



TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

SAMULI HÄMÄLÄINEN
SOSIOEKONOMISTEN MUUTTUJIEN HYÖDYNTÄMINEN
SÄHKÖNKULUTUKSEN ENNUSTAMISESSA
Diplomityö

Tarkastaja: professori Sami Repo
Tarkastaja ja aihe hyväksytty
Tieto- ja sähkötekniikan tiedekunta-
neuvoston kokouksessa 4. joulukuuta
2013

TIIVISTELMÄ

TAMPEREEN TEKNILLINEN YLIOPISTO

Sähkötekniikan koulutusohjelma

HÄMÄLÄINEN, SAMULI: Sosioekonomististen muuttujien hyödyntäminen sähkökulutuksen ennustamisessa

Diplomityö, 93 sivua, 2 liitesivua

Toukokuu 2014

Pääaine: Sähköverkot ja -markkinat

Tarkastaja: professori Sami Repo

Avainsanat: verkkoliiketoiminta, sähkökulutus, mallintaminen, ennustaminen, regressioanalyysi, aikasarjat, sosioekonominen

Nykyaikainen yhteiskunta on entistä riippuvaisempi sähköstä. Sähkö voidaan luokitella perushyödykkeeksi, jonka saatavuuden tulisi olla jatkuvasti turvattu. Tämän takia sähköverkkoyhtiöiden toiminnalle asetetaan entistä suurempia vaatimuksia. Sähkömarkkinalain mukaan verkonhaltijan tulee suunnitella, rakentaa ja ylläpitää sähköverkkoaan siten, että se täyttää sähköverkon toiminnan laatuvaatimukset ja sähkönsiirron sekä -jakelun tekninen laatu on muutoinkin hyvä. Verkonhaltijan on lisäksi laadittava jakelu- sekä siirtoverkkoa koskevat tarkat kehittämissuunnitelmat.

Rajut muutokset Suomen talouden kehityksessä ovat heijastuneet voimakkaasti sähkökulutukseen. Vantaan Energia Sähköverkkojen ennustejärjestelmällä on ollut vaikeuksia reagoida tapahtuneisiin muutoksiin. Diplomityössä tutkitaan mahdollisuutta hyödyntää sosioekonomisia muuttujia sähkökulutuksen keskipitkän ja pitkän aikavälin ennustamisessa, mikä voisi parantaa reagoimiskykyä suhdannevaihteluihin. Muuttujien vaikutuksia sähkökulutukseen tarkastellaan lineaarisen regressiomallin avulla. Ennustettava suure on kulutetun sähköenergian määrä, ja sopivia ennustemalleja pyritään muodostamaan ekonometrisin menetelmin Suomen, Helsingin sekä Vantaan Energia Sähköverkot Oy:n vastuualueen sähkökulutusta varten.

Työssä tehdään aluksi lyhyt kirjallisuuskatsaus, jossa tarkastellaan aikaisempia aiheita käsitteleviä tutkimuksia. Lisäksi perehdytään julkaisuihin, joissa on esitelty pitkän aikavälin ennusteita Suomen sähkökulutuksesta. Tämän jälkeen esitetään työn empiirisessä osassa käytettävät menetelmät ja teoria, jotka liittyvät regressioanalyysin tekemiseen.

Työn empiirisessä osassa muodostetaan sopivat regressiomallit ja regressiodiagnostiikan avulla tarkastetaan, kuvaavatko mallit riippuvuutta tilastollisesti oikein. Mallien kykyä mallintaa ja ennustaa sähkökulutusta arvioidaan tilastollisin menetelmin. Lisäksi arvioidaan riippuvuussuhteiden mahdollisia muutoksia, joita tulisi ottaa huomioon pitkän aikavälin ennustamisessa. Vantaan Energia Sähköverkot Oy:ssä kehitetään työn tulosten perusteella ennustamisjärjestelmä, jonka toiminnallisuutta ja rakennetta esitellään lopuksi.

ABSTRACT

TAMPERE UNIVERSITY OF TECHNOLOGY

Master's Degree Programme in Electrical Engineering

HÄMÄLÄINEN, SAMULI: Forecasting Electricity Consumption Using Socio-economic Variables

Master of Science Thesis, 93 pages, 2 Appendix pages

May 2014

Major: Power systems and market

Examiner: Professor Sami Repo

Keywords: Distribution network business, electricity consumption, modelling, forecasting, regression model, socio-economic

Modern society is increasingly dependent on electricity. Electricity can be classified as a commodity whose availability should be constantly protected. As a result, higher requirements are set to the operation of the distribution network operators. Distribution network operator must design, build and maintain the electrical network in such a way that it meets the requirements for electricity quality. To do so, distribution network operators must provide precise development and operational plans of the distribution network.

Severe changes in Finland's economic development have also affected annual electricity consumption. Vantaan Energia Sähköverkot Oy's forecasting tool for electricity consumption has had difficulties to react to the changes in the operational environment. This master's thesis explores the possibility of using socio-economic variables to model electricity consumption with the intention to use models in long-term consumption forecasting. Socio-economic variables, such as gross domestic product, could improve the forecasting tool's ability to react to business cycles. To examine relations between variables and annual electricity consumption a number of multiple linear regression models are built to model annual consumption in Finland, Vantaa and Helsinki region.

A brief review of the earlier studies on the subject and existing long-term forecasts for Finland's annual consumption is made. After this the thesis presents the methods and theory of the regression analysis.

In the empirical part of the thesis, different regression models are presented. A number of statistical tests are employed to confirm the validity of the models. Usability of the presented models is analyzed and forecasts are obtained on the basis of the regression models.

In the final part of the thesis, a comparison between obtained forecasts and existing forecasts are presented. Also some factors that can cause changes concerning annual consumption forecasting are briefly presented. A simple forecasting tool will be developed to utilize presented models in long-term consumption forecasting. Its functionality and structure are presented in the end of the thesis.

ALKUSANAT

Tämä työ on tehty Vantaan Energia Sähköverkot Oy:ssä syksyn 2013 ja kevään 2014 aikana. Kehityspäällikkö Juho Paloposkea tahdon kiittää tästä mahdollisuudesta.

Työn tarkastajana on toiminut professori Sami Repo ja työn ohjaajina DI Stefan Forsström sekä DI Tuomas Åhlman. Heiltä sain runsaasti asiantuntevia neuvoja ja rakentavia kommentteja työhön liittyen.

Lämpimät kiitokset kuuluvat myös ystäväilleni, jotka ovat tehneet opiskeluvuosistani unohtumattomia. Kiitos perheelleni korvaamattomasta tuesta, jota olen saanut koko pitkän koulutaipaleeni ajan. Erityisesti tahdon kiittää Riikkaa kaikesta kannustuksesta, ymmärryksestä sekä kieliopillisista neuvoista diplomityöni aikana.

Vantaalla 15.5.2014

Samuli Hämäläinen

SISÄLLYS

1	Johdanto	1
1.1	Vantaan Energia -konserni	1
1.2	Aikaisempi tutkimus.....	3
2	Sähkönkulutuksen ennustaminen	5
2.1	Sähkönkulutuksen aikaisempi kehitys ja nykytilanne.....	6
2.2	Ennusteet Suomen sähkönkulutuksen kehittymisestä	8
2.2.1	Elinkeinoelämän keskusliiton ja Energiateollisuus ry:n ennuste	8
2.2.2	VTT:n ennuste	8
2.2.3	Työ- ja elinkeinoministeriön ennusteet	10
2.2.4	Pellervon taloustutkimuksen arvio	11
2.2.5	Ennusteiden vertailu	11
2.3	Sähkönkulutuksen ennustaminen Vantaan Energia Sähköverkot Oy:ssä	12
3	Yleinen lineaarinen regressiomalli.....	15
3.1	Mallin muoto	15
3.2	Regressiokertoimien laskeminen.....	17
3.3	Mallin hyvyys.....	18
3.4	Mallia koskevat standardioletukset	18
4	Selittävien muuttujien valinta	19
4.1	Mallinvalintatestit.....	19
4.1.1	T-testi.....	19
4.1.2	F-testi	21
4.2	Mallinvalintakriteerit.....	23
4.3	Regressiodiagnostiikka.....	25
4.3.1	Poikkeavat havainnot	26
4.3.2	Selittäjien itsenäisyys	26
4.3.3	Jäännöstermien homoskedastisuus	27
4.3.4	Jäännöstermien korreloimattomuus.....	28
4.3.5	Jäännöstermien normaalisuus.....	29
5	Aikasarjat ja ennustaminen	30
5.1	Viiveet	30
5.2	Stationaarisuus.....	31
5.3	Joustot.....	32
5.4	Ennusteen luottamusväli.....	33
6	Aineiston esittely.....	34
6.1	Sähkönkulutus	34
6.2	Bruttokansantuote.....	35
6.3	Väkiluku	37
6.4	Sähkön keskihinta.....	38
6.5	Lämpötila.....	39
6.6	Kotitalouksien käytettävissä olevat tulot.....	40

6.7	Inflaatio ja työttömyysaste	41
6.8	Rakennuskanta ja toimitilojen käyttöaste	42
6.9	Käytettävien aikasarjojen itsenäisyys	43
7	Mallien muodostaminen	45
7.1	Malli Suomen sähkönkulutukselle	45
7.2	Diagnostiset testit mallille S3	46
7.3	Malli Vantaan sähkönkulutukselle	51
7.3.1	Mallit Vantaan sektoreiden sähkönkulutuksille	57
7.4	Malli Helsingin sähkönkulutukselle	60
8	Mallien analysointi	64
8.1	Suomen sähkönkulutusmallin analysointi	64
8.2	Vantaan sähkönkulutusmallien analysointi	66
8.3	Helsingin sähkönkulutusmallin analysointi	69
9	Malleilla ennustaminen	71
9.1	Ennuste Suomen sähkönkulutukselle	71
9.2	Ennuste Vantaan sähkönkulutukselle	74
9.3	Ennuste Helsingin sähkönkulutukselle	76
10	Epävarmuustekijät	79
10.1	Sähköinen liikenne	79
10.2	Sähkön pientuotanto	80
10.3	Lämmitysenergia	81
11	Ennustetyökalu	84
12	Yhteenveto	86
	Lähteet	89
	Liite 1: Korrelaatiomatriisit	94

TERMIT JA NIIDEN MÄÄRITELMÄT

ADF	Augmented Dickey-Fuller
AEG	Augmented Engle-Granger
AIC	Akaike information criterion
AR	Autoregressio
ARIMA	Autoregressive integrated moving average
AVIA	Vantaan Aviaenergia Oy
BKT	Bruttokansantuote
CHP	Combined Heat and Power
EK	Elinkeinoelämän keskusliitto
EKT 95	Euroopan kansantalouden tilinpitojärjestelmä
ET	Energiateollisuus ry
ETLA	Elinkeinoelämän tutkimuslaitos
EV	Electric vehicle, sähköajoneuvo
LM	Lagrange Multiplier
p-arvo	Ilmoittaa nollahypoteesin todennäköisyyden
PHEV	Plug-in hybrid electric vehicle, hybridiajoneuvo
PNS	Pienin neliösumma
Pseudo out-of-sample -analyysi	Ennustaminen otoksen ulkopuolelle
PTT	Pellervon Taloustutkimus
R^2	Selitysaste
SBIC	Schwarzin bayeslainen informatiokriteeri
SSE	Jäännöseliösumma
SSM	Mallineliösumma
SST	Kokonaisneliösumma
TEM	Työ- ja elinkeinoministeriö
VES	Vantaan Energia Sähköverkot Oy
VIF	Varianssin inflaatiotekijä
VTT	Teknologian tutkimuskeskus
α	Kriittinen arvo
β	Regressiokerroin
ε	Mallin jäännöstermi
δ	Sopeutumiskerroin
a_i	Jäännöstermien keskiarvon varianssin ja kovarianssin avulla muodostettu vakio
b	Regressiokertoimen estimaattori
d	Durbin-Watson -testin testisuure
d_L	Durbin-Watson -testin testisuureen kriittinen alaraja

d_U	Durbin-Watson -testin testisuureen kriittinen yläraja
E	Jousto
H	Hypoteesi
k	Velanderin kerroin
R^2	Mallin selitysaste
\bar{R}^2	Mallin korjattu selitysaste
S	Keskivirhe
W	Shapiro-Wilk -testin testisuure
X	Selittävä muuttuja
\tilde{x}	Muuttujan X ennustettu arvo
y	Selitettävä muuttuja

1 JOHDANTO

Nykyaikainen yhteiskunta on entistä riippuvaisempi sähköstä. Sähkö voidaan luokitella perushyödykkeeksi, jonka saatavuuden tulisi olla jatkuvasti turvattu. Tämän takia sähköverkkoyhtiöiden toiminnalle asetetaan entistä suurempia vaatimuksia. Sähkön kysynnän trendi on ollut tasaisen nouseva viime vuosikymmenten ajan, mutta tulevaisuudessa kasvu saattaa hidastua, tai jopa kääntyä laskuun esimerkiksi energiatehokkuuden parantumisen johdosta.

Verkonhaltijan tulee suunnitella, rakentaa ja ylläpitää sähköverkkooaan siten, että se täyttää sähköverkon toiminnan laatuvaatimukset ja sähkönsiirron sekä -jakelun tekninen laatu on muutoinkin hyvä. Verkonhaltijan on lisäksi laadittava jakelu- sekä siirtoverkkoa koskevat tarkat kehittämissuunnitelmat.[1]

Jotta verkkoyhtiö voisi laatia riittävän tarkat suunnitelmat verkon kehittämiseksi, on sillä oltava käsitys sähkön kysynnän kehittymisestä pitkälle, jopa kymmeniä vuosia eteenpäin. Sähkön kysyntä on altis monenlaisille yhteiskunnassa tapahtuville muutoksille, minkä takia verkkoyhtiöiden on jatkuvasti tarkkailtava toimintaympäristöään ja arvioitava muutosten vaikutusta sähkön kysyntään.

Tässä työssä tutkitaan mahdollisuutta hyödyntää sosioekonomisia muuttujia sähkönkulutuksen keskipitkän ja pitkän aikavälin ennustamisessa. Ennustettava suure on kulutetun sähköenergian määrä, ja sopivia ennustemalleja pyritään muodostamaan ekonometrisin menetelmin sekä Suomen että Vantaan Energia Sähköverkot Oy:n vastualueen sähkönkulutusta varten.

Työssä tehdään aluksi lyhyt kirjallisuuskatsaus, jossa tarkastellaan aikaisempia aiheita käsitteleviä tutkimuksia ja julkaisuja, joissa on esitelty pitkän aikavälin ennusteita Suomen sähkönkulutuksesta. Tämän jälkeen esitetään työn empiirisessä osassa käytettävät menetelmät ja teoria.

Työn empiirisessä osassa muodostetaan sopivat regressiomallit ja regressiodiagnostiikan avulla tarkastetaan, kuvaavatko mallit riippuvuutta tilastollisesti oikein. Mallien kykyä mallintaa ja ennustaa sähkönkulutusta arvioidaan tilastollisin menetelmin. Lisäksi arvioidaan riippuvuussuhteiden mahdollisia muutoksia, joita tulisi ottaa huomioon pitkän aikavälin ennustamisessa. Vantaan Energia Sähköverkot Oy:ssä kehitetään työn tulosten perusteella ennustamisjärjestelmä, jonka toiminnallisuutta ja rakennetta esitellään lopuksi.

1.1 Vantaan Energia -konserni

Vantaan Energia Oy:n toiminnan voidaan katsoa alkaneen vuonna 1910, jolloin Malmin Sähkölaitos Oy aloitti toimintansa. Nykyisellä nimellään Vantaan Energia -konserni on

tunnettu vuodesta 1996 lähtien. Emoyhtiö tuottaa sähköä, lämpöä ja palveluita ollen yksi alansa suurimmista toimijoista Suomessa. Vantaan Energia Oy:n omistajia ovat Vantaan kaupunki (60 %), sekä Helsingin kaupunki (40 %).

Emoyhtiön lisäksi konserniin kuuluu Vantaan Energia Sähköverkot Oy (VES), joka on täysin emoyhtiön omistuksessa. Vuoden 2014 alussa aiemmin omana tytäryhtiönä toiminut Vantaan Aviaenergia Oy:n sähköverkkoliiketoiminta liitettiin osaksi Vantaan Energia Sähköverkot Oy:n toimintaa ja kaukolämpöliiketoiminta osaksi emoyhtiön toimintaa. Vantaan Aviaenergia Oy oli aiemmin vastannut lentokenttäalueen sähkönsiirrosta ja kaukolämmön toimituksesta. Muutoksen johdosta VES vastaa sähkönsiirrosta ja -jakelusta nyt koko Vantaan alueella.

Emoyhtiöllä on omistusosuuksia useissa eri energia-alan yhtiöissä. Näitä yhtiöitä ovat tuotanto-osakkuusyhtiöt, kuten Norjassa sijaitseva Svartisenin vesivoimalaitos ja tuulivoimaan erikoistuneet Suomen Hyötytuuli Oy, Innopower Oy sekä EPV Energia Oy. Lisäksi yhtiö omistaa osan Energian Säästöpalvelu Enespa Oy:stä ja Suomen Energia-Urakointi Oy:stä, joka on sähköisen yhdyskuntatekniikan verkkojen ja laitteistojen suunnitteluun, asennukseen ja käyttöön liittyviä palveluita tarjoava yritys.

Vuonna 2012 Vantaan Energia -konsernin liikevaihto oli 434 miljoonaa euroa, josta Vantaan Energia Sähköverkot Oy:n osuus oli 38,5 miljoonaa euroa ja Vantaan Aviaenergia Oy:n osuus 10,3 miljoonaa euroa. Konsernin henkilöstömäärä oli vuonna 2012 352 henkilöä, josta VES:ssä työskenteli 68 ja Vantaan Aviaenergia Oy:ssä seitsemän henkilöä. Taulukossa 1.1 on esitelty muita VES:n sähköverkkoliiketoimintaan liittyviä tunnuslukuja vuodelta 2012.

Taulukko 1.1 Vantaan Energia Sähköverkot Oy:n sähköverkkotoiminnan tunnuslukuja vuodelta 2012

0,4 kV verkon pituus (kaapelointiaste)	20 kV verkon pituus (kaapelointiaste)	110 kV verkon pituus	Sähköasemat	Jakelumuuntajat
2240 km (83 %)	875 km (85 %)	80 km	11 kpl	1061 kpl
Bruttoinvestoinnit	Asiakasmäärä	Keskimääräinen keskeytysaika /asiakas	Keskimääräinen asiakkaan kokema keskeytysmäärä	Siirretty sähkö
22,7 M€	n. 108 000	5,5 min/a	0,21 kpl/a	1752 GWh

Vuonna 2012 Vantaan alueella siirrettiin sähköä 1752 GWh ja lentokenttäalueella 138 GWh. Vantaan kaupungin väkiluvun sekä työpaikkojen lukumäärän oletetaan kasvavan tasaisesti tulevina vuosikymmeninä, minkä seurauksena sähkön kysyntä alueella tulee kasvamaan. Sähkön kysyntä tulee kasvamaan myös lentokenttäalueella, joka tulee kehittämään voimakkaasti vuosina 2014 – 2020.

Vantaan Energian suurin viimeaikainen investointi on ollut uuden jätevoimalan rakentaminen. Uusi voimala tulee korvaamaan Martinlaakson voimalaitoksen yhden tuo-

tantoyksikön. Korvattava CHP-tuotantoyksikkö käyttää polttoaineenaan maakaasua, ja se tulee viimeistään vuonna 2015 käyttöikänsä loppuun. Jätevoimalassa tullaan tuottamaan sekä sähköä että lämpöä, ja sen on määrä aloittaa tuotantonsa keväällä 2014. Lämpöä voimala tulee tuottamaan noin 920 GWh ja sähköä 600 GWh vuodessa. Uuden voimalan tullessa käyttöön Vantaan Energian voimalaitosten fossiilisten polttoaineiden käyttö sähkön ja lämmön tuotannossa vähenee noin 30 % ja yhtiön kokonaispäästöt vähenevät noin 20 % nykyisestä.

Sähköverkkoa on myös vahvistettu viime vuosina mittavilla investoinneilla. Vuonna 2012 otettiin käyttöön uusi Honkanummen 110/20 kV:n sähköasema ja vuonna 2014 on valmistumassa toinen uusi 110/20 kV sähköasema Keski-Vantaalle Tuupakkaan.

1.2 Aikaisempi tutkimus

Sosioekonomisten muuttujien hyödyntämistä sähkönkulutuksen ennustamisessa on tutkittu maailmalla aiemmin. Aiheeseen liittyvää tutkimusta on tehty muun muassa voimakkaasti kehittyvissä maissa, joissa voimakkaan talouskasvun myötä myös sähkön kysyntä on kasvanut rajusti. Tämän tyyppisessä tilanteessa energiayhtiöiden kannalta pahin skenaario olisi se, että sähkön tuotanto- ja siirtokapasiteetti ylittyisi. Tämä voisi johtaa tilanteeseen, jossa sähkön heikko saatavuus rajoittaisi taloudellista kasvua.

Bianco et al. on vuonna 2008 tutkinut lineaarisen regressiomallin sopivuutta Italian sähkönkulutuksen mallintamiseen ja ennustanut sähkönkulutuksen kehittymistä eri regressiomallien avulla. Artikkelissa on pyritty havaitsemaan sosioekonomisten muuttujien vaikutusta sähkönkulutukseen. Tätä tutkittiin lineaaristen regressiomallien avulla, joita muodostettiin mallintamaan koko Italian sähkönkulutusta sekä pelkästään kotitalouksien ja muun yhteiskunnan sähkönkulutusta. Muuttujien joustoja tutkittiin log-lineaarisilla malleilla. Artikkelissa havaittiin, että bruttokansantuotteella, väkiluvulla ja BKT per capitalla on voimakas yhteys sähkönkulutuksen kehitykseen niin kotitalouksissa kuin myös muualla yhteiskunnassa ja niiden avulla voitiin mallintaa sähkönkulutusta hyvin. Sähkön hinnalla puolestaan ei ollut selittävää vaikutusta sähkönkulutukseen. Lisäksi mallien avulla ennustettiin sähkönkulutusta vuoteen 2030 asti. Raportissa todetaan selittävien muuttujien ennustettujen arvojen tarkkuudella olevan merkittävä vaikutus ennusteen luotettavuuteen. [2]

Mohamed ja Bodger ovat tutkineet vuonna 2004 ekonomisten ja demografisten muuttujien soveltuvuutta Uuden-Seelannin sähkönkulutuksen ennustamiseen. He ehdottavat lineaarista regressiomallia, jossa selittävinä muuttujina ovat bruttokansantuote, sähkön hinta sekä väkiluku. Näiden avulla pyrittiin kuvaamaan kotitalouksien, muun yhteiskunnan ja koko maan sähkönkulutusta. Malli selitti kotitalouksien sähkönkulutusta 89 % tarkkuudella, mitä voidaan pitää kohtalaisen hyvänä. Muuta kuin kotitalouksien sekä koko maan sähkönkulutusta malli selitti selvästi paremmin 96 % tarkkuudella. Mallilla myös ennustettiin sähkönkulutusta vuoteen 2015 asti ja ennustetta verrattiin muihin julkaistuihin ennusteisiin. Myös tässä artikkelissa painotettiin selittävien muuttujien ennusteiden tarkkuutta mallin ennustustarkkuuden parantamiseksi. [3]

Yoo ja Lee (2008) tutkivat talouden kasvun ja sähkönkulutuksen syy-seuraussuhdetta. He tutkivat asiaa seuraamalla BKT per capitaan sekä sähkönkulutuksen käyttäytymistä 88 maassa, ja tulivat siihen johtopäätökseen, että kotitalouksien sähkönkulutus kasvaa kotitalouksien tulojen kasvun mukana. Kehittyvissä maissa kasvu on runsaampaa kuin kehittyneissä maissa. [4]

Vuonna 2013 on Helsingin yliopistossa tehty kandidaatintutkielma yhteistyössä Vantaan Energia Sähköverkot Oy:n kanssa, jossa on tutkittu sosioekonomisten muuttujien käyttöä sähkönkulutuksen mallintamisessa ja ennustamisessa. Tutkielmassa muodostettiin malli, jossa sähkönkulutusta VES:n alueella mallinnettiin bruttokansantuotteen, väestön, rakentamisen ja lämpötilan avulla. Mallin avulla onnistuttiin kuvaamaan sähkönkulutusta erittäin hyvin, mutta esiin nousi myös monia aiheita, jotka antavat aiheita jatkotutkimuksen tekoon. Näitä kysymyksiä olivat esimerkiksi mahdollisuus muodostaa malli eri kulutusryhmien sähkönkulutukselle, mallin muodostaminen vuosidatan sijasta vuosineljännes- tai kuukausidatan avulla, ennustamiseen liittyvien virheiden ja luottamusvälien tarkastelu, ARIMA-mallien soveltaminen sekä energiatehokkuuden ja rakennemuutoksen tuomat muutokset sähkön kulutuksen mallintamiseen. [5] Tässä työssä pyritään selventämään osaa näistä aiheista.

2 SÄHKÖNKULUTUKSEN ENNUSTAMINEN

Sähkökulutuksen vuosienergian ennustaminen on oleellinen osa verkkoyhtiöiden suunnittelutoimintaa. Ennusteita tehdään eripituisille ajanjaksoille, ja niiden avulla on tarkoitus selvittää mahdollisimman tarkasti missä, kuinka paljon ja milloin sähköä tulevaisuudessa kulutetaan.

Lyhyen aikavälin ennustamisella tarkoitetaan kulutuksen ennustamista vain tunteja tai päiviä eteenpäin, jotta sähköverkkoyhtiöt kykenisivät ennakoimaan sähköverkon käyttötoimenpiteitä, ja sähkön tuottajat sekä myyjät pystyisivät optimoimaan toimintaansa. Keskipitkän aikavälin ennuste ulottuu joidenkin kuukausien, enimmillään muutamana vuoden päähän. Sen avulla sähköverkkoyhtiöt pyrkivät muun muassa ennustamaan kulutushuippujen ajankohtaa ja suuruutta.

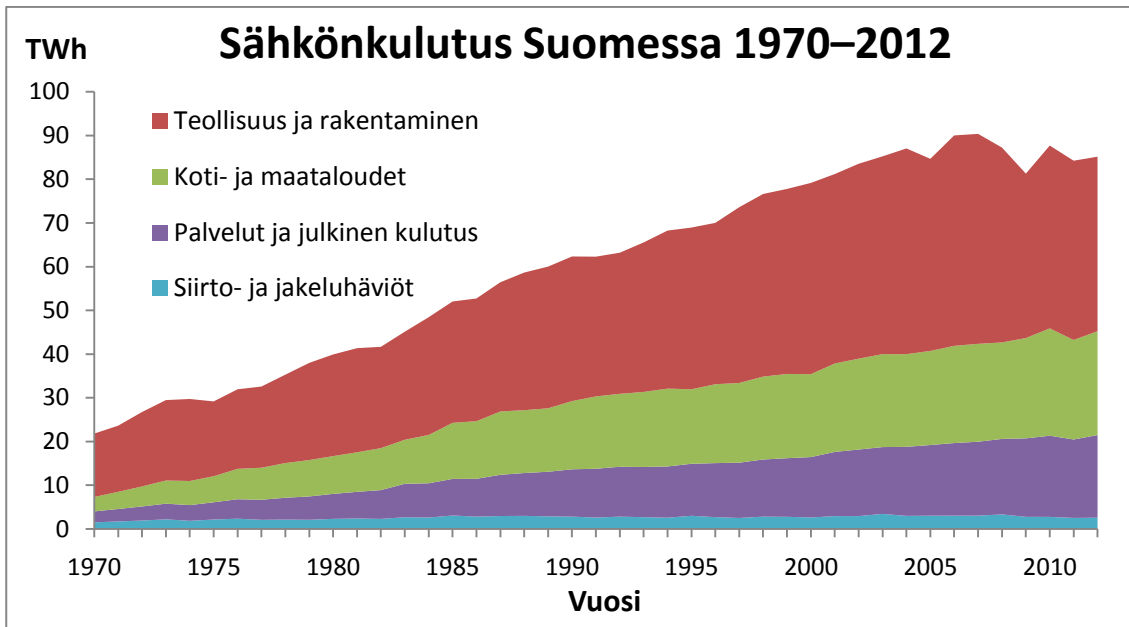
Suurimpien verkkoinvestointien suunnittelu ja toteutus voi pisimmillään kestää yli 20 vuotta. Verkko- ja energiayhtiöiden on pystyttävä ennakoimaan sähkökulutuksen tarve vähintään tuon verran etukäteen, jotta tarvittavat investoinnit ehditään suunnitella ja toteuttaa ajoissa. Pitkän aikavälin ennuste toimii apuna verkon yleissuunnittelussa, jossa pyritään muodostaa kuva sähkön kysynnästä vuosikymmenien päähän ja kartoittaa tulevat siirto- ja keskijänniteverkko-, sähköasema- sekä muut pitkää toteutusaikaa vaativat investoinnit.

Jotta pitkän aikavälin ennuste olisi riittävän tarkka, on ennusteessa otettava huomioon kaikki lähtötiedot ja muuttujat, joiden oletetaan vaikuttavan kulutukseen. Suuri muuttujien määrä voi toisaalta myös lisätä epävarmuutta, sillä useiden muuttujien kehitystä on hyvin vaikea ennustaa. Tämän takia pitkän ajan ennustuksissa on tyypillistä tuoda esille useampia, eri lopputuloksiin johtavia skenaarioita ja tarkastella ennustusten luottamusvälejä.

Kysynnän kehittymistä on hyvä tarkastella myös kulutussektoreittain. Tyypillisesti tarkastellaan erikseen kotitalouksien, teollisuuden ja rakentamisen, liikenteen sekä palvelu- ja julkisalan sähkökulutusta. Tarvittaessa voidaan teollisuus jakaa vielä pienempiin sektoreihin, jolloin ainakin energiaintensiiviset teollisuudenalat, kuten metsä- ja metalliteollisuus tulisivat käsitellyiksi erikseen. Siirto- ja jakeluverkkohäviöiden osuutta sähkökulutuksesta on myös mahdollisuus tarkastella erikseen, jolloin sähkön loppukäytön suuruus tarkentuu. Tarkasteltava alue on syytä pyrkiä jakamaan mahdollisimman pieniin alueisiin, joille on mahdollista luoda oma ennuste. Alueellisen sähkökulutuksen ennustaminen luo pohjan sähköverkkojen yleissuunnittelulle.

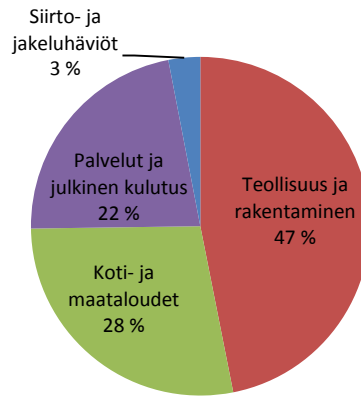
2.1 Sähkönkulutuksen aikaisempi kehitys ja nykytilanne

Sähkönkulutuksen ennustamisen lähtökohta on kulutuksen aikaisempi kehitys ja sen nykyinen tilanne. Sähkönkulutus on kasvanut Suomessa tasaisesti aina 1970-luvulta lähtien. Kuvassa 2.1 on esitetty Suomen sähkönkulutus sektoreittain aikavälillä 1970–2012.



Kuva 2.1. Sähkönkulutuksen kehitys Suomessa vuosina 1970 – 2012 [6]

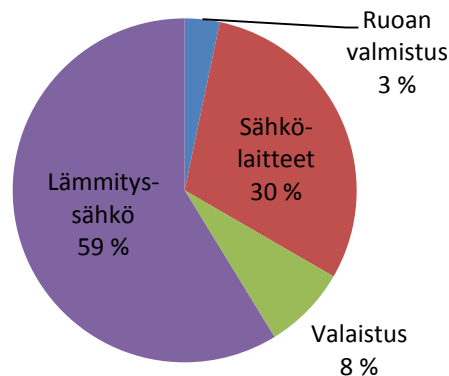
Hieman yli 40 vuodessa suomen sähkönkulutus on yli nelinkertaistunut ollen vuonna 2012 noin 85 TWh. Korkeimmillaan Suomen sähkönkulutus on ollut vuonna 2007, jolloin sähköä käytettiin yhteensä 90,3 TWh. Sähkön kokonaiskulutuksen kasvua on pitänyt yllä maa- ja kotitalouksien, palveluiden sekä julkisen sähkönkulutuksen kohtalaisen tasainen kasvu, jonka voidaan olettaa olevan seurausta väestönkasvusta sekä yhteiskunnan sähköistymisestä. Kuvan 2.2 ympyrädiagrammissa on esitetty eri sektoreiden osuudet Suomen sähkönkulutuksesta vuonna 2012.



Kuva 2.2 Suomen sähkönkulutuksen jakautuminen sektoreittain vuonna 2012 [6]

Vuonna 2012 teollisuus kulutti melkein puolet Suomen sähköstä. Erityisesti metsä- ja metalliteollisuus ovat suuria sähköenergian kuluttajia, jotka yhdessä muodostavat peräti kolmasosan koko maan sähkönkulutuksesta. Näin ollen näissä sektoreissa tapahtuvat muutokset vaikuttavat olennaisesti koko maan sähkönkulutukseen. Esimerkiksi kuvasta 2.1 havaittava vuonna 2005 tapahtunut pudotus selittyy yksiselitteisesti kyseisenä vuonna olleesta metsäteollisuuden pitkäkestoisesta lakosta. Kuvasta 2.1 voidaan nähdä myös vuonna 2008 alkanut jyrkkä pudotus sähkönkulutuksessa, mikä osittain johtuu sinä vuonna alkaneen talouskriisin vaikutuksista Suomen teollisuuden tuotantoon. Toisena merkittävänä syynä vuosien 2008–2009 sähkönkulutuksen laskuun voidaan pitää Suomen teollisuuden rakennemuutosta, josta etenkin paperiteollisuus on kärsinyt. Tästä syystä viimeisimmän laman aikana sähkönkulutuksen lasku on ollut huomattavasti voimakkaampaa kuin 1990-luvun alun laman aikoihin.

Vuonna 2011 on tehty tutkimus, jossa kartoitettiin suomalaisten kotitalouksien sähkönkäyttöä. Kuvassa 2.3 on esitetty, mistä kotitalouksien sähkönkulutus tutkimuksen mukaan koostuu.



Kuva 2.3. Kotitalouksien sähkönkäyttö Suomessa vuonna 2011 [7]

Tutkimushankkeen tuloksista ilmenee, että noin 60 % kotitalouksien sähkönkulutuksesta menee lämmitykseen. Valaistuksen osuus on 8 %, mikä on laskenut kuusi prosent-

tiyksikköä vuodesta 2006 lähinnä uuden valaistustekniikan ansiosta. Sähkölaitteiden muodostama kulutus on hieman kasvanut esimerkiksi tietotekniikan runsaan lisääntymisen myötä. Muiden sähkölaitteiden, kuten televisioiden ja kylmälaitteiden sähkönkulutus on hieman laskenut energiatehokkuuden ansiosta. [7]

2.2 Ennusteet Suomen sähkönkulutuksen kehittymisestä

Monet eri tahot luovat pitkän aikavälin ennusteita Suomen sähkönkulutuksesta. Ennusteita tehdään tilaustöinä päätöksenteon tueksi niin poliitikoille kuin myös energiateollisuuden organisaatioille. Energia-alan investoinnit perustuvat vahvasti energian kysynnän ja hinnan pitkän aikavälin ennusteisiin.

Kaikissa ennusteissa on käytetty perustana ennusteita Suomen talouden ja etenkin teollisuuden kasvusta. Lisäksi on pyritty huomioimaan esimerkiksi tekninen kehitys ja ihmisten käyttäytymisen muutokset.

2.2.1 Elinkeinoelämän keskusliiton ja Energiateollisuus ry:n ennuste

Vuosina 2009 ja 2010 Elinkeinoelämän keskusliito (EK) ja Energiateollisuus ry (ET) ovat julkaisseet raportit, joiden taustalla on ollut laaja pohjatyo koskien eri sektoreiden tulevaisuudennäkymiä ja energiatehokkuuden tehostamispotentiaaleja vuoteen 2050 asti. Raporttien esittelemään visioon on vaikuttanut myös Suomen kansantalouden kehitysnäkymät sekä Turun Kauppakorkeakoulun Tulevaisuuden tutkimuskeskuksen suunnittelemat neljä erilaista energiaskenaariota. Visiossa keskitytään siihen, miten sähkö voi osaltaan olla ratkaisemassa energia- ja ilmastohaasteita. [8,9]

Visiossa on oletettu Suomen teollisuuden toipuvan taantumasta, ja talouskasvu on vuoteen 2050 asti keskimäärin 2 % vuodessa. Metsäteollisuuden rakennemuutos tulee aikanaan johtamaan uusiin tuotteisiin ja prosesseihin. Palvelualan osuus niin bruttokansantuotteen kuin sähkönkulutuksen osalta tulee kasvamaan. Väestö kasvaa yli 6 miljoonaa vuoteen 2050 mennessä ja asuntokuntien keskikoko pienenee, jolloin kotitalouksien määrä kasvaa reilusti. Kotitalouksissa oletetaan käytettävän yhä enemmän sähkölämmitystä sekä energiatehokkaita sähkölaitteita. [8,9]

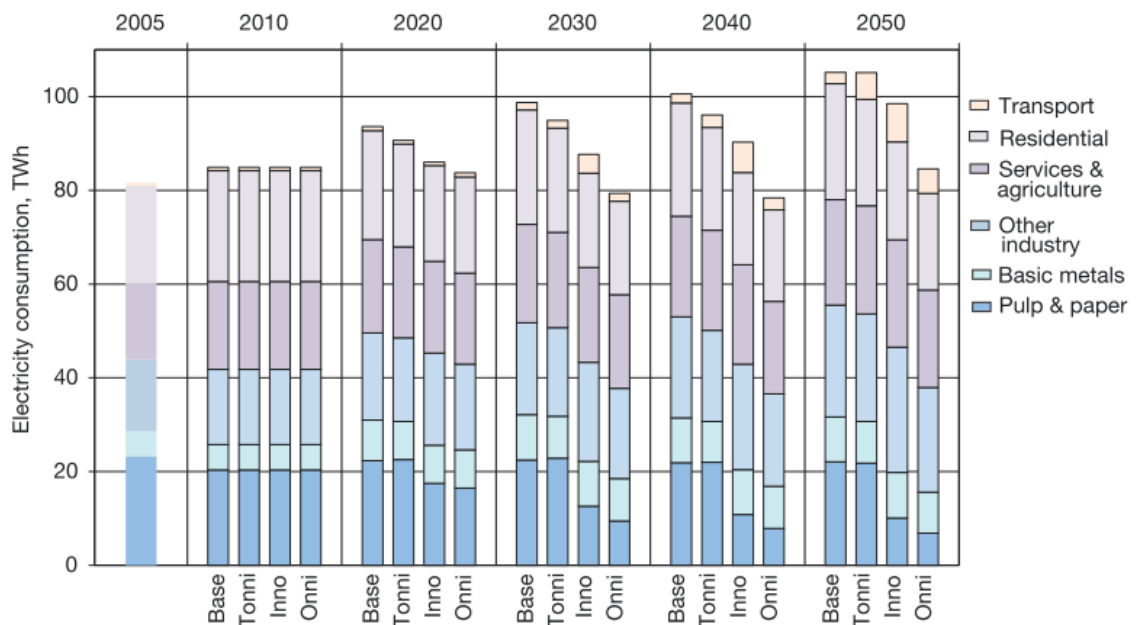
Muun muassa näiden oletusten ja tausta-aineiston perusteella arvioidaan Suomen sähkönkulutuksen olevan vuonna 2030 välillä 100–111 TWh ja vuonna 2050 113–138 TWh. Arvio pohjautuu jo tunnettuun teknologiaan ja sellaisiin ratkaisuihin, jotka ovat suurella todennäköisyydellä käytettävissä tarkastelujakson aikana. [8,9]

2.2.2 VTT:n ennuste

Teknologian tutkimuskeskus VTT:n vuonna 2013 ilmestyneessä julkaisussa on tarkasteltu Suomen energiatulevaisuutta hiilidioksidi- ja muiden kasvihuonepäästöjen vähentämistavoitteiden näkökulmasta kolmen eri skenaarion; Inno-, Tonni- ja Onni-skenaarion sekä perusuran avulla. Kaikissa skenaarioissa on ollut tavoitteena 80 %:n

vähennys päästöissä. Perusura toimii puolestaan vertailukohtana kuvaten tulevaisuutta nykyisten päätösten ja trendien mukaisesti. [10]

Apuna tulevaisuuskuvioiden muodostamisessa on käytetty myös Valtion taloudellisen tutkimuskeskuksen arviota Suomen talouden ja hyvinvoinnin kehittymisestä vuoteen 2050 asti. Arvion mukaan Suomen bruttokansantuote kasvaisi perusurassa vuoden 2004 tasoon (152 Mrd€) verrattuna noin 2,4 -kertaiseksi ollen vuonna 2050 noin 540 Mrd€ vuoden 2004 hintatasolla. Teknologinen kehitys sekä erityisesti palvelualan kasvu siivittäisivät talouden kasvua. [10] Julkaisun ennusteet Suomen sähkönkulutuksesta on esitetty kuvassa 2.4.



Kuva 2.4. VTT:n skenaariot sähkön loppukulutuksesta Suomessa [10]

Tonni-skenaario ei poikkea paljoa perusurasta. Teollisuuden tuotanto ja rakenne jatkuvat trendien mukaisina. Teknologinen kehitys on muihin skenaarioihin nähden maltillista, eikä muuta kuin jo tunnettua teknologiaa huomioida. Ydinvoima toimii pohjana sähköntuotannolle, jota monipuolistetaan uusiutuvilla energianlähteillä. Sähkön hinta pyritään pitämään kilpailukykyisenä, vaikka kysyntä kasvaa. [10]

Inno-skenaariossa eri osa-alueilla tapahtuva teknologinen kehitys on voimakkainta. Energiatehokkuus kasvaa uusien innovaatioiden myötä. Näin tapahtuu myös energia-alalla, jossa energiaa tuotetaan uusien teknologioiden avulla niin suuressa kuin pienessä mittakaavassa. Suurinta kehitys on bio- aurinko- ja tuulivoimateknologioissa. [10]

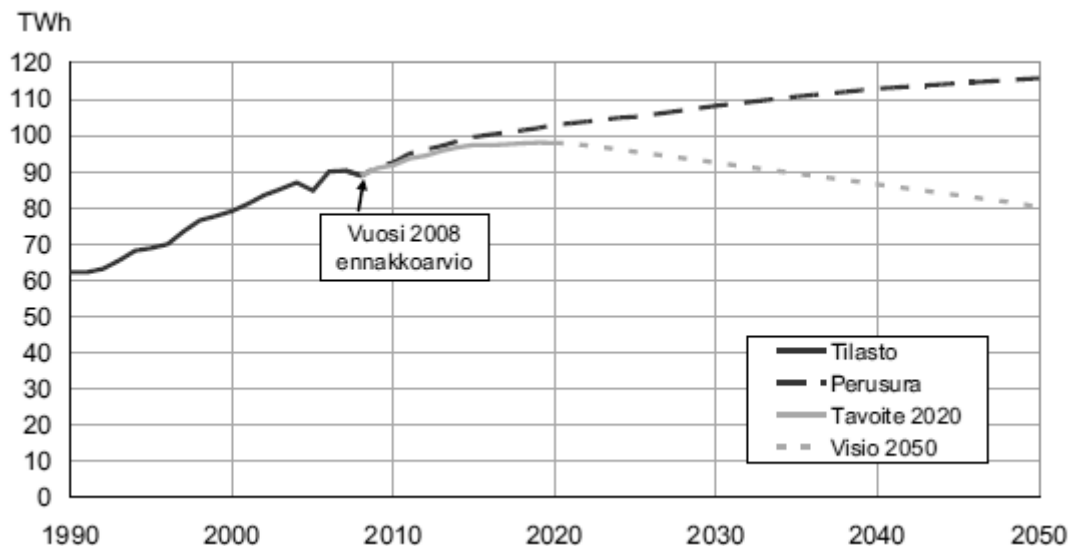
Onni-skenaariossa teollisuuden rakenteen oletetaan muuttuvan rajusti. Perinteinen, energiaintensiivinen teollisuus vähenee ja tilalle tulee runsaasti pienimuotoista, paikallista palveluliiketoimintaa. Sähkö tuotetaan lähellä kulutuspiستettä uusiutuvalla energialla. [10]

2.2.3 Työ- ja elinkeinoministeriön ennusteet

Työ- ja elinkeinoministeriö (TEM) on julkaissut vuonna 2008 Pitkän aikavälin ilmasto- ja energiastrategia -selonteon, jossa ennustetaan sähköenergian käytön kasvun hidastuvan luontaisesti teknologian uudistumisen, energian hinnan nousun sekä päästökaupan vaikutuksesta. Strategiassa esitellään kaksi skenaariota: perusura sekä tavoiteura. Skenaarioissa on otettu huomioon kansantalouden, väestön, energian maailmankauppahintojen ja Suomen energiainfrastruktuurin kehitys sekä politiikkatoimet. Perusoletuksina skenaarioissa on muun muassa, että Suomen väestö kasvaisi keskimäärin 0,2 % vuodessa, mikä perustuu Tilastokeskuksen silloiseen väestöennusteeseen. Kyseisellä kasvuvauhdilla Suomen väestö olisi vuonna 2050 noin 5,7 miljoonaa. Kansantalouden on oletettu kasvavan vuoteen 2020 asti runsaalla kahdella prosentilla vuodessa, minkä jälkeen kasvu hidastuisi vajaan kahteen prosenttiin. [11]

Perusura-skenaario esittelee nykytoimien ja -kehityksen mukaisen tulevaisuuskuvan. Perusurassa sähkönkulutus on vuonna 2020 103 TWh ja vuonna 2030 108 TWh.

Tavoiteura-skenaario on johdettu perusurasta huomioiden EU:n sekä kansalliset tavoitteet. Tavoiteurassa sähkönkulutus on vuonna 2020 98 TWh. Strategia esittää myös vision sähkönkulutuksesta vuodesta 2020 vuoteen 2050, minkä mukaan sähkönkulutus kääntyisi Suomessa laskuun 2020-luvulla. Vision mukaiseen tilanteeseen voitaisiin päästä, mikäli sähkönkulutukseen liittyvät tehostamistavoitteet toteutuvat. [11] Kuva 2.5 esittää vuonna 2008 julkistetun strategiassa esitellyn pitkän aikavälin ennusteen sekä vision vuoteen 2050 asti.



Kuva 2.5. Sähköenergian kysyntä 1990–2007, kysynnän ennusteet 2008–2020 sekä visio vuoteen 2050 asti [11]

Kuvan 2.5 kuvaajissa ei ole huomioitu vuoden 2008 jälkeisiä talouden muutoksia, jotka ovat aiheuttaneet selvän muutoksen sähkön kysynnässä. Vuonna 2012 on julkaistu raportti ”Sähkömarkkinaskenaariot vuoteen 2035”, jossa on käytetty ajankohtaisempaa

kuvaa talouden nykytilanteesta ja kehityksestä. Raportin mukaan sähkönkulutus olisi vuonna 2020 95,8 TWh ja vuonna 2030 103,3 TWh [12].

Vuonna 2013 on julkaistu Kansallisen energia- ja ilmastostrategian päivitetty versio, jossa todetaan talouden kehityskuvan muuttuneen vuodesta 2008 siten, että vuoden 2020 sähkönkulutuksen ennakoitaan jäävän perusuralla 94 TWh:iin [13]. Muutos vuoden 2008 arvioon on peräti 9 TWh, minkä syyksi voidaan arvioida muun muassa energiaintensiivisen teollisuuden tuotannon epäsuotuisa kehitys. Tämän tarkempaa analyysia taloustilanteen muutoksen vaikutuksista sähkönkulutuksen pitkän aikavälin ennustukseen ei raportissa kuitenkaan anneta.

2.2.4 Pellervon taloustutkimuksen arvio

Keväällä 2013 käynnistyi Suomen energia- ja ilmastotiekartan laatiminen vuoteen 2050 saakka. Aloituseminaarissa Pellervon taloustutkimus (PTT) esitti kansantalouden näkökulmia aiheeseen.

Esityksessä on todettu, että talouskasvu vaikuttaa oleellisesti energian ja sähkön kulutukseen. PTT olettaa talouskasvun olevan vuoteen 2035 asti keskimäärin 1,7 % vuodessa, jolloin Suomen sähkönkulutus kasvaisi vuoden 2010 tasosta 16 % vuoteen 2030 mennessä. Kyseinen kasvuvauhti tarkoittaisi sitä, että sähkön vuosikulutus Suomessa olisi tuolloin noin 102 TWh. [14]

2.2.5 Ennusteiden vertailu

Taulukkoon 2.1 on koottu kaikkien esitellyiden ennusteiden arviot Suomen sähkönkulutuksen kehittymisestä. Taulukon lukuarvot on ilmoitettu terawattitunteina.

Taulukko 2.1. Ennusteita Suomen sähkönkulutuksesta terawattitunteina aikavälillä 2020–2050

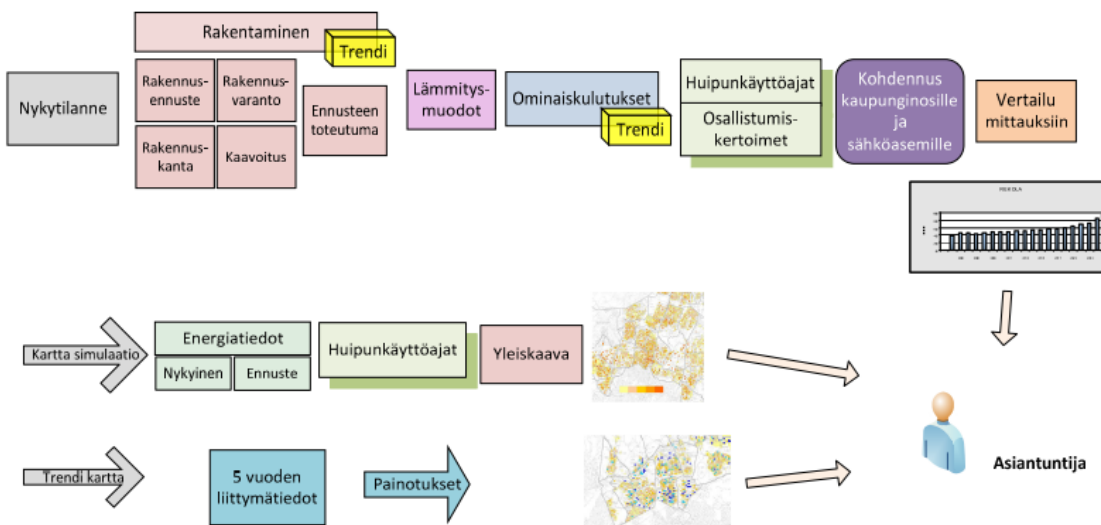
<i>Julkaisija (julkaisuvuosi)</i>	<i>2020</i>	<i>2030</i>	<i>2040</i>	<i>2050</i>
<i>EK/Energiateollisuus ry (2009, 2010)</i>	-	100–111	-	113–138
<i>VTT (2013)</i>	84–94	79–99	78–101	86–110
<i>TEM (2008)</i>	98–103	94–108	88–114	80–116
<i>TEM (2013)</i>	94	-	-	-
<i>PTT (2013)</i>	-	102	-	-

Kuten esitellyistä sähkönkulutuksen kehitystä ennustavista julkaisuista käy ilmi, on taloudella merkittävä vaikutus sähkön kysyntään ja talouden kehitys luo pohjan sähkönkulutuksen muutoksille. Viime vuosina julkaistujen raporttien antamat ennusteet poikkeavat osittain hyvinkin paljon toisistaan, osittain siksi, että kussakin ennusteessa on

käytetty pohjana senaikaisia talousennusteita, sekä osittain erilaisten tulevaisuuskenaarioiden takia. Ennustamiseen liittyvä suuri epävarmuus ilmenee julkaistujen ennusteiden suurina marginaaleina, jotka kasvavat huomattavasti ennustevälin pidentyessä. Vuoteen 2050 asti ulottuvissa ennusteissa marginaalit ovat peräti 36 TWh.

2.3 Sähkönkulutuksen ennustaminen Vantaan Energia Sähköverkot Oy:ssä

VES:ssä on käytössä taulukkolaskentaohjelmiston avulla toteutettu alueellisen sähkönkulutuksen ennustemalli, jonka avulla sähköenergian kulutusta ja tehoa ennustetaan pienalue-, kaupunginosa- ja sähköasematasolla. Koko jakelualueen kattava ennuste voidaan muodostaa esimerkiksi sähköasematason ennusteista summaamalla. Mallin avulla ennustetaan Vantaan sähkönkulutusta enimmillään 25 vuoden päähän. Mallia on käytetty pitkään ja sitä on uudistettu aika ajoin, jotta se pystyisi vastaamaan mahdollisimman hyvin toimintaympäristöä. Kuvassa 2.6 on esitetty pääpiirteisesti ennustemallin laadintaprosessi asiantuntijatyökalujen avulla kuvattuna.



Kuva 2.6. VES:n ennustemallin laadintaprosessi [15]

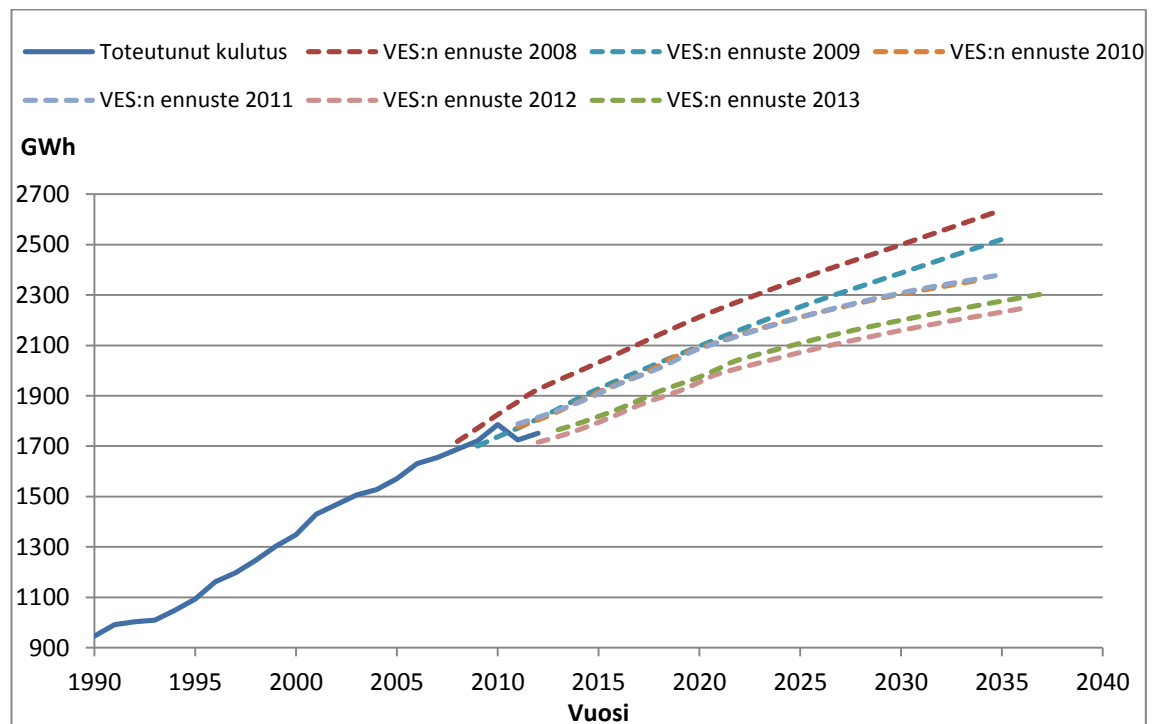
Mallissa ovat lähtötietoina nykyinen sähkönkulutus, kulutuksen tyyppi, sähkölämmitysaste, ominaiskulutus, käyttöpaikan sijainti ja rakentaminen, joka on lähtötiedoista kaikkein tärkein ennustamisen kannalta. Lähtötiedot tulisi päivittää vuosittain, ja se onnistuu kohtuullisella vaivalla. Mallissa käytetään Vantaan kaupungilta saatua rakentamisennustetta, jossa on arvioitu tulevaa rakentamista kymmenen vuoden päähän. Tästä eteenpäin tulevaa rakentamista on arvioitu ekstrapoloimalla. Rakentamisennusteet ovat yleensä optimistisia ja niitä on korjattu kertoimien avulla. Jokaiselle alueelle on myös arvioitu rakentamisen maksimi-arvo, jonka avulla on mahdollista arvioida myös sähkönkulutuksen alueellista maksimi-arvoa.

Ennustemalliin kuuluu myös kahdesta kartasta koostuva karttaosa, joka kuvaa graafisesti kuorman muutosta. Huipputehokartta kuvaa kuormitusennustetta pienaluetasolla ja trendikartta kuvaa puolestaan kuormitusennustetta halutun ajanjakson aikana käyttäen apuna historiatietoja. Koko ennustamisprosessista saatuja tuloksia arvioidaan lopuksi asiantuntijan toimesta. Verkon huipputehon kehitystä voidaan arvioida muuttamalla ennustettu vuosittainen sähkönkulutus huipputehoksi Velanderin kaavalla, joka on tavanomaisin huipputehojen laskentamenetelmä. [16]

$$P_{max} = k_1 \cdot W + k_2 \cdot \sqrt{W}, \quad (1)$$

missä P_{max} = huipputeho [kW]
 k_1, k_2 = Velanderin kertoimet
 W = vuosienergia [MWh]

Mallin avulla saadut ennusteet ovat pitäneet erittäin hyvin paikkansa vuoteen 2007 asti. Vuonna 2008 alkanut talouskriisi ja sitä seurannut lama on aiheuttanut selviä muutoksia rakentamisennusteisiin aikavälillä 2008–2013. Runsaista muutoksista huolimatta ennustemalli on antanut liian optimistisia ennusteita. Kuvassa 2.7 on esitetty Vantaan sähkönkulutus aikavälillä 1990–2012 sekä ennustemallilla laaditut ennusteet vuosina 2008–2013.



Kuva 2.7. VES:n ennustemallilla laaditut ennusteet vuosina 2008–2013

Kuva 2.7 havainnollistaa hyvin, kuinka paljon ennusteet ovat muuttuneet viime vuosien aikana. Vuonna 2008 tehty ennuste arvioi vuoden 2012 Vantaan sähkönkulutukseksi

noin 1925 GWh, joka on peräti 173 GWh suurempi kuin vuoden toteutunut sähkönkulutus 1752 GWh. Vuoden 2008 jälkeen ennustemalliin on tehty muutoksia ja vuonna 2012 laadittu ennuste arvioi Vantaan sähkönkulutuksen kehityksen merkittävästi erilaiseksi kuin vuoden 2008 ennuste. Vuosien 2012 ja 2013 ennusteet eivät poikkea paljon toisistaan. Sähkönkulutusennusteen tarkentuminen on tapahtunut vaiheittain ja kestänyt kokonaisuudessaan 4–5 vuotta, mikä antaa aihetta tutkia mahdollisuutta kehittää nykyistä ennustemallia tai luoda rinnakkainen ennustemalli, joka pystyisi reagoimaan herkemmin toimintaympäristön äkillisiin muutoksiin.

3 YLEINEN LINEAARINEN REGRESSIOMALLI

Lähtökohtana tässä työssä on oletus, että sähkönkulutus on tilastollisesti riippuvainen sosioekonomisista muuttujista. Se tarkoittaa sitä, että sähkönkulutusta ei välttämättä voida tarkasti määrittää sosioekonomisten muuttujien avulla, mutta niitä voidaan käyttää apuna sähkönkulutuksen suuruuden määrittämisessä. Tätä asiaa tullaan tutkimaan tässä työssä tilastollisin menetelmin.

Kun yhden tai useamman muuttujan vaikutusta johonkin toiseen tekijään halutaan tarkastella, voidaan pyrkiä rakentamaan tilastollista riippuvuutta kuvaava malli. Yleisin tilastotieteessä käytettävä malli on nimeltään lineaarinen regressiomalli. Tässä luvussa käydään läpi lineaarisen regressiomallin teoriaa siltä osin kuin työssä tehtävä tutkimustyö edellyttää. Teoria tullaan esittämään sen oletuksen pohjalta, että lukijalla on tuntemusta tilastotieteen perusteista.

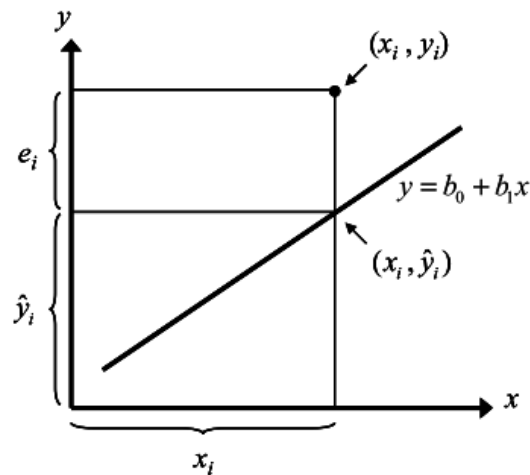
3.1 Mallin muoto

Linearisessa regressiomallissa yhden muuttujan arvo tai arvon muutos pyritään selittämään toisten muuttujien arvojen tai niiden muutosten avulla. Lineaarinen regressiomalli on muotoa: [17, s.337]

$$y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \dots + \beta_k X_{tk} + \varepsilon_t, \quad t = 1, 2, \dots, n, \quad (2)$$

missä y_t = selitettävän muuttujan arvo havaintoyksikössä t
 X_{tj} = selittävän muuttujan arvo havaintoyksikössä t
 β_0 = vakio selittäjän regressiokerroin
 β_j = selittäjän X_j regressiokerroin
 ε_t = jäännös-, eli virhetermin arvo havaintoyksikössä t

Vakio selittäjän regressiokerroin sekä selittävät muuttujat regressiokertoimien muodostavat mallin rakenneosan, jonka lausekkeella lasketaan sovite. Jäännöstermin muodostamaa osaa kutsutaan mallin jäännösosaksi. Jäännöstermi kuvaa sitä osaa selitettävän muuttujan arvosta, mitä malli ei kykene selittämään. Mitä pienempiä jäännöstermit ovat, sitä paremmin malli selittää selitettävän muuttujan y käyttäytymistä. Kuva 3.1 havainnollistaa soviteen ja jäännöstermin tulkintaa yhden selittävän tekijän tapauksessa.



Kuva 3.1. Sovitteen ja jäännöstermin geometrinen tulkinta [17, s.301]

Kuvassa 3.1 piste (x_i, y_i) on selitettävän muuttujan y_i arvo havaintoyksikössä i , ja piste (x_i, \hat{y}_i) sovitteen arvoa samassa havaintoyksikössä. Jäännösarvo ε_i ilmoittaa erotuksen $y_i - \hat{y}_i$. Käsiteltäessä aikasarjoja havaintoyksikkö on ilmoitettu i :n sijasta t :llä. On syytä huomata, että sovitteen kuvaaja on harvoin käytännössä suora, vaikka puhutaankin lineaarisesta mallista. Mallin lineaarisuudella tarkoitetaan, että malli on lineaarinen parametrien suhteen. Selittävien muuttujien suhteen malli voi olla lineaarinen tai epälineaarinen. [17]

Yleinen lineaarinen malli voidaan esittää myös matriisimuodossa seuraavasti:

$$y = X\beta + \varepsilon, \quad (3)$$

$$\text{jossa } y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} 1 & X_{11} & \cdots & X_{k1} \\ 1 & X_{12} & \cdots & X_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & \cdots & X_{kn} \end{pmatrix}, \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}, \text{ ja } \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Tilastollisen mallin muodon ja mallia koskevien oletuksien määrittelemistä kutsutaan mallin täsmennykseksi ja se pitää sisällään seuraavat valinnat: [17, s. 386]

- selitettävän ja selittävien muuttujien valinta
- rakenneosan funktionaalisen muodon ja parametrisoinnin valinta
- selitettävän muuttujan ja selittäjien funktionaalisen muodon valinta
- jäännöstermiä koskevien stokastisten oletusten valinta

Mallin täsmennys pyritään suorittamaan siten, että lopullisen mallin rakenne ja muuttujien muoto vastaisivat parhaalla mahdollisella tavalla tutkimuksen tavoitteita.

3.2 Regressiokertoimien laskeminen

Yleisen lineaarisen mallin regressiokertoimia ei käytännössä pystytä määrittelemään siten, että malli selittäisi y :n täysin, vaan kertoimille pyritään löytämään mahdollisimman tarkat estimaattorit, joiden tarkoitus on sovittaa malli vastaamaan mahdollisimman hyvin selitettävän muuttujan arvojen vaihtelua. Regressiokertoimen β_j estimaattoria merkitään kirjaimella b_j . Estimointiin yleisimmin käytetty menetelmä on nimeltään pienimmän neliösumman (PNS) menetelmä, jossa regressiokertoimet estimoidaan minimoimalla jäännöstermien neliösumma, eli laskemalla lausekkeen

$$N(\beta_0, \beta_1, \dots, \beta_k) = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 X_{i,1} - \dots - \beta_k X_{i,k})^2 \quad (4)$$

pienin arvo. Estimaatit b_0, b_1, \dots, b_k saadaan osittaisderivoimalla kaava (4) regressiokertoimien suhteen, jolloin osittaisderivaatat ovat

$$\begin{aligned} \frac{\delta N}{\delta \beta_0} &= -2 \sum_{t=1}^n 1 \cdot (y_t - \beta_0 - \beta_1 X_{t,1} - \dots - \beta_k X_{t,k}), \\ \frac{\delta N}{\delta \beta_1} &= -2 \sum_{t=1}^n X_{t,1} \cdot (y_t - \beta_0 - \beta_1 X_{t,1} - \dots - \beta_k X_{t,k}), \\ &\vdots \\ \frac{\delta N}{\delta \beta_k} &= -2 \sum_{t=1}^n X_{t,k} \cdot (y_t - \beta_0 - \beta_1 X_{t,1} - \dots - \beta_k X_{t,k}). \end{aligned}$$

Kun osittaisderivaatat merkitään nolliksi, voidaan yhtälöt esittää yhdessä matriisimuodossa ja ratkaista vektori b . [17, s.344]

$$\begin{aligned} 0 &= -2X^T(y - Xb) \\ 0 &= X^T(y - Xb) \\ 0 &= X^T y - X^T X b \\ X^T X b &= X^T y \\ b &= (X^T X)^{-1} X^T y \end{aligned} \quad (5)$$

Jokaiselle estimaattorille voidaan määrittää luottamusväli, joka peittää regressiokertoimen todennäköisyydellä $(1 - \alpha)$. Yleisesti on tapana käyttää 95 % ($\alpha = 0,05$) luottamusväliä, jolloin luottamusväli voidaan esittää muodossa

$$b_j \pm 1,96 \cdot S(b_j), \quad (6)$$

jossa $S(b_j) = b_j$:n keskivirhe $= \left(\frac{\sum_{t=1}^n \varepsilon_t^2}{n - k - 1} \right)^{1/2}$. [18, s. 194]

Estimaattoreiden laskeminen onnistuu kätevimmin käyttämällä ohjelmistoja, jotka laskevat estimaattorit automaattisesti. Tässä työssä mallin muodostamiseen käytetään Excel-taulukkolaskentaohjelmistoa.

3.3 Mallin hyvyys

Tehdyn mallin hyvyyttä voidaan kuvata esimerkiksi selitysasteen R^2 avulla. Selitysaste mittaa mallin selittämää osuutta selitettävän muuttujan havaintoarvojen kokonaisvaihtelusta.

$$R^2 = 1 - \frac{SSE}{SST} = \frac{SSM}{SST}, \quad (7)$$

jossa $SSE = \sum_{t=1}^n \varepsilon_t^2 =$ jäännösneliösumma

$SST = \sum_{t=1}^n (y_t - \bar{y})^2 =$ kokonaisneliösumma

$SSM = SST - SSE =$ mallineliösumma

Kokonaisneliösumma kuvaa selitettävän muuttujan havaittujen arvojen kokonaisvaihtelua, mallineliösumma mallin selittämää osaa ja jäännösneliösumma selittämättä jäänyttä osaa tuosta kokonaisvaihtelusta. Selitysaste saa arvoja välillä 0–1, ja se on tapana ilmaista prosentteina. [17, s. 357]

3.4 Mallia koskevat standardioletukset

Yleisestä lineaarisesta mallista tehdään tavallisesti seuraavat oletukset, joita kutsutaan standardioletuksiksi. Nämä oletukset on oltava voimassa, jotta kaavan (3) mukainen PNS-estimaattori olisi pätevä ja myöhemmin esitettävät estimointi- ja testausmenetelmät olisivat soveltamiskelpoisia mallin analysoimisessa. [17, s. 337]

- (i) Malli on lineaarinen parametrien suhteen.
- (ii) Selittäjien välillä ei ole lineaarisia riippuvuuksia.
- (iii) Kaikilla jäännöstermeillä on odotusarvona 0.
- (iv) Kaikilla jäännöstermeillä on sama varianssi.
- (v) Jäännöstermit eivät korreloi keskenään.
- (vi) Jäännöstermit ovat normaalijakautuneita.

Oletusta (iv) kutsutaan homoskedastisuusoletukseksi. Jos oletus (iv) pätee, jäännöstermejä kutsutaan homoskedastisiksi. Jos oletus (iv) ei päde, kutsutaan jäännöstermejä heteroskedastisiksi. [17, s. 338]

Oletus (vi) pitää sisällään oletukset (iii) ja (iv). Oletuksia voidaan tutkia regressiodiagnostiikan avulla, joka käsitellään myöhemmin. Osaa oletuksista voidaan lieventää tai niistä voidaan jopa luopua. Yhdestä tai useammasta oletuksesta luovuttaessa täytyy käyttää tavanomaisesta poikkeavia estimointi- ja testausmenetelmiä. [17, s. 339]

4 SELITTÄVIEN MUUTTUIJEN VALINTA

Muuttujien valinta on regressioanalyysin keskeisin ja vaikein tehtävä. Malliin pyritään valitsemaan sopiva määrä selitysvoimallisia muuttujia, jotta mallin avulla saataisiin riittävän tarkka arvio selitettävästä muuttujasta. Toisaalta suuri selittävien muuttujien määrä voi tehdä mallista liian monimutkaisen.

Muuttujien valinnassa voidaan aluksi hyödyntää esimerkiksi riippuvuustarkastelua, jossa tutkitaan, onko selittävän ja selitettävän muuttujan välillä lineaarista tilastollista riippuvuutta, eli korrelaatiota. Mikäli muuttuja korreloi riittävästi selitettävän muuttujan kanssa, voidaan muuttuja valita regressiomallin selittäjäkandidaatiksi. Muuttujien sopivuus malliin täytyy varmentaa myös muilla tilastollisilla testeillä ja taustateorian avulla.

Hyvän regressiomallin selitysaste on suuri ja kaikki regressiokertoimet ovat tilastollisesti merkitseviä. Jos regressiomallista puuttuu siihen kuuluvia selittäjiä, mallin selittäjien PNS-estimaattorit voivat olla harhaisia, mikä tarkoittaa että estimaattorit kuvaavat väärin selitettävän ja selittävien suureiden suhdetta. Jos regressiomallissa on turhia selittäjiä, mallin regressiokertoimien PNS-estimaattorit voivat olla tehottomia, jolloin niiden varianssit ovat turhan suuria. Selittävien muuttujien valinnassa on yleisesti käytössä kaksi erilaista menetelmää: mallinvalintatestien ja mallinvalintakriteerien käyttö. [17, s. 374]

4.1 Mallinvalintatestit

Mallinvalintatestien avulla malliin pyritään valitsemaan kaikki tilastollisesti merkitsevät selittäjät johdonmukaista testausstrategiaa noudattaen. Kun käytetään mallinvalintatestejä, on ensin muodostettava jokin lähtömalli, johon tilastollisesti merkitseviä selittäjiä pyritään lisäämään ja josta ei-tilastollisesti merkitsevät selittäjät pyritään poistamaan. Mallinvalintatestit, kuten F- ja t-testit perustuvat hypoteesitestaukseen, jonka tyypillinen muoto tulee esille testien esittelyn yhteydessä.

4.1.1 T-testi

T-testi on tilastollinen testi, jota yleisesti käytetään testaamaan selittäjien merkitsevyyttä. Asetetaan nollihypoteesi

$$H_{0j}: \beta_j = 0, \quad j = 0, 1, 2, \dots, k,$$

sekä vaihtoehtoinen hypoteesi

$$H_{1j}: \beta_j \neq 0,$$

joka pätee, jos nollahypoteesi ei päde. Mikäli nollahypoteesi H_{00} pätee, mallissa ei ole vakioita, ja jos nollahypoteesi H_{0j} pätee, selitettävän muuttujan y ja selittävän muuttujan X_j välillä ei ole lineaarista riippuvuutta. Määritellään t-testisuure: [18, s. 117]

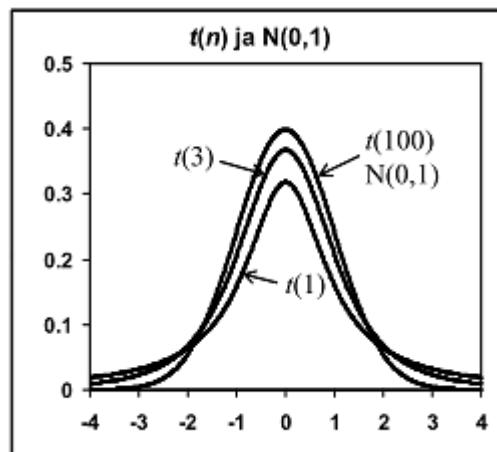
$$t_j = \frac{b_j - \beta_j}{s}, \quad (8)$$

jossa b_j = regressiokertoimen estimaattori

β_j = muuttujan X_j regressiokerroin

$$s = b_j\text{:n keskivirhe} = \left(\frac{\sum_{t=1}^n e_t^2}{n - k - 1} \right)$$

Testisuure noudattaa t-jakaumaa, jonka tiheysfunktion muoto riippuu vapausasteesta n . Tiheysfunktion $t(n)$ muotoutumista on esitetty kuvassa 4.1.



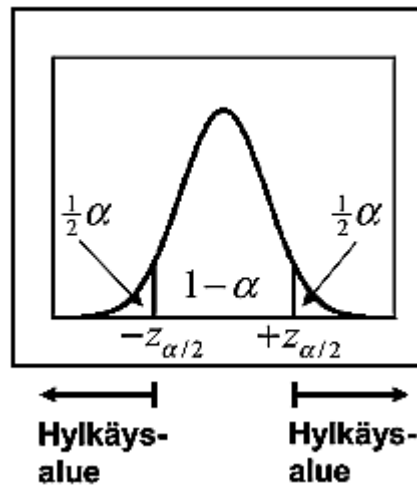
Kuva 4.1. T-jakauman tiheysfunktio [19]

Kuvassa 4.1 on esitetty myös standardoitu normaalijakauma $N(0,1)$. Kuvan avulla voidaan nähdä, että t-jakauma lähestyy standardoitua normaalijakaumaa vapausasteiden lukumäärän kasvaessa. Normaalijakauma-approksimaatio t-jakaumalle on kohtuullisen hyvä jo, kun $n = 30$, ja voidaan olettaa, että $t(100) \sim N(0,1)$. Tätä ominaisuutta voidaan hyödyntää laskettaessa t-arvoa vastaavaa p-arvoa. [18, s. 118]:

$$p = 2\Phi(-|t|), \quad (9)$$

jossa Φ on normaalijakauman kertymäfunktio ja t on t-testin laskettu arvo. P-arvon suuruuteen perustetaan nollahypoteesin mahdollinen hylkääminen. Se kertoo pienimmän merkitsevyytason, jolla nollahypoteesi voidaan hylätä. Itseisarvoltaan suuret t-arvot viittaavat siihen, että H_{0j} ei päde. Yleisesti käytetty kriittinen arvo α merkitsevyytason-

le on 5 %, eli p-arvon ollessa pienempi kuin 0,05 voidaan nollahypoteesi hylätä. Koska vaihtoehtoinen hypoteesi H_{1j} on kaksisuuntainen, voidaan nollahypoteesin hylkäysalue esittää kuvan 4.2 mukaisella tavalla.



Kuva 4.2. Hylkäysalueiden määräytyminen T -testissä [17]

Kaksisuuntaisen vaihtoehtoisen hypoteesin tapauksessa kriittiset arvot ovat molemmin puolin $\frac{1}{2}\alpha$.

4.1.2 F-testi

F-testiä voidaan pitää yleistestinä regression olemassaololle, sillä se kuvaa selittävien tekijöiden yhteistä selitysvoimaa. Siinä määritellään aluksi nollahypoteesi

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0,$$

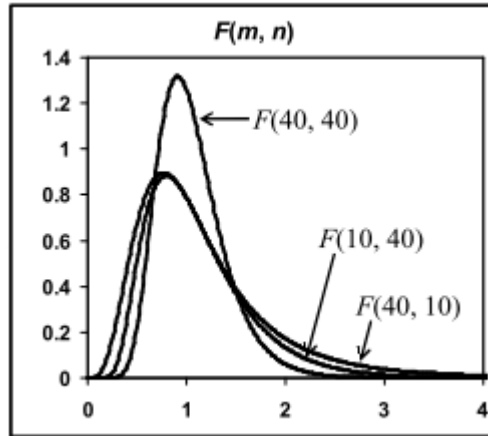
sekä vaihtoehtoinen hypoteesi

$$H_1: \beta_1 = \beta_2 = \dots = \beta_k \neq 0.$$

Jos nollahypoteesi pätee, selitettävällä muuttujalla ei ole lineaarista riippuvuutta yhdenkään selittäjän kanssa, eikä malli näin ollen ole käyttökelpoinen. Määritellään F-testisuure:

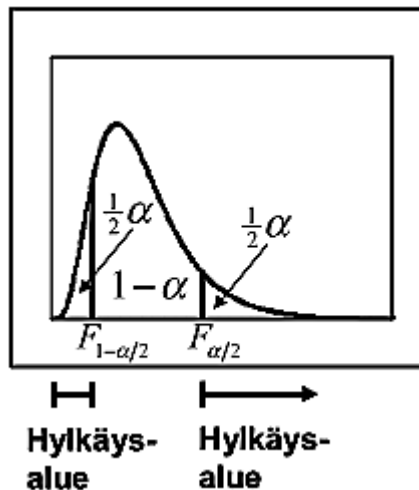
$$F = \frac{n-k-1}{k} \cdot \frac{R^2}{1-R^2}, \quad (10)$$

joka noudattaa F-jakaumaa vapausastein k ja $(n - k - 1)$. F-jakauman tiheysfunktion muoto riippuu sen vapausasteista. Kuvassa 4.3 on esitetty F-jakauman $F(m, n)$ tiheysfunktio välillä $[0, 4]$ muutamilla $m:n$ ja $n:n$ arvoilla.



Kuva 4.3. $F(m,n)$:n tiheysfunktio [19]

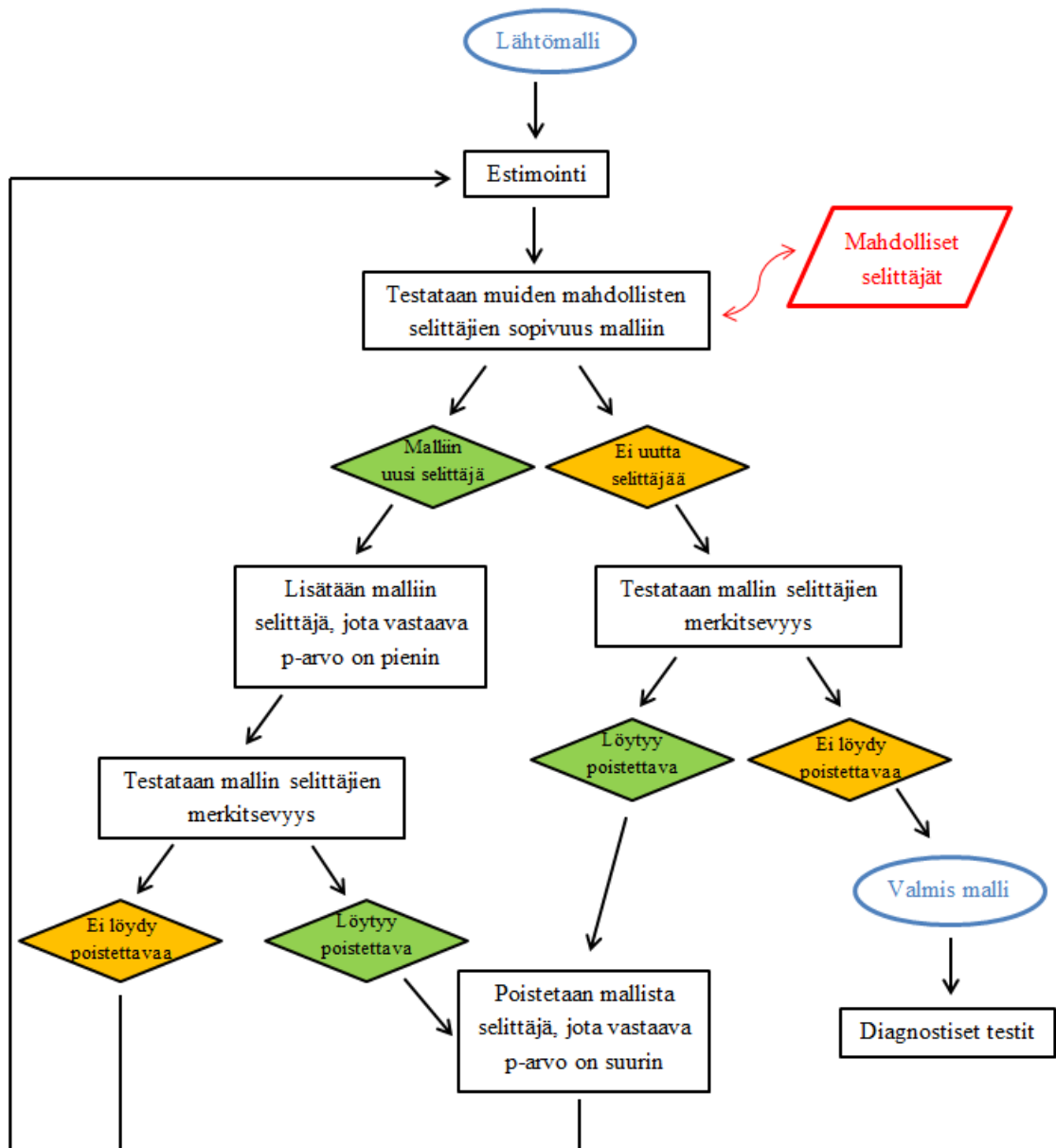
Jos nollahypoteesin F-arvoa vastaava p-arvo on pienempi kuin valittua merkitsevyystasoa vastaava p-arvo, voidaan nollahypoteesi hylätä ja mallia voidaan pitää käyttökelpoisena. Hylkäysalue F-testissä voidaan esittää kuvan 4.4 mukaisella tavalla.



Kuva 4.4. Hylkäysalueiden määräytyminen F-testissä [17]

Suuret F-testin arvot viittaavat siihen, että selittävillä tekijöillä on hyvä yhteinen selitysvaivoima.

Selittäjien lisääminen tai poistaminen mallista ei kuitenkaan ole ongelmattonta. Ei-merkitsevien selittäjien poistamisjärjestys saattaa vaikuttaa lopputulokseen, sillä yhden ei-merkitsevän selittäjän poistaminen mallista saattaa muuttaa aikaisemmin poistetun selittäjän merkitseväksi, jos se otettaisiin takaisin malliin. Toisaalta uuden selittäjän lisääminen malliin saattaa muuttaa toisen mallissa olevan selittäjän ei-merkitseväksi. Tämä ongelma voidaan välttää noudattamalla askeltavan regression mallinvalintastrategiaa, jonka yhden askeleen kulkua kuva 4.5 havainnollistaa.



Kuva 4.5. Askeltavan regression mallinvalintastrategia

Strategia lähtee tilanteesta, jossa on valittu jokin lähtömalli sekä joukko mahdollisia selittäjiä, joiden sopivuutta malliin testataan yksitellen mallinvalintatestien avulla. Valittuja merkitsevyystasoja noudattaen malliin voidaan yhdessä askeleessa lisätä tai poistaa yksi selittävä muuttuja kerrallaan, minkä jälkeen askel suoritetaan uudelleen aloittaen uuden mallin estimoinnilla. Mikäli lisättäviä tai poistettavia muuttujia ei löydy, on malli valmis diagnostisia testejä varten. [17, s. 376]

4.2 Mallinvalintakriteerit

Mallinvalintakriteerien avulla pyritään malliin valitsemaan kaikkien mahdollisten selittäjien joukosta sellainen osajoukko, joka optimoi käytetyn kriteerifunktion arvon. Yksi

esimerkki mallivalintakriteeristä on aikaisemmin esitelty selitysaste R^2 . Hyvän regressiomallin selitysaste on korkea, mikä tarkoittaa sitä, että jäännösneliösumma on pieni. Tällöin estimoidun mallin sovitteet ovat lähellä selitettävän muuttujan havaittuja arvoja ja mallin jäännöstermit ovat pieniä. Aina kun malliin lisätään mikä tahansa muuttuja, jäännösneliösumma pienenee tai pysyy samana. Jäännösneliösumman minimointia ei tästä syystä voida käyttää mallin valinnassa, sillä se johtaisi kaikkien selittäjäehdokkaiden valintaan. [17, s. 378]

Hyvissä mallivalintakriteereissä jäännösneliösummaan lisätään sakkofunktio, jonka arvo riippuu mallissa olevien selittäjien lukumäärästä. Jos lisätyt selittäjät eivät pienennä jäännösneliösummaa tarpeeksi, kasvattaa sakkofunktio kriteerifunktion arvoa. Kaikki kriteerit johtavat harhattomaan malliin, vaikka eri kriteerit saattavat johtaa erilaiseen malliin [16]. Selittävien muuttujien sopiva määrä tullaan tässä työssä valitsemaan Schwarzin Bayeslaisen informaatiokriteerin (SBIC), Akaiken informaatiokriteerin sekä korjatun selitysasteen avulla. Schwarzin Bayeslainen informaatiokriteeri on muotoa:

$$SBIC = \ln \left[\frac{SSE}{n} \right] + (p + 1) \frac{\ln(n)}{n}, \quad (11)$$

jossa p tarkoittaa mallissa olevien selittävien muuttujien lukumäärää. Näin ollen termi $(p + 1)$ ilmoittaa mallissa olevien kertoimien lukumäärän vakioselittäjän regressiokerroin β_0 mukaan lukien. [18, s. 585] Vertailtaessa eri malleja SBIC:n avulla, on kaikissa malleissa oltava havaintoarvoja sama määrä n .

Mikäli epäillään, että SBIC:in avulla saatu selittävien muuttujien määrä on liian pieni, voidaan käyttää Akaiken informaatiokriteeriä (AIC) asian varmistamiseen.

$$AIC = \ln \left[\frac{SSE}{n} \right] + (p + 1) \frac{2}{n} \quad (12)$$

Molempien kriteerien mukaan malli on sitä parempi, mitä pienempi testisuureen arvo on. AIC poikkeaa SBIC:stä vain testisuureen jälkimmäisen termin osalta, jossa $\ln(n)$ on korvattu vakiolla 2. Tämä ero johtaa käytännössä siihen, että malliin voi valikoitua suurempi määrä selittäviä muuttujia. [18, s. 585]

Korjattu selitysaste on nimensä mukaisesti johdettu mallin selitysasteesta.

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1} \cdot \frac{SSE}{SST}, \quad (13)$$

joka saa arvon väliltä 0–1 aivan kuten selitysastekin, ja hyvän mallin korjattu selitysaste on lähellä arvoa 1. Funktion arvo voi kuitenkin pienentyä, jos selittävien muuttujien määrän k kasvaa, mutta jäännösneliösumma ei pienene riittävästi. [18, s. 236]

Vaikka malli olisi kriteerien mukaan paras kaikista vertailluista malleista, on sen myös täytettävä asialogiset kriteerit, eli sen on kuvattava riippuvuutta sisällöllisesti oikein, tarkoittaen että parametrit ovat tulkittavissa ja ne ovat järkevä merkkisiä ja

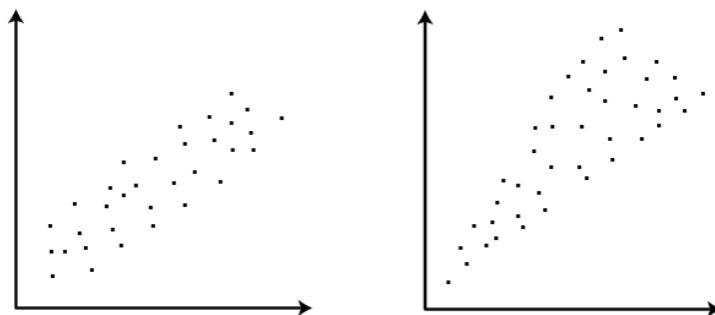
muotoisia. Lisäksi mallin on kuvattava todellisuutta mielekkäällä tavalla ja sen on selvitettävä diagnostisista tarkistuksista.

4.3 Regressiodiagnostiikka

Regressioanalyysin perusongelma on, kuvaako täsmennetty regressiomalli riippuvuutta oikein. Tämän yleisongelman osaongelma on mallista tehtyjen oletusten paikkansapitävyys, joka kertoo, kuvaako malli riippuvuutta tilastollisesti oikein. Regressiodiagnostiikassa tarkistetaan, onko mallista saadut estimointitulokset sopusoinnussa mallista tehtyjen oletusten kanssa. Regressiodiagnostiikassa estimoinnin onnistumista kuvataan diagnostisilla tunnusluvuilla ja havainnollistetaan tilastografiikan avulla. Mallia koskevien oletusten paikkansapitävyys testataan diagnostisilla testeillä. Kaikkien näiden menetelmien avulla tarkistetaan mallista seuraavat asiat: [17, s.392]

- Onko havaintoarvoissa poikkeavia havaintoja?
- Onko malli lineaarinen parametrien suhteen?
- Ovatko selittäjät itsenäisiä?
- Ovatko jäännöstermit homoskedastisia, korreloimattomia ja normaalisia?

Havaitut ja estimoidut arvot on hyvä esittää grafiikkana samassa pistediagrammissa ja estimoidun mallin jäännöstermit omassa pistediagrammissaan. Jäännöstermidia grammista voidaan silmämääräisesti havaita, onko mallin jäännöstermi homoskedastinen, mikä on yksi mallin perusoletuksista. Mikäli jäännöstermisarjan muodostama pistepilvi ei ole tasaleveä, saattaa jäännöstermi olla heteroskedastinen. Se voi tarkoittaa sitä, että mallin täsmennyksessä on tapahtunut virhe. Kuvan 4.6 vasemmanpuoleisessa diagrammissa on esimerkki homoskedastisesta jäännöstermijoukosta, joka on tasaleveä. Oikeanpuoleisessa diagrammissa oleva jäännöstermijoukko ei ole tasaleveä, mikä viittaa jäännöstermien heteroskedastisuuteen.



Kuva 4.6. Homo- ja heteroskedastisten jäännöstermien pistediagrammit [20]

Jäännöstermidia grammi voi paljastaa myös jäännöstermien sarjan autokorrelaatiokentteen sekä poikkeavat havainnot. Jäännössarjan autokorrelaatio tarkoittaa, että jäännöstermit korreloivat keskenään. Perusoletus (v):n mukaan jäännöstermien autokorre-

laatio ei ole toivottavaa. Grafiikan avulla tehtävät havainnot ovat ainoastaan suuntaa antavia ja vasta niiden jälkeen laskettavien diagnostisten tunnuslukujen avulla voidaan tehdä lopulliset johtopäätökset jäännöstermien ominaisuuksista.

4.3.1 Poikkeavat havainnot

Poikkeavalla havainnolla tarkoitetaan havaintoa, joka eroaa merkitsevästi muista havainnoista. Havaintoa voidaan pitää tilastollisesti poikkeavana, jos se vääristää tilastollisen analyysin tuloksia. Yleensä poikkeavat havainnot huomataan jäännösarvodiagrammeista, mutta on olemassa myös niiden havaitsemiseen tarkoitettuja tunnuslukuja. Tässä työssä pyritään löytämään mahdolliset poikkeavat arvot analysoimalla mallin jäännöstermijoukkoa grafiikan avulla sekä studentisoitujen jäännöstermien perusteella.

Jäännöstermien studentisointi on usein käytetty menetelmä poikkeavien havaintojen tunnistamiseen. Ne saadaan muodostettua mallin jäännöstermeistä kaavalla

$$Std(\varepsilon_t) = \frac{\varepsilon_t}{s_{\varepsilon}^2(1-h_{tt})}, \quad (14)$$

jossa h_{tt} on hattumatriisin

$$P = X(X^T X)^{-1} X^T \quad (15)$$

t. diagonaalialkio. Jos PNS-menetelmällä luotu malli kykenee kuvaamaan kaikkia havaintoarvoja, saavat studentisoitujen jäännöstermien itseisarvot vain pienellä todennäköisyydellä suurempia arvoja kuin t-jakauman $T(n-1-k)$ kriittinen raja. Kriittisen rajan ylittävät arvot saattavat viitata poikkeaviin havaintoihin. [17, s. 396]

4.3.2 Selittäjien itsenäisyys

Standardioletuksen (ii) mukaan regressiomallissa olevien selittäjien välillä ei ole lineaarisia riippuvuuksia. Multikollineaarisuus on ilmiö, jossa selittäjien välillä on jonkin verran lineaarista riippuvuutta ja ne selittävät osittain samalla tavalla selitettävän muuttujan käyttäytymistä. Jotta regressiokertoimien estimaattorit voitaisiin laskea kaavan (4) avulla, ei selittäjillä saa olla kovin voimakasta multikollineaarisuutta. Mitä vähäisempää se on, sitä itsenäisemmin selittäjät toimivat osana mallia. Multikollineaarisuus on suhteellinen suure, jota voidaan tutkia esimerkiksi varianssin inflaatiotekijän avulla.

$$VIF_j = \frac{1}{1-R_j^2}, \quad (16)$$

jossa R_j^2 on selitysaste apuregressiomallista, jonka selitettävänä muuttujana on alkupe-
räisen regressiomallin selittävä muuttuja X_j ja selittävinä muuttujina kaikki muut alku-
peräisen regressiomallin selittävät muuttujat. [17, s. 403]

Mitä suurempi VIF_j on, sitä enemmän multikollinearisuutta on havaittavissa selittävien muuttujien välillä. Tulkinan varassa on, kuinka suuri VIF_j voi olla, ennen kuin multikollinearisuus on haitallisella tasolla. Yleisesti voidaan olettaa, että multikollinearisuus on siedettävällä tasolla VIF_j :n arvon ollessa pienempi kuin 5. Tämä tarkoittaa sitä, että kaikki muut selittävät muuttujat pystyvät selittämään X_j :n arvojen vaihtelut alle 80 %:n tarkkuudella.

4.3.3 Jäännöstermien homoskedastisuus

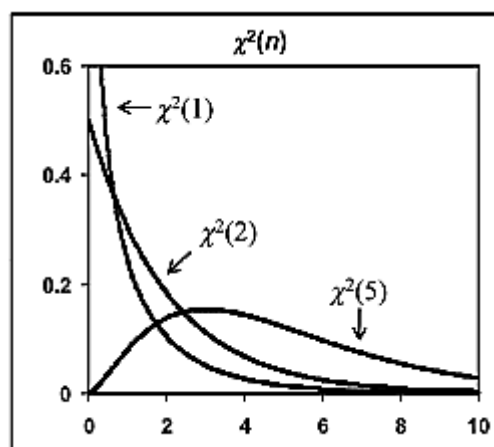
Standardioletuksen (iv) mukaan jäännöstermeillä on oltava sama varianssi. Mikäli tämä homoskedastisuusoletus ei toteudu, tulee regressiokertoimien PNS-estimaattoreiden variansseista suuria, mikä vaikuttaa regressiokertoimien tilastolliseen testaamiseen. Homoskedastisuutta voidaan testata esimerkiksi White-testillä. White-testissä on nollahypoteesinä oletus jäännöstermien homoskedastisuudesta ja se ilmoitetaan muodossa:

$$H_0: \sigma_t^2 = \sigma^2 \text{ kaikille } t.$$

Testi perustuu apuregressioon, jossa jäännöstermin neliö on selitettävänä muuttujana ja selittävinä muuttujina käytetään alkuperäisen mallin selittäjiä, niiden neliöitä sekä niiden ristituloa. Apuregressiomalli on esimerkiksi kahden selittävän muuttujan tapauksessa muotoa:

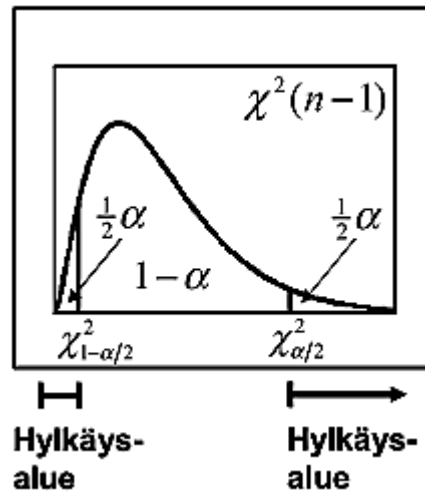
$$\varepsilon_t^2 = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \beta_3 X_{t1}^2 + \beta_4 X_{t2}^2 + \beta_5 X_{t1} X_{t2}, \quad (17)$$

jonka selitysaste kerrotaan havaintojen määrällä n . Testiarvo on siis muotoa nR^2 , joka on jakautunut $\chi^2(k)$ -jakauman mukaisesti, jossa k on apuregression (17) selittävien muuttujien lukumäärä [21]. Kuva 4.7 havainnollistaa, kuinka $\chi^2(n)$ -jakauman tiheysfunktion muoto riippuu voimakkaasti vapausasteesta n .



Kuva 4.7. $\chi^2(n)$ tiheysfunktion muoto [19]

Hylkäysalue voidaan esittää kuvan 4.8 mukaisella tavalla.



Kuva 4.8. Hylkäysalueen määräytyminen χ^2 -jakaumassa [17]

Jos nR^2 :n arvoa vastaava p-arvo ylittää valitun kriittisen arvon, jää nollahypoteesi voimaan ja jäännöstermien voidaan todeta olevan homoskedastisia.

4.3.4 Jäännöstermien korreloimattomuus

Standardioletuksen (v) mukaan mallin jäännöstermit ovat korreloimattomia. Mikäli oletus ei toteudu, ovat vaikutukset samat kuin homoskedastisuusoletuksen kohdalla. Erityisesti aikasarjojen regressiomalleja tutkittaessa voi ilmentyä jäännöstermien autokorreloituneisuutta, jolloin jäännöstermit voidaan ilmoittaa muodossa

$$\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \dots + \rho_p \varepsilon_{t-p} + u_t, \quad (18)$$

jossa ilmenee autokorrelaatiota p . kertaluvun verran. Jäännöstermien 1. kertaluvun autokorrelaatiota $AR(1)$ voidaan testata Durbin-Watsonin d -testillä. Testissä on nollahypoteesina oletus jäännöstermien korreloimattomuudesta:

$$H_0: \rho = 0$$

Testisuureen funktio on muotoa

$$d = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^n \varepsilon_t^2}. \quad (19)$$

Testisuure saa arvonsa väliltä 0 – 4, ja sitä tulkitaan kriittisten ylä- ja alarajojen d_U ja d_L avulla seuraavasti:

Jos $d < d_L$	H_0 voidaan hylätä
Jos $d > d_U$	H_0 pysyy voimassa
Jos $d_L < d < d_U$	testi on tulokseton.

Kriittisten ylä- ja alarajojen arvot riippuvat havaintojen määrästä n sekä selittävien muuttujien määrästä k . Yleisesti voidaan kuitenkin todeta, että nollahypoteesi pysyy voimassa, jos testisuureen arvo on lähellä sen normaaliarvoa 2. [22]

Koska d-testi tarkastelee vain AR(1)-autokorrelaatiota, on jäännöstermien korreloimattomuutta hyvä testata myös Breusch-Godfrey LM (Lagrange multiplier) -testillä, jolla voidaan testata sarjakorrelaatiota $AR(p)$. Tässä testissä on myös nollahypoteesina oletus jäännöstermien korreloimattomuudesta ja siinä käytetään apuregressiota

$$\varepsilon_t = \alpha_0 + \alpha_1 X_{t1} + \dots + \alpha_k X_{tk} + \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \dots + \rho_p \varepsilon_{t-p} + u_t, \quad (20)$$

jonka selitysaste R^2 kerrotaan apuregressiossa käytettävien havaintojen määrällä $(n - p)$. $(n - p)R^2$ on jakautunut $\chi^2(p)$ -jakauman mukaisesti. Mikäli lausekkeen $(n - p)R^2$ arvoa vastaava p-arvo ylittää valitun kriittisen arvon, jää nollahypoteesi voimaan ja voidaan todeta jäännöstermien olevan korreloimattomia. [23]

4.3.5 Jäännöstermien normalisuus

Jos regressiomallin jäännöstermit eivät ole normaalisia, eivät myöskään regressioker-toimien PNS-estimaattoreiden otosjakaumat ole normaalisia. Tällöin t- ja F-jakaumiin perustuva tilastollinen päättely ei välttämättä ole pätevää. Mikäli poikkeamat normaali-suudesta ovat kuitenkin vain kohtuullisen pieniä, voidaan t- ja F-jakaumiin perustuvaa tilastollista päättelyä pitää suuntaa-antavana. Tässä työssä otosten määrä pysyy pienenä, joten normalisuusoletusta ei voida sivuuttaa suurten otosten teoriaan nojaten.

Jäännöstermien normalisuutta voidaan testata Shapiro-Wilkin W-testillä, jossa on nollahypoteesina oletus jäännöstermien normalisuudesta. Testifunktio on muotoa:

$$W = \frac{(\sum_{t=1}^n a_t \varepsilon_i)^2}{\sum_{t=1}^n (X_t - \bar{X})^2}, \quad (21)$$

jossa $\varepsilon(i)$ on järjestetyn jäännöstermijoukon i :s pienin termi ja a_i on jäännöstermien keskiarvon varianssin ja kovarianssin avulla muodostettu vakio. W-arvoa vastaavan p-arvon ollessa suurempi kuin valittu kriittinen arvo 0,05, pysyy nollahypoteesi voimassa ja jäännöstermit ovat normaalisti jakautuneita. [24]

5 AIKASARJAT JA ENNUSTAMINEN

Luvuissa kolme ja neljä käsiteltiin yleistä lineaarista mallia, joka pyrkii kuvaamaan mahdollisimman hyvin selitettävän muuttujan käyttäytymistä. Malli voitiin muodostaa yhden tai useamman selittävän muuttujan havaintoarvoista. Tässä työssä havaintoarvot ovat aikasarjoja, mikä tarkoittaa, että ne ovat aikajärjestyksessä. Aikasarjojen käyttö ei millään tavalla estä käyttämästä lineaarista regressiomallia, mutta mallin muodostuksessa on otettava huomioon aikasarjoille ominaiset piirteet, kuten autoregression yleisyys ja stationaarisuuden testaaminen. Autoregression yleisyys tuo mukanaan myös mahdollisuuden käyttää viiveitä apuna mallin muodostamisessa.

Aikasarjojen avulla muodostettu regressiomalli on muodostettu historia-arvojen avulla kuvastamaan selitettävän muuttujan aikaisempaa käyttäytymistä. Jos oletetaan, että tulevaisuus on historian kaltainen, on mallilla mahdollista ennustaa, mitä arvoja selitettävä muuttuja saa tulevaisuudessa. Ennusteessa voidaan käyttää hyväksi viiveitä sekä selittävien muuttujien ennustettuja arvoja. Selittävien muuttujien valinnan yhteydessä tulee ottaa huomioon arvojen ennustettavuus.

Mallin ennustuskyvyn mittaaminen on haasteellista, kuten tulevaisuuden ennustaminen aina onkin. Eräs menetelmä ennustuskyvyn mittaamiseen on pseudo out-of-sample -analyysi, jossa ennusteen arvoja verrataan niihin selitettävän tekijän toteutuneisiin arvoihin, joita ei käytetty regressiokertoimien estimointiin.

5.1 Viiveet

Kuten todettu, autoregressio on aikasarjojen yleinen ominaisuus. Autoregressiossa selitettävän muuttujan arvo y_t on riippuvainen sen itsensä aikaisemmin saaduista arvoista. Jos myös selittävien muuttujien aikaisemmillä arvoilla on vaikutusta, voidaan näiden kaikkien avulla muodostaa ennustava regressiomalli.

$$\begin{aligned}
 y_t = & \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \dots + \beta_p y_{t-p} \\
 & + \lambda_{11} X_{1t-1} + \lambda_{12} X_{1t-2} + \dots + \lambda_{1q_1} X_{1t-q_1} \\
 & + \dots + \lambda_{k1} X_{kt-1} + \lambda_{k2} X_{kt-2} + \dots + \lambda_{kq_k} X_{kt-q_k} + \varepsilon_t.
 \end{aligned} \tag{22}$$

Kaava (22) on kaavan (2) mukainen lineaarinen regressiomalli, johon pätee samat perusoletukset (i) – (vi) ja jonka regressiokertoimien estimaatit voidaan määrittää kaavan (4) mukaisella tavalla. Viiveiden käyttö voi auttaa kuvaamaan paremmin selitettävän muuttujan käyttäytymistä, mutta kuten kaavasta (22) voidaan nähdä, voi selittävien muuttujien määrä kasvaa helposti hyvinkin suureksi. Jos viivettä on käytetty liian pal-

jon, tarvitsee mallissa estimoida turhia kertoimia, mikä kasvattaa ennusteen virhettä. Jos viivettä on otettu huomioon liian vähän, jää osa selittävästä informaatiosta hyödyntämättä. Viiveiden oikean määrän valinta on samanlainen ongelma kuin luvussa neljä esitetty sopivien muuttujien valintaan liittyvä ongelma, jossa apuna voidaan käyttää SBIC:tä sekä AIC:tä.

5.2 Stationaarisuus

Aikasarjan stationaarisuudella tarkoitetaan, että sarjan arvojen keskiarvot, varianssit ja kovarianssit ovat ajasta riippumattomia. Tällöin sen liikkeitä on helpompi ennustaa olettaen, että tulevaisuus on menneisyyden kaltainen. Aikasarjojen kohdalla esiintyy usein epästationaarisuutta esimerkiksi trendien muodossa. Epästationaaristen sarjojen käyttö regressiomallissa voi johtaa epäpäteviin tuloksiin.

Yksikköjuuriongelmaa pidetään yleisimpänä epästationaarisuutta aiheuttavana tekijänä. Otetaan ensimmäisen asteen autoregressiivinen malli, eli $AR(1)$ -malli

$$y_t = \beta_1 y_{t-1} + \varepsilon_t, \quad (23)$$

jossa ε_t on stokastinen virhetermi, jonka keskiarvo on nolla ja varianssi vakio. Mikäli β_1 on 1, niin kohdataan yksikköjuuriongelma. Yksikköjuuriongelmassa sarjan seuraava arvo on siis yhtä kuin nykyinen arvo ynnättynä satunnaisella tekijällä. Tätä ilmiötä kutsutaan myös nimellä satunnaiskulku (random walk). [18, s. 590]

Yksikköjuuren olemassaolon selvittämiseen käytetään tässä työssä laajennettua Dickeyn ja Fullerin testiä (Augmented Dickey-Fuller Test, ADF), jossa on nollahypoteesina oletus yksikköjuuren olemassaolosta. Testissä käytetään apuna kolmea erimuotoista apuregressiomallia:

$$\text{Vakio + trendi: } \Delta y_t = \beta_0 + \alpha t + \delta Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \gamma_2 \Delta Y_{t-2} + \dots + \gamma_p \Delta Y_{t-p} + \varepsilon_i \quad (24)$$

$$\text{Vakio: } \Delta y_t = \beta_0 + \delta Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \gamma_2 \Delta Y_{t-2} + \dots + \gamma_p \Delta Y_{t-p} + \varepsilon_i \quad (25)$$

$$\text{Ei vakiota: } \Delta y_t = \delta Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \gamma_2 \Delta Y_{t-2} + \dots + \gamma_p \Delta Y_{t-p} + \varepsilon_i \quad (26)$$

Näistä malleista valitaan sopivin mallinvalintakriteerien avulla. Testissä on nollahypoteesina, että valitussa mallissa kerroin δ saa arvon 0. Vastahypoteesina on $\delta < 0$. Merkintä Δy_t kuvaa muutosta aikavälillä $(t-1) - t$. Viiveen suuruus p voidaan valita SBIC:n tai AIC:n avulla. Nollahypoteesin tilastollista merkitsevyyttä testataan laskeamalla t-testin mukainen t-arvo hypoteesille. Koska ADF-testin arvo ei ole normaalijakautunut, täytyy arvoa verrata kriittisiin arvoihin, jotka poikkeavat t-testin vastaavista arvoista [18.] ADF-testin kriittinen arvo on aina negatiivinen, ja kumotakseen nollahypoteesin on testituloksen oltava arvoltaan kriittistä rajaa pienempi. Jos nollahypoteesi voidaan kumota, on sarja stationaarinen. [18, s. 595]

Epästationaarinen sarja voidaan yrittää muuttaa stationaariseksi sarjan differoinnin avulla. Stationaarista sarjaa voidaan merkitä $I(0)$ ja kerran differoitua sarjaa $I(1)$. $I(d)$

kertoo sarjan integroitumisasteen d . Jossakin tapauksessa useamman epästationaarisen sarjan lineaarikombinaatio voi olla stationaarinen, jolloin näitä yhteisintegroituneita sarjoja voidaan käyttää regressiossa. Sarjojen yhteisintegroituneisuutta voidaan testata tekemällä regression virhetermeille AEG-testi (Augmented Engle-Granger). AEG-testissä tehdään mallin jäännöstermeille ADF-testi. Jos virhetermit ovat ADF-testin mukaan stationaarisia, ovat sarjat yhteisintegroituneita ja ne supistuvat kohti pitkän ajan tasapainoa. [18, s. 695]

5.3 Joustot

Selittävien muuttujien vaikutusta selitettävään muuttujaan voidaan tarkastella myös joustojen avulla. Joustojen tarkastelu on kätevää kun halutaan hahmottaa kunkin selittävän muuttujan aiheuttaman muutoksen suunta ja suuruus. Yleisen talousteorian mukainen määritelmä joustolle E on

$$E = \frac{\Delta y/y}{\Delta x/x}, \quad (27)$$

joka kuvaa, kuinka paljon x :n suhteellinen muutos vaikuttaa y :n arvoon. Yhtälön (27) arvoja voidaan tulkita seuraavalla tavalla:

- $E = 0$, x :n muutos ei vaikuta laisinkaan y :n arvoihin
- $E < 1$, y :n suhteellinen muutos on pienempi kuin x :n suhteellinen muutos
- $E = 1$, suhteelliset muutokset ovat yhtä suuret
- $E > 1$, y :n suhteellinen muutos on suurempi kuin x :n suhteellinen muutos

Jousto E voidaan laskea kullekin selittävälle tekijälle myös käyttämällä logaritmuotoista lineaarista mallia

$$\ln y_t = \beta_0 + \beta_{t1} \ln X_1 + \beta_2 \ln X_{t2} + \dots + \beta_k \ln X_{tk} + \varepsilon_t, \quad (28)$$

jonka regressiokertoimille lasketut estimaatit vastaavat kunkin selittävän tekijän jouston arvoa. [18, s.312]

Perinteisen jouston lisäksi on hyödyllistä yrittää laskea joustot niin lyhyelle kuin pitkälle aikavälille. Lyhyellä aikavälillä tarkoitetaan tässä yhteydessä sitä ajanjaksoa, jossa ainakin yhden selittävän tekijän pysyy muuttumattomana. Pitkällä aikavälillä puolestaan tarkoitetaan ajanjaksoa, jossa mallin kaikkien selittävien tekijöiden arvot voivat muuttua. Tässä työssä lyhyenä ajanjaksona voidaan pitää alle vuoden pituista aikaa ja pitkänä ajanjaksona vähintään vuoden pituista ajanjaksoa. Lyhyen ja pitkän aikavälin joustojen laskeminen onnistuu käyttämällä osittaista sopeutumismallia (partial adjustment model), joka olettaa selitettävän muuttujan reagoivan huonosti selittävien muuttu-

jien muutokseen, mutta se reagoi niihin hitaasti supistumalla kohti pitkän tähtäimen tasapainotilaa. Tasapainotilaa kuvataan mallilla

$$\ln y'_t = \beta_0 + \beta_1 \ln X_{t1} + \beta_2 \ln X_{t2} + \dots + \beta_k \ln X_{tk} + \varepsilon_t, \quad (29)$$

jossa y'_t on selitettävän muuttujan tasapainotila, jota kohti kuljetaan yhtälön

$$\ln y_t - \ln y_{t-1} = \delta(\ln y'_t - \ln y_{t-1}) \quad (30)$$

mukaisesti. Kerroin δ kuvaa sopeutumisen nopeutta ja saa arvonsa väliltä 0–1. Mitä lähempänä arvoa 1 kertoimen δ arvo on, sitä nopeampaa on sopeutuminen. Osittainen sopeutumismalli saadaan sijoittamalla yhtälö (25) yhtälöön (26), jolloin osittainen sopeutumismalli saa muodon

$$\begin{aligned} \ln y_t = & \delta\beta_0 + \delta\beta_1 \ln X_{t1} + \delta\beta_2 \ln X_{t2} + \\ & \dots + \delta\beta_k \ln X_{tk} + (1 - \delta)\ln y_{t-1} + \delta\varepsilon_t, \end{aligned} \quad (31)$$

jossa $\delta\beta_1, \delta\beta_2, \dots, \delta\beta_k$ ovat lyhyen ajan sähkökysyntäjoustoja ja $\beta_1, \beta_2, \dots, \beta_k$ ovat pitkän tähtäimen joustoja. Näille kertoimille voidaan laskea PNS-estimaattorit aivan kuten yleisessä lineaarisessa regressiomallissa. [25, s. 404]

5.4 Ennusteen luottamusväli

Ennustemallissa jäännöstermi kuvaa ennustamisessa tapahtunutta virhettä, joka johtuu regressiokertoimien estimaattoreiden virheestä sekä jäännöstermiin liittyvästä epävarmuudesta. Epävarmuus johtuu siitä, että jäännöstermiä ei voida tarkasti mitata kuten mallinnuksessa. Jäännöstermin suuruutta voidaan halutessa kuitenkin arvioida kaavan (32) avulla. [17, s. 364]

$$\pm t_{\alpha/2} s_\varepsilon [1 + \tilde{z}^T (X^T X)^{-1} \tilde{z}]^{1/2}, \quad (32)$$

jossa $\pm t_{\alpha/2} = t(n - k - 1)$ -jakauman luottamustasoon $(1 - \alpha)$ liittyvä kerroin

s_ε = mallin jäännöstermien keskivirhe

\tilde{z} = $(k + 1)$ -vektori $(1, \tilde{x}_1, \tilde{x}_2 \dots \tilde{x}_k)$, jossa $\tilde{x}_1, \tilde{x}_2 \dots \tilde{x}_k$ ovat selittävien muuttujien ennustettuja arvoja.

Kaavalla (32) laskettua arviota jäännöstermin suuruudesta kutsutaan ennustetun arvon luottamusväliksi. Sen mukaan selitettävän muuttujan arvo on $(1 - \alpha)$ % todennäköisyydellä lasketulta väliltä. [17, s.364]

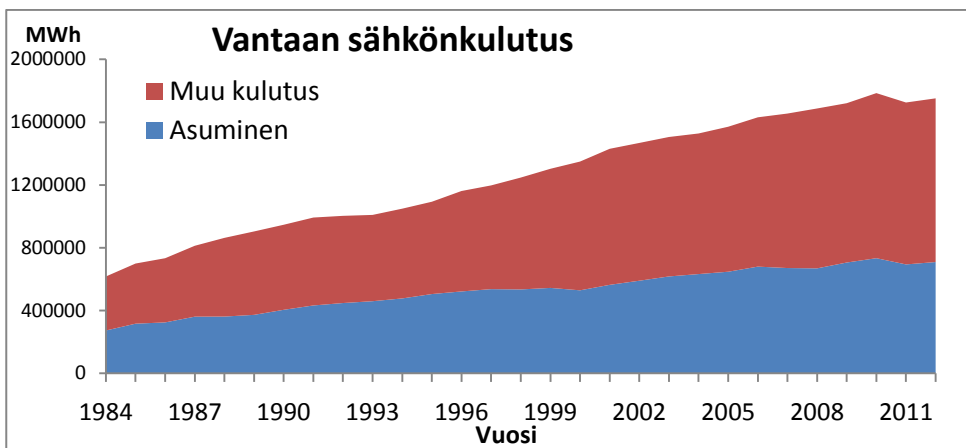
6 AINEISTON ESITTELY

Tässä luvussa tutustutaan selittäviin tekijöihin, joiden avulla sähkönkulutusta pyritään mallintamaan. Joitakin tässä työssä tarkasteltavia muuttujia on perinteisesti käytetty sähkönkulutuksen mallintamiseen. Osa muuttujista koostuu uusista potentiaalisista selittäjäkandidaateista, joiden tuomaa lisäarvoa mallin hyvyyteen ja käytettävyyteen tullaan arvioimaan. Näiden muuttujien joukossa on useita sosioekonomisia muuttujia, kuten bruttokansantuote, kotitalouksien käytettävissä olevat tulot sekä työttömyysaste. Muuttujien arvioinnissa kiinnitetään huomiota erityisesti yksittäisestä muuttujasta saatavan tiedon laatuun, saatavuuteen sekä ennustettavuuteen. Muuttujien käytettävät aikasarjat päättyvät vuoteen 2012, koska vuodelta 2013 saatavan datan laatu ei työn tekovaiheessa ole vielä kaikilta osin varmistettua tai sitä ei ollut saatavissa.

6.1 Sähkönkulutus

Eri alueiden sähkönkulutustiedot toimivat tässä työssä selitettävänä muuttujina. Sähkönkulutustietoja koko Suomen tasolla esiteltiin jo luvussa kaksi (kuva 2.1). Suomen sähkönkulutustilastoja sektoreittain on saatavissa Tilastokeskukselta 1970-luvulta lähtien. Muita Suomen sähkönkulutusdataa julkistavia tahoja ovat ainakin Energiateollisuuden keskusliitto sekä Työ- ja elinkeinoministeriö.

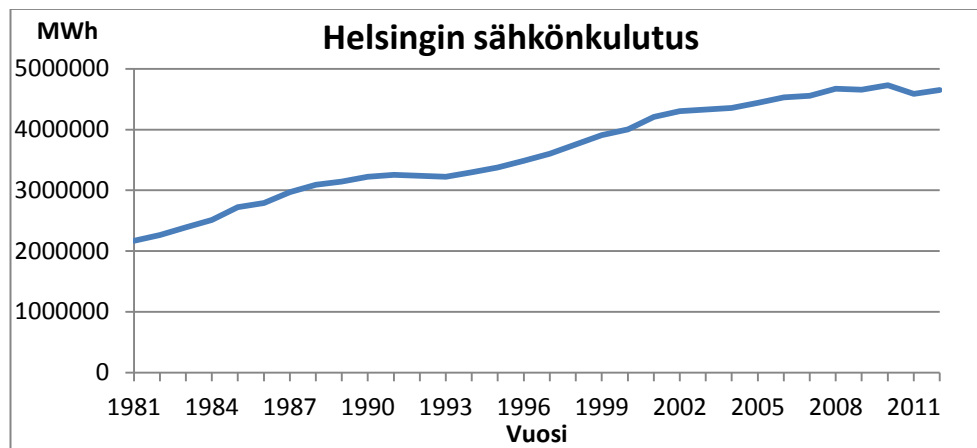
Tässä työssä tullaan käyttämään Vantaan sekä Helsingin alueiden sähkönkulutusdataa. VES on kerännyt kuukausikohtaista tilastoa Vantaan sähkönkulutuksesta vuodesta 1982 lähtien. Käytettävä mittausdata ei sisällä lentokenttäalueen sähkönkulutusta. Kuvassa 6.1 on esitetty sähkönkulutuksen kehittyminen VES:n verkossa aikavälillä 1984–2012.



Kuva 6.1. Vantaan sähkönkulutus aikavälillä 1984–2012

Vantaan sähkönkulutus voidaan tällä hetkellä jakaa kahteen sektoriin; asumisen kulutukseen sekä muuhun kulutukseen, johon sisältyy muun muassa alueen teollisuuden ja palveluiden kuluttama sähkö. Sektorijako on tehty verkkotietojärjestelmän tietojen perusteella. Asuminen kulutti vuonna 2012 noin 40 % koko Vantaalla kulutetusta sähköstä.

Helsingin vuosi- ja kuukausikohtaisia sähkönkulutustietoja on saatu käytettäväksi tutkimusta varten Helen Sähköverkot Oy:ltä. Vuositason aikasarja ulottuu aina 1900-luvun alkuun. Vuosineljänneksen aikasarja alkaa vuodesta 1999. Kuvassa 6.2 on esitetty Helsingin sähkönkulutuksen kehitys aikavälillä 1981–2012.



Kuva 6.2. Helsingin sähkönkulutus aikavälillä 1981–2012

Verrattaessa Vantaan ja Helsingin sähkönkulutuksen kehitystä Suomen sähkönkulutuksen kehitykseen, huomataan, etteivät talouden suuret muutokset ole aiheuttaneet yhtä suurta muutosta alueiden sähkönkulutukseen kuin koko Suomen tapauksessa. Syyinä tähän voidaan olettaa olevan se, että Vantaalla ja Helsingissä ei ole paljon energiaintensiivistä tai suhdanneherkkää teollisuutta verrattuna koko Suomeen.

6.2 Bruttokansantuote

Bruttokansantuote (BKT) on yleisimmin käytetty mittari kuvaamaan kansantalouden kehitystä. Maan bruttokansantuote on kotimaisten tuotantoyksiköiden tuotantotoiminnan lopputulos. Koska sähkö on tuotannontekijänä lähes kaikessa nykypäivän tuotantotoiminnassa, on perusteltua olettaa bruttokansantuotteen olevan hyvä sähkönkulutusta selittävä muuttuja.

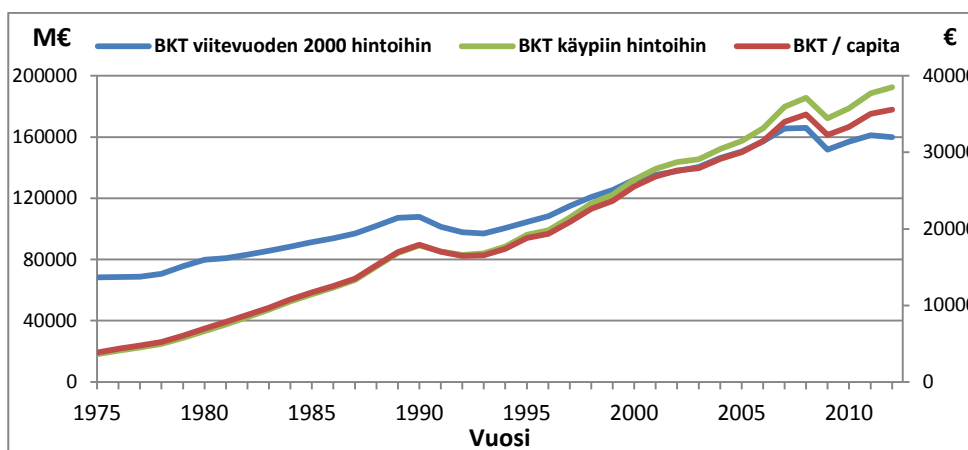
BKT on merkittävä osa kansantalouden tilinpitoa. Aikasarjoja on saatavissa laajasti esimerkiksi tilastokeskukselta. Kansantalouden tilinpidon laskentamenetelmät ovat muuttuneet aikojen saatossa ja vuodesta 1975 alkaen sarjat on laskettu Euroopan kansantalouden tilipitojärjestelmän EKT 95:n mukaisesti. Tässä työssä ei käytetä tämän vanhempia aikasarjoja, joten eri laskentamenetelmistä aiheutuvia ongelmia ei pitäisi ilmentyä. Suomessa bruttokansantuotteen laskennassa käytetään tuotoslähestymistapaa,

joka on yksi monesta eri tavasta laskea BKT. Tuotoslähestymistavassa markkinahintainen BKT saadaan seuraavasti: [26]

$$\begin{aligned} \text{BKT markkinahintaan} &= \text{kansantalouden bruttoarvonlisäys} \\ &+ \text{tuoteverot} \\ &- \text{tuotetukipalkkiot} \end{aligned} \quad (33)$$

Tulevaisuudessa voi tulla käyttöön uusia bruttokansantuotteen laskutapoja. Muutokset on otettava huomioon siten, että eri laskutavoilla muodostettuja sarjoja voitaisiin käsitellä yhtenä sarjana.

Bruttokansantuote voidaan ilmoittaa myös viitevuoden hintaisena sekä alueen väkilukuun suhteutettuna. BKT per capita ilmoittaa käypähintaisen BKT:n suuruuden euroissa suhteessa alueen väkilukuun. Yleisesti sen avulla voidaan vertailla eri maantieteellisten alueiden talouskasvua. Sen potentiaali sähkönkulutuksen mallintamisessa voidaan olettaa olevan yhtä suuri kuin normaalimuotoisen BKT:n sekä väkiluvun. Viitevuoden hintainen BKT ottaa puolestaan huomioon inflaation vaikutuksen, mikä voi tuoda muuttujalle lisää selitysvoimaa. Kuvassa 6.3 on esitetty BKT käypiin hintoihin, vuoden 2000 hintoihin sekä per capita.



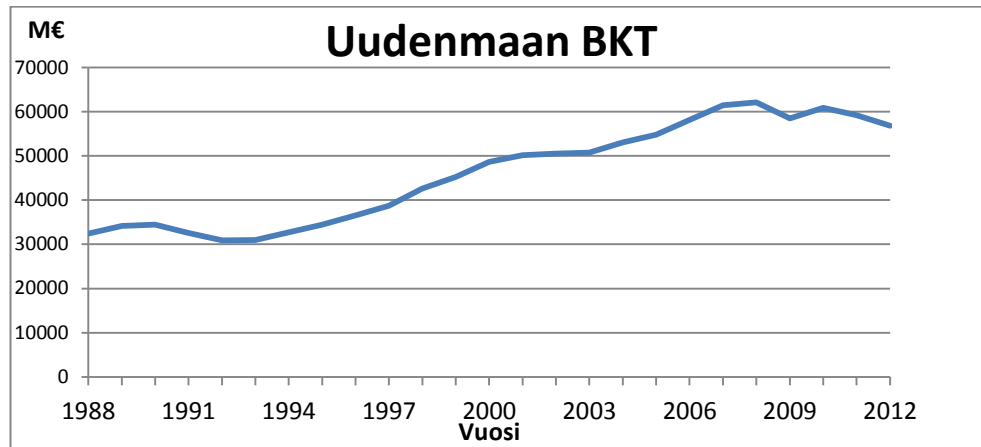
Kuva 6.3. Suomen BKT:n eri ilmoitusmuodot arvoineen aikavälillä 1975–2012 [26]

Kuvasta 6.3 nähdään kuinka 1990-luvun alun lama sekä vuonna 2008 alkanut lama ovat vaikuttaneet BKT:n kehitykseen. BKT:n eri ilmoitusmuodot eivät sarjoina eroa paljoa toisistaan.

Bruttokansantuotteesta on saatavilla aikasarjoja vuodesta 1975 alkaen. Suomen bruttokansantuotteen kehittymistä seurataan aktiivisesti esimerkiksi finanssialalla, joten siitä on saatavissa paljon ennusteita, jotka ulottuvat 1–3 vuoden päähän. Finanssialan keskusliitto julkaisee ajankohtaisia Suomen BKT:n ennusteita kootusti.

Vantaan ja Helsingin sähkönkulutuksen mallintamista varten käytetään Uudenmaan bruttokansantuotetta, josta on saatavilla aikasarja vuodesta 1988 alkaen Elinkeinoelä-

män tutkimuslaitoksen (ETLA) laatimana. Kyseinen aikasarja on esitetty kuvassa 6.4. ETLA on lisäksi laatinut ennusteen Uudenmaan BKT:n kehityksestä vuoteen 2017 asti.



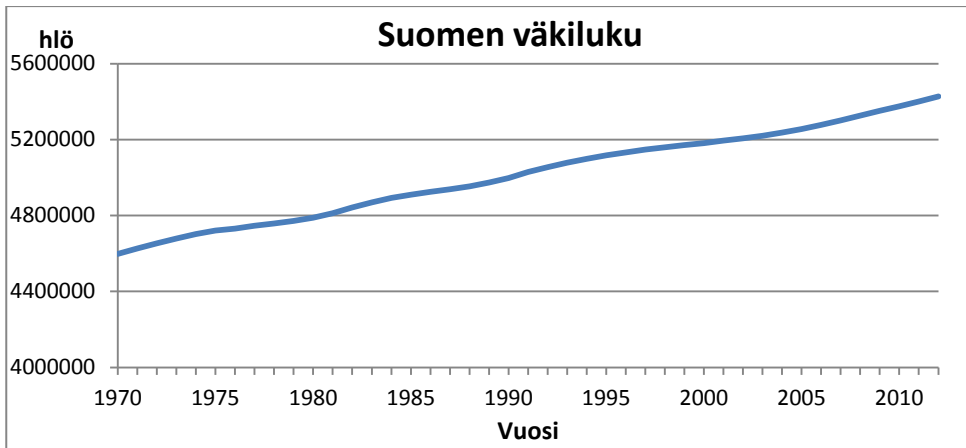
Kuva 6.4. Uudenmaan BKT aikavälillä 1988–2012 [27]

Kuvasta 6.4 nähdään, että Uudenmaan BKT on käyttäytynyt pääsääntöisesti samalla tavalla kuin koko Suomen BKT, mutta keskitettäessä tarkastelu aikasarjan aivan viimeisiin vuosiin huomataan Uudenmaan BKT:n muutosten olleen paljon lievempiä verrattuna Suomen BKT:n muutoksiin. Vantaan ja Helsingin sähkönkulutukselle tullaan lisäksi kokeilemaan selittäviä tekijöitä, joissa Uudenmaan BKT on suhteutettu Vantaan ja Helsingin väkilukuihin. Näin saatuja muuttujia tullaan kutsumaan työssä Helsingin- ja Vantaan BKT per capitaksi.

6.3 Väkiluku

Väkilukua voidaan pitää merkittävänä sähkönkulutusta selittävänä muuttujana. Alueen väestönkasvu lisää luonnollisesti kotitalouksien määrää sekä palveluiden kysyntää, mikä kasvattaa alueen sähkönkulutusta.

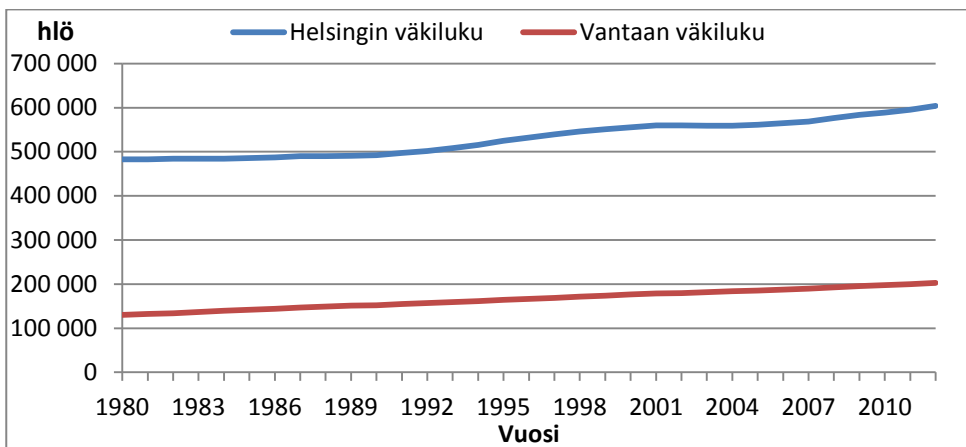
Väestörekisterikeskuksen laatima aikasarja Suomen väestönkehityksestä on saatavissa Tilastokeskukselta [28]. Samassa on julkaistu Suomen väkiluvun kehitys on esitetty kuvassa 6.5.



Kuva 6.5. Suomen väkiluvun kehitys aikavälillä 1970–2012 [28]

Kuvasta 6.5 nähdään, että väkiluvun kasvu on ollut tasaista koko tarkastelujakson ajan. Väestökisterikeskus on laatinut ennusteen Suomen väkiluvun kehitykselle, mikä ulottuu aina vuoteen 2060 asti. [28]

Helsingin ja Vantaan kaupungit julkaisevat väestökisterikeskuksen laatimia aikasarjoja kaupunkien väkiluvuista verkkosivuillaan. Helsingin ja Vantaan väkilukujen kehitys on esitetty kuvassa 6.6.



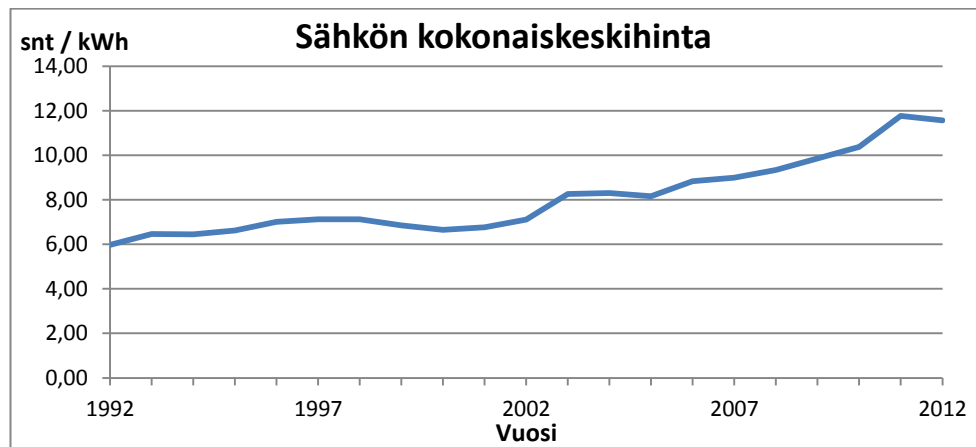
Kuva 6.6. Helsingin ja Vantaan väkilukujen kehitys aikavälillä 1980 – 2012 [29, 30]

Vantaan väkiluku on kehittynyt hyvin tasaista vauhtia. Helsingin väkiluvun kehityksessä on puolestaan havaittavissa pientä epätasaisuutta. Väestökisterikeskus on laatinut ennusteet kaupunkien väkilukujen kehityksestä vuoteen 2040 asti. Ennusteet on saatavilla kaupunkien tilastoista. [29, 30]

6.4 Sähkön keskihinta

Yleisen talousteorian mukaan kaupattavalla hyödykkeellä on kysynnän hintajousto, eli hyödykkeen hinnan muutoksella on vaikutus kyseisen hyödykkeen kysyntään kaavan (23) mukaisesti. Tämän teorian valossa myös sähkön hinnan kehityksellä voisi olla vai-

kutusta sen kokonaiskulutukseen. Kuvassa 6.7 on sähkön kokonaiskeskihinnan kehitys käyvin hinnoin ajanjaksolla 1992–2012.



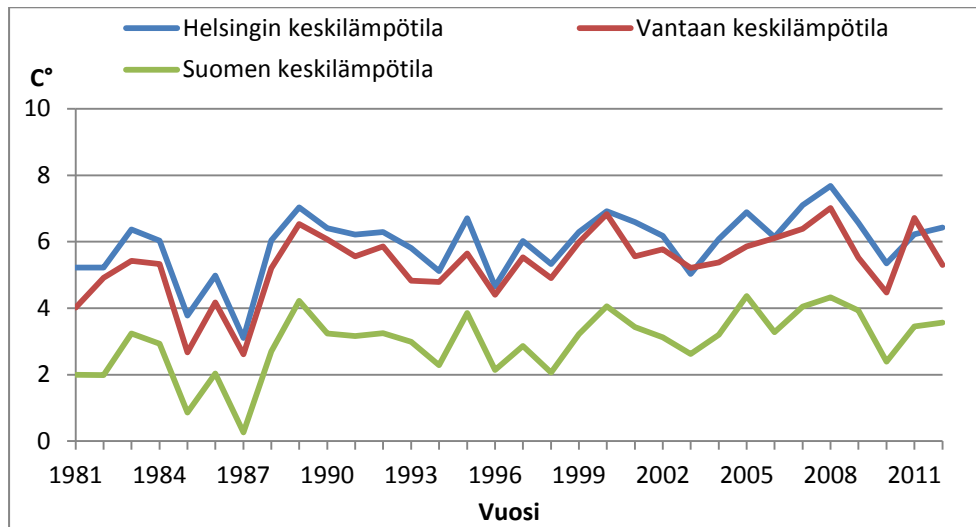
Kuva 6.7. Sähkön keskihinta vuosina 1992–2012 [31]

Sähkön kokonaiskeskihinnan trendi on ollut nouseva viimeiset kaksi vuosikymmentä. On oletettavaa, että sähkön hinnalla voisi olla selitysvoimaa esimerkiksi teollisuuden kuluttaman sähköenergian mallintamisessa, koska erityisesti energiaintensiivinen teollisuus voi jopa säädellä tuotantoaan sähköenergian hinnan mukaan. Kun tarkastellaan sähkön keskihintaa vuositasolla, vaikuttaa hintajoustoon lähinnä vain investointipäätökset. Kotitalouksien sähkönkulutukseen voi sähkön hinnalla olettaa olevan vaikutusta, kun lämmitysmuotojen hintaerot vaikuttavat lämmitysmuodon valintaan. Muuhun kuluuskäyttäytymiseen ei sähkön hinnalla uskota olevan vaikutusta.

Sähkön hinta muodostuu sähköntuotannon kustannuksista, sähkön siirrosta, myynnistä ja veroista. Verojen osuus sähkön kokonaishinnasta on peräti kolmasosa. [32] Sähkön hinnan trendi on ollut nouseva aina tilastoinnin alusta, eli vuodesta 1992 alkaen. Kuvan 6.7 aikasarjan arvot ilmoittavat sähkön kokonaishinnan kaikkien kuluttajatyypien keskiarvona ja se sisältää siirron ja myynnin osuudet sekä kaikki ajankohtana voimassa olleet verot [31]. Ennusteita sähkön hinnan kehityksestä ei ole helposti saatavissa. Veropolitiikan lisäksi sähkön hinta on pitkällä aikavälillä riippuvainen muun muassa energiapolittisista päätöksistä, mikä tekee sähköenergian hinnan pitkän ajan ennustamisesta vaikeaa.

6.5 Lämpötila

Lämpötilalla on kiistaton vaikutus sähkönkulutukseen Suomen olosuhteissa, joissa lämmitykseen käytetty sähköenergia on noin 60 % kotitalouksien kokonaiskulutuksesta (kuva 2.2). Koko Suomen sähkönkulutusta mallinnettaessa on haasteena vuoden keskilämpötilan määrittäminen niinkin maantieteellisesti laajalle alueelle kuin Suomi on. Kuvassa 6.8 on esitetty Helsingin ja Vantaan lisäksi Suomen keskilämpötila, joka on laskettu kolmen eri puolilla Suomea sijaitsevien havaintoasemien keskiarvoista. Mittausasemat sijaitsevat Helsingissä, Jyväskylässä ja Sodankylässä.



Kuva 6.8. Suomen, Vantaan ja Helsingin vuoden keskilämpötilat aikavälillä 1970–2012 [33]

Vantaan tapauksessa tullaan käyttämään Helsinki-Vantaan lentokentän mittausaseman aikasarjaa ja Helsingin tapauksessa Helsingin Kaisaniemessä sijaitsevan mittausaseman aikasarjaa.

Lämpötila voi yksittäisinä vuosina aiheuttaa suuria sähkönkulutuksen vuosimuutoksia, mutta silti lämpötila ei juuri vaikuta verkon mitoittamiseen. Pitkän aikavälin suunnittelussa verkko mitoitetaan kylmimmän lämpötilan mukaan, jolloin kulutushuiput Suomen olosuhteissa tapahtuvat. Tulevaisuudessa myös erityisen lämpimät olosuhteet voivat ainakin alueellisesti aiheuttaa kulutushuippuja, erityisesti jos huomioidaan korkean lämpötilan alentava vaikutus verkkokomponenttien kuormittavuuteen.

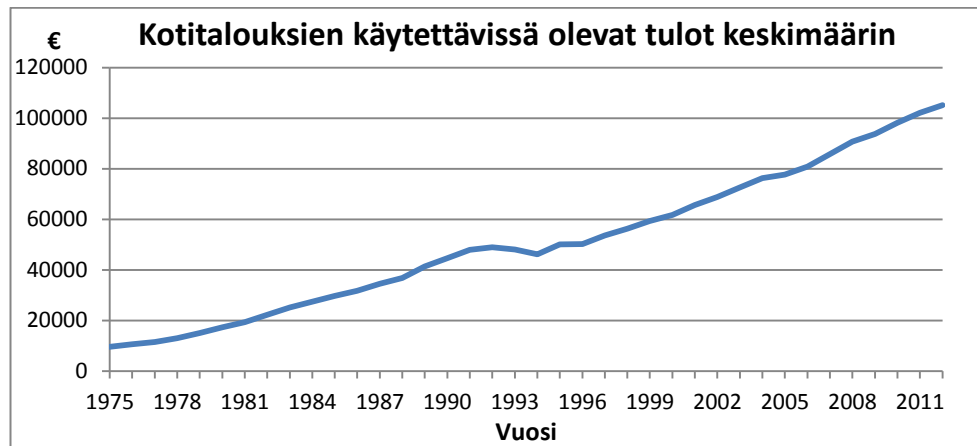
Lämpötilaa on hyvin vaikea ennustaa pitkällä aikavälillä, mutta kehitystä voidaan arvioida Ilmatieteenlaitoksen julkaisemien ilmastoskenaarioiden avulla. Lämpötilan kehityksen aiheuttamaan epävarmuuteen palataan luvussa 10, jossa ilmastoskenaarioihin tutustutaan hieman syvällisemmin.

6.6 Kotitalouksien käytettävissä olevat tulot

Kotitalouksien käytettävissä olevat tulot kertovat sen rahamäärän, joka ihmisille jää verojen jälkeen käytettäväksi. Tämän rahamäärän kotitaloudet voivat käyttää muun muassa hyödykkeisiin, jotka kuluttavat sähköä.

Kotitaloudella tarkoitetaan tässä yhteydessä niitä henkilöitä, jotka asuvat ja ruokailvat yhdessä tai jotka muuten käyttävät tulojaan yhdessä. Tilastointi on muodostettu hallinnollisten rekisterien sekä haastatteluiden avulla. Muuttujasta on tilastoitua tietoa vuodesta 1975 lähtien saatavissa Tilastokeskukselta [34].

Kuvassa 6.9 on esitetty kotitalouksien käytettävissä olevien keskimääräisten tulojen aikasarja välillä 1975–2012.

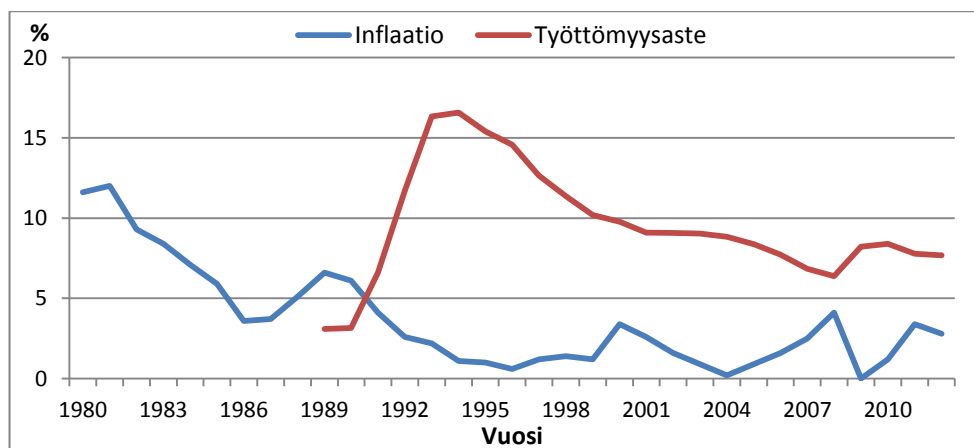


Kuva 6.9. Kotitalouksien käytettävissä olevat tulot keskimäärin aikavälillä 1975–2012 [34]

Käytettävissä olevat tulot kuvaavat elintason kehitystä ja korreloivat vahvasti bruttokansantuotteen kanssa. On oletettavaa, että tulojen lasku ei aiheuta laskua sähkönkulutuksessa, vaan se pikemminkin vain tulisi hillitsemään kulutuksen kasvua.

6.7 Inflaatio ja työttömyysaste

Inflaatiolla tarkoitetaan kuluttajahintaindeksin muutosta, joka kuvaa rahan arvon heikkenemistä. Kansantaloustieteestä tutun teorian mukaan inflaatiolla ja työttömyydellä voidaan katsoa olevan negatiivinen korrelaatio jollakin aikavälillä ja inflaation ja työttömyysasteen muutoksilla on myös yhteys suhdanteisiin. Kuvassa 6.10 on esitetty muutujien kuvaajat aikavälillä 1980–2012.



Kuva 6.10. Inflaatio ja työttömyysaste Suomessa aikavälillä 1980–2012 [35, 36]

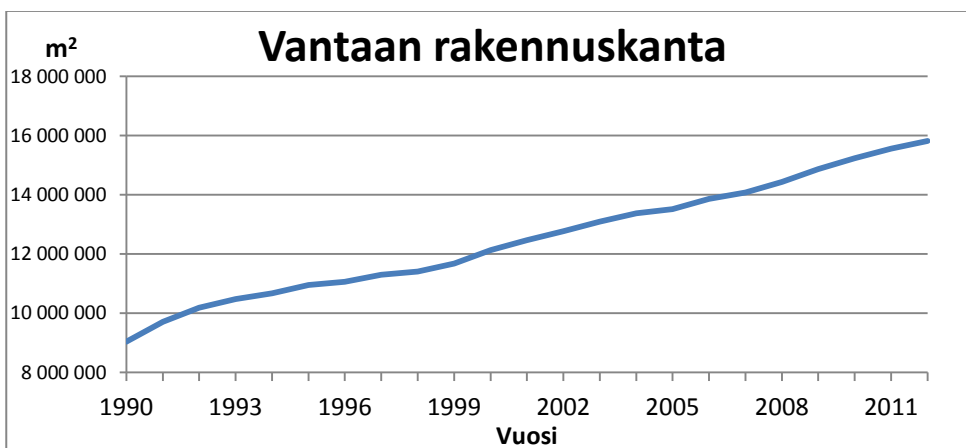
Vantaan ja Helsingin malleihin tullaan sovittamaan näiden alueiden työttömyysasteita. Monet tutkimuslaitokset ja pankit laativat ennusteita Inflaation ja työttömyysasteen kehityksestä. Ennusteet ulottuvat pääsääntöisesti kahden vuoden päähän.

Korkealla työttömyysasteella voi olla kasvattava vaikutus kotitalouksien sähkönkulutukseen, jos ihmisten kotona vietetty aika lisääntyy. Toisaalta korkea työttömyysaste tai korkea inflaatio heikentävät kotitalouksien ostovoimaa, millä saattaa olla negatiivinen vaikutus sähkönkulutukseen. Näitä yhteyksiä tullaan työssä tarkastelemaan.

6.8 Rakennuskanta ja toimitilojen käyttöaste

Sähkönkulutus keskittyy vahvasti rakennuksiin ja niiden lähiympäristöön. Rakennuskanta on tapana ilmoittaa joko rakennusten lukumääränä tai rakennusten kerrosalana. Sähkönkulutuksen mallintamisen kannalta rakennuskantaa on syytä käsitellä kerrosalana, koska sillä on suora vaikutus yksittäisen rakennuksen lämmitystarpeeseen. Pinta-alan lisäksi rakennusten tilavuudella on vaikutusta lämmitystarpeeseen. Vantaan rakennusten tilavuustiedot on kuitenkin todettu osittain puutteellisiksi, mistä johtuen tilavuustietojen sopivuutta kuvaamaan sähkönkulutusta ei lähdetä kokeilemaan. Lisäksi rakennusten ikä vaikuttaa merkittävästi kulutettuun lämmitysenergiaan. Energiatohokkuuden kasvaminen on yksi asia, jota tarkastellaan tulevaisuusskenaarioissa luvussa 11.

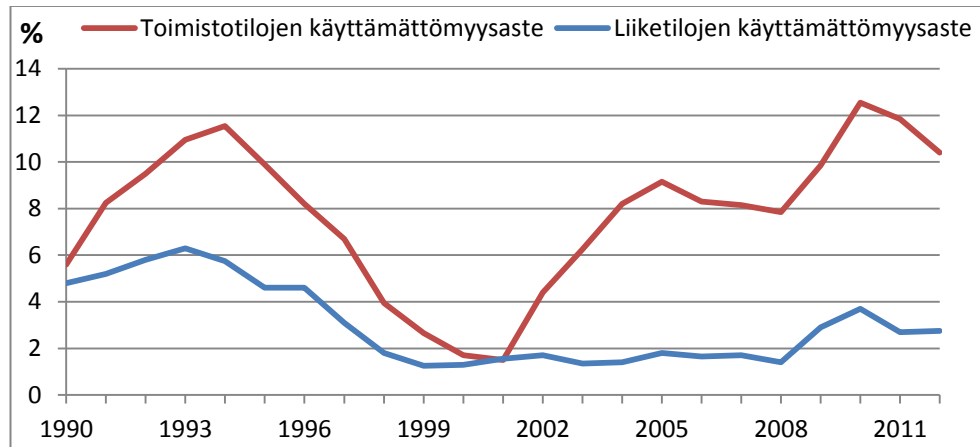
Vantaan rakennuskannasta on saatavissa aikasarja vuodesta 1990 lähtien ja se on esitetty kuvassa 6.11. Helsingin rakennuskannasta yleisesti saatavilla oleva aikasarja alkaa vasta vuodesta 2002, jolloin aikasarjan pituus ei ole riittävä mallin muodostamista varten. Suomen rakennuskantaa ei ole kokonaisuudessaan määritelty kerrosalana. [30]



Kuva 6.11. Vantaan rakennuskanta aikavälillä 1990–2012 [30]

Rakennuskanta kasvaa perustellusti samaa tahtia väkiluvun kanssa. Kunnat ja kaupungit tekevät ennusteita rakennuskannan kasvusta kaavoituksen perusteella. Vantaan kaupungin vuonna 2013 laatima ennuste ulottuu vuoteen 2025 asti.

Myös pääkaupunkiseudun liike- ja toimistotilojen käyttämättömyysasteiden sopivuutta selittäviksi tekijöiksi tullaan kokeilemaan Vantaan sekä Helsingin sähkönkulutusta kuvaaviin malleihin. Pääkaupunkiseudun liike- ja toimistotilojen käyttämättömyysasteiden aikasarjat on esitetty kuvassa 6.12.



Kuva 6.12. Liike- ja toimistotilojen käyttämättömyysaste aikavälillä 1990–2012 [37]

Varsinkin Helsingissä on runsaasti toimitiloja, joten niiden käyttöaste voi vaikuttaa näkyvästi sähkönkulutukseen. Selittävänä tekijänä voidaan kokeilla myös toimitilojen käyttämättömyysastetta, joka on muodostettu laskemalla liike- ja toimistotilojen käyttämättömyysasteiden yhteinen keskiarvo.

6.9 Käytettävien aikasarjojen itsenäisyys

Kun muuttujia sovitetaan selittäviksi tekijöiksi sähkönkulutukselle, saattaa tiettyjen muuttujien sovittaminen samaan malliin tuoda esiin multikollineaarisuusongelman. Tätä asiaa lähdettiin selvittämään jo tässä vaiheessa ennen varsinaisten mallien muodostamista, jotta välttyttäisiin turhalta työltä.

Taulukossa 6.1 on mahdollisten selittäjien korrelaatiomatriisi, jonka avulla voidaan alustavasti tunnistaa ne muuttujat, joiden kohdalla multikollineaarisuus saattaisi nousta ongelmaksi. Arvot on laskettu muuttujien esitetyille aikasarjoille.

Taulukko 6.1. *Multikollineaarisuutta aiheuttavien muuttujien korrelaatiomatriisi Suomen mallin tapauksessa*

	Suomen sähkön- kulutus	Väkilu- ku	BKT 2000 hintoihin	BKT käypiin hintoihin	BKT per capita	Kotitalo- uksien tulot
Suomen kulutus	1,000	0,961	0,956	0,993	0,993	0,987
Väkiluku		1,000	0,972	0,973	0,969	0,976
BKT 2000 hintoi- hin			1,000	0,972	0,970	0,964
BKT käypiin hintoihin				1,000	1,000	0,997
BKT per capita					1,000	0,997
Kotitalouksien tulot						1,000

Matriisin ensimmäisestä sarakkeesta voidaan lukea, kuinka paljon Suomen sähkönkulutus korreloi mahdollisten selittäjien kanssa. Korrelaatioiden perusteella voisi alustavasti olettaa, että BKT -sarjat, väkiluku sekä kotitalouksien tulot sopisivat hyvin selittämään sähkönkulutusta. Matriisista nähdään myös, että kyseiset sarjat korreloivat merkittävästi keskenään. Korrelaation ollessa näiden muuttujien kohdalla hyvin lähellä arvoa 1, päätetään olla sovittamatta näitä muuttujia samaan malliin keskenään. Näin välteään multikollineaarisuusongelma, joka hyvin todennäköisesti nousisi esille mallin diagnostisissa testeissä.

Myös Vantaan ja Helsingin malleihin käytettäville muuttujille tehtiin vastaava tarkastelu, jonka mukaan taulukossa 6.1 esitetyt muuttujat aiheuttaisivat yhdessä multikollineaarisuusongelman hyvin todennäköisesti myös näissä malleissa. Korrelaatiomatriisit on esitetty kokonaisuudessaan liitteessä 1.

7 MALLIEN MUODOSTAMINEN

Tässä luvussa pyritään muodostamaan regressiomallit Suomen, Vantaan ja Helsingin sähkönkulutuksille. Regressiomallit muodostetaan luvussa 4 esitettyjä mallinvalintates-tejä ja -kriteereitä apuna käyttäen. Muodostetuille malleille tehdään diagnostiset testit, joiden tulosten avulla arvioidaan mallien tilastollista pätevyyttä.

7.1 Malli Suomen sähkönkulutukselle

Sopivimmat lineaariset regressiomallit Suomen sähkönkulutukselle on esitetty taulukossa 7.1. Taulukossa on kunkin mallin korjattu selitysaste sekä SBIC-arvo. Vertailtaessa mallien SBIC-arvoja on niissä käytettävä samaa havaintojen määrää. Tässä tapauksessa havaintojen määrää rajoittaa sähkön hinnan aikasarja, jossa on vain 21 havaintoarvoa.

Taulukko 7.1. Parhaimmat mallit Suomen sähkönkulutukselle

Malli	Selittävät tekijät	Korjattu R^2	SBIC
S1	BKT per capita, Sähkön hinta, Lämpötila	0,948	15,46
S2	Väkiluku, Inflaatio, Työttömyys %	0,815	17,06
S3	BKT vuoden 2000 hintaisena, Sähkön hinta, Lämpötila	0,969	14,95

Taulukon arvojen perusteella voidaan valita parhaaksi malliksi Suomen sähkönkulutukselle malli S3, jolla on suurin korjattu selitysaste sekä pienin SBIC -arvo. Malli on muotoa

$$E_{Suomi,t} = \beta_0 + \beta_1 B_t + \beta_2 S_t + \beta_3 T_t + \varepsilon_t,$$

jossa $E_{Suomi,t}$ = Suomen vuosittainen sähkönkulutus [GWh]
 B_t = Suomen bruttokansantuote vuoden 2000 hintaisena [M€]
 S_t = Sähkön keskihinta [snt/kWh]
 T_t = Vuoden keskilämpötila [C°]

7.2 Diagnostiset testit mallille S3

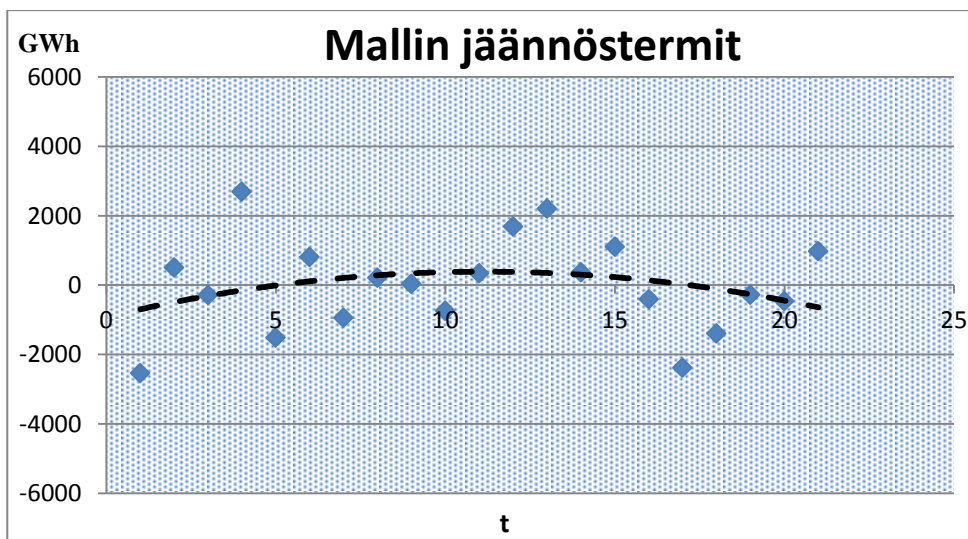
Mallin tilastollista pätevyyttä tarkasteltiin diagnostisten testien avulla. Testien tuloksia on esitetty taulukossa 7.2.

Taulukko 7.2. Diagnostisten testien tulokset mallille S3

W-, D- ja White-testit			LM-testi			VIF-arvot	
Testi	Arvo	p-arvo	p	Arvo	p-arvo	Muuttuja	VIF-arvo
W-testi	0,985	0,977	1	1,01	0,68	B_t	4,11
White-testi	11,59	0,76	2	2,13	0,65	S_t	3,41
D-testi	1,78	-	3	4,84	0,82	T_t	1,40

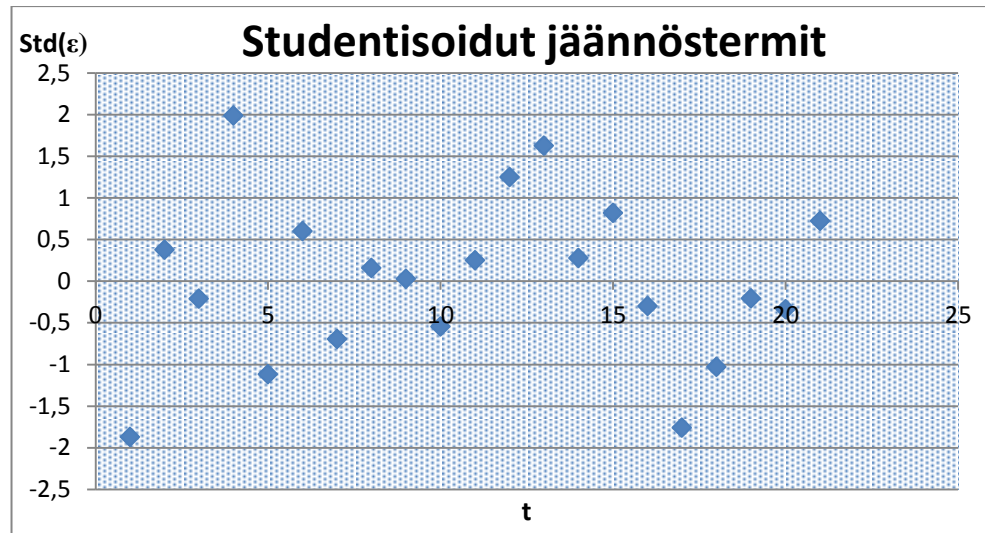
Selittävien tekijöiden VIF-arvot pysyvät kaikki alle arvon 5, joten voidaan todeta, ettei muuttujien välillä ilmene haitallisella tasolla olevaa multikollinearisuutta. D-testin arvo ylittää kriittisen arvon 1,68 ja LM-testin kaikki p-arvot ovat riittävän suuria, joten jäännöstermeissä ei todeta ilmenevän haitallista autokorrelaatiota.

Kuvassa 7.1 on esitetty mallin absoluuttiset jäännöstermit ja kuvassa 7.2 samat jäännöstermit studentisoituina.



Kuva 7.1. Mallin S3 jäännöstermit

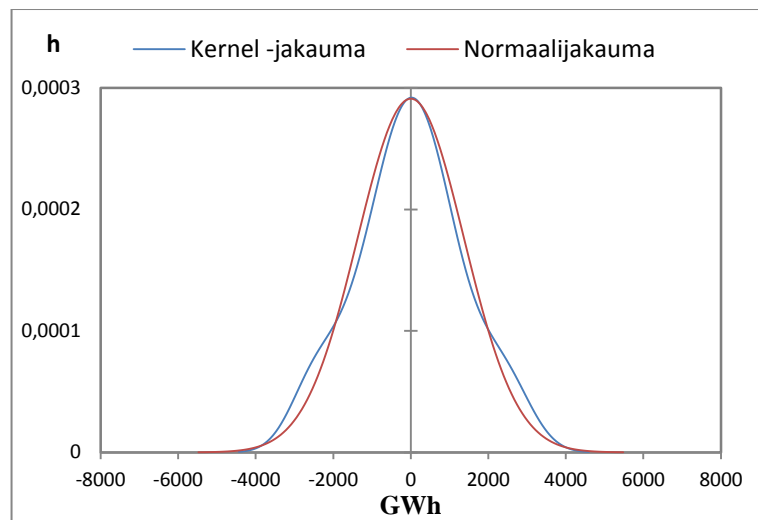
Kuvaan 7.1 piirretty toisen asteen käyrän avulla voidaan silmämääräisesti tarkastella jäännöstermien odotusarvoa, joka on lähellä nollaa käyrän kulkiessa lähellä nolla-akselia.



Kuva 7.2. Mallin S3 studentisoidut jäännöstermit

Kuvasta 7.2 nähdään, että studentisoidujen jäännöstermien itseisarvot pysyvät alle kriittisen arvon 2,08, minkä perusteella ei voida todeta joukossa olevan poikkeavia havain-toja.

Seuraavaksi tarkasteltiin jäännöstermijoukon jakaumaa. Kuvaan 7.3 on piirretty jäännöstermijoukon Kernel-jakauman sekä normaalijakauman tiheysfunktioiden kuvaajat.



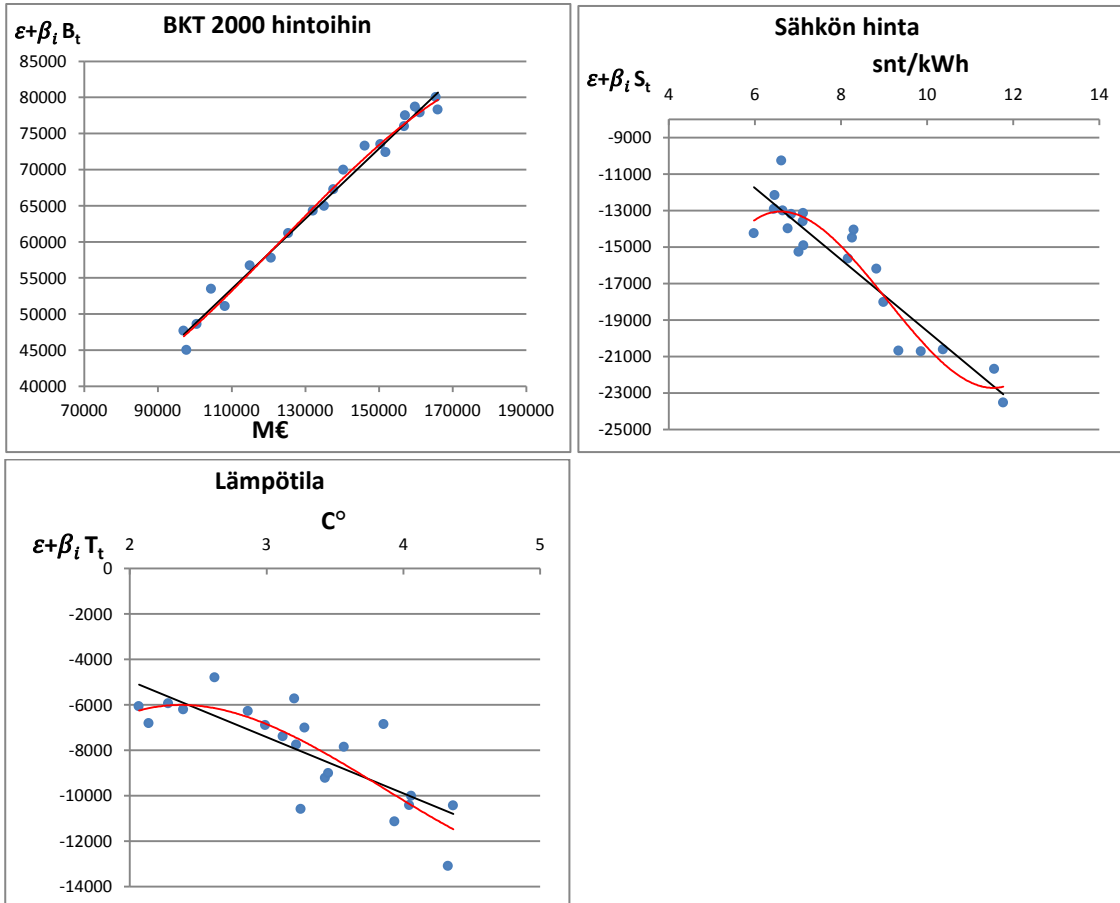
Kuva 7.3. Mallin S1 jäännöstermien Kernel-tiheysfunktio

Kuvasta 7.3 nähdään, että jäännöstermijoukon tiheysfunktio on melko hyvin normaali-jakautunut. Normaalijakautuneisuutta testaavan W-testin p-arvoksi saatiin 0,977. Myös se osoittaa, ettei nollahypoteesia jäännöstermien normaalisuudesta tarvitse hylätä.

Jäännöstermien skedastisuutta voidaan silmämääräisesti arvioida kuvasta 7.1, jossa jäännöstermijoukon voidaan katsoa olevan sen verran tasapaksu, ettei heteroskedastisuutta ole syytä epäillä. Oletus jäännöstermien homoskedastisuudesta testattiin lisäksi

White-testillä, jonka arvoksi saadaan 11,59. Testitulosta vastaava p-arvo on 0,76, jonka mukaan hypoteesi homoskedastisuudesta jää voimaan.

Lineaarisuusoletuksen täytyminen varmistetaan piirtämällä kuvaajat, joissa x-akseleiden arvoina ovat kunkin selittävän muuttujan havaintoarvot ja y-akselilla osittaiset jäännöstermit, jotka ovat muotoa $\varepsilon + \beta_i X_t$. Osittaiset jäännöstermit ovat kuvassa 7.4.



Kuva 7.4. Mallin S3 osittaiset jäännöstermit

Osittaisten jäännöstermien pistejoukkoon on piirretty 1.- ja 3.-asteen sovitteet, jotka kunkin selittävän tekijän kohdalla kulkevat hyvin lähellä toisiaan. Kuvaajien perusteella voidaan lineaarinen riippuvuus selittävän muuttujan ja parametrien välillä todeta todeksi.

Malli läpäisi kaikki diagnostiset testit, minkä mukaan PNS-menetelmällä laskettaviin kertoimien arvoihin voidaan luottaa. Lisäksi täytyy varmentaa mallissa käytettävien aikasarjojen stationaarisuus, mikä suoritetaan ADF-testin avulla.

ADF-testin tulokset on esitetty taulukossa 7.3.

Taulukko 7.3. ADF-testin tulokset mallin S3 sarjoille

	Alkuperäinen sarja			1. differenssi		
	Testiarvo	Parametrit	Kriittinen arvo	Testiarvo	Parametrit	Kriittinen arvo
$E_{Suomi,t}$	0,597	Ei vakiota	-1,952	-6,146	Vakio + trendi	-3,677
B_t	2,568	Ei vakiota	-1,952	-3,998	Vakio + trendi	-3,677
S_t	3,021	Ei vakiota	-1,952	-4,609	Vakio + trendi	-3,677
T_t	-4,444	Vakio + trendi	-3,657	-6,712	Vakio + trendi	-3,677

Taulukosta 7.3 nähdään, että suoritettaessa ADF-testi S3-mallin selittävien tekijöiden aikasarjoille, oli testin arvo riittävän pieni ainoastaan lämpötilan aikasarjalla. Kun sarjat muunnettiin 1. differenssin aikasarjoiksi, olivat kaikkien sarjojen ADF-testiarvot alle kriittisen rajan. Yhteisintegroituvuuden varmistamiseksi suoritettiin mallin virhetermeille ADF-testi, jonka tulokseksi saatiin -4,64, jota vastaava AEG-testin mukainen p-arvo on 0,0514. P-arvo ylittää niukasti kriittisen arvon 0,05. Tämän perusteella mallin sarjojen voidaan olettaa olevan riittävässä määrin yhteisintegroituvia ja tulosten myös tältä osin luotettavia.

Mallille laskettiin kertoimet PNS-menetemällä. Kertoimet sekä mallia muuten kuvaavat ominaisarvot on esitetty taulukossa 7.4.

Taulukko 7.4. Mallin S3 kertoimet ja ominaisarvot

	Arvo	Keskivirhe	t-arvo	p-arvo
θ_0	37956,3	2057,7	18,446	$1,12 \cdot 10^{-12}$
θ_1	0,486	0,028	17,271	$3,24 \cdot 10^{-12}$
θ_2	-1960,50	355,33	-5,517	$3,78 \cdot 10^{-5}$
θ_3	-2475,12	560,09	-4,419	0,00038
R^2	0,973			
Korjattu R^2	0,969			
Mallin F-arvo	208			$1,36 \cdot 10^{-13}$

Taulukon arvoja luettaessa on hyvä kiinnittää huomio kertoimien p-arvoihin sekä mallin selitysasteeseen ja F-arvoa vastaavaan p-arvoon, jonka avulla arvioidaan kertoimien yhteistä selitysvoimaa. Kertoimien p-arvot sekä F-arvoa vastaava p-arvo ovat reilusti alle kriittisen arvon 0,05, mikä tarkoittaa että mallille lasketut kertoimet ovat tilastollisesti päteviä ja mallilla on korkea selitysvoima. Mallin selitysasteella on myös korkea arvo 0,973.

Seuraavaksi muodostettiin log-lineaarinen malli sekä osittainen sopeutumismalli Suomen sähkönkulutusta kuvaavan mallin pohjalta.

Malleille lasketut kertoimet tunnuslukuineen on esitetty taulukoissa 7.5.

Taulukko 7.5. Suomen sähkönkulutuksen log-lineaarisen ja osittaisen sopeutumismallin kertoimet ja ominaisarvot

Log-lineaarinen malli					Osittainen sopeutumismalli				
	Arvo	Keskivirhe	t-arvo	p-arvo		Arvo	Keskivirhe	t-arvo	p-arvo
β_0	2,510	0,483	5,20	$7,2 \cdot 10^{-5}$	β_0	1,995	0,776	2,569	0,021
β_1	0,784	0,048	16,45	$7,1 \cdot 10^{-12}$	β_1	0,669	0,143	4,674	0,00025
β_2	-0,184	0,040	-4,56	0,0003	β_2	-0,170	0,044	-3,854	0,0014
β_3	-0,084	0,022	-3,85	0,0013	β_3	-0,087	0,022	-3,904	0,0013
R^2	0,974				$(1-\delta)$	0,164	0,192	0,851	0,407
Korj. R^2	0,969				R^2	0,975			
F-arvo	212,8			$1,1 \cdot 10^{-13}$	Korj. R^2	0,969			
					F-arvo	157,2			$1,27 \cdot 10^{-12}$

Log-lineaarisen mallin kaikki p-arvot ovat alle kriittisen arvon, eli mallille lasketut kertoimet ovat päteviä ja mallilla on kokonaisuudessaan korkea selitysvaiva. Osittaisessa sopeutumismallissa kertoimelle $(1 - \delta)$ tehdyn t-testin p-arvo on selvästi yli kriittisen arvon. Tästä johtuen osittaisen sopeutumismallin kertoimiin liittyy epävarmuutta, mutta ei ole ehdoton este mallin hyödyntämiselle, sillä muiden kertoimien p-arvot ovat riittävästi pieniä kuin myös mallin F-arvoa vastaava p-arvo.

Malleille tehtiin diagnostiset testit, joiden tulokset on esitetty taulukoissa 7.6 ja 7.7.

Taulukko 7.6. Diagnostisten testien tulokset Suomen sähkönkulutuksen log-lineaariselle mallille

W-, D- ja White-testit			LM-testi			VIF-arvot	
Testi	Arvo	p-arvo	p	Arvo	p-arvo	Muuttuja	VIF-arvo
W-testi	0,987	0,988	1	1,36	0,76	$\ln(B_t)$	4,27
White-testi	12,96	0,84	2	1,68	0,57	$\ln(S_t)$	3,64
D-testi	1,81	-	3	5,59	0,87	$\ln(T_t)$	1,34

Taulukko 7.7. Diagnostisten testien tulokset Suomen sähkönkulutuksen osittaiselle sopeutumismallille

W-, D- ja White-testit			LM-testi			VIF-arvot	
Testi	Arvo	P-arvo	p	Arvo	p-arvo	Muuttuja	VIF-arvo
W-testi	0,973	0,794	1	1,92	0,83	$\ln(B_t)$	37,9
White-testi	15,61	0,66	2	2,11	0,65	$\ln(S_t)$	4,26
D-testi	2,13	-	3	4,95	0,82	$\ln(T_t)$	1,38
						$\ln(E_{Suomi\ t-1})$	28,61

Mallien sarjojen stationaarisuutta testattiin ADF-testillä, jonka tulokset on esitetty taulukossa 7.8.

Taulukko 7.8. ADF-testin tulokset Suomen sähkönkulutuksen log-lineaarisen mallin ja osittaisen sopeutumismallin sarjoille

	Alkuperäinen sarja			1. differenssi		
	Testiarvo	Parametrit	Kriittinen arvo	Testiarvo	Parametrit	Kriittinen arvo
$\ln(E_{Suomi,t})$	-2,324	Ei vakiota	-3,022	-6,197	Vakio + trendi	-3,677
$\ln(B_t)$	-1,624	Vakio	-3,022	-4,191	Vakio + trendi	-3,677
$\ln(S_t)$	-1,44	Vakio + trendi	-3,657	-4,274	Vakio + trendi	-3,677
$\ln(T_t)$	-4,664	Vakio + trendi	-3,657	-7,198	Vakio + trendi	-3,677

Osittaisen sopeutumismallin kohdalla bruttokansantuotteen ja viivästetyn kulutuksen VIF-arvot nousevat hyvin suuriksi, joten mallin kertoimiin liittyy huomattavaa epävarmuutta.

Muiden testien kohdalla ei ilmennyt ongelmia. Kaikki sarjat ovat stationaarisia 1. differenssissään. AEG-testien p-arvoiksi saatiin log-lineaarisen mallin osalta 0,031 ja osittaisen sopeutumismallin osalta 0,027. Arvojen perusteella sarjat kummassakin mallissa ovat yhteisintegroituvia.

7.3 Malli Vantaan sähkönkulutukselle

Vantaan sähkönkulutusta parhaiten kuvaavaa mallia lähdettiin muodostamaan aikaisemman tutkimuksen pohjalta, jossa sähkönkulutusta mallinnettiin BKT per capita, Vantaan rakennuskannan sekä lämpötilan avulla. Eräitä sopivimpia malleja on listattu taulukkoon 7.9.

Taulukko 7.9. Parhaimmat mallit Vantaan sähkönkulutukselle

Malli	Selittävät tekijät	Korjattu R^2	SBIC
V1	BKT per capita, Vantaan rakennuskanta, Lämpötila	0,998	19,47
V2	BKT per capita, Sähkön hinta, Lämpötila	0,986	21,08
V3	Uudenmaan BKT, Sähkön hinta, Lämpötila	0,991	20,58
V4	Väkiluku	0,972	21,57
V5	Rakennuskanta	0,958	22,00

Malleja voidaan jälleen verrata korjatun selityksasteen ja SBIC:n avulla. Taulukon 7.9 kaikki arvot on laskettu 21 havaintoarvolla, jolloin ne ovat vertailukelpoisia. Parhaimmaksi nousee malli V1, joka on muotoa

$$E_{Vantaa,t} = \beta_0 + \beta_1 C_t + \beta_2 R_t + \beta_3 T_{Vantaa,t} + \varepsilon_t,$$

jossa $E_{Vantaa,t}$ = Vantaan vuosittainen sähkönkulutus [MWh]
 C_t = Vantaan bruttokansantuote per capita [M€]
 R_t = Vantaan rakennuskanta [m²]
 $T_{Vantaa,t}$ = Vuoden keskilämpötila Vantaalla [C°]

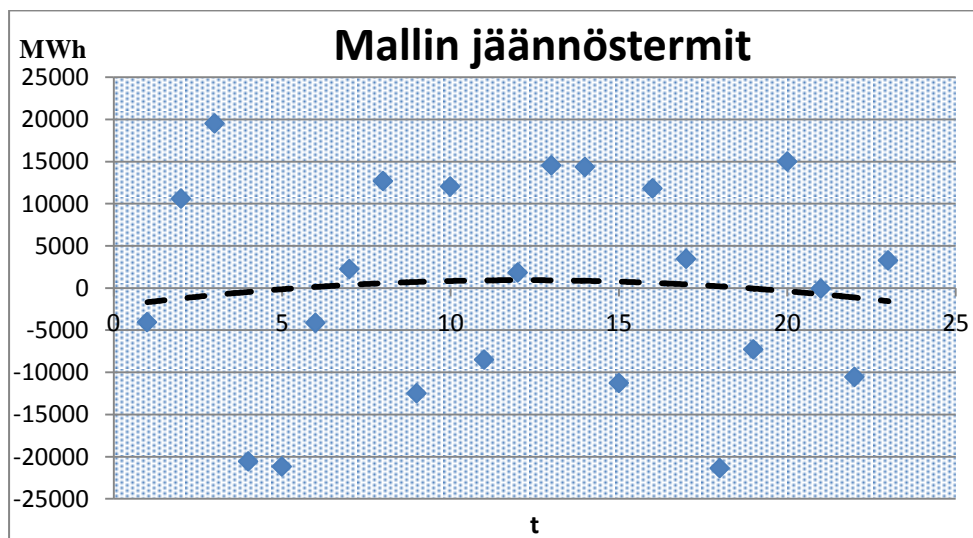
Tälle mallille suoritettiin diagnostiset testit, joiden tuloksia on esitetty taulukossa 7.10.

Taulukko 7.10. Diagnostisten testien tulokset mallille VI

W-, D- ja White-testit			LM-testi			VIF-arvot	
Testi	Arvo	p-arvo	p	Arvo	p-arvo	Muuttuja	VIF-arvo
W-testi	0,939	0,17	1	0,12	0,27	C_t	5,35
White-testi	12,63	0,82	2	3,66	0,84	R_t	4,56
D-testi	2,11	-	3	5,65	0,94	$T_{Vantaa,t}$	1,41

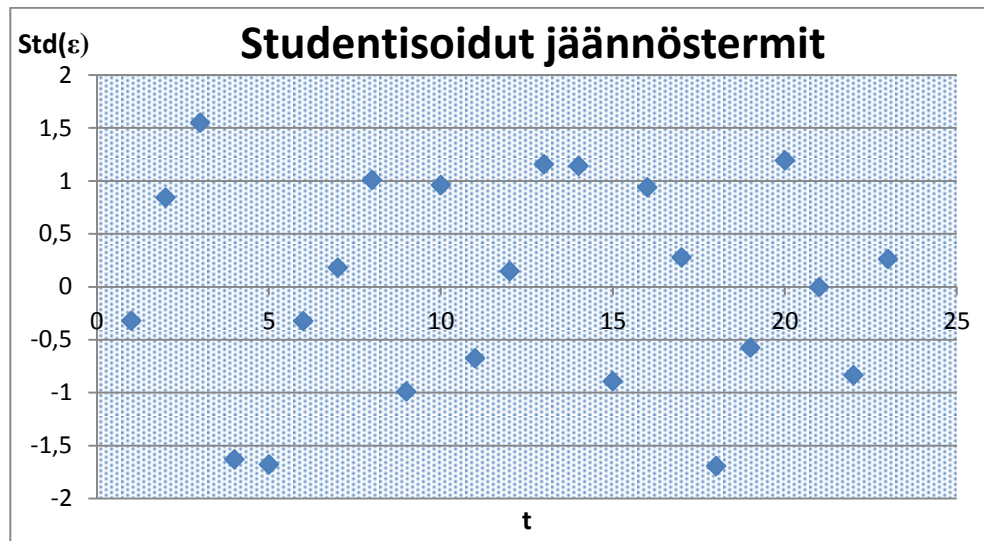
BKT per capitaan VIF-arvo nousee arvoon 5,35. Tätä tulosta voidaan tulkita eri tavoin, sillä jotkin lähteet pitävät lukuarvoa 5 VIF-arvon kriittisenä ylärajana, kun taas toiset lähteet sallivat VIF-arvon nousemisen lukuarvoon 10 ennen kuin multikollinearisuus tulkitaan olevan haitallisella tasolla [16, 17]. Koska VIF-arvo nousee vain hieman lukuarvoa 5 suuremmaksi, ei multikollinearisuutta tulkita ilmenevän niin paljoa, ettei malli olisi käyttökelpoinen. Asia on kuitenkin syytä huomioida, ja rakennuskannan VIF-arvon suuruuden myötä mallin keskivirhe-estimaatteihin saattaa liittyä epävarmuutta.

Muut diagnostiset testit tehtiin tarkastelemalla mallin jäännöstermejä. Kuvassa 7.5 on esitetty mallin jäännöstermit, ja kuvassa 7.6 on esitetty samat jäännöstermit studentisoituina.



Kuva 7.5. Mallin VI jäännöstermit

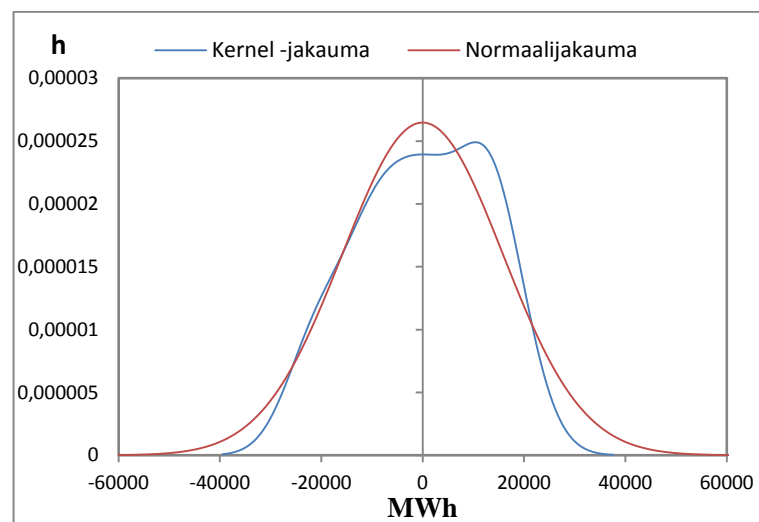
Kuvaan 7.5 piirretyn toisen asteen käyrän avulla voidaan silmämääräisesti tarkastella jäännöstermien odotusarvoa, joka on lähellä nollaa käyrän kulkiessa lähellä nollakselia.



Kuva 7.6. Mallin V1 studentisoidut jäännöstermit

Studentisoitujen residuaalien itseisarvot ovat suurimmillaan 1,69, minkä mukaan joukossa ei ole poikkeuksellisia havaintoarvoja.

Jäännöstermien normalisuutta tarkasteltiin kuvaajan avulla, jossa jäännöstermien jakautumista verrataan normaalijakauman kuvaajaan. Kuvasta 7.7 voidaan havaita jäännöstermijoukon Kernel-tiheysfunktion poikkeavan normaalijakaumasta enemmän kuin Suomen mallin jäännöstermien tiheysfunktio.



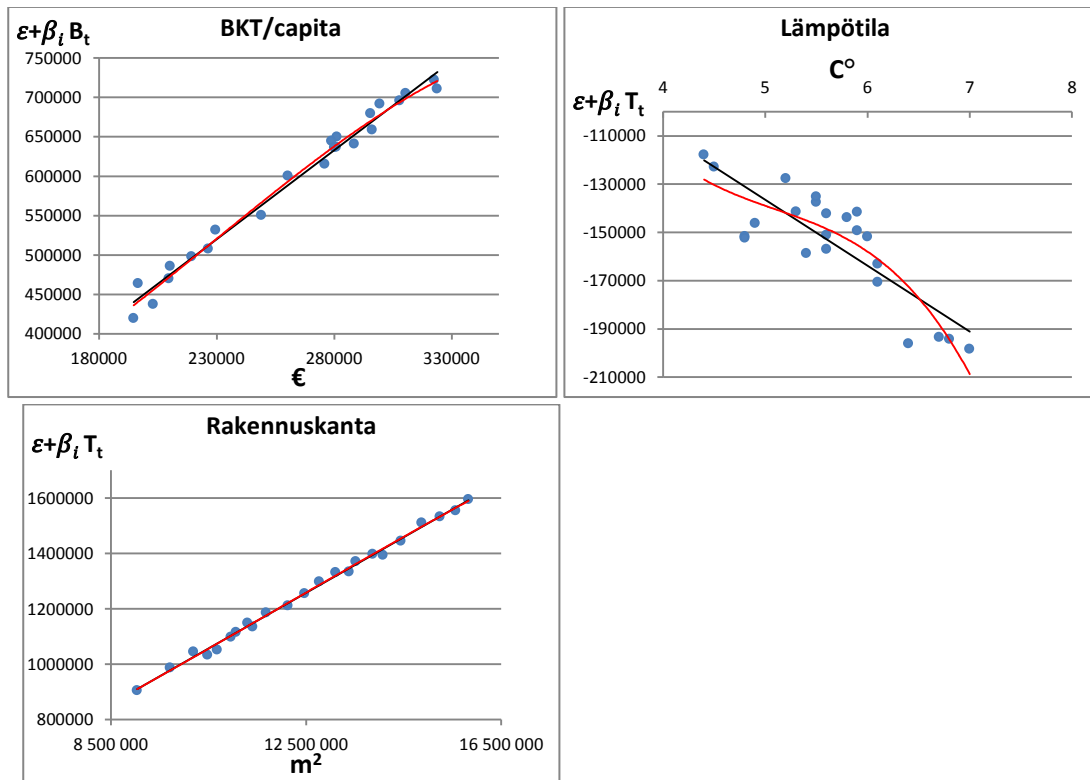
Kuva 7.7. Mallin V1 jäännöstermien Kernel -tiheysfunktio

Normalisuuden varmistamiseksi tehtiin jäännöstermeille myös Shapiro-Wilkin W-testi. Testin arvoksi saatiin 0,939, jota vastaava p-arvo on 0,17. Testituloksen perusteella nolalahypoteesia jäännöstermien normalisuudesta ei tarvitse hylätä.

Jäännöstermien skedastisuutta voidaan havainnoida kuvasta 7.5, jonka mukaan jäännöstermit olisivat homoskedastisia. Homoskedastisuutta testattiin vielä White-testillä, josta saatiin p-arvoksi 0,82. Testituloksen mukaan nollahypoteesi jäännöstermien homoskedastisuudesta jää voimaan.

Jäännöstermien autokorrelaation testaamiseksi suoritettiin d- ja LM-testit. D-testin tulokseksi saatiin 2,11, mikä on hyvin lähellä testin normaaliarvoa 2. LM-testin p-arvot ovat suurempia kuin kriittinen arvo. Testitulosten perusteella nollahypoteesi autokorrelaattomuudesta pysyy voimassa.

Lineaarisuusoletuksen täytyminen varmistettiin piirtämällä osittaiset jäännöstermit kuvaan 7.8.



Kuva 7.8. Mallin V1 osittaiset jäännöstermit

Regression perusoletukset täyttyvät, joten malli V1 on käyttökelpoinen ja sille PNS-menetelmällä laskettavat kertoimet ovat luotettavia.

Taulukkoon 7.11 on koottu regressiomallin ominaisarvot ja PNS-menetelmällä lasketut kertoimet.

Taulukko 7.11. Mallin VI kertoimet ja ominaisarvot

	Arvo	Keskivirhe	t-arvo	p-arvo
θ_0	-369336	28314,14	-13,04	$6,25 \cdot 10^{-11}$
θ_1	2,661	0,157	16,94	$6,33 \cdot 10^{-13}$
θ_2	0,094	0,003	29,80	$2,05 \cdot 10^{-17}$
θ_3	-21757	4894,34	-4,45	0,00028
R^2	0,9980			
Korjattu R^2	0,9977			
Mallin F-arvo	3204			$7,26 \cdot 10^{-26}$

Taulukon 7.11 arvoista nähdään, että p-arvot ovat hyvin pieniä. Mallin selitysaste on erittäin korkea 0,998, minkä mukaan malli kykenee mallintamaan erittäin hyvin Vantaan sähkönkulutusta.

Lopuksi testattiin mallissa käytettävien sarjojen stationaarisuus ADF-testillä. Testitulokset on esitetty taulukossa 7.12.

Taulukko 7.12. ADF-testin tulokset mallin VI sarjoille

Muuttuja	Alkuperäinen sarja			1. differenssi		
	Testiarvo	Parametrit	Kriittinen arvo	Testiarvo	Parametrit	Kriittinen arvo
$E_{Vantaa,t}$	0,808	Vakio	-1,952	-2,642	Ei vakiota	-1,952
C_t	0,808	Ei vakiota	-1,951	-2,642	Ei vakiota	-1,952
R_t	-0,651	Vakio	-1,952	-4,119	Vakio + trendi	-3,646
$T_{Vantaa,t}$	-4,467	Vakio + trendi	-3,636	-6,877	Vakio + trendi	-3,646

Vain lämpötilan aikasarja on stationaarinen ilman differointia. Differoinnin myötä myös muut aikasarjat muuttuvat stationaariksi. Yhteisintegroituuden varmistamiseksi suoritettiin mallin virhetermeille ADF-testi, jonka tulokseksi saatiin -4,862, jota vastaava AEG-testin mukainen p-arvo on 0,033. P-arvo on selvästi alle kriittisen arvon 0,05 ja tämän perusteella mallin sarjat ovat yhteisintegroituja ja tulokset myös tältä osin luotettavia.

Seuraavaksi muodostettiin log-lineaarinen malli sekä osittainen sopeutumismalli Vantaan sähkönkulutusta kuvaavan mallin pohjalta.

Malleille lasketut kertoimet tunnuslukuineen on esitetty taulukossa 7.13.

Taulukko 7.13. Vantaan sähkönkulutuksen log-lineaarisen ja osittaisen sopeutumismallin kertoimet ja ominaisarvot

Log-lineaarinen malli					Osittainen sopeutumismalli				
	Arvo	Keskivirhe	t-arvo	p-arvo		Arvo	Keskivirhe	t-arvo	p-arvo
β_0	-6,148	0,259	-23,75	$1,4 \cdot 10^{-15}$	β_0	-3,725	1,238	-3,008	0,008
β_1	0,520	0,033	15,70	$2,5 \cdot 10^{-12}$	β_1	0,399	0,677	5,896	$1,4 \cdot 10^{-5}$
β_2	0,853	0,033	26,07	$2,4 \cdot 10^{-16}$	β_2	0,517	0,171	3,02	0,007
β_3	-0,082	0,022	-3,65	0,0017	β_3	-0,086	0,21	-4,097	0,0007
R^2	0,998				(1- δ)	0,325	0,163	1,995	0,061
Korj. R^2	0,997				R^2	0,998			
F-arvo	2650			$4,4 \cdot 10^{-25}$	Korj. R^2	0,998			
					F-arvo	2300			$4,1 \cdot 10^{-24}$

Log-lineaarisen mallin kaikkien kertoimien p-arvot ovat alle kriittisen arvon 0,05. Osittaisen sopeutumismallin kohdalla viivetermin kertoimen (1- δ) p-arvo ylittää kriittisen arvon ($0,061 > 0,05$), mutta vain niukasti. Tästä huolimatta voidaan todeta, että mallien kertoimet ovat riittävässä määrin tilastollisesti päteviä ja niitä voidaan käyttää joustojen arviointiin.

Malleille tehtiin myös diagnostiset testit. Testien tulokset on esitetty taulukoissa 7.14 ja 7.15.

Taulukko 7.14. Diagnostisten testien tulokset Vantaan sähkönkulutuksen log-lineaariselle mallille

W-, D- ja White-testit			LM-testi			VIF-arvot	
Testi	Arvo	p-arvo	p	Arvo	p-arvo	Muuttuja	VIF-arvo
W-testi	0,938	0,163	1	1,14	0,71	$\ln(C_t)$	5,46
White-testi	6,53	0,31	2	3,06	0,78	$\ln(R_t)$	4,71
D-testi	1,76	-	3	2,71	0,56	$\ln(T_t)$	1,39

Taulukko 7.15. Diagnostisten testien tulokset Vantaan sähkönkulutuksen osittaiselle sopeutumismallille

W-, D- ja White-testit			LM-testi			VIF-arvot	
Testi	Arvo	p-arvo	p	Arvo	p-arvo	Muuttuja	VIF-arvo
W-testi	0,903	0,029	1	6,21	0,99	$\ln(C_t)$	26,42
White-testi	17,54	0,77	2	10,49	0,99	$\ln(R_t)$	149,20
D-testi	1,90	-	3	9,85	0,98	$\ln(T_t)$	1,41
						$\ln(E_{Vantaa,t-1})$	262,98

Mallien sarjojen stationaarisuuden testaamiseksi niille tehtiin ADF-testit, joiden tulokset on esitetty taulukossa 7.16.

Taulukko 7.16. ADF-testin tulokset Vantaan sähkönkulutuksen log-lineaarisen mallin ja osittaisen sopeutumismallin sarjoille

Muuttuja	Alkuperäinen sarja			1. differenssi		
	Testiarvo	Parametrit	Kriittinen arvo	Testiarvo	Parametrit	Kriittinen arvo
$\ln(E_{Vantaa,t})$	-2,011	Vakio	-3,007	-4,264	Vakio + trendi	-3,646
$\ln(C_t)$	-0,783	Vakio	-3,007	-2,504	Ei vakiota	-1,952
$\ln(R_t)$	-6,850	Vakio + trendi	-3,636	-4,628	Vakio + trendi	-3,646
$\ln(T_t)$	-4,495	Vakio + trendi	-3,636	-7,134	Vakio + trendi	-3,646

Log-lineaarinen malli läpäisee kaikki diagnostiset testit. Osittaisen sopeutumismallin kohdalla W-testin p-arvo jää liian pieneksi, eli jäännöstermien sarja poikkeaa normaalisuudesta sallittua enemmän. Lisäksi monen selittävän tekijän VIF-arvo nousee hyvin suureksi. Näistä tuloksista johtuen mallin kertoimiin liittyy runsaasti epävarmuutta ja myös muihin t- ja F-jakaumiin perustuvia testituloksia voidaan pitää ainoastaan suuntaa antavina.

Kaikki malleissa käytettävät sarjat ovat stationaarisia 1. differenssimuodossaan. AEG-testien p-arvoiksi saatiin log-lineaarisen mallin osalta 0,026 (1. differenssi.) ja osittaisen sopeutumismallin osalta 0,0142 (2. differenssi.). Arvojen perusteella sarjat kummassakin mallissa ovat lopulta yhteisintegroituja.

7.3.1 Mallit Vantaan sektoreiden sähkönkulutuksille

Vantaan sähkönkulutus voidaan jakaa asuntojen kulutukseen ja muuhun kulutukseen (kuva 6.1). Näille sektoreille voidaan tehdä omat sähkönkulutusta kuvaavat mallinsa. Taulukossa 7.17 on listattu eräitä sopivia malleja kuvaamaan Vantaan muuta sähkönkulutusta.

Taulukko 7.17. Parhaimmat mallit Vantaan muulle sähkönkulutukselle

Selittävät tekijät	Korjattu R^2	SBIC
<i>BKT per capita, työttömyys %</i>	0,924	21,96
<i>Vantaan rakennuskanta, lämpötila</i>	0,948	21,58
<i>Väkiluku, työttömyys %</i>	0,982	20,52
<i>BKT per capita, toimistotilojen käyttäm. aste, Vantaan rakennuskanta</i>	0,995	19,38

Korjatun selityksasteen sekä BIC-arvojen perusteella Vantaan muuta kulutusta kuvaamaan ehdotetaan mallia V_{muu} , joka on muotoa

$$E_{Vantaa_muu,t} = \beta_0 + \beta_1 C_t + \beta_2 K_t + \beta_3 R_t + \varepsilon_t,$$

jossa $E_{Vantaa_muu,t}$ = Vantaan muu kulutus [MWh]

C_t = Vantaan bruttokansantuote per capita [M€]

K_t = Pääkaupunkiseudun toimitilojen käyttämättömyysaste [%]

R_t = Vantaan rakennuskanta [m^2]

Mallille tehtyjen diagnostisten testien sekä ADF-testien tulokset on esitetty taulukoissa 7.18 ja 7.19.

Taulukko 7.18. Diagnostisten testien tulokset mallille V_{muu}

W-, D- ja White-testit			LM-testi			VIF-arvot	
Testi	Arvo	p-arvo	p	Arvo	p-arvo	Muuttuja	VIF-arvo
W-testi	0,971	0,712	1	8,33	0,996	C_t	6,96
White-testi	4,26	0,11	2	9,94	0,993	K_t	1,90
D-testi	1,05	-	3	10,73	0,987	R_t	7,50

Taulukko 7.19. ADF- testin tulokset mallin V_{muu} sarjoille

Muuttuja	Alkuperäinen sarja			1. differenssi		
	Testiarvo	Parametrit	Kriittinen arvo	Testiarvo	Parametrit	Kriittinen arvo
$E_{Vantaa_muu,t}$	-0,614	Vakio	-3,007	-3,291	Vakio	-3,014
C_t	-0,844	Vakio	-3,007	-2,642	Ei vakiota	-1,952
K_t	-1,156	Vakio	-3,007	-2,513	Ei vakiota	-1,952
R_t	-2,459	Vakio + trendi	-3,636	-4,119	Vakio + trendi	-3,646

BKT per capitaa VIF-arvo nousee yli arvon 5, mutta pysyy selvästi alle arvon 10, jolloin multikollinearisuus pysyy siedettävällä tasolla. Muut diagnostiset testit malli läpäisee ongelmitta. Kaikki mallin sarjat ovat ADF-testin tulosten mukaan stationaarisia 1. differenssissään ja kun AEG-testin p-arvoksi saatiin 0,0048 (1. differenssi), ovat sarjat myös yhteisintegroituvia. Mallin muuttujille lasketut kertoimet ja mallin ominaisarvot ovat taulukossa 7.20.

Taulukko 7.20. Mallin V_{muu} kertoimet ja ominaisarvot

	Arvo	Keskivirhe	t-arvo	p-arvo
β_0	-390629,96	19934,74	-19,60	$4,62 \cdot 10^{-14}$
β_1	1,590	0,179	8,88	$3,47 \cdot 10^{-8}$
β_2	-3823,19	1255,49	-3,05	0,0067
β_3	0,0645	0,0040	15,95	$1,86 \cdot 10^{-12}$
R^2	0,9954			
Korjattu R^2	0,9947			
Mallin F-arvo	1385			$2,04 \cdot 10^{-22}$

Taulukon 7.20 arvojen mukaan mallille V_{muu} lasketut kertoimet ovat tilastollisesti päteviä ja sen avulla pystytään mallintamaan Vantaan muuta sähkönkulutusta erittäin hyvin.

Vantaan asumisen sähkönkulutusta kuvaamaan kokeiltiin myös useita malleja, joista sopivimmat on esitetty taulukossa 7.21.

Taulukko 7.21. Parhaat mallit Vantaan asumisen sähkönkulutukselle

Selittävät tekijät	Korjattu R²	SBIC
Vantaan väkiluku	0,965	20,07
Uudenmaan BKT	0,871	21,37

Vantaan asumisen sähkönkulutusta voidaan kuvata siis joko Vantaan väkiluvun tai Uudenmaan bruttokansantuotteen avulla. Kun malleihin yritettiin lisätä muita selittäviä tekijöitä, nousivat lisättyjen selittäjien kertoimien p-arvot korkeiksi, eikä selittäjiä näin ollen voitu lisätä malliin.

Taulukossa näkyvien korjattujen selitysasteiden sekä BIC-arvojen perusteella Vantaan kotitalouksien sähkönkulutusta kuvaamaan ehdotetaan mallia V_{asum} , joka on muotoa

$$E_{Vantaa_asum,t} = \beta_0 + \beta_1 V_t + \varepsilon_t$$

jossa $E_{Vantaa_asum,t}$ = Vantaan muu kulutus [MWh]
 V_t = Vantaan väkiluku [hlö]

Mallille tehtyjen diagnostisten testien sekä ADF-testien tulokset on esitetty taulukoissa 7.22 ja 7.23.

Taulukko 7.22. Diagnostisten testien tulokset mallille V_{asum}

W-, D- ja White-testit			LM-testi		
Testi	Arvo	p-arvo	p	Arvo	p-arvo
W-testi	0,938	0,091	1	9,75	0,998
White-testi	6,79	0,966	2	9,60	0,992
D-testi	0,77	-	3	10,3	0,984

Taulukko 7.23. ADF-testin tulokset mallin V_{asum} sarjoille

Muuttuja	Alkuperäinen sarja			1. differenssi		
	Testiarvo	Parametrit	Kriittinen arvo	Testiarvo	Parametrit	Kriittinen arvo
$E_{Vantaa_asum,t}$	-2,821	Vakio + trendi	-3,582	-6,144	Vakio+trendi	-3,589
V_t	-1,883	Vakio + trendi	-3,582	-4,605	Vakio+trendi	-3,589

Diagnostisissa testeissä ei ilmennyt ongelmia. ADF-testin mukaan mallin sarjat ovat stationaarisia 1. differenssissä ja AEG-testin p-arvo 0,001 (1. differenssi) osoittaa sarjojen yhteisintegroituvuuden. Tulosten mukaan malli on käyttökelpoinen.

Mallin muuttujille lasketut kertoimet ja mallin ominaisarvot ovat taulukossa 7.24.

Taulukko 7.24. Mallin $V_{asuminen}$ kertoimet ja ominaisarvot

	Arvo	Keskivirhe	t-arvo	p-arvo
β_0	-668339,5	36706,43	-18,21	$1,09 \cdot 10^{-16}$
β_1	7,004	0,214	32,78	$2,77 \cdot 10^{-23}$
R^2	0,9755			
Korjattu R^2	0,9746			
Mallin F-arvo	1075			$2,2 \cdot 10^{-18}$

Vaikka mallissa on vain yksi selittävä tekijä, nousee mallin selitysaste suureksi ja mallin F-arvoa vastaava p-arvo kertoo mallin selitysvoinman olevan suuri. Sektoreiden sähkönkulutuksille ei tehty log-lineaarisia malleja tai osittaisia sopeutumismalleja.

7.4 Malli Helsingin sähkönkulutukselle

Helsingin sähkönkulutukselle lähdettiin muodostamaan mallia Vantaan sähkönkulutukselle luodun mallin pohjalta. Rakennuskannan aikasarjaa ei Helsingille ollut käytettävissä ja osoittautui, että BKT per capitalla ja lämpötilalla muodostetulla mallilla jäi muuttujille laskettujen kertoimien t-arvot sallittua pienemmiksi. Helsingin sähkönkulutusta ei siis voitu mallintaa samalla mallilla kuin Vantaan sähkönkulutusta. Helsingin sähkönkulutusta parhaiten kuvaavat mallit on lueteltu taulukossa 7.25.

Taulukko 7.25. Parhaimmat mallit Helsingin sähkönkulutukselle

Malli	Selittävät tekijät	Korjattu R^2	SBIC
H1	Uudenmaan BKT, Lämpötila	0,979	9,15
H2	Uudenmaan BKT, Työttömyys %	0,984	8,89
H3	BKT per capita	0,956	9,84
H4	Väkiluku	0,909	10,55
H5	Uudenmaan BKT	0,974	9,29

Mallien luonnissa on käytetty 25 havaintoarvoa, joissa bruttokansantuotteen aikasarjan pituus oli rajoittavana tekijänä. Kaikki taulukon 7.25 mallit läpäisevät mallinvalintatestit. Korjatun selityksasteen ja SBIC-arvon perusteella paras olisi malli H2, jossa selittävinä tekijöinä on Uudenmaan bruttokansantuote ja Helsingin työttömyysprosentti. On kuitenkin perusteltua valita malli H1 kuvaamaan Helsingin sähkönkulutusta, sillä lämpötilan vaikutus sähkönkulutukseen on kiistaton. Malli H1 on muotoa

$$E_{Helsinki,t} = \beta_0 + \beta_1 B_{Uusimaa,t} + \beta_2 T_{Helsinki,t} + \varepsilon_t,$$

jossa $E_{Helsinki,t}$ = Helsingin vuosittainen sähkönkulutus [MWh]
 $B_{Uusimaa,t}$ = Uudenmaan bruttokansantuote [M€]
 $T_{Helsinki,t}$ = Vuoden keskilämpötila Helsingissä [C°]

Mallille H1 tehtiin diagnostiset testit, joiden tulokset on esitetty taulukossa 7.26.

Taulukko 7.26. Diagnostisten testien tulokset mallille H1

W-, D- ja White-testit			LM-testi			VIF-arvot	
Testi	Arvo	p-arvo	p	Arvo	p-arvo	Muuttuja	VIF-arvo
W-testi	0,549	0,966	1	0,12	0,27	B _{Uusimaa,t}	1,09
White-testi	3,767	0,42	2	3,66	0,84	T _{Helsinki,t}	1,09
D-testi	1,33	-	3	5,65	0,94		

W-testin tuloksen mukaan jäännöstermit ovat normaalisti jakautuneita ja White-testi osoittaa jäännöstermien homoskedastisuuden. D-testin tuloksesta ei voida tehdä lopullisia johtopäätöksiä jäännöstermien autokorrelaation suhteen, sillä testiarvo jää kriittisen ylä- ja ala-arvon väliin. Asian varmistamiseksi täytyy tarkastelu kohdistaa LM-testin tuloksiin, joiden mukaan ei ole syytä epäillä jäännöstermien autokorreloituneisuutta. Täten malli H1 läpäisee kaikki diagnostiset testit.

Mallin sarjoille tehdyn ADF-testin tulokset on esitetty taulukossa 7.27.

Taulukko 7.27. ADF-testin tulokset mallin H1 sarjoille

Muuttuja	Alkuperäinen sarja			1. differenssi		
	Testiarvo	Parametrit	Kriittinen arvo	Testiarvo	Parametrit	Kriittinen arvo
E _{Helsinki,t}	-0,880	Vakio	-2,992	-3,679	Vakio + Trendi	-3,625
B _{Uusimaa,t}	-0,746	Vakio	-2,992	-2,546	Ei vakiota	-1,951
T _{Helsinki,t}	-4,197	Vakio + Trendi	-3,615	-7,403	Vakio + Trendi	-3,625

Mallin kaikki aikasarjat ovat stationaarisia 1. differenssimuodossaan. Yhteisintegroituvuuden varmistamiseksi tehtiin vielä mallin virhetermeille ADF-testi. Testissä jäännöstermien sarja muuttui ensimmäisessä differenssissään stationaariseksi testiarvon ollessa -8,582, joka ylitti selkeästi kriittisen arvon -4,589. Testin perusteella mallin sarjat ovat yhteisintegroituvia. Mallille H1 laskettiin kertoimet, jotka näkyvät taulukossa 7.28.

Taulukko 7.28. Mallin H1 kertoimet ja ominaisarvot

	Arvo	Keskivirhe	t-arvo	p-arvo
θ_0	1906868	156349,7	12,2	$2,91 * 10^{-11}$
θ_1	53,35	1,61	33,05	$3,01 * 10^{-20}$
θ_2	-67365,81	25800,83	-2,61	0,016
R^2	0,981			
Korjattu R^2	0,979			
Mallin F-arvo	574			$1,04 * 10^{-19}$

Mallin H1 pohjalta luotiin log-lineaarinen- ja osittainen sopeutumismalli, joiden ominaisarvot on esitetty taulukossa 7.29.

Taulukko 7.29. Helsingin sähkönkulutuksen log-lineaarisen ja osittaisen sopeutumismallin kertoimet ja ominaisarvot

Log-lineaarinen malli					Osittainen sopeutumismalli				
	Arvo	Keskivirhe	t-arvo	p-arvo		Arvo	Keskivirhe	t-arvo	p-arvo
β_0	8,8	0,197	44,74	$4,24 \cdot 10^{-23}$	β_0	3,755	0,686	5,47	$2 \cdot 10^{-5}$
β_1	0,611	0,019	32,24	$5,14 \cdot 10^{-20}$	β_1	0,248	0,05	4,97	$6,4 \cdot 10^{-5}$
β_2	-0,09	0,041	-2,21	0,038	β_2	-0,04	0,023	-1,76	0,093
R^2	0,980				$(1-\delta)$	0,583	0,078	7,44	$2,6 \cdot 10^{-7}$
Korj. R^2	0,978				R^2	0,995			
F-arvo	541			$1,95 \cdot 10^{-19}$	Korj. R^2	0,994			
					F-arvo	1271			$6,8 \cdot 10^{-24}$

Log-lineaarisen mallin kertoimet ovat tilastollisesti päteviä, mutta osittaisen sopeutumismallin kertoimen β_2 p-arvo ylittää arvon 0,05, mutta pysyy niukasti arvon 0,1 alapuolella. Lyhyen- ja pitkän aikavälin laskettuihin joustoihin tulee siis suhtautua pienellä varrauksella.

Mallien tilastollinen pätevyys testattiin diagnostisilla testeillä, joiden tulokset on esitetty taulukoissa 7.30 ja 7.31.

Taulukko 7.30. Diagnostisten testien tulokset Helsingin sähkönkulutuksen log-lineaariselle mallille

W-, D- ja White-testit			LM-testi			VIF-arvot	
Testi	Arvo	p-arvo	p	Arvo	p-arvo	Muuttuja	VIF-arvo
W-testi	0,982	0,914	1	7,16	0,993	$\ln(B_{Uusimaa,t})$	1,08
White-testi	11,5	0,957	2	7,48	0,976	$\ln(T_{Helsinki,t})$	1,08
D-testi	1,08	-	3	8,31	0,960		

Taulukko 7.31. Diagnostisten testien tulokset Helsingin sähkönkulutuksen osittaiselle sopeutumismallille

W-, D- ja White-testit			LM-testi			VIF-arvot	
Testi	Arvo	p-arvo	p	Arvo	p-arvo	Muuttuja	VIF-arvo
W-testi	0,955	0,327	1	16,69	0,999	$\ln(B_{Uusimaa,t})$	25,84
White-testi	5,43	0,204	2	16,03	0,999	$\ln(T_{Helsinki,t})$	1,18
D-testi	1,83	-	3	15,51	0,999	$\ln(E_{Helsinki,t-1})$	25,03

Mallien sarjojen stationaarisuuden testaamiseksi niille tehtiin ADF-testit, joiden tulokset on esitetty taulukossa 7.32.

Taulukko 7.32. ADF-testin tulokset Helsingin sähkönkulutuksen log-lineaarisen mallin ja osittaisen sopeutumismallin sarjoille

Muuttuja	Alkuperäinen sarja			1. differenssi		
	Testiarvo	Parametrit	Kriittinen arvo	Testiarvo	Parametrit	Kriittinen arvo
$\ln(E_{Helsinki,t})$	-1,217	Vakio	-2,992	-3,229	Vakio	-2,999
$\ln(B_{Uusimaa,t})$	-0,840	Vakio	-2,992	-2,302	Ei vakiota	-1,951
$\ln(T_{Helsinki,t})$	-4,336	Vakio + Trendi	-3,615	-7,815	Vakio + Trendi	-3,625

Log-lineaarinen malli läpäisee diagnostiset testit ongelmitta, joten mallin kertoimia voidaan käyttää joustojen arviointiin. Osittaisen sopeutumismallin kohdalla kahden selittävän tekijän VIF-arvot nousevat hyvin korkeiksi. Tästä johtuen mallin kertoimiin liittyy runsaasti epävarmuutta. ADF-testin mukaan mallien sarjat ovat stationaarisia ja AEG-testin p-arvo 0,024 (1. differenssi) osoittaa sarjojen yhteisintegroituvuuden.

8 MALLIEN ANALYSOINTI

Tässä luvussa arvioidaan edellisessä luvussa muodostettuja malleja ja verrataan tuloksia aikaisempiin tutkimuksiin sekä taustateoriaan. Selittävien tekijöiden vaikutuksia sähkönkulutukseen tarkastellaan myös niille laskettujen joustojen avulla. Tässä luvussa keskitytään analysoimaan vain sitä, kuinka hyvin mallit pystyvät selittämään toteutunutta kulutusta toteutuneiden selittävien muuttujien arvoilla. Mallien ennustuskykyä arvioidaan vasta luvussa 9.

8.1 Suomen sähkönkulutusmallin analysointi

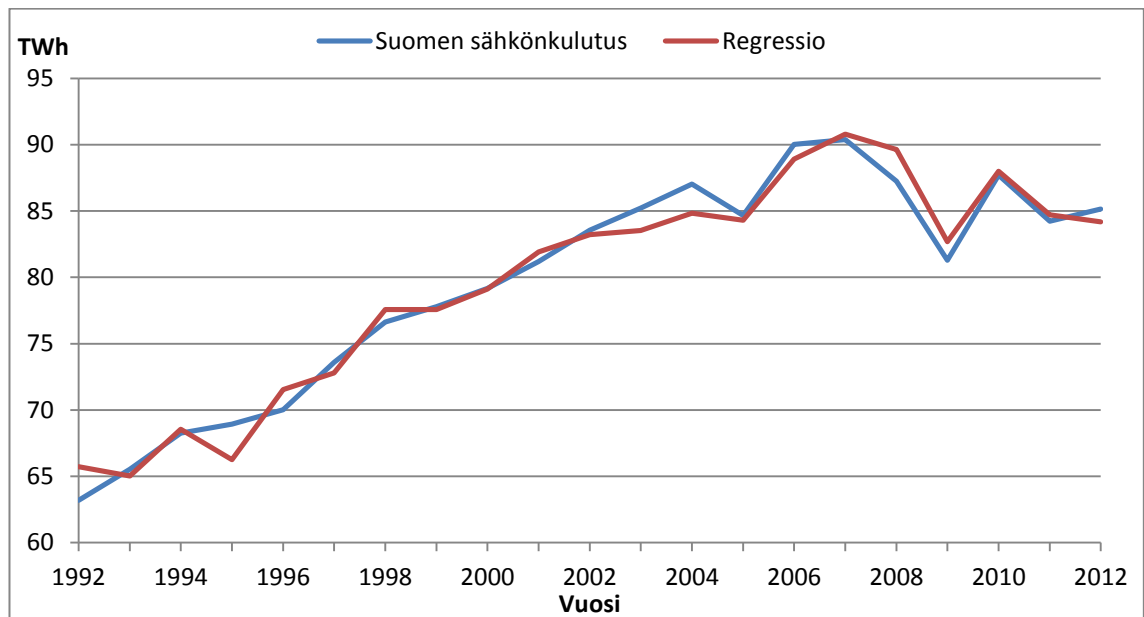
Suomen sähkönkulutusta voidaan mallintaa hyvin usealla erilaisella mallilla (taulukko 7.1). Usean eri mallin muodostuksen jälkeen voidaan todeta, että varsinkin bruttokansantuote soveltuu erinomaisesti selittämään Suomen sähkönkulutusta. Muita hyviksi osoittautuneita selittäjiä olivat sähkön hinta, Suomen väkiluku sekä vuoden keskilämpötila, joista etenkin jälkimmäisellä on kiistatta suuri vaikutus sähkön kokonaiskulutukseen. Viivästettyjä muuttujia sisältäneet mallit saavuttivat usein korkean selitysasteen, mutta mallinvalintatestien tulosten perusteella eivät olleet käyttökelpoisia.

Mallinvalintatestien perusteella Suomen sähkönkulutusta kuvaamaan valittiin malli, jossa selittävinä tekijöinä ovat Suomen bruttokansantuote viitevuoden 2000 hintaisena, sähkön kokonaiskeskihinta ja keskilämpötila, jotka muodostavat myös asialoogisesti järkevän mallin. Testien perusteella malli on käyttökelpoinen ja se voidaan esittää muodossa

$$E_{Suomi,t} = -37956,3 + 0,486B_t - 1960,5S_t - 2475,12T_t + \varepsilon_t,$$

jossa	$E_{Suomi,t}$	= Suomen vuosittainen sähkönkulutus
	B_t	= Suomen bruttokansantuote vuoden 2000 hintaisena [M€]
	S_t	= Sähkön kokonaiskeskihinta [snt/kWh]
	T_t	= Vuoden keskilämpötila [°C]

Kuvassa 8.1 on esitetty Suomen sähkönkulutuskäyrä sekä mallilla laskettu regressiokäyrä aikavälille 1992–2012.



Kuva 8.1. Regressiokäyrä Suomen sähkönkulutukselle

Mallin selitysaste on 97,3 % ja mallin F-arvoa vastaava p-arvo $1,36 \cdot 10^{-13}$. Näiden arvojen perusteella mallia voidaan pitää tilastollisesti erittäin pätevänä ja kuvaavan Suomen sähkönkulutusta hyvin, kuten myös kuvasta 8.1 voidaan nähdä. Malli pystyy hyvin mallintamaan myös vuosina 2008–2010 tapahtuneita rajuja muutoksia sähkönkulutuksessa, minkä perusteella mallin avulla voidaan hyvin arvioida suhdannevaihteluiden vaikutuksia sähkönkulutukseen ja muutoksiin voidaan myös reagoida.

Luvussa 7 muodostettiin Suomen sähkönkulutukselle log-lineaarinen- sekä osittainen sopeutumismalli, joiden avulla saatiin arvioitua mallin selittävien tekijöiden joustot ja ne on esitetty taulukossa 8.1.

Taulukko 8.1. Joustot Suomen sähkönkulutukselle

Muuttuja	Jousto	Lyhyen aikavälin jousto	Pitkän aikavälin jousto
<i>BKT vuoden 2000 hintoihin</i>	0,784	0,669	0,800
<i>Sähkön keskihinta</i>	-0,184	-0,170	-0,203
<i>Vuoden keskilämpötila</i>	-0,084	-0,087	-0,104

Osittaisessa sopeutumismallissa sopeutumisnopeuden δ arvo on 0,836, jonka voidaan tulkita vastaavan ajallisesti noin yhtä vuotta. Joustojen perusteella voidaan todeta, että bruttokansantuotteella on selvästi suurin vaikutus mallin avulla arvioidun sähkönkulutuksen suuruuteen. Lämpötilan jousto on mallin selittäjistä pienin, mutta lämpötilan suuren keskihajonnan takia lämpötilan vaikutus sähkönkulutukseen on käytännössä suurempi, kuin sille laskettu jousto antaa ymmärtää.

Esimerkiksi vuosien 2010 ja 2011 välisessä tarkastelussa keskilämpötilan suhteellinen muutos oli 44 %. Lasketun jouston mukaan tämänsuuruinen muutos laskisi sähkönkulutusta noin 3,7 % ($0,44 \cdot [-0,084] \approx -0,037$). Samalla aikavälillä bruttokansantuotteen muutoksella oli log-lineaarisen mallin mukaan 2,1 %:n ja sähkön keskihinnan muutoksella -2,5 %:n vaikutus Suomen sähkönkulutukseen. Toteutunut Suomen sähkönkulutuksen suhteellinen muutos oli vuonna 2011 noin -3,9 %, minkä luodun mallin mukaan voidaan arvioida aiheutuneen aikaisempaa vuotta korkeammasta keskilämpötilasta.

Mallin merkittävimpana tekijänä voidaan pitää BKT per capita, mikä on vastaava tulos kuin Italian sähkönkulutusta tarkastelevassa tutkimuksessa. Toisin kuin Italian sähkönkulutusta kuvaavassa mallissa, viivetermejä ei voitu käyttää Suomen sähkönkulutusta kuvaavassa mallissa. Toinen poikkeavuus on sähkön hinnan selitysvoima, jota ei ollut Italian kohdalla havaittavissa. Lämpötilan vaikutusta sähköenergian kulutukseen ei Bianco et al.:n tutkimuksessa ole tarkasteltu. Ilmastollisten poikkeavuuksien takia tämä on perusteltua. [2]

8.2 Vantaan sähkönkulutusmallien analysointi

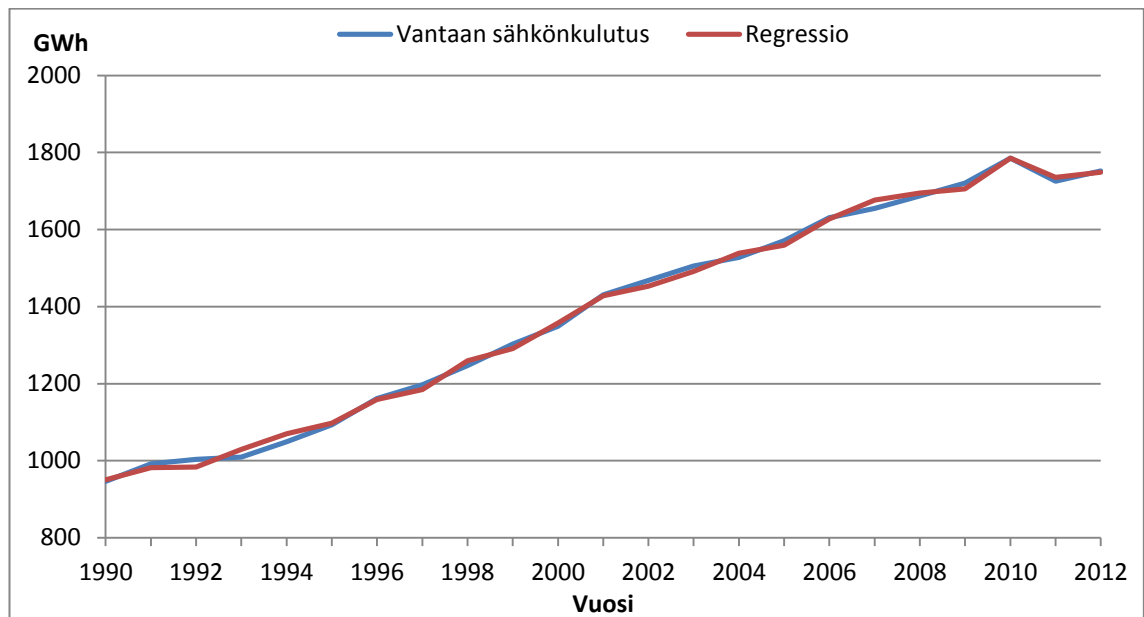
Myös Vantaan sähkönkulutusta olisi mahdollista mallintaa usealla erilaisella mallilla. Parhaimmiksi selittäviksi tekijöiksi paljastui bruttokansantuote, Vantaan väkiluku, rakennuskanta sekä keskilämpötila. Viivästettyjä muuttujia ei voitu hyödyntää mallinvalintakriteerien perusteella. Lopulta Vantaan sähkönkulutusta kuvaamaan valittiin malli, jossa selittävinä tekijöinä ovat BKT per capita, Vantaan rakennuskanta ja keskilämpötila. Malli on muotoa

$$E_{Vantaa,t} = -369336 + 2,661C_t + 0,094R_t - 21757T_{Vantaa,t} + \varepsilon_t,$$

jossa $E_{Vantaa,t}$ = Vantaan vuosittainen sähkönkulutus [MWh]
 C_t = Vantaan bruttokansantuote per capita [€]
 R_t = Vantaan rakennuskanta [m^2]
 $T_{Vantaa,t}$ = Vuoden keskilämpötila Vantaalla [$^{\circ}C$]

Mallin selitysaste on 99,8 % ja mallin F-arvoa vastaava p-arvo $7,26 \cdot 10^{-26}$, minkä mukaan mallilla kyetään kuvaamaan Vantaan sähkönkulutusta hyvin tarkasti. Kaikki mallin selittävät muuttujat ovat myös taustateorian mukaan sopivia selittäjiä sähkönkulutukselle. Aikaisemman tutkimuksen tulokset tukevat vahvasti mallin valintaa [5].

Kuvassa 8.2 on esitetty Vantaan sähkönkulutuskäyrä sekä mallilla V1 laskettu regressiokäyrä aikavälille 1990–2012.



Kuva 8.2. Regressiokäyrä Vantaan sähkönkulutukselle

Kuvasta 8.2 nähdään, kuinka hyvin regressiokäyrä osuu päällekkäin Vantaan toteutuneen kulutuskäyrän kanssa. Regressiomallin jäännöstermien keskihajonta on vain 12324 MWh.

Log-lineaarisen ja osittaisen sopeutumismallin avulla saatiin laskettua joustot Vantaan sähkönkulutusta selittäville tekijöille ja ne on esitetty taulukossa 8.2.

Taulukko 8.2. Joustot Vantaan sähkönkulutukselle

Muuttuja	Jousto	Lyhyen aikavälin jousto	Pitkän aikavälin jousto
<i>BKT per capita</i>	0,520	0,399	0,591
<i>Vantaan rakennuskanta</i>	0,853	0,517	0,766
<i>Vuoden keskilämpötila</i>	-0,082	-0,086	-0,127

Osittaisessa sopeutumismallissa sopeutumismopeuden δ arvo on 0,675, joka ei ole yhtä suuri kuin Suomen mallin sopeutumismopeus. Tästä huolimatta voidaan tulkita arvon vastaavan ajallisesti myös noin yhtä vuotta. Taulukosta 8.2 nähdään, että BKT per capita:n jousto ei ole yhtä suuri kuin koko Suomen tapauksessa. Vantaan rakennuskannalle lasketut joustot ovat puolestaan hyvin suuria, mikä tuloksena on odotettu. Lämpötilan joustot ovat samaa suuruusluokkaa kuin Suomen kohdalla. Aikavälillä 1981–2012 Vantaan keskilämpötilan vaihteluvälin suuruus on ollut 4,4 °C, mikä mallin mukaan voi aiheuttaa sähkönkulutuksessa noin 95,7 GWh:n suuruisen eron kylmän ja lämpimän vuoden välillä. Malli kykenee hyvin selittämään vuonna 2010 tapahtuneen piikin sähkönkulutuksessa. Kyseisenä vuonna BKT per capita aiheutti mallin mukaan 1,47 % muutoksen, rakennuskanta 0,91 % ja lämpötila 1,57 % muutokset sähkönkulutukseen.

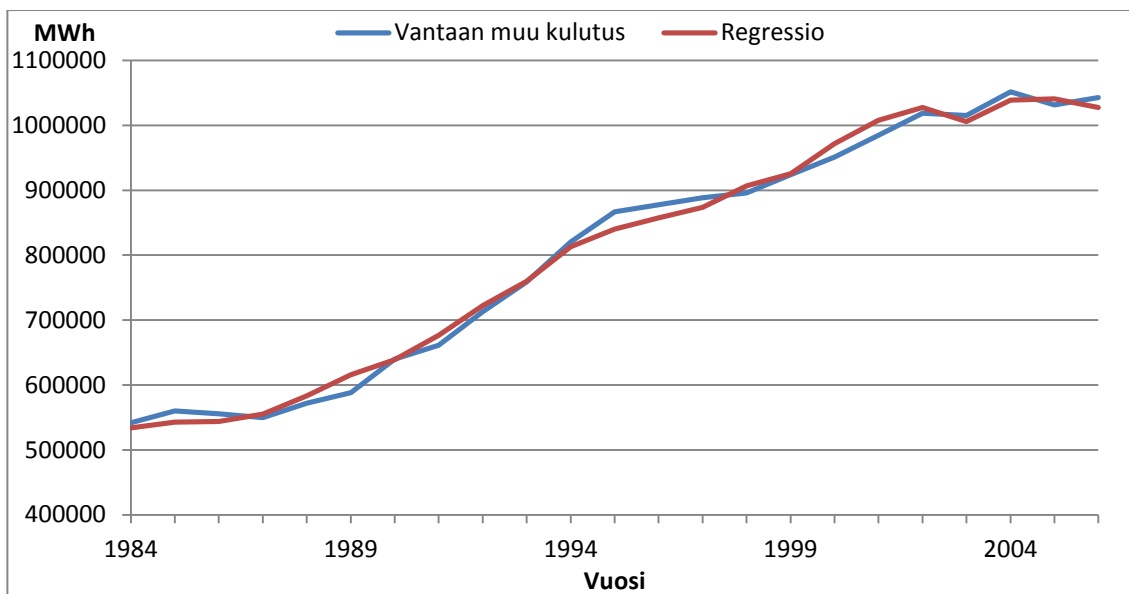
Toteutunut muutos vuonna 2010 oli 3,74 %, joka jää hieman pienemmäksi kuin joustojen avulla arvioitu muutos 3,95 %. Mallin V1 avulla laskettu arvio vuodelle 2010 poikkeaa kuitenkin vain -60 MWh toteutuneesta kulutuksesta.

Vantaan sektoreiden sähkönkulutusta kuvaamaan pystyttiin luomaan myös hyvän selityksasteen omaavat mallit. Vantaan muuta kulutusta voidaan kuvata mallilla

$$E_{Vantaa_muu,t} = -390629,96 + 1,59B_t - 3823,19K_t + 0,0645R_t + \varepsilon_t,$$

jossa $E_{Vantaa_muu,t}$ = Vantaan muu kulutus [MWh]
 B_t = Vantaan bruttokansantuote per capita [M€]
 K_t = Pääkaupunkiseudun toimistotilojen käyttämättömyysaste [%]
 R_t = Vantaan rakennuskanta [m²]

Kuvassa 8.3 on esitetty Vantaan muun kulutuksen sähkönkulutuskäyrä sekä mallilla V_{muu} laskettu regressiokäyrä aikavälille 1984–2012.



Kuva 8.3. Regressiokäyrä Vantaan muulle kulutukselle

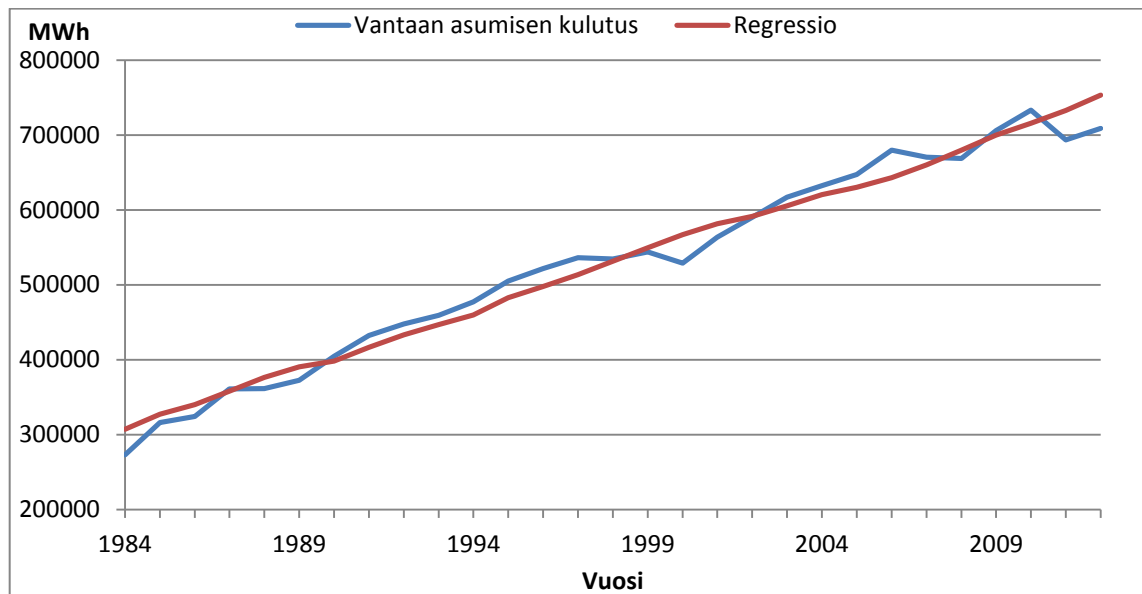
Kuvasta 8.3 nähdään, että malli selittää hyvin Vantaan muun sähkönkulutuksen käyttäytymistä.

Asumisen kuluttamaa sähköä Vantaalla voidaan kuvata mallilla $V_{asuminen}$, joka on muotoa

$$E_{Vantaa_asum,t} = -668339,5 + 7,004V_t + \varepsilon_t$$

jossa $E_{Vantaa_asum,t}$ = Vantaan asumisen kulutus [MWh]
 V_t = Vantaan väkiluku [hlö]

Kuvassa 8.4 on esitetty Vantaan muun kulutuksen sähkönkulutuskäyrä sekä mallilla V_{asuminen} laskettu regressiokäyrä aikavälille 1984–2012.



Kuva 8.4. Regressiokäyrä Vantaan asumisen kulutukselle

Pelkästään väkiluvun avulla malli pystyy selittämään merkittävän osan Vantaan asumisen kuluttaman sähkön varianssista. Kuvasta 8.4 kuitenkin huomataan, että erityisesti yksittäisinä vuosina tapahtuneita piikkejä sähkön kulutuksessa ei mallin avulla kyetä selittämään. Odotusten vastaisesti lämpötilalla ei ollut selitysvaimaa asumisen sähkönkulutusta mallinnettaessa.

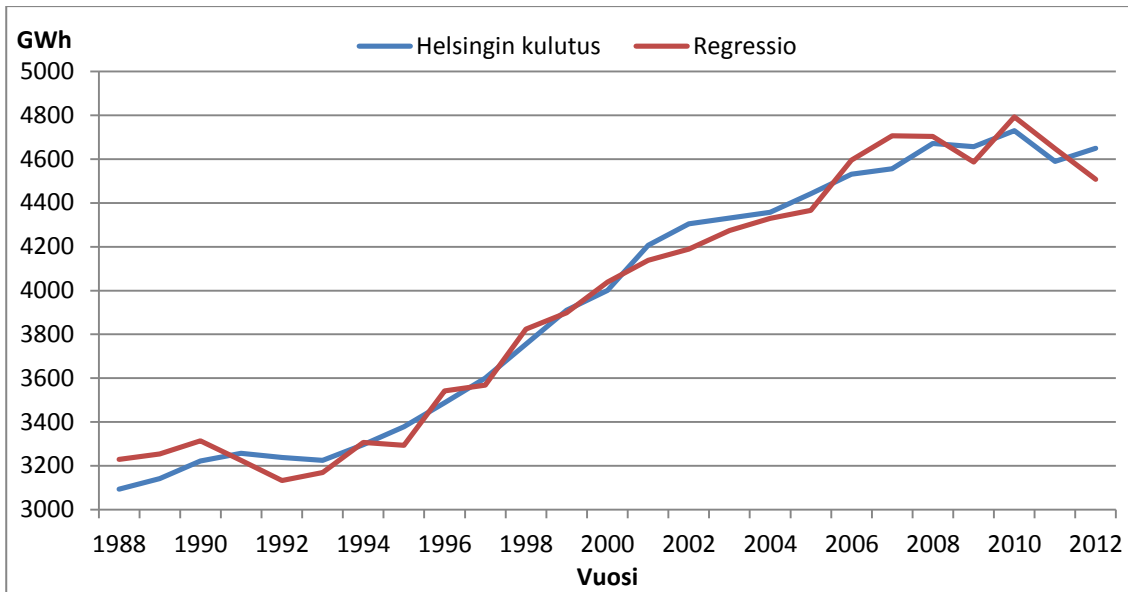
8.3 Helsingin sähkönkulutusmallin analysointi

Helsingin sähkönkulutusta kuvaavassa mallissa oli selittävinä tekijöinä Uudenmaan bruttokansantuote ja Helsingin vuoden keskilämpötila. Malli on muotoa

$$E_{\text{Helsinki},t} = 1906868 + 53,35B_{\text{Uusimaa},t} - 67365,81T_{\text{Helsinki},t} + \varepsilon_t,$$

jossa $E_{\text{Helsinki},t}$ = Helsingin vuosittainen sähkönkulutus [MWh]
 $B_{\text{Uusimaa},t}$ = Uudenmaan bruttokansantuote [M€]
 $T_{\text{Vantaa},t}$ = Vuoden keskilämpötila Helsingissä [°C]

Mallin selitysaste on 98,1 % ja F-arvoa vastaava p-arvo $1,04 \cdot 10^{-19}$, joten myös Helsingin sähkönkulutukselle onnistuttiin luomaan sitä hyvin kuvaava malli. Kuvassa 8.5 on esitetty Helsingin sähkönkulutuskäyrä ja sitä mallintava regressiokäyrä aikavälille 1988–2012.



Kuva 8.5. Regressiokäyrä Helsingin sähkönkulutukselle

Kuvasta 8.5 nähdään regressiokäyrän seuraavan kohtuullisen hyvin toteutunutta kulutusta, mutta suuriakin virheitä esiintyy. Jäännöstermien keskihajonta nousee melko suureksi (80,1 GWh).

Log-lineaarisella ja osittaisella sopeutumismallilla lasketut joustot Helsingin sähkönkulutusta selittäville muuttujille on esitetty taulukossa 8.3.

Taulukko 8.3. Joustot Helsingin sähkönkulutukselle

Muuttuja	Jousto	Lyhyen aikavälin jousto	Pitkän aikavälin jousto
<i>Uudenmaan BKT</i>	0,611	0,248	0,595
<i>Vuoden keskilämpötila</i>	-0,090	-0,04	-0,096

Molempien selittäjien lasketut joustot ovat samaa suuruusluokkaa kuin Suomen ja Vantaan kohdalla. Osittaisessa sopeutumismallissa sopeutumismopeuden δ arvo on 0,417, mikä on selvästi pienempi kuin muiden mallien kohdalla. Vuonna 2007 Uudenmaan bruttokansantuote kasvoi 5,7 % ja keskilämpötilan muutos oli 15,8 % vuoteen 2006 verrattuna. Muutosten myötä sähkönkulutuksen olisi pitänyt kasvaa noin 2 %, mutta toteutunut muutos jäi vain 0,53 %:iin.

9 MALLEILLA ENNUSTAMINEN

Työssä on luotu usean muuttujan regressiomallit, joilla kyettiin mallintamaan eri alueiden sähkönkulutusta varsin hyvin. Samoilla malleilla voidaan myös ennustaa alueiden sähkönkulutusta olettaen, että tulevaisuudessa sähkönkulutus on samalla tavalla riippuvainen selittävästä tekijöistä kuin aikaisemmin.

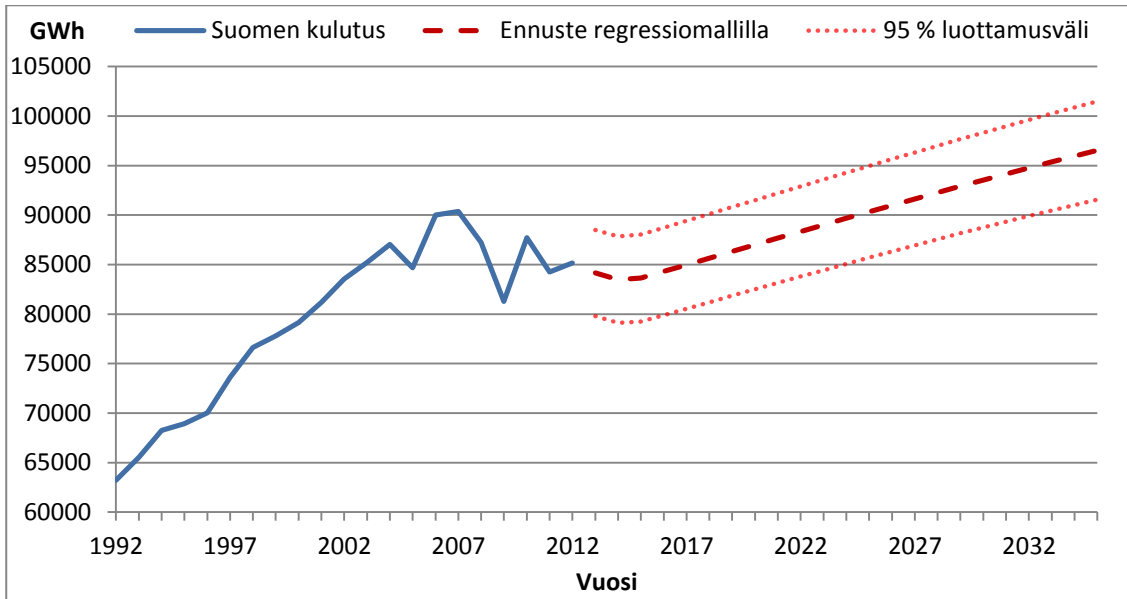
Kaikilla malleilla on todettu olevan korkea selityssaste, mikä ei kuitenkaan tarkoita, että mallien avulla kyettäisiin ennustamaan sähkönkulutusta samalla tarkkuudella. Käytännössä ennustaminen tapahtuu laskemalla sähkönkulutukselle arvoja selittävien muuttujien ennustetuilla arvoilla, joiden tarkkuus vaikuttaa suuresti regressiomalleilla tehtyjen ennusteiden tarkkuuteen. Ennustamisessa voidaan käyttää hyväksi asiantuntijoiden laatimia ennusteita selittävien muuttujien kehityksestä niiltä osin kuin se on mahdollista.

Ennustamiseen liittyvää epävarmuutta kuvataan kaavan (32) avulla lasketuilla 95 %:n luottamusväleillä. Luottamusvälejä on kuitenkin syytä tarkastella varauksella etenkin pitkän aikavälin ennustamisessa. Tässä työssä lasketut luottamusvälit eivät ota huomioon ennusteen pituudesta johtuvaa kasvavaa epävarmuutta.

Kansantalouden pitkän ajan kehityksen vaikutuksia sähkönkulutukseen tarkastellaan kahden yksinkertaisen skenaarion avulla. Toinen skenaario kuvaa nopeaa ja toinen hidasta talouskasvua.

9.1 Ennuste Suomen sähkönkulutukselle

Suomen sähkönkulutuksen ennustamisessa voidaan käyttää hyväksi Suomen bruttokansantuotteen ennusteita. Vuoden 2013 ennusteessa on käytetty bruttokansantuotteen ennakkotietoa, jonka mukaan Suomen kansantalous supistui vuonna 2013 1,45 % [26]. Vuosille 2014 ja 2015 käytetään uusinta Euroopan komission laatimaa ennustetta, joka on julkaistu helmikuussa 2014. Sen mukaan Suomen talous kasvaisi vuonna 2014 0,2 % ja vuonna 2015 1,3 % [38]. Ennusteen lukuja voidaan pitää uskottavina. Tästä eteenpäin arvioidaan Suomen talouden kasvavan keskimäärin 2 %:n vuosivauhtia. Lämpötilan arvona on käytetty mittausjakson 1970–2012 keskiarvoa, joka on 2,8 °C. Sähkön kokonaishinnan oletetaan kasvavan trendinsä mukaisesti. Kuvassa 9.1 on esitetty regressiomallilla tehty ennuste ja luottamusväli Suomen sähkönkulutukselle vuoteen 2035 asti.



Kuva 9.1. Regressiomallilla tehty pitkän aikavälin ennuste Suomen sähkönkulutukselle

Regressiomallilla tehdyn ennusteen mukaan Suomen sähkönkulutus laskisi tai pysyisi vuoden 2012 tasolla 2–3 vuoden ajan ennen uuden kasvuvaiheen alkua. Ennusteen luottamusvälin mukaan Suomen sähkönkulutus nousisi takaisin vuoden 2007 tasolle aikaisintaan vasta noin 6–8 vuoden kuluttua. Vuonna 2020 sähkönkulutus olisi noin 82–92 TWh ja vuonna 2035 noin 92–102 TWh. Regressiomallilla tehty pitkän aikavälin ennuste on hyvin samansuuntainen kuin luvussa kaksi esiteltyt uusimmat ennusteet Suomen sähkönkulutuksen kehityksestä. Ennusteen luottamusväli on leveä, mikä johtuu regressiomallin jäännöstermien suuresta keskihajonnasta (1322 GWh). Luottamusrajat ovat kuitenkin samaa suuruusluokkaa kuin luvussa kaksi esiteltyjen ennusteiden marginaalit.

Suomen sähkönkulutusmallin ennustuskykyä voidaan arvioida lyhyesti pseudo out-of-sample -analyysin avulla laskemalla mallilla arvio vuoden 2013 Suomen sähkönkulutukselle, joka ei ollut mukana mallin muodostamiseen käytetyssä aikasarjassa, mutta siitä on tässä vaiheessa (huhtikuu 2014) kuitenkin saatavissa riittävän tarkkoja arvioita. Energiateollisuus ry:n mukaan Suomen sähkönkulutus oli vuonna 2013 83866 GWh [39]. Kun regressiomalliin sijoitetaan Suomen toteutunut bruttokansantuote (157,6 Mrd. €), arvioitu sähkön kokonaiskeskihinta (11,97 c/kWh) sekä aikajakson 1970–2012 keskimääräinen lämpötila (2,8 °C), arvioksi vuoden 2013 Suomen sähkönkulutuksen suuruudelle saadaan 84947 GWh, jolloin arvion virhe on -1081 GWh (-1,3 %). Arvio on toteutunutta kulutusta suurempi vuoden 2013 tavanomaista korkeamman keskilämpötilan takia. Kun regressiomallissa käytetään toteutunutta keskilämpötilaa (5,7 °C) [33], arvioi regressiomalli vuoden 2013 kulutukseksi 77769 GWh. Lämpötilan korjaaminen saa arvion laskemaan selvästi alle toteutuneen kulutuksen.

Suomen bruttokansantuotteen kehityksestä löytyy ennusteita muutaman vuoden päähän. Esimerkiksi Suomen Pankki julkaisee kaksi kertaa vuodessa ennusteensa Suomen talouden kehityksestä. Alkuvuonna julkaistava ennuste ulottuu kolmen vuoden päähän ja loppuvuonna julkaistava ennuste 2,5 vuoden päähän. Aikavälillä 2003–2013 julkaistujen ennusteiden osuvuutta tarkasteltiin ennustevirheiden itseisarvojen keskiarvon avulla. Tarkastelun tuloksia on esitelty taulukossa 9.1.

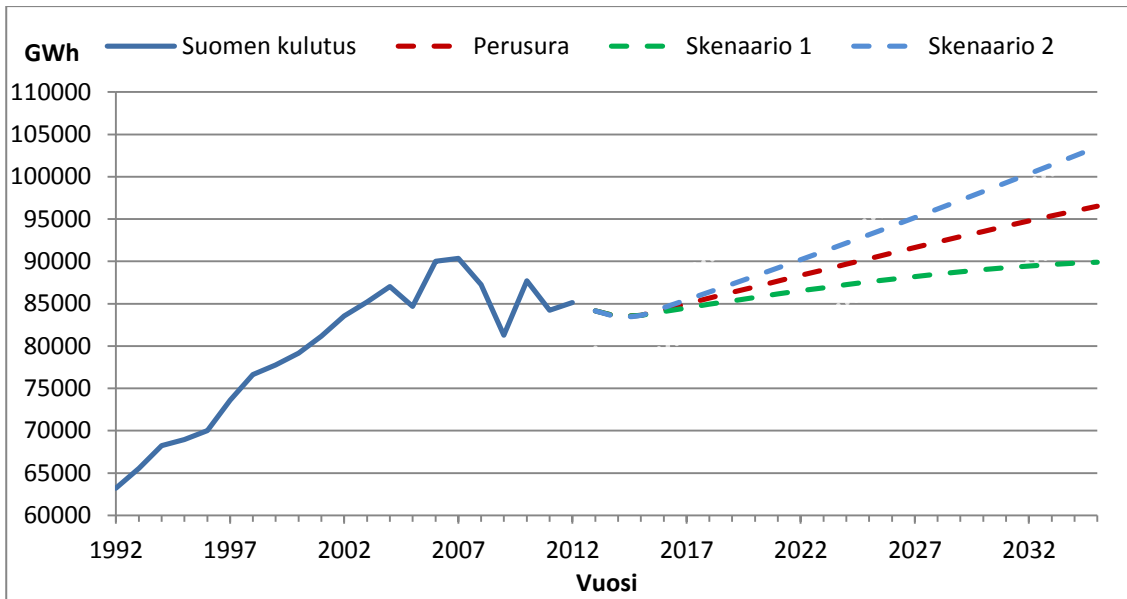
Taulukko 9.1. Suomen Pankin talousennusteiden osuvuus aikavälillä 2003–2013

Ennusteen pituus (vuotta)	Virheen itseisarvo keskimäärin [%-yksikköä]
0,5	0,9
1	1,2
1,5	1,5
2	2,5
2,5	2
3	2

Vuoden 2009 suurta pudotusta ei osattu ennustaa missään tutkimuslaitoksessa. Suomen Pankin vuoden 2008 syyskuussa julkaistu ennuste vuoden 2009 talouskasvulle osoittautui peräti 9,8 prosenttiyksikköä toteutunutta optimistisemmäksi. Taulukon 9.1 lasketuissa virheiden keskimääräisten arvojen laskennassa ei ollut mukana vuodelle 2009 tehtyjä ennusteita, joiden virheet olivat poikkeuksellisen suuria ja olisivat näin ollen vääristäneet keskimääräisten virheiden arvoja. [40]

Pelkästään talouskasvun ennustamisen epävarmuus heijastuu suoraan myös regressiomallin hyödyntämiseen sähkönkulutuksen ennustamisessa. Epävarmuutta voidaan tarkastella esimerkiksi laskettujen joustojen avulla. Bruttokansantuotteelle laskettiin joustoksi 0,784, jonka mukaan esimerkiksi 1,2 prosenttiyksikön virhe seuraavan vuoden talousennusteessa aiheuttaa 0,9 prosenttiyksikön epävarmuuden sähkönkulutuksen ennusteeseen. Pidemmälle ulottuvissa ennusteissa virheet kertaantuvat ja epävarmuus kasvaa ennustevälin pidentyessä.

Kuvan 9.1 ennustetta voidaan pitää perusuran mukaisena ennusteena, jossa muuttujien trendit jatkuvat samanlaisina myös tulevaisuudessa. Kuvassa 9.2 on esitetty perusuran lisäksi kahdella eri skenaariolla laskettua ennustetta Suomen sähkönkulutuksen kehityksestä.



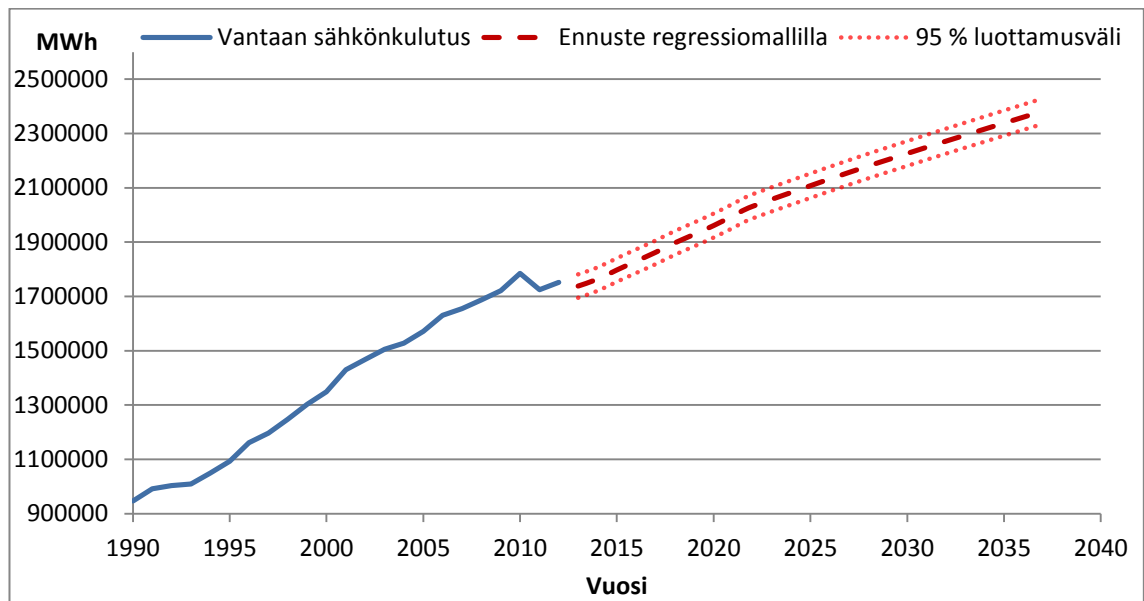
Kuva 9.2. Skenaariot Suomen sähkönkulutukselle

Skenaarioissa on otettu huomioon talousennusteet kahdelle seuraavalle vuodelle, minkä jälkeen skenaario 1 olettaa Suomen talouden kasvavan aiempaa hitaammin, keskimäärin 1,7 % vuodessa. Skenaariossa 2 talouskasvu on puolestaan aiempaa nopeampaa, keskimäärin 2,3 % vuodessa. Skenaarioiden ennustamien sähkönkulutusten erotus vuonna 2035 on noin 13,5 TWh.

9.2 Ennuste Vantaan sähkönkulutukselle

Vantaan sähkönkulutuksen ennusteessa on käytetty ETLAn ennustetta Uudenmaan bruttokansantuotteen kehityksestä vuoteen 2017 asti [27]. Tästä eteenpäin on oletettu, että Vantaan BKT per capita kasvaa keskimäärin 1,1 % vuodessa, mikä on ollut keskimääräinen kasvuvauhti aikavälillä 1988–2012 [30]. Rakennuskannan kohdalla on käytetty Vantaan kaupungin laatimaa ennustetta ja lämpötilan arvona on käytetty mittausjakson 1981–2013 keskilämpötilaa, joka on 5,37 °C.

Ennuste Vantaan sähkönkulutuksen kehityksestä vuoteen 2037 asti on esitetty kuvassa 9.3.



Kuva 9.3. Regressiomallilla tehty pitkän aikavälin ennuste Vantaan sähkönkulutukselle

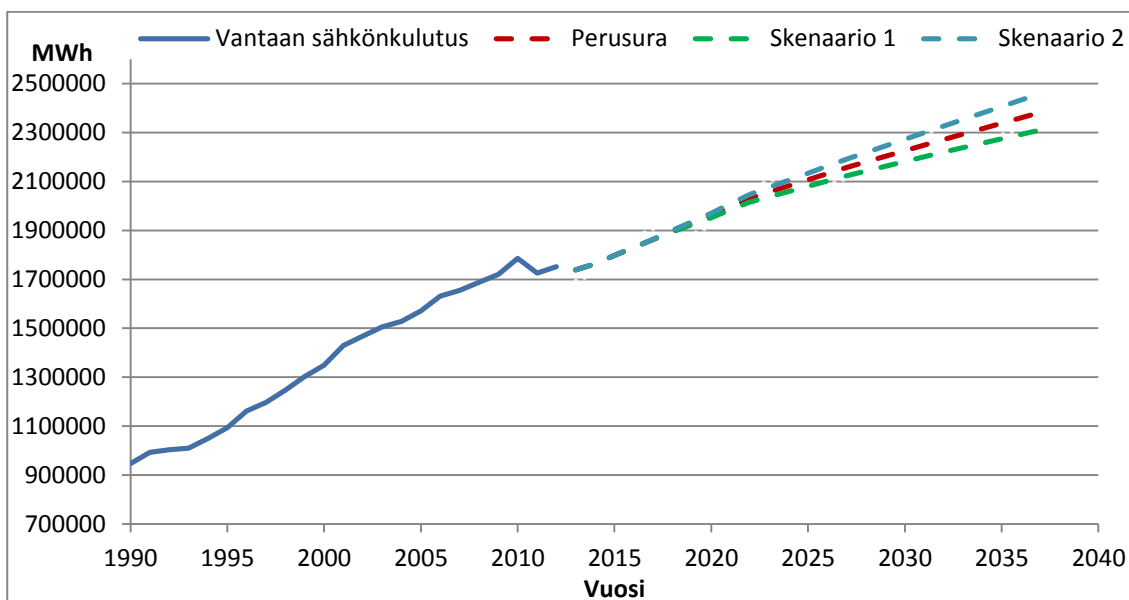
Regressiomallilla tehty ennuste arvioi Vantaan sähkönkulutuksen laskevan hieman vuonna 2013, minkä jälkeen se kasvaisi keskimäärin 1,5 %:n vuosivauhtia vuoteen 2022 asti. Tämän jälkeen kulutuksen kasvuvauhti hidastuisi. Ennusteen mukaan Vantaan sähkönkulutus kasvaisi nykyisestä noin 36 % vuoteen 2037 mennessä. VES:n nykyinen ennustemalli ennustaa samalle aikavälille hieman maltillisempaa 34 %:n kasvua nykyisestä. Regressiomallilla tehty ennuste ja VES:n nykyinen ennuste eivät poikkea paljon toisistaan, koska kummassakin on voimakkaasti läsnä rakennusennusteen vaikutus sähkön kulutukseen. Ennusteen luottamusrajat ovat kapeat, koska mallin selitysaste on hyvin korkea ja jäännöstermien keskihajonta jää suhteellisen pieneksi (12324 MWh).

Vantaan sähkönkulutusmallin ennustuskykyä voidaan myös arvioida lyhyesti pseudo out-of-sample -analyysin avulla. Toteutunut Vantaan sähkönkulutus vuonna 2013 oli 1714311 MWh. Keskilämpötilalla 5,37 °C sekä vuonna 2013 laadittuja talous- ja rakennusennusteiden arvoilla laskettuna ennusteen arvoksi saadaan 1738279 MWh, jolloin virheen suuruus on -23968 MWh (-1,4 %). Ennusteen arvo on siis hieman suurempi kuin toteutunut sähkönkulutus. VES:n nykyisellä ennustemallilla vuonna 2013 laadittu sähkönkulutusennuste vuodelle 2013 on 1765155 MWh, ennusteen jäännöstermin suuruus on -50844 MWh (-3,0 %). Regressiomallilla olisi siis voitu ennustaa vuoden 2013 sähkönkulutus hieman tarkemmin kuin VES:n nykyisellä ennustemallilla. Vuosi 2013 oli Vantaalla selvästi tavanomaista lämpimämpi keskilämpötilan ollessa 6,5 °C [33], mikä pienensi sähkönkulutusta. Jos laskennassa käytettäisiin vuoden 2013 mitattua keskilämpötilaa, saataisiin regressiomallilla lasketuksi Vantaan sähkönkulutuksen arvioksi vuodelle 2013 1713911 MWh ja virhe pienenesi -400 MWh:iin (-0,023 %).

ETLA:n laatimien ennusteiden osuvuutta Uudenmaan bruttokansantuotteen kehityksestä ei voida tässä työssä arvioida, sillä vanhoja ennusteita ei ole vapaasti saatavilla. Kun ennusteessa käytetään vuoden keskilämpötilan keskiarvoa, voidaan lämpötilan vaikutuksen aiheuttamaa epävarmuutta arvioida esimerkiksi sen varianssin avulla. Vantaan keskilämpötilan varianssi aikavälillä 1981–2012 on noin 1 °C, mikä mallin V1 mukaan tuo 21757 MWh:n epävarmuuden ennusteeseen.

Vantaan kaupungin rakennusohjelman toteutumiseen liittyy hieman epävarmuutta, sillä lamavuosien ennusteet ovat olleet optimistisia. Yksittäiselle vuodelle oletetaan rakennusennusteen toteutumisasteen olevan noin 90 %. Ennusteen toteuman virheet yksittäisinä vuosina johtuvat kuitenkin pääosin vain rakennusohjelmassa ilmenneistä viivästyksistä, jotka kompensoituvat seuraavien vuosien aikana. Viimeisen kymmenen vuoden aikana rakennusohjelman kokonaisvaltainen toteutuminen on lähellä 100 %:a.

Kuvassa 9.4 on esitetty perusuran lisäksi kaksi eri skenaariota Vantaan sähkönkulutukselle. Skenaarioissa on otettu huomioon Uudenmaan talousennuste vuoteen 2017 asti.



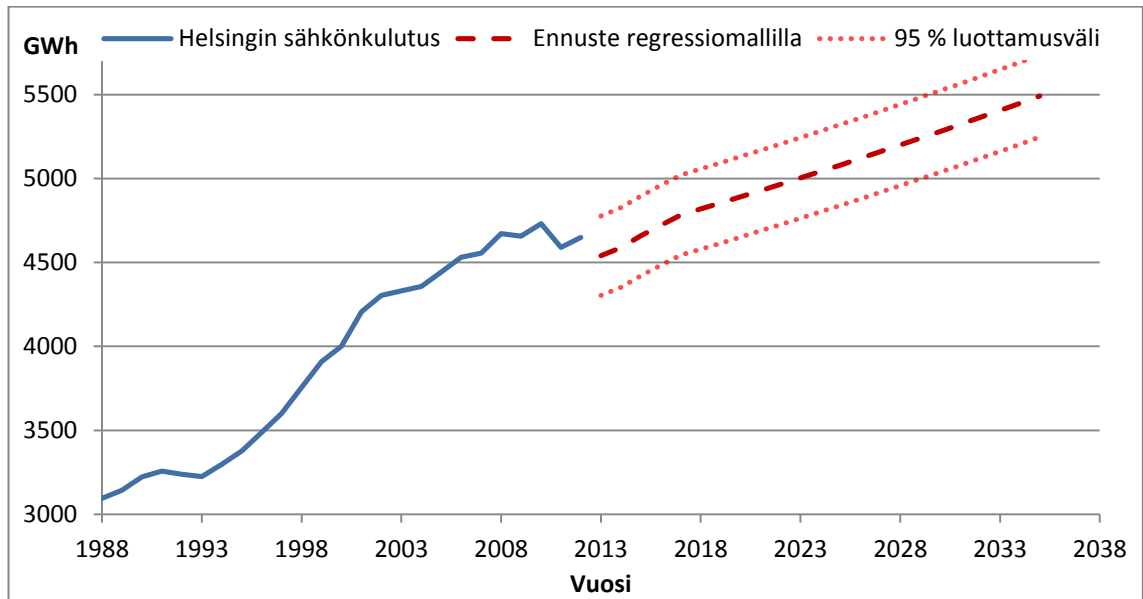
Kuva 9.4. Skenaariot Vantaan sähkönkulutukselle

Skenaario 1 olettaa Vantaan BKT per capita:n kasvavan vuodesta 2017 lähtien aiempaa hitaammin, keskimäärin 0,7 % vuodessa. Skenaariossa 2 kasvu on puolestaan aiempaa nopeampaa, keskimäärin 1,5 % vuodessa. Skenaarioiden aiheuttamat erot Vantaan sähkönkulutukselle ovat havaittavissa, vaikka jäävätkin suhteellisen pieniksi. Vuoden 2037 sähkönkulutusennusteiden ero on noin 150 000 MWh.

9.3 Ennuste Helsingin sähkönkulutukselle

Helsingin sähkönkulutuksen ennusteessa Uudenmaan bruttokansantuotetta koskevat oletukset ovat samat kuin Vantaan ennusteessa. Lämpötilan arvona on käytetty mittaus-

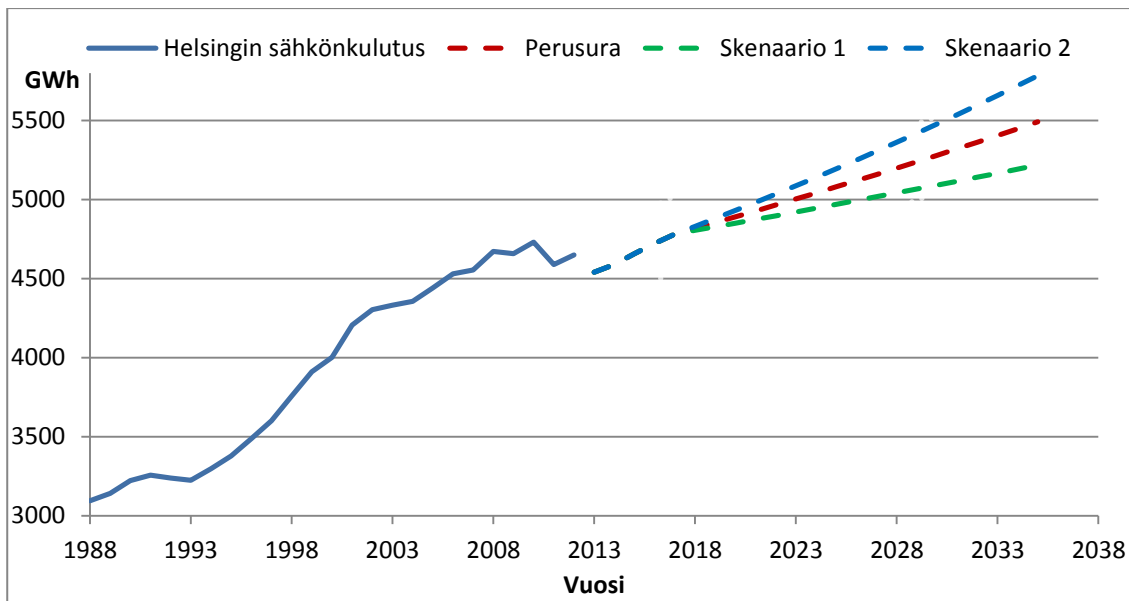
jakson 1970–2012 Helsingin keskilämpötilaa 5,7 °C. Ennuste Helsingin sähkönkulutuksen kehittymisestä on esitetty kuvassa 9.5.



Kuva 9.5. Regressiomallilla tehty pitkän aikavälin ennuste Helsingin sähkönkulutukselle

Regressiomallilla tehty ennuste arvioi Helsingin sähkönkulutuksen laskevan vuonna 2013. Tämän jälkeen sähkönkulutus lähtisi kasvuun alueen talouskasvun mukana. Kasvu olisi keskimäärin 1,3 % vuodessa vuoteen 2017 saakka, jonne asti alueen talouskasvuennuste ulottuu. Tämän jälkeen kasvun on oletettu hieman hidastuvan ollen keskimäärin 1,1 % vuodessa. Helsingin ennusteessa on leveä luottamusväli mallin jäännöstermien keskihajonnan ollessa 80,1 GWh. Vuonna 2013 Helsingin sähkönkulutus oli 4540 GWh [41]. Regressiomallilla laskettu arvio vuodelle 2013 on noin 4541 GWh, joka on lähes täysin sama kuin toteutunut arvo.

Kuvassa 9.6 on esitetty perusuran lisäksi kaksi eri skenaariota Vantaan sähkönkulutukselle. Skenaarioissa on otettu huomioon Uudenmaan talousennuste vuoteen 2017 asti.



Kuva 9.6. Skenaariot Helsingin sähkönkulutukselle

Skenaario 1 olettaa Uudenmaan bruttokansantuotteen kasvavan vuodesta 2017 lähtien aiempaa hitaammin, keskimäärin 0,7 % vuodessa. Skenaariossa 2 kasvu on puolestaan aiempaa nopeampaa, keskimäärin 1,5 % vuodessa. Kuvasta 9.6 voidaan arvioida kansantalouden kehityksen vaikutuksia Helsingin sähkönkulutukseen pitkällä aikavälillä. Skenaarioiden ennusteiden ero vuoden 2035 Helsingin sähkönkulutukselle on noin 566 GWh.

10 EPÄVARMUUSTEKIJÄT

Kulutettava sähköenergian määrä on riippuvainen useasta tekijästä. Tässä työssä on osoitettu esimerkiksi talouden kehityksen, rakennuskannan sekä lämpötilan yhteys sähkönkulutukseen. Tulevaisuudessa voi kuitenkin ilmetä uusia asioita, jotka itsessään vaikuttavat suoraan sähkönkulutukseen tai muuttavat eri tekijöiden merkitsevyyttä. Tässä luvussa käydään lyhyesti läpi eräitä tekijöitä, jotka tulevat todennäköisesti vaikuttamaan sähkönkulutukseen tulevaisuudessa.

10.1 Sähköinen liikenne

Verkosta ladattavien sähköautojen yleistyminen Suomessa näyttää tällä hetkellä hyvin mahdolliselta. Erityisesti sähköisten henkilöautojen yleistyminen voi tulevaisuudessa näkyä sähkönkulutuksen kasvuna. Yleistymisen aikataulun ennustamisessa on turvaututtava erilaisiin selvityksiin. Sähköautojen laajamittaiseen käyttöönottoon vaikuttavat monet tekijät, joista suurimpina kannustimina voidaan pitää EU:n asettamia ilmastotavoitteita sekä polttoainekustannusten nousua. Autonvalmistajat ovat alkaneet panostaa sähköautomallien tuotantoon, minkä johdosta tekniikka kehittyy ja henkilöautoilijoille on tarjolla useita vaihtoehtoja. Suurimpina esteinä sähköautoilun yleistymiselle voidaan pitää autojen rajallista toimintaetäisyyttä ja rajoittunutta pikalatauspaikkojen verkostoa. Lisäksi sähköautojen korkeat hankintakustannukset tekevät monelle sähköautoilun taloudellisesti kannattamattomaksi. [42]

Vuonna 2013 Suomessa oli noin 2,6 miljoonaa henkilöautoa, joista täyssähköautoja oli 169 kpl ja hybridiajoneuvoja 331 kpl. Rekisteröityjä täyssähkö- ja hybridiajoneuvoja oli Suomessa yhteensä 594 kpl. Tällä hetkellä sähköä energianlähteenä hyödynnäviä henkilöautoja on siis häviävän vähän. [43] Niiden yleistymiselle on esitetty esimerkiksi taulukon 10.1 mukaisia skenaarioita.

Taulukko 10.1 Skenaarioita sähköhenkilöautojen yleistymiselle [44]

	Vuosi	Osuus uusista autoista		Kumulatiivinen myyntimäärä (kpl)		Osuus henkilöautojen liikennesuoritteesta	
		PHEV	EV	PHEV	EV	PHEV	EV
Perus-skenaario	2020	10 %	3 %	66 000	13 000	3 %	0,6 %
	2030	50 %	20 %	480 000	160 000	19 %	7 %
Nopea skenaario	2020	40 %	6 %	190 000	26 000	8 %	1 %
	2030	60 %	40 %	960 000	450 000	38 %	19 %
Hidas skenaario	2020	5 %	2 %	38 000	12 000	2 %	0,5 %
	2030	20 %	10 %	207 000	92 000	8 %	4 %

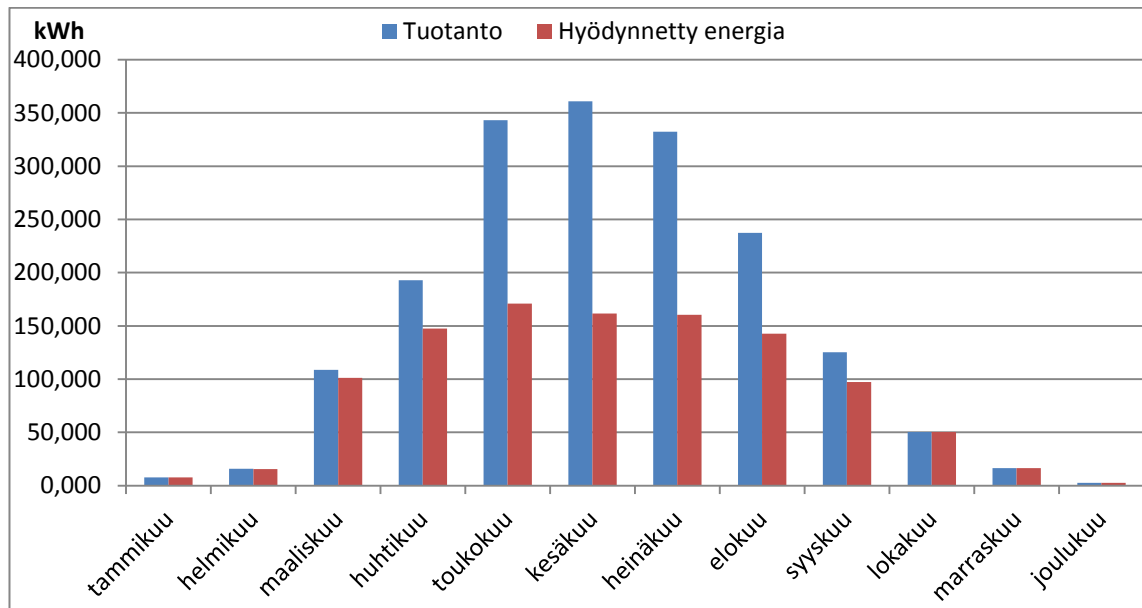
Jos sähköajoneuvon keskimääräiseksi kulutukseksi oletetaan 0,2 kWh/km ja 80 % PHEV-autoilla ajetusta matkasta ajettaisiin sähköllä, kuluttaisi nopean skenaarion mukainen sähköautokanta vuonna 2020 noin 0,6 TWh sähköä vuodessa ja vuonna 2030 noin 3,9 TWh[d]. Tämä tarkoittaisi, että sähköautojen yleistymisen nostaisi sähkönkulutusta vuonna 2020 vain alle prosentin, mutta vuonna 2030 vaikutus voi olla jopa noin 4 %. Laskelmissa on oletettu ajoneuvojen vuotuisiksi ajosuoritteeksi 17 000 km. [44]

Vantaalla oli vuonna 2013 yhteensä noin 109 641 liikennekäytössä olevaa autoa, joista henkilöautoja oli 94 374 kpl (0,46/hlö) [43]. Taulukon 10.1 mukainen nopea skenaario aiheuttaisi noin 1,3 % kasvun Vantaan sähkönkulutukseen vuonna 2020 ja vuonna 2030 vaikutus voisi olla yli 7 %.

10.2 Sähkön pientuotanto

Sähköverkkoyhtiön asiakkaat voivat tuottaa sähköä itse omistamallaan sähkön pientuotantolaitoksilla. Pientuotantolaitoksella tarkoitetaan tässä tapauksessa enintään 50 kVA:n kiinteistökohtaisia järjestelmiä, jotka on sijoitettu lähelle kulutuspistettä ja joiden tarkoitus on tuottaa sähköä ensisijaisesti asiakkaan omaan käyttöön. Järjestelmiä mallinnetaan sähköverkossa negatiivisena kuormana. Pientuotantoa käsiteltäessä huomio kiinnittyy erityisesti pienimuotoiseen aurinko- ja tuulivoimaloihin, jotka on sijoitettu talojen katoille tai pihapiiriin.

Vuonna 2013 VES:issä on tarkasteltu Vantaalla sijaitsevan 2 kW aurinkovoimalan tuottaman sähköenergian määrää. Voimalan sähköntuotanto on esitetty kuvassa 10.1.



Kuva 10.1. Vantaalla sijaitsevan 2 kW aurinkovoimalan energiantuotanto vuonna 2013

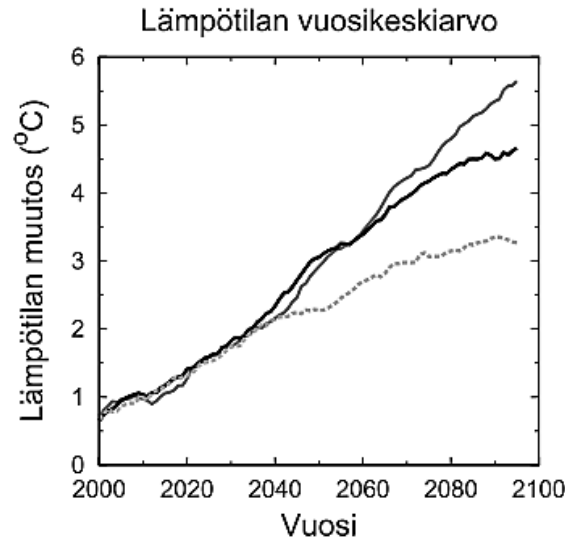
Voimala tuotti sähköä aikavälillä elokuu 2012–elokuu 2013 yhteensä 1794 kWh. Sähköverkkoyhtiön silmin sähkönkulutus laski 1074 kWh, joka on sen osuuden suuruus, jonka kotitalous pystyi itse tuotetusta sähköstä hyödyntämään. Lopun energian voimala syötti sähköverkkoon. Kesäkuukausina voimala pystyi tuottamaan yli 300 kWh, tuosta sähköenergian määrästä asiakas pystyi itse käyttämään vain noin puolet. 71 % voimalan tuottamasta energiasta tuotettiin touko–elokuun aikana, jolloin lämmitykseen ei juurikaan sähköä käytetä. Talvikuukausina voimala pystyi tuottamaan vain murto-osan asiakkaan kuluttamasta sähköstä.

Kaartio on arvioinut mikrovoimaloiden sähköntuotannon potentiaalia Vantaalla. Vaikka mikrovoimaloita otettaisiin käyttöön 20 %:ssa Vantaan asuintaloista, olisi niiden tuottaman sähköenergian määrä vain 0,74 % Vantaan sähkönkulutuksesta. Määrä on hyvin pieni eikä merkittävää pientuotannon yleistymistä ole havaittavissa. Tästä syystä pientuotannon vaikutusten arviointia ei ole syytä ottaa mukaan ennustemalliin. [15]

10.3 Lämmitysenergia

Paljon puhuttu ilmaston lämpeneminen on tilastollinen tosiasia. Vertailukauden 1981–2010 keskilämpötila Suomessa oli lähes 0,4 astetta korkeampi kuin edellisellä vertailukaudella 1971–2000, ja noin 0,7 astetta korkeampi kuin vertailukaudella 1961–1990. [45]

Ilmatieteen laitoksen raportin mukaan Suomen lämpötilan vuosikeskiarvon muutos olisi noin 2 °C vuoteen 2040 mennessä. Vertailukohtana on käytetty vertailukautta 1971–2000. Raportissa esitellyt skenaariot Suomen ilmaston lämpenemisestä on esitetty kuvassa 10.2.



Kuva 10.2. Lämpötilan vuosikeskiarvon muutos -skenaariot aikavälillä 2000–2100 [46]

Raportissa on laadittu ilmastonmuutosennusteita aina vuoteen 2100 asti. Sen mukaan on hyvin todennäköistä, että jo kuluva vuosikymmen on noin 1 °C lämpimämpi kuin vertailukausi 1971–2000 [46]. Talvien keskilämpötilan odotetaan nousevan nopeammin kuin kesien keskilämpötilan ja muutoksen oletetaan olevan voimakkaampi maan pohjoisosissa. Tämä tarkoittaisi, että lämmitykseen kulutettu sähkönkulutus laskisi jonkin verran, mutta etelän suurissa asutuskeskuksissa muutos ei olisi aivan yhtä suurta. Uusien rakennusten paranevan energiatehokkuuden myötä lämmitykseen tarvittava energia pienenee.

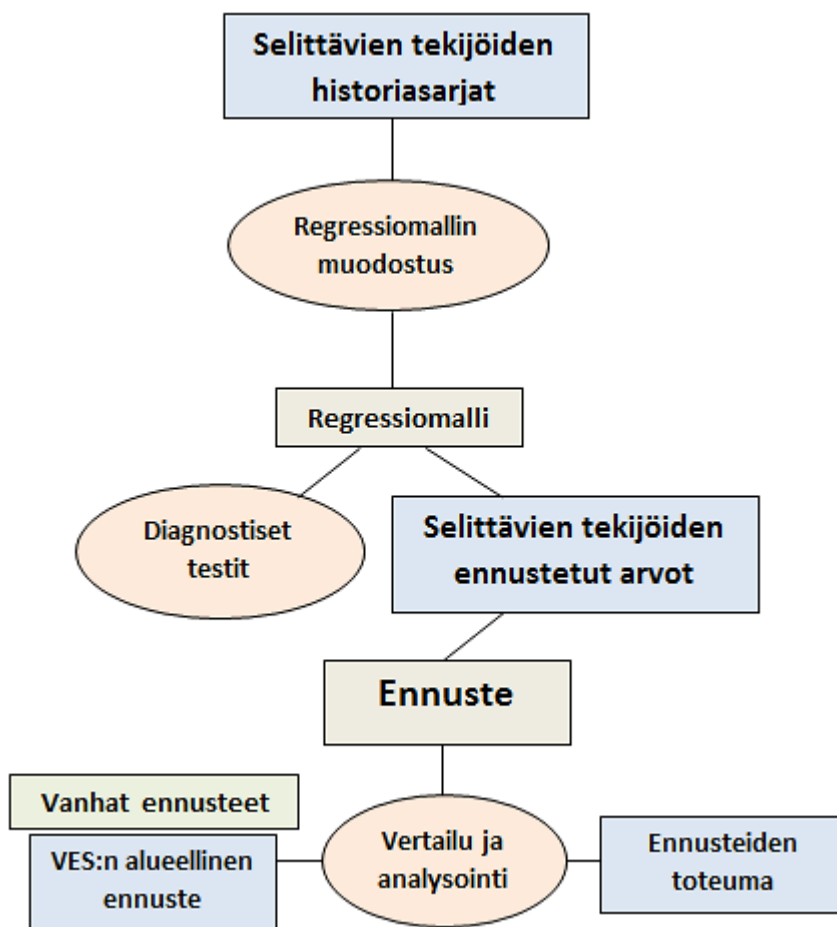
Lämpöpumput ovat yleistyneet kovaa vauhtia viime vuosina. Lämpöpumppujen asentaminen sähkölämmitettyyn rakennukseen pienentää sähkönkulutusta, mutta muihin kuin sähkölämmitteisiin rakennuksiin asennetulla lämpöpumpulla on sähkönkulutusta kasvattava vaikutus. Lämpöpumput ovat pääasiassa ilmalämpöpumppuja, joiden vuotuinen sähkönkulutuksen kasvu Suomessa on aikajaksolla 2006–2011 ollut 37 %. [7] On oletettavaa, että lämpöpumput tulevat kasvattamaan markkinaosuuttaan tulevaisuudessaakin, mutta sähkölämmitys tulee säilyttämään paikkansa yleisimpänä sähkölämmitysmuotona. Käytettävien lämmitysmuotojen yleisyys tulevaisuudessa riippuu esimerkiksi sähkön hinnasta, poliittisista päätöksistä sekä ihmisten arvoista.

Nykyisessä ennustemallissa otetaan huomioon maalämpöpumppujen markkinaosuus rakennettavissa rakennuksissa. Markkinaosuus lasketaan sähkölämmityksen ja kaukolämmityksen osuuksien avulla. Mallissa oletetaan, että noin 30 % Vantaan uusista rakennuksista käyttää lämpöpumppua. Näille rakennuksille on määritelty pienempi omi-

naiskulutuksen arvo, mikä näkyy laskentatuloksissa pienempänä lämmitysenergian kulutuksena. [15]

11 ENNUSTETYÖKALU

Työssä on muodostettu Vantaan sähkönkulutukselle regressiomalli. Luvun 8 tarkastelussa todettiin, että malli kykenee selittämään erittäin hyvin Vantaan sähkönkulutuksen aikaisempaa käyttäytymistä. Luvussa 9 puolestaan arvioitiin mallilla olevan hyvä ennustuskyky vähintään muutamaksi vuodeksi eteenpäin, johon asti Uudenmaan bruttokansantuotteen ennuste ulottuu. Tulosten pohjalta Vantaan Energia Sähköverkot Oy:ssä rakennetaan vanhan ennustejärjestelmän rinnalle toinen ennustetyökalu, joka ennustaa sähkönkulutusta sosioekonomisia muuttujia hyödyntävän regressiomallin avulla. Työkalu toteutetaan taulukkolaskentaohjelmiston avulla, ja sen odotetaan toimivan apuna toiminnan suunnittelussa ja budjetoinnissa. Työkalun rakennetta on havainnollistettu kuvassa 11.1.



Kuva 11.1. Ennustetyökalu kaaviokuvana

Toteutuksessa pyritään pitämään työkalu mahdollisimman yksinkertaisena ja helppokäyttöisenä. Työkalun ytimenä toimii arkisto, joka sisältää useiden eri muuttujien aikasarjoja. Aikasarjojen avulla muodostetaan regressiomalli, jolle työkalu suorittaa oma-toimisesti diagnostiset testit. Testitulosten perusteella arvioidaan mallin tilastollista pätevyyttä.

Muodostettuun regressiomalliin sijoitetaan selittävien tekijöiden ennustetut arvot, jolloin saadaan muodostettua ennusteet Vantaan ja sen sektoreiden sähkönkulutukselle. Ennustetta on hyvä verrata VES:n alueellisen ennustejärjestelmän tuottaman ennusteen kanssa.

On tärkeää, että selittävien tekijöiden aikasarjat pidetään ajan tasalla, jotta myös niiden viimeisimmät arvot olisivat mukana regressiokertoimien laskennassa. Aikasarjojen päivittyessä tulee myös mallien regressiokertoimet laskea uudelleen. Ajan kuluessa mallin regressiokertoimet ja rakenne voivat muuttua. Vanhat mallit säilytetään, jotta tapahuneita muutoksia voitaisiin tarkastella. Tarkastelua ja vertailua voidaan tehdä esimerkiksi kuvaajien ja joustojen avulla. On oletettavaa, että tässä työssä muodostettujen mallien rakenteet tulevat pysymään muuttumattomina useamman vuoden ajan ja muutokset tulevat näkymään ainoastaan mallien kertoimissa.

Sähkönkulutusennusteita verrataan toteutuneisiin kulutuksiin, jolloin voidaan arvioida mallien ennustuskykyä. Myös selittävien muuttujien ennustettuja arvoja verrataan toteutuneisiin arvoihin ja ennusteiden osuvuus tilastoidaan. Tämä auttaa ennusteen luotettavuuden arvioinnissa ja luottamusrajojen luomisessa. Aikasarjojen pidentyessä voidaan muodostetuille malleille suorittaa myös kattavampi pseudo out-of-sample -analyysi.

12 YHTEENVETO

Vuonna 2008 alkanut talouskriisi on vaikuttanut rajusti Suomen talouteen, mikä näkyy erityisesti bruttokansantuotteen laskuna. Vaikutukset ovat näkyneet myös sähkönkulutuksessa, jonka muutokset ovat viime vuosina olleet hyvin samankaltaisia bruttokansantuotteen kanssa. Suuret ja odottamattomat muutokset ovat osoittaneet Vantaan Energia Sähköverkot Oy:ssä käytettävän Vantaan sähkönkulutusta ennustavan järjestelmän heikon sopeutumiskyvyn nopeasti muuttuvassa toimintaympäristössä. Ennustejärjestelmää käytetään apuna sähköverkon yleissuunnittelussa, jonka tarkoituksena on kehittää jakeluverkkoa vastaamaan sähkönkulutuksen tarpeita pitkällä aikajänteellä.

Tässä työssä tutkittiin sosioekonomisten muuttujien hyödyntämistä sähkönkulutuksen ennustamisessa. Tarkoituksena oli luoda lineaarisia regressiomalleja, joiden avulla voitaisiin mallintaa sähkönkulutusta hyvin. Tällöin niitä voisi käyttää hyväksi myös sähkönkulutuksen ennustamisessa. Sosioekonomisten muuttujien hyödyntämisen toisi uuden näkökulman ennustamiseen ja auttaisi reagoimaan suhdanteiden muutoksiin.

Aivan aluksi tutustuttiin aikaisempiin tutkimuksiin, jotka käsittelivät sosioekonomisten muuttujien käyttöä sähkönkulutuksen ennustamisessa. Ilmeni, että regressiomalleilla, joissa selittävinä tekijöinä on käytetty esimerkiksi bruttokansantuotetta ja väkilukua, voidaan mallintaa hyvin sähkönkulutuksen kehitystä.

Luvussa kaksi tutustuttiin Suomen sähkönkulutuksen kehitykseen sekä julkaistuihin pitkän aikavälin ennusteisiin Suomen sähkönkulutuksen kehittymisestä. Suomen sähkönkulutus oli kasvanut kohtalaisen tasaisesti vuoteen 2008 asti, jonka jälkeen kulutus laski huomattavasti vuosina 2009 ja 2010. Tämän jälkeen on tapahtunut pientä, epätaisaista kulutuksen kasvua. Useat eri tahot ovat julkaisseet pitkän aikavälin ennusteita Suomen sähkönkulutuksen kehityksestä. Kaikissa ennusteissa on huomioitu Suomen bruttokansantuotteen vaikutus sähkönkulutukseen, mutta tarkempaa kuvausta käytetyistä malleista ei ole annettu. Ennusteissa on luotu erilaisia skenaarioita, joissa muuttuvina tekijöinä ovat esimerkiksi talouden eri toimialojen kehitys, teknologian kehittyminen sekä uusiutuvan energian käyttö. Ennusteiden perusskenaariot antavat samansuuntaisia arvioita sähkönkulutuksen kehittymisestä, mutta ennusteisiin liittyy suuri epävarmuus, mikä ilmenee suurina marginaaleina. Esimerkiksi VTT ennustaa Suomen sähkönkulutuksen olevan vuonna 2030 79–99 TWh, minkä mukaan sähkönkulutus voisi joko pienentyä 6 % tai kasvaa 18 % nykyisestä tasosta.

Luvussa kolme käytiin läpi regressiomalleihin liittyvää teoriaa siltä osin kuin sitä tässä työssä hyödynnettiin. Lineaarinen regressiomalli on yleisin tilastotieteessä käytettävä malli, jota käytetään kuvaamaan tilastollisia riippuvuuksia. Mallin rakenteen muodostavat selittävät tekijät, joille voidaan laskea kertoimet pienimmän neliösumman me-

netelmällä. Mallin selitystasetta käytetään mallin hyvyyden mittarina, ja se kuvaa kuinka hyvin malli kykenee selittämään selitettävän muuttujan havaittujen arvojen kokonaisvaihtelua.

Keskeisin tehtävä regressiomallin luomisessa on valita malliin sopivimmat muuttujat. Luvussa neljä käytiin läpi mallinvalintatestit, joiden avulla kaikki tilastollisesti merkitsevät selittäjät pyritään valitsemaan malliin. F- ja t-testit ovat hypoteesitestejä, jotka rajaavat pois ne selittäjät ja kokonaiset mallit, jotka eivät kykene selittämään haluttua muuttujaa, tässä tapauksessa sähkönkulutusta. Testit läpäiseviä malleja voi olla useita, joista voidaan valita sopivin mallinvalintakriteerien avulla. Mallinvalintakriteerit suosivat mahdollisimman yksinkertaista mallia, jolla on myös korkea selitystasaste.

Ollakseen käyttökelpoinen on lineaarisen regressiomallin täytettävä riittävässä määrin sitä koskevat standardioletukset. Oletusten täyttymisen varmistamiseksi on mallille tehtävä diagnostiset testit, joissa keskitytään regressiomallin jäännöstermien ominaisuuksiin. Testeissä esiin nousevat epäkohdat on otettava huomioon mallin käytössä. Luvussa viisi nostetaan vielä esille eräitä aikasarjojen käyttämistä koskevia erityispiirteitä kuten viivetermien hyödyntäminen sekä stationaarisuus, joka pitää jokaisen aikasarjan kohdalla varmistaa ADF-testillä. Lisäksi sarjojen yhteisintegroituvuus on testattava AEG-testillä.

Luvussa kuusi esiteltiin regressiomallien muodostamiseen käytettävä aineisto. Esitelyn yhteydessä tarkasteltiin myös aikasarjojen saatavuutta ja ennustettavuutta. Luvussa seitsemän lähdettiin muodostamaan malleja kuvaamaan Suomen, Vantaan ja Helsingin sähkönkulutusta.

Suomen sähkönkulutusta kuvaamaan sopi parhaiten malli, jossa selittävinä tekijöinä oli Suomen bruttokansantuote, sähkön kokonaiskeskihinta sekä vuoden keskilämpötila. Mallin selitystasaste oli yli 97 %. Helsingin sähkönkulutusta kuvaamaan sopi parhaiten malli, jossa selittävinä tekijöinä oli Uudenmaan bruttokansantuote ja vuoden keskilämpötila. Tällä kahden selittäjän mallilla oli yli 98 %:n selitystasaste.

Vantaan sähkönkulutusta kuvaamaan sopi parhaiten malli, jossa selittävinä tekijöinä oli Vantaan bruttokansantuote per capita, Vantaan rakennuskanta sekä vuoden keskilämpötila. Mallin selitystasaste oli 99,8 %. Vantaan sähkönkulutus voitiin jakaa kahteen sektoriin; asumiseen sekä muuhun kulutukseen, joita kuvaamaan luotiin myös regressiomallit. Asumisen kuluttamaa sähköä voitiin kuvata parhaiten pelkän Vantaan väkiluvun avulla, joka itsessään pystyi selittämään yli 97 % Vantaan asumisen kuluttaman sähkön vaihtelusta. Vantaan muuta kulutusta selitti parhaiten malli, jossa selittävinä tekijöinä oli Vantaan bruttokansantuote per capita, Vantaan rakennuskanta ja pääkaupunkiseudun toimistotilojen käyttämättömyystasaste. Mallin selitystasaste oli 99,5 %.

Jokaiselle muodostetulle mallille suoritettiin diagnostiset testit, joissa ei minkään mallin kohdalla ilmennyt suurempia ongelmia. Lisäksi kaikki malleissa käytetyt aikasarjat todettiin stationaarisiksi. Mallien selittäville tekijöille laskettiin joustot log-lineaarisen mallin avulla ja lyhyen ja pitkän aikavälin joustot osittaisen sopeutumismallin avulla.

Luvussa kahdeksan analysoitiin lyhyesti muodostettujen mallien uskottavuutta ja kykyä kuvata sähkönkulutuksia. Yleisesti voidaan todeta, että mallien selityksasteet olivat todella hyviä, mutta mallien muodostamiseen käytettyjen aikasarjojen vaatimaton pituus heikentää hieman mallien uskottavuutta. Sosioekonomisista muuttujista bruttokansantuote, väkiluku, inflaatio ja työttömyysaste osoittautuivat käyttökelpoisiksi selittämään sähkönkulutusta. Parhaissa malleissa kuitenkin vain bruttokansantuotteen eri muodot ja väkiluku olivat mukana. Selittävien tekijöiden viivästettyjä arvoja ei parhaisissa malleissa esiintynyt.

Mallien erinomaisista selityksasteista huolimatta niiden kyky ennustaa sähkönkulutusta pitkällä aikavälillä on rajallinen. Mallien ennustuskyky riippuu suuresti selittävien muuttujien ennustetuista arvoista. Mallista riippuen selittävien muuttujien ennusteet ulottuvat 3–5 vuoden päähän, jonne asti myös malleilla tehtyjä ennusteita sähkönkulutuksesta voidaan pitää jokseenkin luotettavina. Ennustamattomiin talouden muutoksiin eivät myöskään regressiomallit kykene reagoimaan.

Luvussa yhdeksän muodostettiin malleilla ennusteita sähkönkulutukselle. Hitaan ja nopean talouskasvun vaikutuksia sähkönkulutuksen kehitykseen tarkasteltiin yksinkertaisten skenaarioiden avulla. Tuloksista havaittiin kansantalouden pitkäaikaisen kehityksen merkitys sähkönkulutuksen kehitykseen.

Työn lopussa esiteltiin lyhyesti tekijöitä, jotka mahdollisesti tulevat vaikuttamaan sähkönkulutukseen tulevaisuudessa. Liikenteen sähköistyminen voisi joidenkin skenaarioiden mukaan aiheuttaa usean prosentin kasvun vuotuisen sähkönkulutukseen. Pientuotannon vaikutusten puolestaan arvioitiin jäävän marginaalisiksi. Ilmaston lämpeneminen tulee todennäköisesti hieman vähentämään lämmitykseen käytettävän sähkönkulutusta pitkällä aikavälillä. Lisäksi lämmitysenergian kulutus on riippuvainen rakennusten energiatehokkuudesta sekä eri lämmitysmuotojen hyödyntämisestä.

Työn tulosten pohjalta on Vantaan Energia Sähköverkot Oy:ssä tarkoitus kehittää uusi ennustejärjestelmä, aiemman järjestelmän rinnalle, joka tuo toisen, tässä työssä esitellyn näkökulman sähkönkulutuksen ennustamiseen. Järjestelmä toteutetaan taulukolaskentaohjelmana, jonka rakennetta ja toiminnallisuutta esiteltiin lopuksi.

LÄHTEET

[1] L 9.8.2013/588 Sähkömarkkinalaki.

[2] Bianco, V. & Manca, O. & Nardini, S. Electricity consumption forecasting in Italy using linear regression models. *Energy*, 2009. Vol. 34: 9. S. 1413–1421. ISSN 0360-5442.

[3] Mohamed, Z. & Bodger, P. Forecasting electricity consumption in New Zealand using economic and demographic variables. *Energy*, 2005. Vol. 30: 10. S. 1833–1843. ISSN 0360-5442.

[4] Yoo, S. & Lee, J. Electricity consumption and economic growth: A cross-country analysis. *Energy policy*, 2010. Vol. 38: . S. 622–625. ISSN 0301-4215.

[5] Forsström, Stefan. Sosioekonomististen muuttujien soveltaminen sähkönkulutuksen mallintamisessa ja ennustamisessa ekonometrisia menetelmiä hyödyntäen. Kandidaatin-tutkielma. Helsinki 2013. Helsingin yliopisto. Poliitiikan ja talouden tutkimuksen laitos. 42 s.

[6] Suomen virallinen tilasto (SVT): Energian hankinta ja kulutus [verkkójulkaisu]. ISSN=1799-795X. Helsinki: Tilastokeskus [viitattu: 2.4.2014].
Saantitapa: <http://tilastokeskus.fi/til/ehk/index.html>

[7] Adato Energia Oy. Kotitalouksien sähkönkäyttö 2011 [verkkodokumentti]. Julkaistu 26.2.2013 [viitattu 14.1.2014]. Saatavissa:
http://www.tem.fi/files/35856/Kotitalouksien_sahkonkaytto_2011_raportti.pdf

[8] Elinkeinoelämän keskusliitto EK ja Energiateollisuus ry. Arvio Suomen sähkön kysynnästä vuonna 2030 [Verkkodokumentti]. Julkaistu 2009, päivitetty 28.10.2009 [viitattu 12.11.2013]. Saatavissa:
http://www.tem.fi/files/25135/Energian_kysynta_vuoteen_2030_Arvioita_sahkon_ja_energian_kulutuksesta_TEM_EOS_10.11.2009.pdf.

[9] Energiateollisuus ry. Haasteista mahdollisuuksia – sähkön ja kaukolämmön hiilineutraali visio vuodelle 2050 [verkkodokumentti]. Julkaistu 2010, päivitetty 13.2.2011 [viitattu 12.11.2013]. Saatavissa:
http://energia.fi/sites/default/files/haasteista_mahdollisuuksia__ja__hiilineutraali_visio_vuodelle_2050_20091112.pdf

- [10] VTT, Espoo. Low carbon Finland 2050. VTT clean energy technology strategies for society [verkkodokumentti]. Julkaistu 2013, [viitattu 12.11.2013]. Saatavissa: <http://www.vtt.fi/inf/pdf/visions/2012/V2.pdf>
- [11] Työ- ja elinkeinoministeriö. Pitkän aikavälin ilmasto- ja energiastrategia – Valtioneuvoston selonteko eduskunnalle 6. päivänä marraskuuta 2008 [verkkodokumentti]. Julkaistu 2008, päivitetty 16.12.2008 [viitattu 12.11.2013]. Saatavissa: http://www.tem.fi/files/21079/TEMjul_36_2008_energia_ja_ilmasto.pdf
- [12] Työ- ja elinkeinoministeriö. Sähkömarkkinaskaariot vuoteen 2035 [verkkodokumentti]. Julkaistu 2012, päivitetty 5.12.2012 [viitattu 12.11.2013]. Saatavissa: http://www.tem.fi/files/35135/Sahkomarkkinaskaariot_2035_10122012_Final.pdf
- [13] Työ- ja elinkeinoministeriö. Kansallinen energia- ja ilmastostrategia [verkkodokumentti]. Julkaistu 2013, päivitetty 20.3.2013 [viitattu 12.11.2013]. Saatavissa: http://www.tem.fi/files/36221/TEMjul_8_2013_web_20032013.pdf
- [14] Pellervon taloustutkimus. Kansantalouden ja aluetalouden näkökulma. Energia- ja ilmastotiekartta 2050, aloitusseminaari 29.5.2013 [verkkodokumentti]. Julkaistu 2013, päivitetty 10.7.2013 [viitattu 12.11.2013]. Saatavissa: https://www.tem.fi/files/36787/Pasi_Holm.pdf
- [15] Kaartio, Timo. Alueellisen sähkökuorman ennustamisen kehittäminen [verkkodokumentti]. Diplomityö. Espoo 2010 [viitattu 21.11.2013]. Teknillinen korkeakoulu. Sähkötekniikan laitos. 115 s. Saatavissa: <http://lib.tkk.fi/Dipl/2010/urn100210.pdf>
- [16] Partanen, J., Viljanen, S., Lassila, J., Honkapuro, S., Tahvanainen, K. ja Karjalainen, R. Sähkömarkkinat. Lappeenranta 2008. Lappeenrannan teknillinen yliopisto. Opetusmoniste xx s.
- [17] Mellin, Ilkka. Tilastolliset menetelmät [verkkodokumentti]. Opintomoniste. Julkaistu 2008 [viitattu 12.11.2013]. Saatavissa: <https://noppa.aalto.fi/noppa/kurssi/ms-c2128/materiaali>
- [18] Stock, James H. & Watson, Mark M. Introduction to Econometrics. Third edition. Global edition. Essex 2012, Pearson Education Limited. 827 s. ISBN 978-1-4082-6433-1
- [19] Mellin, Ilkka. Todennäköisyysjakaumat [verkkodokumentti]. Luentomateriaali. Julkaistu 2011 [viitattu 15.1.2014]. Saatavissa: https://noppa.aalto.fi/noppa/.../Mat-2_3128_todennakoisyysjakaumat.pdf

- [20] <https://statistics.laerd.com/statistical-guides/img/pearson-10.png>
[viitattu 12.12.2013].
- [21] White, H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 1980. Vol. 48: 4. S. 817–838.
- [22] Durbin, J., Watson, G. S. Testing for Serial Correlation in Least Squares Regression, I. *Biometrika*, 1950. Vol. 37: 3–4. S. 409–428.
- [23] Breusch, T. S. Testing for Autocorrelation in Dynamic Linear Models. *Australian Economic Papers*, 1978. Vol. 17. S. 334–355.
- [24] Shapiro, S.S, Wilk, M. B. An analysis of variance test for normality (complete samples). *Biometrika*, 1965. Vol. 52: 3–4. S. 591–611.
- [25] Dougherty, C. *Introduction to Econometrics*. Fourth edition. New York 2007, Oxford University Press Inc. 573 s. ISBN 978-0-19-956708-9
- [26] Suomen virallinen tilasto (SVT): Kansantalouden tilinpito, Suomen bruttokansantulon menetelmäkuvaus [verkkajulkaisu].
Helsinki 2006 [viitattu:26.2.2014]. Saatavissa:
http://www.stat.fi/til/vtp/2006/vtp_2006_2007-02-08_men_498.html
- [27] ETLA:n alue-ennusteet. [Viitattu 26.2.2014]. Saatavissa:
<http://www2.toimialaonline.fi/>
- [28] Suomen virallinen tilasto (SVT): Väestöennuste [verkkajulkaisu].
Helsinki 2009 [viitattu: 26.2.2014]. Saatavissa:
http://www.stat.fi/til/vaenn/2009/vaenn_2009_2009-09-30_tie_001_fi.html
- [29] Vantaan kaupunki, Tilastot [verkkajulkaisu]. Vantaa 2013 [viitattu 27.2.2014].
Saatavissa: http://www.vantaa.fi/tietoa_vantaasta/tilastot_ja_tutkimukset/tilastot/vaesto
- [30] Helsingin seudun aluesarjat. Helsinki 2013 [viitattu 27.2.2014] Saatavissa:
<http://www.aluesarjat.fi>
- [31] Suomen virallinen tilasto (SVT):Energian hinnat [verkkajulkaisu].
Helsinki 2013 [viitattu: 27.2.2014].
Saatavissa: http://www.stat.fi/til/ehi/2012/02/ehi_2012_02_2012-09-18_tie_001_fi.html

- [32] Energiateollisuus. Helsinki 2013 [viitattu 27.2.2014]. Saatavissa: <http://energia.fi/sahkomarkkinat/sahkon-hinta-ja-sopimukset/mista-sahkon-hinta-muodostuu>
- [33] Suomen virallinen tilasto (SVT). Vuodenaikojen keskilämpötilat [verkkojulkaisu]. Helsinki 2013 [viitattu 27.2.2014]. Saatavissa: [:http://pxweb2.stat.fi/Dialog/varval.asp?ma=010_lampo_tau_101&path=../database/StatFin/ymp/lampo/&lang=3&multilang=fi](http://pxweb2.stat.fi/Dialog/varval.asp?ma=010_lampo_tau_101&path=../database/StatFin/ymp/lampo/&lang=3&multilang=fi)
- [34] Suomen virallinen tilasto (SVT). Kotitalouksien tunnuslukuja 1975-2013 [verkkojulkaisu]. Helsinki 2014 [viitattu 27.2.2014]. Saatavissa: http://pxweb2.stat.fi/Dialog/varval.asp?ma=020_vtp_tau_020&path=../database/StatFin/kan/vtp/&lang=3&multilang=fi
- [35] Suomen virallinen tilasto (SVT): Kuluttajahintaindeksi [verkkojulkaisu]. Helsinki 2014 [viitattu: 27.2.2014]. Saatavissa: http://www.stat.fi/til/khi/2014/01/khi_2014_01_2014-02-19_tie_001_fi.html
- [36] Suomen virallinen tilasto (SVT): Työvoimatutkimus [verkkojulkaisu]. Helsinki 2014 [viitattu: 27.2.2014]. Saatavissa: http://www.stat.fi/til/tyti/2014/01/tyti_2014_01_2014-02-25_tie_001_fi.html
- [37] Helsingin seudun suunnat, Toimitilamarkkinat [verkkojulkaisu]. Helsinki 2014 [viitattu: 27.2.2014]. Saatavissa: <http://www.helsinginseudunsuunnat.fi/graph/hss/hss.aspx?lang=fi&top=2&sub=22#>
- [38] Finanssialan keskusliitto. Suomea koskevat uusimmat talousennusteet [verkkojulkaisu] Helsinki 2014 [viitattu 7.4.2014]. Saatavissa: <http://www.fkl.fi/tilastot/Tilastot/Talousennusteet.pdf>, päivitetty 4.4.2014
- [39] Energiateollisuus. Sähkön hankinta ja kulutus [verkkojulkaisu]. Helsinki 2014 [viitattu 7.4.2014]. Saatavissa: http://energia.fi/sites/default/files/joulukuu_2013.pdf
- [40] Suomen Pankki. Euro ja Talous [verkkojulkaisu] Helsinki 2014 [viitattu 16.4.2014]. Saatavissa: http://www.suomenpankki.fi/fi/julkaisut/euro_ja_talous/talouden_nakymat/Pages/default.aspx
- [41] Helsingin Energia. Helen-konsernin tilinpäätös ja toimintakertomus [verkkodokumentti] Helsinki 27.2.2014 [viitattu 16.4.2014]. Saatavissa: http://www.helen.fi/Documents/Vuosikertomukset%20ja%20osavuosikatsaukset/Helen-konsernin_tilinpaaatos_ja%20toimintakertomus_2013.pdf

- [42] VTT, Espoo. Sähköautojen laajan käyttöönoton skenaarioita ja vaikutuksia sähköjärjestelmään [verkkodokumentti]. Espoo 2010 [viitattu 5.5.2014]. Saatavissa: <http://www.vtt.fi/inf/pdf/workingpapers/2010/W155.pdf>
- [43] Suomen virallinen tilasto (SVT): Moottoriajoneuvokanta [verkkojulkaisu]. Helsinki 2013 [viitattu: 12.5.2014]. Saatavissa: http://www.stat.fi/til/mkan/2013/mkan_2013_2014-03-21_tie_001_fi.html
- [44] Työ- ja elinkeinoministeriö. Sähköajoneuvot Suomessa [verkkodokumentti]. Helsinki 2009, [viitattu 5.5.2014] Saatavissa: https://www.tem.fi/files/24099/Sahkoajoneuvot_Suomessa-selvitys.pdf:
- [45] Ilmatieteen laitos. Ilmastollinen vertailukausi 1981-2010 [verkkojulkaisu]. Helsinki 2014, [viitattu 5.5.2014]. Saatavissa: <http://ilmatieteenlaitos.fi/ilmastollinen-vertailukausi-1981-2010>
- [46] Jylhä, K., Ruosteenoja, K., Räisänen, J., Venäläinen, A., Ruokolainen, L., Saku, S. & Seitola, T.. Arvioita Suomen muuttuvasta ilmastosta sopeutumistutkimuksia varten. ACCLIM-hankkeen raportti 2009 [verkkodokumentti]. Helsinki 2010, [viitattu 5.5.2014] Saatavissa: <https://helda.helsinki.fi/bitstream/handle/10138/15711/2009nro4.pdf?sequence=1>

