

Olli Kivioja

# SÄÄTÖSÄHKÖMARKKINAN SÄÄTÖMÄÄRÄN JA SÄÄDÖN SUUNNAN ENNUSTAMINEN

Diplomityö  
Informaatioteknologian ja viestinnän tiedekunta  
Tammikuu 2022

# TIIVISTELMÄ

Olli Kivioja: Säätosähkömarkkinan säätömäärän ja säädön suunnan ennustaminen  
Tampereen yliopisto  
Sähkötekniikan DI-ohjelma  
Diplomityö  
Tammikuu 2022

---

Tässä työssä tutkittiin Suomen kantaverkon, Fingridin, ylläpitämän säätosähkömarkkinan ennustamista. Kantaverkko pyrkii pitämään sähkön kulutuksen ja tuotannon jatkuvassa tasapainossa, sillä liian suuri kulutus laskee sähköverkon taajuutta 50 Hz viitearvosta ja liian suuri tuotanto sen sijaan nostaa taajuutta yli sallittujen rajojen.

Säätosähkömarkkinoille osallistuvat kantaverkkoyhtiön sähköverkkoon liittyvät tuottajat tai kuluttajat, jotka täyttävät markkinoille asetetut määräykset. Osallistujat välittävät tuntikohtaisia säätosähkötarjouksia ylös- tai alassäädön muodossa. Ylössäätötarjous tarkoittaa tuotannon lisäämistä tai kulutuksen vähentämistä suunnitellusta ja alassäätötarjous tuotannon vähentämistä tai kulutuksen lisäämistä. Säätosähkömarkkinoilta saa korvausta tunnin loputtua määräytyvän säätötyypin mukaan, joka määrittyy säätömäärän mukaan, jos lopullinen säätömäärä poikkeaa nollassa. Tästä syystä säätosähkömarkkinoille osallistuvan tai kantaverkon ylläpitäjän tulisi pystyä ennustamaan seuraavalla tunnilla sähköverkossa tarvittavan säätosähkön määrää sekä säätosähkön suuntaa eli tyyppiä.

Työn tavoitteena oli selvittää, voidaanko säätosähkömarkkinoille aktivoitua säätosähkön määrää ja suuntaa ennustaa. Aluksi käsiteltiin, kuinka säätosähkömarkkinat sijoittuvat kantaverkon ylläpitämiin sähkömarkkinoihin sekä aikaisempaa tutkimusta sähköenergian ja -markkinoiden ennustamisesta. Lisäksi työssä käytiin läpi aikasarjaennustamiseen liittyviä ominaispiirteitä, kuten stationäärisyyttä ja lineaaristen- ja epälineaaristen mallin eroavaisuuksia. Tutkimuksessa vertailtiin ennustamiseen soveltuvia lineaarisia ja epälineaarisia ennustemenetelmiä, joista lineaarisiin menetelmiin kuuluvat ennustemallit valikoituivat käytännön tutkimukseen valikoiduille menneisyyden aikasarjoille.

Tutkimuksessa osoitettiin, että säätosähkömarkkinoille aktivoitujen säätöjen määrä ja suunta on ennustettavissa laskennallisesti tehokkailla lineaarisilla ARIMA-ennustemalleilla ja sen johdannaisilla, kun säätömäärän ennusteesta jalostetaan luokitteluennuste luokittelurajojen avulla. Luokittelurajat määrittelevät alueen, johon kuuluvat säätömäärät lasketaan niin pieniksi ylös- tai alassäädöiksi, että ne voidaan luokitella nollasäätötunneiksi. Haasteeksi tutkimuksessa osoittautui sopivien luokittelurajojen määrittäminen, jotka määrittävät nollatuntien luokittelualueen. Rajojen sopivuus riippuu ennusteen käyttäjän käyttötarkoituksesta. Tiukempi luokittelualue paransi tutkimuksessa ylös- ja alassäätötuntien luokittelutarkkuutta, mutta kasvatti riskiä väärän suuntaiselle luokitteluennusteelle. Jatkotutkimustarpeiksi tunnistettiin muun muassa säätömäärän ja luokitteluennusteen muodostaminen erillisillä ennustemalleilla.

Avainsanat: Säätosähkömarkkina, aikasarja, ennustaminen, lineaarinen mallinnus, ARIMA.

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

# ABSTRACT

Olli Kivioja: Forecasting the amount of regulation and the direction of regulation in the manual frequency restoration reserve electricity market

Tampere University  
Master of Electrical engineering  
Master's Thesis  
January 2022

In this work, the forecasting of the manual frequency restoration reserve electricity market maintained by the Finnish grid, Fingrid, was studied. The grid tends to keep a constant balance between electricity consumption and production, as excessive consumption reduces the frequency of the electricity network from the 50 Hz reference value and excessive production instead raises the frequency beyond the permitted limits.

Producers or consumers connected to the electricity grid of a transmission system operator who meet the market specifications participate in the regulatory electricity market. Participants transmit hourly control electricity offers in the form of up or down adjustment. An up-adjustment offer means an increase in production or a reduction in consumption as planned, and a down-adjustment offer means a reduction in production or an increase in consumption. The control electricity market receives compensation at the end of the hour according to the type of control determined, which is determined by the amount of control if the final amount of control deviates from zero. For this reason, a participant in the manual frequency restoration reserve electricity market or the grid operator should be able to predict the amount of control electricity needed in the electricity network in the next hour, as well as the direction or type of control electricity.

The aim of the study was to find out whether the amount and direction of control electricity activated for the manual frequency restoration reserve electricity market can be predicted. Initially, the position of the manual frequency restoration reserve electricity market in the electricity market maintained by the grid and previous research on electricity energy and market forecasting were discussed. In addition, the characteristics of time series prediction, such as stationarity and differences between the linear and nonlinear models, were reviewed. The study compared linear and non-linear prediction methods suitable for prediction, from which the prediction models included in the linear methods were selected for the past time series selected for the practical study.

The study showed that the number and direction of controls activated for the manual frequency restoration reserve electricity market can be predicted with computationally efficient linear ARIMA forecasting models and its derivatives when the control volume forecast is processed using a classification threshold. The classification limits define the range in which the adjustment quantities are counted as small up or down adjustments that they can be classified as zero adjustment hours. The challenge in the study was to determine the appropriate classification limits that define the zero-hour classification range. The suitability of the boundaries depends on the intended use of the forecast user. A stricter classification range in the study improved the classification accuracy of up and down adjustment hours but increased the risk of a misdirected classification prediction. The need for further research was identified, among other things, to generate the amount of adjustment and the classification forecast using separate forecast models.

Keywords: manual frequency restoration reserve market, timeseries, forecasting, linear modeling, ARIMA

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

# ALKUSANAT

Tämä diplomityö tehtiin Lempäälän lämmön energiayhtiön toiveesta ja halusta edistää tuotantolaitoksien sähköverkon tasapainoa ylläpitävää toimintaa säätösähkömarkkinan ennustamisen avulla. Diplomityössä käsitellään kantaverkon ylläpitämälle säätösähkömarkkinalle tilatun säätösähkömäärän ja määrän mukaan asettuvan säätösuunnan ennustamista. Diplomityö toimi ennen kaikkea oppimisprosessina aikasarjasuureiden ennustamisesta, mikä on merkittävä kokonaisuus koneoppimisen soveltamisen kentässä.

Haluan esittää kiitokset mielenkiintoisen tutkimuksen aloittamisesta Lempäälän Lämpö Oy:n kehityspäällikölle Mikko Kettuselle, joka etsii jatkuvasti uusia tapoja edistämään energiayhtiön optimoitua toimintaa. Lisäksi haluan kiittää Tampereen yliopiston sähkövoimatekniikan professoria Pertti Järventaustaa johdatuksesta säätösähkömarkkinoiden yksityiskohtiin, kuten myös Aalto yliopiston Predictricity - tutkimusryhmän Seppo Sierlaa ja Taneli Hölttää, jotka ovat toimineet merkittävinä uranuurtajina Suomessa säätö- ja reservimarkkinoiden ennustamisessa koneoppimismenetelmin ja jotka ovat halunneet avoimesti keskustella sähkömarkkinoiden ennustamisesta tarjoten työn tekijälle merkittävää tukea työn edistymisessä.

Tampereella, 27.12.2021

Olli Kivioja

# SISÄLLYSLUETTELO

1. JOHDANTO .....	1
1.1 Säättösähkömarkkinat.....	2
1.2 Säättösähkömarkkinoille osallistuminen .....	2
1.3 Säättösähkömarkkinoiden ennustaminen.....	3
1.4 Tutkimuskysymys.....	6
2. AIKASARJAENNUSTAMINEN.....	7
2.1 Aikasarjan tunnusluvut, stationäärisyys ja kausittaisuus.....	7
2.2 Ennustemallien lineaarisuus.....	8
2.3 Lineaarinen SARIMAX-malli ja sen johdannaiset .....	10
2.4 Ennustekyvyn arviointi numeeristen testien avulla .....	11
2.5 Ennusteiden arviointi visualisoinnin ja residuaalin avulla .....	12
2.6 Lineaarisen ennustemenetelmän optimointi .....	14
3. TUTKIMUSMENETELMÄT JA AINEISTOT.....	16
3.1 Tavoitteiden ja rajoitteiden määrittely .....	17
3.2 Datan hankinta.....	18
3.3 Aikasarjan ominaispiirteet .....	19
3.4 Säättösähkömäärän suunnan ennustaminen määrän avulla .....	25
3.5 Tutkimuksen testisarjojen määrittely .....	27
3.6 Ennustemenetelmän valinta .....	30
3.7 Ennustemenetelmän optimointi .....	31
3.8 Ennustemenetelmien tulosten tarkastelu numeeristen testien avulla ..	33
3.9 Ennusteen arviointi visualisoinnin avulla .....	39
3.10 Residuaalin tarkastelu.....	41
4. YHTEENVETO JA PÄÄTELMÄT .....	46
4.1 Jatkotutkimustarpeet.....	47
4.2 Johtopäätökset.....	48
LÄHTEET.....	50
5. LIITE 1. NUMEERISTEN ENNUSTETULOSTEN MITTAUSPÖYTÄKIRJA .....	53

# LYHENTEET JA MERKINNÄT

<i>Ocu</i>	Kantaverkkoyhtiöiden hyväksymistä ylössäätötarjouksista koostuva ylikustannus kantaverkkoyhtiölle
<i>Pu</i>	Säätösähkömarkkinoilta saatu ylössäädön hinta
<i>Ps</i>	Sähkömarkkinoiden yleinen sähkön hinta
<i>Qu</i>	Säätösähkömarkkinoilla hyväksytyn ylössäädön energiamäärä
<i>Ocd</i>	Kantaverkkoyhtiöiden hyväksymistä alassääätötarjouksista koostuva ylikustannus kantaverkkoyhtiölle
<i>Pd</i>	Säätösähkömarkkinoilta saatu alassääädön hinta
<i>Qd</i>	Säätösähkömarkkinoille hyväksytyn alassääädön energiamäärä
ADF	Augmented Dickey Fuller -tilastollinen stationäärisyystesti
KPSS	Kwiatkowski-Phillips-Schmidt-Shin - yksikköjuuritest
AR	Autoregression huomioiva ennustemalli
MA	Liukuvan keskiarvon huomioiva ennustemalli
ARIMA	Autoregression, aikasarjan trendin ja liukuvan keskiarvon huomioiva ennustemalli
ARIMAX	Autoregression, aikasarjan trendin, liukuvan keskiarvon ja apumuuttujien huomioiva ennustemalli
SARIMA	Aikasarjan kausittaisuuden, autoregression, aikasarjan trendin ja liukuvan keskiarvon huomioiva ennustemalli
SARIMAX	Aikasarjan kausittaisuuden, autoregression, aikasarjan trendin, liukuvan keskiarvon ja apumuuttujien huomioiva ennustemalli
LSTM	Long short-term memory aikasidonaisuuden huomioiva neuroverkkomalli
aFFR	Automaattinen taajuudenhallintareservi
<i>p</i>	SARIMAX-mallissa ja sen johdannaisissa käytetty parametri, joka määrittää ennusteessa käytetyn autoregression asteen
<i>d</i>	SARIMAX-mallissa ja sen johdannaisissa käytetty parametri, joka määrittää ennusteessa käytetyn trendin huomioinnin asteen
<i>q</i>	SARIMAX-mallissa ja sen johdannaisissa käytetty parametri, joka määrittää ennusteessa käytetyn liukuvankeskiarvon asteen
<i>c</i>	Ennustemallinnusta nostava vakiotermi
$\phi$	SARIMAX-mallin ja sen johdannaisten muodostava autoregressiotermien muodostama kerroin
$\theta$	SARIMAX-mallin ja sen johdannaisten muodostava liukuvan keskiarvotermien muodostama kerroin
$\varepsilon$	SARIMAX-mallin ja sen johdannaisten virhe mallinnuksessa
<i>m</i>	SARIMAX-mallissa ja sen johdannaisissa käytetty parametri, joka määrittää aikasarjassa huomioidut kausittaisuusjaksot
<i>P</i>	SARIMAX-mallissa ja sen johdannaisissa käytetty parametri, joka määrittää ennusteessa käytetyn autoregression asteen määritellyn kausittaisuuksien välillä
<i>D</i>	SARIMAX-mallissa ja sen johdannaisissa käytetty parametri, joka määrittää ennusteessa käytetyn trendin asteen määritellyn kausittaisuuksien välillä
<i>Q</i>	SARIMAX-mallissa ja sen johdannaisissa käytetty parametri, joka määrittää ennusteessa käytetyn liukuvan keskiarvon asteen määritellyn kausittaisuuksien välillä
$\Phi$	SARIMAX-mallin ja sen johdannaisten muodostava autoregressiotermien muodostama kerroin määriteltyjen kausittaisuusjaksojen välille

$\theta$	SARIMAX-mallin ja sen johdannaisten muodostava liukuvan keskiarvotermien muodostama kerroin määriteltyjen kausittaisuusjaksojen välille
$\beta$	SARIMAX-mallin ja sen johdannaisten muodostava kerroin ulkoisten muuttujien vaikutukselle
$y_t$	Aikasarjan oloarvo hetkellä $t$
$y'_t$	Aikasarjan ennustettu arvo hetkellä $t$
$N$	Aikasarjan arvojen lukumäärä
MAE	Absoluuttinen keskiarvovirhe
MAPE	Prosentuaalinen keskiarvovirhe
RMSE	Neliöllisen virheen keskiarvon neliöjuuri
AIC-arvo	Akaike Information Criteria -arvo arvioi mallin sopivuutta sille osoittamaan aikasarjaan huomioiden mallin monimutkaisuuden
ACF-kuvaaja	Autokorrelaatiokuvaaja
PACF-kuvaaja	Osittaisen autokorrelaation kuvaaja
p-arvo	Tilastollisissa testeissä käytetty todennäköisyysarvo
GA-ARIMA	Geneettistä algoritmia hyödyntävä ARIMA-malli

# 1. JOHDANTO

Suomen sähköverkon kantaverkkoa ylläpitää Suomen valtion ja suomalaisten eläkeyhtiöiden omistama kantaverkkoyhtiö Fingrid Oyj yhteistyössä muiden pohjoismaisten kantaverkkoyhtiöiden kanssa. Yhtiön tehtävänä on ylläpitää verkon toimintavarmuutta huolehtimalla, että verkkoon liitettyjen tuottajien ja kuluttajien toiminta on tasapainossa. Suomen kantaverkko toimii osana yhteispohjoismaista sähköjärjestelmää, mikä parantaa verkon käyttövarmuutta. Lisäksi siirtoyhteydet on muodostettu myös Venäjälle, Viroon ja Keski-Eurooppaan. (Fingrid 2021a.) Fingrid edistää toimintavarmuutta kustannustehokkaasti sähkömarkkinoiden avulla, jotka kannustavat verkkoon liittyneitä kuluttajia ja tuottajia, eli resurssinhaltijoita ylläpitämään kysynnän ja tarjonnan tasapainoa (Fingrid 2021b).

Elaminin ja Fukushigen (2018, 258) mukaan sähköön kulutuksen ennustaminen on merkittävää toiminnan suunnittelun kannalta. Sähkömarkkinoiden ennustaminen mahdollistaisi myös toiminnan suunnittelun kantaverkon tai tuotantoyksikön osalta. Sähkömarkkinat mahdollistetaan, kun kantaverkkoyhtiöt siirtävät sähköön jakeluverkkoyhtiöille, jotka jakavat sähköön omille asiakkailleen. Yleistyvän hajautetun pientuotannon ansiota yksittäinen pienkuluttajakin voi osallistua sähkömarkkinoille (Fingrid 2021b).

Sähkömarkkinat muodostuvat erilaisista markkinapaikoista, kuten finanssi-, vuorokausi-, päivänsisäisistä-, reservi- ja säätösähkömarkkinoista. Markkinoille tarjotaan aktivoitua säätöenergiaa tai reservienergiaa, joita ei välttämättä markkinoille kuitenkaan aktivoida. Tämän lisäksi markkinat poikkeavat toisistaan tarjouten aktivointiperiaatteiden mukaan. Vaikka reservimarkkinoille tarjottua energiaa ei otettaisi verkkoon käyttöön, niin siitä saa sopimuksen mukaisen korvauksen. Säätösähkömarkkinoille tarjotusta energiasta saa korvauksen vain, jos tarjouksen energia otetaan verkossa käyttöön. (Fingrid 2021b.)

Perinteisesti reservimarkkinoille osallistuvat tuotantolähteet koostuvat vesivoimasta tai fossiilisten polttoaineiden tuotantomuodoista. Yleistyvien pientuotantojen aurinkokennot, akustot ja joustavat kuormat eivät yleensä yksistään riitä saatavilla olevan tehonsa puolesta täyttämään markkinoille osallistumisen vaatimuksia. Kuitenkin niin sanotut virtuaalivoimalaitokset voivat yhdistää eri tuotanto- ja kulutusmuodot yhteiskäyttöön ja



kaksisuuntaisen jouston ansiota osallistumaan erityisesti reservimarkkinoille. (Subramanya, Yli-Ojanperä, Sierla, Hölttä, Valtakari, Vyatkin, 2021, 1.)

## 1.1 Säättö sähkömarkkinat

Säättö sähkömarkkinoille osallistuvat resurssinhaltijat muodostavat säättötarjouksia säättökykyisestä tuotannostaan tai kuormastaan (Fingrid 2021d). Aikaisemmin Euroopassa kansalliset kantaverkkoyhtiöt ovat ylläpitäneet säättö sähkömarkkinoita sisäisesti, mutta pohjoismaisen kantaverkkoyhtiöiden yhteistyön tavoin vastaavia kantaverkkoyhtiöiden välisiä yhteistyösopimuksia on muodostettu myös muualla Euroopassa, kuten Portugalin, Espanjan ja Ranskan välillä (Frade, Santana, Shafiekhah, Catalão 2018, 168).

Markkinoiden yhteistyö ja yhteiset markkinamallit on koettu tarpeellisiksi Euroopassa merkittävästi kasvavan uusiutuvan energiatuotannon myötä. Erityisesti yleistyneen tuulituotannon ennustaminen on koettu haastavaksi, jolla on havaittu olevan vaikutuksia markkinoilla myytyyn energiaan. Kansainvälisen sähköenergiatuotannon ja -kulutuksen tasapainon lisäksi säättö sähkömarkkinat pitävät kansainvälisten yhteyksien myötä myös maiden välisten sähkön siirron tasapainoa. (Frade ym. 2018, 168–169.)

## 1.2 Säättö sähkömarkkinoille osallistuminen

Säättö sähkömarkkinoille voi osallistua tekemällä sopimuksen Fingridin kanssa, jos täyttää markkinoille osallistumisen kriteerit, kuten 10 MW tehomuutoksen 15 minuutin aikana. Tarjouksen yhteydessä annetaan tieto säädettävästä tehosta, energian hinnasta, alueesta, jolla säättö tarjoava resurssi sijaitsee sekä resurssin kuvaus. Tarjoukset osoitetaan kantaverkkoyhtiölle ylös- tai alassäättötarjouksina. Ylösäättötarjouksella tarkoitetaan sähkön myyntiä, jolloin tuotantoa lisätään tai suunniteltua kulutusta ollaan valmiita vähentämään tarjouksen aktivoinnin myötä. Vastaavasti alassäättötarjouksen myötä resurssinhaltija ostaa sähkön kantaverkkoyhtiöltä ja on valmis tarjouksen aktivoinnin jälkeen vähentämään tuotantoaan tai lisäämään kulutusta suunnitellusta. Ennen säättötarjousten aktivointia, tarjoukset järjestetään hinnan mukaan ja ylösäättötarjousten yhteydessä aktivoidaan tarpeen mukaan halvimmasta kalliimpaan. Kaikki osallistujat, joiden säättötarjoukset aktivoidaan, saavat korvauksen kalleimman aktivoidun säädön mukaan. (Fingrid 2021d.)

Fingridin säättö sähkömarkkinoilla käytettävää tarjouskäytäntöä kutsutaan Pay as bid -periaatteeksi (Fingrid 2021e). Pay as bid -periaatteinen käytäntö on käytössä myös muualla Euroopassa, kuten Saksan säättö sähkömarkkinoilla, ja se ohjaa tarjouksen

tekijöitä tarjoamaan paljon matalia hintoja pieniä säätöjä varten, jotta pientä säätöä olisi kantaverkolla käytettävissä mahdollisimman edullisesti. Pay as bid -periaate kuitenkin mahdollistaa myös suuremmat tuotot, kun kalliilla hinnalla voi tarjota vähemmän energiaa ja jos korkea tarjous hyväksytään, saa korkean hinnan myös pienemmistä säädöistä, jotka alun perin oli tarjottu halvalla. (Götz & Heim 2021, 5–6.) Kantaverkkoyhtiöiden hyväksymistä ylössäätötarjouksista johtuvaa kustannusta voidaan kuvata ylikustannuksena  $OCu$ . Ylikustannus voidaan laskea aktivoidun säädön hinnan  $Pu$  ja sen hetkisen sähkön hinnan  $Ps$  erotuksen mukaan kerrottuna aktivoidun ylössäädön energiamäärällä  $Qu$  ja on esitetty kaavassa 1 (Frade ym. 2018, 169).

$$OCu = (Pu - Ps) * Qu \quad (1)$$

Alassäätötarjouksia aktivoidaan kalleimmasta halvimpaan ja halvimman tarjouksen hinta jää kaikille aktivoiduille tarjoajille voimaan. Alassäätöhinta on korkeintaan Nord Poolin hinta-alueen mukainen Suomen hinta (Fingrid 2021d). Kuten aktivoidut ylössäätötarjoukset niin myös aktivoidut alassäätötarjoukset aiheuttavat kantaverkkoyhtiölle alassäädöistä johtuvia ylikustannuksia  $OCd$  (Frade ym. 2018, 169). Aiheutuneet kustannukset ovat vastaavia kuin kaavassa 1 kuvattu ylössäätötarjousten kustannus, mutta johtuen alassäätöhinnan päinvastaisesta toiminnasta muodostuu ylikustannus sen hetkisen sähkömarkkinahinnan  $Ps$  ja alimman hyväksytyyn alassäätötarjouksen hinnan  $Pd$  päinvastaisena erotuksena verrattuna ylössäätötarjouksen ylikustannuksen laskentaan. Erotus vastaavasti kerrotaan aktivoitujen alassäätöjen energiamäärällä  $Qd$ , jolloin saadaan kaavan 2 mukainen ylikustannus (Frade ym. 2018, 169).

$$OCd = (Ps - Pd) * Qd \quad (2)$$

Fraden ym. (2018, 174) mukaan kantaverkkoyhtiöiden välisten yhteisten markkinapaikkojen yleistyminen edistäisivät säätösähkömarkkinoiden avulla kustannustehokkuutta. Ylössäätötarjouksiin on asetettu Suomessa 5000 €/MWh maksimihinta. Muodostuneet säätöenergiat huomioidaan automaattisesti reservitoimittajien tasevastaavien taseselvityksessä tuotanto- tai kulutustaseessa. Säätöenergian määrän Fingrid ilmoittaa tasevastaavalle käyttötunnin jälkeen ja viimeistään 13 vuorokauden kuluessa. (Fingrid 2021e.)

### 1.3 Säätösähkömarkkinoiden ennustaminen

Tehokas osallistuminen sähkömarkkinoihin ylläpitää sähköverkon tasapainoa ja edistää kaikkien verkkoon kytkeytyneiden käyttämän sähkön laatua (Fingrid 2021d). Kansainvälisten kantaverkkoyhtiöiden yhtenäistyvät sähkömarkkinamallit mahdollistavat

pohjan hyödyntää säätösähkömarkkinan ennustamisen ja osallistumisen menetelmiä myös muissa Euroopan maissa. Fraden ym. (2018, 169) esittämien kantaverkkojen välisten yhteistyösopimusten lisäksi Subramanya ym. (2021,1) ovat havainneet sähkömarkkinoilla kasvavan kiinnostuksen osallistua markkinoille säädettävän kuorman, tuuli- ja aurinkotuotannon sekä akustojen tai sähköajoneuvojen avulla. Subramanya ym. (2018, 2) esittävät, että näitä tuotantomuotoja yhdistävälle virtuaalivoimalaitokselle olisi edullista energiavirtojen ennustaminen, joista erityisesti tuulivoimatuotannon ennustaminen on koettu haastavaksi. Myös Perez-Diaz, Giusàndez, Chazarra ja Helseth (2020,1) ovat aikaisemmin tutkineet vesivoimatuotannon ennustamista reservimarkkinoille osallistumista varten. Malik ja Haven (2012, 343) esittävät, kuinka kantaverkkojen toimintaa voisi kehittää ennustamalla tulevaa säätömäärän virhettä ja luoda niiden avulla optimoituja tarjouksia säätösähkömarkkinoille. Gellert ym. (2021, 14) ovat esittäneet tapaa, kuinka ennustemalleja voidaan käyttää energian hallinnan päätöksenteossa ennustettaessa kuluttajien sähkön tuotannossa ja kulutuksessa (kuva 1).



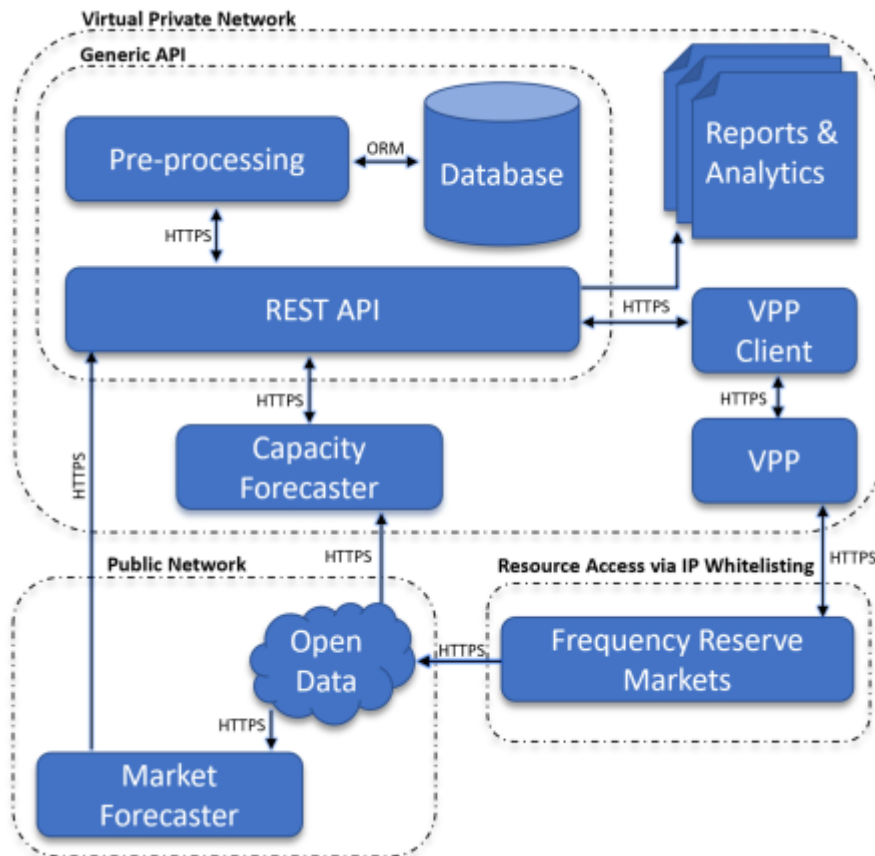
KUVA 1. Ennusteiden käyttäminen energiahallinnan päätöksenteossa (Gellert ym. 2021, 14).

Energiavirtojen ennusteiden avulla voisi olla mahdollista optimoida toimintaa säätösähkömarkkinoilta saadun kustannuksen mukaisesti. Parhaimmillaan resurssinhaltija, jolla on resursseja osallistua molempiin markkinoihin voisi alassäätötunnilla ostaa normaalia halvempaa sähköä varastoon ja myydä sen myöhemmin normaalia sähkön hintaa kalliimmalla ylössäätötarjouksen muodossa. Nitsch, Deissenroth-Uhrig, Schimeczek ja Bertsch (2021, 1) ovat artikkelissaan esittäneet agenttipohjaista sähkömarkkina ennustemallia, joka tarkkailee sähkömarkkinahintoja Saksassa ja akuston avulla luo tarjouksia markkinatilanteen mukaan.

Nopeasti kasvava uusiutuva energiatuotanto luo vaatimuksia joustavaan sähkömarkkinaosallistumiseen. Tällä hetkellä säätösähkömarkkinan tarpeisiin vastataan pääosin perinteisillä voimalaitoksilla tai vesivoimalla, mikä tulevaisuudessa ei mahdollisesti riitä tasapainon ylläpitämiseksi verkon tasapainon nopeista muutoksista johtuen. Tämä korostaa joustavampien tuotantojen ja säädettävien kuormitusten

käyttöönottoa sähkömarkkinoilla sekä tarvetta tarkemmalle sähkömarkkinatasapainon ennustamiselle. (Nitschin ym. 2021, 1–2.)

On kehitetty merkittävä määrä älykkäitä menetelmiä, jotka edistävät hajautetun pientuotannon osallistumista reservimarkkinoille. Pientuotannon tehokas osallistuminen energiareсурseja yhdistävän virtuaalivoimalaitoksen kautta edellyttää energiavirtojen tehokasta ennustamista, mitä varten tarvitaan datan keräysmenetelmiä. Avointen datalähteiden ja tietoteknisten rajapintojen avulla tuotanto- ja kulutusresurssista muodostettu virtuaalivoimalaitos voi hakea tietoja reservimarkkinan viimeaikaisista aktivoinneista ja reservimarkkinaan vaikuttavista tekijöistä. Kerättyjen tietojen avulla virtuaalivoimalaitos voi ennustaa tulevaisuuden tilannetta ja lähettää tiedot markkinan ylläpitäjälle, joka Suomessa on Fingrid (kuva 2). (Subramanya ym. 2018. 6–8.)



KUVA 2. Virtuaalivoimalaitoksen tietotekninen arkkitehtuuri reservimarkkinoihin osallistumiseen (Subramanya ym. 2018, 8)

Kuvassa 2 esitettyä Subramanyan ym. (2018, 8) reservimarkkinoiden ennustamis- ja osallistumismallia voitaisiin mahdollisesti hyödyntää myös säätösähkömarkkinoille osallistumisessa, sillä säätösähkömarkkinoille osallistutaan samankaltaisesti, mutta poikkeavina tekijöinä on ennustettava suure ja ennustettava aikaväli. Reservienergian sijaan säätösähkömarkkinoille tarjotaan aktivoitua säätöenergiaa, ja ajanjakso, jolle

ennuste tehdään, on säätösähkömarkkinoilla seuraava tunti. Reservimarkkinoilla tarjoukset jätetään edeltävänä päivänä, joten reservimarkkinoilla ennustettava aikaväli on seuraavat 24 tuntia. (Subramanya ym. 2018, 6).

Ennustamista varten sähkömarkkinoihin ja voimajärjestelmiin, kuten tuulituotantoon liittyvää dataa saadaan avoimen API-rajapinnan kautta, mistä datakyselyn voidaan tehdä millä tahansa rajapintaa tukevalla työkalulla tai tiedot voidaan ladata sivulta myös taulukkomuodossa (Fingrid 2021f). Sähkömarkkinoiden tiedonhakua ja ennustamista voi tehdä ilmaiseksi API-rajapintoja ja koneoppimismalleja tukevalla ohjelmalla, kuten tässäkin diplomityössä on tehty. Gellertin, Fioren, Florean, Chisin ja Palmierin (2021, 2) mukaan jatkuvasti kasvavan sähkön kulutuksen ja omatuotannon vuoksi on merkittävää löytää toimivia menetelmiä ennustaa energian kulutusta ja tuotantoa.

## 1.4 Tutkimuskysymys

Tässä työssä pyritään ennustamaan säätösähkömarkkinoiden aktivoituneen säätösähkön määrää, jotta yksittäinen kuluttaja ja tuottaja voisi tehostaa osallistumistaan säätösähkömarkkinoille. Säätösähkömarkkinoilla myydään markkinoille energiaa, jolloin vain aktivoituneesta energiasta saa korvauksen. Säätösähkömarkkinat toimivat reaaliaikaisesti ja seuraavalle tunnille tehdyt säätötarjoukset tulee esittää 45 minuuttia ennen tunnin alkua. (Fingrid 2021b.) Sähkömarkkinoille on viime aikoina ilmaantunut niin sanottuja aggregaattoreita eli resurssinhaltijoita, jotka muodostavat pientuotannosta ja kulutuksesta kokonaisuuden. Kokonaisuus voi molempiin suuntiin joustavan kuormituksen ansiota osallistua sähkömarkkinan eri markkinapaikoille (Fingrid 2021c).

Kolehmainen (2016, 59) mukaan säätösähkömarkkinoiden ennustamisen keskeinen tavoite olisi säädöstä saadun hinnan ja säätömäärän lisäksi luotettava arvio säädön suunnasta. Säädön hintaa ei tässä työssä ei ennusteta. Yleensä kuitenkin korkeista säätömääristä saadut korvaukset asettuvat sähkömarkkinan tarjousten aktivointimenetelmästä johtuen korkeisiin hintoihin. Tästä syystä työssä keskitytään säätömäärän ja suunnan ennustamiseen ja tutkimuskysymyksenä on, voidaanko ennustemalleilla ennustaa sähkömarkkinoille aktivoituneiden säätötarjousten määrää sekä määrän mukaan asettuvaa säätösähkötunnin tyyppiä alas- ja ylössäädön välillä. Tutkimuksessa tarkastellaan ennusteen suoriutumista monesta eri näkökulmasta, jotta ennustemenetelmän mahdollisten virheiden seuraamuksia voitaisiin arvioida ennen mallin käyttöönottoa.

## 2. AIKASARJAENNUSTAMINEN

Aikasarjaennustaminen voidaan nähdä laajana kokonaisuutena, jossa ennustetaan aikaa vasten muuttuvaa suuretta, kuten sähkön kulutusta (Clements, Hurn. & Li 2015, 2), sähkömarkkinoiden muutosta (Kemptitya, Sierla, De Silva, Yli-Ojanperä, Alahakoon & Vyatkin 2020, 1), pörssihintojen muutosta (Tsantekidi, Passalis, Tefas, Kannianen, Gabbouj & Iosfidis 2017, 1), eri materiaalien, kuten puun hinnan muutosta (Banas & Utnik-banas 2021) tai asiakasmäärien ruuhkautumista sairaaloissa (Tuominen, Lomia, Oksala, Palomäki, Peltonen, Huttunen & Roine 2021, 1).

Aikasarjan ennustamisen menetelmiä on useita ja niitä luokitellaankin hyvin monipuolisesti. Kemptityan ym. (2020, 3) luokittelevat aikasarjaennustamisen menetelmät viiteen kategoriaan, joita ovat agenttipohjaiset mallit, peliteoriamallit, perinteiset menetelmät, tilastolliset mallit, rajoitettujen muotojen mallit ja älykkäät mallit, kuten neuroverkot ja tukivektorikoneet. Nousi ym. (2017, 1) sen sijaan luokittelevat ennustamista yleisesti mallin lineaarisuuden ja epälineaarisuuden mukaan.

Tuominen ym. (2021, 2) ovat ennustaneet aikasarjamenetelmillä Tampereen yliopistollisen sairaalan ruuhkautumista. Tuomisen ym. (2021, 2) mukaan lineaariset aikasarjamenetelmät voivat suoriutua numeerisesta ennustetehtävästä hyvin, varsinkin jos malli tukee useampien muuttujien käyttöä.

Kuten koneoppimisen ennustemenetelmissä yleisesti, niin myös aikasarjaennustamisessa malli ensiksi harjoitetaan harjoitussarjalla siten, että mallille välitetään myös vastaavina ajanhetkinä toteutuneet arvot. Aikasarjaennustamisessa testisarja on harjoitussarjaa seuraavat ajanhetket, joita esitetään mallille sitä mukaan, kun ennustuksia tehdään (Merten, Rücker, Schoenberger & Uwe Sauer 2020, 10). Aikasarjaennustamisessa on tärkeää huomioida harjoitussarjan arvojen järjestys ja eheys, sillä yleensä ennustettavan suureen edeltäviä arvoja käytetään ennusteen perustana. Ongelmia voi aiheutua, jos tuntipohjaisesta sarjasta puuttuu joitain tunteja, tai monimuuttujamenetelmiä käytettäessä suureet on kerätty eri ajanjaksoilta.

### 2.1 Aikasarjan tunnusluvut, stationäärisyys ja kausittaisuus

Aikasarjaennusteiden muodostamisessa tärkeää on tutustua aikasarjan lisäksi myös sen tilastollisiin tunnuslukuihin, kuten jakaumaan, keskiarvoon tai hajontaan (Cuaresma, Hlouskova, Kossmeier & Obersteinener 2004, 92).

Aikasarjaennusteen vaatimuksena on, että aikasarja on stationääristä, jolloin siinä ei ole havaittavissa kausittaisuutta tai nousevaa trendiä. Jos aikasarja sisältää kausittaisuutta tai nousevaa trendiä, niin ennen mallinnusta ominaispiirteet otetaan huomioon poistamalla ominaispiirteen mallinnuksen ajaksi, jonka jälkeen ne lisätään aikasarjaan takaisin. Muunnoksen voi tehdä erikseen, mutta joskus mallinnusmenetelmissä muokkaustoimenpiteet sisältyvät malliin (Banas & Ulrik-Banas 2021, 3).

Stationäärisyyttä voidaan arvioida visuaalisesti tarkastelemalla, havaitaanko kausittaisuutta tai muuttuuko keskiarvo tai aikasarjan hajonta ajan myötä. Stationäärisyyden määrittämiseksi on myös muodostettu tilastollisia testejä, kuten ADF (*Augmented Dickey-Fuller*) ja KPSS (*Kwiatkowski-Phillips-Schmidt-Shin*) -yksikköjuuritestit (Banas & Ulriks-Banas 2021, 3). Tilastolliset testit perustuvat nollahypoteesiin, jonka kumoamistodennäköisyyttä tutkitaan. ADF-testissä nollahypoteesi olettaa, että sille esitetty aikasarja ei ole stationääristä sen perusteella, että aikasarja sisältää yksikköjuuren. Hypoteesi kumotaan, jos todennäköisyys stationäärisyydelle on tilastollisesti merkittävä. KPSS-testissä kumottava nollahypoteesi on päinvastainen. Kumottava nollahypoteesi olettaa, että aikasarja on stationääristä eikä sisällä yksikköjuurta. (Banas & Ulriks-Banas 2021, 3.) Useimmat stationäärisyytestit nollahypoteesissa olettavat yksikköjuuren läsnäolon. KPSS-testissä yksikköjuuren poissaolo ei vahvista sarjan stationäärisyyttä vaan erityisesti kausittaisuusstationäärisyyttä, mikä on merkittävää tarkastelunäkökulma, sillä sarja voi olla yleisesti epästationäärinen, mutta kausittaisuuden näkökulmasta stationäärinen.

Aikasarjaennustamisessa kausittaisuus on hyvin merkittävä ja mielenkiintoinen piirre, sillä kausittaisuutta voi olla monella tasolla. Tuntitasoisessa aikasarjassa voidaan havaita päivän sisäisen kausittaisuuden lisäksi myös viikko-, kuukausi- tai vuosikohtaista kausittaisuutta, jotka ennustemallin tulee huomioida. (Elamin & Fukushige 2018, 258.)

## 2.2 Ennustemallien lineaarisuus

Epälineaarilla malleilla tarkoitetaan ennustemalleja, jotka pystyvät muodostamaan niille osoitetusta datasta epälineaarisia yhteyksiä. Epälineaariset mallit ovat soveltuvia muun muassa talouden ennustamiseen niiden epäsäännöllisestä ja epästationaarisesta luonteesta johtuen. Yleisiä epälineaarisia malleja ovat neuroverkot ja tukivektorikoneet (Nousi ym. 2017,1).

Kemptitiya ym. (2020, 8) ovat selvityksessään käyttäneet rakenteeltaan yksinkertaista neuroverkkoa reservimarkkinoiden hinnan ennustamiseen, ja ennusteiden avulla tehneet tarjouksia vain merkittäville tunneille.

Lineaariset mallit ovat malleja, jotka perustuvat oletukseen, että asioilla on lineaarinen yhteys keskenään. Tällöin selittävät tekijät asettuvat viivalle tai tasolle ennustettavan suureen ja selittävien tekijöiden muodostamassa avaruudessa. Lineaariset mallit toimivat erinomaisesti silloin kun muuttujalla ja ennustettavalla asialla on olemassa lineaarinen yhteys. Jos lineaarista yhteyttä ei ole niin epälineaaristen mallien käyttöä ennustamiseen kannattaa harkita. (Zhang & Guo 2020, 1)

Lineaarisiiin malleihin kuuluvat tilastollisten mallien mukaelmat, kuten autoregressio AR-, liukuva keskiarvo MA- tai näiden yhdistelmänä syntyvä autoregressio liukuvakeskiarvo malli ARMA-malli, mitä on usein käytetty muun muassa sähkön hinnan ennustamisessa (Vyatkin ym. 2020, 6). Myös Gellertin ym. (2021, 6) mukaan ARIMA on hyvä valinta sähkön kulutuksen ja tuotannon ennustamiseen, sillä se on yhdistelmä aikasarjan autoregressiota hyödyntävästä mallista (AR) ja liukuvaa keskiarvoa hyödyntävästä mallista (MA) lisäksi huomioiden aikasarjan nousevan tai laskevan trendin integroinnilla (I). Näiden lisäksi malliin voidaan yhdistää aikasarjan kausittaisuuden huomioiva osuus (S). Lisäksi myös apumuuttujia voidaan ottaa ARIMA-mallin muodostamaan kaavaan mukaan lisäparametrilla (X). Kaikki laskennalliset ominaisuudet huomioivaa mallia kutsutaan SARIMAX-malliksi. (Gellert ym. 2021, 7.) SARIMA-mallin kausittaisuus parametrilla voidaan ottaa huomioon kausittaisuus, kuten tuntisarjassa 24 tunnin päiväsyklisyys (Clements, Hurn & Li 2015, 523). Jos halutaan 24 tunnin kausittaisuuden lisäksi huomioida myös viikoittainen kausittaisuus, niin SARIMAX-mallin avulla voidaan välittää aikasarjan Fourier-muunnoksen termejä mallin apumuuttujiksi kuvaamaan muita kausittaisuuksia (Gellert ym. 2021, 9).

Gellert ym. (2021, 9) tutkivat sähkönkulutuksessa myös epälineaaristen mallien, kuten neuroverkkojen (NN), konvoluutioneuroverkkojen (CNN) ja aikasidonaisuuden huomioivaa LSTM-neuroverkkomallia, joka suoriutui sähkön kulutuksen ja tuotannon ennustamisesta lineaarisen ARIMA-mallin tavoin myös hyvin. Gellertin ym. (2021, 8) mukaan tilastollisten mallien lisäksi myös epälineaariset mallit ovat osoittaneet selvää kyvykkyyttä yleisesti aikasarjaennustamisessa eikä ilman ongelmakohtaista testausta voida etukäteen olla varmoja suoriutuuko tehtävästä lineaarinen- vai epälineaarinen ennustemalli paremmin. Moonin, Hossainin ja Chon (2021, 1) mukaan kuitenkin yleisesti aikasarjaennustamisessa lineaarisen ennustemenetelmän avulla saavutetaan hyvälaatuisia ennusteita. ARIMA-mallin johdannaiset suoriutuvat usein suosittua epälinearisuuteen perustuvaa LSTM-neuroverkkomallia paremmin, mutta näiden mallinnuksien parhaita puolia voitaisiin ennustamisessa myös yhdistää. Gellert ym. (2021, 6) esittävät, että lineaarinen SARIMA-malli sopii sähkönkulutuksen ennustamiseen, joten mahdollisesti lineaarinen SARIMA malli sopisi myös sähkön



kulutuksen ja tuotannon epätasapainosta johtuvan säätömäärän ennustamiseen. Myös Merten ja muiden (2020, 18) artikkelissa todettiin, että Saksan automaattisia säätösähkömarkkinoita (aFFR) ennustettaessa SARIMA-malli suoriutui kiitettävästi verrattuna toisiin lineaarisiin sekä epälineaarisiin malleihin, kuten perinteiset- tai LSTM neuroverkot, vaikkakin myös muut mallit suoriutuivat ennustamisesta niin hyvin, ettei merkittävää eroa voitu mallien välillä todeta. Elamin ja Fukushige (2018, 265) ovat todenneet sähkön kulutuksen ennustamisen tutkimuksessa, että ARIMA-malli todennäköisesti suoriutuu hyvin kuormituksen kaltaisten suureiden ennustamisessa.

SARIMAX-mallikokonaisuudella tutkimuksen vertailuun saadaan neljä eri mallia, joiden suoriutumisen voidaan saada lisätietoa säätösähkön ominaisuuksista. Edeltävistä syistä ja aikaisemmista tutkimustuloksista johtuen SARIMAX-mallin eri muodot ovat valikoituneet tarkasteltavaksi työhön säätösähkömäärän ennustamiseen. Myös Kolehmainen (2017, 26) on esittänyt, että säätösähkön aikaisemmat oloarvot ennustavat tulevaa, mikä asettaa hyvät edellytykset säätösähkön ennustamiselle SARIMAX-mallilla ja sen johdannaisilla.

### 2.3 Lineaarinen SARIMAX-malli ja sen johdannaiset

ARIMA-malli voidaan kuvata lineaarisena yhtälönä  $ARIMA(p, d, q)$ , jossa  $p$  on autoregressio osan aste,  $d$  esittää, kuinka monen ajanhetken trendiä huomioidaan ajanhetkien erotuksilla ja  $q$  on liukuvan-keskiarvo -osuuden aste. ARIMA-mallin lineaarinen yhtälö on kuvattu kaavassa 3. (Gellert ym. 2021, 9.)

$$y'_t = c + \sum_{i=1}^p \phi_i y'_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t \quad (3)$$

Kaavassa 3  $y'_t$  on  $d$  parametrin mukainen aikasarja, ajanhetkien erotuksen jälkeen. Parametri  $c$  on vakio ja  $\phi_i y'_{t-i}$  on autoregressio (AR) termi, jonka termi  $p$  kuvaa montako ajanhetkeä kertoimen laskentaan huomioidaan kerrallaan ja  $\phi_p$  on lopullinen kerroin mikä lasketaan mallin harjoitusvaiheessa. Kaavassa 3  $\theta_q \varepsilon_{t-q}$  on ARIMA-mallin liukuvakeskiarvo-osuus (MA), jonka termi  $q$  osoittaa montako ajanhetkeä huomioidaan liukuvan keskiarvon kerrointa  $\theta_q$  laskettaessa mallia harjoitusvaiheessa. Jotta lineaarisen mallin kaava pitäisi paikkansa on kaavassa myös jäljelle jäävä virhe  $\varepsilon_t$ . Aikasarjaa ennustettaessa ominaista olisi, että virheiden keskiarvo olisi mahdollisimman lähellä nollaa ja tasaisesti jakautunutta niin sanottua valkoista kohinaa, tai muuten mallinnus tekee säännöllisesti tietynlaista virhettä, joka olisi sopivalla mallilla mallinnettavissa. ARIMA malli on myös laajennettavissa SARIMA-malliksi huomioimalla kausittaisuus neljän parametrin avulla, jolloin mallin yhtälöstä tulee

$SARIMA(p,d,q)(P,D,Q)[m]$ . Tällöin lineaarinen kuvaus sarjasta muuttuu kaavan 4 mukaiseen muotoon. (Gellert ym. 2021, 10.)

$$y'_t = c + \sum_{i=1}^p \phi_i y'_{t-1} + \sum_{h=1}^P \Phi_h y'_{t-m-h} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t + \sum_{k=1}^Q \Theta_k \varepsilon_{t-m-k} + \varepsilon_{T-m} \quad (4)$$

Kaavassa 4 on lisätty kaavaan 3 kausittaisuuden huomiointi termeillä  $P$ ,  $D$  ja  $Q$ , jotka esittävät autoregressiota, erotusmuunnosta ja liukuvaa keskiarvoa kausittaisten ajanjaksojen välillä.  $\Phi_p y'_{t-m-p}$  esittää kausittaista autoregressiota, mikä muodostetaan mallin harjoitusvaiheessa. Kausittaisen liukuvan keskiarvon osuus muodostetaan harjoitusvaiheessa ja on kaavassa 4  $\Theta_Q \varepsilon_{t-m-Q}$ . Kaavassa 4 muuttuja  $m$  kuvaa osoitettua kausittaisuutta, jonka luonnollinen valinta tuntisarjassa voisi olla 24, jos aikasarjassa päivän sisäistä kausittaisuutta havaittaisiin. Apumuuttujia käytettäessä kaavaan lisätään vielä ulkoisten muuttujien vaikutus  $\sum_{i=1}^n \beta_i x_t^i$ , missä  $n$  on ulkoisten muuttujien lukumäärä ja  $\beta$  ulkoisten muuttujien kerroin, mikä harjoitteluvaiheessa muodostetaan. Kaikki ominaisuudet sisältävän SARIMAX-mallin kaava on muotoa (kaava 5). (Gellert ym. 2021, 11.)

$$y'_t = c + \sum_{i=1}^p \phi_i y'_{t-1} + \sum_{h=1}^P \Phi_h y'_{t-m-h} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t + \sum_{k=1}^Q \Theta_k \varepsilon_{t-m-k} + \varepsilon_{T-m} + \sum_{i=1}^n \beta_i x_t^i \quad (5)$$

## 2.4 Ennustekyvyn arviointi numeeristen testien avulla

Säätösähkömarkkinat ovat siltä osin verrattavissa taloudellisten lukujen ennustamiseen, että ne tulee tehdä toistuvasti määritellyn ajanjakson välein eli reaaliaikaisesti. Säätösähkömarkkinoiden osalta ennuste tulee tehdä 45 minuuttia ennen tunnin alkamista (Fingrid 2021b). Tämä tarkoittaa, että ennustemalleja arvioitavissa ennustetarkkuuden lisäksi malleja vertaillaessa tulee huomioida laskennalliset vaatimukset ennustelle. Säätösähkömarkkinoiden aikasarjaa ennustettaessa luontaista on käsitellä suuri määrä dataa, sillä aikasarja muodostuu yksittäisistä tunneista, joita kertyy kuukausissa jo tuhansia. Laskennallista vaatimusta voidaan arvioida ennusteeseen kulunutta aikaa eri mallien välillä. (Nousi ym. 2017.)

Ennusteiden arviointia voidaan tehdä monella eri tavalla ja arviointitapojen monipuolinen käyttö antaa laajan kuvan ennusteiden suorituskyvystä ja käytettävyydestä. Aikasarjoille hyvin ominainen mittari on absoluuttinen keskiarvovirhe MAE (*Mean Absolute Error*), mikä antaa hyvän kuvan odotetusta virheestä, mutta ei ota kantaa virheen suuntaan (Gellert ym. 2021, 15). Absoluuttinen keskiarvovirhe lasketaan kaavalla 6.

$$MAE = \frac{1}{N} \sum_{t=1}^N |y_t - y'_t| \quad (6)$$

Kaavassa 6  $y_t$  kuvaa aikasarjan todellista arvoa ja  $y'_t$  ennustettua arvoa. Keskiarvovirheen lisäksi suosittu mittari on myös absoluuttisen virheen prosentuaalinen suhdeluku MAPE (*Mean Absolute Percentage Error*), joka voidaan laskea kaavan 7 avulla (Lomio & muut 2020, 6)

$$MAPE = \frac{100}{N} \sum_{t=1}^N \frac{|y_t - y'_t|}{y_t} \quad (7)$$

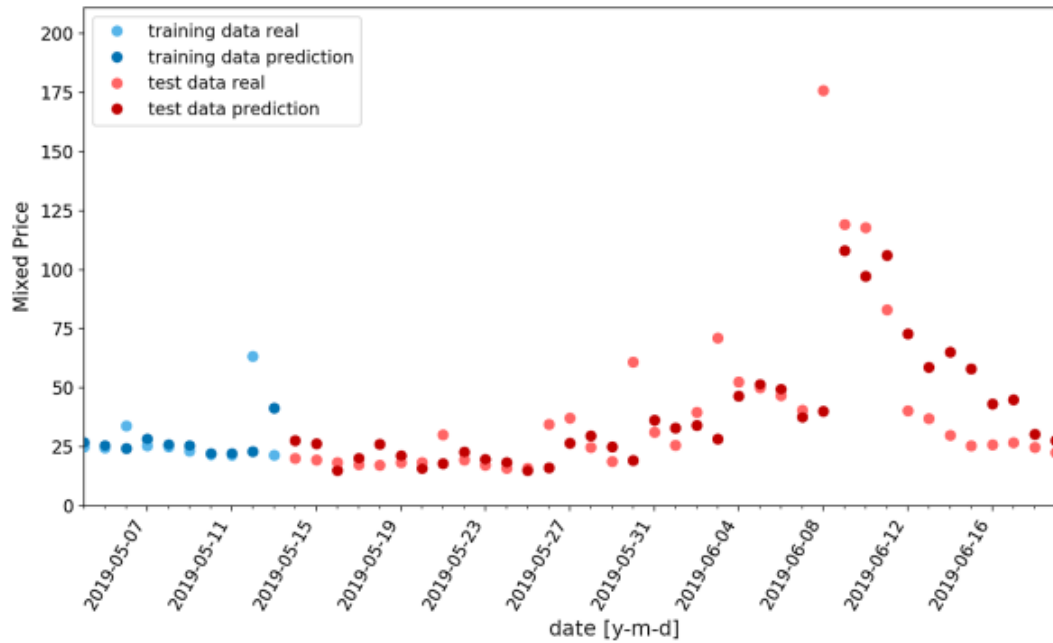
Mallien välisissä vertailuissa yleensä käytetään neliöllisen virheen keskiarvon neliöjuurta RMSE (*Root Means Squared*), sillä se ottaa tehokkaasti huomioon, myös yksittäiset suuret virheet. RMSE-virheen laskenta aikasarjassa on esitetty kaavassa 8 (Zhang & Guo 2020, 8).

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (y'_t - y_t)^2} \quad (8)$$

Zhangin ja Guon (2020, 18) esittivät myös, että mallin valinnan kannalta merkittävää on myös tutkia, kuinka kauan mallin harjoitus- ja ennustevaihe kestää sekunneissa käytettävissä olevalla laskentateholla.

## 2.5 Ennusteiden arviointi visualisoinnin ja residuaalin avulla

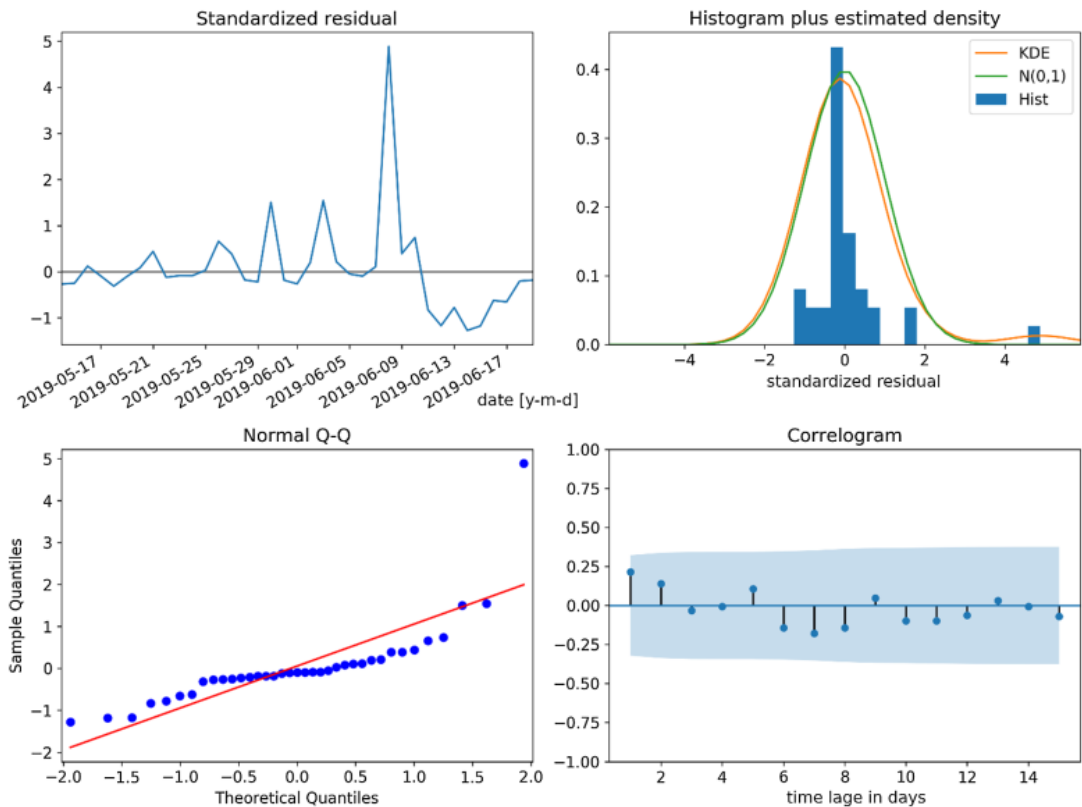
Aikasarjaennustamisessa tyypillistä on, että virheiden laatua voidaan arvioida usealla tilastollisella kuvaajalla (Merten ym. 2020, 14–15). Usealla arviointimenetelmällä voidaan todentaa yhdenmukainen arvio mallien kyvykkyydestä (Elamin & Fukushige 2018, 265). Yleensä aloitetaan vertailemalla aikasarjan ennusteiden ja vastaavien ajanhetkien todellisia arvoja niin sanotun etenevän ennustamisen (*Rolling Forecast*) avulla, jossa ennustuksen jälkeen testisarjan todelliset arvot lisätään seuraavaa ennustamista varten harjoitussarjaan (Dissanayake, Hemachandea, Laksitha, Haputhanthri & Wijayasiri 2021, 566; Merten ym. 2020, 14) (kuva 3).



KUVA 3. Esitys harjoitus- ja testisarjan todellisten tulosten ja ennusteiden vertailusta testiympäristössä (Merten ym. 2020, 14).

Merten ym. (2020, 14) ennustivat säätösähkömarkkinoista saatua hintaa SARIMA-mallilla (kuva 3). Kuvassa 3 on esitettyä sinisellä päivätasolla säätösähkömarkkinoiden ennustettua hintaa harjoitusjakson ja testijakson ajalta. Vaaleammalla sinisellä ja punaisella on todellisen aikasarjan arvot (*training data real* ja *test data real*) ja tummemmilla vastaavilla väreillä SARIMA-mallilla tehdyt ennusteet harjoitus ja testisarjalle (*training data prediction* ja *test data prediction*).

Malli näyttää ennustavan hyvin, sillä suuria poikkeuksia tummempien ja vaaleampien pisteiden välillä ei ole, mutta aikasarjaennustamisessa yleensä käytetään virheen arviointiin visuaalisen arvioinnin lisäksi tilastollisia kuvaajia, kuten aikasidonnaista virhettä eli residuaalia ja residuaalien jakaumaa. Residuaalien jakauman avulla nähdään, kuinka usein tehdään suuria tai pieniä virheitä. Näiden lisäksi voidaan käyttää niin sanottua neljännes kuvaajaa (*Quantile Quantile plot*), jossa residuaalien jakaumaa verrataan normaalijakaumaan. Aikasarjalle ominaisesta aikasidonnaisuudesta johtuen käytetään myös residuaalien autokorrelaatiokuvaajaa, jonka avulla voidaan havaita, jos virhettä tehdään toistuvasti tietyn ajanvälein, mikä on merkki siitä, ettei aikasarjan aikasidonnaisuutta ole täysin onnistuttu mallintamaan ennusteessa (kuva 4). (Merten ym. 2020, 15.)



KUVA 4. Tilastollisia kuvaajia aikasarjan ennusteen residuaalien tarkastelua varten (Merten ym. 2020, 15)

Merten ym. (2020, 15) residuaalien aikasidonnaisesta kuvaajasta (kuva 4 vasen yläkulma) huomataan, yksittäinen suurempi virhe ja kesäkuun puolen välin jälkeen ennusteet ovat olleet alakanttiin verrattuna ennen kesäkuun puoltaväliä ennusteet ovat pääosin olleet toteumia suurempia. Residuaalien jakaumasta nähdään, että suurin osa virheiden lukumäärästä ovat pieniä virheitä, mutta joitain suurempiakin virheitä on tehty. neljännos kuvaajasta nähdään, että residuaalit eivät ole täysin normaalisti jakautuneita, sillä silloin kuvaaja seuraisi punaisella merkittyä tasaisen jakauman viivaa. Kuvan 4 oikean alakulman residuaalien autokorrelaatiokuvaajasta huomataan, ettei merkittävää autokorrelaatiota voida todeta, kun sinisellä alueella on esitetty merkittävyyden raja-arvo, eli mallinnuksessa on onnistuttu mallintamaan kausittaisuusluonnetta, eikä virhettä toistuvasti tehdä. Kuvassa 4 kuvaajien arvot ovat skaalattuja lukuarvoja.

## 2.6 Lineaarisen ennustemenetelmän optimointi

Käytännössä parhaimmat ennustemenetelmät valitaan testisarjan suoriutumiskyvyn mukaan eri mittareilla (Elamin & Fukushige 2018, 264). Tilastollisilla malleilla, kuten SARIMAX-mallin johdannaisilla voidaan kuitenkin tehdä tilastollisiin lukuihin perustuvia arviota mallin sopivuudesta tai käytetyn datan riittävydestä etukäteen jo ennen

ennustetulosten tarkasteluja. Ennustemallille esitettäessä harjoitusdata voidaan laskea AIC-arvo (*Akaiques Information Criteria*), joka arvioi tilastollisesti ennustemallin kyvykkyyttä huomioiden käytettyjen parametrien lukumäärän aiheuttaman monimutkaisuuden. AIC-arvo on sitä pienempi mitä yksinkertaisempi malli on ja mitä paremmin ennustemalli mallintaa ennustettavaa arvoa, mikä ohjaa mahdollisimman kevyen ja tehokkaan mallin valintaan tilastollisin perustein (Elamin & Fukushige 2018, 265).

SARIMAX-mallin johdannaisilla ennustettaessa hyvin merkittävää on oikeiden AR- tai MA parametrien valinta (Moon, Hossain & Chon 2021, 1). Soveltuvia parametreja voidaan valita testaamalla eri parametrien yhdistelmiä ja tutkimalla tilastollisten mittarien, kuten aikaisemmin esitetyn AIC-arvon käyttäytymistä. AIC-arvon laskenta on esitetty kaavassa 9, jossa  $k$  on mallissa käytettyjen parametrien lukumäärä,  $N$  on esitetyn datan lukumäärä ja  $I$  on mallinnuksen aikaisen virhelaskennan tulos (Movahed, Bidgoly, Manes & Mirzaei 2021, 4).

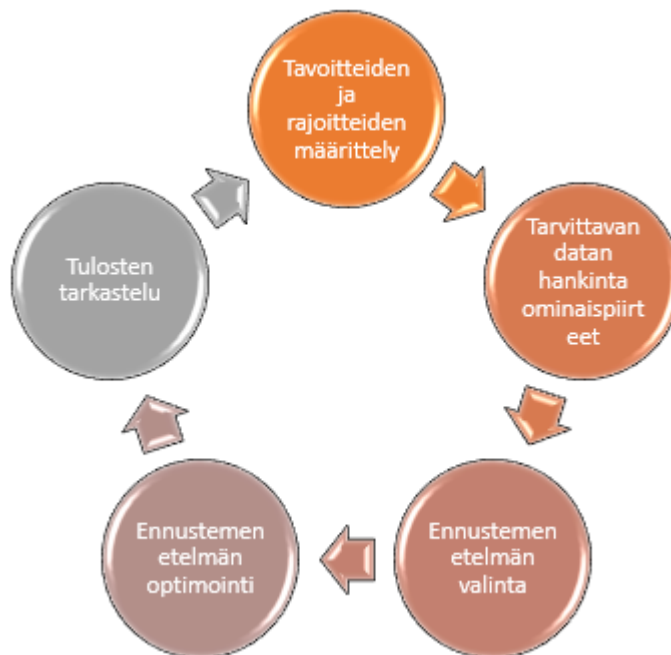
$$AIC = 2k + N \ln(I) \quad (9)$$

Kaavasta 9 huomataan, että AIC-arvon perusteella voidaan myös arvioida, kuinka paljon dataa harjoitusvaiheessa aikasarjan mallintamisessa tarvitaan ilman, että AIC-arvon tulos kasvaa. Movahed ym. (2021, 5–6) mukaan sopivat ARMA-mallin parametrit voidaan valita AIC-arvon kehittymistä tutkimalla ja parametrien vaikuttavuuksien todennäköisyyksien avulla.

AR- ja MA-parametrien lukumäärää voidaan arvioida myös aikaisemmin esitetyn autokorrelaation (*ACF*) ja sekä hieman vastaavan osittaisen autokorrelaation (*PACF*) avulla aikasarjan autokorrelaatio kuvaa edeltävien ajanhetkien korrelointia toteutuneeseen arvoon, joten sen perusteella voidaan arvioida liukuvan keskiarvon MA-parametrin monikertaa. Osittaisen autokorrelaation, missä edeltävien ajanhetkien vaikutuksista otetaan vain ne osat huomioon, jotka eivät vaikuta ajan hetkien välissä oleviin arvoihin avulla voidaan arvioida AR-parametrin monikertaa. (Banas & Ulrik-Banas 2021, 3) Harjoitusvaiheessa AIC-arvon lisäksi sopivan mallin valinnan tueksi saadaan tilastollisia todennäköisyys arvioita p-arvoina siitä mallintaako valitut SARIMAX-mallin parametrit, kuten AR- tai MA-monikerrat tai mahdolliset apumuuttujat aikasarjaa vai ei (Elamin & Fukushige 2018, 262).

### 3. TUTKIMUSMENETELMÄT JA AINEISTOT

Säätösähkömäärän ennustamisen tutkimus toteutettiin kuvassa 5 esitetyn tutkimussuunnitelman mukaisesti. Aluksi tutkimuksessa määriteltiin tavoitteita saavutetuille ennusteille, perehdyttiin ennusteen käytön kannalta keskeisiin ominaispiirteisiin sekä myös mahdollisiin rajoitteisiin, joita oli tunnistettavissa. Seuraavaksi selvitettiin, mistä ennustamista varten tarvittavaa dataa oli saatavissa, ja millaisin toimenpitein data oli tuotavissa ennusteympäristöön. Tässä yhteydessä tutkittiin myös datan ominaispiirteitä, kuten lähteestä saatavan datan eheyttä ja aikasarjaennustamisen kannalta merkittävää stationäärisyyttä. Kun datan haku ja vaadittavat muokkaustoimenpiteet olivat selvillä, pohdittiin työhön soveltuvia ennustemenetelmiä, jonka jälkeen tutkittuja ennustemenetelmiä pyrittiin optimoimaan datan ominaispiirteiden avulla. Tämän jälkeen ennustemallien suoriutumista mitattiin erilaisilla mittareilla, joiden kautta pyrittiin päästä ymmärrykseen soveltuvimmasta mallista.



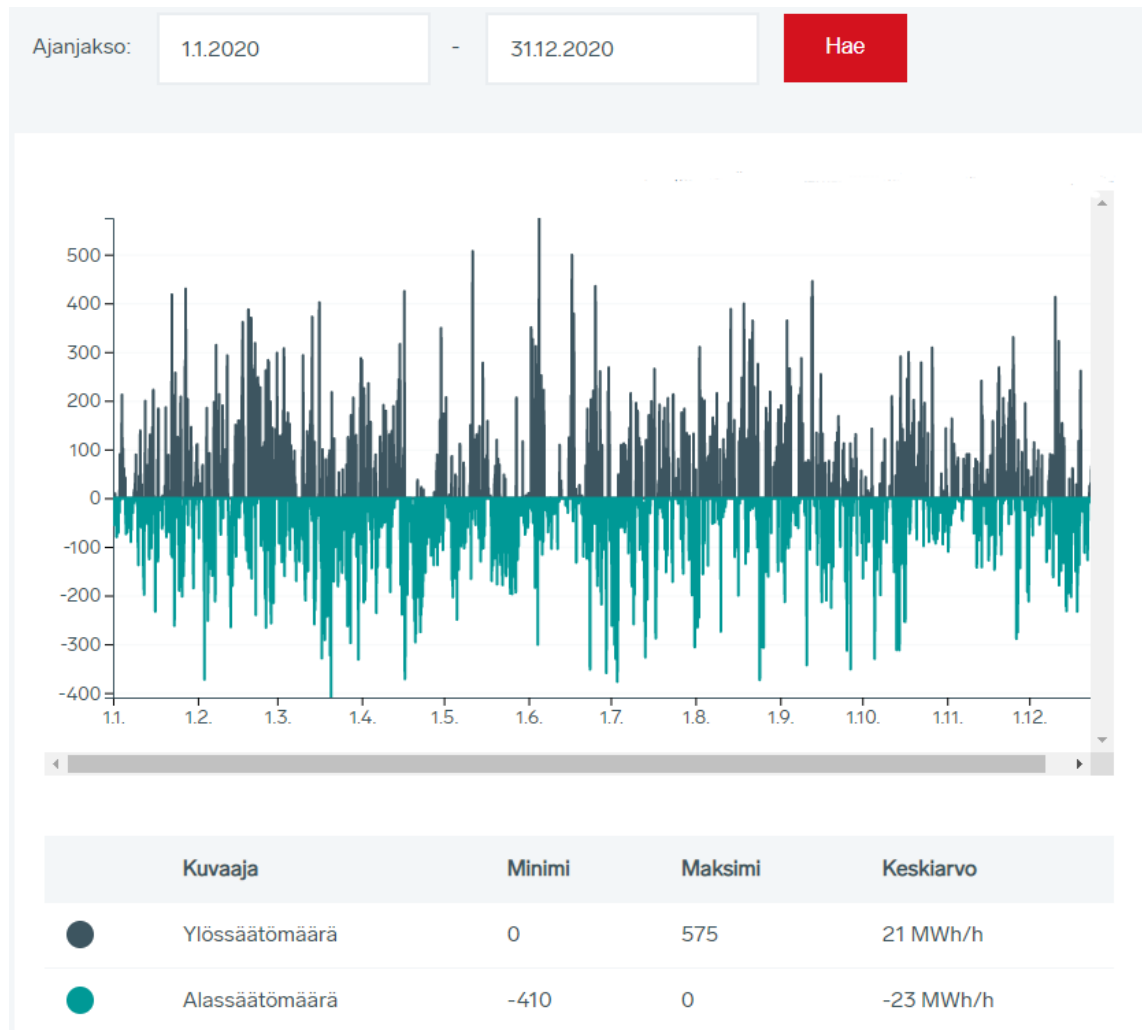
KUVA 5. Työssä kuvattu tutkimussuunnitelma soveltuvimman ennustemenetelmän saavuttamiseksi

Aikasarjaennustamisen kannalta haastavaa saattaa olla ennustemenetelmältä odotettavan riittävän tason määrittely. Tästä syystä viimeisessä vaiheessa tulee pohtia, kuinka ennustetta tullaan käyttämään, ja kuinka liiketoiminnan kannalta kriittisiä virheitä ennuste tekee, ja millaisia virheitä ennusteilta sallitaan. Tulee myös huomioida, että

tutkimussuunnitelma on esitetty kehänä siitä syystä, että tulevaisuudessa tutkimusta voitaisiin jatkaa asettamalla uusia tavoitteita ja rajoitteita ja vaihtaa ennustemenetelmää, kun jonkin menetelmän lakipiste on saavutettu.

### 3.1 Tavoitteiden ja rajoitteiden määrittely

Työn tavoitteena oli muodostaa aikasarjaennustamiseen sopiva malli, jonka avulla voitaisiin ennustaa säätösähkömarkkinoille aktivoituneiden säätöjen määrää sekä samalla näistä muodostuvaa säätösähkötyyppiä. Säätösähkömäärän ennustaminen voidaan siis nähdä numeerisen ennustamisen lisäksi myös luokitteluennusteena. Lisäksi ennustusmenetelmältä toivottiin selitettävyyttä, jotta ennustettavaa suuretta mahdollisesti voitaisiin mallin muodostamisen kautta ymmärtää paremmin. Säätösähkömarkkinoilla säätösähkön määrä on nähtävissä Fingridin sivuilla ja vuoden 2020 säätösähkömäärä on esitetty kuvassa 6 (Fingrid 2021g).



KUVA 6. Vuoden 2020 tilatun säätösähkön määrä säätösähkömarkkinoilla (Fingrid 2021g)



Kuvasta 6 nähdään, että ylössäädön määrä vaihtelee nollan ja 575 MW välillä, kun taas alassäätömäärä nollan ja -410 MW välillä. Vaihteluvälin ääriarvot näin ollen eivät poikke merkittävästi toisistaan, kuten ei myöskään ylössäätömäärän keskiarvo 21 MW/h alassäätömäärän keskiarvosta -23 MWh/h. Alassäädön tapauksessa miinusmerkillä kuvataan sähkön suuntaa verkon näkökulmasta kuluttajalle kuluttajan ylimääräisenä kulutuksena tai suunniteltua pienempänä tuotantona.

Selvästi suurimpia säätöjä pienemmistä keskiarvoista voidaan todeta, että säätömäärien sarjoissa yleisty pienempien säätötuntien lukumäärä sekä mahdollisesti tunnit, joilla säätöä ei tehdä ollenkaan. Säätötyypin luokittelun voisi tehdä määrittelemällä karkeasti, onko tunnilla tapahtunut ollenkaan säätöä kumpaankaan suuntaan. Tässä työssä päätettiin muodostaa luokittelu numeerisen ennusteen jatkotuotteena, jolloin luokittelurajalle päätettiin sallia myös jonkin verran nollasta poikkeamista, että pystyttäisiin luokittelemaan merkittäviä säätötunteja ja pienet säätötunnit voitaisiin yhdistää nollatuntien kanssa samaan luokkaan. Kuvan 6 perusteella merkittävän ylössäätötunnin raja määriteltiin lähtökohtaisesti yli 40 MW ylössäädöksi ja merkittävän alassäätötunnin raja -40 MW alassäädöksi. Näiden arvojen väliin jäävät säätömäärät luokitellaan nollatuntien kanssa yhteneväisiksi nollatunneiksi. Numeerisen ennustamisen tärkeimmäksi tavoitteeksi asetettiin, ettei ennuste olisi toteutuneeseen väärän suuntainen niin paljon, että se ylittäisi aikaisemmin määritellyt merkittävyyden rajat, jolloin sidottaisi suuri määrä energiaressurssia markkinalle, josta aktivointia ei tapahtuisi. Tämän lisäksi, jos ennustetusta alassäätötunnista tulisikin ylössäätö, jouduttaisiin alassäätötunnille suunnitellusta kulutetusta energiasta maksamaan normaalia sähkön hintaa korkeampi ylössäätöhinta.

Säätösähkömarkkinoille voidaan tehdä tuntikohtainen tarjous 45 minuuttia ennen alkavaa tuntia (Fingrid 2021b). Tästä syystä ennuste halutaan muodostaa seuraavalle tunnille ja ennuste tulee pystyä toistamaan joka tunti. Tästä syystä ennustamisen ei haluttu vievän liikaa aikaa tai laskennallisia resursseja. Ennusteille ei asetettu aikasarjaennustamisen suoriutumista mittaavien testien tavoitteita vielä tässä vaiheessa, vaan tavoitteeksi asetettiin ennusteiden monipuolinen mittaaminen, jotta virheitä voitaisiin ymmärtää mahdollisimman monesta näkökulmasta.

### **3.2 Datan hankinta**

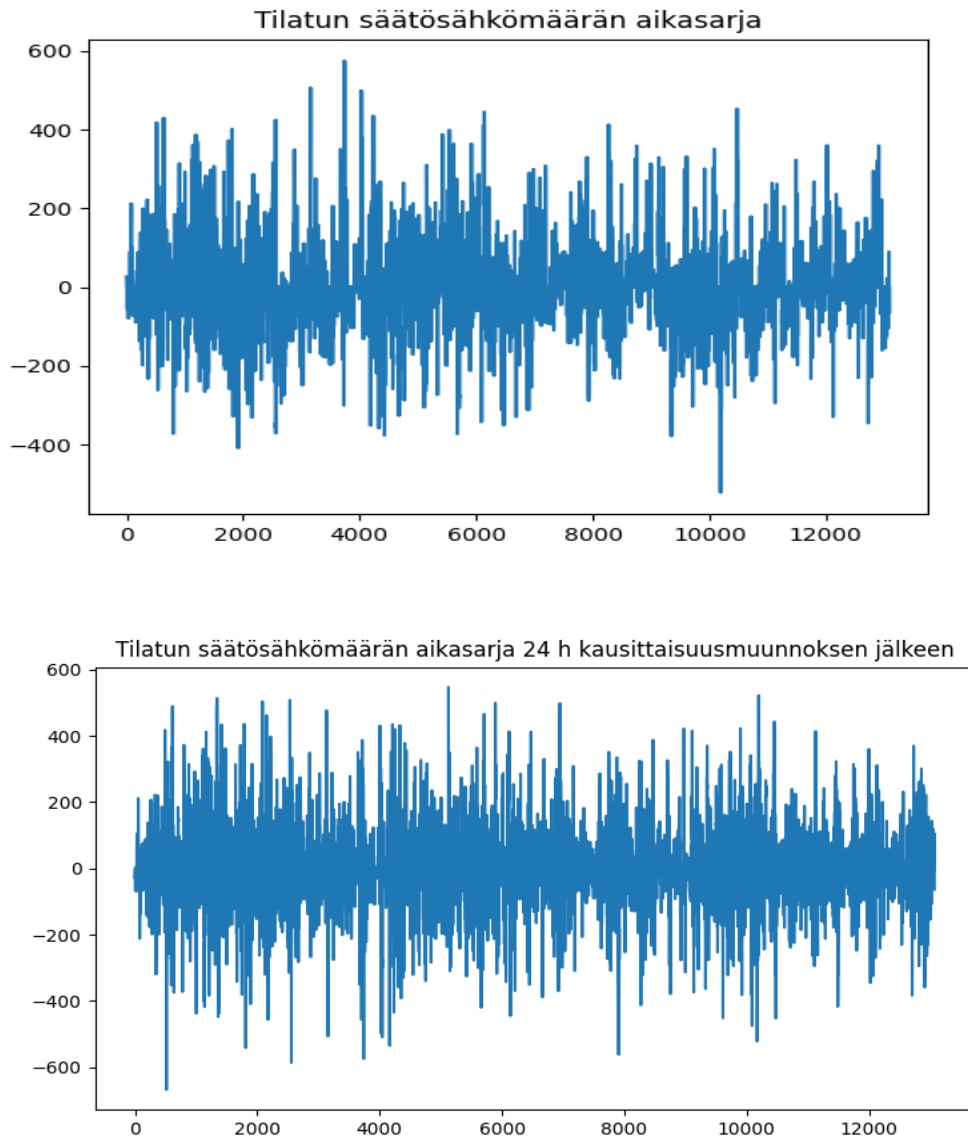
Sähkömarkkinoiden suureiden, kuten tilatun säätösähkömäärän ennustaminen on tehty mahdolliseksi Suomen kantaverkkoyhtiön Fingridin tarjoaman avoimen datan API-rajapinnan avulla, josta on saatavilla monipuolisesti sähköverkon tietoja koneluettavassa muodossa (Fingrid 2021f). Avoimen rajapinnan avulla voidaan suorittaa tiedon haku,

ennustemallin opettaminen ja ennustaminen avoimen lähdekoodin python-kielisen ohjelman avulla, kuten tässä työssä on tehty. Rajapintaan tulee tieto kahden tunnin viiveellä. Viiveestä johtuen ennustaminen on tehty viimeisimmästä saatavilla olevasta tiedosta kolmelle seuraavalle tunnille, joista ensimmäinen kuvaa ennusteajanhetkeä edeltävää tuntia, toinen ennusteajanhetkellä käynnissä olevaa tuntia ja kolmas sitä seuraavaa tuntia. Ennusteiden lopulliseen tallennussijaintiin ei oteta tässä työssä kantaa, sillä avoimen lähdekoodin ohjelmakielet tukevat tiedon tallentamista yleisesti käytettäviin tietovarastoihin, joista tieto on käytettävissä päätöksentekoa tai ohjausta varten.

### **3.3 Aikasarjan ominaispiirteet**

Kuten aikaisemmin kappaleessa kaksi todettiin, tulisi aikasarja todeta stationääriseksi ennen ennustetta, sillä moni lineaarinen aikasarjaennustamisen malli olettaa aikasarjan stationääriseksi. Tällöin aikasarja ei sisällä nousevaa trendiä tai selvää kausittaisuutta, jolloin ennustaminen on riippuvainen ajanhetkestä, milloin se suoritetaan. Aikasarjan sisältäessä kausittaisuutta tai nousevaa trendiä, yritetään kyseiset säännönmukaiset ominaisuudet poistamaan mallinnuksen ajaksi ja lisätä ne mallinnuksen jälkeen muodostuneisiin ennusteisiin, jolloin lopullinen ennuste käyttäytyy nousevan trendin tai kausittaisuuden mukaisesti.

Aikaisemmin tarkasteltiin kantaverkkoyhtiön sivustolta saatavaa sarjaa 2020 vuodelle, mutta kuvassa 7 on esitetty tutkimuksessa käytetty 13 000 tunnin eli n. 1,5 vuoden aikasarja 2020 vuoden alusta 2021 vuoden kesäkuun loppuun. Aikasarjoille saattaa olla luontaista aikasarjan tasosta riippuva kausittaisuus, kuten tuntisarjassa päivittäinen profiili tai päivätasolla viikoittainen. Kuvassa 7 on esitetty myös aikasarjan muunnettu muoto alkuperäisen sarjan alapuolella, jossa oletettu 24 tunnin kausittaisuus on poistettu vähentämällä ajanhetkeä 24 tuntia edeltävä arvo.

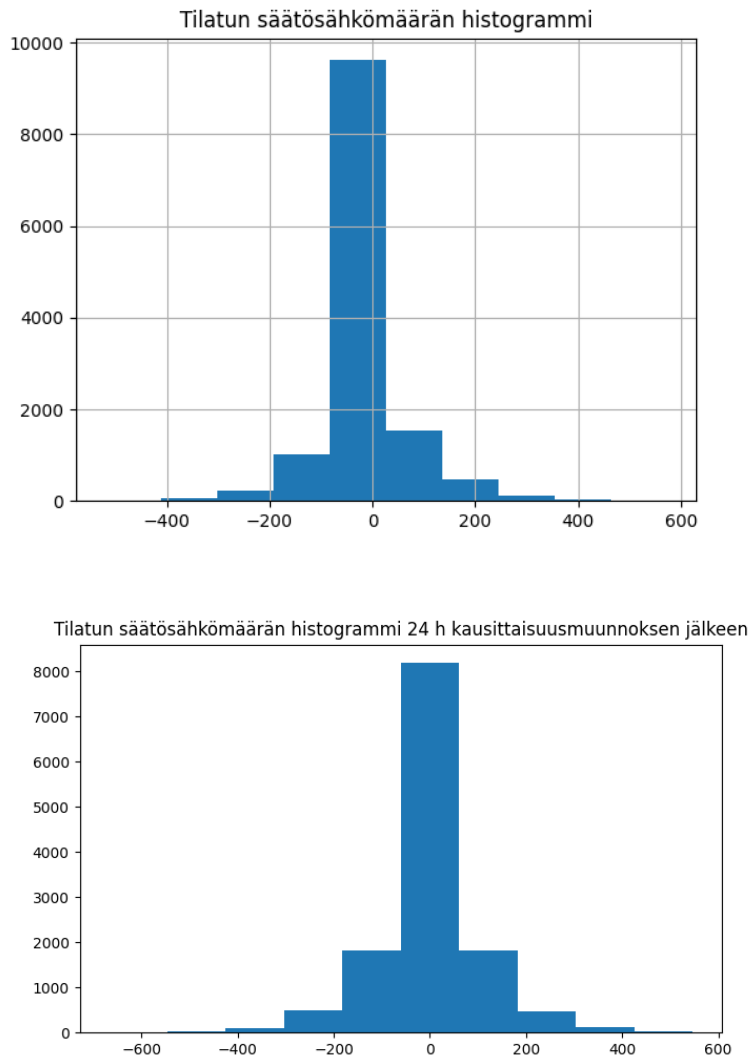


KUVA 7. Alkuperäinen säätösähkömäärän aikasarja sekä aikasarja 24 tunnin kausittaisuusmuunnoksen jälkeen,

Stationäärisyyttä voidaan aikasarjasta arvioida siten, muuttuuko keskiarvo ajan myötä, jos aikasarjan halkaisee kahtia. Vastaavanlaisesti voidaan tarkastella hajonnan muutosta. Tällöin sarjassa olisi havaittavissa nousevaa trendiä tai kausittaisuutta, mutta sellaista ei selvästi ole säätösähkömäärässä havaittavissa, vaikka 24 tunnin kausittaisuusmuunnos jätettäisiin tekemättä. Muunnoksen kanssa aikasarjan pienten arvojen sekä piikkien voidaan havaita hieman yleistyvän eli muunnoksella voi olla merkitystä stationäärisyyteen.

Keskiarvon ja hajonnan lisäksi stationäärisyyttä voidaan tarkastella aikasarjojen histogrammin avulla. Stationääriseen datan histogrammi on normaalisti jakautunut keskiarvon ympäristöön ja molemmin puolin vaihtelu on tasaista ja suurempien arvojen

esiintyminen harvenee tasaisesti. Kuvassa 8 on esitetty histogrammi alkuperäiselle aikasarjalle sekä alkuperäisen alapuolella aikasarjalle 24 tunnin kausittaisuusmuunnoksen jälkeen



KUVA 8. Säätösähkömäärän alkuperäisen aikasarjan sekä 24 tunnin kausittaisuusmuunnoksen aikasarjan histogrammi

Kuvasta 8 huomataan aikaisempaa kuvaa 7 tukeva havainto, että ilman 24 tunnin muunnostakin aikasarja histogrammin mukaan vaikuttaa tasaisesti jakautuneelta eli mahdollisesti stationääriseltä. Kuitenkin 24 tunnin kausittaisuusmuunnoksen jälkeen jakauma on entistä tasaisemmin keskiarvon ympäristöön jakautunut. Molemmissa histogrammeissa huomataan säätösähkömäärälle ominainen piirre, että tunteja, jolloin säätöä ei tehdä on huomattavasti muita enemmän. Nollatuntien suuri lukumäärä on merkittävä havainto myös ennustamisen kannalta.

Aikasarjan lukuarvojen jakaumaa voidaan tarkastella myös perinteisten aikasarjan tunnuslukujen avulla, jotka on esitetty taulukossa 1 alkuperäismuotoiselle aikasarjalle sekä 24 tunnin kausittaismuunnetulla aikasarjalle

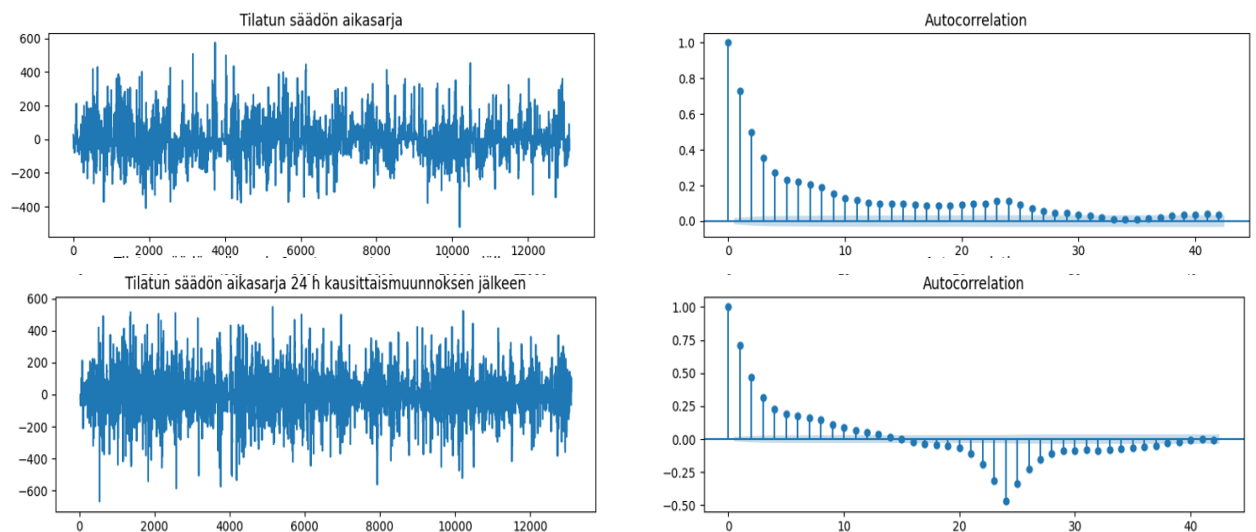
TAULUKKO 1. Säätosähkömäärän alkuperäisen ja kausittaismuunnetun aikasarjan tunnuslukuja

Tilastollinen tunnusluku	Alkuperäinen tilatun säätosähkömäärän aikasarja	24 tunnin kausittaismuunnettu tilatun säätosähkömäärän aikasarja
Aikasarjan tuntien lukumäärä	13 093 kpl	13 093 kpl
Keskiarvo	-2.75 MW	0.012 MW
Keskiarvon muutos, kun aikasarja jaetaan kahteen osaan puolesta välistä	-0.04 MW	0.02 MW
Keskihajonta	78.7 MW	104 MW
Keskihajonnan muutos, kun aikasarja jaetaan kahteen osaan puolesta välistä	2 568 MW	4338 MW
Vaihteluväli [Pienin arvo, Suurin arvo]	-522 MW, 575 MW	-667.8 MW, 547 MW
25 %, 50 %, 75 % Prosentuaalit	-15 MW, 0 MW, 0 MW	-35 MW, 0 MW, 35 MW

Taulukosta 1 voidaan tutkittavan aikasarjan stationäärisyyden kannalta havaita säädön tasapuolisuus ylös- ja alassäädön välillä vaihteluväliä tarkastellessa pienimmän arvon viitatessa suurinta tilattua alassäätömäärää ja suurimman osoittaessa suurinta ylösäätömäärää. Stationäärisyyden kannalta aikasarjojen keskiarvot eivät muutu merkittävästi, tehtiin aikasarjalle kausittaismuunnos tai ei. Ilman kausittaismuunnosta koko sarjan keskiarvo asettuu lähemmäksi nollaa verrattuna alkuperäiseen sarjaan, mikä saattaa olla merkki parantuneesta stationäärisyydestä. Toisaalta kun sarjat jaetaan keskeltä kahteen osaan ja tutkitaan keskihajonnan muutosta, on keskihajonnan muutos

kausittaisuunnetulla aikasarjalla alkuperäistä sarjaa suurempi. Molempien aikasarjojen prosentuaaliluvut viittaavat siihen, että aikasarjan mediaani eli 50 prosentuaali asettuu nollan ympäristöön ja mediaanin ympärillä yleistyy pienet arvot. Kausittaisuunnoksen jälkeen 25 prosentuaalin ja 75 prosentuaalin perusteella aikasarja asettuu nollan ympäristöön. Aikasarjojen ominaislukujen perusteella voidaan arvioida, että säätösähkömäärän aikasarja saattaa olla stationääristä ilmankin 24 tunnin kausittaisuunnosta, mutta muunnos saattaa tehdä siitä hieman stationäärisempää ainakin keskiarvon näkökulmasta.

Aikasarjan kausittaisuuden visuaaliseen tarkasteluun voidaan tunnuslukujen ja histogrammien lisäksi käyttää aikasarjan autokorrelaatiokuvaajaa kausittaisuuden tarkasteluun. Autokorrelaatio kuvaa kuinka edeltävät ajanhetket korreloivat nykyhetken kanssa (kuva 9).



KUVA 9. Säätösähkömäärän alkuperäisen aikasarjan autokorrelaatiokuvaaja sekä 24 tunnin kausittaisuunnetun aikasarjan autokorrelaatiokuvaaja.

Kuvan 9 aikasarjan autokorrelaatiokuvaajat ovat hyvin tyypillisiä autokorrelaatiokuvaajia siltä osin, että korrelaation arvo laskee tasaisesti mitä pidemmälle historiassa mennään. Alkuperäisessä aikasarjassa suurin korrelaation on ensimmäisellä tunnilla, jonka jälkeen korrelaatio pienenee tasaisesti mitä kauemmaksi historiassa mennään. Kausittaisuunnetun aikasarjan autokorrelaatiokuvaajassa ilmiö on vastaava, mutta tämän lisäksi 24 tunnin jälkeisen arvon korrelaatio kasvaa ja muuttuu vastakkaissuuntaiseksi, mikä saattaa olla merkki siitä, ettei 24 tunnin muunnos poista aikasarjasta oletettua kausittaisuutta onnistuneesti niin kuin pitäisi vaan sen sijaan luo aikasarjaan sellaisen. Pelkästään kuvaajien perusteella ei voida stationäärisyyttä varmuudella todeta, joten aikasarjalle tehtiin stationäärisyyden määrittämiseksi

tilastolliset ADF (*Augmented Dickey-Fuller*)- ja KPSS (*Kwiatkowski-Phillips-Schmidt-Shin*) -yksikköjuuritestit, joilla stationäärisyys voidaan tilastollisesti todeta alkuperäiselle sekä 24 tunnin kausittaismuunnetulle aikasarjalle. Testien tulokset on esitetty taulukossa 2.

TAULUKKO 2. Tilatun säätösähkömäärän tilastollisten stationäärisyystestien tulokset alkuperäiselle ja kausittaismuunnetulle aikasarjalle.

Augmented Dickey-Fuller (ADF) -testi			
Tilastollinen tunnusluku	Alkuperäinen aikasarja	säätösähkömäärän	24 tunnin kausittaismuunnettu säätösähkömäärän aikasarja
ADF-arvo	-17.05		-23.72
p-arvo	0.0000		0.0000
Kriittiset arvot [1 %, 5 %, 10 %]	-3.43, -2.86, -2.57		-3.43, -2.86, -2.57
Kwiatkowski-Phillips-Schmidt-Shin (KPSS) -testi			
Tilastollinen tunnusluku	Alkuperäinen aikasarja	säätösähkömäärän	24 tunnin kausittaismuunnettu säätösähkömäärän aikasarja
KPSS-arvo	0.14		0.0018
p-arvo	0.10		0.10
Kriittiset arvot [10 %, 5 %, 2.5 %, 1 %]	0.347, 0.463, 0.574, 0.739		0.347, 0.463, 0.574, 0.793

Taulukon 2 ADF-testitulosten perusteella voidaan todeta alkuperäisen sekä kausittaismuunnetun aikasarjan kumoavan testin stationäärisyyttä vastustava nollahypoteesi, yli 99 % prosentin todennäköisyydellä ADF-arvojen 17,05 ja -23,72 olevan reilusti alle kriittisten arvojen 1 % lukuarvon -3,43. KPSS-testin mukaan alkuperäisen, että kausittaismuunnetun aikasarjan KPSS-testin arvo on pienempi kuin kriittisten arvojen rajat, jotka ADF-testiin verrattuna ovat positiivisia ja ovat sitä

suurempia, mitä todennäköisemmin kumoavat nollahypoteesin eli aikasarjan stationäärisyysoletuksen. Taulukon 2 mukaan aikasarjat eivät p-arvon tilastollisen todennäköisyyden 0,1 mukaan kumoa stationäärisyyden nollahypoteesia vaan aikasarjat testin mukaan ovat KPSS-testin nollahypoteesin mukaisesti stationäärisiä.

### 3.4 Säättö sähkömäärän suunnan ennustaminen määrän avulla

Tutkimuksessa haluttiin tehdä numeerisen määrän ennustamisen lisäksi myös säättö suunnan luokittelua. Säättö sähkömarkkinan tilausten näkökulmasta luokittelu tehdään nollan välittömässä ympäristössä. Tilatun säädön ollessa alle nollan luokitellaan se markkinoilla alassäättö tunniksi, ja ollessa yli nollan luokitellaan se yössättö tunniksi. Taulukossa 3 on tehty aikasarjan säättö määristä luokittelua nollatason mukaisella rajauksella ja tutkittu tutkimussarjaan kuuluvien ylös- ja alassäättö määrrien ominaisarvoja.

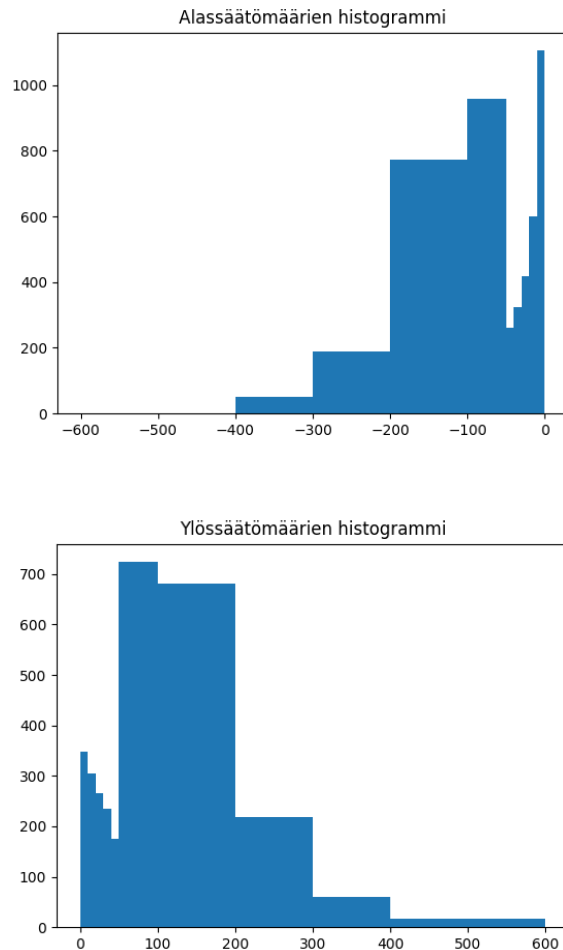
TAULUKKO 3. Ylös- ja alassäättö määrä aikasarjojen tunnuslukuja erikseen

Tilastollinen tunnusluku	Alkuperäinen tilatun säättö sähkömäärän yössättö tuntien aikasarja	Alkuperäinen tilatun säättö sähkömäärän alassäättö tuntien aikasarja
Aikasarjan tuntien lukumäärä	3031 kpl	4681 kpl
Keskiarvo	84.1 MW	-62.2 MW
Keskihajonta	81.1 MW	67.9 MW
Vaihteluväli [Suurin arvo, pienin arvo]	575.00 MW, 0.08 MW	-0.08, -522.00 MW
25 %, 50 %, 75 % Prosentuaalit	22.4 MW, 60.0 MW, 120 MW	-90.5 MW, -37.7 MW, -11.0 MW

Aikaisemmin esitettyssä taulukossa 3 havaittiin, että tutkittavia tunteja oli 13 093 kappaletta. Koostamalla taulukkoon 3 ylös- ja alassäättö tuntien lukumäärät voidaan havaita, että myös nollatunteja on merkittävä osa 5 381 kappaletta, kun yössättöjä on 3 031 ja alassäättöjä 4 681 kappaletta. Keskiarvot molemmissa on samaa luokkaa yössättöjen keskiarvon 81,1 MW ollessa hieman alassäättöjen keskiarvoa -62,2 MW itseisarvoltaan suurempi. Erottamalla taulukosta 3 säättö määrrien aikasarja ylös- ja alassäättö sarjoihin, voidaan todeta vaihteluvälin avulla, että alhaisimpia alas- tai yössättöjä on itseisarvoltaan 0,08 MW tunnit. Prosentuaalien avulla voidaan havaita



mediaanin asettuvan ylössäädöissä 60 MW ja alassäädöissä -37,7 MW arvoihin. Säätosähkömäärää ennustaessa työssä haluttiin antaa numeeriselle ennusteelle tilaa nollan ympäristöön, luokitteluvirheiden välttämiseksi ja keskittyä markkinoille osallistujan näkökulmasta merkittävien säätötuntien luokitteluun, sillä pienten säätötuntien luokittelua nollatunniksi ei pidetty merkittävänä ennusteiden käytettävyyden näkökulmasta. Taulukon 3 prosentuaalien avulla voidaan arvioida sopivaa rajaa nollan ympäristöön. Ylössäätötuntien 25 prosentuaali asettuu 22,4 MW ja vastaavasti jos alassäätötunteja tarkastellaan nollan näkökulmasta 0,008 MW ollessa aikasarjan suurin luku niin alassäätöjen 75 prosentuaali -11 MW antaa hyvää arviota sopivien rajojen valintaan. Tutkimuksessa otettiin luokittelurajan määrittelyn tueksi myös histogrammitarkastelu ylös- ja alassäätötunneista (kuva 10).



KUVA 10. Ylös- ja alassäätötuntien histogrammikuvaajat luokittelurajan määrittystä varten numeerisesta ennusteesta

Kuvan 10 histogrammit ilmaisevat kappalemäärissä, kuinka monta kertaa säätötunnin lukuarvo on tiettyjen rajojen välissä. Tilatun säätosähkömäärän histogrammissa kuvasta 8 havaittiin, että pienet säädöt molemmissa säätösuunnissa ovat yleisimpiä.

Histogrammeja kuvassa 10 on tästä syystä säädetty siten, että pienille lukuarvoille on asetettu pienemmät kymmenen välein olevat rajat, jotta saadaan parempi kuvaus pienien säätöjen jakaumasta sopivan luokittelurajan arviointia varten. Kuvasta 10 huomataan, että molemmissa säätötyypeissä itseisarvoltaan alle 10 MW säätöjä on selvästi eniten, joten nollaluokittelun raja nollan välittömään ympäristöön saattaisi aiheuttaa numeeriselle ennusteelle merkittäviä haasteita, kun otetaan huomioon mahdollinen ennustevirhe. Itseisarvoltaan pienten säätömäärien ilmaantuvuuden huomataan kuvan 10 histogrammissa laskevan 50 MW säätömäärää lähestyttäessä. Tutkimuksessa valittiin kuvan 10 perusteella merkittävän säätötyypin luokittelurajaksi itseisarvoltaan 40 MW säätömäärä säätötyypin luokittelun tutkimuksiin. Lisäksi haluttiin tutkia myös tarkemman luokittelurajan vaikutusta niin testausta tehtiin myös itseisarvoltaan 20 MW luokittelurajalla, jotta nähdään, kuinka luokittelurajan muutos vaikuttaa luokitteluennustetuloksiin.

### **3.5 Tutkimuksen testisarjojen määrittely**

Aikasarjaennustamisen testaus tehtiin tutkimuksessa historian toteutuneisiin tilattuihin säätömääriin perustuen. Tilatut ylös- ja alassäätömäärät ladattiin aikaisemmin mainitulta 1,5 vuoden aikajaksolta taulukkoon. Kerätyistä havainnoista ennustetestaukseen valitaan tietyt ajanjaksot, ja niistä muodostettiin ennustemallille harjoitus- ja testisarjat. Tutkimuksessa ennustetestaus on toteutettu pitkälle-, lyhyelle- ja erikoisajanjaksolle.

Pitkän ja lyhyen ajanjakson harjoitus alkaa 2020 lokakuun puolesta välistä. Lyhyen ajanjakson testissä malli harjoitettiin 2021 kesäkuun puoleen väliin ja ennusteen testisarja toteutettiin tästä hetkestä 200 tuntia eteenpäin eli kesäkuun 2021 loppuun. Lyhyellä testisarjalla nähdään mallin suoriutumista hieman yli viikon ajalta tuntitasolla, kun se harjoitettiin yli puolen vuoden harjoitusdatan avulla. Aikasarjaennustamiselle tyypillistä on, että se voi hetkellisesti suoriutua joinakin ajanjaksoina paremmin kuin toisina. Tästä syystä ennusteen suoriutumista testattiin myös pidemmällä testisarjalla, jonka testisarja tutkimuksessa on huhtikuun 2021 lopusta kesäkuun loppuun, jolloin saadaan tietoa myös kevään suoriutumisesta. Tuntitason ennusteen testisarjan pituutta valittaessa on hyvä huomioida, että testisarjan ajanjakso pitää pitkässä testijaksossa sisällään 1488 tuntia eli 1488 varmennettavaa ennustetta, jonka läpikäynti vie merkittävästi aikaa ja laskentatehoa.

Aikasarjaennusteiden suoriutumista haluttiin tarkastella myös erikoistilanteissa, kuten silloin kuin verkosta putoaa merkittävä energiantuottaja ja verkkoon tarvittava energia on oletettavasti suuri jonkin aikaa. Tällainen tapahtuma tunnistettiin, kun 10.12.2020 Olkiluodossa tapahtui häiriötilanne (Kuukkanen, Konttinen & Meritähti 2020), mikä

aiheutti säätösähkömarkkinoilla hetkellisesti suuren ylössäädön kysynnän säätösähkömarkkinoilla ja säätösähkön ylössäädön hinta oli hetkellisesti 3500 €/MWh, joka oli 66 kertaa suurempi kuin joulukuun 2020 keskihinta 53 €/MWh (Fingrid 2021g). Tästä syystä aikasarjaennustamisen testaamiseen otettiin tutkittavaksi myös tunnistettu erikoisjakso, jonka harjoitusjakso oli heinäkuusta 2020 joulukuun 2020 alkuun ja testijaksona koko joulukuu 2020. Erikoisjakson avulla voitiin arvioida kuinka ennustemenetelmän numeerinen ja luokitteluennuste suoriutuu myös suuren toimijan, kuten ydinvoimalaitoksen pudotessa sähköverkosta.

Ennustamisen tutkimustuloksissa on tärkeää huomioida, onko mallin harjoitukseen käyttämä aikasarja tilastollisilta tunnusluviltaan samanlainen kuin testisarjana toimiva aikasarja. Taulukossa 4 on kuvattu tutkimuksessa käytettävien harjoitus- ja testisarjojen tavanomaisia tilastollisia tunnuslukuja, joita koko aikasarjalle esitettiin taulukossa 1, sekä aikasarjojen stationäärisyyttä arvioivia ADF- ja KPSS-testien tuloksia samoin, kuin koko aikasarjalle kuvatut testitulokset taulukossa 2.

TAULUKKO 4. Harjoitus ja testisarjojen tilastollisia tunnuslukuja

Tilastollinen tunnusluku	Pitkän sarjan harjoitusjakso	Pitkän sarjan testijakso	Lyhyen sarjan harjoitusjakso	Lyhyen sarjan testijakso	Erikoissarjan harjoitusjakso	Erikoissarjan testijakso
Aikasarjan tuntien lukumäärä	4605 kpl	1488 kpl	5893 kpl	200 kpl	3500 kpl	820 kpl
Keskiarvo	-4.7 MW	5.6 MW	-2.3 MW	1.1 MW	4.3 MW	0.7 MW
Keskihajonta	71.8 MW	62.6 MW	69.2 MW	85.9 MW	78.5 MW	68.9 MW
Vaihteluväli [Pienin arvo, Suurin arvo]	-522 MW, 452 MW	-346 MW, 360 MW	-522 MW, 452 MW	-161 MW, 359 MW	-373 MW, 445 MW	-233 MW, 412 MW
25 %, 50 %, 75 % Prosentuaalit	-15 MW, 0 MW, 0 MW	-6 MW, 0 MW, 5 MW	-11 MW, 0 MW, 0 MW	-40 MW, 0 MW, 0 MW	-1 MW, 0 MW, 0 MW	-6 MW, 0 MW, 0 MW
ADF-arvo	-11.6	--9.8	-13.4	-3.8	-16.0	-11,6
ADF p-arvo	0.0000	0.0000	0.0000	0.0028	0.0000	0.0000
ADF Kriittiset arvot [1 %, 5 %, 10 %]	-3.43, -2.86, -2.57	-3.43, -2.86, -2.57	-3.43, -2.86, -2.57	-3.46, -2.88, -2.58	-3.43, -2.86, -2.57	-3.44, -2.87, -2.57
KPSS-arvo	0.86	0.86	0.45	0.63	0.12	0.38
KPSS p-arvo	0.10	0.10	0.05	0.019	0.1	0.086
KPSS Kriittiset arvot [10 %, 5 %, 2.5 %, 1 %]	0.347, 0.463, 0.574, 0.739	0.347, 0.463, 0.574, 0.793	0.347, 0.463, 0.574, 0.739	0.347, 0.463, 0.574, 0.739	0.347, 0.463, 0.574, 0.739	0.347, 0.463, 0.574, 0.739

Tilastollisia tunnuslukuja ja stationäärisyystestien tuloksista ei havaittu merkittäviä poikkeamia harjoitus- ja testijaksoparien välillä, vaan aikasarjat olivat tilastollisten tunnuslukujen mukaan hyvin samankaltaisia keskenään. Aikasarjalle tyypillisesti keskiarvot olivat itseisarvoltaan pieniä, verrattaessa vaihteluvälin esittämiin pienimpiin ja suurimpiin arvoihin eli suurimpiin ylös- tai alassäätömääriin. Keskihajonnat ja prosentuaalit olivat myös hyvin samankaltaisia sarjojen välillä, kuten myös ADF- ja

KPSS-stationäärisyydestien tulokset. Merkittävimmät tunnuslukujen poikkeamat havaittiin tutkittavan lyhyen tutkimusjakson harjoitus- ja testisarjan tulosten välillä, sillä harjoitussarja oli merkittävän pitkä 5893 tunnin sarja, kun taas lyhyt testisarja koostui vain reilu viikon 200 tunnista. Aikasarjojen otantajoukkojen lukumäärästä johtuen vaihteluvälit olivat eri suuruiset, sillä suuret ylös- tai alassäätötunnit eivät osuneet testisarjaan toisin kuin suureen harjoitussarjaan. Stationäärisyydestien mukaan kaikki sarjat voitiin todeta stationäärisiksi, mutta alhaisin KPSS-testin p-arvo ja vastaavasti korkein ADF-testin p-arvo havaittiin lyhyen tutkimusjakson testisarjalla, mikä viittaa stationäärisyyden tilastolliseen epävarmuuteen, ADF- ja KPSS-arvojen kuitenkin todentaen sarjan riittävän stationääriseksi kriittisiin arvoihin verratessa. Taulukon 4 kaltaista tunnuslukujen tarkastelua harjoitussarjoille voidaan hyödyntää tulevaisuudessa ennustemenetelmiä käytettäessä aina ennen ennusteen muodostamista. Tilastolliset tunnusluvut antavat käyttäjälle varmuutta ennusteiden tilastollisista perusteista ennen kuin varsinaisia toteumia on saatavilla. Seurattaessa tilastollisia tunnuslukuja ennusteiden yhteydessä voidaan myös tunnistaa syitä virheellisille ennusteille.

### 3.6 Ennustemenetelmän valinta

Tutkimuksessa säätösähkömarkkinoiden säätömäärän ennustamiseen valittiin lineaarinen tilastollinen ennustemalli ARIMA ja tämän kolme muunnelmaa SARIMA, ARIMAX ja SARIMAX. Mallin valinta perustui ARIMA-mallin aikaisempien tutkimusten osoittamaan potentiaaliin aikasarjaennustamisessa sekä sen lineaarisuudesta johtuvaan selitettävyyteen.

Tilattujen säätösähkömäärien ennustamisesta on tehty vain vähän tutkimusta, eikä suureeseen vaikuttavien ulkopuolisten tekijöiden vaikutusta ole tutkittu, josta syystä ei ensisijaisesti valittu monimuuttujamenetelmään perustuvaa ennustemallia. Kuitenkin ARIMA-mallin X-parametrin käyttöönotto mahdollistaa myös oletuksena lineaaristen suhteiden omaavien apumuuttujien käytön ennusteessa. Aikaisemmin esitetystä aikasarjan stationäärisyyden selvityksessä tultiin johtopäätökseen, että säätösähkömäärän aikasarja alkuperäisessä muodossa voidaan todeta stationääriseksi, mutta huomattiin myös, että 24 tunnin kausittaisuusmuunnos teki aikasarjasta yhä stationäärisempää. Tästä syystä tutkimukseen otettiin myös ARIMA-mallin S-parametri, joka huomioi sille osoitettujen parametrien mukaisesti sille osoitetun kausittaisuuden, jolloin ARIMA-malli mukautuu muotoon SARIMA. S-ominaisuudella voidaan tutkia työssä oletetun 24 tunnin kausittaisuuden mahdollinen vaikutus ennusteeseen.

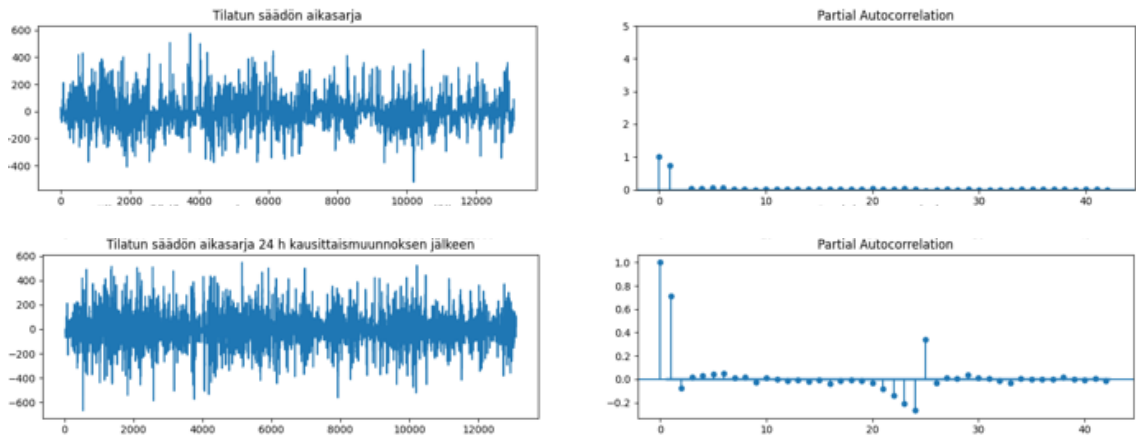
Säätösähkömäärän ja tyyppin ennustamisen tutkittavaksi apumuuttujaksi valittiin kulutustasesähkö, joka määräytyy tavallisen sähkön hinnan mukaan, jos säätöä ei tehdä,

mutta lisäksi säätötyypin osoittaman säätösähkömarkkinan hinnan mukaan, jos säätöä tapahtuu. Apumuuttujien kanssa merkittävä huomio oli, että apumuuttujien ennustearvoja tulee välittää ennustemallille yhtä monta ajanhetkeä kuin ennustetaan. Jos ennusteita ei ole saatavissa, niin voidaan käyttää edeltävien ajanhetkien arvoa ennustavina tekijöinä olettaen, että seuraavat tunnit korreloivat edeltävien kanssa myös apumuuttujien suhteen. Tässä tutkimuksessa on käytetty ennustevaiheen aikana saatavilla olevaa tietoa eli 2 tunnin viiveellä rajapinnasta saatavaa kulutustasesähkön tietoa. Lisäksi SARIMAX-mallilla voidaan testata, paraneeko ennuste, kun sekä 24 tunnin kausittaisuusoletus sekä tieto kulutustasesähkön edeltävistä hinnoista on esitetty mallille.

### 3.7 Ennustemenetelmän optimointi

Tutkimuksessa päädyttiin ARIMA-mallin ja sen johdannaisten valintaan myös siitä syystä, että mallin optimoitavien parametrien lukumäärä on pieni ja parhaimpia parametreja voidaan arvioida tehokkaasti etukäteen tilastollisten testien perusteella, mikä helpottaa optimointiprosessia monimutkaisempien neuroverkkomallien rakennesuunnitteluun verrattuna.

Aikaisemmin esitetystä kaavassa 3 on esitetty ARIMA-mallin rakenne, ja kuinka ARIMA-malli koostuu parametreista  $p$ ,  $d$ , ja  $q$  eli ARIMA-mallin muodostama funktio on  $ARIMA(p,d,q)$ . Kuten kaavan 3 yhteydessä mainittiin, voidaan  $d$ -parametrillä huomioida aikasarjan arvon kasvu ajan myötä muodostamalla ennen ennustetta tuntien erotus ja lisäämällä ne jälkeenpäin ennusteeseen. Säättösähkömäärän aikasarjan kuvaajasta (kuva 7) huomataan, ettei nousevaa trendiä ole, joten  $d$ -parametri voidaan olettaa nolaksi. Kuvasta 9 huomattiin alkuperäisen aikasarjan autokorrelaation olevan suurimmillaan 1-viiveen kohdalla, ja olevan voimakas neljänteen viiveeseen asti, jonka jälkeen korrelaatio hidastaa laskuaan, kunnes kahdeksannen viiveen jälkeen korrelaatio pienenee jälleen voimakkaammin. 24 tunnin kausittaisuunnetun aikasarjan tapauksessa kuvio on hyvin samanmuotoinen, mutta 24 tunnin jälkeen koetaan ympäristöä voimakkaampi vastakkainen autokorrelaatio. Kuvan 9 kuvaajista voitaisiin etukäteen arvioida ARIMA mallin soveltuviksi  $q$ -parametreiksi neljä tai kahdeksan, sillä  $q$ -parametri kaavan mukaan määrittää huomioitavia viiveitä liukuvan keskiarvon MA-kertoimen laskentaan, mikä lasketaan harjoitusdatasta ennen ennustetta. ARIMA-mallin autoregressiota kuvaava parametri  $p$  voidaan vastaavasti määrittää osittaisen korrelaation (*partial autocorrelation*) kuvaajasta, mikä on esitetty kuvassa 11 alkuperäiselle aikasarjalle sekä 24 tunnin kausittaisuunnetulle aikasarjalle.



KUVA 11. Säättö sähkömäärän osittainen autokorrelaatio kuvaaja alkuperäiselle ja 24 tunnin kausittaisuunnnetulle aikasarjalle

Kuvan 11 perusteella voidaan etukäteen määrittää, että soveltuva parametri ARIMA-mallin soveltuva parametri  $p$  olisi yksi niin kuin myös kausittaisuunnoksen huomioivalle SARIMA-mallille. Tämän lisäksi ennen kuin suoritettiin aikaisemmin mainittuja ennustetestejä eri ajanjaksoille, tilastollisia malleja arvioitiin AIC-kriteerin perusteella, jonka tulokset on esitetty taulukossa 5.

TAULUKKO 5. Tutkimuksessa testattavien mallien AIC-testin kriteeriarvot

Tutkittava malli $f(p,d,q)$	AIC-kriteeri
ARIMA (1,0,4)	141336.8
ARIMA (1,0,8)	141315.2
SARIMA (1,0,4)	141236.7
SARIMA (1,0,8)	141214.9
ARIMAX (1,0,4)	140032.1
ARIMAX (1,0,8)	140011.9
SARIMAX (1,0,4)	139900.2
SARIMAX (1,0,8)	139884.2

Taulukosta 5 huomataan, että AIC-kriteerin arvo laskee kasvattamalla  $p$  ja  $q$  parametrien arvoa, jotka määrittävät montako viivettä huomioidaan autoregression (AR) tai liukuvan keskiarvon (MA) kertoimia laskettaessa. Tämän lisäksi myös 24 tunnin kausittaisuuden mukaan ottaminen S-parametrilla sekä kulutustaseen mukaan ottaminen apumuuttujana X-parametrilla parantaa mallin soveltuvuutta huomioiden mallin samanaikainen

monimutkaistaminen. Kaikista alhaisin eli soveltuvin AIC-arvo saadaan monimutkaisimmalla SARIMAX (1,0,8) -mallilla, jossa 24 tunnin kausittaisuus sekä kulutustase apumuuttujana on huomioitu.

Mallin monimutkaisuus vaikuttaa laskennassa käytettävään aikaan ja tarvittavaan laskentaresurssiin. Jos autokorrelaatiokuvaajien sekä testitulosten perusteella on haastava tehdä valintaa soveltuvien parametrien välillä, on suotavaa valita AIC-kriteerin tai muiden testien perusteella tehokkain ja kevyin malli.

### **3.8 Ennustemenetelmien tulosten tarkastelu numeeristen testien avulla**

Tehokkaan optimoinnin ja vähäisten säädettävien parametrien ansiota, voidaan SARIMAX-ennustemallilla ja sen johdannaisilla saavuttaa tehokkaasti piste, jolloin voidaan todeta, ettei ennustetta voida enää parantaa parametreja lisäämällä. Jos tulosten tarkastelun jälkeen voidaan todeta puutteellinen taso, on luonnollista siirtyä toiseen ennustemenetelmään.

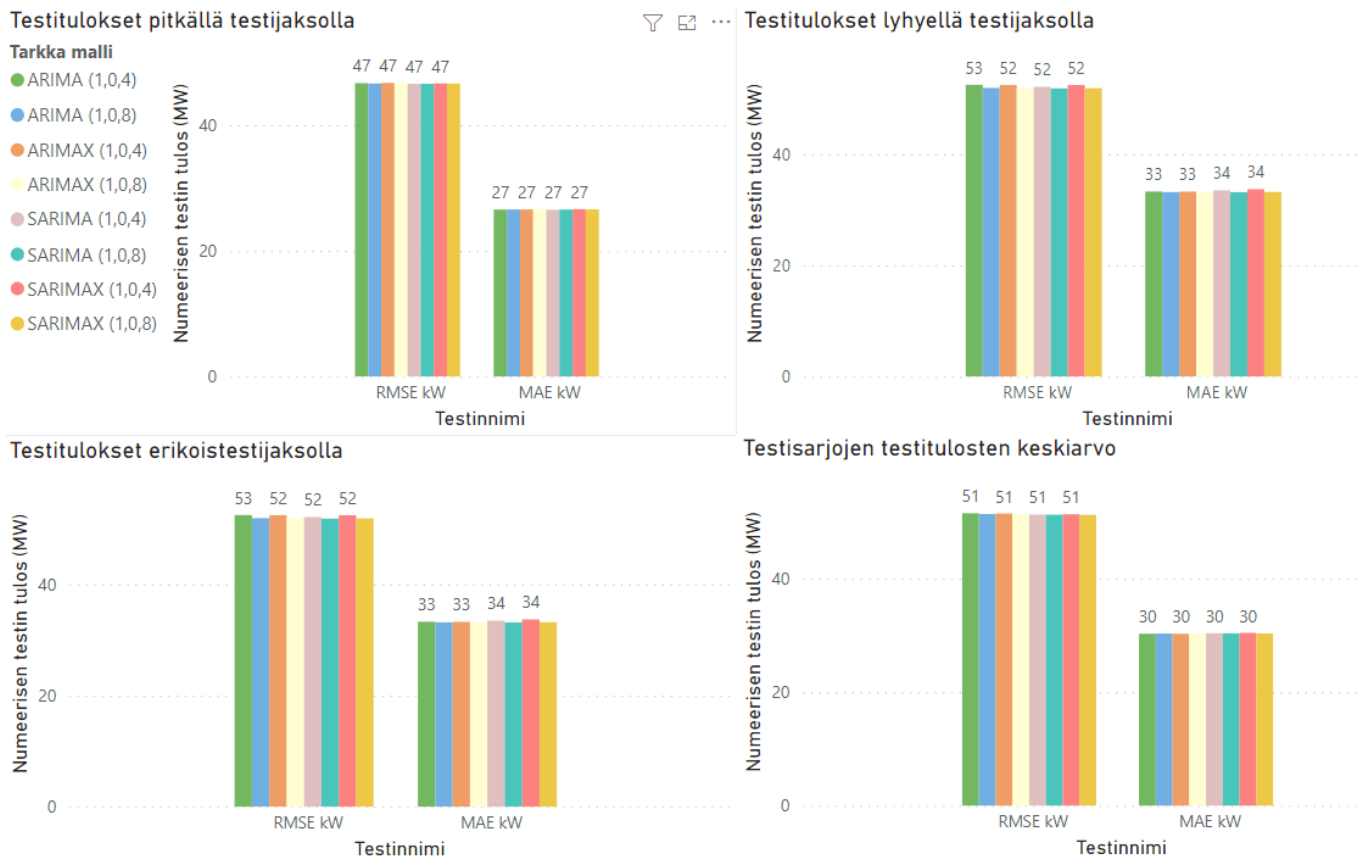
Aikaisemmin mainittu AIC-kriteeri on vain yksi tilastollinen tapa arvioida mallin sopivuutta huomioiden parametrien lukumäärän. Ennusteiden käytettävyyden kannalta on tärkeää arvioida ennusteiden keskivirhettä (MAE) ja neliöllisen keskivirheen neliöjuurta tai luokitteluongelmassa luokitteluvirheen onnistumisen suhdelukua tai luokkakohtaisten ennusteiden suhdelukua. Näin ennustemallin käyttäjän on helpompaa arvioida, millaista virhettä ennusteella on odotettavissa. Kaavassa 7 esitettiin myös prosentuaalisen keskivirheen (MAPE) laskenta, joka antaa numeerisen ennusteen prosentuaalisen virheen. Kaavasta 7 huomataan, että jakajana on aikasarjan keskiarvo. Säättösähkömäärän keskiarvo on lähellä nollaa, kun säätöä tehdään molempiin suuntiin. Tästä syystä MAPE-virhettä ei voida säätösähkömäärän ennustevirheiden arvioinnissa käyttää.

Tutkimuksessa aikaisemmin kuvatuille malleille tehtiin arviointia numeerisen ennustekyvyn mukaan. Tämän lisäksi numeerisen ennusteesta jalostettua luokitteluennustetta tutkittiin luokittelua erinäisten suhdelukujen avulla ja lisäksi ennusteelle tehtiin lisätestejä, jotka vaikuttavat vertailussa olevista malleista soveltuvimman valintaan. Liitteessä 1 on esitetty alkuperäisessä muodossa tulosten mittauspöytäkirja, josta merkittävät tulokset on visualisoitu Power BI -raportointityökalulla tähän kappaleeseen.

Kuvassa 12 on esitetty ennusteen numeerisen ennustekyvyn arviointia MAE- ja RMSE-laskennan avulla tutkimuksen testijaksoille sekä Power BI -raportointityökalun avulla eri



testijaksojen tuloksista laskettiin keskiarvo kuvan oikean alakulman kuvaajaan antamaan vielä pitkän testisarjan kuvaajaa yleistävämpi kuvan ennustemallien suoriutumisesta eri testisarjoilla.



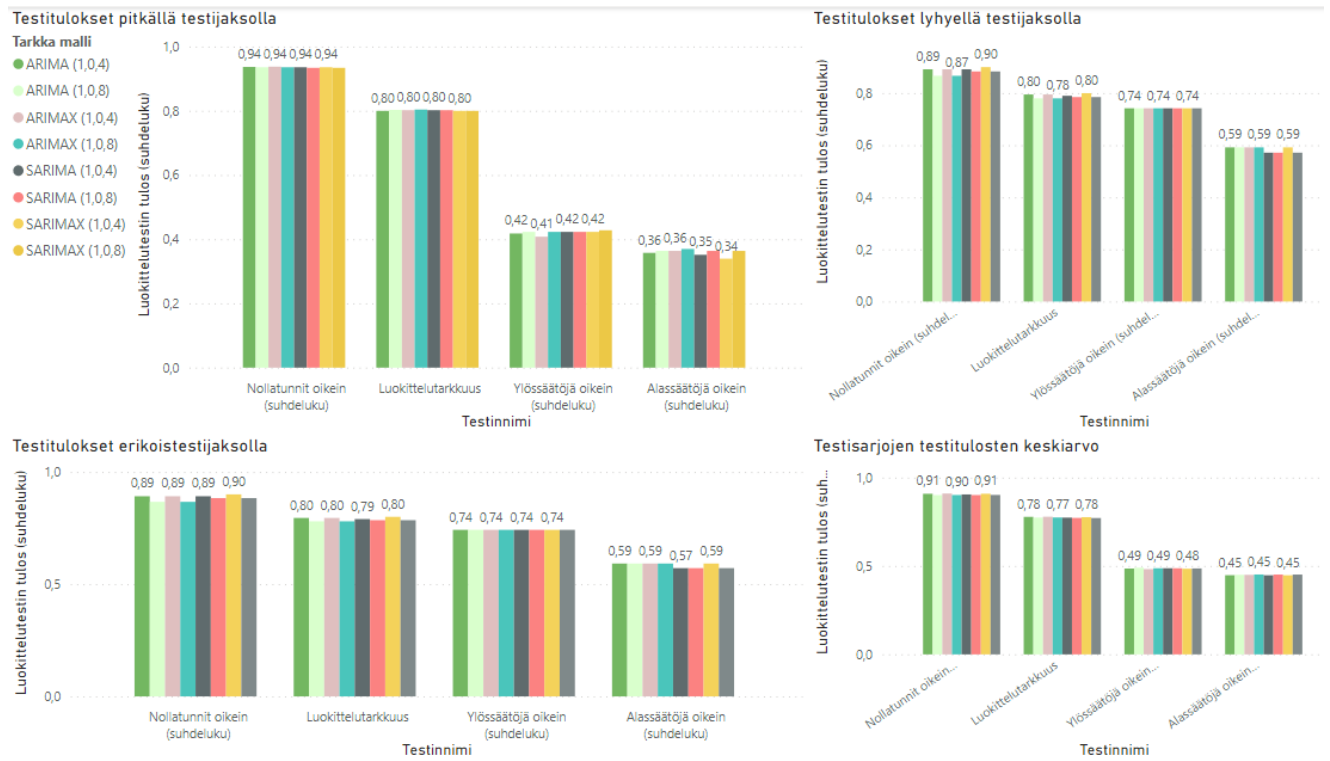
KUVA 12. Ennustemallien numeerisen ennustamisen testitulokset eritestisarjoille ja testisarjojen yhteistulosten keskiarvo

Kuvan 12 perusteella pylväskaavioiden muodosta voidaan todeta mallien yhdenvertaisuus niin MAE- kuin RMSE-virheen mukaan. RMSE-virhe kaavan 8 mukaan laskee neliöjuuren neliöllisten virheiden keskiarvosta, jonka vuoksi sitä yleensä käytetään mallien väliseen vertailuun korostaen suuria poikkeamia MAE-virhettä voimakkaammin. MAE-virhe sen sijaan laskee virheiden keskiarvon kaavan 6 mukaan, jolloin virhearvio saadaan sellaisessa mittakaavassa, mitä käyttäjä voi mahdollisesti ennustemallilta odottaa.

Parametrien vaikutuksia voidaan tarkastella matemaattisesti yksinkertaisimman ARIMA (1,0,4) -mallin (kuvan 12 vihreä pylväs) ja monimutkaisimman SARIMAX (1,0,8) -mallin (kuvan 12 oranssi pylväs) tulosten avulla. Kaikkien mallien virheet asettuvat pitkällä testijaksolla (kuvan 12 vasen yläkulma) 47 MW RMSE- ja 27 MW MAE-virhetasolle mikä osoittaa mallien testitulosten eroavaisuuksien olevan pieniä, mikä osoittaa tasavertaisuuden numeerisessa ennusteessa näillä mittareilla. Myös muilla testijaksoilla

virhelaskentojen tulokset ovat samassa suuruusluokassa, mikä osoittaa mallien tasavertaista suoriutumista.

Numeerisesta ennusteesta muodostettiin merkittävän ylös- ja alassäätötunnin luokittelua 40 MW luokittelurajan mukaisesti. Vastaavien ajanjaksojen testisarjojen toteumilla tehtiin samanlainen luokittelumääritys. Kuvassa 13 on esitetty luokittelutestien tuloksia yleisen luokittelutarkkuuden suhdeluvun avulla sekä myös luokkakohtaisten luokittelusuhdelukujen avulla.



KUVA 13. Ennustemallien luokittelutestien tulokset eri testisarjoille ja testisarjojen yhteistulosten keskiarvo

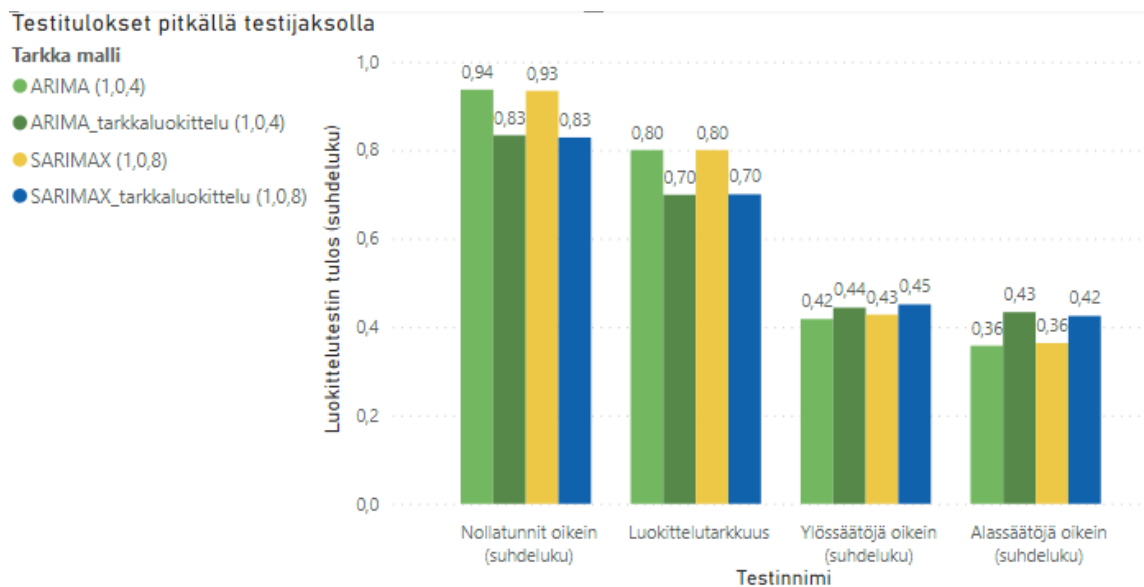
Myös kuvasta 13 huomattiin, ettei luokitteluvirheiden näkökulmasta mallin monimutkaistaminen 24 tunnin kausittaisuuden tai kulutustasehinnan avulla edistä luokittelutuloksia. Suurin luokittelutarkkuuden vaihteluväli ilmeni lyhyen testisarjan ARIMA (1,0,4) ja ARIMA (1,0,8) mallien välillä yleisen luokittelutarkkuuden suhdeluvun ollessa 0,80 ARIMA (1,0,4) -mallilla ja 0,78 ARIMA (1,0,8) mallilla. Sama vaihteluväli huomattiin lyhyellä testijaksolla ARIMAX (1,0,4) -mallilla, jonka tulos oli 0,80 ja ARIMAX (1,0,8) -mallilla, jonka tulos oli 0,78. Pitkällä testijaksolla kaikki mallit suoriutuivat yleisen luokittelutarkkuuden näkökulmasta yhtä hyvin. Kulutustasehinnan ja 24 tunnin kausittaisuuden lisääminen malleihin ei merkittävästi parantanut ennusteita.

Kuvan 13 luokkakohtaisten ennusteiden onnistumisen suhdeluvusta huomataan, että kaikki mallit ennustavat erittäin hyvin 40 MW merkittävyyden luokittelurajalla määritellyt

nollatunnit 0,91 suhdeluvulla, kun tarkastellaan kuvasta 13 oikean alakulman kaikkien testisarjojen keskiarvoa. Ylössäädöt ennustetaan hyvin tasaisesti keskimäärin 49 prosentin luokittelutarkkuudella ja alassäätötunnit keskimäärin 45 prosentin luokittelutarkkuudella.

Selvästi alentuneista ylös- ja alassäätötuntien luokittelutarkkuuksista verrattuna nollatuntien luokittelutarkkuuksiin, voidaan päätellä että 40 MW luokittelurajan ympäristössä säätöä tehdään paljon, ja merkittävän usein ennuste arvioi luokittelun ylös- tai alassäätötunniksi, mutta todellisuudessa säätö jää hieman merkittävyyden rajan alapuolelle. Tulokset antavat myös merkkiä siitä että 40 MW säätörajojen välissä tunteja on paljon verrattuna merkittäviin säätötunteihin, sillä matalat luokkakohtaiset luokittelutarkkuudet eivät alenna kuitenkaan merkittävästi yleistä luokittelutarkkuutta.

Tutkimuksessa haluttiin tarkastella edeltävistä tuloksista johtuen, mikä olisi tiukemman merkittävän säätötunnin raja-arvon vaikutus luokitteluennusteeseen. Tarkaksi luokittelurajaksi valittiin 20 MW ylös- tai alassäätöä, jolloin nollatunniksi laskettujen säätöjen alue pienenee merkittävästi. Taustalla numeerinen ennuste ei muutu, sillä luokitteluennuste muodostetaan jälkeinpäin numeerisen ennusteen arvosta. Muutoksen vaikutusta testattiin pitkällä testijaksolla malleille, joista kuvassa 14 on esitetty yksinkertaisimman mallin ARIMA (1,0,4) ja monimutkaisimman mallin SARIMAX (1,0,8) tulokset.

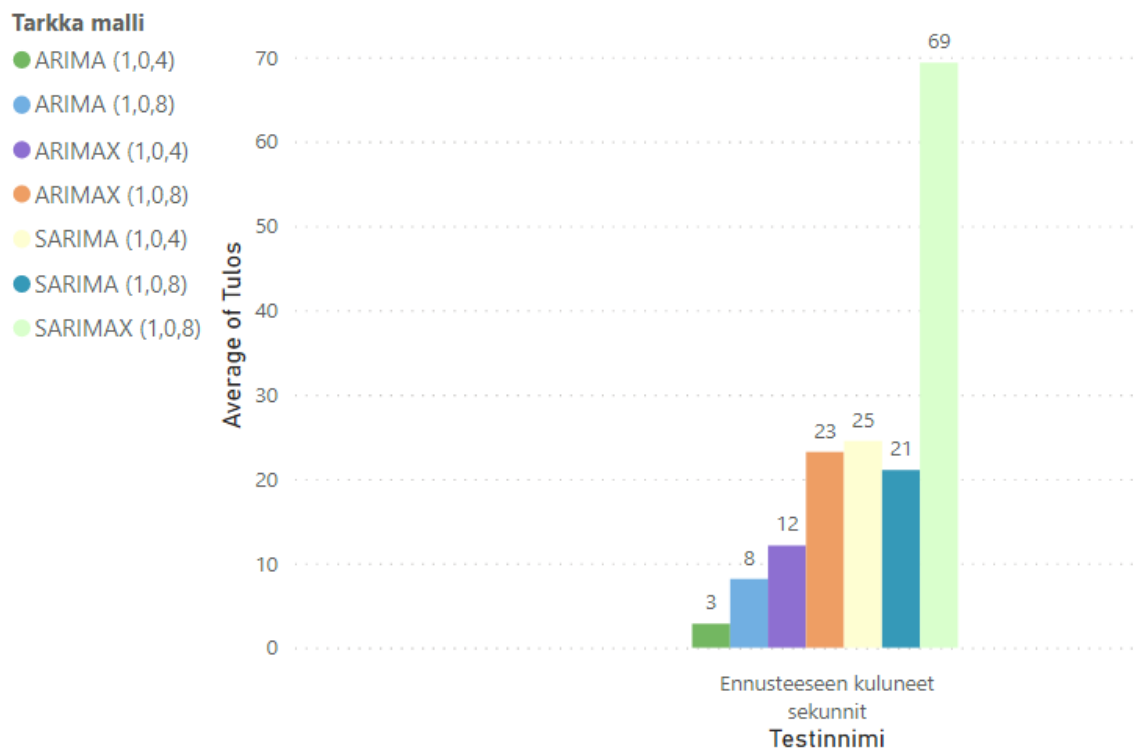


KUVA 14. Luokitteluvirheiden tarkastelu normaalin 40 MW sekä tarkemman 20 MW merkittävän ylös- ja alassäätörajan avulla ARIMA (1,0,4) ja SARIMAX (1,0,8) malleille

Merkittävän säätötunnin rajan vaikutus luokitteluennusteisiin tehtiin yksinkertaisin ja monimutkaisimmalle mallilla siitä syystä, että kaikki mallit voitiin todeta aikaisempien

testien perusteella tasavertaisiksi. Yleistä luokittelutarkkuutta tiukempi luokitteluraja laski molemmilla malleilla 80 prosentista 70 prosentin luokittelutarkkuuteen. Nollatuntien luokittelua tarkemman säätötunnin rajan käyttäminen laski paremmin suoriutuneella ARIMA (1,0,4) mallilla 94 prosentin tarkkuudesta 83 prosenttiin. Sen sijaan ylössäättöjen tarkkuus parani molemmilla malleilla 2 prosenttia ja alassäättöjen tarkkuus 7 prosentilla.

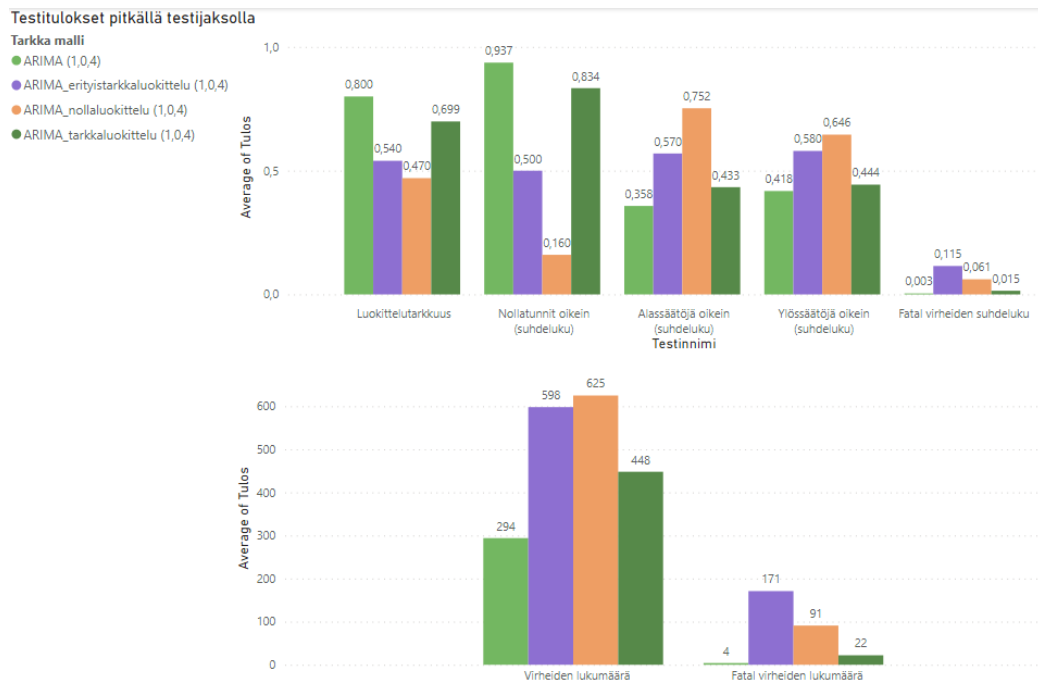
Tutkimuksen testitulokset ovat tähän asti esittäneet, ettei malleilla ole numeeristen tai luokitteluennusteiden perusteella merkittävää eroa. Ennustemallien käytettävyys tulee ottaa huomioon erityisesti, kun ennustetta on tarkoitus suorittaa kerran tunnissa. Kuvassa 15 on esitetty ennustemallien käyttämä aika yhteen ennusteeseen, sillä laskentaresurssilla mitä tutkimuksessa oli käytettävissä.



KUVA 15. Tutkimuksessa käytettävien mallien käyttämä aika ennusteen luomiseen.

Kuvan 15 ennusteisiin kuluneiden sekuntien keskiarvokuvaajasta voidaan havaita, kuinka paljon malliin lisätyt ominaisuudet kasvattavat tarvittavaa laskentatehoa. Laskentatehon pysyessä testissä vakiona kasvaa ennustamiseen kulunut aika. Vaihteluväli on 66 sekuntia raskaimman SARIMAX (1,0,8) - mallin ja kevyimmän ARIMA (1,0,4) mallin välillä, jolloin laskentatehoa tarvitaan ajasta arvioituna 23 kertaa enemmän. Taulukon 5 perusteella sopivin malli olisi AIC-kriteerin perusteella ollut raskain malli SARIMAX (1,0,8), mutta edeltävien tulosten perusteella soveltuvin malli voisi olla käytännöllisyyden ja muiden suorituskykykymittarien näkökulmasta kevyin malli ARIMA (1,0,4).

Kuvan 14 tulosten perusteella voitiin arvioida luokittelurajan tarkentamisen parantavan ylös- tai alassäätötunnin luokittelun tarkkuutta samalla laskien nollatuntien luokittelutarkkuutta, sillä nollatunteihin laskettujen säätöjen alue pienenee. Luokittelurajan vaikutusta tutkittiin vielä lisäämällä erityisentarkka 5 MW luokitteluraja sekä nollaluokitteluraja. Mallien suoriutumisessa eri tutkimuksen testeissä ei huomattu merkittävää eroavaisuutta, joten jatkotutkimuksiin valittiin yksinkertaisin ARIMA (1,0,4) malli. Kuvassa 16 on esitetty ARIMA (1,0,4) -mallille luokittelurajan vaikutusta luokittelutestien tuloksiin. Luokittelurajan tarkentaminen lisää riskiä täysin väärin luokitelluille ennusteille, sillä nollatunneiksi laskettujen tuntien alue pienenee. Täysin väärin luokitellut tunnit kuvataan tässä työssä fatal virheinä. Silloin ylössäätötunti luokitellaankin virheellisesti alassäätötunniksi ja toisinpäin. Kuvassa 16 on esitetty luokittelutestien lisäksi myös virheiden lukumäärä sekä aikaisemmin mainittujen fatal virheiden lukumäärä.



KUVA 16. Luokittelurajan tiukentamisen vaikutus luokitteluennusteisiin ARIMA (1,0,4) ennustemallilla

Merkittävien ylös- ja alassäätötuntien luokittelujen suhdeluvut parantuivat, mitä tarkempaa luokittelurajaa käytettiin. Kuitenkin pienten säätötuntien ja nollatuntien suuri lukumäärä aiheuttaa yleisen luokittelutarkkuuden, sekä nollatuntien ennustettavuuden suhdeluvun laskemisen luokittelurajaa tarkentaessa.

Virheiden kappalemäärä on hyvä mittari konkretisoimaan, monestako virheestä on kyse ennusteen käyttämässä kontekstissa eli tässä työssä kyse on väärin luokitelluista tunneista. Luokittelurajan tarkkuuden valinta on tasapainoilua yleisten sekä fatal

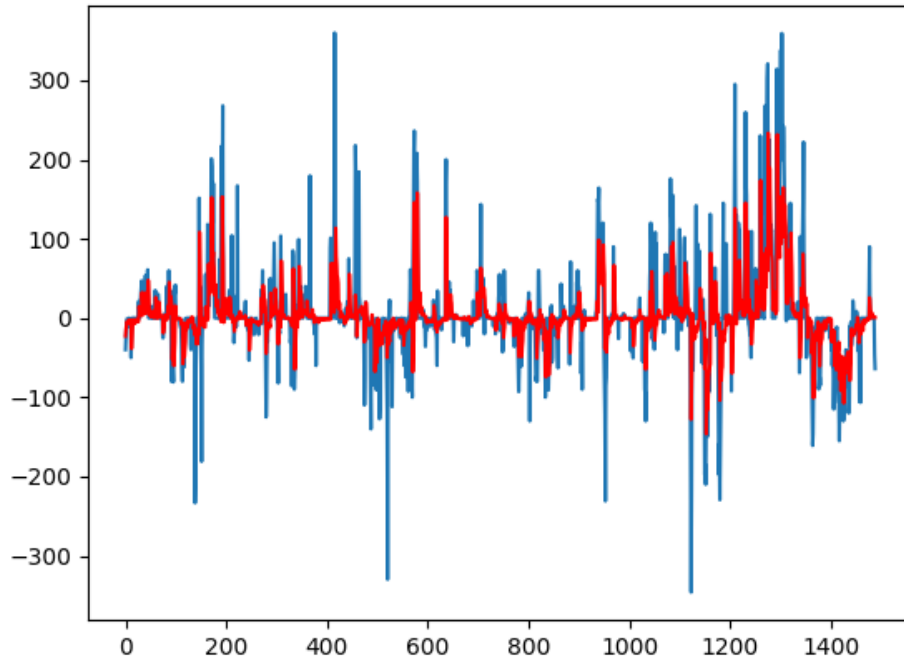
virheiden lukumäärän kasvun sekä parantuneen ylös- tai alassäätöennusteen luokittelun kanssa. Kuvan 16 ylemmästä kaaviosta voidaan arvioida fatal virheiden suhdeluvun avulla, kuinka suuren suhteellisen virheen rajan tiukentaminen aiheuttaa. Yleinen havainto on, että luokittelurajaa tiukentaessa myös fatal virheiden lukumäärä ja suhdeluku kasvaa. Mielenkiintoinen havainto on, että nollaluokittelulla fatal virheiden lukumäärä on pienempi erityisen tarkkaan 5 MW rajaan verrattuna. Käytettäessä tiukinta nollaluokittelurajaa kuusi prosenttia säädöistä ennustetaan päinvastaisiksi säädöiksi. Erityisen tarkan 5 MW luokittelurajalla säätöjen päinvastaisia ennusteita ilmenee 11 prosenttia ennusteista ja aikaisemmin määritellyllä tarkalla 20 MW rajalla virhe laskee 1,5 prosenttiin.

Tiukempi raja mahdollistaa sen, että jos jostain säätötunnista tuleekin ennustettua suurempi ja hinta nouseekin yllättäen niin pienestä säädöstä on mahdollista saada hyvä korvaus säätösähkömarkkinan tarjousten hyväksyntä rakenteen vuoksi. Säätötyypin luokittelurajan valinta riippuu käyttäjän käyttötarkoituksesta ja markkinoille varattavista energiareсурseista. Väärin luokiteltu säädön suunta aiheuttaa sen, että käyttäjä on joutunut varamaan resurssia ennustetun säätötyypin mukaan, sillä säätömarkkinoilla hyväksyty tarjous tulee olla tarvittaessa käytettävissä. Tässä työssä arvioitiin sopivaksi rajaksi 20 MW, vaikka yleinen luokitteluprosentti jäi 70 prosenttiin, niin ylös- ja alassäätöjen luokittelu oli keskimäärin 44 prosenttia nollatuntien oikeinluokittelun ollessa yhä korkea 83 prosenttia. Tarkalla 20 MW luokittelurajalla vältettiin yhä tehokkaasti päinvastaiset ennustevirheet, joita ilmeni vain 1,5 prosenttia ennusteista. Jos käyttäjällä on säätömarkkinoille varattavissa resursseja pienille säätömäärille, niin silloin nollaluokitteluraja voi olla hyvä valinta. Tällöin useasti toistuvilla nollatunneilla todennäköisesti varattaisiin resursseja useammin kuin 20 MW rajalla ilman että niitä aktivoidaan, mutta samaan aikaan toteutuneet ylössäätötunnit voitaisiin arvioida 75 prosentin ja alassäätötunnit 65 prosentin tarkkuudella.

### 3.9 Ennusteen arviointi visualisoinnin avulla

Ennusteita on hyvä tarkastella numeeristen ennustetestien lisäksi, myös visuaalisesti, jotta voitaisiin tunnistaa millaisia yksittäisiä tai toistuvia virheitä ennustemalli todellisuudessa tekee ja mistä virhetestien lukuarvot koostuvat. Kuvassa 16 on esitetty soveltuvimmaksi osoitetun ARIMA (1,0,4) mallin ennusteita pitkällä testisarjalla. Johtuen mallien tasavertaisista onnistumisista niin numeeristen ja luokitteluennusteiden osalta jätetään muiden mallien kuvaajat työstä pois. Kuvassa 17 on sinisellä esitetty säätösähkömäärän toteuma ja punaisella mallin ennustama sarja. Ennustemalli on

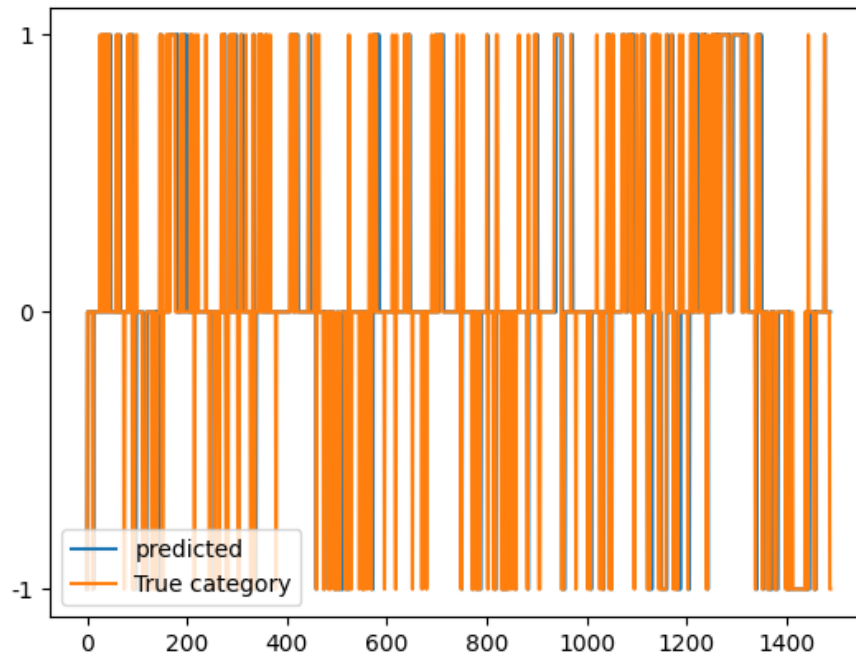
ennustanut kolme tuntia kerrallaan ja aina ennen uutta ennustetta hakenut viimeisimmät aikasarjan toteumat ja harjoittanut mallin uudelleen ennen seuraavaa ennustetta.



KUVA 17. ARIMA (1,0,4) mallin aikasarja ennuste (punainen) 1488 tunnin toteumasarjalle (sininen)

Kuvan 17 ennusteaikasarjasta huomataan ennustemallin suoriutuvan pitkällä aikavälillä ennustuksesta varsin hyvin. Sinistä toteumaa ja punaista ennustesarjaa vertailemalla huomataan, että silloin tällöin on merkittävän suuria, yli 250 MW säätöjä, joista ennuste jää merkittävästi vajaaksi. Tärkeintä ennusteiden käytön kannalta on kuitenkin, että näille tunneille ennuste ylittää merkittävän säädön rajan, jotta korkeaan säätömäärään päästään mukaan, sillä silloin saattaa säätötunnille määräytyvä hintakin nousta. On kuitenkin tärkeää, että korkeille säätötunneille ennustetaan keskimääräistä korkeampaa säätöä, jotta ennusteen käyttäjä voi tehdä isoihin säätöihin varautuvaa toimenpidettä. Ennusteen kannalta hieman todellisuutta maltillisempi ennustaminen on hyvä, kuitenkin niin että myös suuret säätömäärät voidaan ennustaa riittävän usein.

Kuvassa 18 on esitetty kuvaajaa luokittelun onnistumisesta. Kuvassa 17 on sinisellä esitetty ennustettu luokka ja oranssilla todellinen luokka ja kyseisessä kuvassa luokittelu on tehty tiukemman 20 MW luokittelurajan avulla. Kuvassa 18 luvulla 1 tarkoitetaan ylös säätötuntia, luvulla 0 nollatuntia, jolloin säätömäärä on 20 MW luokittelurajojen sisällä, ja luvulla -1 alas säätötuntia.



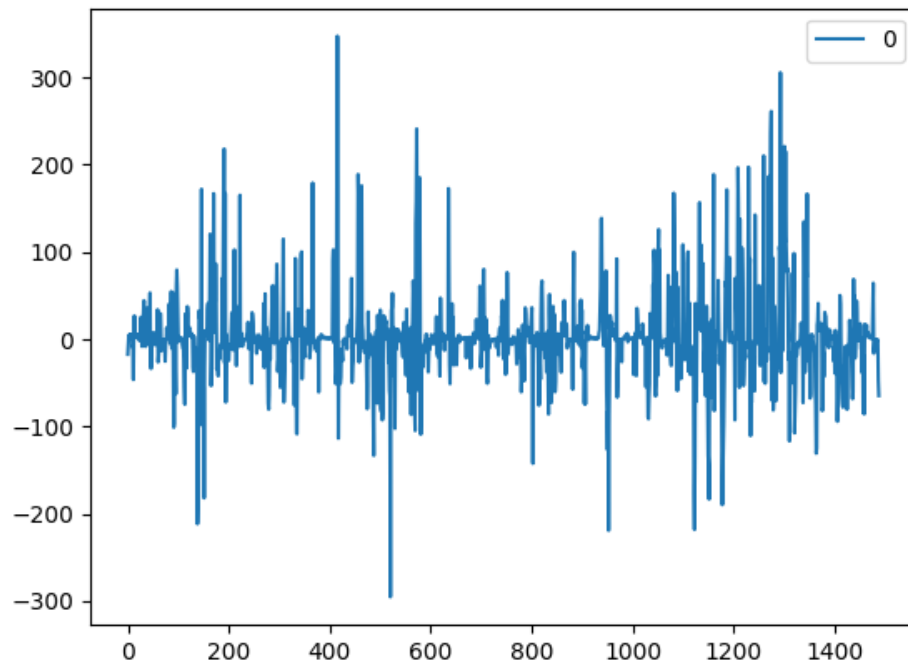
KUVA 18. ARIMA (1,0,4) mallin luokittelun onnistuminen pitkälle testisarjalle 20 MW ylös- ja alassäätömäärän luokittelurajalla

Kuvan 18 yhtenäisestä väristä voidaan havaita, että ennusteen luokittelu seuraa hyvin todellisia luokkia, jolloin ennustaminen pääosin on hyvällä tasolla, kuten kuvan 16 testituloksisakin todettiin. Kuvasta 18 voidaan joitain sinisiä ennusteluokkia havaita, jotka ilmenevät testituloksissa väriä luokkien ennusteina. Kuvan 17 tuloksista tarkasteltiin myös tarkemmin pienemmille aikaväleille, jolloin havaittiin, että säätötunteja hyvin harvoin ennustettiin päinvastaisiksi luokiksi, kuten kuvassa 16 esitetty fatal virheiden lukumääräkin osoitti.

### 3.10 Residuaalin tarkastelu

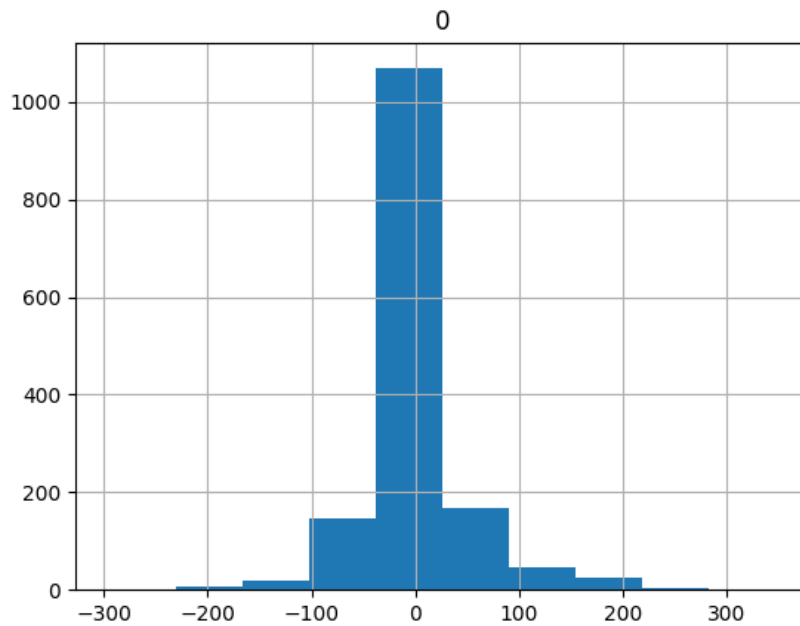
Tarkkaa kuvaa numeerisen ennusteen virheen eli residuaalin toistuvuudesta ja suuruudesta saattaa olla haastava arvioida ainoastaan kuvien 16 ja 17 avulla. Tätä varten numeerisen ennusteen residuaalista on hyvä muodostaa pelkkää residuaalia käsitteleviä kuvaajia. Kuvassa 19 on esitetty mallin ARIMA (1,0,4) residuaali pitkän testisarjan tuloksille.





KUVA 19. ARIMA (1,0,4) mallin residuaali pitkän tutkimusjakson testisarjalle

Kuvassa 19 havaitaan aikasarjaennustamiselle hyviä piirteitä, jolloin virhettä tehdään kohtalaisen tasaisesti molempiin suuntiin, ja virheiden suuruus ei selvästi yleisty jollain ajanjaksolla. Parhaimmillaan residuaali olisi valkoisen kohinan muotoista eli täysin satunnaista, jolloin malli kykenisi ennusteessa mallintamaan kaiken mahdollisen ja jäljelle jäävä virhe olisi ainoastaan valkoista kohinaa. Kuvassa 20 on havainnollistettu residuaalin jakaumaa histogrammin avulla, jossa vaaka-akselilla on eri suuruisia virheitä ja pystyakselilla on virheiden ilmaantuminen testisarjassa kappalemäärinä.

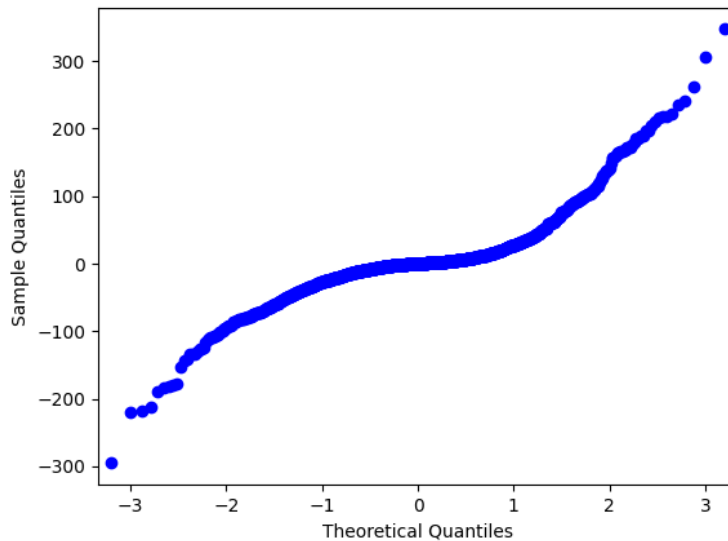


KUVA 20. ARIMA (1,0,4) residuaalien histogrammi

Kuvan 20 histogrammista havaittiin ennustamisen kannalta hyvä asia, että useimmiten toistuivat hyvin pienet virheet, minkä vuoksi ilmaantuminen on suurta nollan ympäristössä. Keskimmäisen palkin keskikohta on hieman negatiivisen puolella, mikä on merkki siitä, että usein ennustetaan hieman vähemmän kuin toteuma on. Sama havainto jälkepäin, voidaan havaita myös ennusteen ja toteuman vertailukuvasta (kuva 16), jolloin punaisella merkitty ennuste on toistuvasti toteumaa pienempi.

Histogrammista huomataan myös, että virheiden suurentuessa niiden ilmaantuvuus pienenee ja keskimääräisiä alle 100 MW virheitä ilmaantuu testisarjassa tasaisesti. Residuaali ei vaikuta merkittävästi vinoutuneelta vaan virheet jakautuvat molemmille puolille saman verran. Kuvassa 21 on esitetty residuaalin kvantaalikuvaaja, minkä avulla

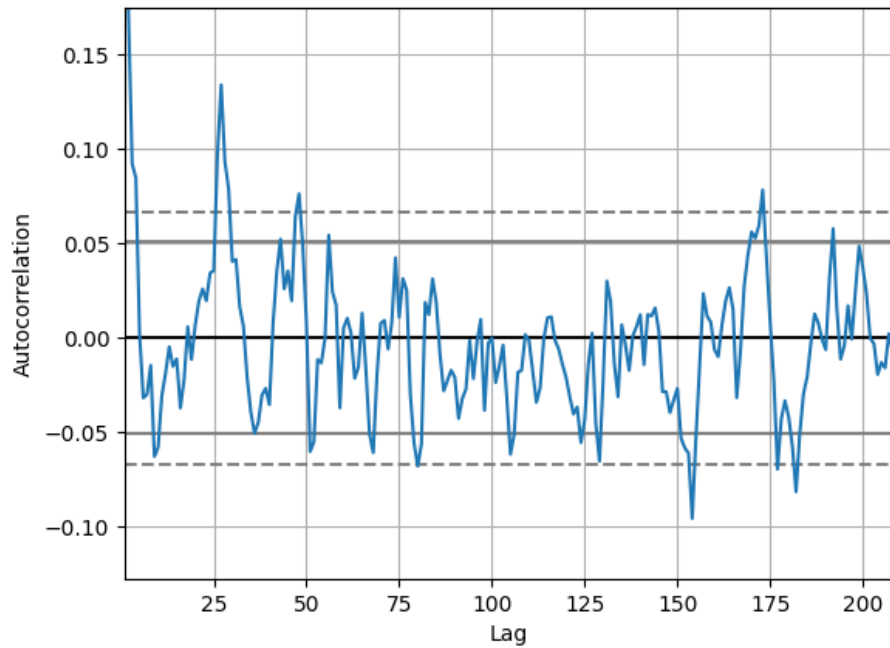
residuaalin jakaumaa voitiin verrata tarkemmin oletettuun normaalijakaumaan.



KUVA 21. ARIMA (1,0,4) mallin pitkän testisarjan residuaalin kvantaalikuvaaja

Kvantaalikuvaajassa vaaka-akselilla on normaalijakauman lukuarvot ja pystyakselilla testattavan residuaalin jakauman arvot. Jos residuaali olisi täysin normaalijakauman mukainen, niin muodostuisi kvantaalikuvaajaan pisteet 45 asteen kulmaan vasemmasta alakulmasta oikeaan yläkulmaan. Ennustemallin residuaalin kvantaalikuvaajasta voidaan huomata, että pisteistä muodostuva suora loiventuu y-akselin -30 ja 30 välillä, mikä osoittaa, että pieniä virheitä tehdään keskiuuria tai suuria useammin. Tuon alueen ulkopuolella kvantaalikuvaajan pisteistä muodostuvan suoran kulmakerroin vaikuttaa kuitenkin hyvin samalta 30 lukua suuremmissa ja -30 lukua pienemmissä y-akselin arvoissa. Kvantaalikuvaajasta huomataan myös, että residuaalien pistesarja ei ole täysin yhtenäinen alusta loppuun vaan alle -300 ja yli 300 residuaalit voidaan todeta normaalista residuaalista poikkeaviksi satunnaisiksi tapahtumiksi.

Residuaalin kannalta merkittävä on myös residuaalin autokorrelaatio kuvaaja, sillä se kertoo, millä varmuudella virhe on riippuvainen edeltävistä ajanhetkistä, mikä taas osoittaa virheen kausittaisuutta. Myös ennustemallien residuaalien autokorrelaatiokuvaajat olivat hyvin samanlaisia keskenään, joten kuvassa 22 on esitetty yksinkertaisuuden ja nopeuden vuoksi parhaimmaksi luokitellun ARIMA (1,0,4) mallin pitkän testisarjan residuaalin autokorrelaatiokuvaaja.



KUVA 22. ARIMA (1,0,4) -mallin pitkän testisarjan ennusteiden residuaalin autokorrelaatiokuvaaja.

Kuvassa 22 on merkitty aikasarjan autokorrelaatiolle määritellyn merkittävyyssrajan 0,06 ylittää 1, 24 ja 25 tunnin aikaviiveet, eikä niistä päästy muillakaan malleilla täysin eroon. Korkein autokorrelaation vaikutusluku on alle 0,15 mikä on hyvin matala, joten selvää kausittaisuutta ei residuaalissa todeta olevan.

## 4. YHTEENVETO JA PÄÄTELMÄT

Tilatun säätösähkömäärän ja sen mukaan aktivoituneen säätötyypin ennustamisesta suoriutui parhaiten tutkimukseen valikoiduista malleista yksinkertaisin ARIMA (1,0,4) -malli 20 MW säätötyypin luokittelurajalla. Yksinkertaisin malli suoriutui numeeristen ennusteiden testeistä monimutkaisempien mallien kanssa tasapuolisesti 27 MW keskivirheellä, mutta laskennallisesti muita tehokkaammin. Numeerisen ennusteen ja 20 MW luokittelurajan avulla voitiin ennustaa ylös- ja alassäädöt keskimäärin 44 prosentin tarkkuudella, luokittelurajojen asettamat nollatunnit 83 prosentin tarkkuudella. Tällöin yleinen luokittelutarkkuus eri luokkien välillä asettui 80 prosenttiin. Ennustemenetelmältä haluttiin suorituskyvyn lisäksi selitettävyyttä, jonka kautta säätösähkömäärän ennustamista voitaisiin ymmärtää paremmin. Harjoittaessa ARIMA (1,0,4) -malli ennen ennustamista esitetyle harjoitussarjalle voidaan muodossa kaavan 3 mukainen malli säätösähkömäärän ennusteelle, jonka tarkempi kuvaus on esitetty kaavassa 10.

$$y'_t = c + \sum_{i=1}^p \phi_i y'_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t \quad (3)$$

$$y'_t = -2.20 + 0.92 \cdot y'_{t-1} - 0.17 \cdot \varepsilon_{t-1} - 0.20 \cdot \varepsilon_{t-2} - 0.13 \cdot \varepsilon_{t-3} - 0.07 \cdot \varepsilon_{t-4} + 2183 \quad (10)$$

Kaavassa 10  $y'_t$  on seuraava ennustettu arvo, autoregressio termejä huomioidaan sarjasta vain yksi  $y'_{t-1}$  ja edeltävien ajanhetkien mallinnuksen virheitä neljän liukuvan keskiarvotermien avulla  $\varepsilon_{t-1}$ ,  $\varepsilon_{t-2}$ ,  $\varepsilon_{t-3}$  ja  $\varepsilon_{t-4}$ . Mallinnuksen arvioidusta 2183 arvioimasta virheestä jäljelle jää vielä todellista virhettä, mitä malli ei ennustuksessa huomioi. Kaavan 10 perusteella säätösähkömäärän ennustamiseen tarvitaan tieto yhdestä ennustetta edeltävästä ajanhetkestä, sekä tieto neljästä edeltävästä virheestä, joita harjoitusvaiheen mallinnuksessa tehtiin, joiden mukaan ennustetta voidaan korjata. Tutkimuksen mukaan soveltuvimmasta mallista voidaan havaita, ettei säätösähkömäärän ennustamista parantanut tieto vuorokauden ajasta. Ennustetta ei myöskään parantanut tieto kulutustasesähkön hinnasta.

Säätösähkömarkkinatiedot ovat koneluetavassa muodossa saatavilla Fingridin avoimesta API-rajapinnasta, joten työ loi edellytykset siirtyä testiympäristöstä reaaliaikaiseen ennustamiseen, kun arvio suorituskyvystä selvitettiin työssä. Testit suoritettiin kolmen tunnin ennusteina, jonka avulla on otettu huomioon rajapinnassa ilmenevä tunnin päivittymisviive sekä vajavainen tieto kuluvan tunnin tilatun säätösähkön määrästä.

Energiahallinnan päätöksen tekoa voidaan kehittää tutkimuksessa luoduilla säätösähkömäärän ennusteilla. Sähköenergiamittauksen sijaan tietoa tilattujen ylös- ja alassäätösähkön määrästä voidaan hakea Fingridin avoimen datalähteen API-rajapinnasta avoimen lähdekoodin python ohjelmalla. Sen jälkeen kerätyn tiedon eheys voidaan tarkistaa ja tarpeen mukaan esikäsitellä ja välittää esikäsitelty tieto ennustusmallille. Tämän jälkeen ennustusmallista saatavaa ennustetta seuraavan tunnin säätömäärästä ja suunnasta voidaan hyödyntää energiahallinnan päätöksen teossa monipuolisesti riippuen siitä, millaisia tuotanto- tai kulutusmuotoja ennusteen käyttäjä omistaa. Mahdollisesti ennusteen tekijä voi olla myös kantaverkkoyhtiö, joka voi ennusteiden avulla optimoida toimintaansa esimerkiksi aktivoimalla ulkomaankauppoja tai käytettävissä olevia resursseja ennusteen mukaisesti.

Tutkimuksessa käytännöllisimmäksi todetun ARIMA (1,0,4) -mallin ennustaminen vei ainoastaan sekunteja, joten lähtökohtaisesti ei ole tarvetta rajoittaa ennustemallille välitetyn tuntisarjan pituutta tietomäärän jatkuvasti kasvaessa. Ennusteen yhteydessä mallista on aina saatavilla myös AIC-kriteeri arvon laskenta, joka kertoo, kuinka hyvin mallinnus käytetyillä parametreilla ja datalla onnistuu, joten se voi myös avustaa riittävän datamäärän valinnassa.

## 4.1 Jatkotutkimustarpeet

Kuvassa 5 kuvattiin aikasarjaennustamisen tutkimussuunnitelmaa kehänä, mitä kierretään, kunnes toivottuihin tavoitteisiin on päästy. Tässä tutkimuksessa saavutetussa numeerisessa ennustamisessa päästiin riittävän korkealle tasolle keskivirheen ollessa 27 MW, jonka käytettävyyttä voidaan arvioida ennusteen ja toteuman visuaalisesta kuvaajasta (kuva 16). Soveltamalla kuvaa 5, voitaisiin ennustemenetelmää vaihtaa tarkempien vaatimusten jälkeen epälineaariseen malliin tai lineaaristen- ja epälineaaristen mallien yhdistelmiin ja pyrkiä yhä parempiin tuloksiin. Tämän tutkimuksen numeerisessa ennusteessa pääosin onnistuttiin, mutta suurien säätömäärien ennustamista voitaisiin parantaa.

Kolehmainen (2016, 25) mukaan säätösähköä on haastava ennustaa. Kuitenkin tekijät, jotka edellisten säätömäärien lisäksi saattavat vaikuttaa säätösähkön ennustamiseen olisivat verkon taajuus, häiriöt ja poikkeustilanteet suurissa voimalaitoksissa ja rajasiirtoyhteyksissä, lämpötila, lämpötilan muutosten sekä lämpötilaennusteiden virheet, tuulivoiman tuotanto ja tuulivoimatuotannossa tehdyt ennustevirheet, vesivoimatuotannon käytettävissä oleva kapasiteetti ja säätötarjosten summa. Monimuuttujamenetelmien jatkotutkimuksissatulisi huomioida, että valittujen muuttujien merkittävyys voi muuttua ajan myötä tai poikkeustilanteista johtuen. Lomion ja muiden

(2021, 5) mukaan monimuuttujamenetelmien valintaa tehostaisi, jos ennusteen vaikuttavien muuttujien valinta automatisoitaisiin. Liu ja Lin (2021, 9) esittivät, että sähkön kulutuksessa on COVID-19 pandemian myötä julkisten rakennusten kulutusprofiili muuttui merkittävästi ja näin vaikutti sähkön kulutuksen ennustamisessa käytettyihin muuttujiin. Muuttujien oikeaa valintaa voitaisiin helpottaa automatisoimalla parametrien asetusten lisäksi myös käytettävät muuttujat harjoitusdataan perustuen.

Jatkotutkimusta voitaisiin kohdentaa luokitteluennusteen parantamiseen pienten säätötuntien osalta. Luokitteluennustaminen tehtiin tässä tutkimuksessa numeerisen ennusteen jatko tuotteena, mutta tulevaisuudessa epälineaaraisia menetelmiä voitaisiin käyttää luokitteluennusteiden muodostamiseen erillään numeerisesta ennusteesta ja näin nostaa luokitteluennusteiden suhdelukua myös pienten säätötuntien osalta.

Aaltonen, Sierla, Subramanye ja Vyatkin (2021, 1) ovat esittäneet kuinka ohjatun oppimisen menetelmiä voitaisiin soveltaa akustojen ohjauksissa optimoiden osallistumista huomioiden akustojen varaukset sekä reservimarkkinan tilanteen. Myös ohjatun oppimisen menetelmien yhdistäminen tämän työn mukaisten säätösähkömäärän ja -tyypin ennusteisiin voisi edistää yhä edelleen ennusteiden muodostamista myös säätösähkömarkkinoilla Aaltosen ym. (2021) tutkimien reservimarkkinoiden lisäksi.

## 4.2 Johtopäätökset

Tutkimuksen lopputuloksena saatiin muodostettua yksinkertainen, selitettävä ja suorituskyvyltään tehokas lineaarinen ennustemalli. Mallin muodostamien säätösähkömäärien avulla pystyttiin muodostamaan myös käyttökelpoista 27 MW keskiarvovirheen numeerista ennustetta sekä luokittelua 70 prosentin tarkkuudella, kun merkittävän säätösähkömäärän luokittelurajaksi asetettiin 20 MW ylös- tai alassäätöä. Loivemmalla 40 MW luokittelurajalla yleiseksi luokittelutarkkuudeksi saatiin 80 prosenttia. Testitulosten avulla pystyttiin arvioimaan, että suurin osa luokitteluvirheistä asettuu pienille säätömäärille. Testitulosten perusteella ennustejärjestelmä tekee luokittelua hyvin säätömäärän korkeille tunneille, joten käyttäjä pääsee ennusteiden avulla näille tunneille osallistumaan. Työn avulla ennusteiden käyttäjä voi arvioida käyttötarpeistaan riippuen sopivan luokittelurajan ja tiedostaa tarkemman luokittelurajan sisältämät mahdollisuudet ja riskit.

Epälineaaristen mallien kehittyessä aikasarjojen, kuten säätösähkömarkkinoiden ennustamisessa voidaan hyvin suoriutuneiden lineaaristen ennustemallien suorituskykyä pitää vertailutasona tulevaisuuden aikasarjaennustamisen monimutkaisemmille koneoppimismalleille, kuten myös Giovanelli, Sierla, Ichise &

Vyatkin (2018, 2) totesivat tutkiessaan reservimarkkinoiden ennustamista. Lisäksi Tang ja Rölliin (2021, 2) ovat esittäneet mahdollisuuksia yhdistää lineaarisia malleja, kuten ARIMA ja epälineaarisia konvoluutioneuroverkkoja, mitä voitaisiin myös säätösähkömäärän ennustamisessa soveltaa. Moon, Hossainin ja Chon (2021, 2) esittävät myös, kuinka koneoppimiseen kuuluvien geneettisten algoritmien avulla voitaisiin valita ARMA-mallille parhaimmat parametrit perustuen ennen ennustetta tehtävään harjoitusdataan, jolloin ARMA-mallista voitaisiin johtaa GA-ARMA hybridiennustemalli. Tämä mahdollistaisi joustavuuden parametrien valinnassa, jos ennustettavan suureen riippuvuudet edeltävistä ajanhetkistä muuttuu merkittävästi eikä testidatan perusteella tehdyt parametrivalinnat pitäisikään enää paikkansa.

Tutkimuksessa käytiin läpi kattavasti aikasarjaennustamiseen liittyviä ominaispiirteitä, joita voidaan soveltaa jatkossa valitusta ennustemenetelmästä huolimatta. Lisäksi tutkimus antoi kuvaa toteutuneen säätösähkömäärän ennustamiseen liittyvistä yksityiskohdista, jotka toivottavasti vievät eteenpäin kohti tasapainoisemman ja parempilaatuisemman sähköverkon tulevaisuutta.



# LÄHTEET

Aaltonen H., Sierla S., Subramanya R. & Vyatkin V. 2021. A Simulation Environment for Training a Reinforcement Learning Agent Trading a Battery Storage. *Energies*. Saatavissa: <https://doi.org/10.3990/en14175587>

Clements A. E., Hurn S. A. & Li Z 2015. Forecasting day-ahead electricity load using multiple equation time series approach. *European Journal of Operational Research*. Elsevier. Saatavissa: <https://elsevier.com/locate/ejor>

Cuaresma J. C., Hlouskova J., Kossmeier S. & Oberstainer M. 2004. Forecasting electricity spot-prices using linear univariate time-series models. *Applied energy*. Elsevier. Saatavissa: <https://www.elsevier.com/locate/apenergy>

Dissanayake B., Hemachandra O., Laksitha N., Haputhantri D. & Wijayasiri A. **2021**. A Comparison of ARIMAX, VAR and LSTM on Multivariate Short-Term Traffic Volume Forecasting. *Proceeding of the 28<sup>th</sup> Conference of Fruct Association*. University of Moratuva.

Elamin N. & Fukushige M. 2018. Modelin and forecasting hourly electricity demand by SARIMAX with interactions. *Energy*. Elsevier. Saatavissa: <https://www.elsevier.com/locate/energy>

Fingrid 2021a. Esittely. Saatavissa: <https://fingrid.fi/sivut/yhtio/esittely/>

Fingrid 2021b. Johdanto sähkömarkkinoihin. Saatavissa: <https://www.fingrid.fi/sahkomarkkinat/markkinoiden-yhtenaisyyys/johdanto-sahkomarkkinoihin/>

Fingrid 2021c. Kysyntäjousto. Saatavissa: <https://www.fingrid.fi/sahkomarkkinat/markkinoiden-yhtenaisyyys/pilottihankkeita/kysyntajousto/>

Fingrid 2021d. Säätosähkö – ja säätökapasiteettimarkkinat. Saatavissa <https://www.fingrid.fi/sahkomarkkinat/reservist-ja-saatokapasiteettimarkkinat/#saatosahkon-hinnoittelu>

Fingrid 2021e. Ehdot ja edellytykset manuaalisen taajuuden palautusreservin (mFRR) toimittajalle.

Fingrid 2021f. Avoin data. Saatavissa: <https://data.fingrid.fi/>

Fingrid 2021g. Säätosähkön määrä ja hinta. Valittu janjakso 1.12.2020-31.12.2020. Saatavissa: <https://www.fingrid.fi/sahkomarkkinat/sahkomarkkinainformaatio/reservimarkkinainformaatio/saatosahkon-maara-ja-hinta/#saatosahkon-maara>

Gellert A., Fiore U., Florea A., Chis R. & Palmeri F. 2021. Sustainable Cities and Society. Forecasting Electricity Consumption and Production in Smart Homes through Statistical Methods. Saatavissa: <https://doi.org/10.1016/j.scs.2021.103426>

- Giovanelli C., Sierla S., Ichise R. & Vyatkin V. 2018. Exploiting Artificial Neural Networks for the Prediction of Ancillary Energy Market Prices. *Energies*. Saatavissa: <https://www.mdpi.com/journal/energies>
- Heim S. & Götz G. 2021. Energy Policy. Do Pay-As-Bid Auctions Favor Collusion? Evidence from Germany's market for reserve power. Elsevier. Saatavissa: [www.elsevier.com/locate/enpol](http://www.elsevier.com/locate/enpol)
- Kemptitya T., Sierla S., De Silva D., Yli-Ojanperä M., Alahakoon D. & Vyatkin V. 2020. Applied energy. An artificial intelligence framework for bidding optimization with uncertainty in multiple frequency reserve markets. Elsevier. Saatavissa: [www.elsevier.com/locate/apenergy](http://www.elsevier.com/locate/apenergy)
- Kolehmainen M. 2016 Palvelutuottajan Elbas-kaupankäynnin kehittäminen tasekustannusten minimoimiseksi perustuen säätösähkön hinnan odotusarvoon. Diplomityö Tampereen teknillinen yliopisto.
- Kuukkanen T. Konttinen M. & Meritähti. 2020. Olkiluodon voimalassa harvinainen häiriötilanne – TVO:lla aavistus ongelman syystä, työntekijät eivät olleet vaarassa, katso tallenne tiedotustilaisuudesta. Yle. Saatavissa: <https://yle.fi/uutiset/3-11690281>
- Liu X. & Lin Z. 2021. Impact of Covid-19 pandemic on electricity demand in the UK basen on multivariate time series forecasting with Bidirectional Long Short Term Memory. *Energy*. Elsevier. Saatavissa: <https://www.elsevier.com/locate/energy>
- Malik O. & Havel p. 2012. Electrical Power and Energy Systems. Decision support tool for optimal dispatch of tertiary control reserves. Elsevier. Saatavissa: [www.elsevier.com/locate/iejeeps](http://www.elsevier.com/locate/iejeeps)
- Merten M., Rücker F., Schoenberger I. & Uwe Sauer D. 2020. Automatic frequency restoration reserve market prediction: Methodology and comparison of various approaches. *Applied Energy*. Elsevier. Saatavissa: <https://www.elsevier.com/locate/apenergy>
- Moon J., Hossain M. B. & Chon K.H. 2021. AR and ARMA model order selection for time-series modeling with ImageNet classification. *Signal Processing*. Elsevier. Saatavissa: <https://www.elsevier.com/locate/sigpro>
- Nitsch F., Deissenroth-Uhrig M., Schimeczek C. & Bertsch V. 2021. *Applied Energy*. Economic evaluation of battery storage systems bidding on day-ahead and automatic frequency restoration reserves markets. Elsevier. Saatavissa: [www.elsevier.com/locate/apenergy](http://www.elsevier.com/locate/apenergy)
- Nousi P., Tsantekidis A., Passalis N., Ntakaris A., Kannianen J., Tefas A., Gabbouj M. & Iosifidis A. 2017. Machine Learning for Forecasting Mid-Price Movements using Limit Order Book Data. *IEEE Access*.
- P.M.S Frade, J.J.E Santana, M. Shafie-khah, J.P.S Catalão 2018. *Utilites Policy* 55 sivut 167-177. Impact of tertiary reserve sharing in Portugal. Elsevier. Saatavissa: [www.elsevier.com/locate/jup](http://www.elsevier.com/locate/jup)
- Pérez-Díaz J.I. Guisàndez I., Chazarra M., Helseth A. 2020. *Electric Power Systems Research*. Medium-term scheduling of a hydropower plant participating as a price-maker in the automatic frequency restoration reserve market. Elsevier. Saatavissa: [www.elsevier.com/locate/epsr](http://www.elsevier.com/locate/epsr)

Subramanya, R. Yli-Ojanperä, M. Sierla, S., Hölttä, T.; Valtakari, J. Vyatkin, V. A. 2021. Virtual Power Plant Solution for Aggregating Photovoltaic Systems and Other Distributed Energy Resources for Northern European Primary Frequency Reserves. *Energies* 2021, 14, 1242. Saatavissa: <https://doi.org/10.3390/en14051242>

Tang W. H. & Röllin A. 2021. Model identification for ARMA time series through convolutional neural networks. *Decision support Systems*. Elsevier. Saatavissa: <https://www.elsevier.com/locate/dss>

Tsantekidis A., Passalis N., Tefas A., Kannianen J., Gabbouj M. & Iosifidis A. 2017. Forecasting Stock Prices from the Limit Order Book using Convolutional Neural Networks. *IEEE 19<sup>th</sup> Conference on Business Informatics*.

Tuominen J., Lomio F., Oksala N., Palomäki A., Peltonen J., Huttunen H. & Roine A. 2021. Forecasting Daily Emergency Department Arrivals Using High-Dimensional Multivariate Data: A Feature Selection Approach. *Research Square*. Saatavissa: <https://doi.org/10.21203/rs.3.rs-907966/v1>

Zhang G. & Guo J. 2020. A novel ensemble method for hourly residential electricity consumption forecasting by imaging time series. *Energy*. Elsevier. Saatavissa: <https://www.elsevier.com/locate/energy>

## 5. LIITE 1. NUMEERISTEN ENNUSTETULOSTEN MITTAUSPÖYTÄKIRJA

Testinimi	SARIMA (1,0,8)-3h ennuste-pitkä	ARIMA (1,0,8)-3h ennuste-pitkä	SARIMA (1,0,4)-3h ennuste-pitkä	ARIMA (1,0,4)-3h ennuste-pitkä
RMSE kW	50,33	50,39	50,21	50,4
MAE kW	28,91	28,93	28,77	28,83
Luokittelutarkkuus	0,80	0,80	0,80	0,8
Ylössäätöjä oikein (suhdeluku)	0,42	0,42	0,42	0,41
Alassäätöjä oikein (suhdeluku)	0,36	0,36	0,35	0,36
Nollatunnit oikein (suhdeluku)	0,93	0,94	0,94	0,94
Virheiden lukumäärä	295	293	295	294
Ennusteeseen kuluneet sekunnit	21,9	7,9	24,5	3,7
Fatal virheiden lukumäärä				4
Testinimi	SARIMA (1,0,8)-3h ennuste-lyhyt	ARIMA (1,0,8)-3h ennuste-lyhyt	SARIMA (1,0,4)-3h ennuste-lyhyt	ARIMA (1,0,4)-3h ennuste-lyhyt
RMSE kW	51,86	51,95	52,15	52,51
MAE kW	33,18	33,2	33,5	33,3
Luokittelutarkkuus	0,79	0,78	0,79	0,80
Ylössäätöjä oikein (suhdeluku)	0,74	0,74	0,74	0,74
Alassäätöjä oikein (suhdeluku)	0,57	0,59	0,57	0,59
Nollatunnit oikein (suhdeluku)	0,88	0,90	0,89	0,89
Virheiden lukumäärä	43	44	42	41
Ennusteeseen kuluneet sekunnit	25,1	7,9	28	7
Testinimi	SARIMA (1,0,8)-3h ennuste-erikois	ARIMA (1,0,8)-3h ennuste-erikois	SARIMA (1,0,4)-3h ennuste-erikois	ARIMA (1,0,4)-3h ennuste-erikois
RMSE kW	59,93	60,27	59,74	60,1
MAE kW	35,05	34,87	34,86	34,65
Luokittelutarkkuus	0,73	0,74	0,73	0,74

Ylössäätöjä oikein (suhdeluku)	0,30	0,31	0,30	0,30
Alassäätöjä oikein (suhdeluku)	0,42	0,40	0,42	0,40
Nollatunnit oikein (suhdeluku)	0,89	0,90	0,89	0,90
Virheiden lukumäärä	219	213	219	216
Ennusteeseen kuluneet sekunnit	33,4	7,9	14,6	0,9
Testinnimi	SARIMAX (1,0,8)-3h ennuste-pitkä	ARIMAX (1,0,8)-3h ennuste-pitkä	SARIMAX (1,0,4)-3h ennuste-pitkä	ARIMAX (1,0,4)-3h ennuste-pitkä
RMSE kW	50,37	50,47	50,32	50,5
MAE kW	28,95	29,05	28,88	28,90
Luokittelutarkkuus	0,80	0,80	0,80	0,80
Ylössäätöjä oikein (suhdeluku)	0,43	0,42	0,42	0,41
Alassäätöjä oikein (suhdeluku)	0,364	0,370	0,339	0,364
Nollatunnit oikein (suhdeluku)	0,93	0,94	0,94	0,94
Virheiden lukumäärä	294	292	297	294
Ennusteeseen kuluneet sekunnit	65,8	23,8	58	12,3
Fatal virheiden lukumäärä	4			
Testinnimi	SARIMAX (1,0,8)-3h ennuste-lyhyt	ARIMAX (1,0,8)-3h ennuste-lyhyt	SARIMAX (1,0,4)-3h ennuste-lyhyt	ARIMAX (1,0,4)-3h ennuste-lyhyt
RMSE kW	51,9	51,92	52,48	52,49
MAE kW	33,22	33,18	33,718	33,28
Luokittelutarkkuus	0,79	0,78	0,80	0,80
Ylössäätöjä oikein (suhdeluku)	0,74	0,74	0,74	0,74
Alassäätöjä oikein (suhdeluku)	0,57	0,60	0,60	0,60
Nollatunnit oikein (suhdeluku)	0,88	0,87	0,90	0,90
Virheiden lukumäärä	43	44	40	41
Ennusteeseen kuluneet sekunnit	90,3	29,8	106,5	15,3
Testinnimi	SARIMAX (1,0,8)-3h ennuste-erikois	ARIMAX (1,0,8)-3h ennuste-erikois	SARIMAX (1,0,4)-3h ennuste-erikois	ARIMAX (1,0,4)-3h ennuste-erikois

RMSE kW	59,7	60,01	59,52	59,87
MAE kW	34,88	34,7	34,7	34,61
Luokittelutarkkuus	0,73	0,74	0,73	0,74
Ylössäätöjä oikein (suhdeluku)	0,29	0,30	0,29	0,30
Alassäätöjä oikein (suhdeluku)	0,42	0,40	0,41	0,40
Nollatunnit oikein (suhdeluku)	0,89	0,90	0,89	0,90
Virheiden lukumäärä	219	213	219	214
Ennusteeseen kuluneet sekunnit	105,9	22,6	56,3	9,9
Testinimi	ARIMA_tarkkaluokittelu (1,0,4)-3h ennuste-pitkä	SARIMAX_tarkkaluokittelu (1,0,8)-3h ennuste-pitkä		
RMSE kW	50,4	50,37		
MAE kW	28,83	28,94		
Luokittelutarkkuus	0,699	0,7		
Ylössäätöjä oikein (suhdeluku)	0,44	0,45		
Alassäätöjä oikein (suhdeluku)	0,43	0,43		
Nollatunnit oikein (suhdeluku)	0,83	0,83		
Virheiden lukumäärä	448	432		
Ennusteeseen kuluneet sekunnit	3,7	39,3		
Fatal virheiden lukumäärä	22	21		
Testinimi	ARIMA_nollaluokittelu (1,0,4)-3h ennuste-pitkä	ARIMA_erityistarkkaluokittelu (1,0,4)-3h ennuste-pitkä		
RMSE kW	50,4	50,4		
MAE kW	28,83	28,83		
Luokittelutarkkuus	0,47	0,54		
Ylössäätöjä oikein (suhdeluku)	0,65	0,58		
Alassäätöjä oikein (suhdeluku)	0,75	0,57		
Nollatunnit oikein (suhdeluku)	0,16	0,50		
Virheiden lukumäärä	625	598		
Ennusteeseen kuluneet sekunnit	5,4	5		
Fatal virheiden lukumäärä	91	171		