

TALLINNA TEHNIKAÜLIKOOL

Majandusteaduskond

Doris Kudre

**PÄIKESEENERGIAST ELEKTRI TOOTMISE JA TARBIMISE
LÜHIAJALINE PROGNOOSIMINE**

Magistritöö

Õppekava majandusanalüüs

Juhendaja: Aleksei Netsunajev, PhD

Kaasjuhendaja: Heili Hein-Sula, MA

Tallinn 2024

Deklareerin, et olen koostanud magistritöö iseseisvalt ja olen viidanud kõikidele selle koostamisel kasutatud teiste autorite töödele, olulistele seisukohtadele ja andmetele, ning ei ole esitanud sama tööd varasemalt ainepunktide saamiseks.

Töö pikkuseks on 9868 sõna sissejuhatusest kuni kokkuvõtte lõpuni.

Doris Kudre 07.05.2024

(kuupäev)

SISUKORD

LÜHIKOKKUVÕTE	4
SISSEJUHATUS	5
1. ELEKTRIENERGIAMAJANDUS.....	7
1.1. Elektrienergiamajandus ning konkurentsivõime.....	7
1.1.1. Elektrienergiamajanduse olemus ning selle olulisus konkurentsivõimele	7
1.1.2. Eesti elektriturg	10
1.2. Elektri tootmist ja tarbimist mõjutavad tegurid	12
1.2.1. Elektri tootmine ning seda mõjutavad tegurid	12
1.2.2. Elektri tarbimine ning seda mõjutavad tegurid	13
1.3. Varasemate empiiriliste uuringute ülevaade.....	15
2. ANDMED JA METOODIKA	20
2.1. Andmed	20
2.2. Mudeli valik.....	25
2.2.1. Mudelite loomine.....	27
2.3. Referentsmudel AR(1).....	29
2.3. Otsustusmetsa regressioon	29
2.4. Tugivektorite regressioon	31
2.5. Mudelite prognoositäpsuse võrdlemine.....	33
3. TULEMUSED JA JÄRELDUSED.....	35
3.1. Mudelite hindamise tulemused	35
3.1.1. Otsustusmetsa regressiooni tulemused	36
3.1.2. Tugivektorite regressiooni tulemused.....	39
3.2. Järeldused ja arutelu	43
KOKKUVÕTE	46
SUMMARY	48
KASUTATUD ALLIKATE LOETELU.....	50
LISAD.....	55
Lisa 1. Lihtlitsents.....	55

LÜHIKOKKUVÕTE

Käesoleva magistritöö eesmärgiks on luua mudelid, mis prognoosivad päev ette väiketootjate elektritootmist ning nende elektritarbimist. Selleks analüüsitakse päikeseenergiast elektri tootmist ning elektritarbimist mõjutavaid tegureid, valitakse sobilikud masinõppe meetodid – otsustusmetsa regressioon ja tugivektorite regressioon – ning luuakse ühesuguste sisendmuutujatega tootmis- ja tarbimismudelid. Meetodide võrdluses valitakse parima prognoositäpsusega mudel põhinedes madalaima prognoosivea tulemustele.

Päikeseenergiast elektri tootmise prognoosimiseks on olulised mitmed tegurid, nendest peamised on ilmaandmed ning päikesepargi süsteemiandmed. Elektritarbimist kujundab lisaks ilmaandmetele suuresti ka tarbija elektrikäitumine. Nii päikeseenergiast elektri tootmist kui tarbimist on varasemalt prognoositud erinevate masinõppe meetoditega, kuid puudub üks kindel parim meetod.

Töö tulemusena luuakse kõrge prognoositäpsusega tugivektorite regressiooni ning otsustusmetsa regressiooni mudelid, kus meetodite omavahelises võrdluses suudab nii tootmist kui tarbimist täpsemini prognoosida tugivektorite regressioon. Masinõppe meetodid suudavad tarbimisväärtusi prognoosida täpsemalt kui tootmisandmeid, mis ühtib ka varasemate empiiriliste uuringutega. Selle põhjuseks on tootmisandmete suurem ebaregulaarsus ning suurem sõltuvust ilmaprognoosi täpsusest. Tarbimise puhul on oluliseks aspektiks käitumislik tarbimine, mis päevade, nädalate ja aastate lõikes nii palju ei erine. Üldiselt on mõlemad mudelid väga täpsed ning võrdluses varasemate empiiriliste uuringutega on prognoosivead madalamad. Kõrgete prognoositäpsuste saavutamiseks on olulised kvaliteetsed ja relevantse andmed, täpsed ilmaprognoosid, sobivate hüperparameetrite valik ning arvutuste tegemine maakonnapõhiste andmetega.

Võtmesõnad: elektri tarbimine, elektri tootmine, päikeseenergia, prognoosimine, masinõpe.

SISSEJUHATUS

Kasvav elektrifitseerimine ja üha rangemad keskkonnasäästlikkuse nõuded suunavad ühiskonda säästlikuma ja optimaalsema elektrienergia tootmise ja tarbimise poole. Viimasel aastakümnel on nii Eestis kui globaalselt kättesaadavaks muutunud ja levinud kodumajapidamistes päikeseenergiast elektri tootmine ning selle oma tarbeks kasutamine. Peamise tõuke on selleks andnud kõrgeenenud elektri hind, millele avaldavad mõju erinevad poliitilised ja majanduslikud tegurid, näiteks koroonakriisi järgne majanduslangus ning Vene gaasitarnete täielik peatamine pärast Venemaa invasiooni Ukrainasse, mis viis gaasihinna Euroopas hüppeliselt kõrgeks (European Commission, 2024).

Päikeseenergia võimsused ja päikeseenergiast elektrit tootvate väiketootjate hulk kasvavad kiire tempoga kogu Euroopas ning agregeeritud kogused mängivad üha enam olulist rolli riikide elektrienergia kogutootmise tasandil. Kuna päikeseenergiast elektri tootmise omapäraks on selle reguleerimise võimaluse puudus, on toodangu prognoosimine hädavajalik päikeseenergia süsteemide optimeeritud tööks (Li *et al.*, 2014). Nendest asjaoludest lähtubki käesoleva magistritöö uurimisprobleem, milleks on päikeseenergiast elektritootmise ning selle tarbimise puudulik prognoosimine, mis üha suuremaid mõõtmeid võttes suurendab riske riigi elektri varustuskindlusele.

Magistritöö eesmärgiks on luua mudelid, mis lühiajaliselt (päev ette) prognoosivad väiketootjate elektritootmist ning nende elektritarbimist. Magistritöö eesmärgi täitmiseks on püstitatud järgmised uurimisülesanded:

1. Kirjeldada elektrienergiamaajanduse olemust ning selle olulisust riigi konkurentsivõimele.
2. Analüüsida elektrienergia tootmist ja tarbimist mõjutavaid tegureid, rõhuasetusega päikeseenergiail.
3. Koostada Eesti andmete põhjal päikeseenergiast elektritootmise ning elektritarbimise lühiajalised prognoosimudelid.

Magistritöö on jagatud kolme peatükki. Töö esimeses peatükis antakse ülevaade elektrienergiamajanduse olemusest ning sellest, kuidas elektrimajandus on läbi energiatõhususe seotud riigi konkurentsivõimega. Lisaks kirjeldatakse täpsemalt Eesti elektriturgu. Peatükis seletatakse ka elektri tootmist ja tarbimist mõjutavaid tegureid, eesmärgiga kirjeldada töös kasutatavate mudelite teoreetilist tausta. Esimese peatüki lõpetab varasemate empiiriliste uuringute ülevaade.

Teises peatükis kirjeldatakse töö analüüsis kasutatavaid andmeid ja prognoosimeetodeid. Andmetest tehakse ülevaade erinevate jooniste ning kirjeldava statistika abil. Varasemate empiiriliste uuringute põhjal tehakse otsus referentsmudeli ning erinevate masinõppe prognoosimudelite valikuks. Mudelite kirjelduses keskendutakse selle arvutuspõhimõttele ning mudeli võimalikele tugevustele ning nõrkustele. Peatükis tuuakse välja ka erinevad statistikud, mille järgi prognoosimudeleid omavahel võrrelda.

Magistritöö kolmandas osas tuuakse välja erinevate prognoosimudelite hindamise tulemused ning otsustatakse, missugune mudel suudab tulevikuväärtusi prognoosida kõige täpsemini. Järelduste ning arutelu alapeatükis võrreldakse saadud tulemusi esimese peatüki teoreetiliste vaatepunktide ja varasemate empiiriliste uuringutega.

Töö autor soovib tänada juhendajaid Aleksei Netšunajevit ja Heili Hein-Sula hea ning asjakohase juhendamise eest. Tänuõnad ka Eesti Energia Kaggle teadusvõistluse eestvedajatele, kelle abil oli võimalik saada ligipääs magistritöös kasutatud andmetele.

1. ELEKTRIENERGIAMAJANDUS

1.1. Elektrienergiamajandus ning konkurentsivõime

1.1.1. Elektrienergiamajanduse olemus ning selle olulisus konkurentsivõimele

Elektrimajandus koondab endast informatsiooni elektrisüsteemi toimimise kohta – infot elektrienergia tootmise kohta, selle ülekandmisest ja jaotamisest elektrivõrkudes, sellega kauplemisest elektriturul ning tarbimisest. Elektrienergia olemasolu tänapäeva ühiskonnas on üks oluline eeldus paljudele tegevustele tööstuses, ettevõtluses ja kodustes majapidamistes. (Elektrimajandus, 2021)

Elektrit on aastakümneid toodetud kasutades taastumatuid energiaallikad, millega kaasneb õhusaaste ja keskkonnareostus. Keskkonnateadlikkuse tõusuga on nii regionaalsel kui globaalsel tasemel orienteerunud elektrimajanduse arendustegevus ühiskonna säästvate arengule. Investeeringud hüdro-, tuule- ja päikeseenergiasse ehk taastuenergiasse on lahenduseks jätkusuutlikule elektritootmisele. Samas on iga-aastaselt näha elektritootmise ning -tarbimise kasvu, kuna majapidamiste kodumasinade ja rahvaarvu suurenemisega kasvab ka elektrinõudlus. (IEA, 2024; European Commission, 2024)

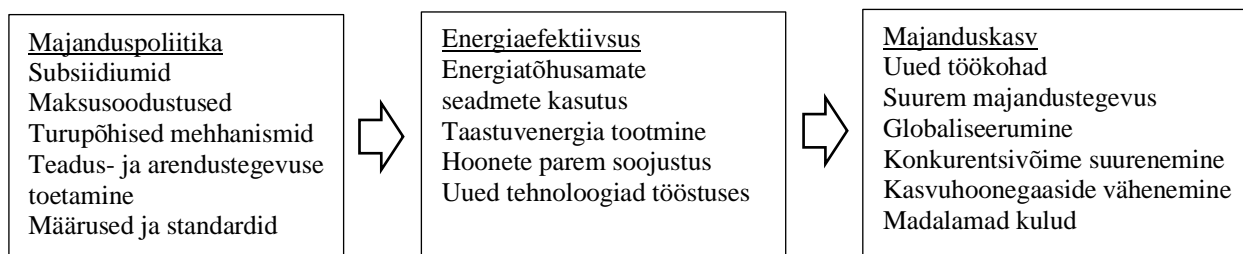
Üha enam kasvav elektrinõudlus ning ranged keskkonnasäästlikkuse nõuded on viinud süsihappegaasi (CO₂) emissioonide hinna tõusuni, mis on üheks oluliseks energiahinna komponendiks ning põhjuseks elektrihindade tõusule. Kõrged hinnad on tekitanud diskussioone riikide konkurentsivõime jätkusuutlikkuse osas ning nii riiklikul kui ka Euroopa tasemel on riigi konkurentsivõime tõstmine oluliseks eesmärgiks, pidades silmas nii kliimaeesmärke kui jätkusuutlikkust. (European Commission, 2024)

Euroopa-üleselt võrreldakse end ka suuremate globaalsete konkurentidega, näiteks Ameerika Ühendriikide ja Hiinaga. Välismaiste konkurentide „eeliseks“ on elektri tootmise meetodid, näiteks erinevate fossiilsete kütuste kõrge kasutus, mida Euroopas peetakse kahjulikuks, kuid mis annab neile eelise majanduslikus konkurentsivõimelisuses Euroopa riikide ees, sest nimetatud

suurriikides puudub süsiniku-maksustamise süsteem (Yin, 2024). Euroopa Liidu negatiivsete konkurentsimeetmete minimeerimiseks nähakse täiendava võimalusena globaliseerumist ning tugeva teadus- ja arendusvaldkonda ekspordi. ELi ettevõtted, kes toovad turule uusi energiasäästlike omadustega tooteid, võivad edukad olla ülemaailmselt (European Council, 2024).

Energiatõhusus tähendab energia ja muude ressursside tarbimise optimaalsust seadme või ehitise kasutamisel (Energialalgud, 2024). Energiatõhususe peamine eesmärk on energia optimaalsem kasutus, kus sama väljundi saamiseks tarbitakse vähem energiat. Seda eesmärki saavutatakse läbi uute tehnoloogiate kasutamise, protsesside parendamise ning hoonete/seadmete disainide täiustamise kaudu. Lisaks on energiasäästu eesmärgiks ka tarbijate igapäevaste harjumuste ja käitumise muutmine. (Mushafiq *et al.*, 2023) Käitumuslik energiasääst on kliima- ja energiaalase arutelu oluline keskkonnasühholoogiline aspekt, millega arvestatakse näiteks hoonete ja transpordi energiakasutuse modelleerimisel ning samuti riiklikul energianõudluse prognoosimisel ja poliitika kujundamisel (Debnath & Mourshed, 2018).

Tark elektritarbimine ja riigi konkurentsivõime on omavahel seotud läbi tehnoloogilise arengu, energiatõhususe, keskkonnasäästlikkuse ning õiguslike ja hariduslike raamistike. Riikidel, kes suudavad neid elemente terviklikult käsitleda, on maailmaturul konkurentsieelis. Riigi majandusarengu ja konkurentsivõimelisuse kasvuks läbi majanduspoliitiliste mehhanismide on välja töötatud raamistik, mis on suunatud energiaefektiivsuse parandamisele. Need majanduspoliitilised mehhanismid mängivad olulist rolli energiatõhususe edendamisel ja energiatarbimise vähendamisel, aidates seeläbi kaasa keskkonnasäästlikkusele ja majanduskasvule. (Faiella & Mistretta, 2022; Mushafiq *et al.* 2023)



Joonis 1. Majanduspoliitiliste mehhanismide raamistik energiaefektiivsuse ja majanduskasvuni
Allikas: (Faiella & Mistretta, 2022; Mushafiq *et al.*, 2023, IEA 2024); autori koostatud.

Majanduspoliitiliste otsuste kaudu on võimalik tõsta tootmisprotsesside ja elamute energiatõhusust, mis soodustab riigi majanduskasvu ning konkurentsivõimet. Subsiidiumid, maksusoodustused ning määruste ja standardite adapteerimine on riigipoliitika tasandil tehtavad otsused, mis saavad otseselt aidata kaasa energiatõhusate meetmete kiirele ja efektiivsele rakendusele. Maksusoodustusi saab kohandada energiatõhusatele toodetele, mis vähendab nende toodete hinda ja soodustab kasutuselevõttu. Toetused taastuvatele energiaallikatele, nagu tuule-, päikese- ja hüdroenergia, aitavad alandada taastuenergia tehnoloogiate hinda, edendades seeläbi ilma heitegaasideta energiaallikate kasutamist.

Rakendusmäärused ja standardid toodete, hoonete ja tööstusprotsesside energiatõhususele on veel üheks mehhanismiks, kuidas aidata kaasa energiaefektiivsuse kasvule, kuna need aitavad kehtestada selged võrdlusalused ja soodustavad tehnoloogilisi arendusi. Turupõhised mehhanismid, näiteks piiramise ja kauplemise süsteemid (*cap and trade*), võivad motiveerida ettevõtteid vähendama oma energiatarbimist ja investeerima energiatõhusamatesse tehnoloogiatesse. Tuntuimaks turupõhiseks mehhanismiks Euroopa Liidus on CO₂-emissioonide hinnakujundus. Teadus- ja arendustegevuse toetamise kaudu paraneb energiatõhusamate tehnoloogiate ja paremate ehitusdisainide kasutuselevõtt ning teadlikumate individuaalsete käitumisharjumuste rakendus. (Faiella & Mistretta, 2022; Mushafiq *et al.*, 2023; IEA, 2024)

Energiatõhusust parandavad meetmed toetavad majanduse stimuleerimise programmide eesmärke, toetades olemasolevat tööjõudu ja luues uusi töökohti, suurendades majandustegevust töömahukates sektorites. (*Ibid.*) Lisaks on leitud ka pikemaajalised eelised: konkurentsivõime suurenemine, kasvuhoonegaaside heitkoguste vähenemine, energia taskukohasuse paranemine ja madalamad arved. (IEA, 2020)

Targa elektritarbimise ja -tootmise kaudu saab tõsta riigi konkurentsivõimet mitmel viisil. Uued tehnoloogiad ja targad süsteemid võimaldavad paremat energiatõhusust nii tarbimise kui tootmise aspektist. Tark elektritarbimine nõuab tehnoloogilist innovatsiooni, mis aitab kaasa uute ettevõtete ja tööstusharude loomisele. Lisaks toetab tark elektrimajandus sageli roheline energia tootmist ja keskkonnasäästlikke lahendusi. Riigid, kes liiguvad jätkusuutlikuma energiapoliitika suunas, omistavad positiivse maine, mis omakorda suurendab võimalusi globaliseerumiseks. Näiteks võib see meelitada investeringuid ja äriettevõtteid, kellele on oluline sotsiaalne ja keskkonnaalne vastutustundlikkus. Edukate uute tehnoloogiate arendamine ja rakendamine soodustab tehnoloogiate ja teadmiste ekspordi võimalusi. Tehnoloogiline koostöö teiste riikidega aitab

suurendada riigi positsiooni rahvusvaheliselt ja suurendada selle konkurentsivõimet. Edukas teadus- ja arendustöö võib kaasa tuua ka uute töökohtade loomise valdkondades, mis on seotud tehnoloogia arendamise, paigaldamise ja hooldusega, mis läbi tööhõive suurendamise aitavad suures pildis kaasa riigi konkurentsivõime paranemisele. (Faiella & Mistretta, 2022)

Taastuvenergiast toodetud energia hind on kasvavate CO₂ emissioonide hindade ning elektri tootmise sisendhindade jätkuva kasvu valguses oluliselt odavam, mis väljendub otseselt ka madalamates kuludes. Lisaks väljendub majanduslik kokkuvõtte ka energiatõhusamate seadmete kasutuses, mis vähendab energiakulusid majapidamiste ja ettevõtete jaoks. Ettevõtete jaoks suurendab see nende konkurentsivõimet ning seda iseäranis energiamahukates sektorites. Lisaks energiatõhusamate seadmete kasutusele majapidamises on igakuiste kulude kokkuvõtte võimaluseks ka elamute energiatõhustamine. (Faiella & Mistretta, 2022; Arenguseire Keskus, 2023)

1.1.2. Eesti elektriturg

Üheks olulisemaks energeetikavaldkonna eesmärgiks Euroopa Liidus on hästi toimiv ühtne elektriturg. Eestis toimub elektriga kauplemine börsil, kus elektrienergia hind kujuneb nõudluse ja pakkumise suhtena. Börsihinna arvutamiseks kogutakse kokku kõigi börsil osalejate pakkumised, mille käigus tekib kaks kõverat, mille lõikumispunkt ehk tarbimise ja tootmise tasakaalupunkt on börsihind, millega turuosalisel elektrit börsil ostavad ja müüvad. Turu pakkumise määravad piirkonna tootmisvõimsused ja ühenduste olemasolu, mis tagavad elektri liikumise nii naaberriikidega kui ka siseriiklikult. Turu nõudluse loovad tarbijad enda elektritarbimisega, mille maht sõltub erinevatest teguritest – näiteks aastaajast, nädalapäevast, kellaajast. (Elering, 2024)

Eesti kuulub ühisesse Nord Pooli elektriturust süsteemi koos Soome, Rootsi, Norra, Läti ja Leeduga. Riikide vahel ning riigisiselt on elektri ülekandeks vajalikud ühendused, mille kaudu toimub elektri eksport ja/või import. Suuremad riigi jagunevad omakorda mitmeks piirkonnaks, kogu Eesti kuulub ühte kauplemisspiirkonda. Kõige olulisemad tegurid Eesti elektrienergia hinna kujunemisel on tootmisvõimsused regioonis, ülekandevõimsused teiste riikidega, ilm (sh hüdroenergia tase Põhjamaades ja Lätis) ning primaarkütuste ja CO₂ hind. (*Ibid.*)

Elektriga kauplemisega tegelevad elektriettevõtted. Kauplemise eesmärgiks ei ole vaid kasumi teenimine, vaid igapäevaselt igaks tunniks klientide elektriga varustamine, mis toimub elektrijaamade ja koormuste valimise kaudu, arvestades ülekande- ja töövõimsuse piirangutega. (Creti & Fontini, 2019) Vastavalt elektriturust toimimise põhimõtetele peab elektriturul olema

tootmine ja tarbimine igal ajahetkel võrdne ning iga turuosaline peab tagama, et tema poolt võrku antud ja/või ostetud elektrienergia kogus oleks igal ajahetkel võrdne tema poolt võrgust võetud ja/või müüdud elektrienergia kogusega. Elektrituru juurde kuulub ka teine, tarbijatele „peidetud“ turuosa – päevasisene turg (*intraday market*) ning sagedusreservide turg, mis moodustavad kogu turu mahust märkimisväärse osa. (Elering, 2023)

Päevasisese elektrikauplemisega tegelevad bilansihaldurid. Bilansihaldurid prognoosivad oma portfellis olevate klientide tarbimist/tootmist ning vastavalt sellele teevad igapäevaselt *Nord Pool Spot* (NPS) elektribörsile oma ostu- ja müügipakkumised. Bilansenergia ostuhind on aga alati kallim kui NPS hind. Seetõttu on kulude madalal hoidmiseks väga oluline prognoosida elektritarbimist ning -tootmist. Klientide tarbimine ning tootmine ei ole alati kergesti prognoositav ning tegelikku tarbimist ja tootmist võivad mõjutada paljud raskesti ennustatavad tegurid, nagu näiteks avariid elektrijaamades või tootmisüksustes, ilm (temperatuur, tuul, päikeseline või pilvine ilm), tootmisvõimsuste suurendamine, puudulik kommunikatsioon. Kuna praktikas pole võimalik igatunniselt täpselt tarbimist ja tootmist prognoosida, siis vahe planeeritud ja mõõdetud tarnete vahel tuleb katta vastavalt bilansenergia ostes või müües. Kuna elektrienergiat tarbitakse ja toodetakse samal ajal ja selle salvestamine suuremates kogustes pole olnud majanduslikult otstarbekas, tuleb elektrienergiat targalt juhtida (Elering, 2024). Eleringi iga-aastases varustuskindluse aruandes on olulist tähelepanu pööratud asjaolule, et taastuenergia tootmise jätkuv kasv põhjustab tootmise prognoosivigade kasvamise, mis toob endaga kaasa kiirete reservide nõudluse kasvu. (Elering, 2023).

Eesti elektritootmine on elektritootmise algusajast alates põhinenud suuresti põlevkivist elektri tootmisel. Põlevkivi-tootmisjaamad on hästi reguleeritavad, kuna tootmisjaamu saab üles- ja alla reguleerida vastavalt päevasele elektrivajadusele. Viimastel aastatel, kui taastuenergia areng on hoogustunud nii Eestis kui teistes Nord Pooli kuuluvates riikides, on juurde tekkinud suures koguses kontrollimatu iseloomuga elektri tootmist. See tähendab, et päikesepaneelid toodavad päikest ainult siis, kui päike paistab ning tuulikud toodavad elektrit ainult siis, kui tuul puhub. Nende kahe tootmisliigi puhul ei ole võimalik elektritoodangut kontrollida nagu see on võimalik reguleeritavate tootmisviiside puhul. (Elering, 2024) Pidades meeles elektrituru põhiprintsiipi, et tarbimine ja tootmine peavad igal ajahetkel olema tasakaalus, muudab taastuenergia kiire levik elektri tootmise prognoosimise kriitiliselt oluliseks.

Eesti on keskmise energiatarbimisega riik – 2022. aastal oli ühe inimese keskmine elektritarbimine 5,8 MWh, mis on enam-vähem Euroopa riikide keskmisel tasemel (Eurostat, 2024). Eesti on energiamahukas nii kodumajapidamistest kui ka tööstuses (Enerdata, 2024), samas on Eesti puhul

näha elektritarbimise kasvu aeglustumist, kuna on rõhku pandud elektritõhususele ning vähem elektrimahukatele majandustegevustele (IEA, 2024).

2022. aastal toodeti Eestis sisemaiselt 7,6 TWh elektrit ning tarbiti 8,5 TWh. Toodetud elektrist 65% toodeti mittetaastuvatest allikatest ja 35% taastuvatest allikatest. Päikeseenergiast toodetud elektri osakaal kogutootmisest oli 2022. aastal 7,4%, mis moodustas umbes viiendiku taastuenergia toodangust. Ekspordi- ja impordimahtude võrdluses oli Eesti 2022. aastal netoimportija mahuga 0,9 TWh. Avatud elektriturul tegutsemine tähendab tihedat elektri liikumist riikide vahel, nii ka Eestis, kus 2022. aastal olid ekspordi- ja impordimahud pea samas mahus kui tegelik tootmine ja tarbimine. 2022. aastal imporditi Eestisse 7 TWh elektrit ning eksportisime 6,1 TWh. (Enerdata, 2024; Elektribilanss, 2024) Ajalooliselt on Eesti importinud elektrit ja gaasi odavalt Venemaalt, kuid pärast 2022. aasta Vene agressiooni Ukrainasse on kõik kaubanduspiirid Venemaaga suletud. See tähendab, et viimastel aastatel on varustuskindlus ning toimiv elektriturg olnud Eesti elektrimajanduses varasemast veelgi olulisem.

1.2. Elektri tootmist ja tarbimist mõjutavad tegurid

1.2.1. Elektri tootmine ning seda mõjutavad tegurid

Ehkki elektrivoolu tootmiseks on teoreetiliselt mitmeid võimalusi, toodetakse suurem osa elektrienergiast kineetilise energia muundamisel elektrienergiaks. Tavaliselt toimub see elektrijaamades turbiinide abil, kuid viimasel aastakümnel on kasvanud taastuenergia tähtsus – tuule- ning päikeseenergia on muutumas üha olulisemateks ja levinumateks väikesemahulisteks tootmistehnoloogiateks. Erinevad elektrigeneraatorid erinevad üksteisest seesuguste omaduste poolest nagu juhitavus, võimsus, tootmiskulud, kasutegur ja keskkonnamõju. (Biggar & Hesamzadeh, 2014)

Elekter toodetakse erinevates elektrijaamades, kus see edastatakse elektrivõrku, mis on jagatud kaheks osaks: põhivõrguks ning jaotusvõrguks. Põhivõrgu ehk kõrgepinge ülekandevõrgu eesmärk on edastada elektrienergiat distantsilt kaugemale ning jaotusvõrgu ehk madalpinge ülekandevõrgu eesmärk on elektrienergia jaotamine lõpptarbijatele. Juhitavate elektritootmisvõimsuste korral reguleeritakse elektritoodangut vastavalt nõudlusele ja turule pääsemisele. (Elektrimajandus, 2021; Cretì & Fontini, 2019)

Päikeseenergiast elektri tootmise üheks oluliseks tunnuseks on aga toodangu üle kontrolli puudumine, mistõttu pole pakkumine nõudlusgraafikule alati vastav. Tulevikus nähakse

päikeseenergiast toodetava elektri võrku jagamise ebastabiilsust võimaliku ohukohana, mis toob endaga kaasa duck curve ehk pardikujulise kõvera tunnipõhisel elektrinõudluse ja -pakkumise graafikul. Päikesepaneelid toodavad majapidamistes ja ettevõtetes päeva jooksul elektrienergia ülejääki, mis suunatakse võrku tagasi, kuid pärast päikeseloojangut ja käitumuslikult energianõudluse suurenemist sõltuvad need majapidamised rohkem võrgust saadavale elektrile, mis toob kaasa energiatarbimise hüppelise kasvu. (Olczak *et al.*, 2021; Padhee & Pal, 2018)

Päikeseenergia on muutunud kodumajapidamistele kättesaadavaks taastuvaks energiaallikaks, mida on uurima ja analüüsima hakatud viimasel aastakümnel. Varasemalt on päikeseenergia aina suureneva elektritoodangu kohta uuringuid tehtud peamiselt eesmärgiga uurida, kuidas see mõjutab kodumajapidamiste tarbimist ning elektrivõrkude töökindlust (Oberst *et al.*, 2018; Pablo-Romero *et al.*, 2014; Padhee & Pal, 2018; Olczak *et al.*, 2021). Saksamaa ja Poola on Euroopas eeskujuks päikeseenergia tootmises ning nende riikide andmete põhjal on läbi viidud uuringuid, kuidas päikeseenergiast toodetud elektri üleriigiline levik on riigi elektritarbimist mõjutanud. Ühelt poolt paistab andmete põhjal välja korrelatsioon madalama energiatarbimise ja päikeseinvesteeringute vahel, kuid sotsiaalmajanduslike ja eluasemega seotud tegurite lisamise korral ei ole valitud uuringutes leitud statistiliselt olulist erinevust tarbimiskäitumises mikrotootjate ja mittetootjate vahel (Oberst *et al.*, 2018; Olczak *et al.*, 2021).

Päikeseparkidel on mitmeid tegureid, mis väljundvõimsust mõjutavad, aga üldiselt jagunevad need kaheks suuremaks valdkonnaks. Ilmaga seotud teguritest on märkimisväärseimad tegurid päikesekiirgus ja temperatuur, kuid seos on leitud ka tuule tugevuse ja suuna, õhurõhu, sademete hulga, pilvisuse ja õhuniiskusega (Wolff *et al.*, 2013; Li *et al.*, 2014; Dolara *et al.*, 2015). Lihtsaimateks süsteemi endaga seotud teguriteks on asukoht, süsteemi ülekande efektiivsus ja paigaldusnurk, kuid süvenedes sügavamale päikesepaneelide tehnilistesse parameetritesse, võib päikesepaneelide tootmiskoguste hulka mõjutada veel mitmed muutujad. Näiteks mõjutavad paneelidel olev füüsiline tolmu või mustus, paneelide vanus, paneelirakkude temperatuur, päikese langemisnurk ja koormuse seisund. Tehnilised aspektid on siiski elektritootjatele üsna piiratud, kuna tehnilisi aspekte kirjeldatakse tavaliselt ainult normaaltingimustel. (Li *et al.*, 2014; Dolara *et al.*, 2015)

1.2.2. Elektri tarbimine ning seda mõjutavad tegurid

Elektrit tarbitakse erinevates seadmetes erinevatel eesmärkidel, nii erinevates suurustes äri- ja tööstuse ettevõtetes kui ka kodumajapidamistes isiklikel eesmärkidel. Kodumajapidamises

kasutatavate elektriseadmete arv on koos sotsiaalmajanduse arenguga kasvanud ning samuti on kasvanud elamute elektritarbimise osakaal kogu energiatarbimises. Kodumajapidamiste energiasäästu edendamine riigi säästva arengu jaoks muutub ressursipiirangute korral üha olulisemaks. (Guo *et al.*, 2018)

Väikesemahuliste tootmistehnoloogiate, nutikate seadmete, elektrisõidukite ja muude elektrisalvestustehnoloogiate areng muudab meie suhtumist elektritööstuse rollist. Kui varem oli elektrimajandus üles ehitatud suurte tootmisüksuste elektritarbimisele väikestele passiivsetele elektritarbijatele, siis tehnoloogia arenguga võib nüüd hoopiski läheneda, et kõik elektritarbijad (nii tootjad kui ka tarbijad) on võimelised tootma, tarbima ja reageerima elektrituru tingimustele. Peamiselt ei eristata enam tootjaid ja tarbijaid, vaid suurtarbijaid (tootjad ja koormused), kes osalevad otseselt elektri hulgimüügiturul, ja väiketarbijaid, kes osalevad kaudselt, võib-olla jaemüüja kaudu. (Biggar & Hesamzadeh, 2014)

Uurimisprobleemiks on kujunenud majapidamiste elektrikäitumine, sest suurenenud elektritarbimisega on vaja leida viise, kuidas energiatarbimist vähendada või energiatõhusust parandada. Peamiseks selgituseks energiatõhususe puudujäägi selgituseks on kirjanduses välja toodud inimeste vähest teadlikkust ja vale käitumist. (Sanstad *et al.* 2006; Wolske *et al.*, 2020) Peamised elektritarbimise käitumist mõjutavad psühholoogilised tegurid on suhtumine, uskumused, väärtused, harjumused, eelistused ja keskkonnateadlikkus. Elektritarbimine majapidamise tasandil on väga varieeruv, kuna see sõltub majapidamise liikmete elustiilist, hoonestusest, ilmast ja kuupäeva ja sesoonsusega seotud teguritest. (Huebner *et al.*, 2016). Ühtlasi on leitud, et inimeste puuduslik info võib olla üheks oluliseks põhjuseks, miks ei tehta investeeringuid energiatõhusamatesse seadmetesse ega taastuvenergia tootmisesse. Sarnaseid mõtteviise on leitud ka tööstuste ja äride vaates, mis viib tagasi alapeatüki 1.1 juurde – kodumajapidamiste ning ettevõtete energiatõhususel on mõju riigi konkurentsivõimele.

Levinuimad meetodid elektritarbimise analüüsimiseks on tegurite-põhised prognoosimeetodid, mis proovivad leida võimalikult täpseid elektritarbimise prognoose, võttes sisendmuutujateks nädalapäeva, ilmategurid (nt temperatuuri, niiskuse), piirkonna teguri ning kalendripühad, mil elektritarbimine on tavapärasest erinev. Nende põhjal on saavutatud hea prognoositäpsus baas- ning agregeeritud andmetele. Ekstreemväärtuste, näiteks päevade tippude ja madalpunktide prognoosimisel on andmetes suured kõikumised, need on madala prognoositäpsusega ning võimalik hinnata ainult tõenäosuslikult. (Fan *et al.*, 2022) Elektritarbimise prognoosimise üheks suurimaks probleemiks on suure hulga sisendmuutujate vajalikkus, millega kaasnevad keerulised arvutused ja suured kulud (Cretì & Fontini, 2019).

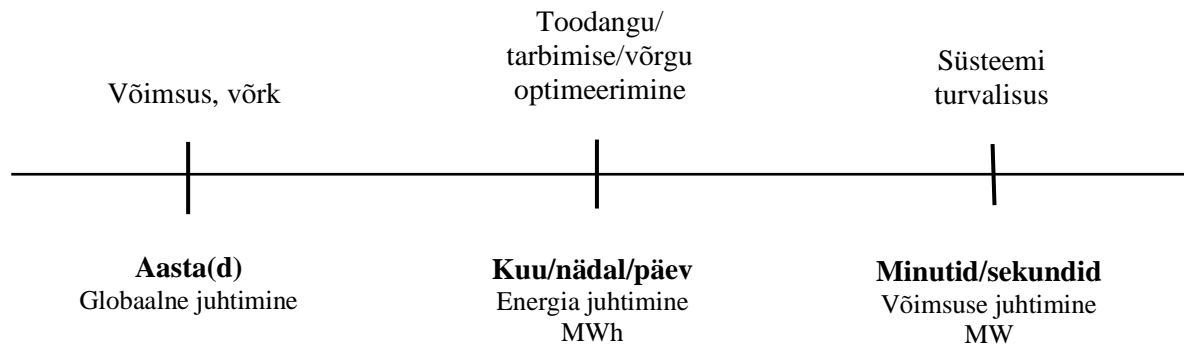
Elektritarbimist analüüsivates regressioonvõrrandites on leitud statistiliselt oluline ja negatiivne seos elektritarbimise ning elektrihinna vahel (Özcan *et al.*, 2013; Xie *et al.*, 2016; Romero-Jordan *et al.*, 2014). Palju on uuritud elektrienergia nõudlust mõjutavaid tegureid Euroopa riikide andmete põhjal, kus on leitud tarbimise statistiliselt oluline seos gaasi hinnaga, eelneva aasta elektrinõudlusega, õhutemperatuuri muutustega ning mitmete majapidamiste tasemel kogutud mikroandmetega (Romero-Jordan *et al.*, 2014; Özcan *et al.*, 2013; Sanquist *et al.*, 2011; Fikru & Gautier, 2015).

Eesti andmeid analüüsides on leitud, et tarbimine muutub vastavalt kellaajale, nädalapäevale, ilmale ja teistele välistele mõjuteguritele ning samuti võivad päevad üksteisest suuresti erineda. Tuleviku tarbimise kasvu ja ühiskonna suuremat elektrifitseerimist arvestades on oodata elektritarbimise paindlikkuse ja hinnaelastsuse kasvu. (Elering, 2023) Mugavad digitaalsed lahendused võimaldaks nihutada elektritarbimist elektrisüsteemi mõttes vähemkoormavale tunnil, muutes traditsioonilist elektritarbimist ning mis muudab elektritarbimise analüüsimise veelgi keerulisemaks.

1.3. Varasemate empiiriliste uuringute ülevaade

Energia planeerimise mudelitel (EPM) on poliitika kujundamisel ja energiasektori arendamisel asendamatu roll. EPM-id on olulised, et aidata erinevatel osapooltel teha teadlikke otsuseid energiasektori arenguks globaalsel, piirkondlikul ja riiklikul tasemel. Enamasti lähtutakse prognoosimismeetodi valikul andmete kättesaadavusest ning lõppeesmärgist, lisaks on meetodi valiku tegemisel oluline ka tulemi lühi- ja pikemaajalisem rakendatavus ning asjakohasus planeerimise ja poliitiliste eesmärkide jaoks. (Debnath & Mourshed, 2018)

Joonisel 2 kujutatud skaala annab ülevaate, missugused elektrienergeetika planeerimise eesmärgid on erinevate intervallidega andmetega töötamisel. Aastaste andmetega töötatakse globaalse juhtimise tasandil, kus peamisteks eesmärkideks on võimsuse ja võrgu töökindlus ehk varustuskindlus. Energiat juhitakse lühema intervalliga andmetel, mis omakorda jaguneb kuu-, nädala- ja päevapõhiseks. Sellel tasandil optimeeritakse ja prognoositakse toodangu, tarbimise ja võrgu tööd. Võimsust juhitakse kõige lühema intervalliga, minuti- ja sekundipõhiselt. Vastavalt elektrituru toimimise reeglitele peab süsteemides tootmine ja tarbimine olema tasakaalus igal sekundil, seega on minuti- ja sekundipõhine energiajuhtimine eelkõige oluline süsteemi toimimise ja turvalisuse eesmärkidel. (Voyant *et al.*, 2017)



Joonis 2. Prognoosimise skaala energia juhtimisel elektrisüsteemides
Allikas: Voyant *et al.* (2017), autori koostatud.

Viimaste aastakümnete areng andmetöötluses on võimaldanud kasutada arvutusmahukaid meetodeid prognoositäpsuse parandamiseks ja arvutusaja lühendamiseks, parandades seeläbi nende rakendatavust. Tehisintellektil põhinevaid tehnikaid kasutatakse nüüd laialdaselt statistiliste mudelite hindamiseks ja kontseptsioonide täiustamiseks. Kui suuremahuline EPMi ülevaade esitab viiekümne erineva meetodi kasutamist viimase 40 aasta jooksul, siis taastuenergia ning eriti päikeseenergia toodangu ja tarbimise prognoosimist on uuritud oluliselt vähem. (Debnath & Mourshed, 2018; Weron, 2014)

EPMi keskmes on tavaliselt energianõudluse ja -pakkumise prognoosimine. Elektrienergia tootmise prognoosimine on saanud taastuenergia kiire levikuga oluliseks uurimisprobleemiks, kuna päikeseenergiast elektri tootmist ei saa reguleerida, aga toodangu prognoosimine on vajalik fotogalvaanika (*photovoltaic*, edaspidi PV) süsteemide optimeeritud tööks. (Li *et al.*, 2014) Päikeseenergia vahelduv iseloom seab taastuenergia süsteemihalduritele palju väljakutseid seoses tegevuse planeerimise ja ajastamisega. Seetõttu on päikeseenergiast elektri prognoosimine oluline elektrisüsteemide töö juhtimiseks ja majandusliku jõudluse hindamiseks. (Ramsami & Oree, 2015)

Varasemad empiirilised uuringud on leidnud, et elektritootmise ja -tarbimise prognoositäpsuse osas näitavad masinõppe meetodid oluliselt paremaid tulemusi kui statistilised meetodid, eriti parameetrite puhul, mille lähteandmetes on suurem varieeruvus. Statistilisi meetodeid kasutatakse ainult lühikese ja keskmise intervalliga andmete puhul, samas kui masinõppe meetodid on eelistatud kõigi intervallide puhul (lühikesed, keskmised ja pikad). (Debnath & Mourshed, 2018)

Päikeseenergia tootmise prognoosimisel on masinõppe meetoditest kasutatud erinevaid närvivõrkude süsteemi meetode (-NN), tugivektorite regressiooni (SVR), üldistatud vähendatud

järgu regressiooni (GRR), mitmest lineaarset regressiooni (MLR), k-lähimaid naabreid (KNN), pikka lühiajalist mälu (LSTM), otsustusmetsa (RF). Erinevates töödes on saadud võimekaid ja väikese veaga prognoosimudeleid, kuid tööde omavahelises võrdluses võib nentida, et elektrienergia prognoosimiseks puudub üks kindel parim meetod. Samuti on leitud vastakaid tulemusi, et kas parem on kasutada eraldiseisvat mudelit või hübriidmudelit. Kõige rohkem on küll kasutatud närvivõrkude meetode, kuid mitmes töös on nenditud, et teised meetodid suudavad NN-mudeli võimekust ületada.

Elektritarbimise prognoosimisel on kasutatud peamiselt järgmiseid mudeleid: erinevaid närvivõrkude süsteemi meetode (-NN), tugivektorite regressiooni (SVR), pikka lühiajalist mälu (LSTM), rekursiivset funktsioonide kõrvaldamist (RFE), otsustuspuud ja otsustusmetsa (DT/RF), XGBoosti (XBG), Gradiendi võimendamise mudelit (GBC). Tihti on mitut meetod kasutatud koos, levinuim on klassifitseerimise ja regressioonmudeli koos kasutamine. Sarnaselt tootmise prognoosimisele on ka tarbimise puhul leitud, et puudub üks kindel meetod, mis annaks erinevate andmete puhul alati täpsemaid tulemusi. Tarbimismudelitel on välja toodud mitme töö puhul otsustusmetsa hea võimekus ning mitte vaid tunnuste valikul, vaid regressiooni läbiviimisel. (Trivedi *et al.*, 2022; Wang *et al.*, 2018), mille eeliseks on võimas tunnuste olulisuse funktsioon. Ülevaatic tabel varasematest empiirilistest uuringutest on esitatud tabelis 1.

Enamikes tabelis 1 esitatud töödes on kasutatud mudelite omavaheliseks võrdluseks RMSE (*root mean squared error*) ehk keskmise ruutvea väärtust, mis on hea indikaator andmestiku-siseste prognooside põhjal võrdlemiseks, kuid universaalseks võrdlusstatistikuks on normaliseeritud RMSE ehk nRMSE. nRMSE protsendiline väärtus varasemates töödes nii üksikmudelitel kui hübriidmudelitel varieerub 5%-st kuni 26%-ni. Madalamaid prognoosivigasid on leitud näiteks GA-SVM-ANN-ARIMA hübriidmudeliga üks tund ette prognoosides ning SVR üksikmudeliga, kus nRMSE väärtused vastavalt 5% ja 6,2%. (Fan *et al.*, 2022; Wolff *et al.*, 2013) Enamasti pole tootmist ning tarbimist koos prognoositud, vaid Trivedi *et al.* (2022) vaatas tootmis- ja tarbimisandmeid koos ning kasutas neil samu prognoosimeetode koos erinevate tunnuste ja mudelite ülesehitustega.

Tabel 1. Varasemad masinõppe analüüsid elektri tootmise ja tarbimise prognoosimiseks

Viide	Prognoositaav tunnustav	Sisendmuutujad	Prognoosipikkus	Meetodid	Peamised tulemused
Li <i>et al.</i> (2014)	PV toodang	aeg ja kuupäev, ilmaandmed (temperatuur, sademete hulk, päikesepaiste kestus, niiskus)	24 tundi ette	ARIMA, ARMAX, NN	ARMAX mudel koos ilmaandmetega on oluliselt täpsem prognoosimudel kui ARIMA ja tehisnärvivõrkude mudel
Ramsami & Oree (2015)	PV toodang	aeg ja kuupäev, ilmaandmed (temperatuur, õhurõhk, niiskus, sademed, päikesekiirgus, tuule suund ja kiirus)	24 tundi ette	stepwise-NN hübriidmudel, GRR, FFNN, MLR	Hübriidmudelil parem prognoosivõime kui eraldiseisvatel, võimekas meetod ainult ilmaandmete põhjal prognoosimiseks
Dolara <i>et al.</i> (2015)	PV toodang	aeg ja kuupäev, ilmaandmed (temperatuur, tuule tugevus, päikesekiirgus), päikesepaneelide karakteristikud.	1 tund ette	<i>Erinevad tehnilise analüüsi mudelid*</i>	Prognoosimudelid võimekad, prognoosivead peamiselt seotud ilmaprognoosidega. Hübriidmudelitel pole üksikmudelite ees eelist.
Wolff <i>et al.</i> (2013)	PV toodang	aeg ja kuupäev, suhteline võimsus, võimsus pilvitu ilmaga, temperatuur, päikesekiirgus, pilvisus)	1 tund ette	KNN, SVR	SVR mudel parem kui KNN. Prognoositäpsus sõltub andmete kvaliteedist, tunnuste valikust ning mudelite parameetritest.
Trivedi <i>et al.</i> (2022)	PV toodang, tarbimine	aeg ja kuupäev, tootmis- ja tarbimisandmed viiteajaga, ilmaandmed (temperatuur, päikesekiirgus)	24 tundi ette	GRU, LSTM, RNN, SNN, ARIMA	Võimekas mudel, kus SFS tunnuste valikuks ja RF prognoosimiseks. Tunnuste valik aastaaja-põhiselt täpsustas prognoose.
Wang <i>et al.</i> (2018)	hoone elektritarbimine	aeg ja kuupäev, ilmaandmed (temp., niiskus, õhurõhk, sademed, tuule kiirus, päikesekiirgus), elanike arv	24 tundi ette	RF, RT, SVR	RF annab täpsemad prognoosid kui SVR ja RT. RF mudeli eeliseks on võimas tunnuste olulisuse funktsioon.
Zahid <i>et al.</i> (2019)	elektrihind, elektritarbimine	<i>Esitletud vaid tunnused elektrihinna prognoosimiseks*</i>	24 tundi ette	XGB, DT, RFE, RF, ECNN, ESVR	CNN ja SVR mudelid sarnase prognoosivõimekusega. Mudeli täpsus kõrge nii väikse kui suure testandmestiku valimiga.
Fan <i>et al.</i> (2022)	elektritarbimine	aeg ja kuupäev, piirkond, ilmaandmed, elektrinõudlus	1 tund ette	SVR-GC-RF hübriidmudel	Kõrge prognoosivõimekusega SVR-GC-RF mudel

Allikad: (Wolff *et al.*, 2013; Li *et al.*, 2014; Ramsami & Oree, 2015; Dolara *et al.*, 2015; Das *et al.*, 2018; Wang *et al.*, 2018; Zahid *et al.*, 2019; Fan *et al.*, 2022; Trivedi *et al.*, 2022); autori koostatud.

**pole käesolevas töös relevant*

Ilm on üks olulisemaid tunnuseid päikeseenergiast elektri tootmise ja elektrinõudluse prognoosimisel. Ilmaandmete olulisust on märgitud kõikides varasemates empiirilistes uuringutes ning kõige olulisemate tunnustena on välja toodud temperatuur, päikeseikiirgus, sademed. Lisaks on valikuliselt kasutatud ka niiskuse, õhurõhu, tuule kiiruse ja suuna tunnuseid. Alles hiljutistes töödes on tunnusega prognoosimudelitesse lisatud ka viiteajaga tootmis- ja tarbimisandmed, varasemates töödes pole neid peaaegu üldse kasutatud. Olenevalt prognoosiperioodi pikkusest (1 tund või 24 tundi ette) on töös kasutatud kas tunnuseid või päevaseid andmeid, kuid kõikidesse töödesse on lisatud erinevaid kuupäeva-tunnuseid sesoonsuste tuvastamiseks, näiteks nädalapäev (mõnes töös eraldatud ka riigipühad) ning kuu aastas või aastaaeg. On leitud, et tunnuste valik aastaaja-põhiselt täpsustas prognoose. Seletavate tunnuste valik on oluline ning hea valikuga võivad tugivektorite regressioon ja otsustusmetsad prognoosida väga sarnaseid tulemusi, ilma statistiliste erinevusteta, (Voyant *et al.*, 2017).

Oluline aspekt, milleni varasemates töödes on korduvalt jõutud, on järeldus, et prognooside täpsus ei sõltu mitte nii väga meetodi valikust, vaid hoopis treeningandmete kvaliteedist (Voyant *et al.*, 2017). Päikeseenergiast elektri tootmise prognoosimist käsitlevates töödes on leitud, et suurem osa prognoosivigadest on seotud ilmaprognooside täpsusega (ehk eksogeensete tunnuste enda kvaliteediga), mitte päikesepaneelide enda tootlikkuse prognoosiga (Wolff *et al.*, 2013; Dolar *et al.*, 2015; Voyant *et al.*, 2017).

Varasemate empiiriliste analüüside põhjal ei saa välja tuua kindlaid prognoosimudeleid, mis annavad kõige täpsemad mudelid. Parima mudeli valik sõltub suuresti töö eesmärgist, hüperparameetrite õigest reguleerimisest ning võimalikult täpsete andmete kasutamisest. Pidades silmas neid aspekte, tehti käesolevas töös otsus kasutada päikeseenergiast elektri tootmise ning elektritarbimise lühiajaliseks (24 tundi ette) prognoosimiseks otsustusmetsa regressiooni ning tugivektorite regressiooni mudeleid. Lisaks agregeeritakse andmed maakondade põhiseks, et vähendada üksikute ebaregulaarse sündmuste (näiteks suurte ilmamuutuste) mõju üksiktarbijate tootmisele ja tarbimisele.

2. ANDMED JA METOODIKA

Töö eesmärgi saavutamiseks analüüsitakse erinevaid arvandmeid ning koostatakse varasemate sarnaste empiiriliste uurimuste põhjal päikeseenergiast elektritootmise ja -tarbimise prognoosimudelid. Analüüsi metoodika sisaldab endas kirjanduse põhjal enim kasutust leidnud mudelite ülevaadet ning võrdlusmodelina kasutatakse AR(1) mudelit. Prognoosimudelitest kasutatakse tugivektorite regressiooni ning otsustusmetsa regressiooni. Vastava metoodika alapeatükkides räägitakse lähemalt mudelite eeldustest, võimalikest probleemkohtadest ning nende lahenduste leidmistest.

2.1. Andmed

Andmete valik sisaldab endas uuritavate ehk prognoositavate tunnuste ning sisend- ehk sõltumatute tunnuste valikut. Sõltuvaid tunnuseid on kaks, mis on käesoleva analüüsi uurimisobjektideks – päevane päikeseenergiast toodetud elektri kogus ning päevane elektritarbimise kogus.

Andmed pärinevad Eesti Energia poolt korraldatud teadusvõistluse „*Enefit – Predict Energy Behavior of Prosumers*“ andmekogust. Kaggle keskkonnas 1. novembril 2023 aastal alanud teadusvõistluse eesmärgiks on luua prognoosimudel päikeseenergiast elektrit tootvate majapidamiste toodetud ja tarbitud elektri kogustele. Andmekogus on juurdepääs reaalsele andmetele ilmastiku, energiahindade ja valitud Eesti Energia klientide installeeritud päikeseпаркide võimsuste kohta. Kõik andmed on antud ühetunnise intervalliga ning ajaperioodiks on september 2021 kuni mai 2023. Kliendid on eraldatud äri- ning eraklientideks ehk ettevõteteks ning majapidamisteks, kus lisaks maakonnapõhisele asukohale on teada ka päikeseпаркide installeeritud võimsused. Klientide kohta on teada iga tunni tootmis- ning tarbimiskogused, mis on käesoleva töö olulisemad andmed prognoosimiseks. Ilmaandmetest on olemas nii ajaloolised andmed kui kuni 48-tundi-ette ilmaennustused temperatuuri, sademete, päikesekiirguse, pilvisuse ning tuule suuna ja tugevuse kohta. Ilmastikuandmed on maakonnapõhised ehk kogutud on vastavalt igas maakonnas asuva ilmajaama andmed. (Eljand *et al.*, 2023) Suur andmemaht annab

võimaluse ühendada kõik erinevatest allikatest päritud andmed omavahel ning viia prognoos läbi võimalikult asjakohaste andmetega.

Andmed jaotatakse treeningandmestikuks, mille põhjal koostatakse mudel, ning testandmestikuks, mille põhjal prognoosimudeli täpsust arvutatakse. Prognoosi täpsuse huvides leitakse analüüsi käigus optimaalsed treeningandmestiku ning testandmestiku perioodid. Töö autor on saanud Eesti Energialt loa Kaggle keskkonnas olevate andmete kasutamiseks oma magistritöös. Järgnevalt on välja toodud analüüsis kasutatavate tunnuste kirjeldavad statistikud.

Tabel 2. Tunnuste kirjeldus

Tunnus	Keskväärtus	Mediaan	Standardhälve	Miinumum	Maksimum
Tarbimine päevas (MWh)	47,89	24,53	72,75	0,42	469,58
Tootmine päevas (MWh)	9,27	2,49	19,12	0,00	231,09
Tarbimine 1 päev varem (MWh)	47,90	24,56	72,75	0,42	469,58
Päeva keskmine temperatuur (°C)	5,39	5,10	8,08	-23,00	26,95
Päeva keskmine pilvisus (%)	68,85	76,13	29,67	0,00	100,00
Otsene päikesekiirgus kogu päeva jooksul (MW/m ²)	3,50	2,52	3,39	0,00	12,36
Sademetede hulk (mm)	1,91	0,42	3,39	-0,02	58,57
Installeeritud võimsus 1 päev varem (MW)	6,36	4,08	7,79	0,32	44,64
Keskmine elektri hind päevas (€/MWh)	157,72	136,64	96,30	11,46	682,05

Allikas: Enefit (2023); autori arvutused.

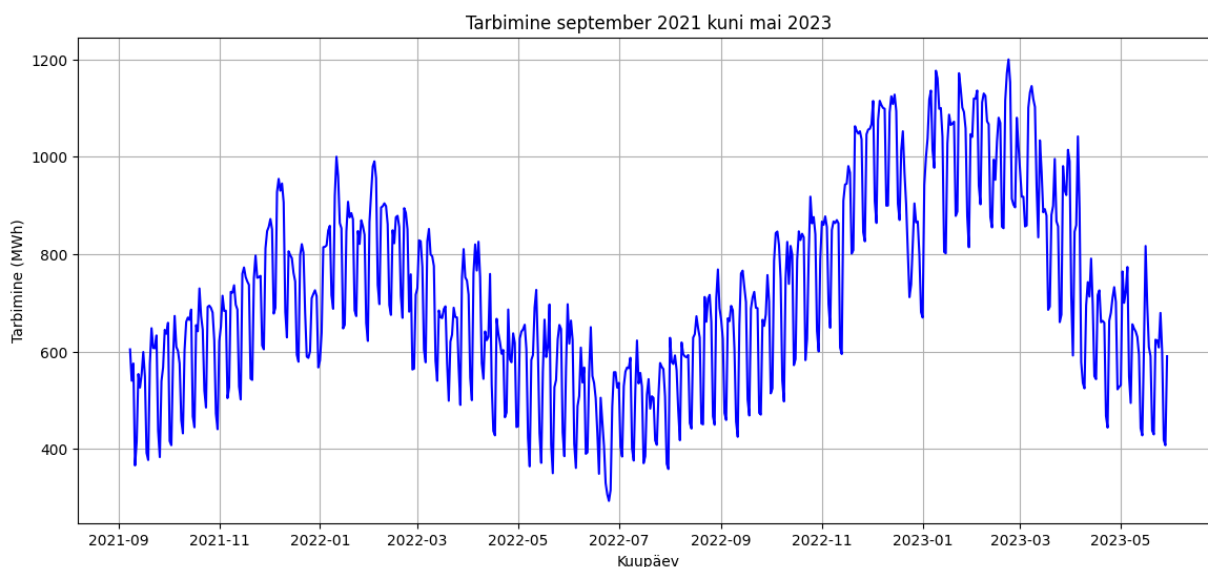
Mittearvulised tunnused:

- Päev nädalas (1-7)
- Kuu aastas (1-12)
- Maakond (1-15)

Kõik tunnused on eraldatud maakondadeks ning iga maakonna kohta on agregeeritud päevasteks andmeteks nelja saadaoleva päikesepaneeli omava kliendi andmed. Kliendid on igas maakonnas jagunenud viisil, et neist kaks on kodumajapidamised ning kaks on ettevõtted. Ilmaennustuse andmed on võetud pikkus- ja laiuskraadide järgi, olles igas maakonnas mõõdetud vastava maakonna ilmapunktis. Kokku on mudelite sisendiks andmed 9435 vaatluse kohta ehk töös on kasutatud 15 maakonna andmeid 629 päeva kohta. Treening- ning testandmete perioodid selguvad modelleerimise käigus.

Tunniste andmete päevasteks agregeerimisel kasutati nii summeerimist kui keskvärtuse arvutamist. Tootmisväärtustel, tarbimisväärtustel, installeeritud võimsusel, päikesekiirguse ning sademete hulga päevaste andmete jaoks on ööpäeva kõik 24 väärtust omavahel liidetud. Keskmise temperatuuri, pilvisuse ning elektrihinna jaoks on võetud ööpäeva keskmine väärtus. Ilmaandmete ning elektrihinna korral on tegemist prognoosiandmetega 24 kuni 47 tundi ette, mida kasutatakse prognoosimudelites.

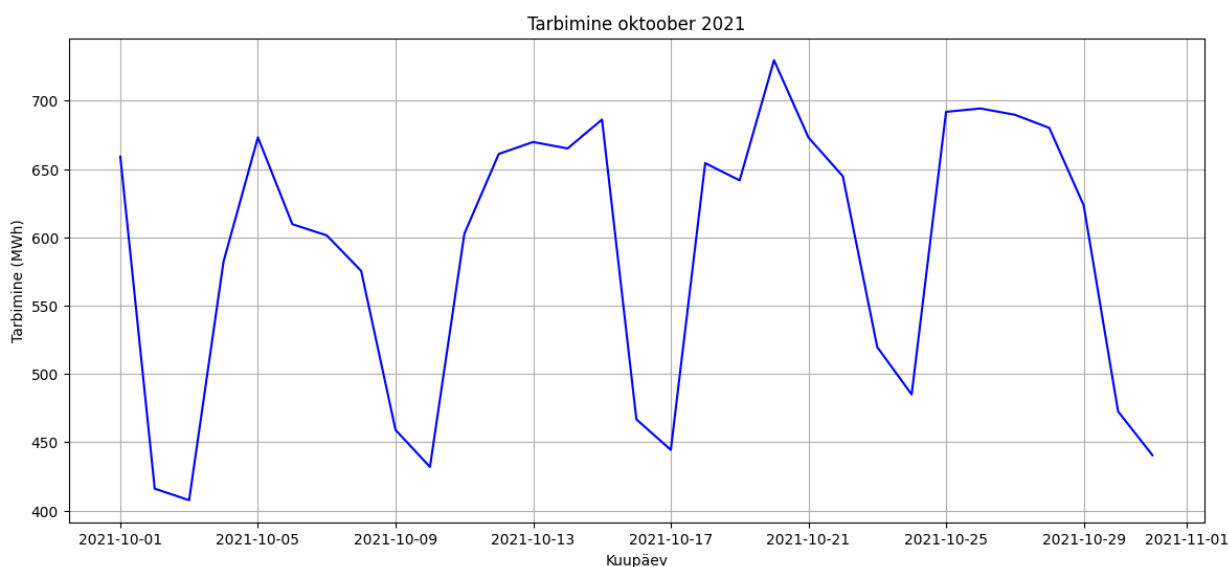
Järgnevalt on kirjeldatud prognoositavaid andmeid (tarbimist ning tootmist) graafiliselt. Joonisel 3 on kirjeldatud kõikide maakondade nelja kliendi kogutarbimist päevade lõikes. Vertikaalteljel on kogu andmestiku agregeeritud tarbimine megawatt-tundides (MWh) ning horisontaalteljel kogu andmeperiood september 2021 kuni mai 2023. Jooniselt paistab välja kaks erinevat sesoonsust. Talveperioodil tarbitakse Eestis märkimisväärselt rohkem elektrit kui suveperioodil. Eristuvad ka pühad, näiteks jõulud, kus mõlemal andmestikku jääval aastal on näha madalamaid energiatarbimise väärtuseid. Suveperioodil tarbitakse elektrit oluliselt vähem: suve- ning talveperioodi elektritarbimise erinevus on kuni kahekordne.



Joonis 3. Agregeeritud tarbimisandmed
Allikas: autori koostatud.

Teine sesoonsus on nädalapõhine ning illustreeritud joonisel 4, kus tööpäevadel tarbitakse rohkem elektrit kui nädalavahetusesti. Kuna andmetes on esindatud ka ettevõtete tarbimisandmed, siis puhkepäevade vähesem energiatarbimine on olulisel määral sellest tingitud. Nädalasisest sesoonsust kirjeldab täpsemalt ka joonis 4, kus on presenteeritud 2021. aasta oktoobrikuu andmed

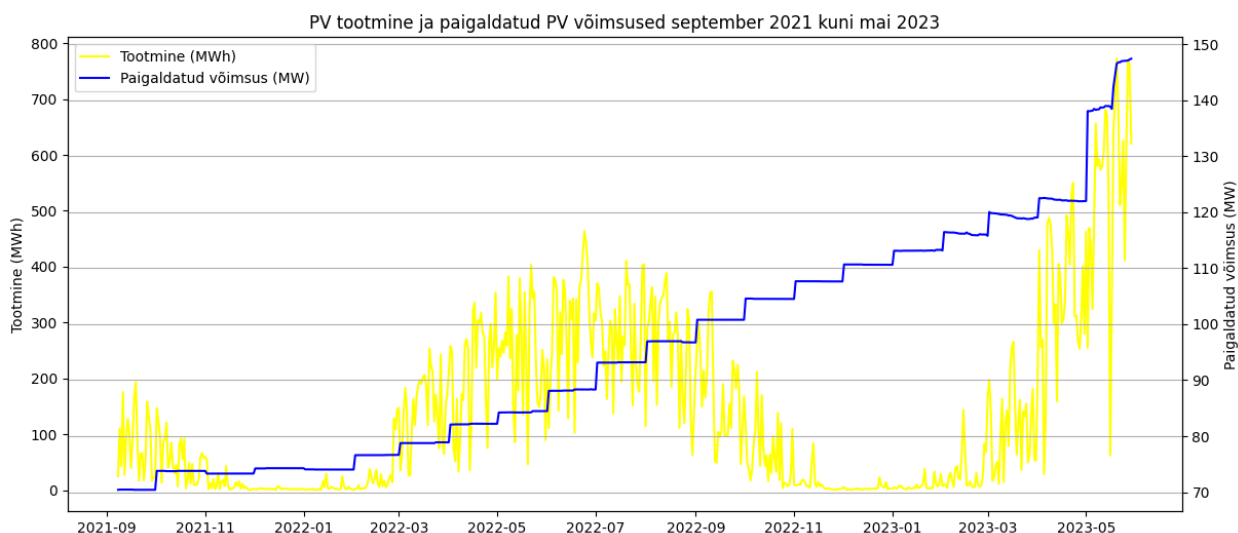
ning selgub, et nädalavahetusesti on agregeeritud tarbimine ligikaudu kolmandiku võrra madalam kui tööpäeviti.



Joonis 4. Agregeeritud tarbimisandmed oktoober 2021

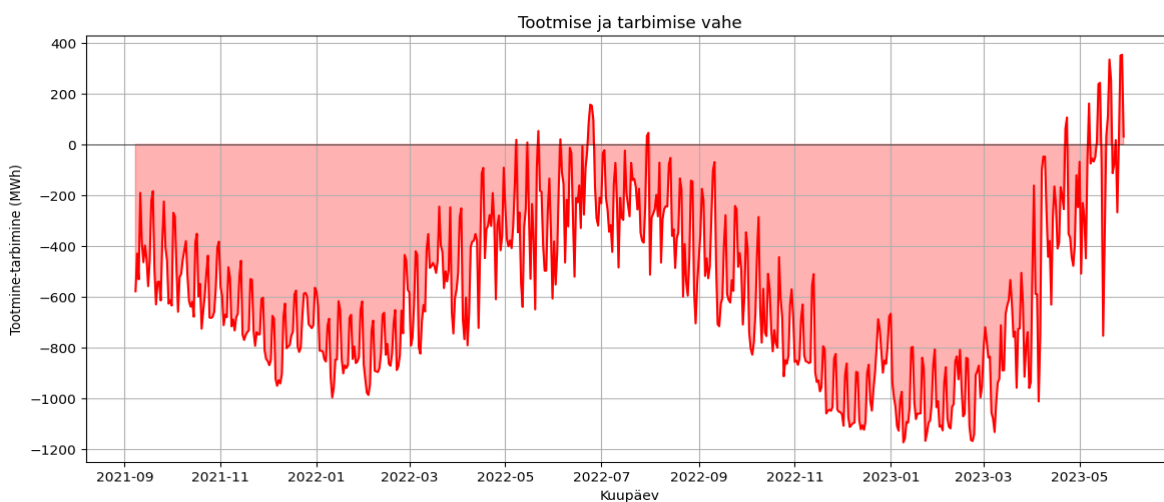
Allikas: autori koostatud.

Tootmisandmed on agregeeritud ja joonisel 5 kujutatud sarnaselt tarbimisandmetele. Vasakpoolsele vertikaalteljele on paigutatud päevased agregeeritud päikeseenergiast elektri tootmise andmed (skaala MWh-des), parempoolsele vertikaalteljele installeeritud päikeseenergia võimsused (skaala MW-des) ning horisontaalteljel on ajaperiood september 2021 kuni mai 2023. Jooniselt paistab koheselt silma päikeseenergia omapära – elektritootmine on talvekuudel madal (kui päikest nii palju ei paista) ning suvekuudel kõrge. Samuti on näha väga suurt toodangu erinevust kahe suveperioodi vahel. 2022. aasta suvel jäid toodangu väärtused vahemikku 100 kuni 400 MWh, kuid järgmisel suvel tõusid need kuni 800 MWh-ni. Sinine joon näitab päikesepaneelide agregeeritud võimsusi, mille põhjal on näha, et kliendid on laiendanud oma päikesepaneelide parke ning installeeritud võimsus kasvas pooleteise aastaga 70 MW-lt 140 MW-le ehk kahekordseks. Graafikult võib veel välja lugeda, et nädala ega muu pikema ajaperioodi lõikes sesoonsust ei esine. Seega võib juba andmete visualiseerimise pealt järeldada, et tootmisandmed ei sõltu nädalapäevast, vaid päikesekiirguse tugevusest ja muudest ilmakomponentidest.



Joonis 5. Agregeeritud tootmisandmed ja paigaldatud PV võimsused
Allikas: autori koostatud.

Kui elektrimajanduse põhieesmärk on hoida elektri tootmine ja tarbimine igal hetkel võrdsed ja tasakaalus, siis päikeseenergiast tootmise ning tarbimise graafikute põhjal peab tõdema, et talveperioodil, kui päikeseenergia tootmismahud on minimaalsed, on elektritarbimine aasta kõrgeim. Lisaks on suveperioodil, kui päikesekiirgus on oluliselt tugevam ning paistab rohkematel tundidel, just tarbimine madalam. Joonis 6 kirjeldab punaselt tarbimise ning tootmise omavahelist seost, kust selgub, et üsna suurel osal aastast kajastab see puudujääki. 2022. aastal ületas päikesetoodang tarbimist ainult üksikutel suvepäevadel ning päikesepaneelide ning taastuenergia suurema levikuga on 2023. aasta aprilli- ja maikuu samuti näha mõnda päeva, kus tootmine ületab tarbimist.



Joonis 6. Tootmise ja tarbimise erinevus
Allikas: autori koostatud.

Päikeseenergiast elektri tootmise ehk pakkumise kasvu ning tarbimise ehk nõudluse vahel on vastuolu. Lisandunud võimsused on riigi elektrimajanduse kindlustamiseks vajalikud, kuid päikesepaneelid toodavad rohkem elektrit just neil perioodidel, kui tarbimine on madal, seega ei aita päikesepaneelide paigaldamise trend tasakaalustada tarbimismahukate perioodide elektrinõudluse vajadust. Kirjeldavas statistikas kirjeldati vaid analüüsi algandmeid, töö empiirilises osas ja prognoosimudelites arvestatakse andmete juures ka nende sesoonsusega.

2.2. Mudeli valik

Prognoosimine on oleviku- ja minevikuandmete analüüsil põhinevate tulevikuprognoside loomine, kus andmed jagunevad kolmeks: sisendmuutujad (mineviku- ja olevikuandmed), prognoosimis-/hinnangumeetodid (trendide analüüs) ja väljundmuutujad (tulevikuennustused). Trendide analüüsimiseks kasutatud tehnikate põhjal võib uuritud meetodid jagada kahte põhitüüpi: eraldiseisvad ja hübriidsed. Eraldiseisvad meetodid kasutavad analüüsimiseks ühte tehnikat, kuid hübriidmeetodid integreerivad suurema prognoositäpsuse huvides mitut eraldiseisvat tehnikat.

Sõltuvalt analüüsitehnika tüübist jagatakse eraldiseisvad meetodid kolme kategooriasse: statistiline, masinõpe (*Computational Intelligence*) ja matemaatiline programmeerimine (*Mathematical Programming*). Statistilistest meetoditest on aegridade analüüsis enim levinud ARIMA ning regressioonmudelid, masinõppes tugivektorite regressioon, tehisnärvivõrkude analüüs ning otsustusmetsa meetodid. (Debnath & Mourshed, 2018; Wolff *et al.*, 2013; Fan *et al.*, 2022; Trivedi *et al.*, 2022) Matemaatilise programmeerimise hulka kuuluvad erinevad dünaamilised programmeerimised on neist kolmest kõige vähemlevinumad meetodid (Fouladgar *et al.*, 2013).

Prognoosimise analüüsimeetodeid saab grupeerida ka sisendinfo põhjal:

1. Hinnanguline prognoosimine – põhineb subjektiivsel hinnangul, intuitsioonil, siseinfo või mõne muu olulise info kasutamisel;
2. Ühemõõtmeline meetod – prognoos ja selle täpsus sõltub ainult prognoositava tunnuse enda oleviku- ja minevikuväärtustest ning omab lineaarset trendi;
3. Mitmemõõtmeline meetod – prognoositava tunnuse väärtus sõltub mitme aegrea väärtustest. (*Ibid.*)

Varasemalt on kasutatud elektritootmise ning -tarbimise prognoosimiseks erinevaid ühe- ja mitmemõõtmelisi meetodeid ning levinud on nii eraldiseisvad mudelid kui hübriidmudelid kahe meetodi parimatest mudelitest (Dolara *et al.*, 2015; Trivedi *et al.*, 2022; Debnath & Mourshed, 2018; Sepasi *et al.*, 2016). Tänapäeval on kõige levinum meetod aegridade tulevikuväärtuste prognoosimiseks masinõppe meetodite kasutamine. Need meetodid õpivad treeningandmeid kasutades ära tundma mustreid andmetes. Taastuvenergia tootmine, mis sõltub suuresti ilmastikust, on tihti ettearvamatu ning vajab prognoosimiseks masinõpet. Taastuvenergia tootmise prognoosimudel nõuab treeningandmeteks ajaloolisi andmeid ilmaennustuse, tegelike elektritootmise koguste ja keskkonna kohta. (Dolara *et al.* 2015; Sepasi *et al.*, 2017)

Põhinedes varasematele empiirilistele analüüsidele taastuvenergia tootmise ja tarbimise prognoosimise kohta, otsustati käesolevas töös kasutada analüüsimeetodiks otsustusmetsa regressiooni meetodi (RF - *Random Forest Regression*) ning tugivektorite regressiooni (SVR – *Support Vector Regression*). Prognoosimudelite võrdlusmudelina kasutatakse AR(1) mudelit, mis on lineaarne mudel ühemõõtmelise aegrea jaoks. AR(1) on lihtsaim autoregressiivse mudeli vorm, mis kasutab aegrea tulevikuväärtuste arvutamiseks vaid ühte tegurit – aegrea minevikuväärtust.

Kuigi tehnikarvorkude analüüsi (ANN) meetod pakub märkimisväärseid võimalusi andmete põhjal keerukate mustrite õppimiseks, pole käesolevas töös seda meetodi kasutatud mitmel põhjusel. Ülesobitamise (ing. k. *overfitting*) oht ja vajadus suure hulga andmete järele süsteemi treenimiseks on peamised põhjused, mis muudavad NN-meetodite rakendamise mõne lühiajalise prognoosimisprobleemi korral keeruliseks (Kavousi-Fard *et al.*, 2014). ANN on nõudlik mudel nii andmete mahu, arvutuskäikude keerukuse kui hüperparameetrite tundlikkuse osas (Li *et al.*, 2014; Zahid *et al.*, 2019).

Otsustusmetsa regressiooni ja tugivektorite regressiooni mudelitel on mitmeid eeliseid. Need annavad hea ülevaate sellest, kuidas sisendfunktsioonid väljundit mõjutavad. Lisaks on need ka arvutuslikult tõhusamad, mistõttu sobivad need mudeli treenimiseks piiratud arvutusressursside ja aja tingimustes. Otsustusmetsa regressioon annab ülevaate tunnuste olulisusest, mis aitab prognoositavate ja sisendmuutujate vahelistest seostest paremini aru saada. (Wolff *et al.*, 2013; Wang *et al.*, 2018; Trivedi *et al.*, 2022)

2.2.1. Mudelite loomine

Tootmismudeli tunnuste esmane valik koostati varasemate uuringute põhjal. Päikeseenergia tootmismahd sõltub peamiselt installeeritud võimsusest ning kui tugev on päiksekiirgus ehk erinevatest ilmaandmetest. Kuna päikesepaiste (eriti parasvöötmes olevas Eestis) on sesoonne ja paistab suveperioodil oluliselt pikemalt, siis oli oluline lisada mitteamvulise tunnusega ka kuupäevaga seotud tunnuseid. Kõik väärtused on päevased agregeeritud andmed maakonniti. Prognoositavaks väärtuseks on päevased tootmisandmed, samuti maakonniti. Mudeli hindamise käigus võidakse tunnuste arvu ja koosseisu muuta.

Algelisse tootmismudelisse lisati seletavate tunnustena: maakond, kuu aastas, päev nädalas, päeva keskmise temperatuuri prognoos maakonnas, päeva keskmise pilvisuse prognoos maakonnas, otsese päikesekiirguse prognoos maakonnas, kogusademete hulga prognoos maakonnas, installeeritud päikesevõimsus maakonnas.

Tarbimismudeli tunnuste esmane valik tehti samuti varasemate uuringute põhjal. Tarbimine sõltub teistsugustest teguritest kui tootmine ning esineb rohkem sesoonsust, seega oli oluline lisada mitteamvulise tunnusega nii nädalapäev, päev kuus kui ka kuu aastas tunnused. Kuna kodumajapidamistes ja ka ettevõtetes on tarbimisharjumused nädalast nädalasse sarnased, siis lisati prognoosimise algsesse mudelisse ka tarbimiskogused 1 ning 7 päeva varem. Tarbimine sõltub ka ilmast, seega lisati mudelisse ka ilmaprognoosi andmeid. Varasemate uuringute põhjal sõltub tarbimine ka elektri hinnast, mis on samuti mudelisse lisatud. Kõik väärtused on päevased agregeeritud andmed maakonniti. Prognoositavaks väärtuseks on päevased tarbimisandmed, samuti maakonniti. Mudeli hindamise käigus võidakse tunnuste arvu ja koosseisu muuta.

Seletavateks tunnusteks on: maakond, kuu aastas, päev nädalas, päeva keskmise temperatuuri prognoos maakonnas, päeva keskmise pilvisuse prognoos maakonnas, otsese päikesekiirguse prognoos maakonnas, kogusademete hulga prognoos maakonnas, installeeritud päikesevõimsus maakonnas, tarbimine maakonnas 1 päev tagasi, päeva keskmine elektri hind.

Järgnevalt on mudelitesse valitud seletavatest tunnustest ülevaatlükuma pildi saamiseks kõik tunnused tabelisse 3 koondatud. Igale tunnusele on antud lühend ning lühend on märgitud vastavasse veergu kui see vastavas mudelis esineb.

Tabel 3. Seletavate tunnuste valik erinevates mudelites ja muutujate lühendid

Tunnused	Tootmine	Tarbimine
maakond	C	C
kuu aastas	D ^M	D ^M
nädalapäev	D ^W	D ^W
temperatuur	W ^T	W ^T
pilvisus	W ^C	
päikesekiirgus	W ^S	W ^S
sademete hulk	W ^P	
installeeritud võimsus	PV	PV
tarbimine 1 päev tagasi		Y-1
elektrihind		E

Allikas: autori koostatud

Kuna prognoosimisel on esmatähtis, kuidas mineviku andmete põhjal tulevikku prognoosida, ei anna juhuslike valimite abil läbi viidud ristvalideerimine õiget hinnangut mudeli toimimisele tulevikuväärtuste prognoosimisel. Tootmise ja tarbimise aegridade andmed ei vasta juhuslikele katsetele, vaid on sageli ajas korreleeritud ning seetõttu tuleb kasutada teistsugust taasvaliku meetodit. Aegrea andmete korral saavad treeningandmestiku moodustada vaid testandmestikule eelnevad andmed ja testandmestikuna ei kasutata perioodi algust. (Sethi, 2024) Käesolevas töös reguleeritakse treening- ja testandmete manuaalselt, leides optimaalse tasakaalu prognoosivea minimeerimiseks.

Erinevate mudelite omavaheliseks võrdluseks peavad aegread vastama kõikide mudelite eeldustele. Peamised nõudmised andmetele panevad AR(1) mudel ning tugivektorite regressioon, kus esimese puhul peavad aegread olema statsionaarsed ning teise puhul standardiseeritud. Aegridade statsionaarsuse testimiseks kasutatakse Dickey-Fuller ühikjuure testi (ADF). Vajadusel muudetakse aegrida statsionaarseks, kasutades transformatsiooni ehk teisendust. Tugivektorite regressioonid eeldavad ka seletavate tunnuste standardiseerimist ehk väärtused peavad olema samas skaalas.

Kogu andmeanalüüs viiakse läbi Google Colabortory platvormil kasutades Pythoni programmeerimiskeelt.

2.3. Referentsmudel AR(1)

AR(1) on autoregressiivse mudeli lihtsaim vorm, mis annab lihtsa, tõlgendatava ja tõhusa lähtepunkti aegridade analüüsiks ja prognoosimiseks. Mudel kasutab aegrea tulevikuväärtuste arvutamiseks vaid ühte tegurit – aegrea minevikuväärtust. AR(1) mudeli koefitsiente on lihtne tõlgendada ning see aitab mõista varasemate vaatluste vahetut seost aegrea praeguse väärtusega. See mudel võimaldab hinnata keerukamate mudelite vajadust - kui masinõppe mudelid ei suuda AR(1) mudeli täpsust ületada, viitab see sellele, et täiendav keerukus ei pruugi vastavate andmete puhul olla õigustatud.

AR(1) mudeli eelduseks on andmete statsionaarsus. Peatükis 2.2. läbi viidud ADF testi põhjal ei ole nii tootmis- kui tarbimisandmed statsionaarsed, seega tuleb sesoonsuse eemaldamiseks täiendada AR(1) mudelit fiktiivsete muutujatega – nädalapäev ning kuu aastas.

Autoregressiivse mudeli AR(1) matemaatiline mudel ja tegurite kirjeldus:

$$Y_t = \phi Y_{t-1} + \beta_1 D_{1,t} + \beta_2 D_{2,t} + \epsilon_t \quad (1)$$

kus

Y_t on prognoositav tunnus ajahetkel t ,

ϕ on autoregressiivne parameeter, mis kirjeldab eelmise väärtuse efekti praegusele,

Y_{t-1} on prognoositava muutuja väärtus ajahetkel $t-1$,

$D_{1,t}$ ja $D_{2,t}$ on fiktiivsed muutujad,

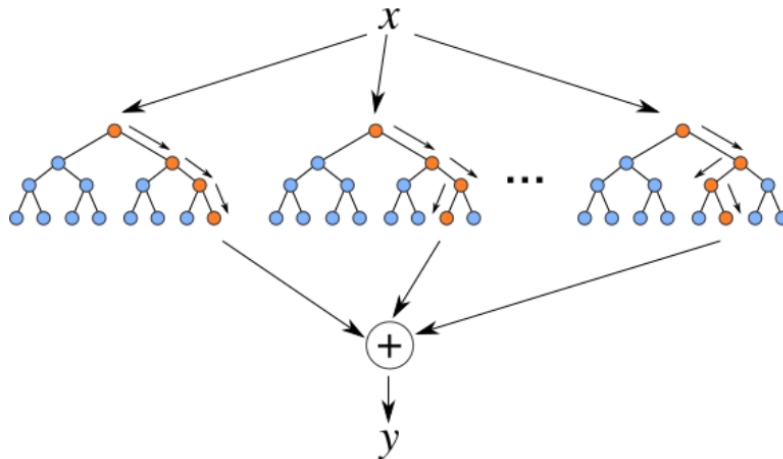
β_1 ja β_2 on fiktiivsete muutujate koefitsiendid,

ϵ_t on vealiige.

Referentsmudeli eesmärk on võrrelda, kas ja kui palju on masinõppe meetodid parema prognoositäpsusega kui lihtne ühemõõtmeline meetod.

2.3. Otsustusmetsa regressioon

Otsustusmetsa (*Random Forest* – RF) regressiooni meetodi kasutatakse eesmärgiga prognoosida seletavate tunnuste komplekti põhjal pideva tunnusega sihtmuutujat. See on meetod, mis loob otsustuspuude kogumi ja ühendab nende prognoosid lõpliku ennustuse tegemiseks.



Joonis 7. Otsustusmets
Allikas: (Stacc, 2022)

Otsustusmetsa regressiooni matemaatiline mudel ja tunnuste kirjeldus:

$$\hat{y}(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N f_i(\mathbf{x}) \quad (2)$$

kus
 $\hat{y}(\mathbf{x})$ on prognoositud väärtus seletavate tunnustele \mathbf{x} ,
 N on otsustuspuude hulk metsas.

Otsustusmetsa regressioon on võimas ja mitmekülgne algoritm, mis sobib hästi paljude prognoositavate modelleerimisülesannete jaoks. Matemaatiliselt võib otsustusmetsa regressiooni mudelit pidada klassifikatsiooni- ja regressioonipuude arenduseks, millel on eeliseks on kõrge ennustustäpsus mittelineaarsete seoste korral, vähene ülesobitamise oht, võimekus andmeid üldistada, suurte andmemahtude arvutamisevõimekus ning tunnuste olulisuse hinnangu andmine, mis aitab mudelit paremini tõlgendada. Sellel meetodil on ka omad puudused, millega tuleb arvestada – mudel tekitab „musta kasti“, mis väljendub keerulisemas tõlgendatavuses võrreldes lihtsamate lineaarsete mudelitega. Teiseks kitsaskohaks on see, et regressioonipuud ei saa luua äärmuslikumaid tulemusi kui need, mis olid treeningandmestikus. (Wang *et al.*, 2018; Fan *et al.*, 2022)

Mudelis esineb juhuslikkust vaatluste ja parameetrite valikul, mille tulemusel väheneb sõltuvus andmete muutumisest, andmetes olevast müra ja eranditest. See meetod ei eelda andmete eeltöötlemist ega parameetrite valikut. Lisaks ei pea muretsema andmete omavahelise korreleerimise pärast, kuna otsustusmetsa mudel suudab toime tulla korreleeritud tunnustega. Juhuslikkuse tõttu ei ole mudelis stabiilseid koefitsiente, mudeli tulemus võib olla igal algoritmi rakendamisel erinev. Otsustusmetsa meetodi kasutamisel on oluline leida kõige optimaalsemad

väärtused hüperparameetritele: puude arv, optimaalne treening- ja testandmete suhe ning miimumarv juhuslikult valitud tunnuseid. (Wolff *et al.*, 2013; Fan *et al.*, 2022)

Otsustusmetsa treenimiseks ning testimiseks prooviti erinevaid tunnuste kombinatsioone ning mõne tunnuse teisendamist pidevast arvulisest väärtusest järjestustunnusteks ja kategoorilisteks tunnusteks. Lisaks katsetati treeningandmete mahu suurendamist ning vähendamist, testimaks, kas see täpsustab prognoositulemusi.

Mudelite prognoositäpsuse võimalikuks parandamiseks kasutatakse ka LASSO meetodi, mis aitab leida üles olulised tunnused ning seeläbi saavutada mudeli parem prognoositäpsus ja tõlgendatavus. LASSO on kasulik suuremõõtmeliste andmekogumite korral, kus tunnused võivad olla korreleeritud või mitteolulised.

2.4. Tugivektorite regressioon

Tugivektorite regressioon (SVR - *Support Vector Regression*) on edasiarendus tuntud tugivektorite masina (SVM - *Support Vector Machine*) klassifitseerimise masinõppe meetodist, mille eesmärk on leida kõige optimaalsem funktsioon seletavate tunnuste ja pideva sihtmuutuja (prognoositava tunnuse) vahel, minimeerides prognoosiviga.

Tugivektorite regressioon kasutab masinõpet, et leida lineaarne hüperatasand N-dimensionaalses ruumis (N on seletavate tunnuste arv), mis sobib kõige paremini pideva ruumi andmepunktidega. SVR kasutab tugivektoreid, mis on loodud treeningandmete punktidest, et määratleda hüperatasand, mis maksimeerib prognoositud väärtuste ümber oleva marginaali. Selle jaoks genereeritakse esialgsest hüperatasandist prognoosivea kaugusel asuvad otsuse piirid nii, et hüperatasandile kõige lähemal asuvad andmepunktid oleksid sellel piirijoonel.

SVR võimaldab käsitleda mittelineaarseid seoseid seletavate ja prognoositavate tunnuste vahel tuumafunktsiooni abil, mis muudab mudeli võimsaks tööriistaks keeruliste seostega regressiooniülesannete jaoks.

Järgnevalt on esitatud tugivektorite regressiooni matemaatiline mudel ja tunnuste kirjeldus. Treeningandmetega $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$, kus \mathbf{x}_i on treeningandmete seletavate tunnuste vektor ja y_i on prognoositav väärtus, lahendab SVR optimeerimisprobleemi:

$$\min \quad \frac{1}{2} \|\boldsymbol{\omega}\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \quad (3)$$

otsustuspiiride piirangutega:

$$\begin{cases} y_i - \boldsymbol{\omega} * \varphi(\mathbf{x}_i) - b \leq \varepsilon + \xi_i \\ \boldsymbol{\omega} * \varphi(\mathbf{x}_i) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (4)$$

ja otsustusfunktsiooniga:

$$f(\mathbf{x}) = \sum_{i=1}^N (a_i - \alpha_i^*) k(\mathbf{x}_i, \mathbf{x}) + b \quad (5)$$

kus

N on seletavate tunnuste arv,

$\varphi(\mathbf{x}_i)$ on mittelineaarne kaardistamisfunktsioon,

C on regularisatsiooniparameeter,

k on tuumafunktsioon,

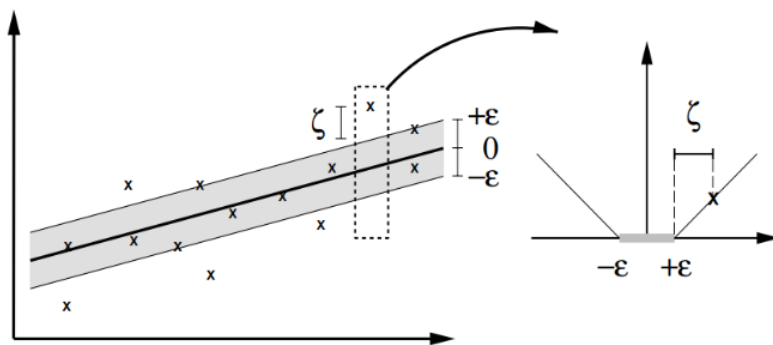
$\boldsymbol{\omega}$ on kaaluvektor,

b on vabaliige,

ε on vealiige,

ξ_i, ξ_i^* on epsilonist suuremaid prognoosivigasid iseloomustavad muutujad.

Uusi väärtusi prognoositakse viisil, et sisendandmete punktid, mis jäävad otsustuspiiridesse ja mille veamäär on madalaim või mis jäävad lubatud piiridesse, annavad parima mudeli ning parim sobivusjoon on hüpertasand, millel on maksimaalne arv punkte. Otsuse piirid on jooned, mis tõmmatakse hüpertasandist kaugusele $-\varepsilon$ ja $+\varepsilon$. Graafiliselt on SVR arvutusmeetod kujutatud joonisel 8. (Sethi, 2024)



Joonis 8. Ühemõõtmelise lineaarse tugivektorite regressiooni piirkonna kujunemine graafiliselt
Allikas: (Sethi, 2024)

Üheks enimkasutatavaks tuumafunktsiooniks on polünoomiaalne tuumafunktsioon. Lisaks on levinud ka Gaussi tuumafunktsioon ning Sigmoidi tuumafunktsioon (Sethi, 2024). Polünoomiaalne tuumafunktsioon, kus otsitavatest parameetriteks on d ehk polünoomi aste, konstant c ning tõus α esitatakse järgmiselt:

$$k(x, y) = (\alpha \langle x, y \rangle + c)^d, d = 1, 2, \dots \quad (6)$$

Tulevikuväärtuste prognoosimiseks arvutatakse mudel treeningandmete abil. Mudeli täpsuse parandamiseks saab kohandada erinevaid hüperparameetreid – tuumafunktsiooni tüüp, regularisatsiooniparameeter ja tuumafunktsiooni-spetsiifilised parameetrid. Pärast SVR-i mudeli treenimist saab seda kasutada testandmetega prognoosimiseks. Mudel ennustab vastavaid prognoositavate tunnuste väärtusi, milleks käesolevas töös on päevased tootmiskogused ning tarbimiskogused maakonniti. Prognooside täpsuse hindamiseks võrreldakse prognoositud väärtusi testimisandmete tegelike väärtustega ning leitakse erinevad mudeli võimekuse hindamise mõõdikute väärtused. (*Ibid.*)

2.5. Mudelite prognoositäpsuse võrdlemine

Analüüsi eesmärk on leida parim meetod päikeseenergiat elektri tootmise ning elektri tarbimise hinnangu loomiseks minimaalse prognoosiveaga. Mudelite prognoositäpsust on varasemates empiirilistes töodes hinnatud erinevate kriteeriumite abil, käesolevas töös hinnatakse mudeleid kahe peamise statistiku abil:

1. RMSE ehk keskmine ruutviga (*root mean squared error*):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

2. nRMSE ehk normaliseeritud keskmine ruutviga (*normalized root mean squared error*):

$$nRMSE = \frac{RMSE}{y_{max} - y_{min}} \quad (8)$$

kus y_i tähistab toodangu või tarbimise tegelikku väärtust ja \hat{y}_i on toodangu või tarbimise prognoos ajahetkel i . nRMSE mõõdab mudeli prognoosiviga võrreldes prognoositava tunnuse vahemikuga, mis võib olla kasulik mudeli võimekuse võrdlemiseks erinevate andmekogumitega.

Varasemates empiirilistes uuringutes on prognoosiviga leitud RMSE, absoluutse vea (MAE), ning nRMSE veastatistikuid. Esimese kahe näitaja puuduseks on mõjutatavus erinditest ehk üksikutest suurtest prognoosimisvigadest. (Wolff et al, 2013; Debnath & Mourshed, 2018; Trivedi *et al.*, 2022) Samas on nRMSE ainus näitaja, mida saab kasutada erinevate andmestike ja meetodidega tehtud prognooside võrdlemiseks, kuna see on normaliseeritud suurus keskmise ruutvea kohta.

3. TULEMUSED JA JÄRELDUSED

3.1. Mudelite hindamise tulemused

Referentsmudel AR(1) tootmismudeli puhul saadi RMSE tulemuseks 25,38 MWh, mis näitab, et keskmine prognoosiviga on ligikaudu 11%. Tarbimismudeli puhul on keskmine prognoosiviga pea kahekordne – 19,4%. See näitab, et lihtne autoregressiivne mudel suudab eelmise väärtuse põhjal prognoosida järgmise ajahetke väärtust tootmise puhul oluliselt täpsemalt kui tarbimise puhul. Autoregressiivsel mudeli puhul on oluline ka eelduste täitmine. Dickey-Fulleri ühikjuure test viidi läbi nii tootmisandmete kui tarbimisandmete põhjal ning eraldi kõikide maakondade lõikes. Andmetes leiti osalist positiivset autokorrelatsiooni, mille põhjal võib järeldada, et andmetes esineb sesoonsust ning hooajalisi trende. ADF-i testi tulemuste põhjal ei saa ümber lükata nullhüpoteesi ehk nii tootmis- kui tarbimisandmetes esineb mittestatsionaarsust. Samuti ei saa väita, et AR(1) mudel oleks käesolevas töös kasutatud andmete puhul parim, kuna üks viiteaeg ei suuda kogu autokorrelatsiooni eemaldada.

Optimaalseima treening- ja testperioodi vahemike leidmiseks prooviti manuaalselt erinevaid kombinatsioone. Paralleelselt arvutati AR(1), otsustusmetsa ja tugivektorite regressiooni mudeleid ning kõikide mudelite põhjal selgus, et mudel suudab kõige täpsemini prognoosida testandmete väärtusi, kui treeningandmed moodustavad 70% andmestikust ning testandmed ülejäänud 30%, mis teeb treeningperioodiks 8. september 2021 kuni 22. november 2023 ning testandmete perioodiks 23. november kuni 31. mai 2023.

Tabelisse 4 on kogutud referentsmudeli ning kahe erineva prognoosimudeli parimad tulemused nii tootmise kui tarbimise andmete kohta. Kõikide mudelite hindamisel on kasutatud sama andmeperioodi ning masinõppe meetodide puhul ka samu selgitavaid tunnuseid. RMSE väärtused on MWh-tes.

Tabel 4. Prognoosimeetodite parimad tulemused testperioodil

	Tootmine		Tarbimine	
	RMSE	nRMSE	RMSE	nRMSE
AR(1)	25,38	10,98%	90,70	19,36%
RF	12,04	5,22%	17,23	3,79%
SVR	9,12	3,94%	13,94	3,05%

Allikas: autori arvutused

Parimaks prognoosimudeliks on väiksema ruutveaga mudel, mis nii tootmise kui tarbimise prognoosimisel on tugivektorite regressioon. Siiski ei jää otsustusmetsa regressiooni prognoositäpsus palju parimale tulemusele alla. AR(1) mudelist on mõlemad masinõppe meetodid oluliselt võimekamad.

Tootmismudeli puhul on tugivektorite regressiooni normaliseeritud ruutvea väärtuseks testandmete puhul 3,94% ning tarbimismudeli puhul 3,04%. Seega võib öelda, et masinõppe suutis tarbimisväärtusi prognoosida täpsemalt kui tootmisandmeid, mis ühtib ka varasemate empiiriliste uuringutega. Selle põhjuseks on toodud tootmisandmete suuremat ebaregulaarsust ning tarbimise puhul on oluliseks aspektiks käitumislik tarbimine, mis ei erine päevade, nädalate ja aastate lõikes nii marginaalselt. Üldiselt on mõlemad mudelid väga võimekad ning võrdluses varasemate empiiriliste uuringutega on nRMSE väärtused madalamad. Mõlema masinõppe meetodi tulemusi kirjeldatakse täpsemalt järgmistes alapeatükkides, kus kirjeldatakse hüperparameetrite reguleerimist ning parima mudeli leidmist.

3.1.1. Otsustusmetsa regressiooni tulemused

Otsustusmetsa (RF) meetodi minimaalse RMSE leidmiseks leiti kõige optimaalsemad väärtused hüperparameetritele: puude arv, optimaalne treening- ja testandmete suhe ning LASSO abil parim kombinatsioon seletavatest tunnustest. Tabelis 5 on välja toodud tulemused tootmise ja tarbimise mudelite parimate hüperparameetrite kohta.

Tabel 5. Otsustusmetsa hüperparameetrite valik

Tootmine			tarbimine		
Hüperparameeter	RMSE	hinnang	hüperparameeter	RMSE	hinnang
n=50	12,04	parim	n=50	17,37	kehvem
n=100	12,10	kehvem	n=100	17,23	parim
n=200	12,10	kehvem	N=200	17,34	kehvem
LASSO	13,47	kehvem, tunnused: C, D ^M , D ^W , W ^T , W ^C , W ^S , W ^P	LASSO	17,90	kehvem, tunnused: D ^W , W ^S , Y-1, E

Allikas: autori arvutused

Hüperparameetrite hindamiseks viidi esialgu läbi kõige optimaalsema puude arvu leidmine. Algmudelis oli puude arv valitud automaatselt n=100 ning n=50 ja n=200 testimine näitas, et tootmismudeli puhul on parimaks n=50 ning tarbimisel n=100. RMSE erinevused polnud küll märkimisväärselt erinevad, kuid puude arvu vähendamine lühendab oluliselt mudeli arvutusaega.

Otsustusmetsa mudeli võimalikuks parendamiseks viidi läbi LASSO meetod, mis aitab leida üles olulised tunnused ning parandada mudeli prognoosivõimet. LASSO tulemused olid aga kehvemad kui esialgne mudel, andes keskmiseks veaks tootmisel 13,47 MWh ja tarbimisel 17,90 MWh ning pidades olulisteks tunnusteks vähem tunnuseid kui alguses mudelis. Tootmismudeli puhul ei pidanud LASSO oluliseks installeeritud päikesepaneelide võimsust ning tarbimise puhul maakonda, nädalapäeva ja installeeritud päikesepaneelide võimsust. Kuna LASSO tulemused olid mõlema mudeli puhul kehvemad kui esialgne mudel, jäädigi parima mudeli valikul siiski esialgsete tegurite juurde.

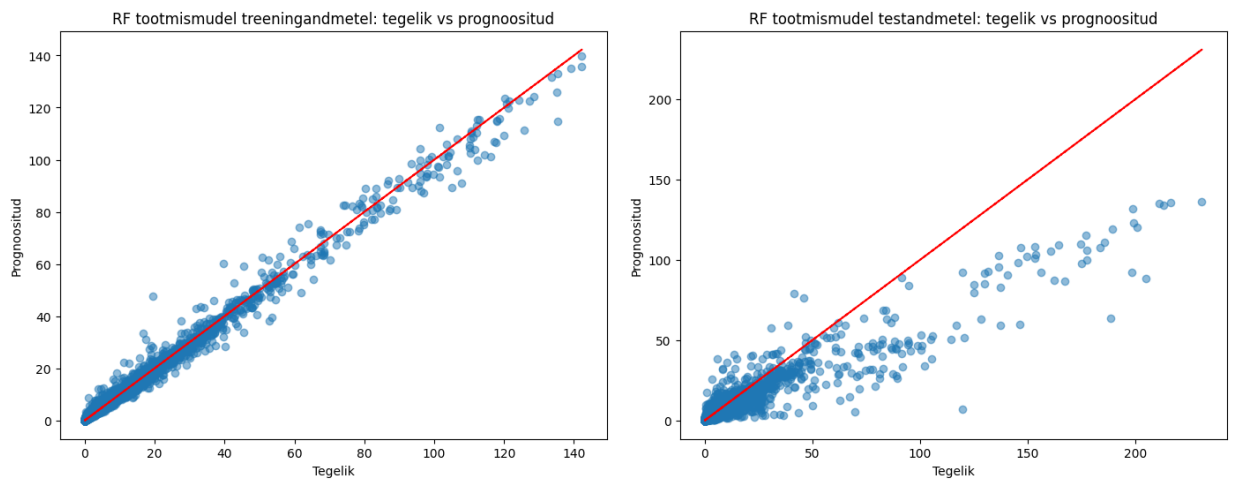
Võimaliku üle- või alasobitamise avastamiseks arvutati parimate mudelite prognoosistatistikud nii treening- kui testandmestike kohta. Suured erinevused treening- ja testandmete prognoositäpsuste vahel võivad viidata ülesobitamisele, kus mudel jätab treeningandmed meelde, mitte ei õpi üldisi mustreid. Tabelis 6 on parimate mudelite treening- ja testandmete statistikud, mille põhjal leiti võimalik ülesobitamine tootmismudeli puhul, kus treeningandmete mudeli võimekus on väga kõrge, üle 99%, kuid testandmete puhul on prognoositäpsuseks vaid 77%. Tarbimismudeli puhul üle- ega alasobitamist ei ole.

Tabel 6. Parimate RF mudelite treening- ning testandmestiku statistikud

Statistik	tootmine		tarbimine	
	treening	test	treening	test
RMSE	1,54	12,04	2,26	17,23
R ²	99,05%	77,08%	99,87%	96,29%

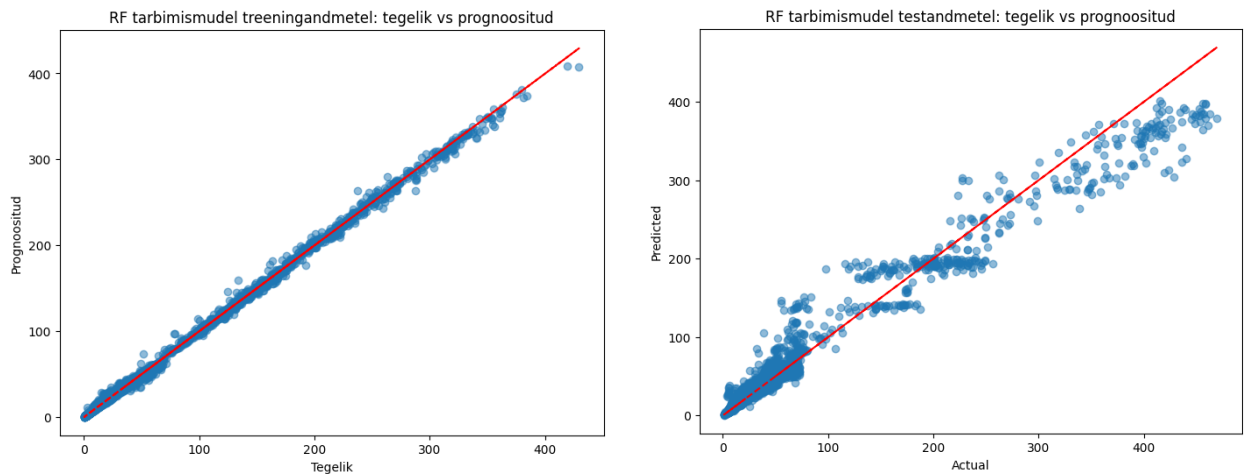
Allikas: autori koostatud.

Võrreldes treening- ning testandmete prognoositulemusi võrreldes tegelike väärtusega, on näha, et treeningandmetel on prognoositulemused väga lähedal reaaltulemusele, kuid testandmete puhul on alates 30 MWh väärtuste juures näha tugevat alaprognoosimist.



Joonis 9. Otsustumetsa (RF) tootmismudeli tegelikud vs prognoositud väärtused
Allikas: autori koostatud.

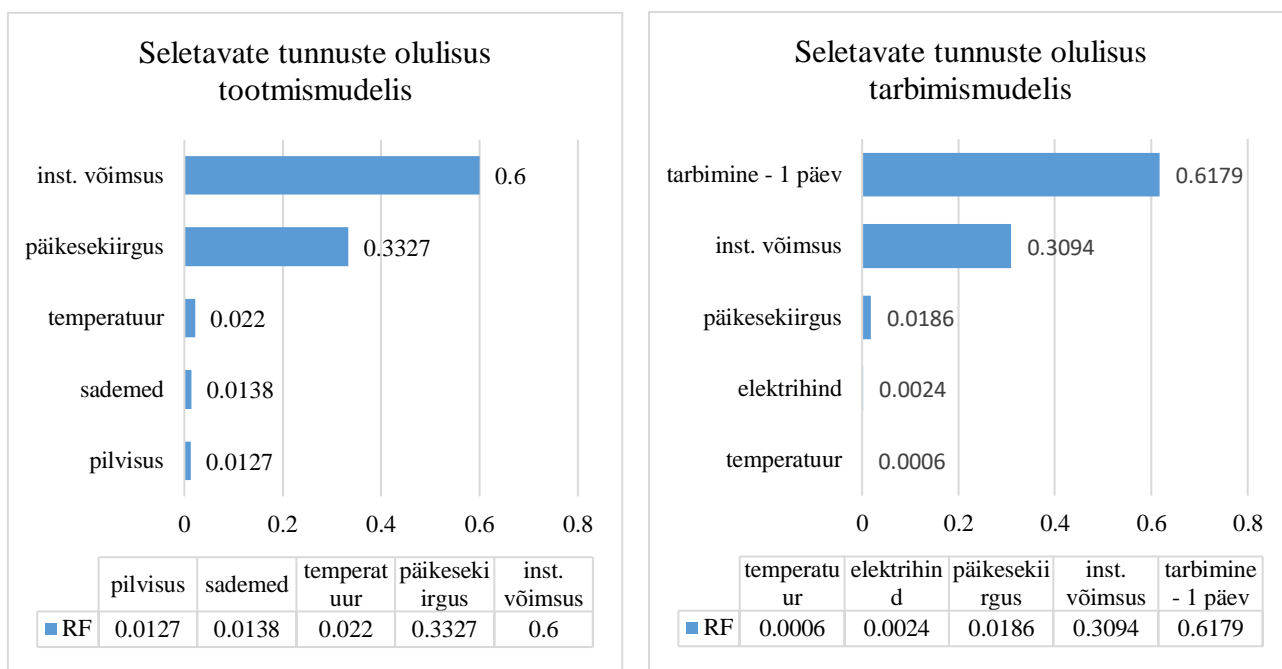
Tarbimismudeli puhul on sarnaselt tootmismudeliga näha treeningandmete väga head prognoositäpsust, kus kõik väärtused on reaallähedased. Testandmete puhul on näha nii üle- kui alaprognosimist. Kõrgemate väärtuste juures (üle 300MWh) prognoosib mudel pigem tegelikkusest madalamaid väärtusi.



Joonis 10. Otsustumetsa (RF) tarbimismudeli tegelikud vs prognoositud väärtused
Allikas: autori koostatud

Otsustumetsa mudeli tugevuseks on varasemates empiirilistes uuringutes välja toodud seletavate tunnuste olulisuse funktsiooni. Siiski on see mudelispetsiifiline ning sõltub peamiselt prognoosi aluseks oleva algoritmi omadustest. Otsustumetsa mudelites tähistavad tunnuste olulisuse osakaalud iga tunnuse panust mudeli prognoosi võimekusse. Mõlema mudeli tunnuste olulisused on normaliseeritud nii, et nende summa on 1 (ehk 100%) kõigi tegurite lõikes, mis muudab mudeli

tõlgendamise lihtsamaks. Sarnane funktsioon puudub tugivektorite regressiooni meetodi puhul, mistõttu ei saa kahe erineva prognoosimudeli põhjal tunnuste olulisust võrrelda.



Joonis 11. Seletavate tunnuste olulisused tootmis- ja tarbimismudelites
Allikas: autori koostatud

Tootismudeli puhul on suurim panus mudelisse installeeritud võimsusel – 60%. Koos otsese päikese kiiruse prognoosiga kirjeldavad need kaks selgitavat tegurit ligikaudu 93% kogu mudeli väärtustest. Prognoosiväärtused temperatuuri, sademete ja pilvisuse kohta panustavad mudelisse 1,2 kuni 2%. Tootismudeli puhul tasub märkimist, et 40% mudeli tegurite olulisusest moodutavad tegurid, mis on ilmaprognoosid, millest võib järeldada, et tootismudeli puhul on väga oluline kvaliteetsete ilmaprognooside olemasolu. Märkimist väärrib ka see, et otsustusmetsa mudelis tunnuste olulisuse arvutamisel ei osutunud elektrihind oluliseks tunnuseks tarbimise prognoosimisel. Selle põhjuseks võib olla asjaolu, et elektrihind ei mõjuta tarbimist agregeeritud tunnuseks. Tarbijad reguleerivad tarbimist päevasiseselt, kuid ei lükka kalli energiahinna pärast tarbimist järgmisele päevale.

3.1.2. Tugivektorite regressiooni tulemused

Tugivektorite regressioonimudelil on mitu hüperparameetrit, mis võivad mudeli võimekust märkimisväärselt mõjutada – tuumafunktsiooni kuju, regularisatsiooniparameeter (c-väärtus) ja tuumafunktsiooni parameetrid (nt polünoomiaalse puhul selle aste). Parameetrite õige regulatsioon on kriitiliselt oluline täpsete prognooside arvutamiseks ning mudeli õigeks tõlgendamiseks.

Tabelis 12 on kirjeldatud hüperparameetrite optimeerimise protsessi tulemusi, kus on väikseimate keskmiste ruutvigade järgi leitud parimad mudelid tootmisandmete ning tarbimisandmete prognooside jaoks.

Tabel 7. SVR hüperparameetrite valik

Tootmine			tarbimine		
hüperparameeter	RMSE	hinnaang	hüperparameeter	RMSE	hinnaang
tuumaf. polü, c=10, d=3	9,12	parim	polü, c=10, d=3	22,48	kehvem
polü, c=50	9,18	arvutus 5x pikem	polü, c=50	18,68	kehvem
polü, c=1	12,18	kehvem	polü, c=100	26,66	arvutus 4x pikem
tuumaf. rbf	19,61	kehvem	tuumaf. rbf	95,39	kehvem
tuumaf. lineaar	17,89	kehvem	tuumaf. lineaar	18,62	kehvem
d=2, c=10	11,36	kehvem	d=2, c=50	13,94	parim, treening kehvem kui d=3
d=4, c=10	10,89	kehvem	d=4, c=50	21,60	kehvem

Allikas: autori arvutused

Parimaks tuumafunktsiooniks sobib treeningandmete põhjal polünoomiaalne tuumafunktsioon. Sellega kaasas käib ka optimaalseima tuumafunktsiooni astme (d) leidmine. Kõrgem funktsiooni aste võimaldab mudelil leida andmetes keerukamaid seoseid, kuid see võib põhjustada ka ülesobitamist. Madalam aste tähendab lihtsamad mudelid. Tavaliselt on see vahemikus 1 kuni 10, sõltuvalt andmete keerukusest. Tootmismudeli jõudlus on optimaalne astme 3 korral, mille puhul oli keskmine ruutviga väiksem. Kõrgema astme puhul oli tegemist ülesobitamise ja madalamate puhul alasobitamisega. Tarbimismudel andis parimaid tulemusi d=2 korral, samas selgus, et treeningandmete keskmine ruutviga oli parem d=3 korral, kuid mudeli arvutused testandmete põhjal andis täpsemad prognoosid d=2 korral.

Ala- ja ülesobitamist aitab hinnata c-väärtus ehk regularisatsiooniparameeter, mis määrab SVR mudelis vektorite marginaali suuruse. Kõrgemad c-väärtused vähendavad reguleeritavust, võimaldades mudelil treeningandmetega paremini sobitada, kuid liiga suure väärtuse korral on oht ülesobitamisele. Väiksemad c-väärtused regulariseerivad treeningandmeid rohkem, soodustades suuremat marginaali. Tavaliselt on see vahemikus 0,01 kuni 1000. Tootmisandmete puhul on kõige optimaalsemaks c-väärtuseks c=10, millest kõrgemate ja madalamate väärtuste puhul on mudeli keskmine ruutviga kõrgem. Tarbimisandmete puhul osutus kõige optimaalsemaks c-väärtuseks c=50. Kõrgema c-väärtuse korral kulub arvutuse läbi viimiseks oluliselt rohkem aega. Kui c=10

korral kulub mudeli arvutamiseks ligikaudu 25 sekundit, siis $c=50$ korral on arvutusaeg 2 minutit ja 20 sekundit.

C-väärtuse suurendamine võimaldab mudelil olla treeningandmete sobitamisel paindlikum, mis võib olla vajalik, kui esialgne mudel oli liiga piiratud. Kõrgema c -väärtusega mudelil on rohkem vabadust andmete sobitamiseks, arvestades rohkem selle keerukusega. Mudel võib olla jätnud osa infost reguleerimise ranguse tõttu kasutamata. Kuigi c -väärtuse suurendamine võib viia testandmete põhjal parema prognoositäpsuseni, on oluline tagada, et paranemist täheldatakse nii treening- kui ka testandmete korral, mis näitab, et mudel tabab paremini andmete aluseks olevaid seoseid ja mustreid ilma ülesobitamisetä. (Sethi, 2024)

Sarnaselt otsustusmetsa tulemuste hindamisel, kontrolliti ka tugivektorite regressiooni mudeli ala- või ülesobitamist eraldi treening- ja testandmete hindamise põhjal. Tulemused tabelis 8 näitavad, et mõlema mudeli puhul on treeningandmete puhul võimekus parem kui testandmete puhul, samas tarbimisel on mudeli täpsuse protsendi erinevus üsna väike. Tootmisandmete puhul on mudelite täpsuses 8-protsendipunktiline erinevus, mis võib esmapilgul viidata ülesobitamisele. Samas kirjeldava statistika joonise 5 järgi suurenes testperioodil installeeritus võimsus koos tootmismahutudega oluliselt, samas kui tarbimismahud jäid testperioodil enamvähem samaks. Üldiselt on mudeli ennustused üsna lähedased tegelikele väärtustele nii treening- kui testandmete puhul ning võrdluses otsustusmetsa meetodiga on tugivektorite regressiooni tulemused oluliselt paremad mõlema mudeli puhul.

Tabel 8. SVR parimate mudelite treening- ning testandmestiku statistikud

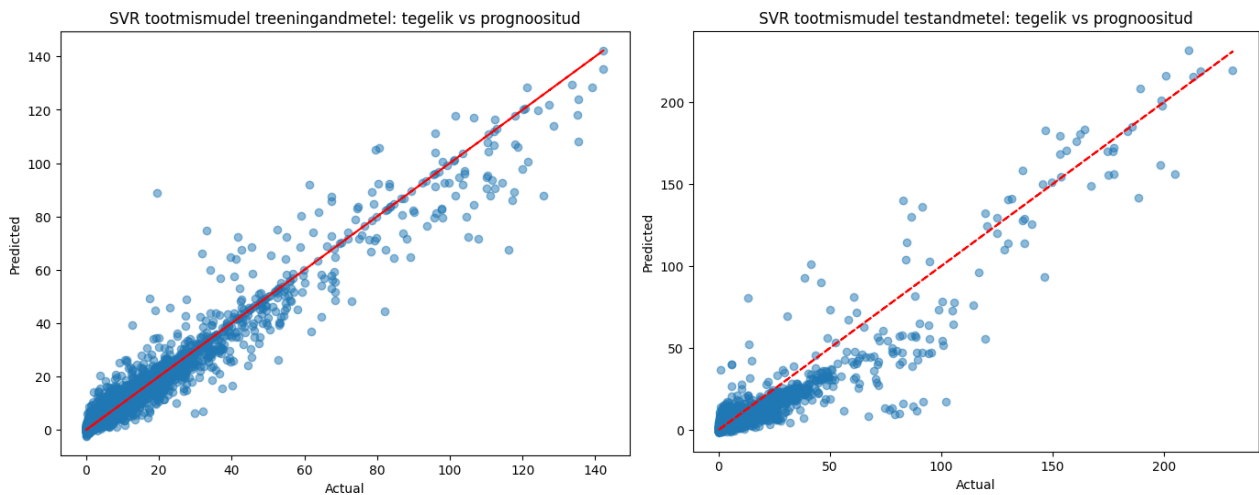
Statistik	Tootmine		tarbimine	
	treening	test	treening	test
RMSE	3,49	9,12	6,15	13,94
R^2	95,06%	86,96%	99,07%	97,57%

Allikas: autori koostatud.

R -ruudu (R^2) skoor mõõdab mudeliga seletatava prognoositava tunnuse dispersiooni osakaalu. R^2 väärtused jäävad vahemikku 0 kuni 1, kõrgemad väärtused näitavad paremat sobivust. Tootmismudeli puhul on treeningandmete R^2 95,06% väga hea tulemus, samas testandmete R^2 86,96% näitab, et mudel suudab testandmestiku puhul tootmisväärtusi prognoosida väiksema täpsusega. Tarbimismudeli puhul on nii treeningandmete R^2 99,07% kui ka testandmete R^2 97,57% väga lähedal ideaalsele tulemusele, mis näitab, et mudel selgitab väga suure osa prognoositava

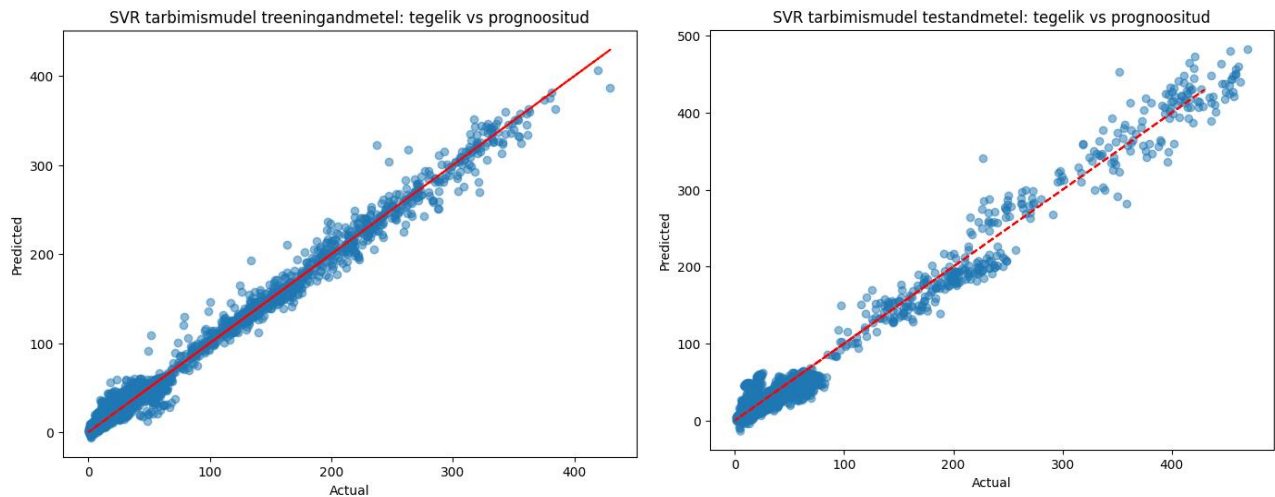
tunnuse dispersioonist mõlemas andmestikus. Treeningu ja testi tulemuslikkuse mõõdikud on suhteliselt lähedased, mis näitab, et mudelis ei esine märkimisväärset ala- ega ülesobitamist.

Võrreldes treening- ning testandmete prognoositulemusi tegelike väärtusega graafiliselt, on näha, et treeningandmete puhul paiknevad prognoosid üsna ühtlase jaotusega ümber ideaalväärtuste. Testandmete puhul paistab silma 70-100 MWh tegelike väärtuse juures üsna tugev alahindamine. Kõrgemate väärtuste juures on prognoosimudel üsna võimekas. Ära märkimist tasub ka see, et treening- ja testandmete vertikaal- ja horisontaaltelgede pikkused on erinevad – kui treeningandmed on kuni 150MWh vahemikus, siis testandmed küündivad peaaegu 250MWh-ni. Kõrgete väärtuste korral on mudeli prognoositäpsus väga hea.



Joonis 12. Tugivektorite regressiooni (SVR) tootmismudeli tegelikud vs prognoositud väärtused
Allikas: autori koostatud

Tarbimismudeli puhul joonisel 13 on treeningandmete prognoosimise graafik väga ideaalväärtuste lähedale kogunenud – see näitab väga relevantset ja täpset mudelit. Testandmete puhul esineb nii üle- kui alaprognoosimist. Vahemikus 170-250 MWh esineb rohkem alaprognoosimist ning 250-330 MWh väärtuste juures esineb rohkem üleprognoosimist. Kuna tulemustes on kõikide maakondade andmed, siis võib see olla ka üksikute maakondade erilisus või andmete ebatäpsus.



Joonis 13. Tugivektorite regressiooni (SVR) tarbimismudeli tegelikud vs prognoositud väärtused
Allikas: autori koostatud

Erinevalt otsustusmetsast ei anna tugivektori regressioon oma olemuselt otsest mõõtmist funktsioonide tähtsuse kohta, nagu otsustuspuud või otsustusmets. SVR keskendub tugivektorite alamhulga alusel andmetele kõige paremini sobiva maksimaalse varuga tugivektori leidmisele. Üksikute tunnuste mõju SVR-i prognoosidele ei ole nii hästi tõlgendatavad kui otsustusmetsa tunnuste olulisused.

3.2. Järeldused ja arutelu

Töös esitletud mudelite võrdluses suutis kõige täpsemaid prognoose arvutada tugivektorite regressiooni meetod. Samas ei jäänud otsustusmetsa regressiooni prognoositäpsus palju parimale tulemusele alla. AR(1) referentsmudelist olid mõlemad masinõppe meetodid oluliselt võimekamad.

Võrreldes omavahel otsustusmetsa ja tugivektorite regressioonide mudeleid, on mõlemal omad tugevused ja nõrkused. Tugivektorite regressioon suutis testandmestiku põhjal prognoosida päevaseid päikeseenergiast toodetud elektri koguseid keskmiselt 4% prognoosiveaga, mis on väga hea tulemus. Sama andmestiku ja samade klientide elektritarbimise põhjal suutis tarbimismudel prognoosida tarbimiskoguseid koguni 3% prognoosiveaga. Mõlemad tulemused on kõrgemad kui varasemates empiirilistes uuringutes leitud keskmised prognoosivead. Selle meetodi ainsaks miinuseks on see, et arvutusviis tekitab „musta kasti“, kus sisendandmetest saadakse

prognoosiväärtused nii, et päris täpselt ei ole aru saada, mis tunnused on mudelis kõige olulisemad ja kuidas üks või teine tunnus prognoosi mõjutab. Seega on üldist prognoosimudelit üsna keeruline pikemalt lahti seletada.

Otsustusmetsa regressioon ei suutnud luua nii võimekaid prognoosimudeleid kui tugivektorite regressioon, kuid samas ei olnud nende kahe erineva meetodi prognoosivead väga märkimisväärselt erinevad – tootmismudeli puhul oli keskmine viga vaid 1,3% ning tarbimismudeli puhul vaid 0,7% kõrgem. Varasemates uuringutes välja toodud otsustusmetsa tugevus – tunnuste olulisuse hindamine oli ka käesolevas töös suureks abiks. Otsustusmetsa meetodi tootmismudeli puhul jäi mudeli nõrkusena silma treeningandmete selgeks õppimine, mitte mustrite loomine, mis viis oluliselt kehvamate prognoositäpsusteni testandmete puhul.

Sarnaselt mitme varasema uuringu tulemustega leiti käesolevas töös, et päikeseenergiast toodetud elektri prognoosimudelis on olulistest seletavateks tunnusteks nädalapäev, aastaaeg, ilmaandmed (temperatuur, pilvisus, päikeseikiirus, sademete hulk) kui ka päikesepaneelide süsteemi installeeritud võimsus. Elektritarbimise prognoosimudelis on olulisteks seletavateks tunnusteks samuti nädalapäev ja hooaeg, ilmaandmetest vaid temperatuur ja päikeseikiirus, kuid lisaks veel ka eelmise päeva tarbimine ja elektri hind. Ka installeeritud päikesepargi võimsus on tarbimismudelis oluline seletav muutuja.

Võrreldes omavahel tootmise ning tarbimise prognoosimudeleid, leiab palju erinevusi. Masinõpe suutis tarbimisväärtusi prognoosida täpsemalt kui tootmisandmeid, mis ühtib ka varasemate empiiriliste uuringutes leitud. Selle põhjuseks on toodud tootmisandmete suuremat ebaregulaarsust ja sõltuvust ilmast ning toodangu üle kontrolli puudumist. Tootmise puhul sõltub prognoosi täpsus palju ka ilmaennustuse täpsusest. Otsustusmetsa tootmismudeli tunnuste olulisuse põhjal seletasid ilmaprognoosid 40% prognoositäpsusest, samas kui tarbimise puhul oli see vaid 2%. Seega on oluline, et ilmaprognoosid oleksid võimalikult täpsed ja relevantid. Tarbimise puhul on oluliseks aspektiks käitumuslik tarbimine, mis ei erine päevade, nädalate ja aastate lõikes nii marginaalselt, mis võis ka põhjustada täpsemad prognoosid. Üldiselt on aga mõlemad mudelid väga võimekad ning võrdluses varasemate empiiriliste uuringutega on nRMSE väärtused madalad.

Tootmismudeli madalam prognoositäpsus puhul kõrgete väärtuste juures võib osaliselt olla põhjustatud installeeritud võimsuste ajas kasvamisest ning nende jätkuvast kasvust pärast

treeningandmete perioodi lõppu. Koos installeeritud võimsustega kasvasid ka tootmismahud ning andmed olid väljaspool treeningandmete väärtusi. Otsustusmetsa üks puudustest on, et regressioonipuud ei saa luua äärmuslikumaid tulemusi kui need, mis olid treeningandmestikus. Samas tugivektorite regressiooni puhul oli tootmismudeli täpsus oluliselt kõrgem, mis näitab, et ühe sisendmuutuja (installeeritud võimsuse) olulise kasvu korral suutis prognoos sellega arvestada ja mudel ei jätnud meelde treeningandmeid vaid suutis luua algoritmi, mida järgida.

Käesoleva töö kõrged prognoositäpsuse väärtused sõltuvad heade andmete ning relevantsete ilmaprognooside olemasolust, põhjalikest hüperparameetrite valikust ning arvutuste tegemisest maakonnapõhiste andmetega. Samuti päevaste agregeeritud väärtuste kasutamisest, mis tasandab ära päevasisesed erandid või välistest tingimustest põhjustatud ebaregulaarsused.

Töös saadud tulemused pakuvad lahendusi elektrimajanduses oluliste tegurite prognoosimiseks. Kuna prognoositud on päevaseid andmeid maakondade lõikes, siis sobivad väljatöötatud prognoosimudelid võrgu töö juhtimiseks ning tulevikus aina kasvavate tootmis- ja tarbimismahtudega arvestamiseks.

Töö edasiarendusena pakub autor välja tulemuste esitamise maakonnapõhiselt, mille kaudu võib avastada erisusi erinevate maakondade tootmis- ja tarbimismudelites. Käesolevas töös oli küll kategoorilise tunnusega maakond lisatud, kuid prognoositäpsuse ja tunnuste olulisuste leidmine maakonnapõhiselt võib anda informatsiooni maakondade erinevustest, mis võimaldab luua veelgi täpsemad mudelid iga maakonna kohta eraldi. Lisaks võimaldavad töös kasutatud andmed luua prognoose ka tunnipõhiselt. See on väga arvutus- ja andmemahukas, kuid teema aktuaalsust ja olulisust hinnates oluline uurimisobjekt.

KOKKUVÕTE

Käesoleva magistritöö eesmärgiks on luua prognoosimudelid, mis lühiajaliselt (päev ette) prognoosivad väiketootjate elektritootmist ning nende elektritarbimist. Töö eesmärgi saavutamiseks püstitati kolm uurimisülesannet:

1. Kirjeldada elektrienergiamaajanduse olemust ning selle olulisust riigi konkurentsivõimele.
2. Analüüsida elektrienergia tootmist ja tarbimist mõjutavaid tegureid, rõhuasetusega päikeseenergia olulisusel.
3. Koostada Eesti andmete põhjal päikeseenergiast elektritootmise ning elektritarbimise lühiajalised prognoosimudelid.

Esimese uurimisülesande lahendamiseks kirjeldatakse erinevate teooriate ning varasemate uurimuste põhjal põhjalikult elektrienergiamaajanduse olemust, energiatõhususe olulist rolli energiamaajanduses ning riigi konkurentsivõime tõusu läbi targa elektritarbimise ja -tootmise. Eesti elektriturg, kus toimiv elektriturg on alati olnud väga olulisel kohal, seisab silmitsi taastuvenergia hoogustumisest põhjustatud probleemidega, mida põhjustavad nii Eestis kui teistes ühisesse elektriturust süsteemi kuuluvates riikides juurde tekkinud suures koguses kontrollimatu iseloomuga elektri tootmise mahud.

Elektri tootmist päikeseenergiast ning elektri tarbimist on varasemalt uuritud ülemaailmselt ning mudeleid nende prognoosimiseks on mitmesuguseid. Erinevad prognoosimudelid vajavad erinevaid sisendeid, kuid üldiselt paistab välja sarnane sisendmuutujate valik. Päikeseenergiast elektri tootmise hulka mõjutavad mitmed tegurid, aga üldiselt jagunevad need kaheks suuremaks valdkonnaks: ilmaandmed ning päikesepargi süsteemiandmed. Elektritarbimist kujundavad lisaks ilmaandmetele ka tarbija elektrikäitumine. Mõlema andmestiku puhul esineb sesoonsust nii nädala, aastaaja kui ka aastate lõikes.

Varasemate empiiriliste uuringute põhjal selgus, et nii päikeseenergiast elektri tootmist kui tarbimist on prognoositud erinevate masinõppe meetoditega, kuid puudub üks kindel parim meetod. Kolmanda uurimisküsimuse lahendamiseks loodi kaks erinevat masinõppe prognoosimeetodi – otsustusmetsa regressiooni ning tugivektorite regressiooni. Prognoosimudelite võrdlusmudelina kasutati AR(1) mudelit ühemõõtmelise aegrea jaoks. Tootmise ning tarbimise prognoosimise jaoks kasutati üksteisest erinevaid siendmuutujaid ning loodi erinevad mudelid.

Parimaks prognoosimudeliks saadi nii tootmise kui tarbimise prognoosimisel tugivektorite regressiooni mudel, millel oli väga kõrge prognoositäpsus. Masinõppe suutis tarbimisväärtusi prognoosida täpsemalt kui tootmisandmeid, mis ühtib ka varasemate empiiriliste uuringutega. Selle põhjuseks on toodud tootmisandmete suuremat ebaregulaarsust ning sõltuvust ilmast. 40-protsendiline ilmategurite olulisus tootmismudelil näitab kõrget ilmaprognooside täpsuse olulisust kogu mudeli prognoositäpsusele. Tarbimise puhul on oluliseks aspektiks käitumuslik tarbimine, mis ei erine päevade, nädalate ja aastate lõikes nii marginaalselt. Üldiselt on mõlemad mudelid väga võimekad ning võrdluses varasemate empiiriliste uuringutega madalamate prognoosivigadega. Kõrged prognoositäpsused aitasid saavutada kvaliteetsed ja relevantseid andmed, täpsed ilmaprognoosid, põhjalik hüperparameetrite valik ning arvutuste tegemine maakonnapõhiste andmetega. Väljatöötatud prognoosimudelid sobivad võrgu töö täpsemaks juhtimiseks ning tulevikus aina kasvavate tootmis- ja tarbimismahtude paremaks arvestamiseks.

Töö edasiarendusena pakub autor välja mudelite võimekuste esitamise maakonnapõhiselt, mille abil võib leida olulist informatsiooni maakondade erinevustest, mis võimaldab luua veelgi täpsemad mudelid iga maakonna kohta eraldi. Lisaks oleks tunnipõhiste prognooside loomine teema aktuaalsust ja olulisust hinnates väga oluline.

SUMMARY

SHORT-TERM FORECASTING OF PHOTOVOLTAIC POWER GENERATION AND CONSUMPTION

Doris Kudre

The aim of this master's thesis is to develop precise models for short-term forecasting of photovoltaic generation and electricity consumption. To achieve this goal, significant features essential for the models are selected and integrated based on findings from previous empirical studies. The chosen appropriate machine learning methods include Support Vector Regression and Random Forest Regression, and models for PV generation and consumption are developed accordingly. The selection of the most optimal forecasting model is determined based on forecast accuracy, specifically focusing on minimizing forecast errors.

Several factors influence photovoltaic generation, notably weather data and the characteristics of PV systems. Electricity consumption is heavily influenced by consumer behavior in addition to weather features. While various machine learning methods have been applied in previous studies to forecast PV generation and consumption, no single method has proven to be universally superior.

The outcomes of this research include Support Vector Regression and Random Forest Regression models with high forecasting performance. In comparative analysis, Support Vector Regression demonstrates more accurate results in both PV generation and consumption models. Machine learning was able to predict consumption values more accurately than generation data, which aligns with previous research findings. The difference can be explained by the natural variability in generation data, which relies heavily on precise weather forecasts. In contrast, consumption forecasting is more influenced by consistent behavioral patterns that show less fluctuation over time.

Overall, the developed models exhibit high accuracy, with lower normalized forecast errors than those reported in previous studies. The achievement can be attributed to the use of high-quality and relevant data, accurate weather forecasts, careful selection of hyperparameters, and analysis incorporating location-specific data. These elements collectively enhance the forecasting capabilities significantly.

KASUTATUD ALLIKATE LOETELU

- Arenguseire Keskus. (2023). Rohepöörde trendid ja stsenaariumid Eestis. Raport. Tallinn.
- Aydin, M. (2019). Renewable and non-renewable electricity consumption–economic growth nexus: Evidence from OECD countries. *Renewable Energy*, 599-606. doi:<https://doi.org/10.1016/j.renene.2019.01.008>.
- Bedir, M., Hasselaar, E., & Itard, L. (2012). Determinants of electricity consumption in Dutch dwellings. *Energy and Buildings*, 58, 194-207. doi:<https://doi.org/10.1016/j.enbuild.2012.10.016>.
- Biggar, D. R., & Hesamzadeh, M. R. (2014). *The Economics of Electricity Markets*. New York: John Wiley & Sons Ltd.
- Carlson, D. R., Matthews, H. S., & Bergés, M. (2013). One size does not fit all: Averaged data on household electricity is inadequate for residential energy policy and decisions. *Energy and Buildings*, 64, 132-144. doi:<https://doi.org/10.1016/j.enbuild.2013.04.005>.
- Commission, E. (2024, aprill 17). European Solar Charter. Retrieved from European Commission: https://energy.ec.europa.eu/topics/renewable-energy/solar-energy/european-solar-charter_en
- Council, E. (2024, aprill 4). Energy efficiency and competitiveness. Retrieved from European Council for an Energy Efficient Economy: <https://www.ecee.org/policy-areas/energy-efficiency-and-competitiveness/>
- Cretì, A., & Fontini, F. (2019). *Economics of Electricity*. Cambridge: Cambridge University Press.
- Das, U. K., Tey, K. S., M., S., Mekhilef, S., Idris, M. Y., Deventer, W. V., . . . Stojcevski, A. (2018, jaanuar). Forecasting of photovoltaic power generation and model optimization: A review. *Renewable and Sustainable Energy Reviews*. .Volume 81, Part 1, 912-928.
- Debnath, K. B., & Mourshed, M. (2018). Forecasting methods in energy planning models. *Renewable and Sustainable Energy Reviews*, 297-325.
- Dolara, A., Leva, S., & Manzolini, G. (2015). Comparison of different physical models for PV power. *Solar Energy*, 83-99.
- Elektrimajandus. (28. aprill 2021). Allikas: Elektrialgud: <https://energiatalgud.ee/Elektrimajandus>

- Elering. (17. 02 2023. a.). Taastuvenergia toetus. Allikas: Elering:
<https://elering.ee/taastuvenergia-toetus>
- Elering. (2024). Elektribilanss.
<https://dashboard.elering.ee/et/balance/total?interval=hours&period=search&start=2021-12-31T22:00:00.000Z&end=2022-12-31T21:59:59.000Z&show=graph>
- Eljand, K., Laid, M., Scellier, J.-B., Dane, S., Demkin, M., & Howard, A. (2023). Kaggle. Retrieved from Enefit - Predict Energy Behavior of Prosumers:
<https://kaggle.com/competitions/predict-energy-behavior-of-prosumers>
- Enerdata. (2024). Estonia energy report. <https://www.enerdata.net/estore/country-profiles/estonia.html>
- Energialgud. (2024, aprill 1). Energiatõhusus.
https://energialgud.ee/Peaartikkel%3A_Energiat%C3%B5husus
- European Council. (14. Detsember 2022. a.). Infographic - Energy crisis: Three EU-coordinated measures to cut down bills. Allikas: European Council:
<https://www.consilium.europa.eu/en/infographics/eu-measures-to-cut-down-energy-bills/>
- Faiella, I., & Mistretta, A. (2022). The Net Zero Challenge for Firms' Competitiveness. *Environ Resource Econ* 83, 85-113.
- Fan, G.-F., Zhang, L.-Z., Yu, M., Hong, W.-C., & Dong, S.-Q. (2022). Applications of random forest in multivariable response surface for short-term load forecasting. *Electrical Power and Energy Systems*.
- Fan, G.-F., Yu, M., Dong, S.-Q., Yeh, Y.-H., & Hong, W.-C. (2021, detsember). Forecasting short-term electricity load using hybrid support vector regression with grey catastrophe and random forest modeling. *Utilities Policy*. Volume 73.
- Fikru, G. M., & Gautier, L. (2015). The impact of weather variation on energy consumption in residential houses. *Applied Energy*, 144, 19-30.
doi:<https://doi.org/10.1016/j.apenergy.2015.01.040>.
- Fouladgar, M. M., Yazdani-Chamzini, A., Zavadskas, E. K., & Moini, S. H. (2013). Selecting the optimal renewable energy using multi criteria decision making. *Journal of Business Economics and Management*, 14(5), 957-978.
- Guo, Z., Zhou, K., Zhang, C., Xinhui, L., Chen, W., & Yang, S. (2018). Residential electricity consumption behavior: Influencing factors, related theories and intervention strategies. *Renewable and Sustainable Energy Reviews*, 399-412.
- Howarth, R. B., & Sanstad, A. H. (1995). Discount rates and energy efficiency. *Contemporary Economic Policy*, 13(3), 101-109.

- Huebner, G., Shipworth, D., Hamilton, I., Chalabi, Z., & Oreszczyn, T. (2016). Understanding electricity consumption: A comparative contribution of building factors, socio-demographics, appliances, behaviours and attitudes. *Applied energy*, 177, 692-702.
- IEA. (2020, aprill 8). Energy efficiency and economic stimulus. Retrieved from <https://www.iea.org/articles/energy-efficiency-and-economic-stimulus>
- IEA. (2024). Estonia. Retrieved from IEA: <https://www.iea.org/countries/estonia/electricity#how-is-electricity-used-in-estonia>
- Inglesi-Lotz, R., & Blignaut, J. N. (2011). South Africa's electricity consumption: A sectoral decomposition analysis. *Applied Energy*, 88(12), 4779-4784. <https://sciencedirect.com/science/article/pii/S0306261911004065>
- Ishaq, H. (2023, juuli 10). Time Series Forecasting with Support Vector Regression (SVR). Retrieved from Medium: <https://medium.com/@hassanishaq/time-series-forecasting-with-support-vector-regression-svr-aab08da8cd77>
- Kavousi-Fard, A., Samet, H., & Marzbani, F. (2014, oktoober). A new hybrid Modified Firefly Algorithm and Support Vector Regression model for accurate Short Term Load Forecasting. *Expert Systems with Applications*. Volume 41, 6047-6056.
- Konkurentsiamet. (detsember 2021. a.). Konkurentsiamet. Allikas: Elektrienergia hinnatõusu analüüs 2021: https://www.konkurentsiamet.ee/sites/default/files/Dokumentide-failid/elektrienergia_hinnatõusu_analüüs_2021_dets.pdf
- Li, R., & Leung, G. C. (2021). The relationship between energy prices, economic growth and renewable energy consumption: Evidence from Europe. *Energy Reports*, 1712-1719.
- Li, Y., Su, Y., & Shu, L. (2014). An ARMAX model for forecasting the power output of a grid. *Renewable Energy*, pp. 78-89.
- Mandot, P. (2017, september 9). What is the Significance of C value in Support Vector Machine? Retrieved from Medium: <https://medium.com/@pushkarmandot/what-is-the-significance-of-c-value-in-support-vector-machine-28224e852c5a>
- Maniatis, G. I., & Milonas, N. T. (2022). The impact of wind and solar power generation on the level and volatility of wholesale electricity prices in Greece. *Energy Policy*, 170, 113-143. doi:<https://doi.org/10.1016/j.enpol.2022.113243>
- Martínez, A., Valero, S., Velasco, E., & Senabre, C. (2015, aprill). Trends in electricity prices and their relation to the competitiveness and efficiency in EU countries. *Renewable Energies and Power Quality*.
- Mushafiq, M., Arisar, M., Tariq, H., & Czapp, S. (2023). Energy Efficiency and Economic Policy: Comprehensive Theoretical, Empirical, and Policy Review. *Energies*. doi:<https://doi.org/10.3390/en16052381>

- Oberst, C. A., Schmitz, H., & Madlener, R. (2018). Are prosumer households that much different? Evidence from stated residential energy consumption in Germany. *Ecological Economics*, 158, 101-115.
- Olczak, P., Jaško, P., Kryzia, D., Matuszewska, D., Fyk, M. I., & Dyczko, A. (2021). Analyses of duck curve phenomena potential in polish PV prosumer households' installations. *Energy Reports*, 7, 4609-4622.
- Pablo-Romero, M. d., Pozo-Barajas, R., & Yñiguez, R. (2017). Global changes in residential energy consumption. *Energy Policy*, 342-352.
- Padhee, M., & Pal, A. (2018). Effect of Solar PV Penetration on Residential Energy Consumption Pattern. *North American Power Symposium (NAPS)*, 1-6.
- Ramsami, P., & Oree, V. (2015). A hybrid method for forecasting the energy output of photovoltaic. *Energy Conversion and Management*, 406-413.
- Reihani, E., Sepasi, S., & Ghorbani, R. (2016). Scheduling of price-sensitive residential storage devices and loads with thermal inertia in distribution grid. *Applied energy*, 183, 636-644.
- Romero-Jordan, D., Peñasco, C., & del Rio, P. (2014). Analysing the determinants of household electricity demand in Spain. An econometric study. *Power and Energy Systems*, 63, 950–961. doi:<https://doi.org/10.1016/j.ijepes.2014.06.005>.
- Sanquist, T. F., Orr, H., Shui, B., & Bittner, A. C. (2011). Lifestyle factors in U.S. residential electricity consumption. *Energy Policy*, 42, 354–364. doi:<https://doi.org/10.1016/j.enpol.2011.11.092>
- Sanstad, A. H., Hanemann, W. M., & Auffhammer, M. (2006). End-use Energy Efficiency in a “PostCarbon” California Economy: Policy Issues and Research Frontiers. *Managing greenhouse gas emissions in California*, 6-32.
- Sethi, A. (2024, veebruar 9). Support Vector Regression Tutorial for Machine Learning. Retrieved from Analytics Vidhya: <https://www.analyticsvidhya.com/blog/2020/03/support-vector-regression-tutorial-for-machine-learning/>
- Stacc. (2022). Otsustuspuul põhinevad masinõppemeetodid. <https://stacc.ee/et/otsustuspuul-pohinevad-masinoppemeetodid/>
- Statistikaamet. (7. detsember 2022. a.). Tarbijahinnaindeksi suurimaks mõjutajaks on tõusmas toid. Allikas: Statistikaamet: <https://www.stat.ee/et/uudised/tarbijahinnaindeksi-suurimaks-mojutajaks-tousmas-toit>
- Zachariadis, T., & Pashourtidou, N. (2006). An empirical analysis of electricity consumption in Cyprus. *Energy Economics*, 29, 183-198. doi:<https://doi.org/10.1016/j.eneco.2006.05.002>

- Zahid, M., Ahmed, F., Javaid, N., Abbasi, R., Zainab Kazmi, H., Javaid, A., . . . Ilahi, M. (2019). Electricity Price and Load Forecasting using Enhanced Convolutional Neural Network and Enhanced Support Vector Regression in Smart Grids. . *Electronics*, 8.
- Trivedi, R., Patra, S., & Khadem, S. (2022). A Data-Driven Short-Term PV Generation and Load Forecasting Approach for Microgrid Applications. *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, 911-919.
- Wang, N., & Mogi, G. (2017). Industrial and residential electricity demand dynamics in Japan: How did price and income elasticities evolve from 1989 to 2014? *Energy Policy*, 106, 233-243.
- Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International journal of forecasting*, 30(4), 1030-1081.
- Wolff, B., Lorenz, E., & Kramer, O. (2013). *Statistical Learning for Short-Term Photovoltaic Power Predictions*. Oldenburg: Carl von Ossietzky University, Institute of Physics, Energy and Semiconductor Research Laboratory.
- Wolske, K. S., Gillingham, K. T., & Schultz, P. W. (2020). Peer influence on household energy behaviours. *Nature Energy* 5(3), pp. 202-212.
- Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F., & Fouilloy, A. (2017, mai). Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*. Volume 105, pp. 569-582.
- Özcan, K. M., Gülay, E., & Üçdoğruk, Ş. (2013). Economic and demographic determinants of household energy use in Turkey. *Energy Policy*, 60, 550–557.
doi:<http://dx.doi.org/10.1016/j.enpol.2013.05.046>
- Xie, Q., Ouyang, H., & Gao, X. (2016). Estimation of electricity demand in the residential buildings of China based on household survey data. *International Journal of Hydrogen Energy*, 41, 15879-15886. doi:<http://dx.doi.org/10.1016/j.ijhydene.2016.03.152>
- Yin, I. (2024, jaanuar 17). *Commodities 2024: China's domestic carbon market set for revamp*. Retrieved from S&P Global Commodity Insights:
<https://www.spglobal.com/commodityinsights/en/market-insights/latest-news/energy-transition/011724-chinas-domestic-carbon-market-set-for-revamp-in-2024-article-6-in-limbo>

LISAD

Lisa 1. Lihtlitsents

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks¹

Mina Doris Kudre

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Päikeseenergiast elektri tootmise ja tarbimise lühiajaline prognoosimine“, mille juhendajateks on Aleksei Netšunajev ja Heili Hein-Sula,

1.1 reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2 üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. Olen teadlik, et käesoleva lihtlitsentsi punktis 1 nimetatud õigused jäävad alles ka autorile.

3. Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

07.05.2024

¹ Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal vastavalt üliõpilase taotlusele lõputööle juurdepääsupiirangu kehtestamiseks, mis on allkirjastatud teaduskonna dekaani poolt, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil. Kui lõputöö on loonud kaks või enam isikut oma ühise loomingulise tegevusega ning lõputöö kaas- või ühisautor(id) ei ole andnud lõputööd kaitsvale üliõpilasele kindlaksmääratud tähtajaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. jq 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.