheuttavat järjestelmien tehonjakolaskentaan ja KAH-laskentaan virheitä sekä heikentävät laskennan tuloksista verkkoyhtiöiden saamaa hyötyä.

Tämän työn sähkönkulutusta ennustavien uusien indeksisarjojen luominen tehtiin sähköverkkoyhtiö Kajave Oy:ltä saatujen tuntimittausten pohjalta. Ensin tuntimitatusta aineistosta suodatettiin sellaiset käyttöpaikat, joiden tuntimittaussarjat olivat puutteelliset ja joissa virheellisten mittausten lukumäärä oli suuri. Asiakkaiden kuluttajaryhmiin luokittelun kriteerinä käytettiin ainoastaan sähkönkulutuksen tuntimittaussarjoja ilman vanhojen asiakastyypitietojen hyödyntämistä.

Luokittelun kannalta hankalien käyttöpaikkojen suodattamisen jälkeen ennen varsinaista klusterointia, suoritettiin aineiston dimensioiden redusointi. Tämä tehtiin sen vuoksi, että klusterointi olisi helpompi suorittaa. Dimensioiden redusointiin kokeiltiin aineiston jakamista vuosijaksoihin vuodenaikojen perusteella sekä pääkomponenttianalyysia. Kokeilluista menetelmistä valittiin pääkomponenttianalyysi, koska menetelmän antamat tulokset olivat huomattavasti paremmat. Tämän jälkeen asiakkaiden kuluttajaryhmiin luokitteluksi käytettiin K-means -menetelmää. Klusteroinnin suorittaminen pelkkiin asiakkaiden tuntimittaussarjoihin perustuen on hyödyllistä, koska asiakastyypitiedot ovat suurelta osin peräisin liittymäsopimuksen tekemisen ajoilta. Asiakkaiden sähkön kulutustottumukset ja asuntojen lämmitysmuodot ovat voineet muuttua merkittävästi

alkuperäisestä. Klusteroinnin tuloksena voitiin kuluttajaryhmäkohtaisten tunnittaisten keskiarvosarjojen muodostamisen perusteella erottaa neljä selvää kuluttajaryhmää. Lämpöpumppujen erottaminen aineistosta olisi vaatinut muita menetelmiä ja se jätetään myöhemmäksi tavoitteeksi.

Kuluttajaryhmien muodostamisen jälkeen määritettiin ryhmäkohtaiset lämpötilariippuvuuskertoimet. Näitä käyttämällä poistettiin lämpötilan vaikutus tuntimittauksista mitatun aikasarjan normalisoinnilla, jotta eri vuosien mittaukset olivat keskenään vertailukelpoisia. Näin korjattuja tuntitehoja käyttämällä muodostettiin kuluttajaryhmäkohtaiset indeksisarjat. Saatujen uusien Kajave20 -sarjoiksi nimettyjen indeksisarjojen sekä verkkoyhtiö Kajavella nykyisin käytössä olevien GGK062003 -sarjojen vertailemiseksi laskettiin molemmista sarjoista tuntitehot tarkasteltaville neljälle vuoden päivälle. Laskettuja tunnittäisiä tehoja verrattiin kuluttajaryhmistä valittujen käyttöpaikkojen vuoden 2019 todellisista tuntimittauksista laskettuihin summakuormiin. Kolmanneksi vertailun kohteeksi otettiin Kajave20 -sarjoista lasketut tuntitehot, joiden ulkoinen indeksi lämpötilakorjattiin. Yksittäisten käyttöpaikkojen todellista tuntimittausten vertailudataa ei käytetty vertailussa, koska yksittäisen asiakkaan sähkönkulutus on hyvin ennalta-arvaamatonta.

Saatujen tulosten perusteella voidaan sanoa, että Kajave20 -indeksisarjoista lasketut tuntitehot mallintavat Kajavella nykyisin käytössä olevia GGK062003 -indeksisarjoja paremmin todellista tuntimitattua kulutusta valittuina neljänä vuoden päivänä. Tämä pätee luultavasti myös muille vuoden päiville, koska tarkastelun kohteeksi valittiin vuoden kylmin ja kuumin päivä, joina tulosten tarkkuus todennäköisimmin poikkeaa todellisista mittauksista. Voidaankin sanoa, että uudet verkkoyhtiökohtaisesti luodut indeksisarjat ovat potentiaaliset parantamaan tehonjakolaskennan ja KAH-laskennan tarkkuutta ja näin ollen tuottamaan verkkoyhtiöille hyötyä. Työssä saadut tulokset vaikuttavat lupaavilta, mutta tulosten hyödyntäminen vaatii vielä lisätutkimusta, jotta luodut indeksisarjat voidaan ottaa käyttöön. Klusterointi suoritettiin suhteellisen pienellä määrällä käyttöpaikkoja. Tämän vuoksi tulisi ryhmittelyn ulkopuolelle jätetyille käyttöpaikoille määrittää saaduista kuluttajaryhmistä sopivin esimerkiksi tekoälyn avulla. Tekoäly voisi asiakaskohtaisesti laskettuja ulkoisia indeksejä vertailemalla määrittää jokaiselle klusteroinnin ulkopuolelle jätetyille käyttöpaikalle

sopivimman kuluttajaryhmän. Tämän jälkeen olisi syytä vielä laskea indeksisarjoista tuntitehot näille käyttöpaikoille ja verrata todelliseen tuntimittausten vertailudataan. Jos saadut tulokset vastaisivat tässä työssä saatuja tuloksia eli Kajave20 -sarjoista lasketut tuntitehot mallintaisivat todellista tuntimittausten vertailudataa GGK062003 -sarjoja paremmin, olisi selvää, että uudet indeksisarjat kannattaa ottaa käyttöön. Sellaisille käyttöpaikoille, joita mikään muodostetuista kuluttajaryhmistä ei kuvaa hyvin, voitaisiin luoda vielä omat tyyppikäyrät. Tällä tavoin tehonjakolaskennasta ja KAH-laskennasta saatavat tulokset olisivat nykyiseen verrattuna huomattavasti tarkemmat ja todellisemmat.

Tulevaisuudessa sähkönkulutuksen ennustaminen indeksisarjojen avulla tulee vielä huomattavasti vaikeutumaan johtuen muun muassa hajautetun tuotannon lisääntymisestä. Myös sähköautojen latauksen seurauksena syntyvät kuormitushuiput, sähkönkäytön siirtäminen edullisempien tuntien ajalle sekä sähkön varastointi tekevät ennustamisesta entistä vaikeampaa.

Lähteet

- [1] Rimali, V. Etäluettavan energiamittaustiedon hyödyntäminen alueellisissa kuormitusennusteissa. Diplomityö. Aalto Yliopisto, sähkötekniikan korkeakoulu. Espoo, 2011. (11 + 103) s.
- [2] Jalonen, M. Sähkökuorman mallintaminen jakeluverkon sähköasemamittauksista. Diplomityö. Tampereen Teknillinen Korkeakoulu, sähkötekniikan osasto. Tampere, 1998. (8 + 95) s.
- [3] Lakervi, E. Sähkönjakeluverkkojen suunnittelu. Helsinki, Suomi: Otatieto Oy, 1996. 110 s. ISBN 951-672-220-2.
- [4] Meldorf, M. & Täht, T. & Lehtonen, M. & Seppälä, A. & Jalonen, M. A method for assessing the maximum load flows in distribution systems. Tallinn University of Technology, VTT Energy. Tallinn, Estonia, Espoo, Finland, 1999. 6s.
- [5] Seppälä, A. Load research and load estimation in electricity distribution. Väitöskirja. VTT Energy. Helsinki, 1996. 137 s. ISBN 951-38-4947-3. [viitattu 5.1.2020]. Saatavissa: <http://www.vtt.fi/inf/pdf/publications/1996/P289.pdf>.
- [6] Jalonen, M. & Ruska, M. & Lehtonen, M. Kuormitustutkimus 2003. Projektiraportti. Valtion Teknillinen Tutkimuskeskus. Espoo, 2003. 79 s.
- [7] Trimble. Trimble NIS Power System Analysis Theory Guide For Finland. [sähköinen materiaali]. [viitattu 18.1.2020].
- [8] Sähköenergialiitto. Sähkön käytön kuormitustutkimus 1992. Helsinki, Suomi: Sähköenergialiitto ry, 1992. 172 s. ISSN 0786-7905.
- [9] Koreneff, G. Kuormituskäyrien hyödyntäminen tulevaisuudessa. Tutkimusraportti. Valtion Teknillinen Tutkimuskeskus. Espoo, 2010. 38 s.
- [10] Koivisto, M. & Heine, P. & Mellin, I. & Lehtonen, M. Clustering of Connection Points and Load Modelling in Distribution Systems. IEEE Transactions on Power Systems. [verkkolehti] Vol. 28, no.2. 2013. S. 1255-1265. [viitattu 10.2.2020]. ISSN 1799-4934 (painettu). ISSN 1799-4942 (sähköinen). Saatavissa: <https://doi.org/10.1109/TPWRS.2012.2223240>.
- [11] Kajave. Investoinnit [verkkosivusto]. [viitattu 12.2.2020]. Saatavissa: <https://www.kajave.fi/tietoa-meista/tama-kajave/investoinnit>.
- [12] Mellin, I. Pääkomponenttianalyysi. Tilastolliset monimuuttujamenetelmät opetusmoniste. Aalto Yliopisto, systeemi analyysin laboratorio, PL 00076 Aalto. Espoo, 2004. [viitattu 17.2.2020]. Saatavissa: http://www.sal.tkk.fi/vanhat_sivut/Opinnot/Mat2.112/pdf/PCOMP10.pdf.

- [13] Sharma, S. *Applied Multivariate Techniques*. New York, New York, USA: Wiley, 1996. 512 s. ISBN 978-0-471-31064-8.
- [14] Jolliffe, I.T. *Principal Component Analysis*. New York, New York, USA: Series: Springer Series in Statistics, 2002. 487 s. ISBN 0-387-95442-2.
- [15] Kerppola, J. *Datan klusterointi: K-means [verkkoaineisto]*. [viitattu 20.2.2020]. Saatavissa: <http://datalouhimo.fi/kmeans/>.
- [16] Anton, H. & Rorres, C. *Elementary Linear Algebra*. Canada: John Wiley & Sons, 2014. 713 s. ISBN 978-1-118-43441-3.
- [17] De Amorim, R.C. & Hennig, C. Recovering the number of clusters in data sets with noise features using feature rescaling factors. *Information Sciences*. [verkkolehti] Vol. 324. 2015. S. 126-145. [viitattu 24.2.2020]. Saatavissa: <https://doi.org/10.1016/j.ins.2015.06.039>.
- [18] Jussi Niskanen. 5.3.2020. Sähköposti. Kajave Oy.
- [19] Tiuraniemi, J. *Alueellinen sähkökuormien kehitysennuste Rovaniemen verkko Oy:n jakelualueella*. Diplomityö. Lappeenrannan Teknillinen Yliopisto, sähkötekniikan koulutusohjelma. Lappeenranta, 2016. (8 + 95) s.
- [20] Koskela, J. *Sähköenergian varastoinnin hyödyntämismahdollisuudet kotitalouksien energian käytön hallinnassa Suomessa*. Diplomityö. Tampereen Teknillinen Yliopisto, sähkötekniikan diplomi-insinöörin tutkinto-ohjelma. Tampere, 2016. (10 + 75) s.