

TALLINNA TEHNIKAÜLIKOOL  
Infotehnoloogia teaduskond  
Tarkvarateaduse instituut

Aleksandr Tsõganov 143032 IABB

AEGRIDADE PÕHINE KADUDE  
TUVASTAMINE TARKADE ARVESTITE  
VÕRGUS KASUTADES *WAVELET*  
TEISENDUSI

Bakalaureusetöö

Juhendajad: Tarvo Treier

Kristjan Kuhi

Tallinn 2017

## **Autorideklaratsioon**

Kinnitan, et olen koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

Autor: Aleksandr Tsõganov

22.05.2017

## Annotatsioon

Käesolevas bakalaureusetöös analüüsitakse *wavelet* teisenduste kasutamise võimalusi kadude tuvastamise protsessi parandamise eesmärgil. Valdkond millele töö tulemuste saamisel tuginetakse on seotud tarkade arvestite võrguga elektridomeenis ning lähtub kadude tuvastamisel elektritarbimise aegridadest.

Peamiseks töö eesmärgiks on parandada *wavelet* teisenduste rakendamisega kadude piirkonna visuaalset eristumist aegridadel. Visuaalse kao piirkonna tuvastamise lihtsuse tagamine on aluseks ennustusalgoritmide efektiivseks tööks. Antud töö tulemus on sisendväärtuseks edaspidisele arendusprotsessile, mille lõppeesmärgis on luua automatiseeritud kadude tuvastamise lahendus.

Töö käigus käsitletakse aegridade ja *wavelet*'ite teooriat, tuuakse välja erinevad meetodid tulemuse saavutamiseks ning rakendatakse neid näidisandmete peal. Tulemusena tehakse järeldus, et *wavelet*'ite kasutamine on antud eesmärgi saavutamiseks raskendatud ning tulemus on efektiivne teatud iseloomuga kadude tuvastamise parandamisel.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 55 leheküljel, 7 peatükki, 38 joonist, 4 tabelit.

## **Abstract**

### **Time series based fraud detection in smart meter grid using wavelet transforms**

The objective of this bachelor thesis is to analyse wavelet transform possibilities to improve fraud detection process. The approach of this thesis is based on fraud detection in time series data in electricity domain of smart meter grid.

Main purpose of this thesis is to improve visual fraud area recognition by using wavelet transforms. Visual fraud detection improvements are the basis for prediction model efficiency improvements. The result of this thesis analysis will be used as input value for development of automated fraud detection solution.

Current thesis explores time series options, wavelet theory, usage of wavelet transforms and implementation methods. The methods are applied in experiments where electricity consumption time series data were used as an example. As a conclusion of analysis and experiments the author points out that reaching the main purpose for all example datasets is difficult and the result is effective for detection of specific type of fraud time series.

The thesis is written in Estonian and contains 55 pages of text, 7 chapters, 38 figures, 4 tables.

## Lühendite ja mõistete sõnastik

<i>Wavelet</i>	Matemaatiline funktsioon, mida kasutatakse andmete töötlemise eesmärgil.
Aegrida, signaal	Aja ja väärtuse, antud töö kontekstis numbriliste, paaridest koosneva arvandmete rida.
Andmete visualisatsioon	Andmete ehk töö kontekstis aegridade, signaalide visuaalne kujutamine graafilises formaadis, kasutatakse parema andmete sisu arusaama loomiseks
DWT	<i>Discrete Wavelet Transform</i> – Diskreetne wavelet teisendus, mida kasutatakse andmete filtreerimise, lihtsustamise, dekomponeerimise eesmärgil.
CWT	<i>Continuous Wavelet Transform</i> – Pidev wavelet teisendus, mida kasutatakse andmete detailsemaks analüüsimiseks
MSE	<i>Mean squared error</i> – Statistikas kasutatav meetod aegridade vahelise erinevuste võrdlemiseks.
Aegride ennustamine	Aegridade prognoosimine kasutades eelnevalt õpitud aegridade väärtusi ja omadusi.
ARIMA, SARIMA	<i>Autoregressive integrated moving average</i> – mudel mida kasutatakse aegridade ennustamiseks kasutades automaatset regressiooni. Eesliide 'S' tähistab mudeli alaliiki, mis on loodud sesoonsete signaalide ennustamiseks.
JPEG	Laialt levinud piltide pakkimise formaat, mis teisendab pildi originaali kompaktsemasse vormi, säilitades head kvaliteeti.
Teek	Programmeerimiskeeles kasutatav funktsioonide, klasside ja muu vajaliku kogum, mis aitab lahendada konteksti spetsiifilisi funktsioonide ja tegevuste vajadusi.

# Sisukord

<b>1 Sissejuhatus .....</b>	<b>10</b>
<b>2 Kaod tarkade arvestite võrgus.....</b>	<b>13</b>
2.1 Targad arvestid.....	13
2.2 Kaod.....	14
2.3 Kadude mõju.....	14
2.4 Hetkeolukord Elektrilevi näitel.....	15
2.5 Tausta kokkuvõte .....	17
<b>3 Wavelet'id ja aegread .....</b>	<b>19</b>
3.1 Wavelet .....	19
3.2 Aegread.....	20
3.3 Wavelet'ite teooria ehk teisendused.....	22
3.3.1 Pidev wavelet teisendus.....	23
3.3.2 Diskreetne wavelet teisendus .....	24
<b>4 Lahenduse väljatöötamine .....</b>	<b>27</b>
4.1 Diskreetsete wavelet teisenduste kasutamine .....	28
4.1.1 Ühe taseme teisenduste kasutamise meetod rekonstrueerimiseta .....	29
4.1.2 Mitmetasandilise teisenduse kasutamise meetod rekonstrueerimisega.....	29
4.2 Pideva wavelet teisenduste kasutamine.....	30
<b>5 Juhtumiuring .....</b>	<b>31</b>
5.1 Näidisandmete analüüs .....	31
5.1.1 Näidisandmetes välja toodud kadude analüüs ja valik.....	31
5.1.2 Erinevad lähenemised andmete kuvamisele.....	33
5.2 Rakendatava teegi valik .....	35
5.3 Lahenduste realiseerimine .....	37
5.3.1 Diskreetse wavelet teisenduse meetodid ja visuaalne võrdlemine .....	37
5.3.2 Katsetused rakendades ennustusalgoritmi.....	45
5.3.3 Pideva wavelet teisenduse lahenduse realiseerimine .....	46
<b>6 Tulemus ja diskussioon .....</b>	<b>48</b>
<b>7 Kokkuvõte.....</b>	<b>50</b>
<b>Kasutatud kirjandus.....</b>	<b>52</b>
<b>Lisa 1 – Parimad wavelet funktsioonide ühtimised näidete aegridadega MSE alusel ühe aasta pikkusel perioodil.....</b>	<b>55</b>

## Jooniste loetelu

Joonis 1 Elektrilevi tehniliste ja mittetehniliste kadude osakaal .....	15
Joonis 2 Aegridade erinevus keskmise väärtuse konstantsuse järgi (Srivastava, 2015).	21
Joonis 3 Aegridade erinevus koherentsuse järgi (Srivastava, 2015).....	22
Joonis 4 Aegridade erinevus kovariatsiooni järgi (Srivastava, 2015).....	22
Joonis 5 <i>Wavelet</i> funktsioon 'Mexican hat' .....	24
Joonis 6 Signaal enne CWT teisendust.....	24
Joonis 7 Signaal peale CWT rakendust .....	24
Joonis 5 Mitmetasandiline diskreetse <i>wavelet</i> teisenduse joonis (Ogorodov, 2009).....	25
Joonis 9 <i>Wavelet</i> funktsioon 'Daubechies 2' .....	25
Joonis 10 Diskreetne <i>wavelet</i> teisendus kasutades "Daubechies 2" funktsiooni .....	26
Joonis 11 Ühe taseme teisenduste kasutamise meetodit selgitav joonis.....	29
Joonis 12 Kolme tasandilist <i>wavelet</i> teisenduse kasutamist selgitav joonis .....	30
Joonis 13 Pideva <i>wavelet</i> teisenduse kasutamist iseloomustav skeem .....	30
Joonis 14 Näide 3 - aegread kõikide tunniandmetega eraldi .....	33
Joonis 15 Näide 3 - aegread ainult öötundidega .....	34
Joonis 16 Näide 5 - aegread kõikide tunniandmetega eraldi .....	34
Joonis 17 Näide 5 - aegread tunni täpsusega .....	34
Joonis 18 Näide 3 - aegread kõikide tunniandmetega eraldi .....	35
Joonis 19 Näide 3 - aegread päeviti summeeritud väärtustega .....	35
Joonis 20 Funktsiooni 'rbio3.7' kuju teisendusel (pywavelets, 2008-2017) .....	38
Joonis 21 Funktsiooni 'rbio3.7' kuju rekonstrueerimisel (pywavelets, 2008-2017) .....	38
Joonis 22 Näide 1 – 'rbio3.7' teisendus.....	39
Joonis 23 Funktsiooni 'bior1.1' kuju teisendusel ja rekonstrueerimisel (pywavelets, 2008-2017).....	40
Joonis 24 Näide 2 – 'bior1.1' teisendus.....	40
Joonis 25 Näide 3 – 'db36' teisendus .....	41
Joonis 26 Funktsiooni 'rbio3.3' kuju teisendusel (pywavelets, 2008-2017) .....	42
Joonis 27 Funktsiooni 'rbio3.3' kuju rekonstrueerimisel (pywavelets, 2008-2017) .....	42
Joonis 28 Näide 5 – 'rbio3.3' teisendus.....	42

Joonis 29 Funktsiooni 'rbio3.5' kuju teisendusel (pywavelets, 2008-2017) .....	43
Joonis 30 Funktsiooni 'rbio3.5' kuju teisendusel (pywavelets, 2008-2017) .....	43
Joonis 31 Näide 6 – 'rbio3.5' teisendus.....	44
Joonis 32 Näide 6 – 'db35' teisendus 1. meetodi korral.....	44
Joonis 33 Näide 1 - ennustuse võrdlus lähteandmete ja 3 tasemelise diskreetse <i>wavelet</i> teisenduste vahel .....	46
Joonis 34 Näide 4 lähteandmed 1 aastasel ajavahemikul .....	47
Joonis 35 Näide 4 - 12. tunni näidud 'gaussian1' .....	47
Joonis 36 Näide 4 - 12. tunni näidud 'gaussian2' .....	47
Joonis 37 Näide 4 - 12. tunni näidud 'haar' .....	47
Joonis 38 Näide 4 - 12. tunni näidud 'morlet' .....	47



## Tabelite loetelu

Tabel 1 Arvestite eeliste võrdlus tüübi järgi .....	13
Tabel 2 Kadude näited aegridadel koos põhjustega.....	32
Tabel 3 Erinevate programmeerimiskeelte <i>wavelet</i> teekide võrdlus .....	36
Tabel 4 Visuaalsete katsete tulemuste kokkuvõte .....	45

# 1 Sissejuhatus

Tänapäeva ühiskonnas on väga aktuaalne luua lahendusi, mis soodustavad inimkonna poolt kasutatavate ressursside säästlikku kasutust olgu selleks maavarad, energia, raha, aeg vms.

Tehnoloogia areng võimaldab aina täpsemalt jälgida ning koguda informatsiooni tarbitava ressursi kohta. Suurtes kogustes varundatud andmetega kaasnevad mitmed asjaolud, esiteks on selleks tarbijate privaatsuse tagamine, andmete mõistlik varundamine ja efektiivne andmanalüüs. Efektiivseks võib andmeanalüüsi nimetada, kui selle tulemus toob kasu, lahendab probleeme või aitab neid ette näha.

Antud töös analüüsitakse *wavelet* teisenduste rakendamisevõimalusi selleks, et lihtsustada aegridu ja muuta efektiivsemaks protsesse, mida tehakse probleemide ettenägemise eesmärgil. Probleemiks on selle töö kontekstis kaod, mis tekivad ressursside fikseerimata tarbimisest. Täpsemalt käsitletakse selles töös ressursitarbimise aegridade põhist kadude tuvastamist tarkade arvestite võrgus kasutades *wavelet* teisendusi. Kadude tuvastamise all on mõeldud kao piirkondade visuaalse eristumise parandamist.

Visuaalne kadude tuvastatavuse parandamine on aluseks edaspidiseks protsessi automatiseerimiseks, milleks on ennustusalgoritmi rakendamine tarbimise aegridadel. Ennustusalgoritmide rakendus on antud valdkonnas üheks viisiks, mis aitab reaajas tuvastada tarbimisprofiili muutusi või kõrvalekallet seniks olnud harjumuspärasest tarbimisest. Selle töö kontekstis käsitletakse seda konstantse funktsioonina, mida ainult rakendatakse lõpptulemuse hindamiseks.

*Wavelet* teisenduste kasutamist õigustab antud valdkonnas tarbimistäitade keerukus ning mitmekülgsus. Vaadeldavad aegread ei oma ranget korrapärasust ning on pidevalt muutuvad. *Wavelet* teisenduste rakendamine aitab teisendada aegridu nii detailsemale kui lihtsamale kujule säilitades samal ajal lähteandmete omadusi ja lokaliseeritavust ajateljel.

Antud töö eesmärgiks on rakendada *wavelet* teisendusi saavutamaks lähteandmed kujule, mis võimaldab paremini tuvastada kao piirkondi ning samal ajal säilitada ülevaate ajateljest.

Esimeseks alamülesandeks on teha selgeks *wavelet*'ide olemus, analüüsida erinevaid *wavelet* teisenduste omadusi ning kasutusalasid. Samuti vaadatakse lähteandmete paremaks mõistmiseks aegridade omadusi ja liigitusi.

Teiseks alamülesandeks on analüüsida ja välja tuua meetodid kuidas on võimalik kasutada *wavelet* teisendusi selleks, et jõuda peaesmärgi saavutamiseni.

Kolmandaks alamülesandeks on analüüsida lähteandmeid ja vaadata erinevaid lähenemisvõimaluste eeliseid. Uurida ning välja valida sobiv programmeerimiskeel ja teek, mis võimaldab lahendusi teostada. Teha katseid rakendades valitud meetodeid ning hinnata tulemust.

Antud bakalaureusetöö on aluseks automatiseeritud kadude tuvastuse süsteemi arendamiseks tarkade arvestite valdkonnas, milles autor osaleb arendaja rollis. Töö tulemus aitab hinnata *wavelet* teisenduste kasutamise tasuvust ning erinevaid rakendusvõimalusi selles valdkonnas. Püsitatud hüpoteesid seoses *wavelet* teisenduste kasutamisega:

- Arvestitega fikseeritud andmeid on võimalik teisendada *wavelet*'ide kujule.
- Kaod on aegridadel alati tuvastatavad .
- *Wavelet* teisendused aitavad muuta uuritavad andmed kergemini analüüsitavaks.
- Aegridade ennustamine kasutades *wavelet* teisendusi on korrektsem kui muutmata lähteandme puhul.
- *Wavelet*'ite lai valik aitab tagada teisendatud andmete täpsust.

Töö koosneb viiest sisulisest peatükist millele lisandub sissejuhatus ja kokkuvõte.

- Teises peatükis tehakse sissejuhatus valdkonda ning käsitletakse kao probleemi olemust ja mõju tööstusele üldiselt. Lisaks tuuakse välja kadude teema elektri valdkonnast Elektrilevi näitel.
- Kolmandas peatükis tuuakse välja *wavelet*'ide teoreetilised alused ja aegridade omadused ja liigitused. Lisaks tuuakse välja *wavelet* teisenduste selgitused ja kasutusvaldkonnad.

- Neljandas peatükis kirjeldatakse lahenduse meetodeid ning selgitatakse nende valiku põhjused.
- Viiendas peatükis tehakse juhtumiuuringuna lähteandmete ja lahenduse teostamiseks vajalike vahendite analüüs ning tehakse katsetusi lähteandmete peal.
- Kuuendas peatükis tehakse tulemuse ülevaade ja diskuteeritakse lahenduse kasutamise võimalusi edaspidiselt.

## 2 Kaod tarkade arvestite võrgus

Järgnevas peatükis tuuakse välja teoreetilised alused mida on tarvis probleemi põhjalikumaks selgitamiseks.

### 2.1 Targad arvestid

Tark arvesti kujutab endast elektroonset seadet, mille ülesandeks on mõõta ressursi tarbimist ning edastada saadud info kesksesse süsteemi. Tarkade arvestite funktsionaalsus seisneb peamiselt tihedas ning automatiseeritud kahesuunalises kommunikatsioonis. Teisisõnu peamine eelis mehaaniliste arvestite ees on kaugloetavus, mis tagab ülevaatlikkuse ja kontrolli toimuva üle, mis omakorda aitab ühtlustada makstava ressursi ja tarbitava ressursi vastavust. Tabelis 1 on välja toodud peamised erinevused tarkade ja mehaaniliste arvestite vahel.

<b>Arvesti tüüp</b>	Tark elektriarvesti	Mehaaniline arvesti
<b>Näitude tihedus</b>	seadistatav (üldjuhul iga tund)	kord kuus
<b>Veatuvastus</b>	kiire	aeglane
<b>Näitude edastus</b>	automaatne	käsitsi
<b>Omavoliline arvesti mõõtetulemuste mõjutamise tõenäosus</b>	väike	suur
<b>Hind</b>	kallis	odav

Tabel 1 Arvestite eeliste võrdlus tüübi järgi

Ettevõtted, kellel on kasutuses targad elektriarvestid kulutavad vähem tööjõu ressursi selleks, et kontrollida arvesteid, sest enamjaolt probleemid tuvastatavad eemalt. Samuti pole tarkade arvestite võrgus välistatud kõiksugused küberrünnakud, sest tegu on ikkagi andmevahetusega mööda erinevaid kommunikatsiooniahelaid. (Elster & McCullough, 2011)

Eesti näitel on kõige rohkem tarku arvesteid kasutusel elektrivõrgu valdkonnas. Mida on tänaseks päevaks kasutusele võetud aastatega 2013 - 2016 ligi 630 000 ühikut. (Elektrilevi OÜ, 2013)

Targad veemõõtjad on järgmine trend mida mõned vee-ettevõtjad on juba hakanud EL abiga paigaldama, kuid siiani pole globaalset väljavahetamise projektist avalikult kuulda olnud. (Kose Vesi, 2015)

## **2.2 Kaod**

Kadudeks peetakse igasugust ressursi kasutust, mis ületab arvesti poolt fikseeritud tarbimise väärtusi. Erinevate ressursside kasutuse puhul on olemas konkreetsed normid kui suur on normaalne kadu, näiteks tehnilistel põhjustel või mingeil muul loomulikel põhjustel. Antud lõputöös keskendutakse sellistele kadudele, mis ilmnevad lõpptarbija valdustes ning siinkohal välistatakse kõiksugused loomulikud kaod. Tarbija poolel võivad kaod tekkida päris mitmetel põhjustel, näiteks arvesti riknemine, arvesti kulumine, tahtlik arvesti näitude mõjutamine, halb ühendus, tahtlik ümberühendus vms.

Vaadates mehaanilisi arvesteid, siis kadude teke võib tulla peamiselt arvesti kulumisest, sest mehaanilistel osakel on omadus ajapikku kuluda. Samuti pole harv juhus, kui arvestit on meelega mõjutatud väliste teguritega eesmärgiga saada sellest omapoolne kasu.

Tarkade arvestite puhul on väliste teguritega mõjutamine raskendatud, kuna näitude edastus toimub üldjuhul ~700 korda tihedamini ning paljudel nutikatel arvestitel on funktsioon teavitada vigade tuvastustest. Kaod tarkades võrkudes ei ole aga tihtipeale tuvastatavad sellistes kommunikatsiooniahelates, kus tarbitava ressursi mõõtepunkt asub omaette harus, mida mõõdetakse ainult ühes otspunktis. Selline nähtus on väga tõenäoline näiteks elektrivõrkudes.

## **2.3 Kadude mõju**

Igal vähegi väärtuslikul ressursil on tänapäeval oma hind. Elekter, vesi ja gaas on enamasti kõige globaalsemalt kasutuses olevad ressursid, mida tarbitakse regulaarselt kõikjal. Selleks, et nimetatud ressursid jõuaks korrapäraselt tarbijateni teeavad ettevõtted palju tööd. Üheks tähtsaks tegevuseks on kindlasti kommunikatsiooniliinide korrashoid

ning korrektne tarbimise fikseerimine. Majanduslikult tähendab see pidevat kulutust tootja/vahendaja jaoks, kuna peale korrashoidu panustatakse ka uuendustesse, mis aitavad tõsta tarbitava ressursi kättesaadavust ning samas hoida hinda tarbijate suhtes mõistlikuna. Kulutuste alla lähevad uued tehnoloogiad, arendused, tööjõu tasustamine jpm.

Tarbija kui lõppkasutaja on peamine sissetulek tootja/vahendaja jaoks, seega tarbija korrektne maksustamine on ettevõttele väga tähtis. Paraku ei ole selle korrektsuse tagamine väga lihtne kuna enamuse fikseeritavast tarbimisest toimub tarbija valduses paigaldatud mõõtepunktidest. Siinkohal ei saa kunagi kindel olla, et esiteks mõõtja fikseerib korrektset tarbimist, teiseks ei saa välistada, et kadu ei teki enne arvestit. Paljudel juhtudel põhjustavad kadusid just tarbijad ise, kuna neil on näiteks majanduslikud raskused või lihtsalt hea võimalus kasu saamiseks, mida kasutatakse ära ressursi varguse näol.

Kõiksugused kaod, mida ressursi tootja/vahendaja kaotab tekitavad tegelikkuses kahjumit kõige enam tarbijatele endile. Aeglustub sellel põhjusel ettevõtte panus arengusse, mille eesmärgiks on säästlikkuse tagamine, tekib kvaliteedi halvenemise oht või tarbija jaoks kõige kurvem, tõuseb ressursi hind.

## 2.4 Hetkeolukord Elektrilevi näitel

Järgnev tekst on kokkuvõtte Elektrilevi OÜ tarbimise järelevalve osakonna juhataja Toomas Kasega läbi viidud suulisest intervjuust.

Suur osa Elektrilevi tegevusest on suunatud sellele, et hoida madalaid kulusid. Kaod jaotuvad tehnilisteks ja mittetehnilisteks kadudeks ning nende proportsionaalsust sümboliseerib järgnev joonis.



Joonis 1 Elektrilevi tehniliste ja mittetehniliste kadude osakaal

Mittetehniliste kadude põhjuseks on peamiselt mõõtmata ahel ning peamised kadude tekkepõhjuste tüübid on järgnevad:

- Omavolitsemine:
  - Harud mõõtmata ahelas
  - Vargus liinilt
  - Peakaitsme suurendamine
  - Plommid eemaldatud
  - Arvesti rikkumine
  - Mõõtesüsteemide vargus
  - Plommid rikutud
- Kommunikatsioonivead

Peamiselt tegeletakse mittetehniliste kadude vähendamisega, mille alandamine on kordades odavam kui tehniliste kadude vähendamine. Iga aastane kadude tõus on ligi 2,5% seega kadude osakonna töö eesmärk on olla sellest tõusust kiiremad, milleks on ligikaudu 33GWh aastas. Kogu kadu, mis elektrivõrgus tekib ostetakse börsilt ette, börsihind sõltub aga kadude suuruselt, mida vähem on kadusid seda odavam on hind nii sisseostul kui lõpptarbijal (näiteks 2013 vähenes tariif 0.3% võrra kadude vähendamise tõttu).

Eesti peale on kokku ligikaudu 630 000 mõõtepunkti, mida kontrolliti üks kord aastas kuni aastani 2007, millest alates rakendati järkjärgult bilansimõõtjaid, mida nüüdseks on ligi 10 000. Bilansimõõtjad loevad bilanssi alajaama punktist, mis annab võimaluse analüüsida tulemuse vastavust mõõtepunktide näitude summaga, mis kasutavad vaadeldavat alajaama. Antud mõõtjate kasutuselevõtt tingis asjaolu, kus ettevõtte vajab pea 10 korda vähem inimesi sama hulga töö ära tegemiseks kui varem.

Normaalseks ehk loomulikuks kaoks loetakse 2-3% alajaama kohta. Kadusid leiti varem *inspect-detect* meetodiga, ehk enne oletati siis leiti üles, nüüd 2007. a alates on vastupidi *detect-inspect* (vähendas ressursi 20 inimese võrra). Keskmiselt leitakse üks kadu 2h jooksul, kuid tavaliselt ainult kadude leidmisega protsess ei piirdu, järgnevalt tuleb



tegeleda selle likvideerimisega ning tõestamisega, millele kulub kordades rohkem aega. Eesmärgiks on hetkel aastaga korrastada bilanss 500 alajaamas.

Täpsemaks meetodiks loetakse tarbimise mõõtmist ning praeguseks hetkeks toimub kadude tuvastus mõõteandmete võrdlemise teel, kus kahe analüütiku töö tulemusena tuvastatakse tarbimise käitumise muutusi eelnevalt fikseeritud mõõtmistega. Bilansimõõtjatega varustatud alajaamades kulub kao tekkepõhjustaja leidmiseks aega päris vähe, küll aga läheb suures mahus ajalist ressursi neile, kellel eelnimetatud mõõtja puudub. Selliseid objekte on üle 20 000 ning nende ülevaatamine võtab aega ligi 3 kuud, ehk aastaga jõutakse üle vaadata selliseid tarbijaid ~4 korda. Teoorias on kasulik ka tarbijate tausta tundmine, mis annaks lisavõimalusi varguste kindlaks tegemistel, kuid väliste andmete kasutamine ei ole õiguslikult lubatud.

## **2.5 Tausta kokkuvõte**

Arvestades, et enamus Eesti elektriarvestitest on vahetatud tarkade lahenduste vastu, vee arvestite maailm on samuti liikumas tarkade lahenduste poole ning kaugel pole ka digitaalsete gaasi mõõtjate kasutuselevõtt.

Tarkade arvestite kasutuselevõtu tulemusena saavad ettevõtted kordades rohkem andmeid ning nende töötlemine muutub palju keerulisemaks kui üksikute muutuste analüüsimine. Mida suuremaks ja detailsemaks andmehulgad muutuvad seda rohkem tekib analüüsitava osas müra ning käsitsi on selle loetavus raskendatud. Lähtudes Elektrilevi, suurimale tarkade võrkude kasutajale Eestis, näitest võib selgelt välja tuua, et tehtud uuendused ei ole täielikult lahendanud kadude probleemi. Seda tõendab ka Tarmo Mere öeldes: "... Lisaks aitavad uued arvestid kokku hoida kulusid – näiteks on viimase kolme aasta jooksul vähenenud elektrivõrgu kaod ligi viiendiku võrra" (Rudi, 2016). Kaod on küll saanud osades juhtudes kergemini lokaliseeritavad, kuid osadel juhtudel pole endiselt tuvastusprotsess lihtne. Arvestades aastaseks elektritarbimise mahuks Eestis 7000GWh, on 180GWh endiselt märkimisväärne number, mis läheb kaduma ainult tarbijate poolt tekitatud varguste näol (Elektrilevi OÜ, 2017). Vaadates 2017 maikuu seisuga viimase aasta keskmist kWh hinda, milleks on 4.02 senti, võib öelda, et aastane kadu antud andmete põhjal ulatub 7,2 miljoni euroni.

Kui suuremates alajaamades on kadude tuvastus muutunud lihtsamaks tänu bilansimõõtjatele, siis üksikutes mõõtepunktides rakendatakse statistilist andmete analüüsi käsitsi, võrreldes sama kasutaja tarbimist eelnevate perioodidega. Eesti näitel on bilansimõõtjatega kaetud pea kõik suuremad alajaamad, kelle tarbijate arv on suurem kui 4, siis Läti ja Leedu puhul on bilansimõõtjad kasutuses väga vähestes kohtades, mis tähendab, et käsitsi tehtavat analüüsi on vaja teha pea kogu riigi tarbijate kohta.

Antud projekti skoobis on käsitleda üksikute tarbijate poolt tekitatud kadude tuvastamise protsessi lihtsustamist lähtudes Elektrilevi OÜ poolt välja pakutud reaalsele andmetele.

### 3 *Wavelet*'id ja aegread

Järgnevatel punktides on toodud välja *wavelet*'ide teooria ning erinevad aegride omadused ja eripärad. *Wavelet*'ide mõistmine on vajalik selleks, et mõista selle sisu ning osata rakendada lahenduse meetodite väljatöötamisel, mida kasutada kadude tuvastamiseks. Aegride omadused ja eripärad on kirjeldatud eesmärgiga paremini mõista ja formuleerida aegride erinevust.

#### 3.1 *Wavelet*

*Wavelet*'i kasutamine on hea vahend nii signaali detailsemaks muutmiseks, sisendandmete hulga vähendamiseks kui ka müra eemaldamiseks, jättes samal ajal võimaluse taastada esialgne kuju minimaalsete kadudega. Visualiseerimise poolelt annab see võimaluse ühel juhul tuvastada raskemini märgatavad signaali osad ning teisalt muuta seda loetavamaks. Peale signaali teisendamist säilib võimalus lokaliseerida väärtusi ajas. *Wavelet* teisenduste abil on võimalik tuvastada järsked muutused signaali väärtustes. (Zorin, 1997).

Lähtudes (Köhler & Lorenz, 2005) tehtud erinevatest müra eemaldamise katsetustest on välja toodud kõige edukamaks meetodiks aegridadest *wavelet*'ide kasutamine.

*Wavelet* ehk "laineke" kujutab endast võnkliikuvat matemaatilist funktsiooni, millel eksisteerib nii amplituud, kui võnkumised positiivsele ning negatiivsele poolusele. *Wavelet*'i perekonda liigitatav funktsioon  $\psi$  peab rahuldama kolme tingimust :

- Funktsioon peab olema tsentreeritud amplituudi null väärtusel

$$\int_{-\infty}^{\infty} \psi(t) dt = 0$$

- Peab omama lõplikku energiat. Selle tulemusena on ta lokaliseeritav ajas

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty$$

- Piisav tingimus pöördvõrdelise teisenduse jaoks

$$C = \int_{-\infty}^{\infty} \frac{|\psi(t)|^2}{\omega} d\omega < \infty$$

*Wavelet*'e nimetatakse üldjuhul mitmuses, sest nende olemus hõlmab järjestikust liikuvaid väikseid lainekehi. Omadustelt on nad ajaliselt ning sageduse poolest lokaliseeritavad ning nende peamiseks omadusteks on skaleeritavus võnkeperioodi ning amplituudi sihis. *Wavelet*'id on kasulikud suurest andmehulgast erinevat tüüpi informatsiooni eraldamiseks, müra filtreerimiseks jpm. Peamiselt toimub kogu signaali analüüsi protsess *wavelet* koefitsientideks teisendatud tasapinnal (mastaap – aeg – amplituud), peamiseks eeliseks näiteks tuntud Fourier' teisendusele<sup>1</sup> on sageduste täpne ülevaade kogu aja vältel. (Wikipedia, 2017; Goreva T.S., 2015)

Antud töö raames vaadeldakse mõistet 'signaal' peamiselt aegridade definitsioonina, s.o ajaliselt järjestatud numbriliste väärtuste informatsioon, mis on antud töö kontekstis tarkade arvestite näidud.

### 3.2 Aegread

Aegridade ehk signaali paremaks analüüsiks ning ennustuseks on tarvilik selle omaduste mõistmine. Selleks on tarvilik teada järgmiseid mõisteid.

**Trendiks** peetakse aegridade pikemaajalist muutuse suunda, mis näitab väärtuste muutuste liikumist kas positiivsel või negatiivsel suunal, samuti võib see püsida konstantsena. Trend ei pea olema lineaarne vaid võib vabalt olla ajas muutuv erineva kiirusega näiteks mingist hetkest järsult tõusev ning hiljem aeglaselt langev.

**Tasand** ehk sirgjooneline konstantne väärtus, mis iseloomustab käsitletavat aegrida vaadeldava aja vältel. (Brownlee, 2016)

**Müraks** võib nimetada selliseid aegridade osi millel pole vaadeldavas kontekstis olulist rolli. Aegridade analüüsil püütakse üldjuhul signaali hoida võimalikult puhtana, et rakendatav mudel saaks paremat tulemust anda.

---

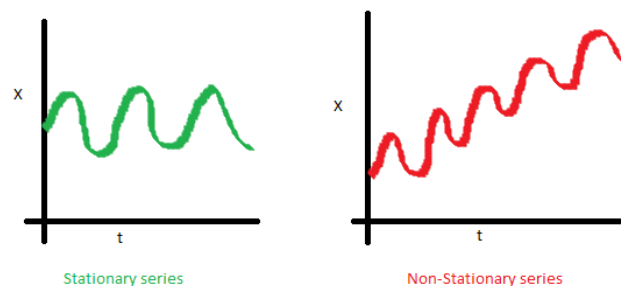
<sup>1</sup> Fourier' teisendus on integraalteisenduste hulka kuuluv lineaarne operaator, mis teisendab funktsiooni selle sagedusspektrit iseloomustavale kujule. (Wikipedia, 2017)

**Sesoonsed muutused** näitavad väärtuste muutust tihtipeale väga sarnaselt perioodilistele funktsioonidele. Muutused on korduvad teatud perioodi tagant, näiteks numbriline väärtus suvel suureneb ja talvel väheneb või päeval suureneb ja öösel väheneb. Siinkohal ei tähenda see väärtuste identseid kordusi vaid pigem kindla aja tagant taas esinevat trendi muutust sarnasel suunal. (Rob J Hyndman, 2012)

**Mittestatsionaarsed aegread** on sellist tüüpi signaalid, mille keskmine väärtus ei ole ajas konstantne st kõikumine positiivsuse ja negatiivsuse suunas ei ole keskmiselt sümmeetriline. Dispersioon ning vastastikused lineaarsed seosed arvmuutujate vahel võivad olla ajas muutuvad. Mittestatsionaarseid aegridu on reeglina raske ette ennustada. (Srivastava, 2015)

**Statsionaarsed aegread** omavad kogu ajatelje ulatuses oma omadusi konstantsetena. Esiteks ei muutu signaali keskmine väärtus ehk perioodiliste muutuste keskmine väärtus püsib muutumatuna. Signaali perioodiliste lainete kooskõlalikus ehk jäävus võnkesageduses ja võnkefaaside diferentsis. Samuti kovariatsiooni<sup>1</sup> väärtuste püsimine null väärtuses. (Srivastava, 2015)

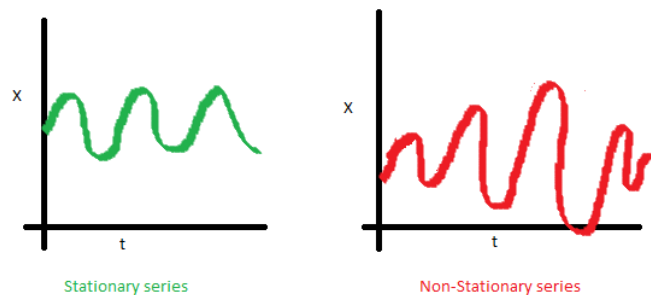
Järgnevalt on välja toodud kolmel joonisel peamised erisused statsionaarsete ja mitte statsionaarsete aegridadele vahel.



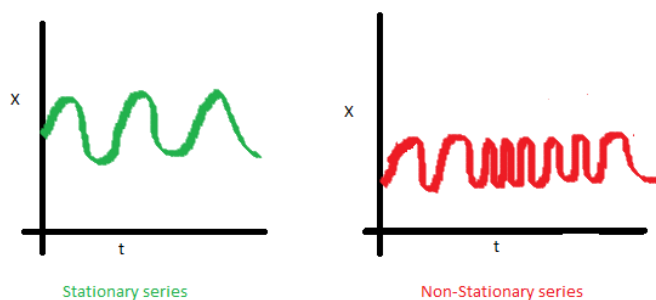
Joonis 2 Aegridade erinevus keskmise väärtuse konstantsuse järgi (Srivastava, 2015)

---

<sup>1</sup> Kovariatsioon — vastastikune lineaarne seos arvmuutujate vahel. (Kalle Remm, 2012)



Joonis 3 Aegridade erinevus koherentsuse järgi (Srivastava, 2015)



Joonis 4 Aegridade erinevus kovariatsiooni järgi (Srivastava, 2015)

Signaali tundmine ning korrektne modifitseerimine on tähtis aegridade analüüsi osa, mille lõppeesmärgiks on tihti peale võimalikult täpne prognoosimine. Erinevat tüüpi signaalide ennustamiseks loodud lahendusi on tänapäeval väga palju. Vanimaks ning populaarseimaks võib nimetada ARMA (inglise keeles *Auto Regressive Moving Average*) mudelit, mida on võimalik rakendada vaid statsionaarsetele aegridadele. Reaalsuses on aegridad pigem mitte statsionaarsed ning see muudab ennustamisprotsessi palju keerulisemaks, sest täielikuks ennustuseks on tarvis signaali dekomponeerida või modifitseerida eesmärgiga saavutada andmete erinevad komponendid. Nendeks on statsionaarne kuju lähteandmete sesoonsetest muutustest, üldised trendi muutused ja suvaliste muutuste ehk *random* osasi iseloomustav. (Amit Konar, 2017)

### 3.3 Wavelet'ite teooria ehk teisendused

*Wavelet*'id on matemaatilised funktsioonid, mida kasutatakse sisendiks antud funktsiooni või aegsignaalide eraldamiseks eritasemeliseks mastaabi komponentideks. *Wavelet* teisenduseks nimetatakse sisendi representatsiooni *wavelet*'ide abil milleks on

lähtefunktsiooni (ema-*wavelet*'i) eksemplaride (tütar-*wavelet*'ite) kogum, mida on skaleeritud amplituudi või aja sihis ühtimaks sisendväärtusi. (Wikipedia, 2017)

Fourier'-teisenduse, mille on signaali teisendus siinus ja koosinus tüüpi eritasemelised komponendid, nõrkus seisneb selles, et väljundi komponendid on lokaliseeritavad ainult sageduse domeenis. *Wavelet* teisenduste eelis seisneb tulemuse lokaliseeritavusel nii sageduse kui aja domeenis ning samuti säilib ka võimalus teisendada andmed esialgsele kujule. Lähtudes (Polikar, 1996) tehtud katsetele võib samuti tuua *wavelet* teisenduste eeliseks Fourier' teisenduste ees võimet paremini toime tulla mitte statsionaarsete aegridadega milleks on ka tarbitavate ressursside näidud.

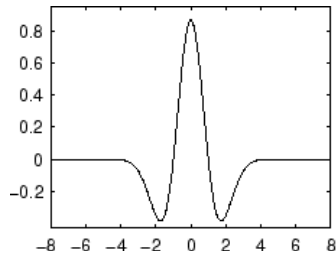
*Wavelet* teisendusi on võimalik jaotada mitmete erinevate *wavelet* funktsiooni omaduste alusel, kõige suurem eristus on ortogonaalsuse alusel. Ortogonaalsuse tunnust omavaid *wavelet* teisendusi nimetatakse diskreetseteks *wavelet* teisendusteks ning mitte ortogonaalsuse puhul pidevateks *wavelet* teisendusteks. (Klapetek, et al., 2004–2007, 2009–2016)

### 3.3.1 Pidev *wavelet* teisendus

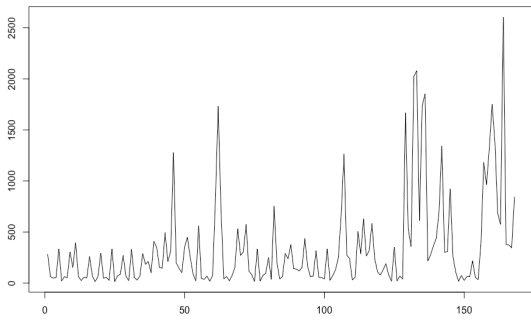
Pidev *wavelet* teisendus (inglise.k *continuous wavelet transforms* ehk CWT) peamine kasutusala on teaduslik andmete analüüs. Peale teisenduse rakendust saadakse ühe dimensiooni võrra rikkamad andmed, mis võimaldab tuua välja detaile mida varasemalt andmetest ei olnud võimalik tuvastada. CWT lähtub järgnevast funktsioonist, kus  $S$  on pidev *wavelet* teisendus,  $a$  on mastaap,  $b$  on aja nihe, on  $\psi$  on *wavelet* funktsioon. (Chaovalit, et al., 2011)

$$S(b, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi' \left( \frac{t - b}{a} \right) s(t) dt,$$

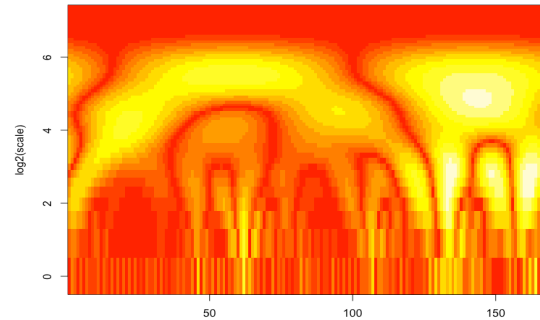
Järgnevatel joonistel on näha 'Mexican hat' funktsiooni, mida on rakendatud näidisandmete peal ning ülevaadet näidis signaalist enne ja peale pidevat *wavelet* teisendust.



Joonis 5 *Wavelet* funktsioon 'Mexican hat'



Joonis 6 Signaal enne CWT teisendust



Joonis 7 Signaal peale CWT rakendust

Pideva *wavelet* teisenduse halvimaks omaduseks on andmete liiasuse tekitamine, mis andmete detailsemaks muutmise kõrval võib tekitada väga palju liigset müra.

Antud teisendus leidnud rakendust näiteks EKG<sup>1</sup> analüüsi valdkonnas, kus on väga oluline tabada spetsiifiliste sageduste väärtusi, mis võivad olla väga lühiajalised kõrgetel sagedustel ning pikaajalised madalatel sagedustel. (Moga & Victor Dan Moga, 2005; Addison, 2005)

### 3.3.2 Diskreetne *wavelet* teisendus

Diskreetne *wavelet* teisendustel (inglise.k *discrete wavelet transforms* ehk DWT) on mitmeid kasulikke rakendusi aegriidade andmekaeve vallas. Antud teisenduse eesmärgiks on vähendada signaali andmete kogust, välja filtreerida müra või tuvastada ebakorrapäraseid käitumisi. Teisendust läbinud andmed on dekomponeeritud erinevatele sageduste alusel kaheks: ligikaudsed väärtused (inglise.k *approximations*) ja detailsed väärtused (inglise.k *details*). Ligikaudsed hõlmavad endas madalama sageduse väärtusi ning detailsed kõrgema sageduse väärtusi. Diskreetne teisendus võimaldab samuti

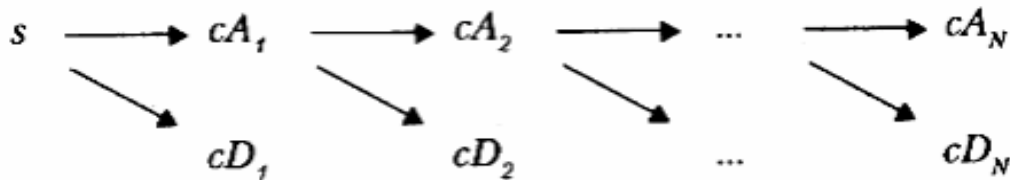
---

<sup>1</sup> EKG – ehk elektrokardiogramm on südamelöökide elektrilisel mõõtmisel saadud signaali visualisatsioon.



mitmetasandilist teisendust, mis tähendab ligikaudsete väärtustele teisenduse rakendamist mitmekordselt nii nagu näidatud all oleval joonisel.

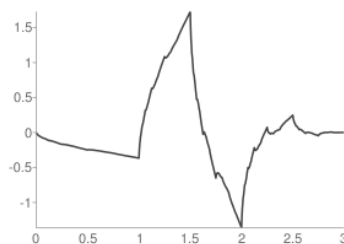
$S$  – lähtesignaal,  $cA_1 - cA_n$  ligikaudsed väärtused erinevatel tasanditel,  $cD_1 - cD_n$  detailsed väärtused erinevatel tasanditel.



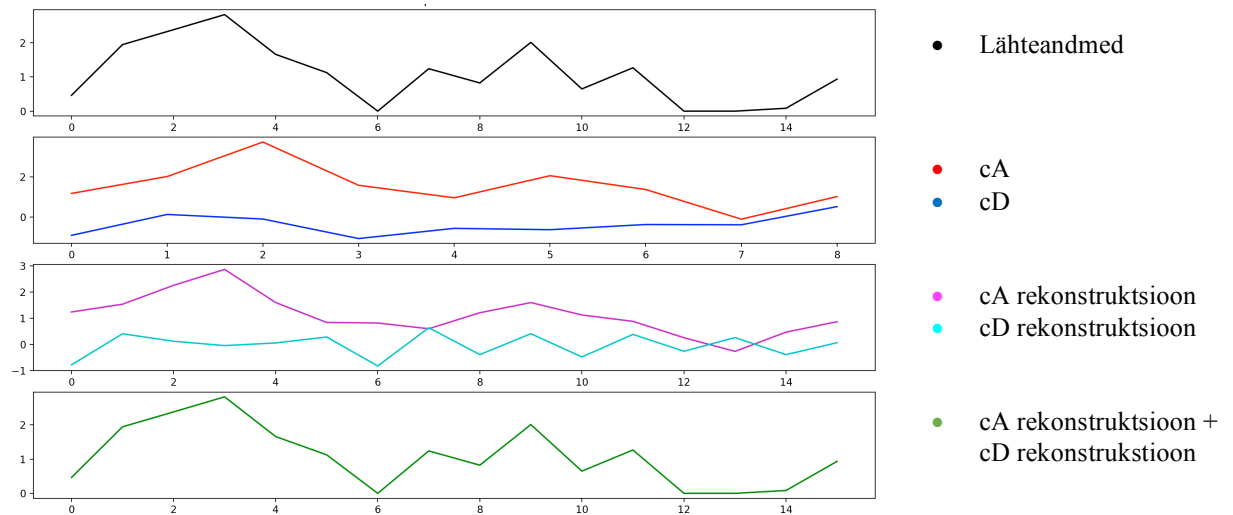
Joonis 8 Mitmetasandiline diskreetse *wavelet* teisenduse joonis (Ogorodov, 2009)

Antud teisenduse puhul on negatiivseks küljeks range seos filtri mastaabi ja asukoha vahel, mis võib spetsiifiliste signaaliosade tuvastamise keerulisemaks muuta.

Järgnevatel joonistel on toodud ‘*Daubechies 2*’ *wavelet* funktsioon ning signaali diskreetne teisendus kasutades eelnimetatud funktsiooni. Teisenduse joonisel on näidatud, esimeses kastis aegrida esialgsel kujul ning teisel on näha koefitsiente  $cA$  ja  $cD$ , mis tekivad peale 1 taseme teisendust. Kolmandas on kujutatud eraldi koefitsientide vastupidine teisendus ehk rekonstrueeritud kuju, kus on näha, et andmed omandavad esialgse väärtuste hulga. Viimases on näidatud summa  $cA$  ja  $cD$  rekonstruktsioonidest, mis tõestab, et andmed on teisendatavad esialgsele kujule.



Joonis 9 *Wavelet* funktsioon ‘*Daubechies 2*’



Joonis 10 Diskreetne *wavelet* teisendus kasutades "Daubechies 2" funktsiooni

Diskreetsete teisenduste kasutusala on väga lai tänu *wavelet* funktsioonide suurele hulgale. Kõige laiemalt on leitud kasutust näiteks pilditöötluses ja pildituvastuses. Üldtuntud pildiformaat JPEG 2000 põhineb diskreetsel *wavelet* teisendusel, mille tulemusena JPEG formaadis olev pilt on lähteformaadist kordades väiksema mahuga.

## 4 Lahenduse väljatöötamine

Kontseptsioonis on kasutada *wavelet* teisendusi aegridade teisendamiseks kujule, milles on andmeid kergem analüüsida. Kergema analüüsimise all on mõeldud ennekõike kadude visuaalse tuvastatavuse parandamist, mis tähendab ühel juhul müra eemaldamist signaalist teisel juhul signaali teisendamist sageduste domeeni.

Müra eemaldamine on vajalik, sest lokaalsed miinimum ja maksimum väärtused ei oma erilist tähtsust juhul kui otsitav piirkond, antud töö kontekstis kadu, on nähtav järskude muudatuste näol erinevate aegridade omaduste poolest.

Kui *wavelet* teisenduste tulemusena saadud koefitsiendid on väiksema müra sisaldusega ning visuaalselt on teadaolev kao piirkond endiselt selgelt tuvastatav, siis võib järeldada, et käsitletava näite jaoks on *wavelet* teisendus kasutatav ning edaspidi teisendatud kuju on rakendatav ennustamisalgoritmi sisendiks. *Wavelet*'ite kasutamine ennustamise tulemuse parandamise eesmärgil on õigustatud mitmetes teaduslikes väljaannetes (Fryzlewicz, 2003).

Ennustuse kasutamine on hea võimalus kadude tuvastamiseks kuna see annab võimaluse võrrelda reaajas tulevaid andmeid ennustatud väärtustega, mis tuginevad eelnevalt fikseeritud ajalool. Siinkohal on väga tähtis teada, et ennustusalgoritmi jaoks antud sisendandmetes ei esineks kadusid. Kui tekib liialt suur vahe ennustatud väärtuse ja reaalse väärtuse vahel võib öelda, et antud piirkonnas on kao kahtlus. Samuti on oluline, et ennustus oleks võimalikult täpne.

Antud töö raames ei käsitleta ennustusalgoritmi detailset ülesehitust ning seadistust. Kasutusel olev ennustusalgoritm on antud töö kontekstis konstantne funktsioon, mida kasutatakse tulemuste võrdlemiseks. Ennustusalgoritmina kasutatakse SARIMA (*seasonal autoregressive integrated moving average*) ja *time series lineaar regressioon model* meetodeid. Tulemuste all on mõeldud ennustatud väärtusi, mis on ühel juhul saadud lähteandmete ja teisel juhul diskreetsest *wavelet* teisendusest saadud andmete õppimise tulemusena.

Aegridade teisendus sageduste domeeni ehk pideva *wavelet* teisenduse kasutamine on mõeldud kasutamiseks juhtudel, kus visuaalselt ei ole võimalik kao piirkonda eristada ka esialgsetel andmetel. Tulemuste võrdlemine toimub ainult visuaalsel tasemel.

Peamiseks eesmärgiks on realiseerida kasulik andmete teisendus tingimustega:

- Säilib või paraneb kao piirkonna eristumine
- Säilib ülevaade ajast

Katsetuste etapid:

- Vastava *wavelet* teisenduse rakendamine kasutades erinevaid meetodeid ning tulemuse visualiseerimine
- Hinnangu andmine teada olnud kao piirkondade tuvastamisele

Lisaks positiivse visuaalse hinnangu puhul üksiku(te) tulemuste rakendamine ennustamise sisendiks ning saadud tulemuse hindamine.

#### **4.1 Diskreetsete *wavelet* teisenduste kasutamine**

Lähtudes diskreetsete *wavelet* teisenduste omadustest ning rakendusvaldkonnast on eesmärgiks teisendada aegread kujule, mis võimaldab lihtsamini tuvastada kadude piirkonda. Kuna näidisandmed mille peal meetodid rakendatakse on oma profiili poolest erinevad, kasutatakse erinevaid *wavelet* funktsioone uurimaks, kas erinevate kadude profiilide tuvastamisel kujunevad välja paremat ja halvemat tulemust andvad funktsioonid. Esialgne *wavelet* funktsiooni valik on tehakse rakendades teisendusi kõikide teegis leiduvate funktsioonidega lähteandmetel, seejärel teisendatakse  $cA$  koefitsient esialgsele kujule, et tagada andmehulga arvu võrdsus ning rakendatakse keskmise ruudu vea meetodit ( $MSE^1$ ) tuvastamiseks parimini ühtivad funktsioonid, mille valikust lähtuvalt tehakse edaspidi visuaalset hinnangut.

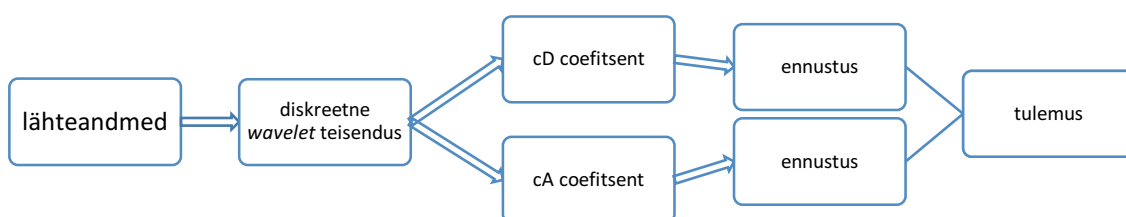
Kuna töö eesmärgiks on muuta andmed paremini analüüsitavaks ning ennustatavateks, siis kasutatavate meetodite valik on inspireeritud erinevatest *wavelet* teisenduste rakendustest ennustamise eesmärgil.

---

<sup>1</sup>  $MSE$  ehk *mean square error* – statistikas kasutatav meetod, mis arvutab väärtuste paaride vigade keskmist ehk teisisõnu vahe kahe arvrea väärtuste vahel (Wikipedia, 2017)

#### 4.1.1 Ühe taseme teisenduste kasutamise meetod rekonstrueerimiseta

Lähtudes Evgenii Mirotini (Mirotin, 2011) näitest on võimalik kasutada ennustust esimese taseme diskreetse *wavelet* teisenduse koefitsientidele ilma teisendamata esialgsele kujule. Tulemuse täpsustamiseks kasutab ta kahe koefitsiendi liitmist. Mirotini põhjendus antud lähenemisele seisneb selles, et teatud juhtudel võib koefitsientide rekonstruktsioon esialgsele kujule tekitada moonutusi signaali otspunktides, mille tulemusena halveneb ennustuse tulemus. Järgneval joonisel on näidatud, kuidas toimub andmete teisendus ning ennustuse rakendamine.



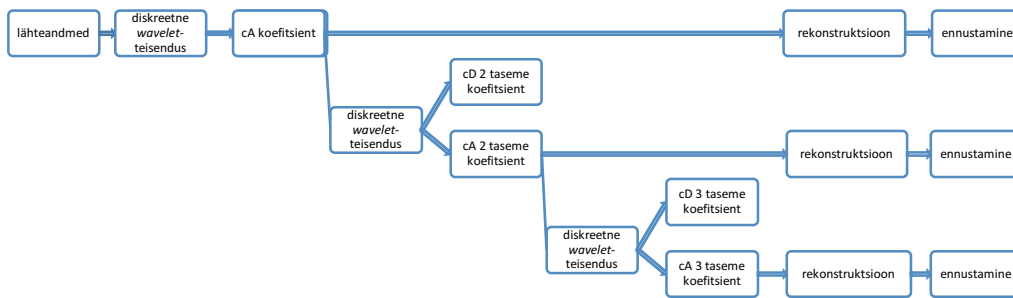
Joonis 11 Ühe taseme teisenduste kasutamise meetodit selgitav joonis

Teades, et paljude *wavelet* funktsioonide rakendamise tulemusena muutub lähteandmete hulk väiksemaks, on see suureks eeliseks näiteks andmete varundamise korral. Negatiivseks küljeks on ajatelje mastaabi muutumine, mis tähendab et peab säilitama esialgsete aegridade pikkust ning täpseks aja lokaliseerimiseks peab andmeid teisendama.

#### 4.1.2 Mitmetasandilise teisenduse kasutamise meetod rekonstrueerimisega

*Wavelet* transformatsioon võimaldab mitmetasandilist dekompositsiooni, ehk mitmekordset signaali eraldamist erinevate sageduste järgi, mis tagab andmete säilimise ning võimaluse teha vastupidise teisenduse esialgsele kujule. Rakendamiseks kasutatakse vaid cA koefitsiente, kuna need omavad väiksemat müra taset esialgsest signaalist. Erineval teisenduse tasemel saadud signaalid rekonstrueeritakse esialgsele kujule, et tagada täpsem ülevaade ajas, kuna iga taseme teisendusel saadakse üldjuhul kaks korda väiksem hulk andmeid. (Chaovalit, et al., 2011)

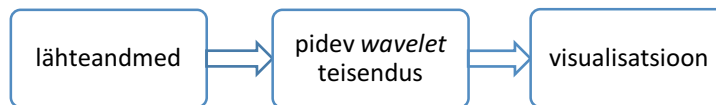
Järgneval joonisel on kujutatud kolme tasandiline diskreetse *wavelet* teisenduse joonis.



Joonis 12 Kolme tasandilist *wavelet* teisenduse kasutamist selgitav joonis

## 4.2 Pideva *wavelet* teisenduste kasutamine

Pideva *wavelet* teisenduse kasutamine on mõeldud kasutamiseks juhtudel, kus visuaalselt ei võimalik kao piirkonda eristada ka esialgsetel andmetel. Teades, et pidev *wavelet* teisendus muudab andmed detailsemaks, siis antud meetodit rakendatakse ainult eesmärgiga hinnata kao visuaalse tuvastamise lihtsust. Alljärgneval joonisel on toodud teisenduse rakendamise joonis.



Joonis 13 Pideva *wavelet* teisenduse kasutamist iseloomustav skeem

## 5 Juhtumiuuring

Juhtumiuuringus analüüsitakse ja valitakse lähteandmed mille peal rakendatakse peatükis 4 välja toodud meetodeid testandmete peal. Testandmeteks on saadud 8 komplekti reaalseid elektrinäitude väärtuseid, mis on kogutud ühe kuni kahe ja poole aasta ajavahemikus erinevatelt tarbijatelt, kelle tarbimistäitude analüüsi tulemusena on tuvastatud kadusid.

### 5.1 Näidisandmete analüüs

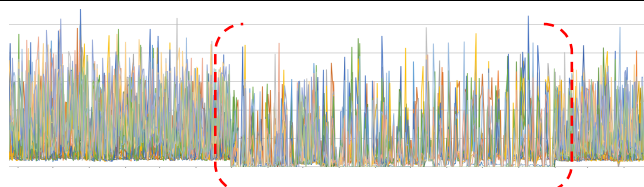
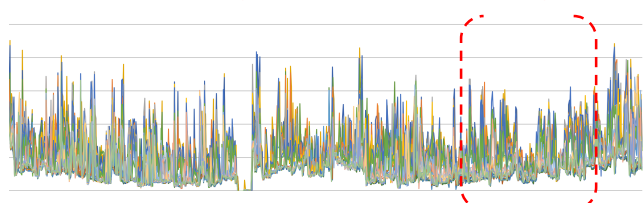
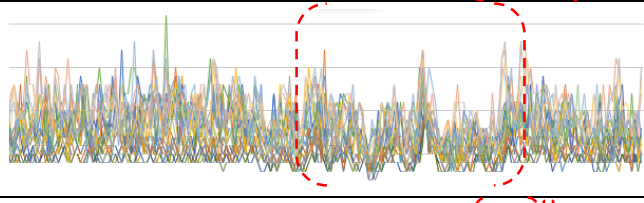
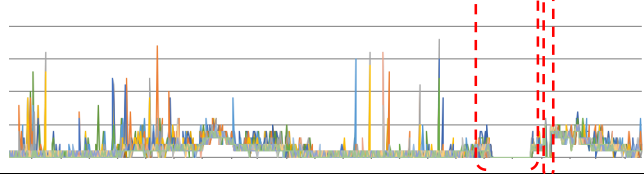

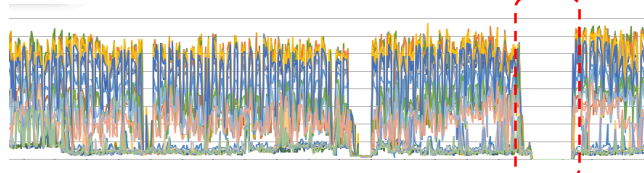
Näidisandmete analüüs on vajalik, et saada ülevaade näidisandmetest ning tuvastada kadude piirkondi visuaalselt ja tuua välja erinevate kadude omadusi aegridade vaates.

#### 5.1.1 Näidisandmetes välja toodud kadude analüüs ja valik

Elektrilevi OÜ poolt saadud elektrinäitude igatunnised andmed on reaalsest olukordadest, kus praeguseks hetkeks on kaod tuvastatud. Järgnevas tabelis on toodud 8 näidet koos põhjuste ja aegridade illustatsioonidega. Vertikaalne telg visualiseerib tarbimise mahtu, horisontaalteljel olevad väärtused sümboliseerivad kuupäevi, iga erivärvi joon (24 tk) näitab erinevat tunni väärtust ning punase punktiirjoonega märgitud alad on teada olevad kao piirkonnad.

Näide	Kao põhjus	Kao piirkonna visualisatsioon
1	“L2 <sup>1</sup> pingemuhe oli lahti”	
2	“Arvesti rikkumine 2 faasi ei mõõda.”	

<sup>1</sup> L1, L2, L3 – on üldkasutatav faaside tähistus kolmefaasilises elektrisüsteemis.

3	“Klient juhtis herkonlülitiga <sup>1</sup> arvesti tarbimise mõõtmist”	
4	“Arvestil L2 pinge puudus, AR 5 mõõtmisel viga - 29,8%”	
5	“L2 väljuv pinge puudus (halb ühendus)”	
6	“mõõteskeem vale, primaarahela kaabli vale paigutus, arvesti luges tarbimise valesti.”	
7	“KLG rike”	
8	“KLG rike”	

Tabel 2 Kadude näited aegridadel koos põhjustega

Välja toodud näidete pealt võib näha, et kõik tarbimise andmed on teineteisest erinevad ning omapärased trendi ja sesoonsuse poolest. Kõigi näidete signaalid on mitte statsionaarsed kuna koherentsus ei ole tagatud ning väärtuste keskmine on aja jooksul muutuv. Lähtudes illustatsioonidest on näha, et näidetel 6, 7, 8 iseloomustab kadusid väärtuste täielik kadumine aegridade vaatest. Näidet 1 iseloomustab maksimaalsete väärtuste langemine kogu kao perioodi ulatuses. Väärtuste koondumine ühe aegridade tasandi ümber ning samuti väärtuste langemine on iseloomulik näitele number 2. Kolmandas näidisjuhtumis on esmapilgul erinevust raske märgata, kuid peamine eristus tekib visuaalselt hõreduse näol mis tähendab, et päevade vaates toimus tarbimise

<sup>1</sup> Herkon lüliti – keelrele ehk herkonrelele on vooluahelate lülitamise seade.



märgatav vähenemine kindla perioodi tagant. Samuti esineb ka minimaalsete väärtuste langemine madalamale tasemele kui kadudeta perioodidel, mis esineb ka näitel number 5. Neljanda näite kadude leidmine on visuaalselt kõige keerulisem ehk illustratsiooni põhjal täiesti märkamatu.

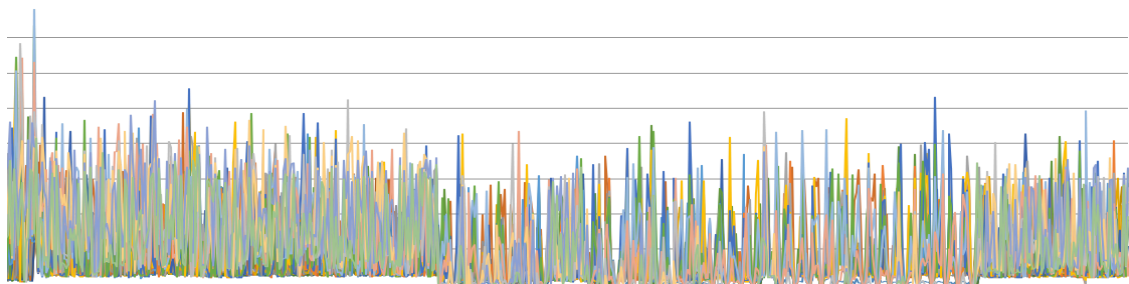
Lähteandmete analüüsi kokkuvõtteks võib järeldada, et elektritarbimise andmed on tarbijati väga erinevad ning igale kao põhjuse liigile on iseloomulikud erinevat tüüpi muutused tarbimise andmetes. Kõige enam iseloomustab kadusid vastavates piirkondades fikseeritud väärtuste järsud langused või muutused hälbe ulatuses keskmise väärtuse ümber.

Lähtuvalt näidisandmete ülevaatamisel välja paistnud signaali omadustele kasutatakse edaspidiselt diskreetse *wavelet* teisenduste rakendamisel näiteid 1, 2, 3, 5 ja 6 kuna nendel näidetel on omapärased, teineteisest erinevad omadused ning visuaalselt on kao piirkond tuvastatav. Näitele number 4 rakendatakse pidevat *wavelet* teisenduse meetodit eesmärgiga tuvastada kao piirkond visuaalselt.

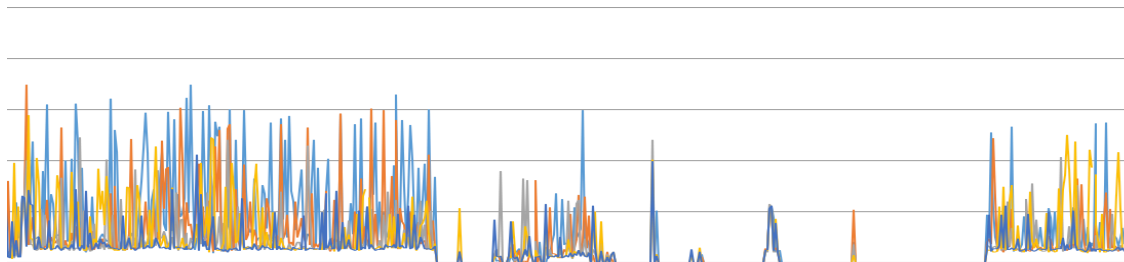
### 5.1.2 Erinevad lähenemised andmete kuvamisele

Antud peatükis on kirjeldatud kolme võimalust, kuidas on võimalik lähteandmeid kuvada. Iga võimaluse kohta on välja toodud illustreeriv näide.

1. **Konkreetses tunni näitude võrdlemine päevade kaupa** annab võimaluse analüüsida ühe tunni väärtuste muutust kogu ajaloo vältel. Säilib võimalus vaadelda igat tundi eraldi ning teades millal on tarbimine koordinaarselt erinev võib andmed liigitada mitmeks ühiseks osaks. Järgneval kahel joonisel on toodud võrdlus kõigi tunniväärtuste kuvamisel korraga ning ainult öise (üks ühine osa sarnaste näitudega üle kogu aja) aja väärtused samal ajavahemikul.

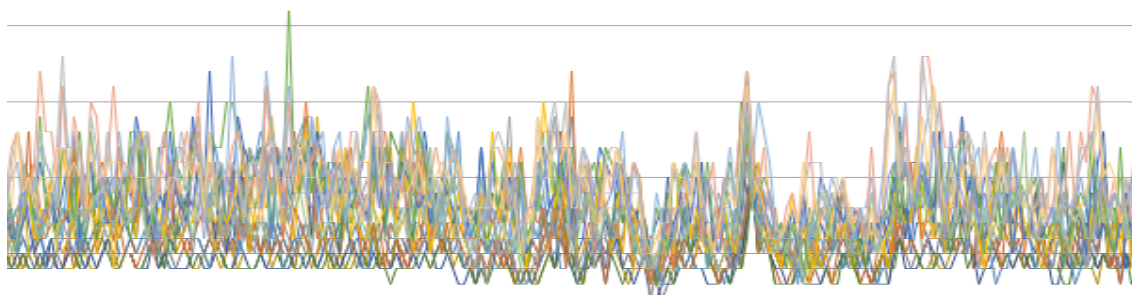


Joonis 14 Näide 3 - aegread kõikide tunniandmetega eraldi

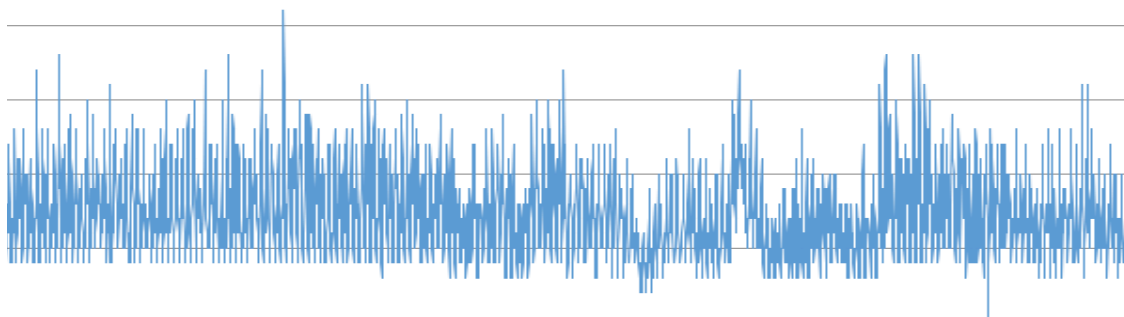


Joonis 15 Näide 3 - aegread ainult öötundidega

2. **Järjestikune päevade võrdlemine igatunnise täpsusega** annab parima ülevaate väärtuste esinemise kohta. Antud vaate halvaks küljeks on 24 korda suurem väärtuste maht ajateljel, mis võib mõjutada aeglustavalt nii andmetöötlust programmides kui ka visualiseerimisel. Visuaalselt on mõlemal joonisel võrdväärset eristatav signaali kuju. Järgnevalt on esitatud visuaalne võrdlus kahe lähenemise vahel.

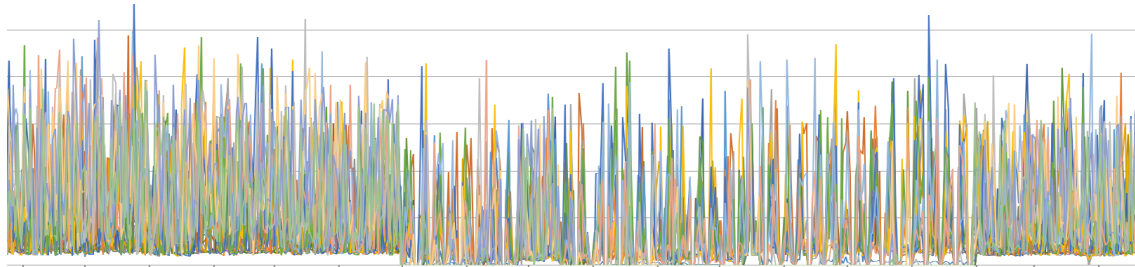


Joonis 16 Näide 5 - aegread kõikide tunniandmetega eraldi

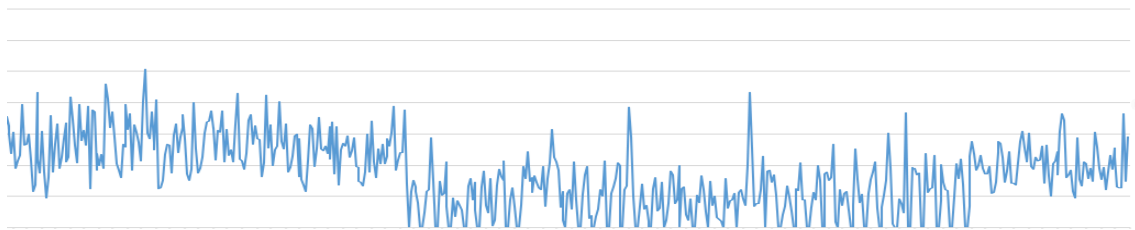


Joonis 17 Näide 5 - aegread tunni täpsusega

3. **Päeva tarbimiste summa võrdlemine** annab palju parema ülevaate juhul kui toimub tarbimise mustri kardinaalne muutus. Andmete hulk väheneb 24 korda samal ajavahemikul. Halb külge kaob suur hulk andmeid ning juhtudel nagu näites 5 muutub otsitav kao piirkond hoopis ähmasemaks. Järgneval kahel joonisel on kuvatud aegridade vaade kõikide väärtustega korraga ning teisel joonisel sama ajavahemiku summeeritud väärtused päeva kohta.



Joonis 18 Näide 3 - aegread kõikide tunniandmetega eraldi



Joonis 19 Näide 3 - aegread päeviti summeeritud väärtustega

Kõige kasulikumateks lähenemisteks andmete kuvamisele tuginedes vaadeldavatele näidetele on variandid 1 ja 3. Lähtudes kao mõiste olemusest (ehk fikseeritava tarbimise väärtuse langemisest) on teoreetiliselt kõige optimaalsem variant number 3, kuid näites 3 välja toodud olukorras, kus tegelikkuses summaarne väärtus võib olla sama, on antud variant kehv. Selleks, et säilitada võimalik andmete täpsus kasutatakse edaspidi 1. lähenemist andmete kuvamisele.

## 5.2 Rakendatava teegi valik

Lahenduse realiseerimiseks kasutatav programmeerimiskeele ja teegi valik on otsustatud lähtuvalt teekide võimalustest, peamiseks kriteeriumiks on edaspidise kasutamise lihtsus. Valikus välja toodud programmeerimiskeeled R ja Python on ühed kõige enam kasutatavad keeled andmeanalüüsis (Piatetsky, 2014; Pointer, 2016). Alljärgnevas tabelis on välja toodud teegid ning nende omaduste võrdlus.

Programmeerimiskeel	R	R	R	Python
Teek	Wavelets	wmtsa	WaveletComp	pyWavelets
CWT teisendus	puudub	olemas	olemas	olemas
Teisenduste visualisatsioon spektrogrammil		olemas omaette funktsioonina	olemas omaette funktsioonina	peab ise leiutama
CWT funktsioonide arv	-	4tk	1tk	21tk
CWT edaspidise kasutamise võimalus	-	teegi spetsiifiline klass, keeruline välja võtta	teegi spetsiifiline klass, võimalik saada väärtused aegreana.	lihtne, väärtused arvreana.
DWT teisendus ja rekonstruktsioon	olemas	olemas	puudub	olemas
DWT rekonstruktsioon	olemas, cA ja cD korruga	olemas, cA ja cD korruga	-	Olemas, cA ja cD korruga ja eraldi
DWT funktsioonide arv	25tk	32tk	-	106tk
DWT edaspidise kasutamise võimalus	teegi spetsiifiline klass	teegi spetsiifiline klass, keeruline välja võtta	-	lihtne, väärtused arvreana.

Tabel 3 Erinevate programmeerimiskeelte *wavelet* teekide võrdlus

Wavelet funktsioonide valiku ja edaspidise kasutamise lihtsuse poolest domineerib Pythoni teek pyWavelets nii CWT kui ka DWT teisenduste puhul, kuid pyWavelets puhul on keeruline visualiseerida CWT tulemust, mis on antud teisenduse juures peamine vajadus. Edaspidi kasutatakse diskreetsete *wavelet* teisenduste rakendamisel Pythoni teeki pyWavelets ja pideva *wavelet* teisenduse tarbeks R keele wmtsa'd, kuna see omab laiemat funktsioonide valikut kui WaveletComp.

### 5.3 Lahenduste realiseerimine

Lahenduse realiseerimise jaoks vajalike ettevalmistavate tegevuste loetelu programmeerimisel:

1. Andmete importimine .xlsx failist
2. Andmete teisendus mugavamasse formaati, et oleks mugavam valida andmeid vajalikus ajavahemikus kuupäeva põhisedelt.
3. Sobiva piirkonna valimine (näidete visualiseerimise jaoks on kasutatud 1 aasta pikkust vahemikku)

Järgnevad tegevused on välja toodud peatükkide 5.3.1 ja 5.3.2 alguses.

#### 5.3.1 Diskreetse *wavelet* teisenduse meetodid ja visuaalne võrdlemine

1. Lähteandmete võrdlemine MSE meetodi abil *wavelet* teisenduse rekonstrueeritud 1 taseme väärtustega. Parimate *wavelet* funktsioonide välja toomine iga aegrea (kokku 24 tk) jaoks lähtudes vähimast vea väärtusest.
2. Diskreetse *wavelet* teisenduse rakendamine, koefitsientide teisendus esialgsele kujule. Operatsiooni tehakse kasutades eelnevalt valitud *wavelet* funktsiooni.
3. Visuaalseks võrdlemiseks graafikute genereerimine iga parimaks valitud *wavelet* funktsiooniga, mis on välja toodud tabelis (Lisa 1).

#### Näide 1

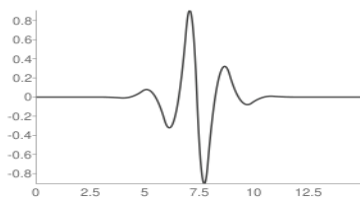
Lähtudes parimatest *wavelet* funktsioonide valikust MSE meetodi abil (Lisa 1) on näha, et funktsioonid on iga tunni aegride kohta väga erinevad. See põhjendab signaalide sesoonsuse, trendi ja keskmise väärtuse mitmekülgust igal vaadeldaval ajatunnil ning näidisandmete keerukust üleüldiselt.

Iga 24 tunni aegrea peale kokku sobis 13 erinevat *wavelet* funktsiooni. Vaadates kõige enam esinenud funktsioonide perekondi on enim välja paistvamad 'rbio' (inglise. k

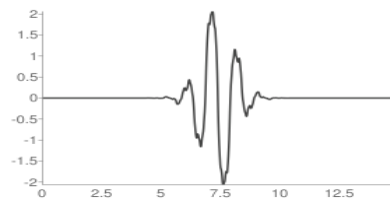
*Reverse biorthogonal wavelet*) ning 'db' (inglise. k *Daubechies<sup>1</sup> wavelet*). Siinkohal on tähtis märkida, et öise aja tarbimisel, kus on elektri tarbimismaht märgatavalt väiksem kui päevasel ajal, on paremini ühtiv 'db' perekond ning päevasel ajal domineerivad 'rbio' perekonna funktsioonid.

Erinevate *wavelet* teisenduste tulemuste visuaalsel võrdlemisel selgus, et kõigi 13 erineva funktsiooni puhul on visuaalne kao piirkonna eristumine säilinud. Esimese meetodi rakendusel saadud cA ja cD koefitsientide summa säilitab kõige detailsema kuju ning 1, 2, 3 taseme koefitsientide konstruktsioonid on sujuvamad. Esimese meetodi tulemuste kohal on oluline märkida, et erinevate *wavelet* funktsioonide kasutamine muudab aegridade pikkust erinevalt ning tulemuseks saadud aegridade pikkus erineb funktsiooniti ligi  $\pm 4\%$  võrra. Teise meetodi puhul säilib aegridade pikkus konstantselt muutumatuna olenemata teisenduse tasemest, mis tõestab diskreetsete *wavelet* teisenduste omadust säilitada ajatelje mastaapi võrdväärsena lähteandmetega.

Visuaalselt kõige paremini kao piirkonna eristumist on välja toonud funktsiooni 'rbio3.7' kasutamine. Järgnevatel joonistel on välja toodud funktsiooni 'rbio3.7' kuju teisendamisel ja rekonstrueerimisel.



Joonis 20 Funktsiooni 'rbio3.7' kuju teisendusel  
(pywavelets, 2008-2017)

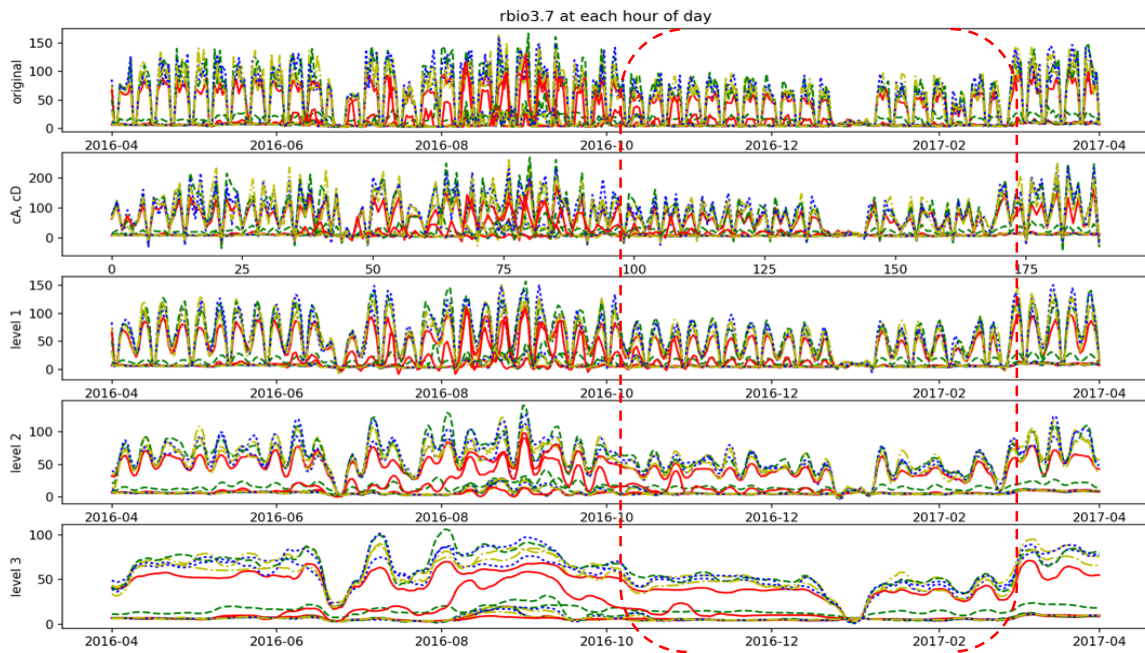


Joonis 21 Funktsiooni 'rbio3.7' kuju  
rekonstrueerimisel (pywavelets, 2008-2017)

Alljärgneval joonisel on välja toodud visuaalselt parimaks valitud *wavelet* funktsiooni 'rbio3.7' rakendus 1. näite andmetel, kus esimesel real on näha lähteandmeid, teisel real 1. meetodi tulemit ning 3.-5. real 2. meetodi tulemit, kao piirkond visualiseeritud punase punktiirjoonena.

---

<sup>1</sup> *Daubechies* – on *wavelet*'ite perekond, mis on saanud oma nimetuse nende looja USA matemaatiku Ingrid Daubechies järgi.



Joonis 22 Näide 1 – ‘rbio3.7’ teisendus

Kokkuvõttes on kõigi 13 erineva funktsiooni kasutamine andnud hea tulemuse. Kõige paremini tekib visuaalne kao piirkonna eristumine teise meetodi 3. taseme tulemis, kus aegriidade sesoonsus on muutunud väga ühtlaseks ning kao piirkond on eristuv tänu madalamale keskmise taseme väärtusele.

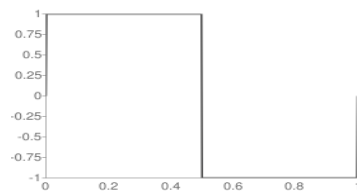
## Näide 2

Tuginedes tabelile (Lisa 1) on parimate funktsioonide valiku tulemusena 2. näite jaoks sobinud 18 erinevat *wavelet* funktsiooni, mida on võrreldes 1. näitega rohkem, ehk andmed on keerukamad kui 1. näite puhul. Kõige enam on esinenud ‘db’ perekonna funktsioonide ühtivust. Kuna antud näite lähteandmed on keerukad, siis funktsioonide valiku osas ei ole märgata mustrit, millele võib leida põhjendust.

*Wavelet* teisenduste tulemuste visuaalsel võrdlemisel lähteandmetega selgus, et *wavelet* teisenduste rakendamine on umbes pooltel juhtudel muutnud kao piirkonna eristumise hoopis keerulisemaks. Kõige paremini suutis otsitava piirkonna esile tuua funktsiooni ‘bior1.1’ (inglise.k *biorthogonal wavelet*) kasutamine. Huvitav on fakt, et antud funktsioon on üks kõige harvemini esinenud funktsiooni sobivus. Põhjuseks miks paljud tulemused, muutsid andmeid kehvemaks on piiride hägusaks muutumine. Siinkohal on oluline märkida, et kao piirkonnal on väga erilised omadused. Kao tekkimisel toimub küll järsk tarbimise vähenemine võrreldes varasemate perioodiga, kuid edaspidi on

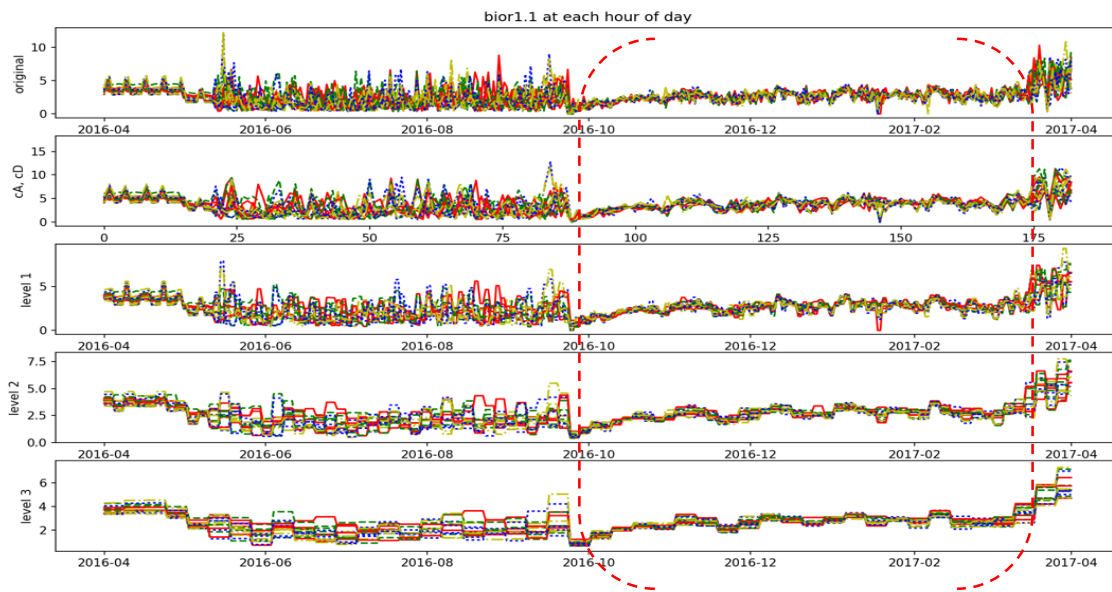
aegridade tase väga sarnane normaalsele käitumisele. Kao piirkonda iseloomustab kõige rohkem kõigi kellaegade väärtuste koondumine sarnasetele sesoonsetele ja ta trendi omadustele.

Visuaalselt on diskreetse *wavelet* teisenduse tulem kasutades funktsiooni 'bior1.1' kõige efektiivsem, kuna selle puhul säilis kõige paremini järsu muudatuse eristumine tarbimise profiilis. Antud valikut põhjendab selle funktsiooni diskreetne kuju, mis on toodud alljärgneval joonisel.



Joonis 23 Funktsiooni 'bior1.1' kuju teisendusel ja rekonstrueerimisel (pywavelets, 2008-2017)

Alljärgneval joonisel on välja toodud visuaalselt parimaks valitud *wavelet* funktsiooni 'bior1.1' rakendus 2. näite andmetel lähtuvalt kahe meetodi rakendusest. Aegridade paigutus on identne esimesele näitega.



Joonis 24 Näide 2 – 'bior1.1' teisendus

Kokkuvõtteks on 2. näite kao piirkonna omadustest sõltuvalt väga keeruline valida õiget *wavelet* funktsiooni, mis aitab teisendada lähteandmed antud töö eesmärgi kontekstis paremale kujule. Sobiva funktsiooni valikul on kao piirkonna tuvastus tagatud kõigi

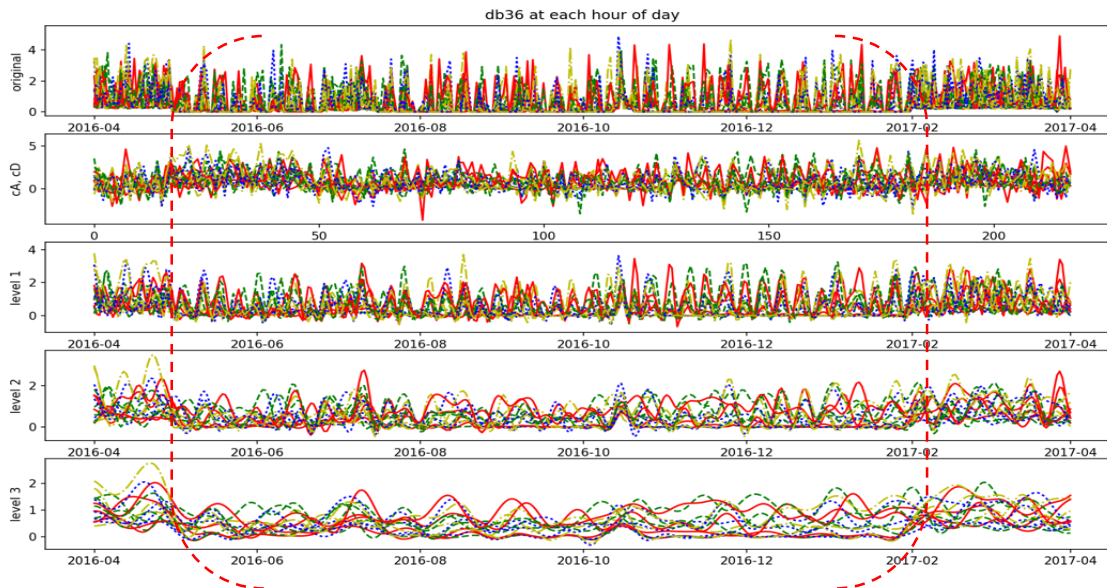


meetodite puhul. Ajatelje mastaap on sarnaselt 1. näitele säiliv 2. meetodi puhul ning funktsiooniti muutuv 1. meetodi rakendamise korral.

### Näide 3

Antud näite jaoks osutus sobivaimateks funktsioonideks (Lisa 1) alusel 15 erinevat kandidaati, kõige tihedama esinemissagedusega on 'db' perekonna funktsioonid. Sarnaselt 2. näitele on *wavelet* funktsioonide valikut erineva tunni näitude puhul keeruline selgitada ning märgatavat mustrit ei ole võimalik tuvastada.

Näite 3 puhul on tegu kõige raskemini eristuva kao piirkonnaga ning kõige keerukamate lähteandmetega üle kogu ajatelje. Esimese meetodi rakendamisel ei ole ühelgi funktsiooni puhul võimalik kao piirkonda eristada ning teise meetodi puhul on kao piirkonna tuvastus enamuse funktsioonide korral väga keeruline. Visuaalse eristumise parandus ei ole antud näite puhul tagatud ning järgneval joonisel on välja toodud näite 3 *wavelet* teisendus kasutades 'db36' funktsiooni, kus kao piirkond on nähtav vaid teise meetodi kolmanda taseme teisendusel. Aegridade paigutus on identne eelnevate näidetega ja kao piirkond on näidatud punase punktiirjoonega.



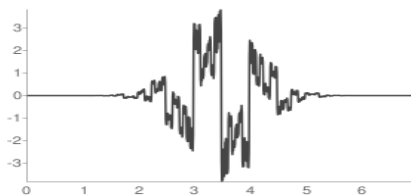
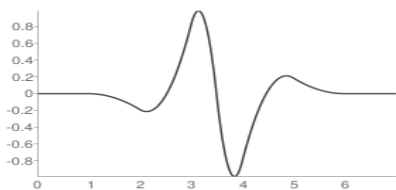
Joonis 25 Näide 3 – 'db36' teisendus

### Näide 5

Lähtudes parimatest *wavelet* funktsioonide valikule, mis on toodud tabelis (Lisa 1) on kokku sobinud 13 erinevat funktsiooni, mis viitab sellele, et aegread on rohkem

ühtlasemad erinevate tundide vaates kui eelneval kahel näitel. Kõige rohkem esineb sarnaselt esimesele näitele 'rbio' ja 'db' perekonnale sobivamaid funktsioone.

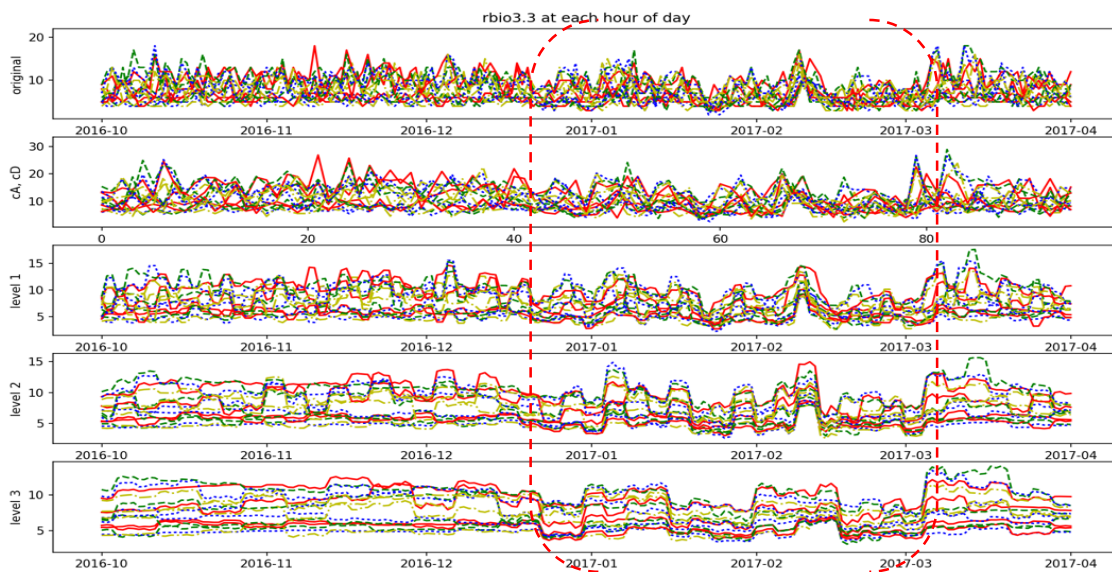
Visuaalsel võrdlemisel selgus, et kõige paremini on toimunud teise meetodi rakendamine, kus on alates teise taseme diskreetsest *wavelet* teisendusest märgatavalt näha kao piirkonna eristumist tavapärasest kadudeta signaalist. Suurim visuaalselt märgatav eristumine tekkis aegridade teisendusel kasutades 'rbio3.3' funktsiooni. Antud funktsioon tekitab kao piirkonnas suuremad sesoonsed muutused kui puhta signaali alal, ning samuti on nähtav ka tavapärasest madalama tasandi esinemine. Järgneval joonisel on toodud funktsiooni 'rbio3.3' kuju.



Joonis 26 Funktsiooni 'rbio3.3' kuju teisendusel (pywavelets, 2008-2017)

Joonis 27 Funktsiooni 'rbio3.3' kuju rekonstrueerimisel (pywavelets, 2008-2017)

Alljärgneval joonisel on kujutatud lähteandmed ja nende diskreetsed *wavelet* teisendused kasutades 'rbior3.3' funktsiooni. Kao piirkond on märgitud punase punktiirjoonega ning esimesel aegreal on kujutatud lähteandmed, teisel real teisendatud kuju kasutades 1. meetodit ning read 3 kuni 5 on 2. meetodi mitme tasandilise *wavelet* teisenduse tulem.



Joonis 28 Näide 5 – 'rbio3.3' teisendus

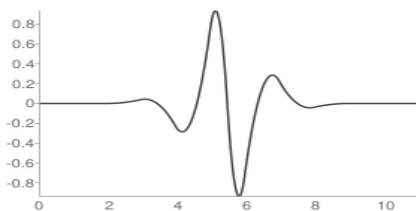
Kokkuvõttena on 5. näite tulemusel meetodite rakendus efektiivne, suurem progress on tagatud 2. meetodi puhul kolmanda taseme teisendusel. Visuaalne eristumine on tagatud ning ajatelje mastaap säilib, seega kao piirkond on ajas hästi lokaliseeritav.

## Näide 6

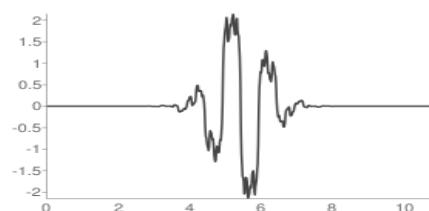
Lähtudes tabelis (Lisa 1) välja toodud *wavelet* funktsioonide valikust sobis antud näite jaoks 18 erinevat *wavelet* funktsiooni. Kõige enam esinenud funktsioonide perekonnad on 'db' ja 'rbio'. Näite lähteandmed omavad kõige lihtsamini eristuvat kao piirkondi, mida on antud juhul kaks. Lisaks peaesmärgile on selle katse eesmärgiks vaadata kuidas toimub andmete teisendus, juhul kui otsitav piirkond on niigi väga hästi nähtav ning vaadeldava kao ajavahemiku konstantselt madala väärtuse ja sesoonsuse poolest. Teiseks on tegemist kahe kao piirkonnaga, millest teine on toimunud väga lühikest aega.

Visuaalse võrdlemise tulemusena on võimalik järeldada, et aegridade teisendatud kuju võib lihtsa näite puhul olla nii halvustav kui ka visuaalselt mitte mingit lisaväärtust andev. Selle põhjuseks on ennekõike antud näite kao piirkonna omapära, kus kogu piirkond on konstantselt null väärtuse juures. Mitmetasandilise teisenduse puhul läheb paljudel katsetel kaduma teine kao piirkond.

Visuaalse hinnangu parima funktsiooni valik lähtus sellest, et säilitada võimalikult hästi mõlema kao lokaliseerimist ajas. Kõige parem on selles olukorras diskreetne *wavelet* teisendus kasutades 'rbio3.5' funktsiooni. Järgneval joonisel on kujutatud selle funktsiooni kuju, mis on esmapilgul identne näites 1 esitatud funktsiooniga 'rbio3.7', kuid sellegipoolest erinev väärtuste detailsuse poolest rekonstrueerimisel.

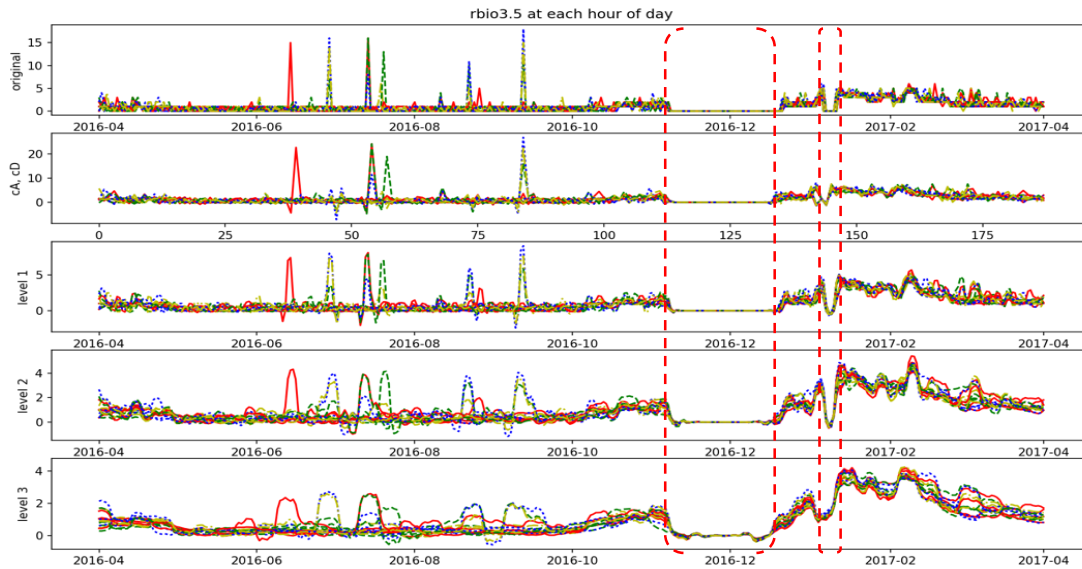


Joonis 29 Funktsiooni 'rbio3.5' kuju teisendusel (pywavelets, 2008-2017)



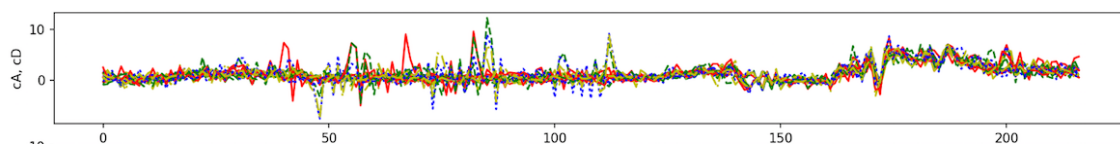
Joonis 30 Funktsiooni 'rbio3.5' kuju teisendusel (pywavelets, 2008-2017)

Alljärgneval joonisel on näha lähteandmeid ning tulemusi kasutades 'rbio3.5' *wavelet* funktsiooni. Kao piirkonnad on näidatud punase punktiirjoonega.



Joonis 31 Näide 6 – ‘rbio3.5’ teisendus

Antud jooniselt on näha, et kõigi meetodite puhul on esimene kao piirkond tuvastatav ning teise puhul on näha, et eristumine muutub halvemaks 1. meetodi ja teise meetodi 3 taseme teisendusel. Teisenduste tulemusena kaotab teine piirkond lokaliseeritavuse ajas ning lähteandmetel nähtavad null väärtused omandavad suuremaid näitajaid. Rakendatud katsetuste juures on oluline välja tuua ka tulemuse näide olukorras, kus kao piirkond on muutunud väga raskesti eristatavaks. See põhjendab, et *wavelet* teisendused ei ole alati efektiivsed selliste andmete töötlemiseks, väga oluline on teada millist tüüpi andmeid on mõistlik teisendada ja milliseid mitte. Järgneval joonisel on toodud antud tulemus rakendades 1. meetodit kasutades ‘db35’ funktsiooni.



Joonis 32 Näide 6 – ‘db35’ teisendus 1. meetodi korral.

Kokkuvõttena on antud näite puhul mõlema meetodi rakenduse tulemused eesmärgi saavutamise seisukohalt suhtelised. Teise meetodi tulemused säilitavad lokaliseeritavuse ajas mitte täielikult (tuginedes teise kao piirkonna tuvastamisele) ning esimese meetodi tulem ei oma erilist väärtust lähteandmete kõrval. Katsetuse tulemuse hinnang on neutraalne.

Diskreetse *wavelet* teisenduste tulemused 5 erinevate näidisandmete kohta on toodud alljärgnevas tabelis.

Näide	Funktsiooni valik	Visuaalne tulemus	Ajatelje mastaabi säilimine	
1	'rbio3.7'	Piirkonna tuvastus tagatud ning visuaalselt parandatud.	Meetod 1	raskendatud
			Meetod 2	tagatud
2	'bior1.1'	Piirkonna tuvastus tagatud, visuaalne hinnang - neutraalne	Meetod 1	raskendatud
			Meetod 2	tagatud
3	'db36'	Piirkonna tuvastus raskendatud, tulemus pigem negatiivne	Meetod 1	raskendatud
			Meetod 2	tagatud
5	'rbio3.3'	Kao piirkonna tuvastus tagatud, tulemus positiivne	Meetod 1	raskendatud
			Meetod 2	tagatud
6	'rbio3.5'	Piirkonna tuvastus tagatud, tulemus neutraalne	Meetod 1	raskendatud
			Meetod 2	tagatud 1 ja 2 taseme teisendusel, 3 tasemel raskendatud

Tabel 4 Visuaalsete katsete tulemuste kokkuvõte

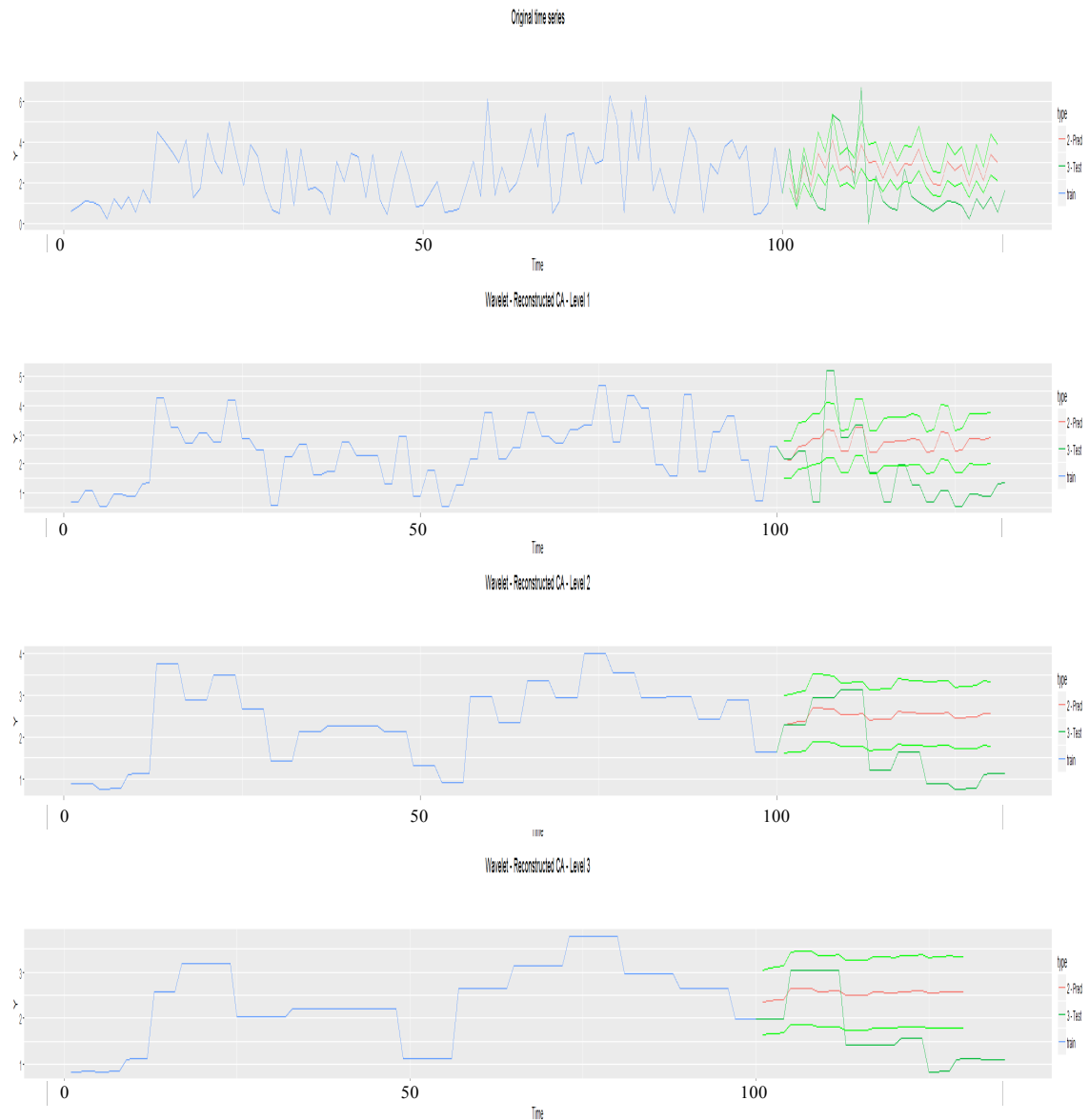
### 5.3.2 Katsetused rakendades ennustusalgoritmi

Ennustusalgoritmi rakendamise katsetel on eesmärgiks fikseerida ennustuskvaliteedi erinevust erinevate sisendandmete korral, milleks on lähteandmed või diskreetse *wavelet* teisenduste tulemusena saadud aegread.

Detailsemaks uurimiseks on valitud näide number 5 ning selle teisenduse tulemus, mis on saadud punktis 5.3.1. Aegread on valitud pikkusega 130 päeva, millest ligikaudselt 80% on ilma kadudeta ning ülejäänud osa on signaali lõpupoole jääv teadaolev kao piirkond. Päeva mahu valik on tingitud sellest, et paljud lahendused ei hoia rohkem kui 3 kuu andmeid süsteemis, vaid eraldi seisvas arhiivis. Ennustusalgoritmi sisendiks kasutatakse õppimiseks 70% andmetest selleks, et jääks võimalus võrrelda ennustuse täpsust 10% ulatuses ning kao piirkonna tekke eristamiseks 20%.

Ennustusalgoritmi rakenduse tulemused on lähteandmete ja *wavelet* teisenduste korral erinevad. Tänu mitmetasandilisele teisendusele, mille tulemusena väheneb signaali müra, muutub ka ennustatud ala järjest täpsemaks reaalsele väärtustele. Kao piirkond jääb ennustatud väärtuste vahemikust välja mis tähendab, et tulemust on võimalik kasutada kao identifitseerimiseks. Alljärgneval joonisel on näha lähteandmete ja teise

meetodi teisenduste tulemusena saadud 3 taseme aegridade ennustatud piirkondi, mis on visualiseeritud ainult ühe ajatunni vaatele (kella 13 ajal). Väärtused 0-100 õpitav ala, 101-110 ennustuse täpsuse võrdlus puhta signaaliga, 111-130 kao piirkond.



Joonis 33 Näide 1 - ennustuse võrdlus lähteandmete ja 3 tasemelise diskreetse *wavelet* teisenduste vahel

### 5.3.3 Pideva *wavelet* teisenduse lahenduse realiseerimine

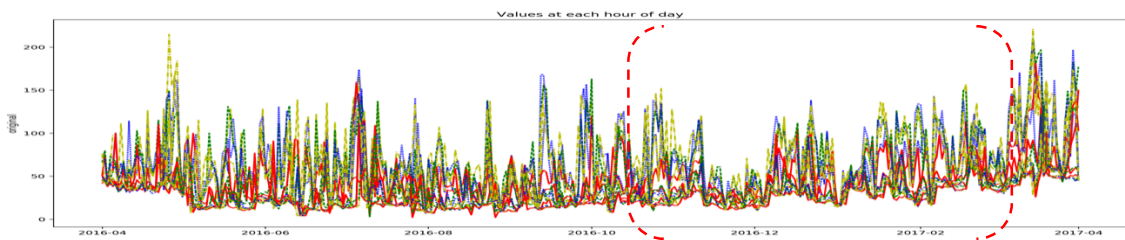
Lahenduse realisatsiooni jätkamisprotsesside järjestus:

1. Pideva *wavelet* teisenduse rakendamine nelja erineva *wavelet* funktsiooniga
2. Visuaalsete tulemuste eksportimine

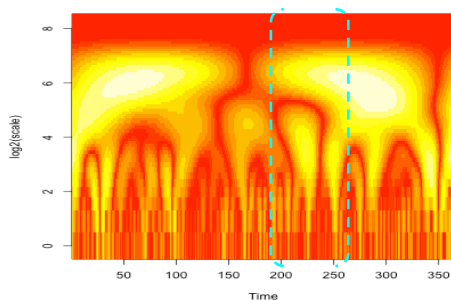
## Näide 4

Katsetuste tulemusena on saadud 96 (4 funktsiooni igal 24'l tunnil) erinevat visuaali pideva *wavelet* teisenduste rakendustest. *Wavelet* funktsioonid mida rakendati on järgnevad 'haar', 'gaussian1', 'gaussian2' ja 'morlet'. Kõigi 96 tulemuse ülevaatamisel ei olnud võimalik märgata teada olevas kao piirkonnas märgatavaid erisusi muudest piirkondadest. Antud meetodit võib antud näite puhul lugeda kasutuskõlbmatuks, ehk pideva *wavelet* teisenduse rakendamine ei anna soovitud tulemust juhul kui lähteandmetes ei ole võimalik eristada kao piirkonda. Lähteandmetele sageduse domeeni lisamine ei too soovitud tulemusi.

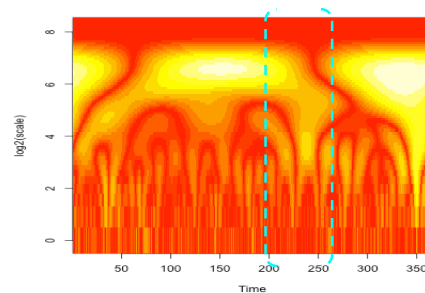
Järgneval joonisel on toodud ülevaade lähteandmete signalidest ning pideva *wavelet* teisenduse tulemustest kõigi nelja funktsiooni korral. Kao piirkond on märgitud punase punktiirjoonega lähteandmetel ning helesinise joonega teisenduste visualisatsioonidel.



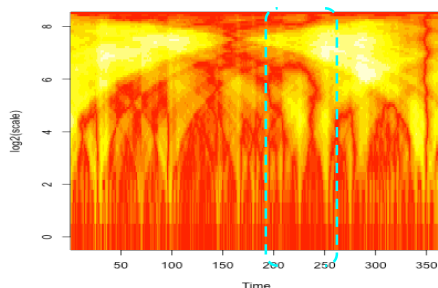
Joonis 34 Näide 4 lähteandmed 1 aastasel ajavahemikul



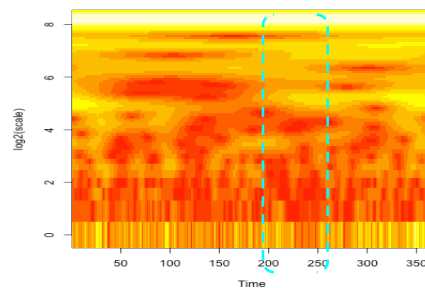
Joonis 35 Näide 4 - 12. tunni näidud 'gaussian1'



Joonis 36 Näide 4 - 12. tunni näidud 'gaussian2'



Joonis 37 Näide 4 - 12. tunni näidud 'haar'



Joonis 38 Näide 4 - 12. tunni näidud 'morlet'

## 6 Tulemus ja diskussioon

Katsete tulemustest on näha, et diskreetsete *wavelet* teisenduste kasutamine aegridade lihtsustamiseks on õigustatud üksikutel juhtudel. Kao piirkonna eristumise tagamine on rangelt seotud kao aegridade lõigu omadustest ning erinemisest kadudeta aegridadest. Keeruliseks teeb diskreetsete *wavelet* teisenduste kasutamist sobiva funktsiooni valimine, kuna erinevate kao piirkondade tuvastamiseks on paremad erinevad funktsioonid. Nende valik on väga lai nagu seda on ka näidisandmete sesoonsuse, trendi ja keskmiste väärtuste variatsioon. Paremaks andmeanalüüsiks on kasulik filtreerida funktsioonide valikut lähtuvalt otsitava kao piirkonna iseloomule. Katsetuste tulemustele tuginedes on kõige sagedamini sobiv 'rbio' tüüpi *wavelet* funktsioonide perekond. Funktsioonide valik on oluline tagamaks vähimat andmete kadu teisenduste rakendamiste korral ning aegridade profiili muutuste tabamiseks.

Parema visuaalse tuvastatavuse tagamiseks teatud ajaperioodi jaoks on tarvis teha mitmetasandilisi teisendusi. Teisenduste tasemete arv sõltub ajaperioodi pikkusest ning andmete keerukusest. Näiteks aasta pikkuse perioodi (~365 aja-väärtuse paari jaoks) oli katsete korral piisav 2-3 taseme teisenduse tulek. Lähtuvalt sellest on 2. meetodi kasutamine eesmärgi saavutamiseks efektiivsem. Ajatelje mastaabi säilitamise kriteeriumi puhul on samuti õigustatud teise meetodi kasutamine. Esimese meetodi puhul on tarvis jälgida aegridade pikkuse muutusi ning seejärel kalkuleerida õiged ajahetke väärtused, mis on tekitab nõudluse lisaprotsesside ning täiendavate andmete varundamise järele. Täpsemal teisendatud aegridade mastaabi muutuste uurimisel võib selguda ka täpne muutumise koefitsient, mis on oletatavasti seotud lähteandmete mahust ning *wavelet* funktsiooni valikust. Kuna töö katsetuste eesmärgiks oli tagada aja mastaabi võrdväarsus lähteandmetega, siis nende täpsustuste välja selgitamine ei olnud antud töö skoobis.

Diskreetsete *wavelet* teisenduste kasutamine kadude visuaalse tuvastamise eesmärgil on õigustatud ainult õigete meetodite ja funktsioonide valiku korral. Valiku tegemine



nõuab suuremate katsetuste tegemist, mis nõuab väga palju aega ja pingutust. Kõige paremad tulemused on saavutatavad järskude muudatuste identifitseerimisel.

Lisaks visuaalse poole katsetustele on proovitud teha katseid ennustusalgoritmiga. Saadud tulemused on väga erinevad ning siinkohal on *wavelet* teisenduste kasu raske hinnata, sest mõnedel juhtudel on ennustatud ala korrektsem lähteandmete korral ning teiste puhul teisendatud andmete korral. Kindlasti peab siinkohal mainima, et ennustusalgoritmi valik ja seadistus oli konstantne, mis ei ole kõige parem viis erineva iseloomuga andmete ennustamise tegemisel, sest diskreetseid *wavelet* teisendusi läbinud andmed omavad vähem hägusamat sesoonsuse omadust. Edaspidiseks rakendamiseks on kasulik otsida või välja töötada ennustusalgoritm, mis teevad parimaid tulemusi valitud *wavelet* lähte funktsiooni ennustamiseks, sest reaalsete andmete teisendatud kujud omandavad iga taseme teisenduse järel aina rohkem lähtefunktsiooni omadusi. Samuti on korrektsemate ennustuste tegemiseks tarvis edaspidi katsetada erinevate sisendandmete mahu valikutega.

Pideva *wavelet* teisenduse kasutamine oli rakendatud katsetuse eesmärgil leida visuaalselt nähtamatu kao piirkond. Nelja funktsiooni rakendamine ei andnud oodatud tulemust. Katsetuste nurjumise põhjuseks on lähteandmete korrapäratus, keerukus ning paljude mitte statsionaarsetele aegridadele iseloomulike omaduste olemasolu.

Üldine hinnang *wavelet*'ite tasuvusele antud kadude probleemi lahendamise juures on suhteline. Antud töös kasutatud näideteks olid väga keeruliste ja erinevate iseloomudega aegread, mis olid suureks ajakulu tekitajaks. Kui kasutatud aegridadele oleks lisaks kasutada ka muud sündmused, mida arvestid fikseerida võimaldavad, saaks tulemusele anda suuremat väärtust. Samuti mängivad elektri valdkonnas tarbimisnäitude juures suurt rolli ka paljud teised muutujad, milleks on globaalsed muutused, mis on tingitud ilmast ja erinevatest sündmustest, mis teadaolevalt mõjutavad tarbimist. Selliste andmete varundamine ja taaskasutamine aitaks identifitseerida tarbimise muutuste normaalsust aegridades ning omakorda mõjutada teisenduste ja ennustuste täpsust.

## 7 Kokkuvõte

Teema valik tugineb hüpoteesil, et *wavelet* teisendused aitavad dekomponeerida tarkade elektriarvestite produtseeritud tarbimisandmeid selliselt, et tulemuses eristuvad erinevad kadude mustrid paremini ning dekomponeeritud andmed on sobivad sisendiks automatiseeritud kadude tuvastussüsteemi loomiseks. Automatiseeritud kadude tuvastamiseks on oluline lihtsustada sisendandmeid, säilitades lähteandmete ajaväärtuste täpsuse. Töö eesmärgiks oli analüüsida erinevaid *wavelet* teisenduste kasutusvõimalusi, et teisendada tarbimisandmete aegread kujule, mis lihtsustaks kadude piirkonna tuvastamist visuaalselt. Töö skoobis oli analüüsida *wavelet* teisenduste võimalusi tuginedes reaalse elu näidetele ning parandada kao piirkondade eristumist visuaalselt, automatiseerimist siin töös ei käsitletud ning see jääb edasiseks uurimiseks. Kasutatud näideteks olid erinevad elektritarbimise aegread, millel olid analüüsi käigus tuvastatud kaod.

Selleks, et jõuda antud eesmärgini vaadati üle *wavelet*'ite olemus, erinevad rakendusala, uuriti võimalikke aegridade omadusi ning toodi välja võimalikud meetodid mida rakendada. Enne nende rakendust tehti lähteandmete analüüs, käsitleti erinevaid võimalusi kuidas aegridadele on võimalik läheneda ning tehti valik realiseerimiseks kasutatava programmeerimiskeele ja teegi osas.

Valitud diskreetse *wavelet* teisenduste meetodite rakendamisel viiel erineval näitel, oli märgatav kadude eristumise paranemine kahe näite puhul. Lähteandmete keerukus ja kao piirkondade iseloomude mitmekülgus on põhjuseks, miks visuaalse kao piirkonna eristumine ei olnud kõigi *wavelet* teisenduste tulemusena parandatud. Diskreetsete *wavelet* teisenduste tulemuste analüüsil selgus, et erinevate *wavelet* funktsioonide valik mõjutab tulemuse kvaliteeti erinevalt. Sobivate funktsioonide valik aitab tagada paremat tulemust kao piirkonna väljatoomise puhul. Mitmetasandiliste teisenduste kasutamine on taganud visuaalse kao piirkonna eristamise parandamise. Teisendatud andmete kasutamine ennustusalgoritmi rakendamisel on parandanud täpsust mitte kõikidel juhtudel, mis tähendab, et ennustusalgoritmi kasutamine ja valik nõuab paremat seadistust lähtuvalt *wavelet* funktsiooni valikule.

Pideva *wavelet* teisenduste kasutamine ei tõstnud esile visuaalselt mitte eristuvat kao piirkonda, mis tähendab, et ainult elektrinäitude aeGRIDade põhjal ei ole võimalik kõiki kao piirkondi tuvastada. Kao piirkondade efektiivseimaks tuvastamiseks on vajalik kasutada ka teiste andmete kasutamist, mida targad arvestid võimaldavad fikseerida, näiteks vigade ja katkemiste teated.

Sissejuhatuses püstitatud eesmärgid said täidetud mitte kõikide katsetuste korral. Töö käigus selgus, et kõige kasulikum *wavelet* teisenduste rakendus on mitmetasandilise diskreetse *wavelet* teisenduse kasutamine. Kao tuvastamine tuginedes ainult aeGRIDadele on raskendatud, sest elektri tarbimise muutused võivad olla tingitud ka ilmast ja pühadest, mida antud töös ei kasutatud. Pikemas perspektiivis on *wavelet* teisenduste kasutamine õigustatud väga detailse seadistamise korral ning saadud tulemus ei saa täielikult asendada lähteandmeid. *Wavelet* teisenduste tulemused on soovituslikud kasutamaks toeks ennustusalgoritmi tulemustele, mis on saadud lähteandmetest.

## Kasutatud kirjandus

- [1] Addison, P. S., 2005. *Wavelet transforms and the ECG: a review*. [Võrgumaterjal]  
Available at:  
[http://www.robots.ox.ac.uk/~gari/teaching/cdt/A3/readings/ECG/Addison\\_0967-3334\\_26\\_5\\_R01.pdf](http://www.robots.ox.ac.uk/~gari/teaching/cdt/A3/readings/ECG/Addison_0967-3334_26_5_R01.pdf)  
[Kasutatud Mai 2017].
- [2] Amit Konar, D. B., 2017. *Time-Series Prediction and Applications: A Machine Intelligence Approach*. s.l.:s.n.
- [3] Antoniadis, A. & Sapatinas, T., 2003. Wavelet methods for continuous-time prediction using Hilbert-valued autoregressive processes. rmt.: *Journal of Multivariate Analysis*. s.l.:s.n., pp. 133-158.
- [4] Brownlee, J., 2016. *What Is Time Series Forecasting?*. [Võrgumaterjal]  
Available at: <http://machinelearningmastery.com/time-series-forecasting/>  
[Kasutatud 8 Mai 2017].
- [5] Chaovalit, P., Aryya, G., George, K. & Zhiyuan, C., 2011. *Discrete Wavelet Transform-Based Time Series Analysis and Mining*. [Võrgumaterjal]  
Available at: <http://mlsurveys.s3.amazonaws.com/102.pdf>  
[Kasutatud 10 Mai 2017].
- [6] Elektrilevi OÜ, 2013. *Kauglugemine*. [Võrgumaterjal]  
Available at: <https://www.elektrilevi.ee/kauglugemine>  
[Kasutatud 4 Mai 2017].
- [7] Elektrilevi OÜ, 2017. *Elektrivarguste kohtulahendid*. [Võrgumaterjal]  
Available at: <https://www.elektrilevi.ee/elektrivarguste-kohtulahendid>  
[Kasutatud 9 Mai 2017].
- [8] Elster & McCullough, J., 2011. *Deterrent and detection of smart grid meter tampering and theft of electricity, water, or gas*. [Võrgumaterjal]  
Available at: [http://publicsde.regie-energie.qc.ca/projets/34/DocPrj/R-3770-2011-C-ROE%C3%89-0058-PREUVE-AUTRE-2011\\_12\\_16.pdf](http://publicsde.regie-energie.qc.ca/projets/34/DocPrj/R-3770-2011-C-ROE%C3%89-0058-PREUVE-AUTRE-2011_12_16.pdf)  
[Kasutatud 5 Mai 2017].
- [9] Fryzlewicz, P., 2003. *FORECASTING NON-STATIONARY TIME SERIES BY WAVELET PROCESS MODELLING*. [Võrgumaterjal]  
Available at: [http://www.ism.ac.jp/editsec/aism/pdf/055\\_4\\_0737.pdf](http://www.ism.ac.jp/editsec/aism/pdf/055_4_0737.pdf)  
[Kasutatud 10 mai 2017].
- [10] Goreva T.S., P. N., 2015. *ПРОГРАММНОЕ И АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЯ АППАРАТНЫХ КОМПЛЕКСОВ ПОВЫШЕНИЯ КАЧЕСТВА ПИТАЮЩЕЙ ЭЛЕКТРОЭНЕРГИИ В РАСПРЕДЕЛИТЕЛЬНЫХ СЕТЯХ*. Venemaа: Академия Естествознания.
- [11] Joala, V., 2011. *Aegridade analüüsi ajaloost*. [Võrgumaterjal]  
Available at:

- <https://aegread.wikispaces.com/8.+Aegridade+anal%C3%BC%C3%BCsi+ajaloost>  
[Kasutatud 7 Mai 2017].
- [12] Köhler, T. & Lorenz, D., 2005. *A Comparison of Denoising Methods for One Dimensional Time Series*. [Võrgumaterjal]  
Available at:  
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.565.1807&rep=rep1&type=pdf>  
[Kasutatud mai 2017].
- [13] Kalle Remm, J. R. A. K., 2012. *Ruumiliste loodusandmete statistiline analüüs*. Tartu: s.n.
- [14] Klapetek, P., Nečas, D. & Anderson, C., 2004–2007, 2009–2016. *Wavelet Transform*. [Võrgumaterjal]  
Available at: <http://gwyddion.net/documentation/user-guide-en/wavelet-transform.html>  
[Kasutatud 11 Mai 2017].
- [15] Kose Vesi, 2015. *Kaugloetavad veemõõtjad*. [Võrgumaterjal]  
Available at: <http://www.kosevesi.ee/kliendile/veearvesti-ja-veemootesolm/kaugloetavad-veemootjad/>  
[Kasutatud 4 Mai 2017].
- [16] Mirotin, E. A., 2011. *Aegridade ennustamine kasutades jaotamata wavelet teisendusi ning ühesuunalisi filtreid*. [Võrgumaterjal]  
Available at: <http://elib.bsu.by/bitstream/123456789/14093/1/127-132.pdf>  
[Kasutatud 10 Mai 2017].
- [17] Moga, M. & Victor Dan Moga, G. M., 2005. *Continuous Wavelet Transform in ECG Analysis. A Concept or Clinical Uses*. [Võrgumaterjal]  
Available at:  
[https://www.researchgate.net/publication/228546367\\_Continuous\\_Wavelet\\_Transform\\_in\\_ECG\\_Analysis\\_A\\_Concept\\_or\\_Clinical\\_Uses](https://www.researchgate.net/publication/228546367_Continuous_Wavelet_Transform_in_ECG_Analysis_A_Concept_or_Clinical_Uses)  
[Kasutatud Mai 2017].
- [18] Ogorodov, A., 2009. [Võrgumaterjal]  
Available at: [http://fetmag.mrsu.ru/2009-3/pdf/Financial\\_transient\\_series.pdf](http://fetmag.mrsu.ru/2009-3/pdf/Financial_transient_series.pdf)  
[Kasutatud 3 Mai 2017].
- [19] Piatetsky, G., 2014. *Four main languages for Analytics, Data Mining, Data Science*. [Võrgumaterjal]  
Available at: <http://www.kdnuggets.com/2014/08/four-main-languages-analytics-data-mining-data-science.html>  
[Kasutatud 1 mai 2017].
- [20] Pointer, I., 2016. *Which freaking big data programming language should I use?*. [Võrgumaterjal]  
Available at: <http://www.infoworld.com/article/3049672/application-development/which-freaking-big-data-programming-language-should-i-use.html>  
[Kasutatud 1 mai 2017].
- [21] Polikar, R., 1996. *The wavelet tutorial Second edition part 1*. [Võrgumaterjal]  
Available at: <http://web.iitd.ac.in/~sumeet/WaveletTutorial.pdf>  
[Kasutatud 11 Mai 2017].
- [22] pywavelets, 2008-2017. *Wavelet Reverse biorthogonal 3.7 (rbio3.7)*. [Võrgumaterjal]  
Available at: <http://wavelets.pybytes.com/family/rbio/>  
[Kasutatud 14 mai 2017].

- [23] Rob J Hyndman, G. A., 2012. *Time series components*. [Võrgumaterjal]  
Available at: <https://www.otexts.org/fpp/6/1>  
[Kasutatud 8 Mai 2017].
- [24] Rudi, H., 2016. *Alates uuest aastast peavad kõik elektriarvestid olema kaugloetavad*. [Võrgumaterjal]  
Available at: <http://tarbija24.postimees.ee/3900247/alates-uuest-aastast-peavad-koik-elektriarvestid-olema-kaugloetavad>  
[Kasutatud 9 Mai 2017].
- [25] SINGH, R., 2013. *Wavelet Transforms In Time Series Analysis*. [Võrgumaterjal]  
Available at: <http://iaud.aydin.edu.tr/makaleler/ciltxsayix/singh.pdf>  
[Kasutatud 10 Mai 2017].
- [26] Srivastava, T., 2015. *A Complete Tutorial on Time Series Modeling in R*. [Võrgumaterjal]  
Available at: <https://www.analyticsvidhya.com/blog/2015/12/complete-tutorial-time-series-modeling/>  
[Kasutatud 8 Mai 2017].
- [27] Sterling, L. S., 2009. *The Art of Agent-Oriented Modeling*. London: The MIT Press.
- [28] Wikipedia, 2017. *Fourier' teisendus*. [Võrgumaterjal]  
Available at: [https://et.wikipedia.org/wiki/Fourier%27\\_teisendus](https://et.wikipedia.org/wiki/Fourier%27_teisendus)  
[Kasutatud 8 Mai 2017].
- [29] Wikipedia, 2017. *Mean squared error*. [Võrgumaterjal]  
Available at: [https://en.wikipedia.org/wiki/Mean\\_squared\\_error](https://en.wikipedia.org/wiki/Mean_squared_error)  
[Kasutatud 9 mai 2017].
- [30] Wikipedia, 2017. *Wavelets*. [Võrgumaterjal]  
Available at:  
<https://ru.wikipedia.org/wiki/%D0%92%D0%B5%D0%B9%D0%B2%D0%BB%D0%B5%D1%82#.D0.A2.D0.B5.D0.BE.D1.80.D0.B8.D1.8F.D0.B2.D0.B5.D0.B9.D0.B2.D0.BB.D0.B5.D1.82.D0.BE.D0.B2>  
[Kasutatud 6 Mai 2017].
- [31] Zorin, D., 1997. *Topics in Computer Graphics, Mathematical Models for Computer Graphics*. [Võrgumaterjal]  
Available at: [http://cva.stanford.edu/classes/ee482a/docs/lect01\\_sample.pdf](http://cva.stanford.edu/classes/ee482a/docs/lect01_sample.pdf)  
[Kasutatud 2017].

**Lisa 1 – Parimad *wavelet* funktsioonide ühtimised näidete aegridadega MSE alusel ühe aasta pikkusel perioodil**

Kellaeg	Näide 1	Näide 2	Näide 3	Näide 5	Näide 6
1	db2	db36	rbio2.2	rbio2.2	rbio2.4
2	db2	bior1.1	db2	rbio2.2	db4
3	rbio5.5	db19	rbio6.8	rbio2.2	sym5
4	db3	db27	rbio2.2	rbio2.2	db17
5	db2	db35	db3	db5	rbio2.4
6	db2	db5	rbio2.2	rbio2.2	db18
7	db2	db5	rbio2.2	rbio6.8	db33
8	rbio2.4	db5	sym4	rbio2.2	sym4
9	rbio3.7	db5	coif1	rbio3.9	db5
10	rbio3.7	db11	db24	rbio2.2	db25
11	sym7	db17	db36	bior1.1	rbio2.2
12	rbio3.7	sym2	db7	db5	rbio3.5
13	rbio3.7	db34	db3	rbio3.3	rbio2.2
14	rbio6.8	db4	sym2	coif1	rbio2.2
15	rbio3.7	db7	sym7	rbio5.5	db2
16	sym4	sym5	db35	db3	db37
17	rbio6.8	rbio2.4	db34	db3	rbio2.2
18	sym11	rbio2.4	db2	coif1	coif1
19	sym15	db33	db2	bior1.1	rbio2.2
20	bior1.1	db3	db2	db20	rbio2.2
21	rbio2.4	db2	sym5	db23	bior2.4
22	db17	sym3	db20	bior1.1	db7
23	db17	db11	db2	db11	db35
24	rbio2.2	db5	db2	db2	db3