

Aalto Yliopisto

MS-C2133 - Operaatiotutkimuksen laboratoriota I

**Työselostus:**  
**Sähköinkulutuksen ennustaminen ja han-**  
**kinnan optimointi**

Timi Turpeinen, 100740833  
Teemu Turpeinen, 100582499

Palautettu: 2. toukokuuta 2024

# Sisällys

<b>1 Johdanto</b>	<b>4</b>
1.1 Motivointi . . . . .	4
1.2 Tehtävän ja työselosten taustoitus . . . . .	4
<b>2 Ennustaminen ja aikasarja-analyysi</b>	<b>4</b>
2.1 Aikasarja-analyysin teoria . . . . .	4
2.2 Data-analyysi . . . . .	6
2.3 Aikasarjamallin muodostaminen . . . . .	9
2.3.1 Kertalukujen identifionti . . . . .	9
2.3.2 Parametrien estimointi . . . . .	10
2.4 Herkkyysanalyysi, diagnostiikka ja ennustaminen . . . . .	12
<b>3 Sähköhankintasuunnitelma ja selvitys</b>	<b>16</b>
3.1 Kaupankäynti ympäristö ja sopimukset . . . . .	16
3.2 Optimoointi . . . . .	17
3.2.1 Tehtävän formalisointi . . . . .	17
3.2.2 Optimoinnin teoria ja tarvittavat työkalut . . . . .	18
3.2.3 Optimoinnin tulokset ja hankintastrategia . . . . .	19
<b>4 Yhteenveto</b>	<b>22</b>
4.1 Johtopäätökset . . . . .	22
4.2 Mahdolliset parannukset . . . . .	22
<b>5 Liitteet</b>	<b>23</b>
<b>A Lähdekoodi aikasarja-analyysille</b>	<b>23</b>
<b>B R-funktioiden selitteet</b>	<b>33</b>
<b>C Excel pohja</b>	<b>33</b>

## Merkinnät ja lyhenteet

(S)ARIMA(X) Integrated Seasonal ARMA-model, Integroitu kausivaihtelu-ARMA-malli

AR Autoregressive, Autoregressiivinen

ARIMA Autoregressive Moving Average Model, Autoregressiivinen liukuvan keskiarvon malli

MA Moving Average, Liukuva keskiarvo

# 1 Johdanto

## 1.1 Motivointi

Sähköntuotannon, hankinnan ja kulutuksen ennustaminen ja optimointi ovat yhteiskunnan ja monien yritysten toiminnan elinehtoja. Energiakustannusten kasvaessa ja kriisitilanteiden kehittyessä on kiinnostus sähkönhintaa ja hankintaa kohtaan noussut sekä kotitalouksilla, että yrityssektorilla. Sähkönhankinnan optimointia voidaan pitää paitsi yleissivistävänä myös arvokkaana valuuttana nykyisillä ja tulevilla työmarkkinoilla. Tässä raportissa muodostetaan ennuste yrityksen sähkökulutukselle ja optimaalinen sähkönhankintastrategia tämän ennusteen pohjalta.

## 1.2 Tehtävän ja työselosteenvausta

Työn tavoitteena oli muodostaa perusteltu ja luotettava ennuste yrityksen sähkökulutukselle yhden vuorokauden ajaksi käyttämällä tilastollista R ohjelmointikieltyä. Tämän ennusteen pohjalta yrityksen sähkönhankintaa optimoitiiin hyödyntämällä Excelin Solver-työkalua, jolla voidaan ratkaista lineaarisia ja epälineaarisia optimointiongelmia.

Ennuste muodostettiin hyödyntämällä yrityksen sähkökulutusdataa ajanjaksoita 31.10. - 3.12.2014 ja ulkolämpötiladataa ajanjaksoita 31.10. - 4.12.2014 [1]. Aikasarjadatan pohjalta suoritettiin perustellinen aikasarja-analyysi, jonka tuloksena saimme muodostettua SARIMAX-mallin sähkökulutuksen ennustamiseksi. Aikasarja-analyysin teoriaa ja analyysin vaiheita käsittelemme työselosten osiossa 2.

Yrityksen sähkönhankinnassa on kaksi merkittävää osa-aluetta: vuorokaudenajat ja kauppangumpaanit. Sähkökauppaa voidaan käydä Nord Poolissa ja sähkömeklarien tai sähkömyyntiyhtiö SVK:n kanssa. Jokaisella kaupankäynnin osa-puolella on solmittuna erilaiset sopimukset, joiden pohjalta sähkönhankitaa pyritään optimoimaan. Sähkönhankintasuunnitelmaa ja kaupankäyntiympäristöä käsittelemme tarkemmin työselosten osiossa 3.

Työselosten osiossa 4 tarkastelemme analyysin onnistumista, ehdotamme tarvittavia parannuksia ja pohdimme erilaisten lähestymistapojen meriittejä.

# 2 Ennustaminen ja aikasarja-analyysi

## 2.1 Aikasarja-analyysin teoria

Tässä työssä tavoitteena oli muodostaa aikasarjamalli sähkökulutukselle, millä voitaisiin muodostaa yhden vuorokauden sähkökulutusennuste. Sähkökulutuksen ennustamisen perusteina käytimme multiplikatiivisia SARIMAX aikasarjamalleja, jotka voidaan esittää analyttisesti mudoissa

$$\Phi_p(B)\Phi_p(B^S)\nabla^d\nabla_S^D z_t = C(B)\nabla^d\nabla_L^D x_t + \Theta_q(B)\Theta_Q(B^S)a_t, \quad (1)$$

missä  $z_t$  on mallitettava suure (sähkökulutus),  $B$  viiveoperaattori ( $Bz_t = z_{t-1}$ ,  $B^S z_t = z_{t-s}$ ),  $\nabla = (1 - B)$ ,  $a_t$  valkoista kohinaa,  $\nabla_S = (1 - B^S)$ ,  $x_t$  selittävä muuttuja (tässä aika-sarja analyysissa ulkolämpötilan vaihtelu),  $C(B)$  viivepolynomi selittävälle muuttujalle ja polynomit  $\Phi_p(B) = (1 - \phi_{1,1}B - \phi_{1,2}B^2 - \dots - \phi_{1,p}B^p)$ ,  $\Theta_q(B) = (1 - \theta_{1,1}B - \theta_{1,2}B^2 - \dots - \theta_{1,q}B^q)$ ,  $\Phi_P(B^S) =$

$(1 - \phi_{1,1}B^S - \phi_{1,2}B^{2S} - \dots - \phi_{1,p}B^{PS}, \Theta_Q(B^S) = (1 - \theta_{1,1}B^S - \theta_{1,2}B^{2S} - \dots - \theta_{1,q}B^{QS}$ , jotka on esitetty aloitusluennon [2] ja työohjeen [1] notaation mukaisesti. Parametri  $S$  kuvaa mallinnettavan suureen ja  $L$  selittävän muuttujan jaksollisuuden arvoa.

Aikasarja-analyysin tavoitteena on mallintaa stokastisen prosessin<sup>1</sup> kehitystä ja ennustaa rittäväällä tarkkudella tämän prosessin tulevaisuuden kehitystä. Stokastisessa prosessissa voidaan havaita erilaisia säännöllisyksia, jotka ovat prosessille ominaisia. Näitä säännöllisyksiä voidaan mallintaa matemaattisesti aikasarjamalleilla. Alla erittelemme SARIMAX aikasarjamallin eri komponentteja.

## Ulkoinen muuttuja

Erityisesti yllä esitetyn mallin (1) polynomit ja viiveoperaattorit kuvaavat prosessissa esiintyviä säännöllisyksiä. Jos näiden säännöllisyksien lisäksi havaitsemme, että stokastinen prosessimme korreloivat voimakkaasti jonkin toisen stokastisen prosessin kanssa, voimme tarkastella näitä prosesseja niiden riippuvuuussuhteiden kautta. SARIMAX -mallissa (1) tätä ilmiötä kuvaava parametri  $x_t$  ja sen viivepolynomi  $C(B)$ , joka kertoo millä viivellä mallinnettavan prosessin suure  $z_t$  korreloivat selittävän prosessin tekijän  $x_t$  kanssa.

## Valkoinen kohina $a_t$

Satunnaisuutta SARIMAX -mallissa kuvaava termi  $a_t \sim WN(0, \sigma^2)$ , jota kutsutaan valkoiseksi kohinaksi<sup>2</sup>. Jos malli on hyvä, residuaalitermi  $a_t$  seuraa valkoisen kohinan jakaumaa. Tällöin termin odotusarvolle pätee  $\mathbb{E}[a_t] = 0$ , varianssi on ajasta riippumaton vakio  $\text{Var}[a_t] = \sigma^2$  ja termin arvot ovat eri ajanhetkillä toisistaan riippumattomia  $\text{Cov}[a_i, a_j] = 0$ , kun  $i \neq j$ .

## Liukuva keskiarvo MA

Aikasarjamallissa satunnaisuutta pyritään residuaalitermin lisäksi mallintamaan liukuvalla keskiarvolla (MA). Liukuva keskiarvo kuvailee aikasarjan nykyarvoon on lineaarisesti riippuvainen residuaalitermin tämän hetkisen ja menneistä arvoista eri painoilla  $\theta_{1,i}$ . Stokastisessa prosessin MA -osa kannattaa ottaa huomioon, jos MA -osan polynomien juuret  $\theta_i$  ovat suuruusluokkaa  $|\theta_i| \geq 1$ . Tällöin juurien arvot ovat painoiltaan merkittäviä prosessin kulkuun nähden. Mallin MA -osa on aina stationaarinen. Mallin MA -osan juurien määrää kuvaavat parametrit  $q, Q$ , joita pyrimme optimoimaan työselosteeseen kohdassa 2.3.2.

## Autoregressio AR

SARIMAX -mallissa (1) stokastisen prosessin jaksollisuutta kuvaava autoregressiivinen osa, jossa stokastisen prosessin nykyistä arvoa mallinnetaan aikasarjan aikasemman datan avulla tietyillä viiveillä  $d$  ja  $D$ . Parametri  $d$  differoi prosessin trendiä ja  $D$  differoi prosessin jaksollista osaa.

Mallin AR -osaa kuvaavat myös polynomit  $\Phi_p(B)$  ja  $\Phi_P(B^S)$ . MA -osaa vastaavalla tavalla, jos polynomien juuret ovat suuruusluokkaa  $|\phi_i| \geq 1$ , kannattaa AR -osa ottaa huomioon aikasarjamallin muodostamisessa. Mallin AR -osan juurien määrää kuvaavat parametrit  $p, P$ , joita pyrimme myös optimoimaan työselosteeseen kohdassa 2.3.2. Differoinnin arvot  $d$  ja  $D$  pyrimme määrittämään aikaisemmin kohdassa 2.3.1. [3] [1]

---

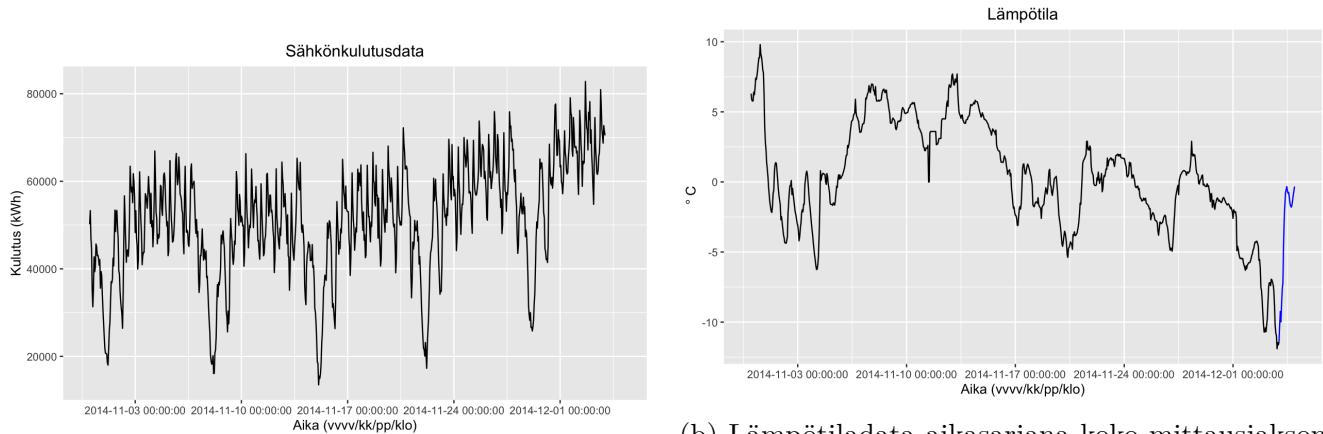
<sup>1</sup>Stokastisella prosessilla kuvataan jatkuva-aikaista satunnaisesta kehittyvää prosessia.

<sup>2</sup>Valkoisen kohinan termi  $a_t$  vastaa resiudaaleja ja pyrkii estimoimaan virhetermejä, joita ei voida todellisuudessa mitata.

Lopullisessa aikasarjamallissa kaikki mallin säännöllisyys ja satunnaisuus tulisi olla kuvattuna mallin (1) mukaisesti eri parametrien  $p, d, q, P, D, Q$  ja selittävän tekijän  $x_t$  arvoilla. Tällöin olemme saaneet kuvattua stokastisen prosessin stationaarisenä prosessina, johon jää vain residuaalien muodostama satunnaisuus.

## 2.2 Data-analyysi

Aikasarja-analyysin pohjana toimi 816 datapistettä tunnittaista sähkökulutusdataa ajanjaksoalta 31.10. – 3.12.2014 ja 840 datapistettä tunnittaista lämpötiladataa ajanjaksoalta 31.10. – 4.12.2014 [1]. Data-analyysin ensimmäisessä vaiheessa suoritimme datalle järkevyystarkastuksen poissulkemalla selkeät mittausvirheet kuten negatiiviset kulutuksen arvot. Selkeitä mittausvirheitä ei esiintynyt datassa, joten siirryimme tarkastelemaan datasta muodostettuja aikasarjoja.



(a) Sähkökulutusdata aikasarjana koko mittausjakson ajalta. x-akselin aika on formatoitu R-ohjelmiston tarjoaman konvention mukaisesti.

(b) Lämpötiladatata aikasarjana koko mittausjakson ajalta. Sinisellä korostettuna ennustettavan ajanjakson lämpötila. x-akselin aika on formatoitu R-ohjelmiston tarjoaman konvention mukaisesti.

Sähkökulutuksen aikasarjasta 1a on havaittavissa, että sähkökulutuksessa esiintyy selkeää viikottaisista ja mahdollista vuorokautista vaihtelua. Aikasarjadatata tarkastellessa selkeät kulutuksen putoamiset esiintyvät alla esitetyn taulukon mukaisina päivinä ja kellonaikoina.

Taulukko 1: Sähkökulutuksen selkeät viikottaiset pudotukset päivän referenssitasosta.

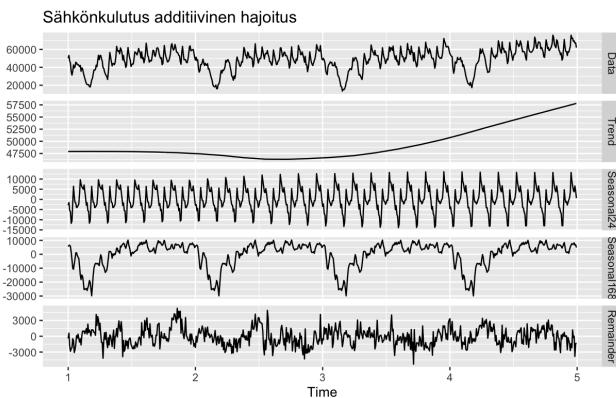
Päivämäärä	Kellonaika	Sähkökulutus (kWh)	Lämpötila (°C)
01.11.2014	05:00:00	18 011	-1.48
08.11.2014	01:00:00	18 199	5.95
15.11.2014	02:00:00	13 487	4.42
22.11.2014	05:00:00	17 281	-0.22
29.11.2014	04:00:00	25 804	-0.23

Taulukossa 1 esiintyvät luvut vastaavat sähkökulutuksen selkeimpiä pudotuksia. On huomattavaa, että pudotukset tapahtuvat lauantaiöisin. Yrityksen toiminta on siis vikkotasolla erittäin jaksollista ja saatta perustua tehdastuotantoon. Vuorokausivaihtelua tarkastellessa huomataan, että sähkökulutus keskimäärin nousee päiväsaikaan, mutta kulutuksen nousu ei ajoitu selkeään kellonaikaan ja nousu saatta venyä aamupäivästä iltaan. Vuorokautiseen vaihteluun tulee siis suhtautua varauksella.

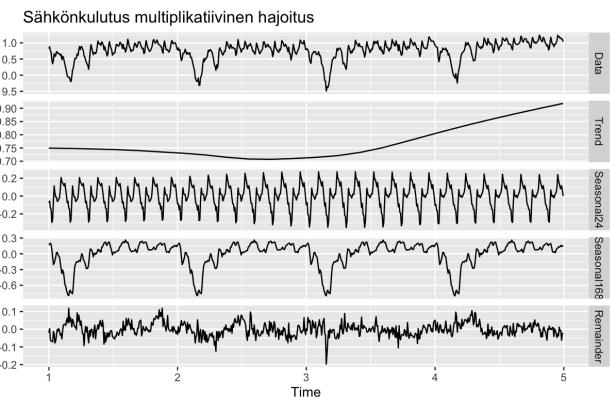
Kuvassa 1b on nähtävissä lämpötiladatan aikasarja, jonka keskiarvo on noin  $0,36$   $^{\circ}\text{C}$ . Täten lämpötilan aikasarja on oleellisesti nollakeskiarvoista ja mahdollista lämpötilariippuvuutta voidaan mallintaa työohjeen [1] mukaisesti kaavalla  $\alpha T_{\text{ulko}}$ .

Kuvissa 1a ja 1b on myös nähtävissä lineaarisia trendejä. Lineaariset trendit ovat vastakkaisia, mikä viittaa negatiiviseen korrelaatioon kyseisten prosessien välillä. Lämpötilan laskiessa voidaan havaita sähkökulutuksen keskimääräistä nousua. Aikasarjoissa ei ole havaittavissa selkeää amplitudin kasvamista tai pienemistä, joten logaritminen skaalaus on tarpeetonta.

Jotta näitä visuaalisia havaintoja voitaisiin varmistaa, voidaan R-ohjelmistolla hajoittaa aikasarjat trendiin, kausittaisiin osiin ja residuaaleihin.



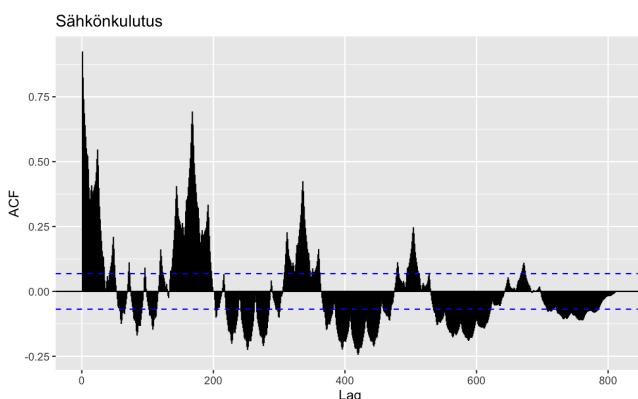
(a) Sähkökulutusdatan hajoitus olettamalla additiiviset vuorovaikutukset.



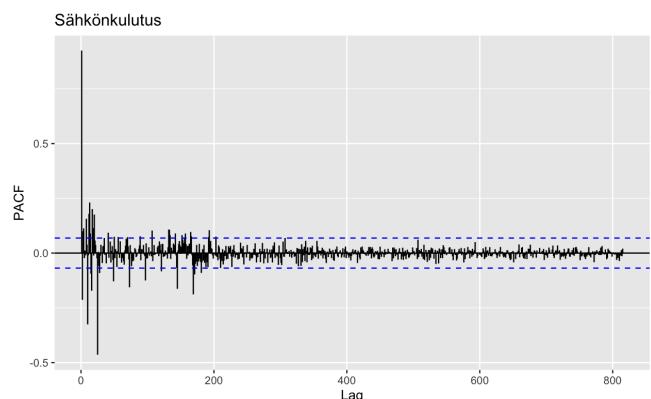
(b) Sähkökulutusdatan hajoitus olettamalla multiplikatiiviset vuorovaikutukset.

Kuvista 2a ja 2b on havaittavissa, että aikasarjat hajoittuvat molemilla oletuksilla vuorokautiseen jaksollisuuteen ja viikottaiseen jaksollisuuteen. Datassa havaittava trendi ei ole täydellisen lineaarinen, mutta etenkin havaintojakson loppupuoliskolla trendi korostuu. Hajoituksen jäännökset muistuttavat valkoista kohinaa, mutta jäännökset jäävät myös kohtuullisen suuriksi.

Hajoitettujen aikasarjojen lisäksi voidaan tarkastella aikasarjojen autokorrelaatiofunktioita ja risikokorrelaatiofunktioita.



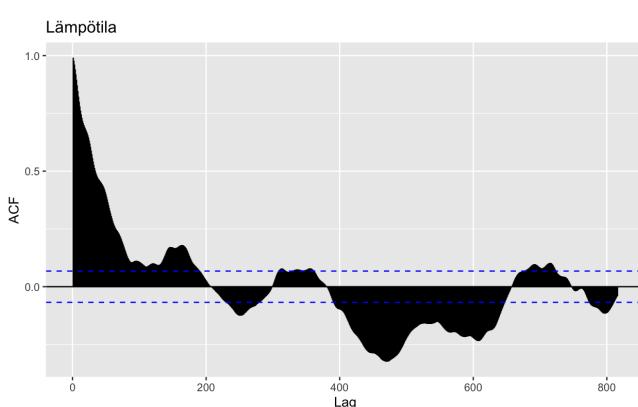
(a) Sähkökulutusdatan autokorrelaatiofunktio.



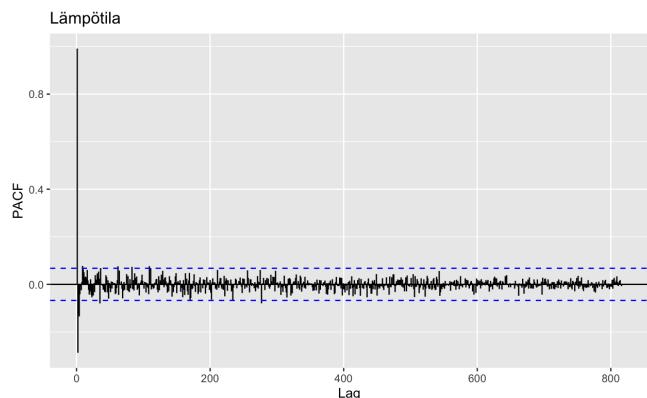
(b) Sähkökulutusdatan osittaisautokorrelaatiofunktio

Sähkökulutuksen autokorrelaatiofunktion kuvaajasta 3a on havaittavissa jaksollisuutta, joka tu-

kee havaintoa aikasarjan jaksollisuudesta. Osittaisautokorrelaatiofunktion kuvaajasta 3b on havaittavissa suhteellisen nopea heikentyminen nollaan.

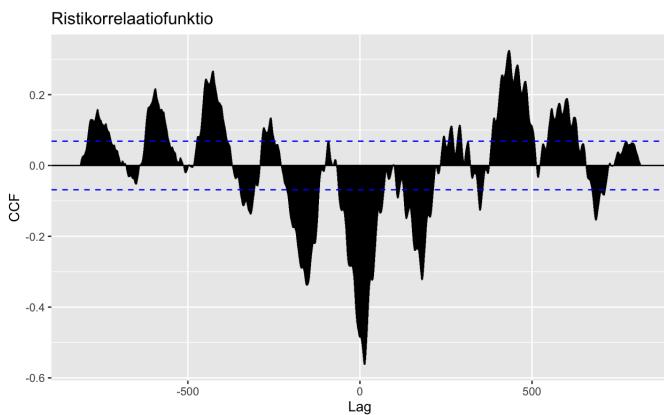


(a) Lämpötiladatan autokorrelaatiofunktio.

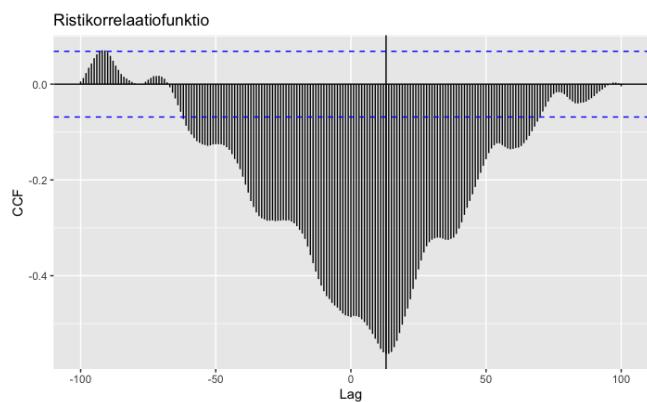


(b) Sähkökulutuksen hajoitus additiivisten suhteiden oletuksella.

Lämpötiladatan autokorrelaatiofunktion kuvaajasta 4a ei ole havaittavissa selkeää jaksollisuutta ja osittaisautokorrelaatiofunktion kuvaaja 4b terminoituu välittömästi nollaan. Nämä havainnot osoittavat, että lämpötiladata muistuttaa valkoistakohinaa. Selkein korrelaatio havaitaan viivellä 1 ja selkeästi rajautuneet positiiviset ja negatiiviset arvot viittaavat lineaariseen trendiin.



(a) Sähkökulutuksen ja lämpötilan ristikorrelaatiofunktio.



(b) Tarkennettu ristikorrelaatiofunktio, jossa merkintä viiveellä  $\text{lag}=13$ .

Ristikorrelaatiofunktiossa on havaittavissa jaksollisuutta ja selkeä negatiivisen ja positiivisen korrelaation muutos. Ristikorrelaatiofunktio saa suurimman amplitudin suhteellisen pienillä viiveillä, jolloin amplitudi on noin  $-0,8$ . Kuvasta 5b havaitsemme, että tällöin viive on noin  $\text{lag}=\{12, 13, 14\}$ . Tulemme siis erityisesti keskittymään näihin lämpötilan viiveisiin analyysissamme.

Yrityksen sähkökulutus on vahvasti jaksollista ja sen suuruusluokka kasvaa tasaisesti lämpötilan madaltuessa, mutta kulutuksen pääpiirteet säilyvät samana. Tämän perusteella sähkökulutuksen kasvaminen perustuu todennäköisesti erilaisten toimitilojen lämmittämiseen ja suurin osa kulutuksesta pohjautuu yrityksen tuottannolliseen toimintaan. Mahdollinen viive voi johtua lämmön varastointumisesta toimitiloihin, jolloin sähkökulutusta mahdollisesti kasvatetaan vasta useamman tunnin kuluttua, kun toimitilan sisälämpötila alkaa laskemaan ulkolämpötilan seuraksena.

Sähkökulutuksesta ja lämpötilasta tehtyjen visuaalisten havaintojen perusteella päädyimme todennäköisesti multiplikatiiviseen aikasarjamalliin, jossa ulkoisen selittäjänä toimii lämpötilan vaihtelu. Täten todennäköinen mallimme on SARIMAX -malli,  $(\text{SARIMAX}(p, d, q)(P, D, Q))$  joka on esitelty tarkemmin osiossa 2.1. Multiplikatiivisuutta voidaan pitää oikeutettuna, sillä havaitsimme sähkökulutuksessa viikottaista ja vuorokautista jaksollisuutta. Koska sähkökulutusdata oli tunnittaista ja havaitsimme erityisen vahvaa viikottaista jaksollisuutta, tulee jaksollisuuden arvoksi todennäköisesti  $S = 7 \cdot 24 = 168$ . Vuorokautista jaksollisuutta voidaan myös pitää oikeutettuna, joten tulemme kokeilemaan aikasarjamallinnusta jaksollisuuden arvoilla  $S = \{24, 168\}$ .

Koska havaitsimme sähkökulutuksessa lineaarista trendiä, voidaan aikasarjan differointia  $d$  pitää perusteltuna. Tähän tulokseen tulee suhtautua varauksella, sillä työohjeen [1] mukaisesti aikasarjamallin signaali-kohina suhde saattaa heikentyä liiallisesta differoinnista. Differointia voidaan perusteella kuitenkin analyyttisesti tarkastelemalla yksinkertaisen AR(1) -mallin juuren  $\phi_1$  arvoa aikasarja-analyysin teorian mukaisesti. Palaamme tähän asiaan kohdassa 2.3.1.

Aikasarjamallissa lämpötilaa voidaan pitää todennäköisenä ulkoisen-selittäjänä kuvasta 5a tehtyjen havaintojen perusteella. Tämä johtopäätös voidaan varmistaa tarkastalemalla aikasarjamallia ilman ulkoista tekijää ja ulkoisen tekijän kanssa. Tällöin ulkoisen-selittäjän merkittävyyttä voidaan tarkastella erilaisilla tilastollisilla testeillä. Tämän analyysin suoritamme myös osiossa 2.3.

## 2.3 Aikasarjamallin muodostaminen

Aikasarjamallin muodostamisessa noudatimme Box-Jenkinsin menetelmää

1. Aikasarjasta muodostetaan stationaarin aikasarja identifioimalla differoinnin kertaluvut.
2. Stationaarisen aikasarjan perusteella määritetään muut aikasarjamallin parametrit.
3. Aikasarjamallin validoinniksi suoritetaan tilastollisia testejä.
4. Jos malli ei läpäise tilastollisia testejä, palaa kohtaan 1 tai 2.
5. Vertaile kaikkia testit läpäisseitä malleja.

### 2.3.1 Kertalukujen identifionti

Data-analyysimme perusteella aikasarjassa esiintyy selkeää jaksollisuutta ja mahdollinen lineaarinen trendi. Varmistaaksemme differoinnin tarpeellisuuden suoritimme kuvaajien ja erinnäisten korrelaatiofunktoiden tarkastelemisen lisäksi työohjeen [1] mukaiset laskennalliset testit.

Lineaarisen trendin differentointi voidaan oikeuttaa tarkastelemalla AR(1) - mallin paramerin  $\phi_1$  arvoa. Jos parametrin arvolle pätee  $|\phi_1| > 0.90$  lineaarista differointia voidaan pitää perusteltuna, sillä aikasarjassa esiintyy selkeää korrelatiota. AR(1) - malli on muotoa

$$(1 - \phi_1 B) z_t = a_t, \quad (2)$$

mikä voidaan myös kirjoittaa SARIMAX -mallina  $\text{SARIMAX}(1,0,0)(0,0,0)_1$ , jossa ulkoista tekijää ei ole.

Vastaavasti kausidifferointia voidaan perustella tutkimalla mallin

$$(1 - \phi_S B^S)z_t = a_t \quad (3)$$

parametria  $\phi_S$ . Kausidifferointia voidaan vastaavasti pitää perusteltuna, kun  $|\phi_S| > 0.90$  ja mallin voi implementoida SARIMAX- mallina  $SARIMA(0,0,0)(1,0,0)_S$ , missä ei esiinny ulkoista tekijää.

Sovitimme molemmat mallit sähkökunkulutusdataan ja saimme seuraavat tulokset.

Taulukko 2: AR(1) ja SAR(1) mallien kertoimien  $\phi_1$  ja  $\phi_S$  arvot.

Parametri	Arvo	Keskivirhe
$\phi_1$	0.9260	0.0131
$\phi_{S=168}$	0.9055	0.0099
$\phi_{S=24}$	0.5810	0.0291

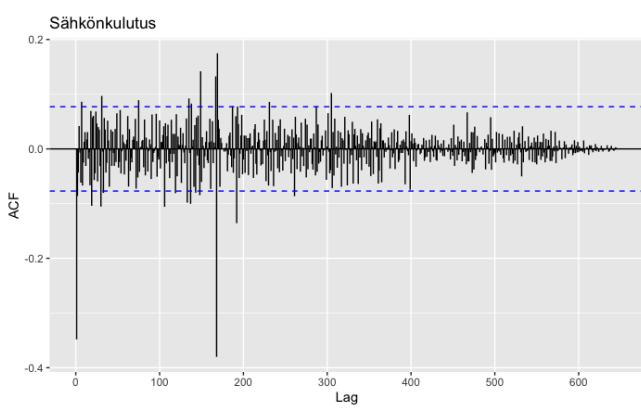
Taulukosta 2 on nähtävissä, että parametrit  $\phi_1$  ja  $\phi_{S=168}$  toteuttavat testin ja täten voimme pitää lineaarista ja jaksollista differointia perusteltuna näissä tapauksissa. Tämä testi osoittaa, että jaksollisuudella  $S = 24$  differointiin tulee suhtautua varauksella.

Kausivaihtelun ja lineaarisen differoinnin kertalukuina käytimme arvoja  $d = \{1, 2\}$  ja  $D = 1$ . Lineaarisen trendin kanssa kokeilimme kahta differennoinnin kertalukua, sillä visuaalisen tarkastelun perusteella epälineaarinenkin trendi voisi olla mahdollinen. Kausivaihtelun pituksina käytimme arvoja  $S = \{24, 168\}$ , jotka perustuvat havaitsemaamme päivittäiseen ja viikottaiseen jaksollisuuteen. On huomattavaa, että tähän mennessä olemme havainneet enemmän viikottaista jaksollisuutta puoltaviaasioita, mutta kattavan analyysin tuottamiseksi kokeilemme myös jaksollisuutta  $S = 24$ . Kokeilimme kausivaihtelulle myös kaksinkertaista differointia molemmilla jaksollisuuskolla, mutta tällöin signaali-kohina-suhde heikentyi merkittävästi ja päätimme jatkaa tarkastelua kertaluvulla  $D = 1$ .

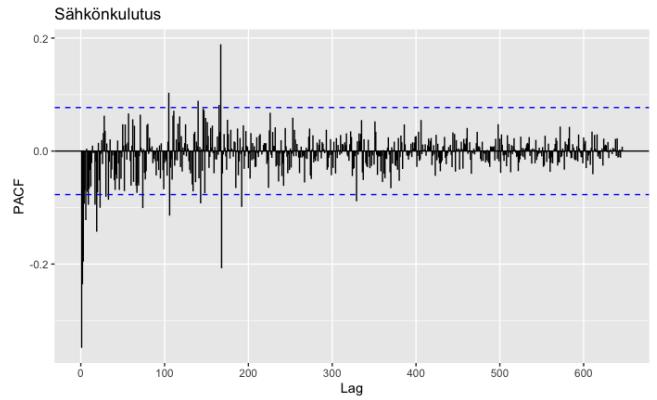
Differointien jälkeen siirryimme tarkastelemaan stationaarisia aikasarjoja, jotta voisimme estimoida aikasarjamallin parametrit ( $p, q, P, Q$ ) ja ulkoisen tekijän mahdollisen viiveen arvon  $L$ . Tämä prosessi oli iteratiivinen ja hyödynsi tilastollisia testejä.

### 2.3.2 Parametrien estimointi

Osioissa 2.2 suorittamamme data-analyysin perusteella päädyimme multiplikatiiviseen aikasarjamalliin, jossa todennäköisesti esiintyy ulkoisena selittäväänä muuttujana lämpötila. Voimme siirtyä havaitsemaan differoitujen aikasarjojen ACF ja PACF funktioita parametrien estimoimiseksi. Epälineaarisen differoinnin kuvaajia emme esitä, sillä molemmat differoinnit muodostivat samankaltaiset ACF ja PACF funktiot. Erittelemme differennoinnin kannattavuutta osiossa 2.4.



(a) Sähkökulutuksen differoidun mallin autokorrelaatiofunktio. Kuvaaja tuotettu komennolla `ggAcf()` ja täten piikki arvolla `lag=0` ei esiinny kuvaajassa.



(b) Sähkökulutuksen differoidun mallin osittaisautokorrelaatiofunktio. Kuvaaja tuotettu komennolla `ggAcf()` ja täten piikki arvolla `lag=0` ei esiinny kuvaajassa.

Kuvasta 6a havaitsemme, että differoinnin jälkeen mallin autokorrelaatiofunktio on lähes statioonarinen. Autokorrelaatiofunktion ensimmäisten viiveiden jälkeen aikasarja vaimenee välittömästi nollan virherajoihin. Osittaisautokorrelaatiofunktion kuvaajasta 6b havaitaan eksponentiaalista vaimenemista ja muutamia merkittäviä piikkejä.

Havaitsemme myös, että molemmat korrelaatiofunktiot saavat merkittäviä piikkejä viiveen arvolla  $lag=168$ , mikä vastaa viikottaisista viivettä. Aikasarjoilla on siis merkittävä viikottainen korrelaatio, mikä viittaa jaksollisuuden arvoon  $S = 168$ .

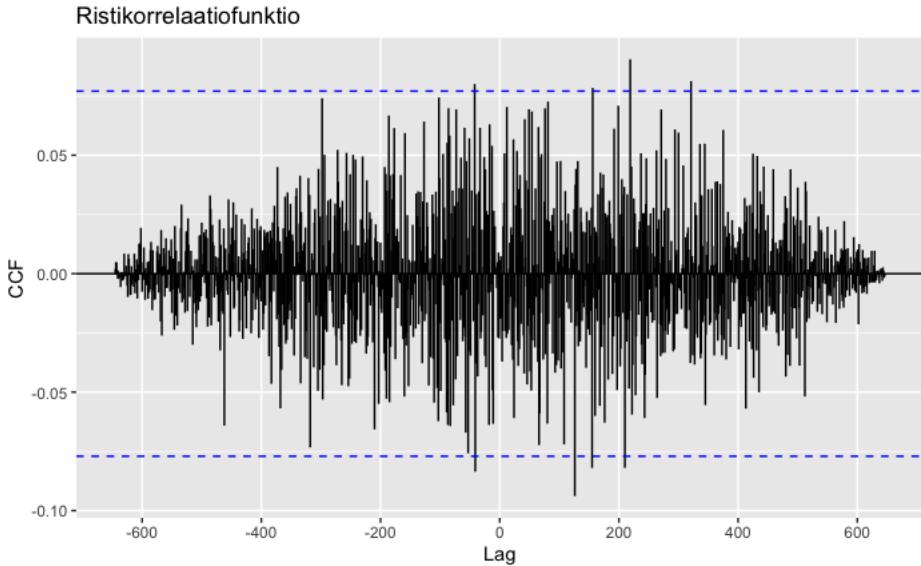
**AR( $p$ )** Parametrin  $p$  arvo voidaan päätellä tarkastelemalla kuinka monta merkittävää piikkiä havaitaan osittaisautokorrelaatiofunktion kuvaajassa ennen kuin piikeistä tulee merkityksellömiä. Tarkastelemalla kuvaajaa 6b parametrin  $p$  mahdolliksi arvoiksi saadaan  $p \in \{1, 2, 3, \dots, 6\}$ .

**MA( $q$ )** Parametrin  $q$  arvo voidaan  $p$  arvoa vastaavalla tavalla päätellä tarkastelemalla autokorrelaatiofunktion kuvaajaa. Kuvaajassa on havaittavissa merkittäviä piikkejä noin 1-2 kappaletta, jolloin saamme parametrin  $q$  arvojoukoksi  $q \in \{1, 2\}$ .

**SAR( $P$ )** Osittaiskorrelaatiofunktion eksponentiaalinen vaimeneminen jaksolla  $S = 168$  ja autokorrelaatiofunktion suhteellisen nopea terminoituminen viittaavat jaksolliseen SMA malliin [4]. Tällöin parametrille pätee  $P = 0$ .

**SMA( $Q$ )** Parametrin  $Q$  arvo voidaan todennäköisessä SMA( $Q$ ) mallissa päätellä autokorrelaatiofunktion terminoitumisenä nollaan jakson  $Q_s$  välein. Tarkastelemalla kuvaajaa 6a saadaan todennäköiseksi parametrin arvoksi  $Q = 1$ .

Lämpötilasta differoimme lineaarisen trendin, sillä kelit selkeästi kylmenivät. Tämän lisäksi koelimme lämpötilan differoimista vuorokausittaisella jaksovaihtelulla, mutta kaikissa tapauksissa sähkökulutuksen ja lämpötilan ristikorrelaatio säily samanlaisena.



Kuva 7: Differoitujen aikasarjojen ristikorrelaatiofunktio, kun  $d = 1$ ,  $D = 1$  ja  $S_{\text{sähkö}} = 168$ ,  $S_{\text{lämpö}} = 24$ .

Lämpötilan viiveen arvoiksi kokeilimme data-analyysi osiossa muodostettujen havaintojen vuoksi  $L = \{12, 13, 14\}$ , sillä tällöin lämpötilan ja sähkökulutuksen välinen korrelaatio on suurinta.

Alle olemme koonneet mallin parametrien ja kertalukujen mahdolliset arvot.

Taulukko 3: Mahdolliset parametrien arvot ja differoinnin kertaluvut.

Parametri	Mahdolliset arvot	lkm
$p$	{1,2,3,4,5,6}	6
$d$	{1,2}	2
$q$	{1,2}	2
$P$	{0}	1
$D$	{1}	1
$Q$	{1}	1
$S$	{24,168}	2
$L$	{12,13,14}	3

Taulukosta 3 saadaan mahdolisten mallien lukumääräksi  $6 \cdot 3 \cdot 2^3 = 144$ . Todellisuudessa emme tule estimoimaan jokaista näistä malleista, vaan karsimme nopeasti selkeästi huonommat vaihtoehdot.

## 2.4 Herkkyysanalyysi, diagnostiikka ja ennustaminen

Malleilla suoritetut diagnostiset tarkastukset koostuvat lähinnä Ljung-Boxin  $\chi^2$ -testin  $p$ -arvojen tarkastelusta ja AIC pisteytyksistä. Ljung-Boxin  $\chi^2$ -testillä voidaan tarkastella ovatko muodostetun mallin ja todellisen datan väliset residuaalit valkoistakohinaa. Jos residuaalit muistuttavat valkoistakohinaa ja niiden autokorrelaatiot ovat nollia, on malli hyvä, ja kohinassa ei esiinny merkittävää autokorrelaatiota. Ljung-Boxin  $\chi^2$ -testisuure voidaan laskea työhjeeen [1] mukaisesti

kaavalla

$$Q_k = \frac{n \cdot (n+2)}{n-k} \cdot \sum_{i=1}^k r_i^2, \quad (4)$$

missä  $n$  on datapisteiden lukumäärä,  $r_i$  autokorrelaation arvo viiveellä  $i$  ja  $k$  testattavien viiveiden lukumäärä. Testin  $p$ -arvot saadaan  $\chi^2$ -jakaumasta vapausasteella  $k-m$ , missä  $m$  on mallin parametrien lukumäärä. Tarkastelussa tulemme yleisen nyrkkisäännön mukaisesti käyttämään arvoa  $k=20$ .

Testin hypoteeseina toimivat

$$\begin{aligned} H_0 : & \text{ Data on riippumattomasti jakautunut,} \\ H_1 : & \text{ Data ei ole riippumattomasti jakautunut eli datassa esiintyy korrelaatiota,} \end{aligned} \quad (5)$$

joen hyvällä mallilla testi noudattaa nollahypoteesia ja  $p$ -arvot ovat kohtuullisen suuria.

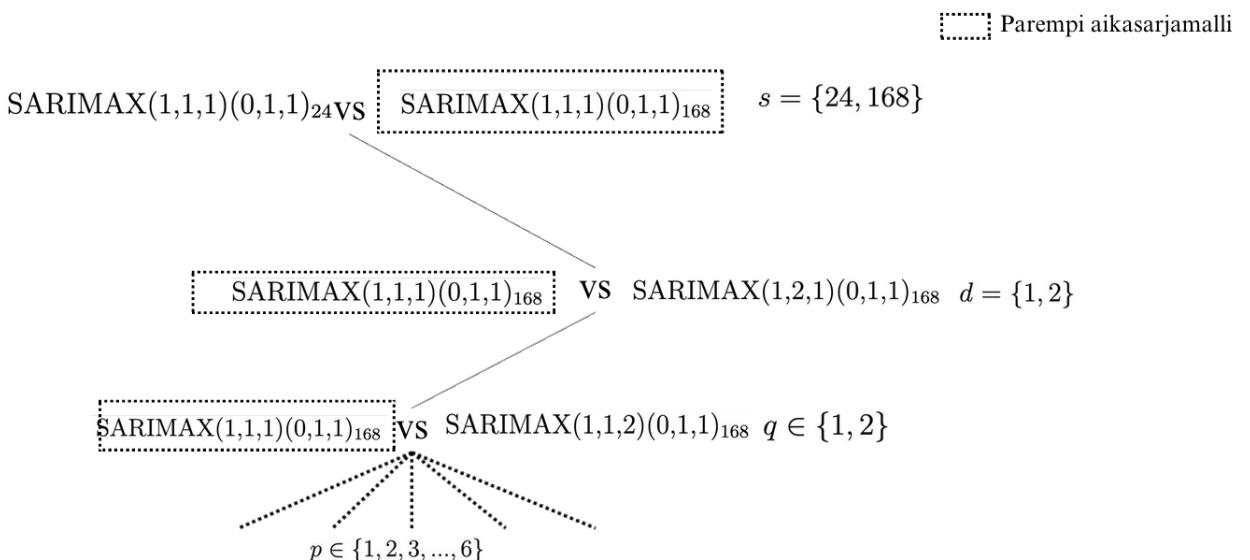
Residuaalien lisäksi tarkastelimme AIC (*Akaike Information Criterion*) pisteytystä. AIC pisteytys välittää ylisovittamista ja liian monimutkaisia malleja rankaisemalla mallin pisteytystä termien lukumäärällä. AIC pisteytyksellä voimme siis ottaa huomioon niukkuusperiaatteen ja mallin ex post -ennusteen tarkkuuden. AIC laksitaan seuraavasti

$$\text{AIC} = -2 \ln(\hat{L}) + 2k, \quad (6)$$

missä  $\hat{L}$  vastaa uskottavuusfunktion (likelihood function) maksimiarvoa ja  $k$  estimoitujen parametrien lukumäärää.

Aloitimme iteratiivisen prosessin määrittämällä optimaalisen differoinnin arvon  $d$  ja jaksollisuuden arvon  $s$  yksinkertaiselle mallille. Näin mahdollisten mallien lukumäärä oli enää 36. Tämän jälkeen optimoimme parametreja  $p$  ja  $q$  ja viimeisenä määritimme lämpötilan mahdollisen viiveen  $L$ . Vaikka mallin muodostaminen oli pääsääntöisesti porrasmaista, kokeilimme myös ajoittain muita kombinaatioita, jotta parametrien poissulkeminen ei johtaisi liian rajattuun ja mahdollisesti hunoon tulokseen.

Kuvasta 8 voi havaita mallin valitsemisprosessin pääpiirteet.



Kuva 8: Päätöspuu mallin muodostamiselle. Aina malleja verratessa otetaan huomioon Ljungin-Box  $\chi^2$ -testi ja AIC arvot.

Yllä esitetyn prosessin tuloksena saimme rajattua mahdollisten mallien joukkoa merkittävästi. Esimerkiksi jaksollisuus ja differennoinnin kertaluvut rajautuivat hyvin nopeasti arvoihin  $s = 168$  ja  $d = D = 1$ .

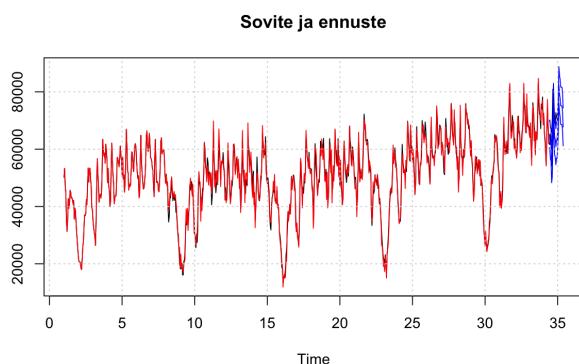
Alla esitämme joitain SARIMAX-malleja. Joukossa esiintyy myös pari referenssitason mallia, joilla muodostimme kokonaiskäsitystä AIC arvon suuruusluokasta.

Taulukko 4: Potentiaaliset SARIMAX-mallit ja niiden diagnostiset arvot. Taulukkoon lisättynä referenssiksi mallit arvoilla  $p = \{5, 9\}$ .

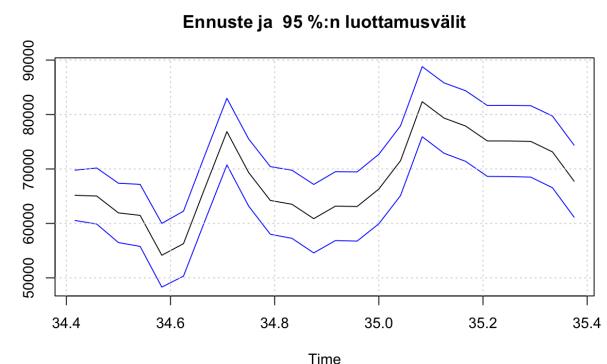
Malli	$L$	Ljung-Box testin $p$ -arvo	AIC
SARIMAX(9,1,1)(0,1,1) <sub>168</sub>	12	0.3998	11669
SARIMAX(5,1,1)(0,1,1) <sub>168</sub>	12	0.1684	11863
SARIMAX(1,1,1)(0,1,1) <sub>168</sub>	12	0.00129	11675
SARIMAX(4,1,2)(0,1,1) <sub>168</sub>	14	0.0604	11632
SARIMAX(4,1,2)(0,1,1) <sub>168</sub>	13	0.2471	11657
SARIMAX(4,1,2)(0,1,1) <sub>168</sub>	12	0.2106	11674
SARIMAX(4,1,1)(0,1,1) <sub>168</sub>	14	0.0250	11633
SARIMAX(4,1,1)(0,1,1) <sub>168</sub>	13	0.2455	11655
SARIMAX(4,1,1)(0,1,1) <sub>168</sub>	12	0.2193	11673
SARIMAX(3,1,1)(0,1,1) <sub>168</sub>	14	0.1763	11636
SARIMAX(3,1,1)(0,1,1) <sub>168</sub>	13	0.1784	11653
SARIMAX(3,1,1)(0,1,1) <sub>168</sub>	12	0.1602	11669
SARIMAX(2,1,1)(0,1,1) <sub>168</sub>	14	0.0307	11636
SARIMAX(2,1,1)(0,1,1) <sub>168</sub>	13	0.0342	11652
SARIMAX(2,1,1)(0,1,1) <sub>168</sub>	12	0.0325	11665

Taulukosta 4 havaitaan, että riittävän suuret  $p$ -arvot saavutetaan malleilla, jossa termien lukumäärä on suurempi. Toisaalta, termien lukumäärän kasvaessa AIC arvot kasvavat.

Näiden arvojen perusteella malleista selkeästi parhaimmat AIC arvot saavutetaan lämpötilan viiveellä  $L = 14$ . Suurin osa malleista saa kohtuullisen suuren  $p$ -arvon, joten tarkastelussa korostamme AIC arvon suuruutta. Täten päädyimme malliin SARIMAX(3,1,1)(0,1,1)<sub>168</sub> lämpötilan viiveellä  $L = 14$ . Tällä mallilla on pienin AIC kaikista Box-Ljungin  $\chi^2$ -testin läpäisseistä malleista. Alla on nähtävissä sähkökulutusennuste ja 95% luottamusväli.



(a) Sähkökulutusdata ja sähkökulutusennuste merkittynä sinisellä.



(b) Sähkökulutusennuste ja 95% luottamusväli vuorokaudelle 4.12.2024

Kuvasta 9a on nähtävissä, että mallin ja toteutuneen kulutuksen välillä ei näy merkittäväitä poikkeamia. Sähkönhankintastrategiamme tulee siis pohjautumaan yllä esitettyyn ennusteeseen. Osioissa 3 pureudumme tarkemmin lopulliseen sähkönhankintastrategiaan, sillä hankinnassa voidaan ottaa huomioon erilaisia riskitekijöitä.

### 3 Sähkönhankintasuunnitelma ja selvitys

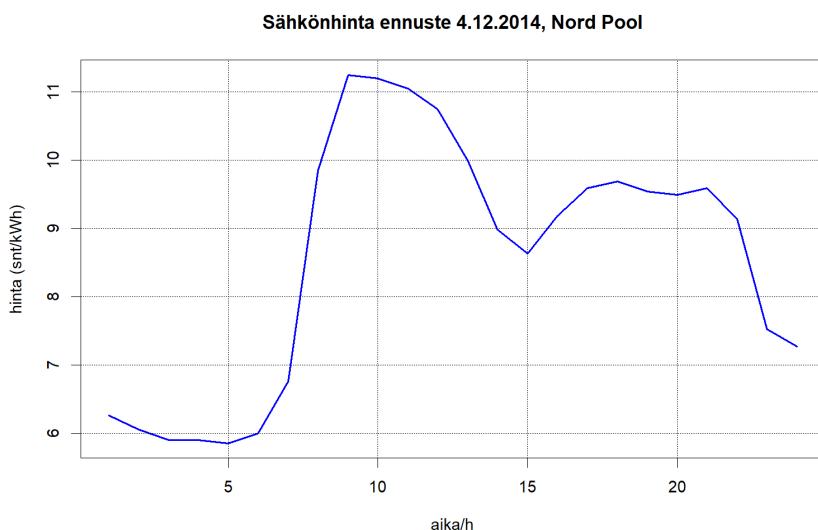
#### 3.1 Kaupankäynti ympäristö ja sopimukset

Kohdassa 2 saimme muodostettua sähkökulutukselle ennusteen vuorokauden 4.12.2014 ajaksi. Pyrimme tämän sähkökulutuksen ennusteen pohjalta optimoimaan kuvitteellisen yrityksen sähkönhankitaa. Tavoitteena on hankkia sähkökulutusennusteen mukaiset määärät sähköä siten, ettei sähkönhankinnan kustannukset minimoituvat. Yrityksen kaupankäynti tapahtuu Suomessa eli sähkötoimitukset tapahtuvat Suomen kantaverkossa. Työohjeen [1] mukaisesti yrityksemme on tehnyt hankintasopimukset kolmen eri osapuolen kanssa.

**1. Nord Pool:** Työohjeen [1] mukaisesti yrityksemme voi hankkia sähköä pohjoimaisesta Nord Poolin Elspot-markkinalta.

Elspot-markkinalla käydään paljon pieniä sähkönhankinnan kauppoja ja hintatasosta ollaa yksimielisiä, jolloin alle 100 000 kilowattitunnin tunnittaisille ostoiille ja myynneille on kiinteät hinnat. Näiden kiinteiden hintojen lisäksi yritys maksaa tai saa välityspreemion  $0,009 \frac{\text{snt}}{\text{kWh}}$ . [1]

Kuvaajassa 10 on nähtävissä Elspot-markkinan kiinteät hinnat vuorokaudelle 4.12.2014.



Kuva 10: Sähkönhinnan ennuste ajalle 24.12.2014 Nord Pool -markkinoilla

Täten Nord Poolin Elspot-markkinoilta ostettaessa sähköön tunnettainen osto- ja myyntihinta määrätyy seuraavasti

- **Ostohinta/kWh:** Tunnittainen sähkönhinta + välityspreemio
- **Myyntihinta/kWh** Tunnin ennustettu sähkönhinta - välityspreemio

Nämä hinta arviot ovat luotettavia vain jos ostettu ja myyty sähkömäärä on vähemmän kuin 100 000 kWh. Koska vaihtoehtoista hinta-arviota ei materiaaleissa ole esitetty, voidaan 100 000 kWh kattoa pitää ehdottomana. Koska kiinteät hinnat ovat varsin epätavallisia pörssiympäristössä, osiossa 3.2.3 erittelemme tarkemmin sähkönhankintastrategiaan kohdistuvia muutoksia, jos epävarmuustekijöitä huomioitaisiin hinnoittelussa.

**2. Meklari:** Työohjeessa [1] on myös määritelty Nord Poolissa toimiva meklari, joka tarjoaa vuorokaiden ajaksi blokkituotteita. Blokkituotteissa yritys sitoutuu ostamaan tunnittaisen va-

kiomääärän sähköä päivisin ja öisin. Vakiomääät määritellään erikseen päiväsajalle ja yöajalle. Sähkönhinta määräytyy seuraavasti

- $\text{Sähköhintapäivä} : 10,00 \frac{\text{snt}}{\text{kWh}}$
- $\text{Sähköhintayö} : 6,00 \frac{\text{snt}}{\text{kWh}}$

Meklarita voi ostaa tunnittain sähköä maksimissaan 50 000 kWh. Jos päädymme ostamaan yhtään sähköä meklarilta, joutuu yritys maksamaan kiinteän 2 000 € sopimuspreemion. Tämä preemio ei riipu sähkönhankinnan määrästä

**3. Sähkömyyntiyhtiö SVK:** Yrityksemme voi myös työohjeen [1] perusteella hankkia sähköä sähkömyyntiyhtiö SVK:n kautta. SVK:n sopimuksessa sähkönhinta määräytyy tunneittain seuraavasti

- $\text{Sähköhintapäivä} : \text{Perushinta } 10,08 \frac{\text{snt}}{\text{kWh}}, \text{ huippuhinta } 10,23 \frac{\text{snt}}{\text{kWh}}$
- $\text{Sähköhintayö} : \text{Perushinta } 5,85 \frac{\text{snt}}{\text{kWh}}, \text{ huippuhinta } 6,06 \frac{\text{snt}}{\text{kWh}}$

Molemmissa tapauksissa perushintaa käytetään, jos tunnittainen sähkön ostomääärä on alle 40 000 kWh. Tämän ostomääärän ylittävästä osasta yritys joutuu maksamaan huippuhintaa.

Kaikilla toimijoille päivääika on määritelty 7.00-22.00 väliseksi ajaksi ja muut vuorokauden tunnit yöjaksi. Sähköä voi siis vain myydä Nord Poolin Elspot-markkinoilla. Yrityksemme ei pysty myöskään varastoimaan ylimääräistä sähköä, vaan ylimäären sähkö on myytävä välittömästi. Ylimääriinen sähkö voidaan myydä Elspot-markkinalle tai tasaajana toimivalle Assivoima Oy:lle. Tasaajana toimivaa Assivoima Oy:tä ei oteta huomioon optimoinnissa. Assivoima Oy toimii todellisen sähkönhankinnan ja kulutuksen tasaajana, jolloin yrityksemme kustannukset määräytyvät sähkönhankintastrategiamme mukaisesti.

## 3.2 Optimointi

### 3.2.1 Tehtävän formalisointi

Tavoitteemme on siis minimoida vuorokauden sähkönhankinta kustannukset, kun otamme huomioon edellä mainitut sopimusehdot ja sähkökulutusennusteemme. Ennen kuin siirrymme optimimaan laskennallisilla työkaluilla, formalisoimme optimointiongelman analyyttisesti. Optimoitavat päätösmuuttujat ja selittävät muuttujat määrittelemme seuraavasti:

- $x_{i,1} :=$  ostetun sähkön määrä Nord Pool Elspot-markkinoilta tunnilta  $i$ .
- $x_{i,2} :=$  ostetun sähkön määrä Nord Poolissa toimivalta meklarilta tunnilta  $i$
- $x_{i,3} :=$  ostetun sähkön määrä sähkömyyntiyhtiö SVK:lta tunnilta  $i$
- $k_i :=$  sähkökulutusennuste tunnilta  $i$ .
- $y_i :=$  myydyn sähkön määrä Nord Pool Elspot-markkinoille tunnilta  $i$
- $c_i :=$  Sähkön hinta Nord Pool Elspot-markkinoilta tunnilta  $i$ .

Parametri  $i$  saa arvoja joukossa  $i \in \{0, 1, 2, \dots, 23\}$ . Merkitsemme parametreilla  $z_{i,1}, z_{i,2}, z_{i,3}$ <sup>3</sup> tunnittaisia sähköhintoja kullekin kaupankäynnin osapuollelle. Saamme hankintasopimusten perus-

---

<sup>3</sup> $z_{i,1}$  vastaa Elspot-markkinoiden hintoja,  $z_{i,2}$  vastaa meklarin hintoja ja  $z_{i,3}$  vastaa SVK:n hintoja.

teella näille tunnittaisille hinnoille seuraavat yhtälöt

$$\begin{aligned} z_{i,1} &= \begin{cases} (c_i + 0.09) \cdot x_{i,1} & x_{i,1} \leq 100\,000 \end{cases} \\ z_{i,2} &= \begin{cases} 10.00 \cdot x_p & x_{i,2} = x_p \leq 50\,000, \text{ kun } i \in \{7, 8, \dots, 22\} \\ 6.00 \cdot x_y & x_{i,2} = x_y \leq 50\,000, \text{ kun } i \in \{23, 0, \dots, 6\} \end{cases} \\ z_{i,3} &= \begin{cases} 10.08 \cdot x_{i,3} & x_{i,3} \leq 40\,000, \text{ kun } i \in \{7, 8, \dots, 22\} \\ 10.08 \cdot 40\,000 + 10.23 \cdot (x_{i,3} - 40\,000) & x_{i,3} \geq 40\,000, \text{ kun } i \in \{7, 8, \dots, 22\} \\ 5.85 \cdot x_{i,3} & x_{i,3} \leq 40\,000, \text{ kun } i \in \{23, 0, \dots, 6\} \\ 5.85 \cdot 40\,000 + 6.06 \cdot (x_{i,3} - 40\,000) & x_{i,3} \geq 40\,000, \text{ kun } i \in \{23, 0, \dots, 6\} \end{cases} \end{aligned} \quad (7)$$

Tunittain Elspot-markkinoille myydyn sähkön kokonaishinnaksi  $a_i$  saamme vastaavasti

$$a_i = (c_i - 0.09) \cdot y_{i,1}, \quad y_{i,1} \leq 100\,000 \quad (8)$$

Näiden yhtälöiden avulla saamme tunnittaiseksi sähkön kokonaishinnaksi

$$z_i = z_{i,1} + z_{i,2} + z_{i,3} - a_i, \quad i \in \{0, 1, 2, \dots, 23\} \quad (9)$$

Kun minimoimme tätä kokonaishintaa, tulee meidän myös ottaa huomioon meklarisopimuksen mahdollinen sopimuspreemio.

$$z = \begin{cases} \sum_{i=0}^{23} (z_{i,1} + z_{i,2} + z_{i,3} - a_i) + 2000 \cdot 100 & (\text{snt}), \quad \text{jos } \sum_{i=0}^{23} x_{i,2} > 0 \\ \sum_{i=0}^{23} (z_{i,1} + z_{i,2} + z_{i,3} - a_i) & (\text{snt}), \quad \text{jos } \sum_{i=0}^{23} x_{i,2} = 0 \end{cases} \quad (10)$$

Tällöin lopullinen optimointiongelma voidaan esittää päätösmuuttujien ja selittävien muuttujien avulla seuraavasti

$$\text{Min } z = \begin{cases} \sum_{i=0}^{23} (z_{i,1} + z_{i,2} + z_{i,3} - a_i) + 2000 \cdot 100 & (\text{snt}), \quad \text{jos } \sum_{i=0}^{23} x_{i,2} > 0 \\ \sum_{i=0}^{23} (z_{i,1} + z_{i,2} + z_{i,3} - a_i) & (\text{snt}), \quad \text{jos } \sum_{i=0}^{23} x_{i,2} = 0 \end{cases}$$

s.e.

$$\begin{aligned} k_i &= x_{i,1} + x_{i,2} + x_{i,3} - y_i, \quad \forall i = 0, 1, 2, \dots, 23 \\ x_{i,1} &\leq 100\,000, \quad \forall i = 0, 1, 2, \dots, 23 \\ y_i &\leq 100\,000, \quad \forall i = 0, 1, 2, \dots, 23 \\ x_{i,2} &\leq 50\,000, \quad \forall i = 0, 1, 2, \dots, 23 \\ x_{i,3} &\leq 70\,000, \quad \forall i = 0, 1, 2, \dots, 23 \\ x_{i,1}, x_{i,2}, x_{i,3}, y_i &\geq 0, \quad \forall i = 0, 1, 2, \dots, 23 \end{aligned} \quad (11)$$

Seuraavassa osiossa pyrimme ratkaisemaan tämän optimointiongelman laskennallisesti. Taustotamme ensiksi optimoinnin teoriaa ja työkaluja, joilla suoritamme lopullisen kustannusten minimoinnin.

### 3.2.2 Optimoinnin teoria ja tarvittavat työkalut

Optimointi ongelman ratkaisemiseksi meidän täytyy tunnistaa ongelman ominaisuudet: Onko ongelma lineaarinen vai epälineaarinen?

Lineaarinen optimointiongelma on muotoa

$$\begin{aligned} \text{Max } z &= c^T x \\ Ax &\leq b \\ x &\geq 0, \end{aligned} \tag{12}$$

missä vektorin  $x$  elementit ovat päätösmuuttuja,  $c$  kustannusvektori ja pistetulo  $c^T x$  tavoitefunktio. Epäyhtälöt  $Ax \leq b$  ja  $x \geq 0$  määrittävät optimointiongelman rajoitteet ja tavoitteena on maksimoida tavoitefunktion arvo. Yksinkertaisia lineaarisia optimointiongelmia voidaan ratkaista esimerkiksi Simplex-algoritilla tai Branch-and-bound-metodilla<sup>4</sup>.

Jos optimointiongelma on epälineaarinen, on mahdollisia ratkaisumenetelmiä useita. Yleisimpiä ratkaisumenetelmiä sileisiin epälineaarisiin ongelmiin ovat esimerkiksi puolittamismenetelmä ja gradienttilaskumenetelmä. Jos ongelma ei ole sileä ja differentiaalien laskeminen on mahdotonta, voidaan ongelma ratkaista evoluutioalgoritmeilla. Alla esitämme yleisen gradienttilaskumenetelmän formaalin esityksen

$$\theta^{n+1} = \Pi_{\Theta}[\theta^n - a_n \hat{\nabla} J(\theta^n)], \tag{13}$$

missä  $\theta^n$  vastaa tämän hetkistä ratkaisua,  $a_n$  askeleen pituutta,  $\Pi_{\Theta}$  projisointia mahdolliseen arvojoukkoon ja  $\hat{\nabla} J(\theta^n)$  gradienttiestimaattia.

Erityisesti Excelin **Solver**-työkalu ratkaisee optimointiongelmia käyttäen Simplex-algoritmia, gradienttilaskua ja evoluutioalgoritmeja.

### 3.2.3 Optimoinnin tulokset ja hankintastrategia

Sähkökulutuksen optimointiongelmamme (11) vaikuttaa lineaariselta ongelmalta, sillä rajoitteet ovat lineaarisia ja päätösmuuttujat ovat jatkuvia. Ongelman mahdollinen arvojoukko on myös konveksi ja ensimmäinen ratkaisumenetelmämme näyttäisi pohjautuvan Excelin Simplex LP -metodiin. Huomaamme kuitenkin, että paloittain määritellyt tavoitefunktio ja päätösmuuttujat tekevät optimointi ongelmasta epälineaarisen. Täten epälineaarinen optimointiongelma voidaan ratkaista käytämällä gradienttilaskua (13). Erityisesti Excelin Solver työkalun ratkaisumetodi **GRG Nonlinear** ratkaisee epälineaarisia optimointiongelmia käyttäen gradienttilaskuun perustuvaa menetelmää. Ratkaisemme siis optimointiongelman Excelissä käyttäen tästä metodia.

Optimointiongelman pohjana toimii muodostamamme sähkökulutusennuste, joka voi tunnettaihin olla liian matala tai suuri suhteessa toteutuneeseen kulutukseen. Työohjeen [1] mukaisesti toteutuneen kulutuksen ja sähköhankinnan erotuksen tasaajana toimii Assivoima Oy. Assivoima Oy ostaa ylijäämää sähköä Nord Poolin tuntihinnasta  $-4\%$  päivääikaan ja  $-3\%$  yöäikaan. Vastaavasti Assivoima Oy myy sähköä Nord Poolin tuntihinnasta  $+5\%$  päivääikaan ja  $+4\%$  yöäikaan. Tämän perusteella sähköhankinnan alijäämästä koittuu enemmän kustannuksia kuin ylijäämästä.

Alijäämäkustannuksista johtuen sähköhankintastrategiamme tulee pohjautumaan muodostamaamme sähkökulutusennusteeseen, jota puskuroimme ylimääräisellä sähköllä. Puskurointi saavutetaan käytämällä sähkökulutusennusteen 55,625% luottamusvälin ylärajajan arvoja. Tämä vastaa

---

<sup>4</sup>Simplex-algoritilla voidaan ratkaista konvekseja ongelmia, missä päätösmuuttujat ovat jatkuvia. Branch-and-bound -metodia käytetään erityisesti, jos päätösmuuttujat ovat diskreettejä.

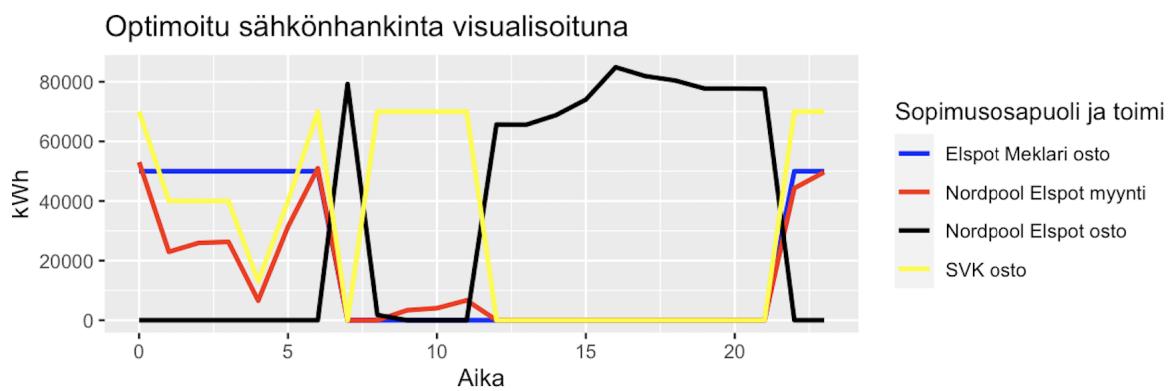
kahdeksasosaa 95% luottamusvälin ylärajan ja ennusteen välisistä arvoista. Tällä saavutimme kohtuullisen puskurin sähköhankinnalle. Sähköhankintastrategiamme on siis riskejä minimoiva, sillä emme laskeneet kaikkea tarkan sähkökulutusennusteen varaan. Strategiassa korostuu sähkön alijäämän välttäminen.

On myös mahdollista, että Nord Poolin sähköpörssin hinnissa esiintyisi epävarmuutta. Tällöin sähköhankintastrategiassa voitaisiin hyödyntää esimerkiksi simulointia. Jokaisen tunnin sähköhintaan voitaisiin mallintaa todennäköisyysjakauumalla, jolloin simuloinnilla voitaisiin mallintaa tuhansia päiviä ja valita kustannukset minimoiva strategia. Tämä strategia olisi laskennallisesti huomattavasti raskaampi verrattuna kiinteiden hintojen optimointiin.

Alla olemme esittäneet optimoinnin tulokset ja havainnollistavan kuvan Excelin työpohjasta.

Taulukko 5: Sähköhankintastrategia jokaiselle tunnille

Tunti	Nordpool Elspot -markkinat osto (kWh)	SVK osto (kWh)	Elspot Meklari osto (kWh)	Nordpool Elspot -markkinat myynti (kWh)
0:00	0	70000	50000	53034,80993
1:00	0	40000,03464	50000	22958,63296
2:00	0	39999,99469	50000	25934
3:00	0	39999,97951	50000	26298,25351
4:00	0	13011,6502	50000	6570,03291
5:00	0	40000,01376	50000	31384,84703
6:00	0	70000	50000	50970,33006
7:00	79251,11062	0	0	0
8:00	1766,718211	70000	0	0
9:00	0	70000	0	3351,988824
10:00	0	70000	0	4044,805555
11:00	0	70000	0	6670,535047
12:00	65643,68092	0	0	0
13:00	65589,7332	0	0	0
14:00	68806,72219	0	0	0
15:00	74020,37603	0	0	0
16:00	84871,22289	0	0	0
17:00	81881,96935	0	0	0
18:00	80416,70549	0	0	0
19:00	77690,96164	0	0	0
20:00	77692,74838	0	0	0
21:00	77625,96906	0	0	0
22:00	0	70000	50000	44285,72318
23:00	0	70000	50000	49666,86583



Kuva 11: Optimoitu sähköhankintastrategia

Taulukosta 5 on nähtävissä tunnittaiset sähköhankintamme ennustettavalle vuorokaudelle. Taulukon arvot selkeästi noudattavat optimaalista sähköhankintaa, sillä esimerkiksi ajankohtina 8:00-11:00 sähköhinta Norpoolin Eslpot -markkinoilla on suurempi kuin SVK:lta ostettu sähköhinta,

joten sähköä kannattaa ensitarpeeseen ostaa sähkömyyntiyhtiö SVK:lta ja tämän jälkeen Nordpoolin Elspot -markkinoilta. Vastaavasti päivisin Meklarilta ostettu sähkö on keskimäärin kalliimpa suhteessa Nord Poolin pörssistä tai SVK:lta ostettu sähkö. Tulokset ovat siis luotettavia ja ratkaisu on optimaalinen.

## 4 Yhteenveton

### 4.1 Johtopäätökset

Tämän työn tavoitteena oli minimoida yrityksen sähköhankintakustannukset muodostamalla luotettava sähkökulutusennuste aikasarja-analyysilla ja optimaalinen sähköhankintastrategia. Päättöksenteon tukena toimi yrityksen sähkökulutusdata, lämpötiladata ja sopimukset eri kaupan-käynnin osapuolien kanssa.

Sähköhankintastrategiamme pohjana toimi aikasarja-analyysilla muodostettu sähkökulutusennuste. Kattavan aikasarja-analyysin, diagnostisten testien ja ennusteiden perusteella parhaiten sähkökulutusta ennusti multiplikatiivinen SARIMAX(3,1,1)(0,1,1)<sub>168</sub> -malli lämpötilan viivellä  $L = 14$ . Aikasarjamallin muodostamisessa otettiin huomioon 35 edellisen päivän sähkökulutus ja ulkolämpötila tältä ja ennustettavan vuorokauden ajalta. Aikasarja-analyysin perusteella lämpötilaa voidaan pitää merkittäväänä tekijänä sähkökulutuksen ennustamisessa.

Sähköhankintastrategian toinen osa perustui sähköhankinnan optimoimiseen pörssihintojen ja sopimusten perusteella. Koska sähkön alijäämästä koitui enemmän kustannuksia, päättimme puskuroida sähkökulutusennustetta. Tätä lopullista sähkökulutusennustetta optimointiin Excelin Solver työkalulla hyödyntäen Excelin gradienttilaskua GRG Nonlinear. Optimoinnin tuloksena vuorokaudessa sähkön ennakkotilauskohtaisesti kokonaiskustannus olisi 141 111,4284 €.

### 4.2 Mahdolliset parannukset

Sähkökulutuksen aikasarjamallinnuksessa esiintyy useita epävarmoja tekijöitä. Jos haluaisimme esimerkiksi ennustaa mallilla pidemmälle tulevaisuuteen, tulisi mallin tunnittaisiin ennustusarvoihin suhtautua varauksella. Toisaalta pitkälle ennustettaessa voidaan havaita suuremman aikavälin trendejä, joiden avulla saadaan osviittaa tulevasta kulutuksesta. Jos mallilla oltaisiin pyritty ennustamaan ensi viikonlopuun kulutus, tulisi ennusteessa ottaa huomioon sähkökulutuksen pohjautuminen yrityksen todelliseen toimintaan, sillä sähkökulutuksessa on havaittavissa selkeimmät kuopat juuri launtaiöisin. Pitkällä aikavälillä malli myös hyötyisi suuremmasta otannasta. Kaikkea ylimmäräistä otantaa ei välttämättä käytettäisi mallin muodostamiseen, sillä mallista tulisi melko monimutkainen, mutta ylimmäräistä dataa voidaan käyttää esimerkiksi ennusteiden validoimiseen.

Sähkökulutuksen ennustamiseen voitaisiin myös ottaa muita lähestymistapoja aikasarjamallinnuksen sijasta. Sähkökulutuksen ennustamiseen voitaisiin käyttää esimerkiksi erilaisia neuroverkkoja tai koneoppimisalgoritmeja, mitkä olisivat erityisesti hyödyllisiä, jos datassa esiintyisi aukkoja.

Yleisesti sähköhankinta on monimuotoinen kokonaisuus, jonka estimointi on todella haastavaa. Intuition, data-analyysin ja matemaattisten testien avulla voidaan kuitenkin muodostaa kannattavia sähköhankintastrategioita. Todellisuudessa sähköhankintaympäristö ja kulutuksen ennustaminen ovat molemmat paljon monimutkaisempia kokonaisuuksia. Tästä huolimatta kirjoittajat ovat päässeet toteuttamaan perusteellista aikasarjamallinnusta ja hyödyntämään optimointityökaluja.

## 6 viitteet

- [1] “Työohje, Sähköhankinnan ennustaminen ja hankinnan optimointi”. *MS-C2133 - Operaatiotutkimuksen laboratoriöt I* (). URL: [https://mycourses.aalto.fi/pluginfile.php/2181077/mod\\_resource/content/9/Ty%C3%B6ohjeTy%C3%B62\\_V2023.pdf](https://mycourses.aalto.fi/pluginfile.php/2181077/mod_resource/content/9/Ty%C3%B6ohjeTy%C3%B62_V2023.pdf).
- [2] “Sähkönkulutuksen ennustaminen aikasarjamallin avulla Sähkön hankinnan optimointi”. *MS-C2133 - Operaatiotutkimuksen laboratoriöt I* (). URL: [https://mycourses.aalto.fi/pluginfile.php/2181081/mod\\_resource/content/4/LabraI\\_Tyo2\\_2024.pdf](https://mycourses.aalto.fi/pluginfile.php/2181081/mod_resource/content/4/LabraI_Tyo2_2024.pdf).
- [3] Gregory C. Reinsel George E. P. Box Gwilym M. Jenkins. *Time series analysis, forecasting and control*. 4. painos. John Wiley Sons, 2008. Luku 9, s. 353–382, 408–411. URL: [https://mycourses.aalto.fi/pluginfile.php/2181079/mod\\_resource/content/2/Time\\_Series\\_Analysis\\_ch9.pdf](https://mycourses.aalto.fi/pluginfile.php/2181079/mod_resource/content/2/Time_Series_Analysis_ch9.pdf).
- [4] *MS-C2128 Prediction and Time Series Analysis, Characteristics of the ARMA models*. URL: [https://mycourses.aalto.fi/pluginfile.php/2150017/mod\\_resource/content/1/lecture-week-4.pdf](https://mycourses.aalto.fi/pluginfile.php/2150017/mod_resource/content/1/lecture-week-4.pdf).

## 5 Liitteet

### A Lähdekoodi aikasarja-analyysille

```
library(ggplot2)
library(forecast)
library(gridExtra)
library(dplyr)
library(tidyr)
library(quantmod) # for getSymbols for share price data
library(ggseas)  # for nzbop data
library(Quandl)
library(TSstudio)
library(LSTS)

alkuPVM <- "2014-10-31 00:00:00"
loppuPVM <- "2014-12-04 23:00:00"
pvm_seq <- seq(as.POSIXct(alkuPVM), as.POSIXct(loppuPVM), by="hour")
pvm <- format(pvm_seq, "%Y-%m-%d %H:%M:%S")
length(pvm)

eletemp = read.table(file = "sahko.csv",
                      sep = ";",
                      dec = ",",
                      skip = 1,
                      col.names = c('kWh','Celcius'))

eletemp$pvm = as.POSIXct(pvm)
eletemp <- data.frame(eletemp)
eletemp

eletemp_c <- eletemp[816:840,]
ele = ts(eletemp$kWh[1:816], start = 1, frequency = 24)
temp = ts(eletemp$Celcius, start = 1, frequency = 24)
```

```

temp816 = ts(eletemp$Celcius[1:816], start = 1, frequency = 24)
temp24 = ts(eletemp$Celcius[816:840], start = c(35,1), frequency = 24)

ele_day = ts(eletemp$kWh[1:100], start = 1, frequency = 24)
ele_week = ts(eletemp$kWh[1:169], start = 1, frequency = 24)
ele_month = ts(eletemp$kWh[1:673], start = 1, frequency = 24)

ele %>%
  tail(24*7*4) %>%
  decompose(type = 'mult') %>%
  autoplot()

temp %>%
  tail(24*7*4) %>%
  decompose(type = 'mult') %>%
  autoplot()

# Voidaan vertailla molempia additiivista mallia

ele %>%
  tail(24*7*4) %>%
  decompose(type = 'add') %>%
  autoplot()

temp %>%
  tail(24*7*4) %>%
  decompose(type = 'add') %>%
  autoplot()

msts_cons <- eletemp$kWh %>% msts( seasonal.periods = c(24, 24*7))
msts_cons %>% head( 24 *7 *4 ) %>% mstl() %>% autoplot()
+ gtitle("Sähkökulutus additiivinen hajoitus")

msts_cons <- log(eletemp$kWh) %>% msts( seasonal.periods = c(24, 24*7))
msts_cons %>% head( 24 *7 *4 ) %>% mstl() %>% autoplot()
+ gtitle("Sähkökulutus multiplikatiivinen hajoitus")

msts_cons <- eletemp$Celcius %>% msts( seasonal.periods = c(1))
msts_cons %>% head( 24 *7 *4 ) %>% mstl() %>% autoplot()

mean(eletemp$Celcius)

ggplot(eletemp, aes(pvm,kWh)) + geom_line() +
  labs(x = "Aika (vvvv/kk/pp/klo)", y = "Kulutus (kWh)") +
  gtitle("Sähkökulutusdata") +
  theme(plot.title = element_text(hjust = 0.5)) +
  scale_x_datetime(date_labels = "%Y-%m-%d %H:%M:%S")

ggplot() +
  geom_line(data = eletemp[1:816,], aes(pvm[1:816],

```

```

Celcius[1:816]), color = "black") +
geom_line(data = eletemp[816:840,], aes(x = pvm, y = Celcius),
           color = "blue") +
labs(x = "Aika (vvvv/kk/pp/klo)", y = expression(~degree~C)) +
ggtitle("Lämpötila") +
theme(plot.title = element_text(hjust = 0.5)) +
scale_x_datetime(date_labels = "%Y-%m-%d %H:%M:%S")

kulutus <- ggplot(eletemp[1:96,], aes(pvm,kWh)) + geom_line() +
  labs(x = "Aika (vvvv/kk/pp/klo)", y = "Kulutus (kWh)") +
  ggtitle("Sähkökulutusdata") +
  theme(plot.title = element_text(hjust = 0.5)) +
  scale_x_datetime(date_labels = "%Y-%m-%d %H:%M:%S")

t <- ggplot() +
  geom_line(data = eletemp[1:96,], aes(pvm,Celcius), color = "red") +
  labs(x = "Aika (vvvv/kk/pp/klo)", y = expression(~degree~C)) +
  ggtitle("Lämpötila") +
  theme(plot.title = element_text(hjust = 0.5)) +
  scale_x_datetime(date_labels = "%Y-%m-%d %H:%M:%S")

grid.arrange(kulutus,t,ncol = 1)

plot(ele)

ele_df <- data.frame(time = time(ele), kWh = ele)
temp_df <- data.frame(time = time(temp816), Celcius = temp816)
temp24_df <- data.frame(time = time(temp24), Celcius = temp24)
ele_day_df <- data.frame(time = time(ele_day), kWh = ele_day)

ggplot(ele_df, aes(x = time, y = kWh)) +
  geom_line(color = "black") +
  labs(x = "aika/vrk", y = "kulutus/kWh") + ggtitle("Sähkökulutusdata") +
  theme(plot.title = element_text(hjust = 0.5))

ggplot() +
  geom_line(data = temp_df, aes(x = time, y = Celcius), color = "black") +
  geom_line(data = temp24_df, aes(x = time, y = Celcius), color = "blue") +
  labs(x = "aika/vrk", y = expression(~degree~C)) + ggtitle("Lämpötila") +
  theme(plot.title = element_text(hjust = 0.5))

ele_acf <- ggAcf(ele, lag.max = 816)
ele_acf <- ele_acf + ggtitle("Sähkökulutus") + xlab("Lag") + ylab("ACF")
ele_pacf <- ggPacf(ele, lag.max = 816)
ele_pacf <- ele_pacf + ggtitle("Sähkökulutus") + xlab("Lag") + ylab("PACF")

temp_acf <- ggAcf(temp, lag.max = 816)
temp_acf <- temp_acf + ggtitle("Lämpötila") + xlab("Lag") + ylab("ACF")
temp_pacf <- ggPacf(temp, lag.max = 816)

```

```

temp_pacf <- temp_pacf + ggtitle("Lämpötila") + xlab("Lag") + ylab("PACF")
ele_acf
ele_pacf
temp_acf
temp_pacf

grid.arrange(ele_acf, ele_pacf, temp_acf, temp_pacf, ncol = 2)

ccf <- ggCcf(ele,temp,lag.max = 100)
ccf <- ccf + ggtitle("Ristikorrelaatiofunktio tarkennettuna") + xlab("Lag")
+ ylab("CCF")
ccf <- ccf + geom_vline(xintercept = 13)
ccf

AR1 <- arima(ele, order=c(1,0,0))
summary(AR1)

AR_seasonal_kk <- arima(ele, order=c(0, 0, 0), seasonal=list(order=c(1, 0, 0),
, period = 7*24))
AR_seasonal_pv <- arima(ele, order=c(0, 0, 0), seasonal=list(order=c(1, 0, 0),
period = 24) )
summary(AR_seasonal_kk)

summary(AR_seasonal_pv)

d = 1 # Differoinnin kertaluku d
S = 24*7 # Kausidifferoinnin jakso S
D = 1 # Kausidifferensoinnin kertaluku D
dele = ele
dtemp = temp

if (d > 0) {
  dele = diff(dele, lag = 1, differences = d)
  dtemp = diff(dtemp, lag = 1, differences = d)
}

if (D > 0) {
  dele = diff(dele, lag = S, differences = D)
  dele = diff(dele, lag = 24,1)
  dtemp = diff(dtemp, lag = 24, differences = 1)
}

dele_acf <- ggAcf(dele, lag.max = 24)
dele_acf <- dele_acf + ggtitle("Sähkökulutus") + xlab("Lag") + ylab("ACF")
dele_pacf <- ggPacf(dele, lag.max = 816)
dele_pacf <- dele_pacf + ggtitle("Sähkökulutus") + xlab("Lag") + ylab("PACF")
Acf(dele, lag.max=816)
acf(dele, lag.max = 816)

```

```

ccf_dele <- ggCcf(dele, dtemp, lag.max = 816)
ccf_dele <- ccf_dele + ggtitle('Ristikorrelaatiofunktio')
ccf_dele
dele_acf
dele_pacf
acf(dele, lag.max=816)
acf(dele, lag.max=816, type = "partial")
acf(dtemp, lag.max=816)
acf(dtemp, lag.max=816, type = "partial")
ccf(dele,dtemp, lag.max=816)

p = 3
q = 1
P = 0
Q = 1

malli = arima(ele,
               order = c(p,d,q),
               seasonal = list(order = c(P, D, Q), period = S),
               method = "CSS")
enne = predict(malli, n.ahead = 24)

# Estimoidaan malli loppumisajan kanssa. Määritetään loppumisajan mahdollinen viive L.
L = 14
tempestimointi = eletemp$Celcius[1:(816-L)]
tempennuste = eletemp$Celcius[(816-L+1):(816-L+24)]
eleestimointi = ts(eletemp$kWh[(1+L):816], start = 1, frequency = 24)
malli2 = arima(eleestimointi,
                order = c(p,d,q),
                seasonal = list(order = c(P, D, Q), period = S),
                xreg = tempestimointi,
                method = "CSS")

enne2 = predict(malli2,
                 n.ahead = 24,
                 newxreg = tempennuste)

mallinSovitus <- function(ele, p, d, q, P, D, Q, s, L) {

  malli <- arima(ele,
                 order = c(p, d, q),
                 seasonal = list(order = c(P, D, Q), period = s),
                 method = "CSS")

  enne <- predict(malli, n.ahead = 24)

  tempestimointi <- eletemp$Celcius[1:(length(ele) - L)]

```

```

tempennuste <- eletemp$Celcius[(length(ele) - L + 1):(length(ele) - L + 24)]
eleestimointi <- ts(ele[(1 + L):length(ele)], start = 1, frequency = s)

malli2 <- arima(eleestimointi,
                  order = c(p, d, q),
                  seasonal = list(order = c(P, D, Q), period = s),
                  xreg = tempestimointi,
                  method = "CSS")

enne2 <- predict(malli2, n.ahead = 24, newxreg = tempennuste)

return(malli2)
}

aic_f <- function(malli) {
  tulos = malli$loglik * -2 + 2 * length(coef(malli))
  return(tulos)
}

# Referenssimallit
sarima911_12 = mallinSovitus(ele, 9, 1, 1, 0, 1, 1, 168, 12)
sarima511_12 = mallinSovitus(ele, 5, 1, 1, 0, 1, 1, 168, 1)
sarima111_12 = mallinSovitus(ele, 1, 1, 1, 0, 1, 1, 168, 12)

#SAIMAX(4,1,2)(0,1,1) (L = 12, 13, 14)
sarima412_14 = mallinSovitus(ele, 4, 1, 2, 0, 1, 1, 168, 14)
sarima412_13 = mallinSovitus(ele, 4, 1, 2, 0, 1, 1, 168, 13)
sarima412_12 = mallinSovitus(ele, 4, 1, 2, 0, 1, 1, 168, 12)

#SARIMAX(4,1,1)(0,1,1) (L=12,13,14)
sarima411_14 = mallinSovitus(ele, 4, 1, 1, 0, 1, 1, 168, 14)
sarima411_13 = mallinSovitus(ele, 4, 1, 1, 0, 1, 1, 168, 13)
sarima411_12 = mallinSovitus(ele, 4, 1, 1, 0, 1, 1, 168, 12)

#SARIMAX(3,1,1)(0,1,1) (L=12,13,14)
sarima311_14 = mallinSovitus(ele, 3, 1, 1, 0, 1, 1, 168, 14)
sarima311_13 = mallinSovitus(ele, 3, 1, 1, 0, 1, 1, 168, 13)
sarima311_12 = mallinSovitus(ele, 3, 1, 1, 0, 1, 1, 168, 12)

#SARIMAX(2,1,1)(0,1,1) (L=12,13,14)
sarima211_14 = mallinSovitus(ele, 2, 1, 1, 0, 1, 1, 168, 14)
sarima211_13 = mallinSovitus(ele, 2, 1, 1, 0, 1, 1, 168, 13)
sarima211_12 = mallinSovitus(ele, 2, 1, 1, 0, 1, 1, 168, 12)

box911_12 <- Box.test(sarima911_12$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

```

```

box511_12 <- Box.test(sarima511_12$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

box111_12 <- Box.test(sarima111_12$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

box412_14 <- Box.test(sarima412_14$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

box412_13 <- Box.test(sarima412_13$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

box412_12 <- Box.test(sarima412_12$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

box411_14 <- Box.test(sarima411_14$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

box411_13 <- Box.test(sarima411_13$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

box411_12 <- Box.test(sarima411_12$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

box311_14 <- Box.test(sarima311_14$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

box311_13 <- Box.test(sarima311_13$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

```

```

box311_12 <- Box.test(sarima311_12$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

box211_14 <- Box.test(sarima211_14$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

box211_13 <- Box.test(sarima211_13$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)
box211_12 <- Box.test(sarima211_12$residuals,
                        lag = 20,
                        type = "Ljung-Box",
                        fitdf = p + q + P + Q)

aic911_12 = aic_f(sarima911_12)
aic511_12 = aic_f(sarima511_12)
aic111_12 = aic_f(sarima111_12)

aic412_14 = aic_f(sarima412_14)
aic412_13 = aic_f(sarima412_13)
aic412_12 = aic_f(sarima412_12)

aic411_14 = aic_f(sarima411_14)
aic411_13 = aic_f(sarima411_13)
aic411_12 = aic_f(sarima411_12)

aic311_14 = aic_f(sarima311_14)
aic311_13 = aic_f(sarima311_13)
aic311_12 = aic_f(sarima311_12)

aic211_14 = aic_f(sarima211_14)
aic211_13 = aic_f(sarima211_13)
aic211_12 = aic_f(sarima211_12)

names <- c("Malli", "L", "p-arvo", "AIC")

mallit <- c("SARIMAX(9,1,1)(0,1,1)[168]", "SARIMAX(5,1,1)(0,1,1)[168]",
           "SARIMAX(1,1,1)(0,1,1)[168]", "SARIMAX(4,1,2)(0,1,1)[168]",
           "SARIMAX(4,1,2)(0,1,1)[168]", "SARIMAX(4,1,2)(0,1,1)[168]",
           "SARIMAX(4,1,1)(0,1,1)[168]", "SARIMAX(4,1,1)(0,1,1)[168]",
           "SARIMAX(4,1,1)(0,1,1)[168]", "SARIMAX(3,1,1)(0,1,1)[168]",
           "SARIMAX(3,1,1)(0,1,1)[168]", "SARIMAX(3,1,1)(0,1,1)[168]",
           "SARIMAX(2,1,1)(0,1,1)[168]", "SARIMAX(2,1,1)(0,1,1)[168]",
           "SARIMAX(2,1,1)(0,1,1)[168]")

```

```

p_arvot <- c(box911_12$p.value,box511_12$p.value,box111_12$p.value,box412_14$p.value,
              box412_13$p.value,box412_12$p.value,box411_14$p.value,box411_13$p.value,
              box411_12$p.value,box311_14$p.value,box311_13$p.value,box311_12$p.value,
              box211_14$p.value,box211_13$p.value,box211_12$p.value)

aic_arvot <- c(aic911_12,aic511_12,aic111_12,aic412_14,aic412_13,aic412_12,aic411_14,
                  aic411_13,aic411_12,aic311_14,aic311_13,aic311_12,aic211_14,aic211_13,
                  aic211_12)

lampoviive <- c(12,12,12,14,13,12,14,13,12,14,13,12,14,13,12)

Diagnostiikka <- data.frame(Malli = mallit,L = lampoviive, 'p-arvo' = p_arvot,
                               AIC = aic_arvot)

library(Hmisc)
latex(Diagnostiikka, file="") # If you want all the data

alpha = 1 - 0.55625
z_55625 <- qnorm(1 - alpha / 2)
z_55625

Box.test(malli$residuals,
          lag = 20,
          type = "Ljung-Box",
          fitdf = p + q + P + Q)

Box.test(mallit2$residuals,
          lag = 20,
          type = "Ljung-Box",
          fitdf = p + q + P + Q)

enne2

par(mfrow=c(1,1))

# Plotataan kuva sahkonkulutusaikasarjasta, mallin (1) sovitteesta,
# ennusteesta ja ennusteen 95 %:n eli 1,96*sigma-luottamusv?leist?.
ts.plot(ele,
        ele - mallit$residuals,
        enne$pred,
        enne$pred + 1.96*enne$se,
        enne$pred - 1.96*enne$se,
        col = c("black", "red", "blue", "blue", "blue"),
        main = "Sovite ja ennuste")

ts.plot(ele,
        ele - mallit2$residuals,
        enne2$pred,
        enne2$pred + 1.96*enne2$se,
        enne2$pred - 1.96*enne2$se,
        col = c("black", "red", "blue", "blue", "blue"),
        main = "Sovite ja ennuste")

```

```

main  = "Sovite ja ennuste")
grid()

ts.plot(enne2$pred,
        enne2$pred + z_55625*enne2$se,
        enne2$pred - z_55625*enne2$se,
        col = c("black", "blue", "blue"),
        main = "Ennuste ja 95 %:n luottamusvälit")
grid()

ts.plot(enne$pred,
        enne$pred + 1.96*enne$se,
        enne$pred - 1.96*enne$se,
        col = c("black", "blue", "blue"),
        main = "Ennuste ja 95 %:n luottamusvälit")

plot(eleestimoointi)
lines(fitted(mallit2), col = 'red')

z_6125 = qnorm(0.19375, lower.tail = FALSE)
nordpool_osto <- c(0,0,0,0,0,0,79251.11062,1766.718211,0,0,0,65643.68092,
                    65589.7332,68806.72219,74020.37603,84871.22289,81881.9635,
                    80416.70549,77690.96164,77692.74838,77625.96906,0,0)
svk <- c(70000,40000.03464,39999.99469,39999.97951,13011.6502,40000.01376,70000
         ,0,70000,70000,70000,0,0,0,0,0,0,0,0,0,70000,70000)
meklari <- c(50000,50000,50000,50000,50000,50000,50000,50000,0,0,0,0,0,0,0,0,0,0,0,0,
            0,0,0,0,0,50000,50000)
nordpool_myynti <- c(53034.80993,22958.63296,25934,26298.25351,6570.03291,
                      31384.84703,50970.33006,0,0,3351.988824,4044.805555,
                      6670.535047,0,0,0,0,0,0,0,0,0,44285.72318,49666.86583)
aika <- seq(0,23)
df_optimointi <- data.frame(Nordpool_Elspot_osto_kWh = nordpool_osto,
                               SVK_kWh = svk, Elspot_meklari_kWh = meklari,
                               Nordpool_Elspot_myynti_kWh = nordpool_myynti
                               ,Aika = aika)
df_optimointi <- tidyr::gather(df_optimointi, key = "kategoria",
                                  value = "kWh", -Aika)
df_optimointi

ggplot(df_optimointi, aes(x = Aika, y = kWh, color = kategoria)) +
  geom_point() +
  geom_line()

ggplot(df_optimointi, aes(x = Aika, y = kWh, color = kategoria, group = kategoria)) +
  geom_line(size = 1) +
  scale_color_manual(values = c("blue", "red", "black", "yellow"),
                     labels = c("Elspot Meklari osto","Nordpool Elspot myynti" ,
                               "Nordpool Elspot osto","SVK osto")) +
  labs(x = "Aika", y = "kWh", color = "Sopimusosapuoli ja toimi") +
  ggtitle("Optimoitu sähköhankinta visualisoituna") +

```

```

coord_fixed(ratio = 0.0001) +
theme(plot.margin = margin(0.01, 0.01, 0.01, 0.01, "cm"))

output = cbind(enne2$pred,
               enne2$pred + z_55625*enne2$se,
               enne2$pred - z_55625*enne2$se)
write.csv2(output, file = 'ennuste.csv')

```

## B R-funktioden selitteet

- `as.POSIXct()`, `format()`
  - Funktioita ajan formatointiin kuvaajissa.
- `msts()`, `mstl()`, `autoplot()`
  - Funktiot aikasarjan hajoittamiseen ja hajoitettujen osien kuvaamiseen.
- `ggplot()`
  - Funktio paremmin formatoitujen ja kauniimpien kuvaajien muodostamiseen. Funktiolla on useita alikomentoja, joilla kuvaajat muodostetaan.
- `ggAcf()`, `ggPacf()`, `ggCcf()`
  - Funktioita autokorrelaatio-, osittaisautokorrelaatio- ja ristikorrelaatiofunktioiden kuvaamiseen `ggplot` tyyliä käyttäen.
- `mallinSovitus()`
  - Itse muodostamamme funktio lämpötilaviivellisen aikasarjamallin muodostamiseen. Tarkoitettu nopeuttamaan useamman aikasarjamallin vertailua.

## C Excel pohja

[Sähköhankinta optimointi](#)

Aikateima	Hinta (snt/kWh)	SÄHKÖN KULUTUS ENNUSTE (kWh)	OSTOMÄÄRÄ NORDPOOL	MYYNTIMÄÄRÄ NORDPOOL	OSTOMÄÄRÄ SVK	OSTOMÄÄRÄ MEKLARI	ASSIVOIMA MYYNTI	ASSIVOIMA OSTO	YHTEENSA
4.12.2014 0.00	6,26	66965,19007	0	53034,80993	70000	50000			66965,19007
4.12.2014 1.00	6,05	67041,40168	0	22958,63296	40000,03464	50000			67041,40168
4.12.2014 2.00	5,9	64065,99469	0	25934	39999,99469	50000			64065,99469
4.12.2014 3.00	5,9	63701,726	0	26298,25351	39999,97951	50000			63701,726
4.12.2014 4.00	5,85	56441,61729	0	6570,03291	13011,6502	50000			56441,61729
4.12.2014 5.00	6	58615,16673	0	31384,84703	40000,01376	50000			58615,16673
4.12.2014 6.00	6,76	69029,66994	0	50970,33006	70000	50000			69029,66994
4.12.2014 7.00	9,84	79251,11062	79251,11062	0	0	1,42109E-13			79251,11062
4.12.2014 8.00	11,25	71766,71821	1766,71821	0	70000	1,42109E-13			71766,71821
4.12.2014 9.00	11,2	66646,01118	0	3351,988824	70000	1,42109E-13			66648,01118
4.12.2014 10.00	11,05	65955,19445	0	4044,805555	70000	1,42109E-13			65955,19445
4.12.2014 11.00	10,75	63329,46495	-1,13687E-13	6670,535047	70000	1,42109E-13			63329,46495
4.12.2014 12.00	9,99	65643,68092	65643,68092	0	0	1,42109E-13			65643,68092
4.12.2014 13.00	8,98	65589,7332	65589,7332	0	0	1,42109E-13			65589,7332
4.12.2014 14.00	8,63	68806,72219	68806,72219	0	0	1,42109E-13			68806,72219
4.12.2014 15.00	9,18	74020,37603	74020,37603	0	0	1,42109E-13			74020,37603
4.12.2014 16.00	9,59	84871,22269	84871,22269	0	0	1,42109E-13			84871,22269
4.12.2014 17.00	9,69	81881,96935	81881,96935	0	0	1,42109E-13			81881,96935
4.12.2014 18.00	9,54	80416,70549	80416,70549	0	0	1,42109E-13			80416,70549
4.12.2014 19.00	9,49	77690,96164	77690,96164	0	0	1,42109E-13			77690,96164
4.12.2014 20.00	9,59	77692,74838	77692,74838	0	0	1,42109E-13			77692,74838
4.12.2014 21.00	9,13	77625,96906	77625,96906	0	0	1,42109E-13			77625,96906
4.12.2014 22.00	7,52	75714,27682	0	44285,72318	70000	50000			75714,27682
4.12.2014 23.00	7,27	70333,13417	0	49666,86583	70000	50000			70333,13417
		44	OSTOHINTA NORDPOOL	MYYNTIHINTA NORDPOOL	OSTOHINTA SVK	OSTOHINTA MEKLARI	ASSIVOIMA MYYNTI	ASSIVOIMA OSTO	YHTEENSA
			0	331520,5969	415800	300000			384279,4031
			0	138693,1017	234000,2098	300000			395307,1082
			0	162777,194	233099,9669	300000			381222,7749
			0	154923,0114	233999,8861	300000			379076,8687
			0	38375,56223	76118,15367	300000			337742,5914
			0	188026,6185	234000,0834	300000			345973,4648
			0	344100,6982	415800	300000			371699,3018
			780544,1885	0	0	1,42109E-12			780544,1885
			19891,49034	0	710100	1,42109E-12			729991,4803
			0	37512,10693	710100	1,42109E-12			672587,8931
			0	44658,69913	710100	1,42109E-12			665441,3019
			-1,22316E-12	71648,21694	710100	1,42109E-12			638451,7831
			656371,1655	0	0	1,42109E-12			656371,1655
			589586,1118	0	0	1,42109E-12			589586,1118
			594421,273	0	0	1,42109E-12			594421,273
			680173,2354	0	0	1,42109E-12			680173,2354
			814678,8685	0	0	1,42109E-12			814678,8685
			794173,2207	0	0	1,42109E-12			794173,2207
			767899,1207	0	0	1,42109E-12			767899,1207
			737986,4446	0	0	1,42109E-12			737986,4446
			745772,6917	0	0	1,42109E-12			745772,6917
			709423,7312	0	0	1,42109E-12			709423,7312
			0	332630,0668	415800	300000			383169,9332
			0	360631,1128	415800	300000			355168,8872
							KOKONAISHINTA		14111142,84
							EUROISSA		14111142,84

Kuva 12: Sähkönhankinta optimointi - Excel