

DATAANALYS OCH AVANCERADE ALGORITMER

RAPPORT 2018:537



Dataanalys och avancerade algoritmer

Möjligheter med utökad mätinfrastruktur

MATTIAS PERSSON

CLAES SANDELS

ANDREAS NILSSON

ISBN 978-91-7673-537-4 | © Energiforsk oktober 2018

Energiforsk AB | Telefon: 08-677 25 30 | E-post: kontakt@energiforsk.se | www.energiforsk.se

Förord

Projektet *Dataanalys och avancerade algoritmer – Möjligheter med mätinfrastruktur* har granskat hur distributionsägare kan använda informationen från ett elnät med en utökad mätinfrastruktur och fler mätparametrar i energimätare för att effektivisera sin drift. Projektet har verifierat att högre mätkvalité ger ökad analysmöjlighet. Kunskap från stamnät och tidigare resultat har överförts till distributionsnivå.

Projektet, som tillhörde programmet Smarta Elnät, har:

- Påvisat nytta med utökade mätparametrar hos energimätare i distributionsnät med nätstationsmätning.
- Dokumenterat nytta med god mätkvalité sett till förbättrad lokalisering av icke tekniska förluster.
- Specificerat krav på nästa generations mätare för att kunna möjliggöra föreslagna analysverktyg.
- Undersökt förändringar i kostnad för elnätsägaren vid ökad mätnytta, d.v.s. mätare med fler mätvärden (på grund av fler mätparametrar och ökad samplingsfrekvens).

Mattias Persson på RISE har varit projektledare för projektet. Energiforsk vill framföra ett stort tack till referensgruppen, som bidragit med sina erfarenheter och synpunkter under projektets gång. Referensgruppen har bestått av:

- Anders Lindskog, RISE
- David Westerlund, Rejlers Embriq
- Lars Olsson, Seniorit AB

Programmet Smarta Elnäts programstyrelse, som initierat, följt upp och godkänt projektet, består av följande ledamöter:

- Peter Söderström, Vattenfall Eldistribution AB (ordförande)
- Torbjörn Solver, Mälarenergi AB (vice ordförande)
- Mira Rosengren Keijser, Svenska kraftnät
- Patrik Björnström, Sveriges ingenjörer (Miljöfonden)
- Kristina Nilsson, Ellevio AB
- Björn Ållebrand, Trafikverket
- Ferruccio Vuinovich, Göteborg Energi AB
- Per-Olov Lundqvist, Sandviken Energi AB
- Claes Wedén, ABB AB
- Daniel Köbi, Jämtkraft AB
- Hannes Schmied, NCC AB
- Mats Bergström, Umeå Energi Elnät AB
- Magnus Sjunnesson, Öresundskraft AB
- Henric Johansson, Jönköping Energi AB
- David Håkansson, Borås Elnät AB
- Peter Addicksson, HEM AB
- Anders Fredriksson, Energiföretagen Sverige (adjungerad)

Följande bolag har deltagit som intressenter till projektet. Energiforsk framför ett stort tack till samtliga företag för värdefulla insatser.

Vattenfall Eldistribution AB
Ellevio AB
Svenska kraftnät
Göteborg Energi AB
Elinorr ekonomisk förening
Mälarenergi Elnät AB
Krafttringen Nät AB
Jämtkraft Elnät AB
Umeå Energi Elnät AB
Öresundskraft AB
Karlstads Elnät AB
Jönköping Energi Nät AB
Halmstad Energi & Miljö Nät
AB Falu Elnät AB
Borås Elnät AB
AB Borlänge Energi
C4 Elnät AB
Luleå Energi AB
ABB AB
Akademiska Hus
NKT Cables AB
NCC Construction Sverige AB
Trafikverket
Storuman Energi AB

Stockholm i september 2018

Susanne Stjernfeldt
Energiforsk AB
Programområde Elnät, Vindkraft och Solel

Här redovisas resultat och slutsatser från ett projekt inom ett forskningsprogram som drivs av Energiforsk. Det är rapportförfattaren/-författarna som ansvarar för innehållet och publiceringen innebär inte att Energiforsk har tagit ställning till innehållet.

Sammanfattning

Incitament för att nätbolagen skall effektivisera sin drift föreslås av Energimarknadsinspektionen.

Sett till effektivitet har 54 av 159 nätbolag högre nätförluster än 4% på årsbasis. Om dessa 54 hade gjort förbättringar i sina elnät och nått ned till en nätförlust på 4%, så skulle det innebära en minskning på totalt 143 GWh/år. Antas en kostnad på 50 öre/kWh motsvarar det 72 miljoner kr/år. Detta kan vara dels onödig förlust att behöva betala i anslutningspunkten men också en saknad inkomst för distributionselnätsägaren.

Enkät svar från elnätsbolag visar vilka framtida utmaningar som man ser är allt från ökade energimätarkrav, förstärkningar och att värdesäkra nätet. Enkät svaren visar att det finns stor spridning hur nätbolagen hanterar sina nätförluster, hur man i nuläget hanterar icke-tekniska förluster (ITF) och dess lokalisering och identifiering. Jämfört med kostnaden för nya generationens mätare kommer lagringskostnader för extra mätvärden och ökad mätupplösning inte vara den stora kostnadsdrivaren.

I denna rapport utvärderas korrelation, effektflödesanalys och maskininlärning för att detektera och lokalisera olika typer av ITF. Med hjälp av maskininlärning kan förluster lokaliseras utan perfekt kunskap om elnätets uppbyggnad så länge dessa förluster följer förbrukningsmönstret och perioder utan ITF finns för upplärning av algoritmen. För slumpmässiga förluster har 3 olika metoder utvecklats och utvärderats (SiM, K:SE och K:V) som visar lovande resultat när det gäller lokalisering och detektering med hjälp av spänningsvärden hos energimätare. Upplösningen hos energimätvärdena och dess påverkan på lokaliseringen har även studerats.

Summary

An incentive for distribution companies to improve the efficiency of their operation is being suggested by Energimarknadsinspektionen.

In terms of efficiency, 54 out of 159 network companies have higher distribution losses than 4% on an annual basis. If these 54 had made improvements in their power grids and reached an energy loss of 4%, this would mean a reduction of 143 GWh / year overall. Assuming a cost of 50 öre / kWh this would be equivalent to 72 million kr / year. This may be unnecessary loss of having to pay at the connection point, but also a missing income for the distribution network owner.

Surveys from electricity companies show what future challenges are ranging from increased energy metering requirements and reinforcements. The questionnaire shows that there is a large spread of how network companies manage their network losses, how to handle non-technical losses (ITF) and how they go about to locate and identify them. Compared to the cost of the new generation meters, storage costs for additional metrics and increased measurement resolution will not be the major cost driver.

In this report, correlation, power flow analysis and machine learning are evaluated in order to detect and locate different types of ITF. With the help of machine learning, losses can be located without perfect knowledge of the power grid structure as long as these losses follow the usage pattern, and periods without ITF are available for the learning of the machine learning algorithm. For random losses, 3 different methods have been evaluated (SiM, K: SE and K: V) that show promising results with regard to localizing and detecting ITF using voltage values of energy meters. The resolution of the energy meters and its influence on the possibility to localize a ITF has also been studied.

Innehåll

1	Introduktion	9
1.1	Syfte	9
1.2	Mål	10
1.3	Avgränsingar	10
1.4	Struktur	10
2	Resultat från enkät till elnätsbolag: Nya mätare, kostnader och icke-tekniska förluster	11
2.1	Inledning	11
2.1.1	Kapitelsammanfattning	11
2.2	Elnätsbolagens utmaningar	13
2.3	Icke-tekniska förluster - nuläge	14
2.3.1	Generellt	14
2.3.2	Mätinfrastruktur nuläge	14
2.3.3	Uppföljning av nätförluster	15
2.3.4	Lokalisering - räknexempel	17
2.4	Tackla icke-tekniska förluster med en utökad mätinfrastruktur	18
2.4.1	Behov av utökad mätinfrastruktur	18
2.4.2	Behov av lokalisering	19
2.4.3	Nyttor	20
2.4.4	Kostnader, hinder	20
3	Lokalisering av icke-tekniska förluster med systemkunskap	22
3.1	Metod	22
3.1.1	Avgränsningar	22
3.1.2	Modellering av icke-teknisk förlust	23
3.1.3	Nätområden	24
3.2	Fallstudie från Lokalisering av icke-teknisk förlust	25
3.2.1	Lokalisering av icke-tekniska förluster	27
3.2.2	Lokalisering i större expelelnät	30
3.2.3	Påverkan från varierande mätkvalitet	33
4	Sammanställning: Maskininlärning för lokalisering av icke-tekniska förluster	36
4.1	Vad är maskininlärning?	36
4.2	Hur funkar det tekniskt?	37
4.3	Vilka metoder ingår i ML:s verktygslåda?	38
4.4	Hur kan man använda AI för att identifiera och lokalisera icke-tekniska förluster i distributionsnät?	39
4.5	Vad är potentialen och fördelen med att använda ML i jämförelse med andra metoder (exempelvis load flow)?	39
4.6	Vad är utmaningarna med att använda ML i området?	42

5	Exempel: Maskininlärning för lokalisering av icke-tekniska förluster	45
5.1	Metod	45
5.2	Fallstudie: Ett litet elnät med stationsmätning	48
5.3	Resultat	49
5.3.1	Basfallet: ingen icke-teknisk förlust	49
5.3.2	Olika nivåer av elstöld	51
5.3.3	Systematisk validering	54
5.3.4	Korrelerade fel	54
5.3.5	Slumpmässiga ITF	56
6	Slutsatser och diskussion	57
7	Citerade arbeten	59
Bilaga A:	Intervjufrågor till elnätsföretagen	60

1 Introduktion

Den ekonomiska avskrivningen av energimätare är snart klar för de mätare som installerades i samband med den första mätaruutrullningen, samtidigt finns nya förslag på krav för energimätning från Energimarknadsinspektionen (Energimarknadsinspektionen, 2015).

I distributionsnät har det visat sig att med hjälp av timvärden från energimätare och kunskap om elnätparametrar går att identifiera olika problem i distributionsnät. Detta hanterades i (Hagmar & Lindskog, Bedömning av nätstatus baserad på dataanalys och avancerade algoritmer (2017:443), 2017) där mätningen i distributionsnätet kompletterades med nätstationsmätning för att identifiera fel som icke-tekniska förluster (p.g.a. linjära, korrelerade, konstanta eller slumpmässiga-fel) men också att lokalisera dessa fel. Begränsningar i mätkvalité (tidsynkronisering, mätupplösning och avsaknad av data) identifierades och resulterade i delvis svårtolkade resultat. Med en bättre mätkvalité och fler mätparametrar, såsom spänning hos slutkund, borde kunna förbättra analysen avsevärt.

I Sverige diskuteras nu att ställa nya krav på slutkundsmätning av el. Dessa krav kommer sannolikt att påverka nästa generations mätare som kommer att installeras under de närmaste åren. I samband med detta är det intressant att om möjligt undanröja de problem som identifierats i (Hagmar & Lindskog, Bedömning av nätstatus baserad på dataanalys och avancerade algoritmer (2017:443), 2017). Att identifiera och lokalisera problem i elnäten är av stort värde för elnätföretagen. Därför är det intressant att undersöka om de identifierade problemen undanröjs då begränsningarna i mätkvalité tas bort. Dessutom är det intressant att undersöka hur utökade krav samt mätparametrar påverkar kostnaden om det görs i samband med en ny utrullning av energimätare.

1.1 SYFTE

Att verifiera och utveckla tidigare framtagna analysverktyg i ett distributionselnät med hög mätkvalité, utvärdera kostnader samt tillgodose kunskapsspridning och dokumentation av projektet.

Huvudsyftet med projektet är att skapa ökad kunskap om hur mätdata från energimätare kan användas för att underlätta och minska kostnaderna i den dagliga driften i distributionsnäten. Tidigare projekt har identifierat begränsningar i mätinfrastrukturen som försvårar vissa typer av analyser. Detta projekt syftar till att analysera nyttan av god mätkvalité för distributionselnätsägaren, och som en följd av detta att rekommendera vilka krav som kan eller bör ställas på mätarinfrastruktur för att kunna identifiera och lokalisera fel.

För att lösa ovanstående mål syftar projektet att:

Dokumentera och visa vad god mätkvalité och extra mätparametrar hos energimätare kan leverera för mätnytta i distributionsnätet.

Beskriva nödvändig kvalitet på upplösning och tidssynkronisering hos energimätare för att kunna utföra beskrivna matematiska algoritmer.

- Dokumentera och visa hur algoritmer kan sortera ut och lokalisera icke-tekniska förluster med hjälp av extra mätvärden, sett till de som primärt används i dagsläget, energimätare.
- Utvärdera eventuell förändring i kostnad för elnätbolagen för extra funktionalitet som krävs för de analyser och verktyg gällande mätkvalitet och extra mätparametrar hos energimätare.

1.2 MÅL

Att verifiera och utveckla tidigare framtagna analysverktyg i ett distributionsnät med hög mätkvalitet, utvärdera kostnader samt tillgodose kunskapsspridning och dokumentation av projektet.

Huvudsyftet med projektet är att skapa ökad kunskap om hur mätdata från energimätare kan användas för att underlätta och minska kostnaderna i det dagliga driften i distributionsnäten. Tidigare projekt har identifierat begränsningar i mätinfrastrukturen som försvårar vissa typer av analyser. Detta projekt syftar till att analysera nyttan av god mätkvalitet för distributionsnätägaren, och som en följd av detta att rekommendera vilka krav som kan eller bör ställas på mätarinfrastruktur för kunna identifiera och lokalisera fel.

1.3 AVGRÄNSNINGAR

Endast analyser av distributionsnät kommer att utföras.

Ett begränsat antal elnätbolag kommer att nyttjas i enkätresultaten, dessa kommer att anonymiseras för att så öppet som möjligt kunna tala om deras syn på förluster.

Begränsningar i tillgänglighet hos elnätdata, såsom resistanser, induktanser och kabellängder, kommer ej att urvärderas.

1.4 STRUKTUR

Rapporten börjar med i Kapitel 2 en sammanfattning och presentation av enkätsvar på elnätbolagens syn på nätförluster, kostnader och deras önskan om en lokalisering funktion för elnätsförluster. Delvis baserat på dessa enkätsvar och resultaten från (Hagmar & Lindskog, Bedömning av nätstatus baserad på dataanalys och avancerade algoritmer (2017:443), 2017) kommer metoder baserade på god systemkunskap att presenteras i Kapitel 3. Sedan introduceras Maskininlärning för lokalisering av icke-tekniska förluster i Kapitel 4 för att sedan utvärderas i Kapitel 5. Rapporten avslutas med slutsatser från arbetet.

2 Resultat från enkät till elnätsbolag: Nya mätare, kostnader och icke-tekniska förluster

2.1 INLEDNING

Sju elnätsföretag har intervjuats med två syften; dels att utvärdera eventuell förändring i kostnad för extra funktionalitet som krävs för de analyser och verktyg som föreslås ovan, gällande mätkvalitet och extra mätparametrar hos energimätare, dels att undersöka kostnader för elnätsföretag sett till felsökning av icke-tekniska förluster. Frågorna finns i bilaga A.

Vid urvalet av företag har vi försökt få en viss spridning i storlek och stadsnät/landsbygdsnät. Vissa av företagen har vi kontaktat baserat på att vi känt till att de redan arbetar med en utbyggd mätinfrastruktur och därför kunde förväntas vara bekanta med de aktuella frågeställningarna och utmaningarna.

2.1.1 Kapitelsammanfattning

Resultaten av kostnadsanalysen har varit svåra att kondensera ned till generella nyckeltal och slutsatser. Vi ser tre anledningar till detta: att intervjupersonerna och de sju företagen de representerar har verksamheter som skiljer sig från varandra, att flera av frågeställningarna handlar om kostnadsförändringar i ett tänkt scenario samt att flera av företagen anser att andra utmaningar än nätförluster är viktigare. Se kapitel 2.2 "Elnätsbolagens utmaningar".

Med detta sagt, så har vi ändå fått mycket intressant information av elnätsföretagens syn på dessa frågor.

Första syftet, att utvärdera eventuell förändring i kostnad för förbättrad mätkvalitet och extra mätparametrar för analyser, adresseras i kapitel 2.4 "Tackla icke-tekniska förluster med en utökad mätinfrastruktur".

Andra syftet, att undersöka kostnader för elnätsföretag sett till felsökning av icke-tekniska förluster, beskrivs i kapitel 2.3 "Icke-tekniska förluster - nuläge".

Tabellen nedan är en översikt av de sju intervjuade elnätsföretagen och deras karaktär med avseende på mätinfrastruktur och nätförluster.

Tabell 1: Översikt av de sju intervjuade elnätsföretagen. Statistik för år 2016 från [3]. Siffrorna är avrundade.

Företag	A	B	C	D	E	F	G
Storlek (GWh/år)	200	1000	4 500	70	800	600	18 000
Abonnemang	10 000	60 000	250 000	3 000	40 000	40 000	1 000 000
Ledningslängd (km)	2000	3000	7000	200	2000	1500	130 000
Stadsnät? (m ledning/abonnemang)	180	50	30	70	40	35	130
Nätförluster	3,2 %	2,9 %	2,8 %	2,9 %	2,4 %	4,1 %	3,3 %
Största utmaningar	Byta mätare .	Förnya och bygga ut nätet.	Hantera energistöld civilrättsligt Mätarskåp ute i gatan	Förnya gamla delar Få bukt med icke-tekniska förluster.	Reinvestera. Hur rationellt hantera alla larm och händelser.	Nytt högspänningsnät. Bygga ut elnätet.	Höja driftsäkerheten. Nästa utrustning av elmätare.
Nätstationer	800	1 000	1 600	50	400	500	50 000
Nätstationer med mätning	1	50	ca 1000	49	150	2	0
Följer upp onormala nätförluster genom nätstationsmätning	Nej	Ja	En kampanj år 2014	Ja	Nej	Nej	Nej
Ett tänkt analysverktyg för icke-tekniska förluster - önskad lokaliseringsgrad	-	-	Beror på kostnader	Vilken kabel ut från nätstationen	Nätstation	-	Mellan nätstation och matande kunder för stationen
Mätning hos kund, vilka storheter	Energi dygnsvärden ¹ , vissa kunder med timvärden	Energi timvärden	Energi timvärden, spänning och ström 2 ggr/dygn	Energi timvärden, förberett för spänning, ström och 15 min-intervall	Energi timvärden	Energi timvärden (dagens mätare kan mer ²)	Energi månadsvärden, timvärden på vissa kunder
Mätning hos kund, upplösning energi (kWh)	0,01	0,01	0,01	0,1 på 80 % av mätarna, 1 på resten	0,001	-	0.001-1 (många olika mätartyper)

¹ Energi i kWh / dygn² Företagets nuvarande elmätare kan mer. De uppfyller t.ex. flera av de kommande funktionskraven i [2] som mätning av spänning, ström och effekt samt slå av och på elanläggningen på distans. Kundgränssnitt och uppgradering på distans klarar de dock inte.

2.2 ELNÄTSBOLAGENS UTMANINGAR

Elnätsföretagen fick frågan om vilka som var deras främsta utmaningar. Här redovisas dessa utmaningar kortfattat för att ge perspektiv åt utmaningen med icke-tekniska förluster som detta projekt fokuserar på.

- A. Byte av elmätare m.fl. komponenter p.g.a. felfunktion eller om krav på timmätning ställts av kunden.
- B. Förnyelse av nätet (avskrivningstider, kapitalbas); Utbyggnad av elnätet i takt med befolkningsökningen. Nätet växer med ca 50 % på 10 år.
- C. Att förnya gamla delar. Elnätet finns i en avfolkningsbygd och växer därför inte; ordinarie drift och underhåll; under åren 2011–2013 genomfördes ett större projekt för att få bukt med icke-tekniska förluster.
- D. Att flytta elmätarna ut till mätarskåp som står utanför kundens fastighet för att eliminera möjligheten till strömstölder; Att hantera energistölder civilrättsligt, d.v.s. att stämma kunden eftersom åklagarmyndigheten ofta är hårt belastade med att hantera allvarligare brott. Staten genom åklagarmyndigheten har ju annars åtalsplikt om man vet att ett brott är begånget och det kan bevisas.
- E. Reinvesteringar i den takt nätet åldras, framförallt nätstationer som börjar närma sig 40 år; att hitta verktyg och arbetssätt för att rationellt hantera alla larm och händelser från de allt mer finessrika elmätarna.
- F. Ett flerårigt projekt att bygga ett nytt, starkare högspänningsringnät i staden (befintligt 30 kV-nät har för dålig kapacitet); utbyggnad av elnätet eftersom staden expanderar; byte till nytt system för fjärrstyrning av högspänningsnätet.
- G. Att vädersäkra nätet och minska avbrotten, med kablifiering och isolerade luftledningar; Kommande utrullning av nya generationen elmätare. De menar att nätstationsmätning är önskvärt i framtiden för mer optimerad dimensionering och nyttjande av elnätet och att möjligheten att undersöka förluster får ses som en bonus.

Företag F uttrycker sig så här på frågan om en utökad mätinfrastruktur:

"I ett framtida nytt system så kommer en del nya funktioner kanske att kräva tex 40-faldigt mer mätdata men det är fortfarande smalbandigt, alltså smådatamängder att överföra, jämfört med den omfattande datakommunikationen i ett modernt, uppkopplat samhälle. Det är inte mätningen och mätdatainsamlingen som kommer att vara utmaningen eller det kostnadsdrivande. Det kommer istället att vara att bygga upp en infrastruktur som stödjer våra mål och som har en hög tillgänglighet och låga drift- och underhållskostnader. Likaså bemanning – hur driftar vi organisationen? Hur mycket utrustning sätter man ute i fält?"

Sammanfattningsvis så är det bland dessa sju företag tre stycken som aktivt arbetat med nätförluster och icke-tekniska nätförluster de senaste åren. De andra företagen

anser att deras nätförluster är på en så pass låg nivå att de valt att inte aktivt fokusera på och arbeta med att minska dem.

2.3 ICKE-TEKNISKA FÖRLUSTER - NULÄGE

2.3.1 Generellt

Här redogörs för hur företagen ser på och arbetar med nätförluster och i synnerhet de icke-tekniska förlusterna, med en del kostnadsuppskattningar.

Ett elnäts nätförluster är differensen mellan inmatad och utmatad energi i ett elnät. Nätförlusterna kan delas in i tekniska förluster och icke-tekniska förluster.

Tekniska förluster är den energi som blir till värme i transformatorer och ledningar i ett elnät. Icke-tekniska förluster är förluster som har andra än tekniska orsaker, t.ex. feldokumenterade nät, elstölder, och annan omätt förbrukning.

I detta projekt försöker vi belysa möjligheterna och kostnaderna med att komma till rätta med icke-tekniska förluster. Hur stora är då de icke-tekniska förlusterna – vad skulle vinsten kunna vara med att identifiera, lokalisera och åtgärda dem?

För att få en känsla av denna eventuella vinst gör vi en räkneövning där vi tittar på potentialen för de nätbolag som har en hög rapporterad nätförlust. Baserat på (Energimarknadsinspektionen, Sarskild_rapport_teknisk_data.xlsx, (2018-06-13)) så ligger årsmedelvärdet av de totala nätförlusterna i Sverige på 3,6 %. Vi ansätter att ett elnät borde kunna ha en nätförlust på max 4 %. Dock har 54 av 159 bolag högre nätförluster än så. Om dessa 54 hade gjort förbättringar i sina elnät och nått ned till en nätförlust på 4 %, så skulle det innebära en minskning på totalt 143 GWh/år. Antas en kostnad på 50 öre/kWh (Energimarknadsinspektionen, Ei R2015:07 Incitament för effektivt utnyttjande av elnätet, 2015) motsvarar det 72 miljoner kr/år.

Kostnaden för nätförlusterna och därmed incitamentet att försöka minska dem, beror förstås på elprisutvecklingen.

2.3.2 Mätinfrastruktur nuläge

Här beskrivs dagens mätinfrastruktur i och under varje nätstation. Detta görs för att relatera till den utökade mätinfrastruktur som skulle behövas för att leverera data till verktyg för förlustanalyser, se kapitel 2.4.1.

Hushållsmätare

Fem av de sju intervjuade företagen har timvärden på aktiv energi, medan två av dem har timvärden enbart på vissa kunder.

I företagens mätinsamlingssystem finns det möjlighet att fråga delmängder av mätarna i nätet om spänning, ström m.m, inom ramen för särskilda insatser.

Förutom den mätningen finns hantering av larm av olika slag från mätarna.

Tre av sju företag har energimätning med upplösning på 0,01 kWh eller bättre. Ett företag har sämre upplösning och två företag har inte svarat.

Nätstationsmätning

Fyra av bolagen har någorlunda omfattande mätning i sina nätstationer och en av dessa företag har till och med mätning i alla sina nätstationer.

De tre återstående företagen har mätning bara i någon eller några enstaka nätstationer på prov.

Mätningen i nätstationer består till övervägande delen av elmätare, medan vissa nätstationer är bestyckade med elkvalitetsinstrument.

2.3.3 Uppföljning av nätförluster

Företagen B, C och D arbetar eller har arbetat med att följa upp och åtgärda nätförluster på nätstationsnivå:

Företag B

50 av 400 nätstationer har mätning. Proceduren med dessa är sådan att en månad efter att en nätstationsmätning satts upp så görs en analys med Excel av medelvärdet av totala nätförlusten den månaden. Därefter görs inte systematiskt återkommande analyser av nätförlusterna i nätstationerna.

Nätstationsmätningarna kommer till användning när någon händelse eller upptäckt triggar en kontroll av något slag.

Hittills har företaget inget lyckat exempel på där de med hjälp av nätstationsmätning och sedan uppsättning av temporära mätningar, borrar sig ner i nätet och hittat en strömstöld.

Företaget arbetar i dagsläget snarare med följande metoder:

- Historiskt upptäcks icke-tekniska förluster vid övriga aktiviteter, som t.ex. vid nedplockning av mätare p.g.a. bristande betalning.
- Halvårsvis kontrolleras spänningsvärdena per fas hos hushållselmätarna. Massfråga ställs till de som rapporterar 0 V på en fas. Därefter åker man ut och kontrollerar de som är onormala. Ex: byggsåp som kopplats fel eller en otillåten strömvägledning i form av att I1 kopplats förbi strömtrafon. 30–50 brukar man hitta. Kostnad totalt ca 1,5 tim/fall.
- Kontroll av nollförbrukare via mätarfråga, en gång per år. Ca 30–50 upptäcks. Kostnad 1,5 tim/fall.
- Kontroll av elmätarens register för exporterad energi. Ifall det registervärdet växer så är mätaren felkopplad. Detta görs årsvis eller halvårsvis. 30–50 fall per år upptäcks. Kostnad ca 1,5 tim/fall.
- Maxlarm (säkringsuppgifter i insamlingen). Åker ut och kollar (har kunden i smyg bytt säkring till 35-50A?) 30–50 fall/år. 1,5 tim/fall.
- Fel upptäcks av ackrediterade kontrollanter vid kontroll i fält ibland.
- Ibland kallas om postadress för räkning stämmer med abonnemangsadressen. Det tyder på åretrunt-boende och då är en förbrukning under 10 000 kWh/år onormalt.
- På birthday.se kan man se hur många som bor på en adress och dra vissa slutsatser.

- Årsförbrukning vatten och fjärrvärme v kan korreleras med elanvändningen för att se nåt skumt, eftersom företaget levererar både el, fjärrvärme och vatten.

Företag C

Under 2014 beräknades förlusterna per nätstation. Nätstationerna med högst förluster undersöktes vidare. Denna satsning ledde till att man utvecklade nya, proaktiva arbetssätt för att hålla koll på sitt elnät med bl.a. bättre rutiner för kontroll av anläggningar med bristande betalning.

En viktig metod är undersökningskampanjer i ett område under en begränsad tid. Typiskt arbetar två personer i två veckor (160 mantimmar) i en stadsdel eller ett industriområde. De fall av omätt förbrukning som upptäcks under en sådan här tvåveckors kampanj är typiskt i storleksordningen 40 000 kWh/år men det har även funnits extremfall på 180 000 kWh/år.

I de här kampanjerna kontrolleras inte bara omätt förbrukning. Man kontrollerar även anläggningens skick vilket ibland leder till att elcentralen döms ut etc. Dessa kampanjer höjer personalens arbetssäkerhet, leveranssäkerheten och kundens inre säkerhet. Att jaga strömstöld är i första hand säkerhetsfråga. Detta arbete är viktigt eftersom man ju nuförtiden inte åker ut och kontrollerar elanläggningarna ens i närheten lika ofta som förr när det var månadsavläsning av energiförbrukningen.

De arbetssätt som utvecklats består bl.a. av följande:

- Utbildning av teknikerna, framtagning av en manual som beskriver hur man ska okulärbesikta, dokumentera, kommunicera till kund, kommunicera internt, utföra mätningar på sätt som håller rättsligt m.m.
- Anläggningar som är fränkopplade (ej betalt, saknas avtal, farliga) kontrolleras så att de inte är driftsatta och omätta.
- Anläggningens förbrukning i förhållande till tidigare förbrukning och byggnadens storlek undersöks. Om förbrukningen verkar orimlig, kontrolleras den med amperetång-mätning.
- Nollförbrukare undersöks (gränsen är satt till 5 kWh/år). Ex: Närlivs eller cykelverkstad med 0 kWh.
- Mindre och medelstora verksamheter – där finns en överrepresentation av förbikopplade mätare.

Företag D

Företag D arbetade aktivt med att få ned nätförlusterna i sitt elnät under 2011-2013, då de satte upp nätstationsmätning och med hjälp av bl.a. denna hittades 25 strömtjuvar och nätförlusterna minskade med 1,2 procentenheter från 4 % till 2,8 %. Företaget bedömer att de 25 strömtjuvarna stulit mellan 6000 och 40 000 kWh/år i gissningsvis 10-20 år. Det motsvarar grovt räknat 0,7 % av energin i nätet. Resterande minskning på 0,5 procentenheter kan tänkas bero på signalvärdet i att bolagets tekniker är ute och mäter i nätet och att folk ser detta och hör talas om andra i trakten som åker fast för strömstöld. Det kan också bero på andra felaktigheter som upptäckts och rättats till under perioden, utöver de 25 strömstölderna.

Metoden har i stora drag bestått i att de har en automatisk beräkning som löpande övervakar nätförlusterna per nätstation. Med hjälp av erfarenhet, så märker man om något inte står rätt till. Därefter påbörjas en period med temporära mätningar ute i kabelskåp. En sådan mätning pågår typiskt i en vecka och därefter analyseras mätdata och flyttas den temporära mätningen till en ny plats i nätet. I genomsnitt har ca sex sådana här temporära mätningar behövts innan en strömtjuv lokaliserats helt. Varje sådan mätning tar ca 1 timmas arbetstid för upp- och nedmontering samt 1 timma för analys. Andra metoder för lokalisering finns också. Exempelvis om man har ringat in en strömstöld till 10 abonnenter så kan man titta på förbrukningssiffrorna och ana vilka som inte värmer sina hus med el. Om någon av dessa ändå har snö på sin skorsten så är det misstänkt.

Övriga företag

Övriga företag arbetar inte systematiskt med nätförluster. De svarar att de är relativt nöjda med den totala nivån av nätförluster i deras nät och att de har andra utmaningar som de behöver lägga sin tid och kraft på.

De upptäcker förstås då och då icke-tekniska förluster i samband med andra aktiviteter, t.ex. vid ombyggnation eller den regelbundna (6 år) verifieringen av energimätsystem. Exempel på sådant som upptäcks är felkopplingar av faser, inkoppling av last före mätning, strömstölder, att man efter kontroll av mätaren glömt koppla tillbaka strömmen, att kunden bytt till större säkring utan lov.

Ett speciellt fall som skulle kunna klassas som "avsiktlig" icke-teknisk förlust ur ett mätperspektiv, är följande: Ett av de intervjuade företagen som har inte har någon mätning av stadens gatubelysning utan istället debiterar kommunen för märkeffekten multiplicerat med den tid som lamporna lyser (tidmätning). Skulle man försöka analysera förluster per nätstation där så skulle man få ta hänsyn till denna osäkerhetsfaktor.

2.3.4 Lokalisering - räknexempel

Här följer ett räknexempel på kostnad och besparing vid lokalisering av en icke-teknisk förlust när en sådan upptäcks under en nätstation:

Exempel: En nätstation med 5 underliggande grupper och 10 abonnenter i varje grupp. Mätning finns i nätstationen.

1. Analys av nätförlusterna med hjälp av mätning i nätstation och i hushållsmätarna. Nätförlusterna visar sig vara 7 %.
2. Kontroll av historiska data ger ingen förklaring. Arbetstid 1 tim.
3. Temporär mätning sätts upp på en av grupperna i en vecka. Arbetstid 2 tim. Nätförlusterna i gruppen är låga, 3 %.
4. Den temporära mätningen flyttas till nästa grupp och upprepas. Låga nätförluster. Arbetstid 2 tim.
5. På tredje gruppen upptäcks onormalt höga nätförluster på 9 %. Arbetstid 2 tim.
6. Temporär mätning sätts nu upp i kabelskåpet vid en abonnent som företagets tekniker misstänker. Låga nätförluster. Arbetstid 2 tim.

7. Temporär mätning på nästa abonnent. Här är stora skillnader mellan denna mätning och vad kundens elmätare rapporterar. Den icke-tekniska förlusten är lokaliserad. Arbetstid 2 tim.
8. Elanläggningen undersöks och en T-koppling upptäcks som går till ett bilgarage/verkstad. Stulen energi har gissningsvis uppgått till 20 000 kWh/år i flera år. Arbetstid 2 tim.
9. Fallet avslutas med att kunden ställs till svars ekonomiskt. Arbetstid 3 tim.

Total tidsåtgång: 16 tim. Med en timkostnad på 500 kr blir det 8 000 kr.

Besparing: 20 000 kWh/år i upphörd strömstöld.

2.4 TACKLA ICKE-TEKNISKA FÖRLUSTER MED EN UTÖKAD MÄTINFRASTRUKTUR

Här följer en beskrivning av vad som krävs i form av mätinfrastruktur för att kunna hitta icke-tekniska förluster. Företagens syn på nyttor, kostnader och hinder för en utökad mätinfrastruktur redogörs också för.

2.4.1 Behov av utökad mätinfrastruktur

Mätdata från en utökad mätinfrastruktur ger förutsättningar för automatiserade analysverktyg att peka ut icke-tekniska förluster i lågspänningsnätet under en nätstation. Mätning i nätstationen är då önskvärd, liksom mätning av fler storheter än energi i hushållselmätarna, tillräcklig upplösning i tid och storlek samt god tidssynkning.

Nätstationsmätning

Nätstationsmätning är önskvärd då det bidrar till att nätet kan sektioneras och noggrant analyseras utan inverkan av överliggande nät. De två huvudtyperna av nätstationsmätning som förekommer, består av mätningar med ordinära elmätare eller mätningar utförda med elkvalitetsinstrument.

Hushållselmätare

I [1] föreslås följande krav på abonnenternas elmätare och för stödsystemen:

- Tidhållning: Mätdata är korrekt tidsstämplat ± 1 sekund
- Upplösning energi: åtminstone 0,01 kWh
- Mätning av spänning: Det enskilt viktigaste mätvärdet för att förbättra möjligheterna till att utföra diverse nätanalyser. Spänningsvärden ökar både noggrannheten och möjligheten till lokalisering i nätanalyserna.
- Mätning av ytterligare storheter: Mätvärden för ström, energi, samt aktiv och reaktiv effekt bedöms höja noggrannheten i analyserna, men är inte avgörande för dess möjligheter.
- Avläsning var 15 minut: Då förluster i elnätet i princip är kvadratisk beroende på uttagen effekt kan effektlödesberäkningar och liknande bli missvisande om endast timvärden används. En mer frekvent avläsning var 15:e minut förbättrar även noggrannheten i det avseendet, men höjer som tidigare nämnts kraven på tidssynkronisering.

- Min-, max- samt medelvärden: Alternativ till en mer frekvent avläsning kan vara att även min-, max, samt medelvärden registreras.
- Stödsystem: Stödsystem ska inte förvanska mätvärden t.ex. genom avrundning. Tidsstämpling ska göras på ett sätt som möjliggör att mätvärden kan sammanfogas på ett enkelt och standardiserat sätt.
- Mätnoggrannhet på ovan nämnda storheter: Både i nuvarande och i de föreslagna, kommande föreskrifterna för energimätning ställs noggrannhetskrav enbart på just mätningen av energi. Extra krav utöver de som föreskrivs i (Energimarknadsinspektionen, Ei R2017:08 Funktionskrav på elmätare – Författningsförslag, 2017) bör därför ställas på elmätarna om man ämnar använda dessa storheter för analyser i framtiden. I funktionskraven nämns inte krav på mätnoggrannhet för de olika mätvärden som föreslås. Det finns i dagsläget som sagt endast lagstadgade krav för elmätarna angående aktiv energi, och därmed inte för någon annan storhet. Detta innebär även att noggrannheten för övriga tillgängliga mätvärden som eventuellt kan erhållas ur mätaren ej kan garanteras eller verifieras. Erfarenheten är att det kan finnas väsentliga fel i dessa storheter.

Swedac är den myndighet som utfärdar föreskrifter med krav på noggrannhet i mätning. Swedac arbetar nu med förslag till författningsändringar till följd av EI:s rapport om funktionskrav för elmätare (Energimarknadsinspektionen, Ei R2017:08 Funktionskrav på elmätare – Författningsförslag, 2017). Gällande upplösningen i energi föreslås i remiss 2018/653 från 2018-05-25 att kravet ska vara 0,001 kWh. Mätnoggrannheten för övriga mätstorheter nämns inte i remissen.

2.4.2 Behov av lokalisering

Analyser som körs i framtiden på mätvärden från en utökad mätinfrastruktur, kan hjälpa till med att kategorisera mätfel och förluster samt att identifiera och lokalisera icke-tekniska förluster. När det gäller lokalisering är det intressant att fråga sig hur bra den måste vara. Lokalisering kan delas i följande fyra nivåer:

- Nätstationsnivå (låt oss säga att 7 kablar går ut från denna)
- Kabelnivå (ett antal kopplingsskåp på varje kabel)
- Kopplingsskåpsnivå (2-6 kunder)
- Kundnivå

Företag C anser att Good enough är väldigt beroende på storleken på förlusten. Kostnaden för att hitta och åtgärda förlusten måste ju i alla fall vara i som högst av samma storleksordning som kostnaden per år för själva förlusten.

Företag D anser att mycket är vunnet om lokaliseringen kan ske ner till kopplingsskåpsnivå. Vissa kunder kan uteslutas direkt och då återstår kanske bara 2-3 kunder. Det är också troligt att man direkt kan förstå exakt vilken kund det är.

Företag E menar att lokalisering till nätstationsnivå duger, på grund av deras relativt låga nätförluster.

Företag F resonerar att man bara vill identifiera icke tekniska förluster mellan nätstation och matande kunder för stationen. Sedan får man avgöra huruvida man

anser det värt att agera och göra djupare analyser. Att ta in kvalificerade system för inhämtning, bearbetning etc. kostar ju också pengar.

2.4.3 Nyttor

Hur omfattande är problemet? Vad är "normal" nätförlustnivå? Hur stor andel kan utgöras av icke-tekniska förluster?

I räkneövningen i kap. 2.3.1 ansatte vi 4 % som en övre gräns för nätförlusterna i ett elnät. Det gav att elnätsföretagen skulle kunna minska sina nätförluster med totalt 143 GWh/år vilket motsvarar 3,9 % av de totala nätförlusterna i Sverige. Antas en kostnad på 50 öre/kWh [4, sid 34] motsvarar det 72 miljoner kr/år.

Av företag D får vi ytterligare en bild av nätförlusternas storlek (kapitel 2.3.3): Under en treårig kampanj eliminerades 25 strömstöldar à 25 MWh/år = 625 MWh/år, vilket är ca 0,7 % av företagets totala inmatade energi. Totalt minskade nätförlusterna från 4 % till 2,8 % under kampanjen. Om vi antar en kostnad på 50 öre/kWh [4, sid 34] så har företag D med ca 3 000 abonnenter sparat in ca 420 000 kr/år.

Det finns ett signalvärde i detta också, som skulle kunna uttryckas ungefär så här: "Vi har en modern mätinfrastruktur för monitorering av förluster i vårt elnät. Det bidrar till rättvisa och sänker kostnaderna för oss och våra kunder."

Övriga nyttor med nätstationsmätning, enligt de intervjuade företagen, är att man får möjlighet till bättre kontroll på elkvalitet, kapacitet (kablar, transformatorer) och bättre underlag för investeringsbeslut. Företag C anser att jaga strömstöldar i första hand är en säkerhetsfråga (arbets säkerhet, leveranssäkerhet och kundens inre elsäkerhet). Företag G menar att nätstationsmätning framförallt är önskvärt för att åstadkomma en mer optimerad dimensionering och bättre nyttjande av elnätet och att möjligheten att undersöka nätförluster får ses som en bonus.

Både i detta arbete och i (Hagmar & Lindskog, Bedömning av nätstatus baserad på dataanalys och avancerade algoritmer (2017:443), 2017) vittnar de intervjuade elnätsföretagen alltså om att det i hög grad förenklar och effektiviserar arbetet samt att informationen börjat användas i flera delar av verksamheten. De beskriver att det kan vara svårt att värdera nyttan rent ekonomiskt men de skulle aldrig vilja eller ens kunna gå tillbaka till tiden utan nätstationsmätning.

2.4.4 Kostnader, hinder

Installation av nätstationsmätning med elmätare kostar 3 500–6 000 kr/station, enligt uppskattningar av de intervjuade företagen.

När det gäller kundmätarna så kommer en ny generation elmätare att rullas ut de närmaste åren som en följd av de kommande nya funktionskraven på elmätare [2]. Man kommer i många fall att behöva byta till nytt insamlingssystem också i samband med detta. I samband med dessa stora utrustningsprojekt som ligger framför dem, så har elnätsföretagen svårt att uppskatta särkostnaden för extra mätdata, bättre tidshållning och för att övergå från timmätning till 15 minuters mätning. Företag C resonerar som så att största hindret för (tim)mätning av

fasspänning är att det krävs omfattande utveckling av insamlingssystemen: head-end system, konzentrorer, kommunikationsmoduler, i princip allt. Därtill kommer testning. För deras största mätinsamlingssystem skulle kostnaden bli någonstans mellan 1 och 2 miljoner kronor. Ett annat hinder de har är att spänningsmätningen på en av deras mätartyper inte fungerar tillfredställande.

Även om mängden mätdata mångfaldigas så är det ändå små datamängder jämfört med annan mobil- och internetkommunikation i samhället framöver. (Ett av företagen, företag A, gjorde ändå en enkel uppskattning av kostnadsökningen med nuvarande kommunikationslösning. Företag A har i dagsläget kostnader för mobildata på 1,50 kr/år och kund. Med allt annat lika skulle datakostnaden för att få med de tre fasspänningarna utöver energin, öka med en faktor 4, d.v.s. till 6 kr/år och kund.)

Sammanfattningsvis bedöms att, mätartekniskt och bandbreddsmässigt sett, så kommer inte mätning och insamling av mer och vassare mätdata att vara den stora utmaningen eller det som är kostnadsdrivande i den kommande utrullningen av nya generationens elmätare.

3 Lokalisering av icke-tekniska förluster med systemkunskap

Syftet här är att utvärdera möjligheten med utökade mätparametrar såsom spänning samt ökad samplingsfrekvens hos värdena hos energimätarna. Möjligheten till lokalisering kommer att påverkas av mätkvaliten hos energimätarna, därför kommer även upplösningen hos energimätarna att utvärderas för att visa deras påverkan på resultatet från lokaliseringen av de icke-tekniska förlusterna.

3.1 METOD

Metoden baseras på att följande information finns tillgänglig.

- En nätstation med spänning och energimätning.
- Slutkundsmätning med energimätare med en samplingsfrekvens på minst en timma för huvudspänningens RMS-värde.
- Information om elnätets struktur och uppbyggnad.
- Elnätsparametrar såsom kabellängd, resistans, induktans och kapacitans per meter.

Andra metoder som baseras på korrelation, ett statistiskt samband mellan parametrar, kommer även att utvärderas.

3.1.1 Avgränsningar

Metoden syftar inte till att skapa larm och därmed i realtid lokalisera icke-tekniska förluster. Detta leder till att det inte spelar så stor roll om mätdata kommer in snabbt eller långsamt, däremot är som tidigare diskuterats (se Kapitel 1), att mätkvaliten viktig.

Metoderna kommer att utvärderas på två faktiska elnät, Detta skall ses som ett första steg till lokalisering, även om tankar om generalisering finns under projektets gång.

Då mätdata av denna sort (energi och spänningsmätvärden) inte funnits tillgängliga under projektets gång har mätdata simulerats fram. Denna data har dock baserats på verkliga laster för respektive elnät, detta för att få så verklighetstroga spänningsvariationer som möjligt i de olika elnäten.

Påverkan av obalanserade spänningar har inte undersökts. Sådana spänningar bör underlätta lokaliseringen; vid mätningar av individuella faser och deras spänningsfall samt individuella energimätningar per fas.

Mängden data som behövs för att lokalisera olika typer av icke-tekniska förluster kommer inte undersökas, detta påverkar tiden som verktyget tar att köra och dess möjlighet att lokalisera kortvariga icke-tekniska förluster.

3.1.2 Modellering av icke-teknisk förlust

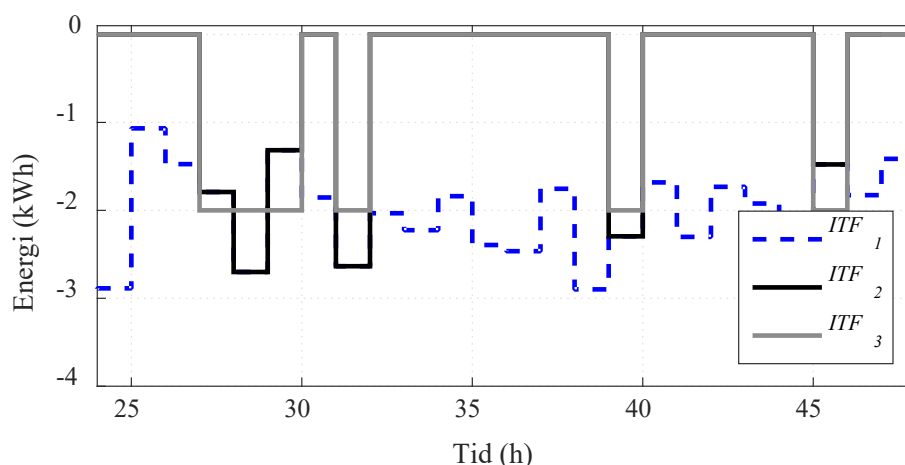
Icke-tekniska förluster som detekterats tidigare (Hagmar & Lindskog, Bedömning av nätstatus baserad på dataanalys och avancerade algoritmer (2017:443), 2017) har varit allt från belysningscentraler som varit omätta till, pumpar och värmeelement. Beteendet hos dessa har varierat med säsong, ljus och ibland även tid. Problematik relaterat till otillåten strömvledning (strömstöld) har ännu inte funnit så många direkta fall i Sverige även om tack vare enkätresultaten det har visat sig att de strömstölder som funnits har varit större laster.

Då slumpmässiga icke-tekniska förluster (ITF) verkar vara de svåraste att upptäcka enl. (Hagmar & Lindskog, Bedömning av nätstatus baserad på dataanalys och avancerade algoritmer (2017:443), 2017) har sådana slumpmässigt räknats fram från en normaldistribution med ett medelvärde, $\mu = 2$, och en standard avvikelse, $\sigma = 0.5$. Detta utfördes med hjälp av;

$$ITF = \text{normrnd}(\mu, \sigma, \text{antal timmar})X \quad (3.1.1)$$

i Matlab, där X är ett slumpmässigt värde (antingen 1 eller 0) för att bestämma om lasten är inkopplad eller ej. Funktionen används sedan för att skapa tre typer av slumpmässiga ITF för utvärderingen av lokaliseringsalgoritmen.

- ITF_1 , modelleras som slumpmässiga energiförluster som är aktiva under hela utvärderingsperioden, varje timma.
- ITF_2 , modelleras som slumpmässiga förluster som är slumpmässigt inkopplade under utvärderingsperiodens olika timmar.
- ITF_3 , modelleras som 2 kW, slumpmässigt inkopplad under utvärderingsperioden olika timmar.



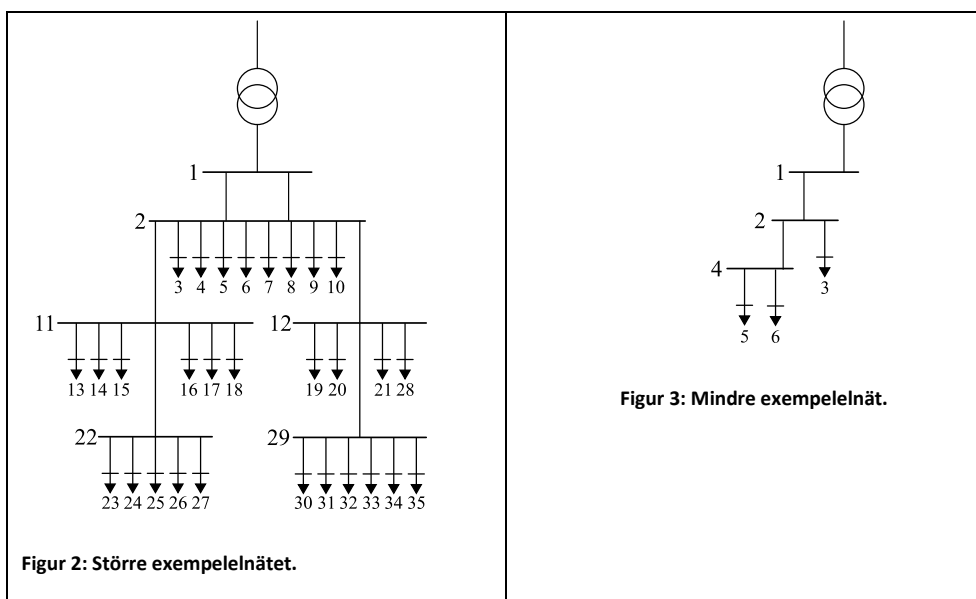
Figur 1: Visar de olika slumpmässiga ITF som kan användas för lokalisering av ITF.

Figur 1 visar dels den ITF_1 som enl. en normaldistribution varierar men är aktiv hela tiden under utvärderingsperioden, detta skulle representera en mer kontinuerlig ITF. ITF_2 och ITF_3 är sedan slumpmässigt aktiva under perioden men

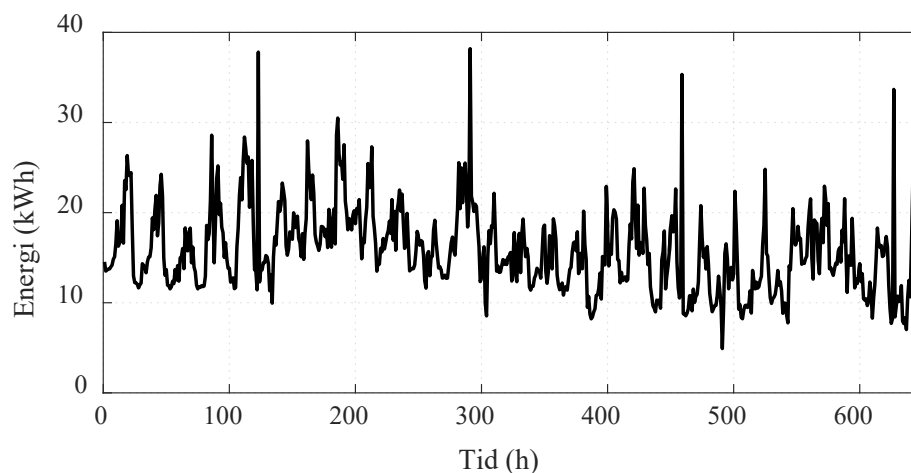
där effekten antingen är konstant 2 kW (hos ITF_3) under timmen eller med en varierande storlek ITF_2 , likande den hos ITF_1 .

3.1.3 Nätområden

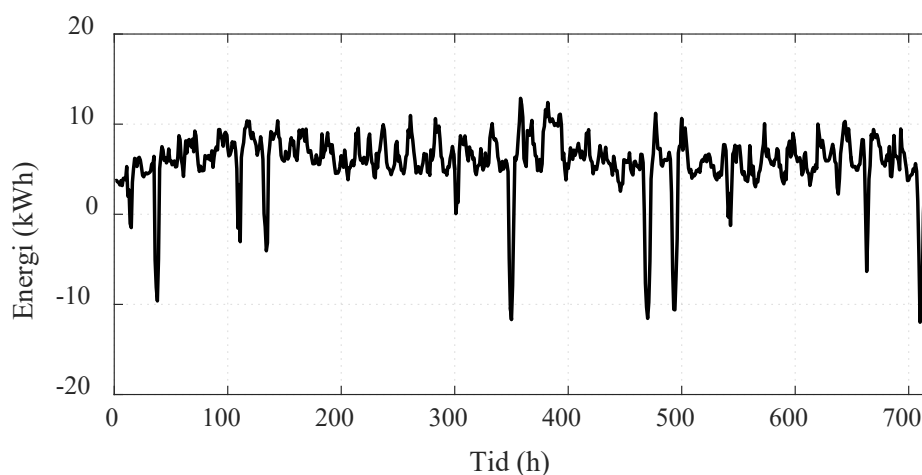
Två existerande nätområden, med verkliga laster har utvärderats. Ett mindre med aningen längre avstånd samt ett större stadsnät. Lasterna är de faktiska timvärdena från respektive område. Med hjälp av nätmodeller har sedan nätstationsmätningen simulerats fram.



Flödena av energi är väldigt olika, men båda elnäten är delvis uppbyggt av kunder med förnybar elproduktion, vilket gör att effektflödet kan vara dubbelriktat under underutvärderingstiden. Detta kommer vara ett vanligt förekommande i framtida elnät och något en lokaliseringens algoritm bör klara av. Transformatorbelastningen för det större elnätet under februari ser ut enl. Figur 4 och januari för det mindre elnätet enligt Figur 5.



Figur 4: Energiflöde genom transformatorn på det större exempelnetet (se Figur 2) under utvärderingsperioden i februari.



Figur 5: Energiflöde genom transformatorn på det mindre exempelnetet under utvärderingsperioden i januari.

Från Figur 4 kan ytterligare en nytta ses av att aggregera mätdata, effekt topparna som syns 4 gånger under de utvärderade perioderna visade sig vara legionellaskydd hos varmvattenberedare som aktiverats baserat på en timer i flera av husen i elnätet.

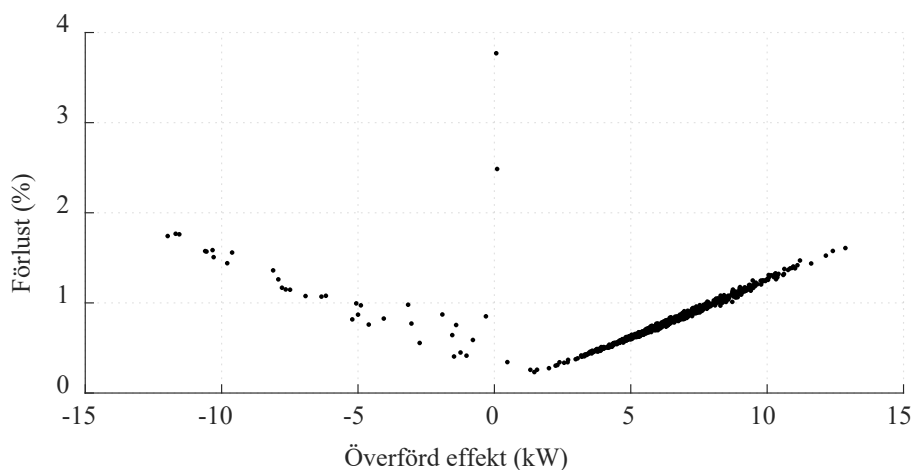
3.2 FALLSTUDIE FRÅN LOKALISERING AV ICKE-TEKNISK FÖRLUST

För att studera hur nätet påverkas av en ITF kördes två stycken effektfödesberäkningar (eng. power flow analysis) för att dels beräkna nätförlusterna i näten, och dels för att kunna estimerar spänningarna i respektive lastpunkt (Saadat, H. (1999)). Den första effektfödesberäkningen innehåller ingen ITF och ger då de värden som skulle rapporterats in till nätägaren från de energimätarna hos kund. Denna data är i form av energiförbrukningar per timma,

h , per kund, i , $E_{i,h}$. Den andra effektlödesberäkningen innehåller en ITF, detta resulterar i spänningsvärden som skulle bli uppmätta hos varje kund, varje tidssteg, $V_{i,h}^{meas}$. Nätstationsmätning fås sedan genom summan av energin som rapporteras in, $E_{i,h}$, och förlusterna i elnätet men innehåller även en ITF. Med hjälp av samma procedur kan sedan de estimerade spänningarna som borde vara hos kunderna, $V_{i,h}^{est}$, beräknas genom att använda nätstationens spänning som antas vara 400 V huvudspänning, parametrarna för elnätet och energiförbrukningarna hos kunderna. Detta resulterar i simulerade spänningsvärden, baserade på verkliga laststorlekar med eller utan ITF och en nätstationsmätning.

Systemets förluster utan ITF

För att undersöka metoden användes primärt det mindre exempelelnätet. Utan någon ITF såg förlusterna ut enligt figuren nedan.



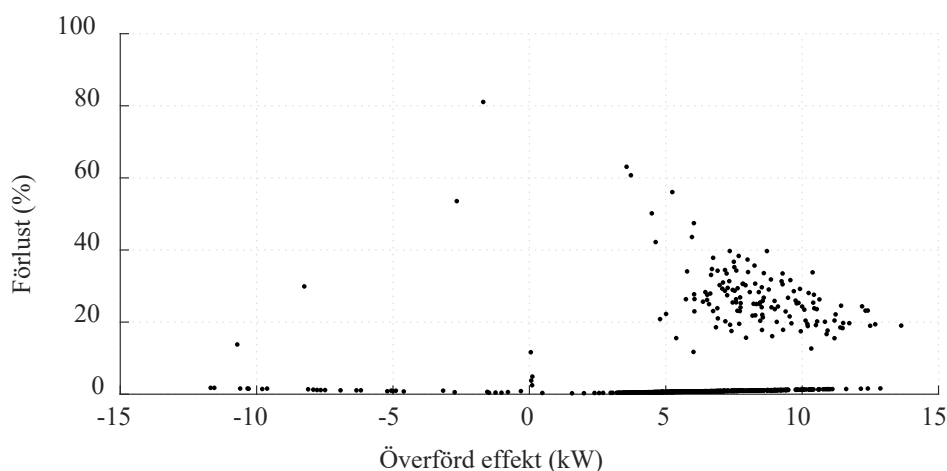
Figur 6: Förluster i elnätet som en funktion av överförd effekt då med ordinarie lastdata i det mindre exempelelnätet.

Figuren visar det sambandet mellan förluster och överförd effekt i nätstationen utan någon ITF. Då elnätet även har en större andel solproduktion ses även negativ effekt under några få soliga dagar med. Samt några punkter med låg överförd effekt som, pga. upplösningen hos energimätarna på 10 Wh, skapar förluster som är höga. Detta är dock inte faktiska förluster, men det är viktigt att vara medveten om att detta kan ske i verkliga applikationer. Det är även intressant att se att vid låg negativ överförd effekt så är spridningen större än motsatta överföringsstorlekar, det är möjligt att det finns en offset innan produktionsregistret aktiveras hos mätaren. Detta har inte undersökts ytterligare i denna rapport.

Systemets förluster med ITF

Då en ITF simulerades vid inkopplingspunkt #6 enl. ITF_2 i Figur 1, det vill säga slumpvis aktiv och till slumpvis storlek *detekterades* dessa relativt enkelt då dess storlek jämfört den överförda effekten är signifikant i det lilla elnätet. Motsvarande

förlust gentemot överförd effekt, som i Figur 6, blir således annorlunda sett till en nätstation med en ITF.



Figur 7: Förluster i elnätet som en funktion av överförd effekt då en ITF är placerad vid lastpunkt #6 i det mindre exampelnätet.

Från Figur 7 kan det observeras de tidpunkter som den ITF är aktiv som skapar förluster i procent av överförd effekt som är väsentligt mycket högre än de förväntade förlusterna som ligger som ett pärlband upp mot 2% på liknande sätt som i Figur 6. Det är således ganska klart att det är något som är fel och möjligheten att *detektera* detta är definitivt något som kan utföras. Frågan som nu blir aktuell blir att se huruvida *lokaliseringen* av den detekterade ITF är möjlig.

3.2.1 Lokalisering av icke-tekniska förluster

Två strategier för lokalisering har utvärderats och en kort sammanfattning och beskrivning över dessa listas nedan;

1. Korrelation: Spänning – Energi (K:SE); Om vi antar att det finns spänningsmätning i energimätarna kan korrelationen undersökas mellan en effektökning och en effektminskning vid en energimätare och huruvida den möts med en motsvarande förändring i spänning.
2. Spänningsfel i mät punkt (SiM); Givet spänningsmätning i nätstation och lastflödesanalys med hjälp av data från energimätare kan spänningen räknas ut baserat på elnätsparametrarna. Ett fel mellan uppmätt och beräknad spänning skulle då kunna indikera en lokalisering av ITF.

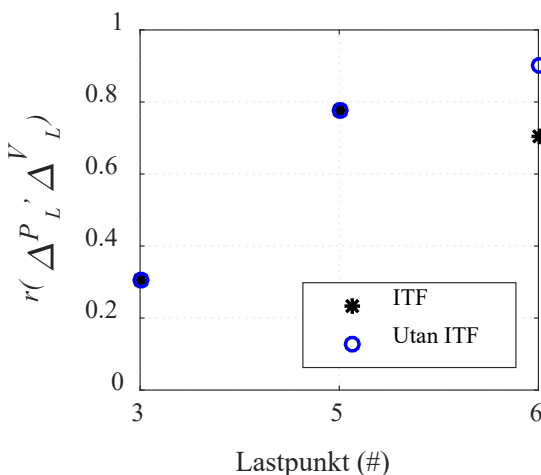
Generellt kan sägas att under antagandet att en ITF inte är konstant aktiva bör förlusterna vara en inparameter för att påbörja strategin för SiM. Därför sätts en gräns på nätförluster på 10%, över denna gräns kommer SiM utföras. Däremot, när det gäller K:SE handlar det om att ha en så hög korrelation över all data som möjligt, om sedan den korrelationen ändras så borde K:SE försämrats (gå längre från 1 mot 0) och då påvisa en eventuell ITF.

K:SE

$$r(A, B) = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{A_i - \mu_A}{\sigma_A} \right) \left(\frac{B_i - \mu_B}{\sigma_B} \right)$$

Där, $r(A, B)$ är korrelationskoefficienten mellan värdena i A och värdena i B, A_i och B_i är varje observation, i . μ och σ är medelvärdet respektive standardavvikelsen hos A och B som totalt har N -antal observationer totalt. $r(A, B)$ kan då variera mellan -1 till 1 varpå 1 och -1 indikerar positiv eller negativ korrelation samtidigt som värden runt 0 ± 0.5 kan ses som en svag korrelation. En svag korrelation påvisar då att exempelvis A inte till någon större utsträckning påverkar B. En korrelationskoefficient på ± 1 indikerar en direkt linjär påverkan mellan de två.

Om vi då ersätter A och B med energimätvärdet och spänningsvärdet för en specifik timme borde de två korrelera ganska starkt mot varandra.

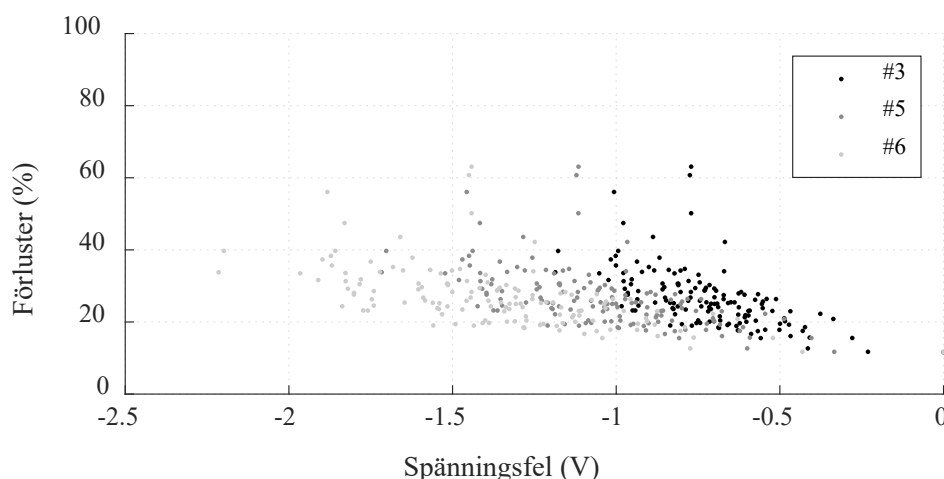


Figur 8: Korrelationskoefficienten mellan värden i energi och spänning med och utan ITF_2 vid de olika lastpunkterna i det mindre exemplnätet.

Resultaten från K:SE är aningen svaga för att påvisa en tydlig skillnad, exempelvis kan det ses att nära inkopplingspunkten som i detta fallet anses ha en konstantspänning så sker givetvis en lägre korrelation, detta då det är mindre resistans i denna delen av nätet. Det är intressant att det ändå är en markant skillnad i K:SE vid lastpunkt #6 sett med eller utan ITF. Problematiken blir om elnätet ovan transformatorstationen är svagt eller på något sätt (justering av lindningsomkopplare, inkoppling av kondensatorbatterier och omkopplingar i nätet) påverkar transformatorsspänningen avsevärt. En sådan påverkan skulle då kunna ge enstaka felaktiga värden hos lastpunkter, då korrelationen blir motsatt det förväntade. Fördelen är att vid ett mindre nät kan det vara möjligt att bara observera spänning och energi från enskilda mätare, och om de vid perioder korrelerar dåligt så kan detta vara en indikation på en ITF. Men enstaka förändringar kan vara på grund av förändringar i transformatorspänningen. Sådana förändringar i spänning påverkar däremot inte SiM strategin för lokalisering som presenteras nedan.

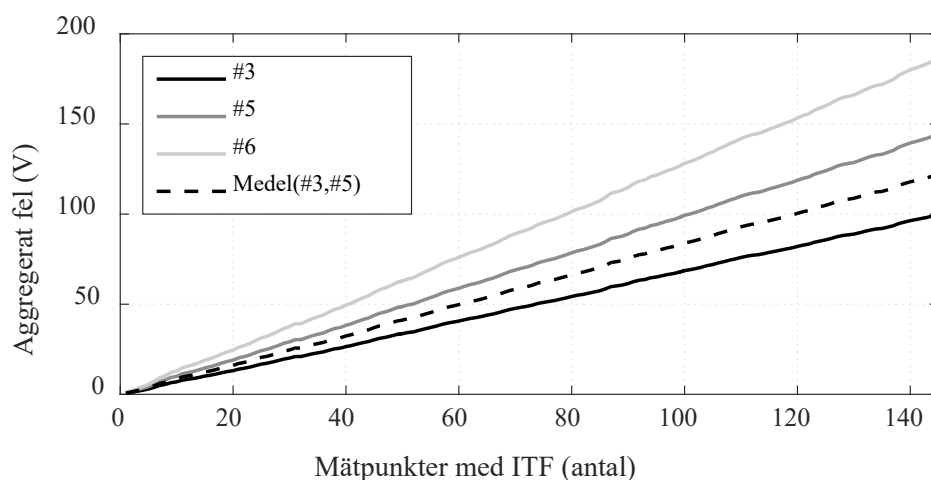
SiM

Här utvärderas spänningsavvikelsen mellan estimerad spänning, $V_{i,h}^{est}$, och uppmätt spänning, $V_{i,h}^{meas}$, i det lokala elnätet. Denna avvikelse fås med hjälp av spänningsmätning i transformatorstationen och med kunskap om elnätsparametrarna, samt spänning och energimätning hos kunderna.



Figur 9: Spänningsavvikelsen mellan uppmätta och beräknadespänningsvärden hos olika lastpunkter vid olika förlustnivåer med ITF_2 i det mindre exemplarnätet.

Ur Figur 9 går det att särskilja spänningsfelet eller spänningsavvikelsen mellan det uppmätta (från energimätaren i lastpunkten) och det beräknade spänningen (från effektflödesberäkningar, spänning i nätstation och energimätning hos kund). Problematiken då ITF nu är lokaliserad långt ut i nätet blir att den givetvis påverkar andra närliggande lastpunkter också. Oavsett så avviker lastpunkt #6 mest. Ett sätt att illustrera detta är att varje timma som förlusterna är över 10% aggregera ihop spänningsfelen för varje lastpunkt.

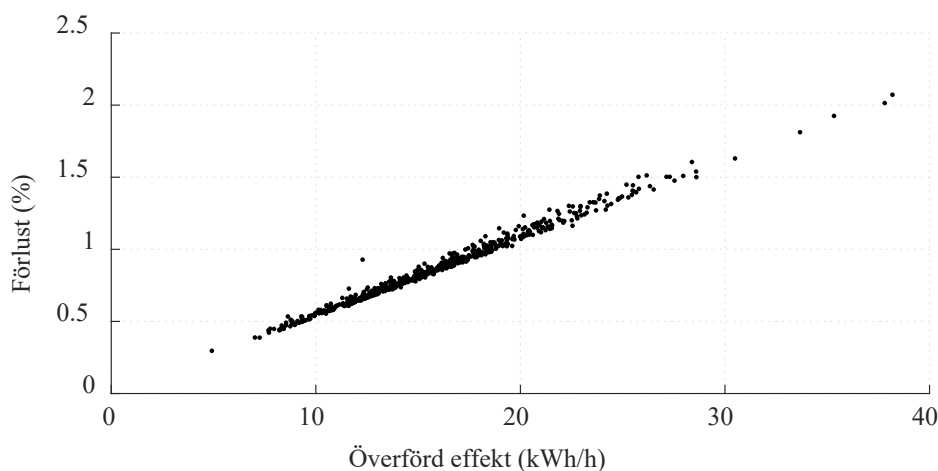


Figur 10: Resultat från SiM-metoden i det mindre elnätet tagna då de elektriska förlusterna överstiger gränsen, på 10 % av överförd effekt, 145 av 721 timmar.

Figur 10 visar det aggregerade spänningsfelet för varje lastpunkt (#3,#5 och#6) i det lilla elnätet. Det aggregerade felet syns på y-axeln och x-axeln representerar varje timma med förluster högre än 10%. Det kan observeras att lastpunkt #6 avviker från de andra lastpunkterna, men ännu tydligare medelvärde mellan de andra. Detta aggregerade fel (som presenteras på y-axeln i Figur 10) hade troligtvis avvikit ännu mer om inte ITF_2 haft så små variationer i toppvärde över dygnet. Samtidigt går det att observera att alla laster påverkas av den ITF som finns men *lokaliseringen* är möjlig. Vid mer komplexa elnät kanske den inte lyckas identifiera en specifik lastpunkt med ITF, men den bör kunna ge en fingervisning om var man skall börja leta. Även om målet med projektet har varit att undersöka möjligheterna till att lokalisera en specifik lastpunkt så har det i intervjuer med elnätsföretagen kommit fram att om man kan isolera ett kabelskåp där det kan finnas en ITF så kan mycket tid besparas i lokaliseringsförfarandet, för ytterligare information se Kapitel 2. Ytterligare en fördel med att jämföra varje individuell lastpunkt på detta vis är att mätosäkerheten hos mätarna och mindre fel i elnätsparametrar aggregeras ihop. Eventuellt kan däremot en dåligt åtskruvad kabelanslutning öka resistanser etc. men detta kan ev. fångas upp. Kabelresistans vid temperaturökningar i XLPE, 90 C° vs 20 C° ger en ökning på 28% (Bergh, Kadurek, Cobben, & Kling, 2011), detta är en möjlig felkälla i detta fallet. Frågan huruvida detta är ett problem i svenska distributionselnät kvarstår.

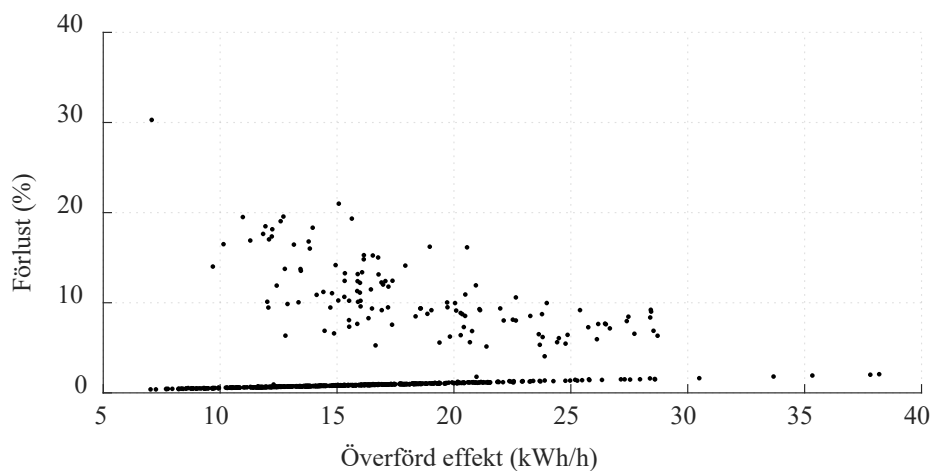
3.2.2 Lokalisering i större exempelelnät

För att utvärdera SiM i ett större elnät applicerades samma stöld (ITF_2) elnätet i Figur 2. Utan ITF ser förlust distributionen ut som i Figur 11.



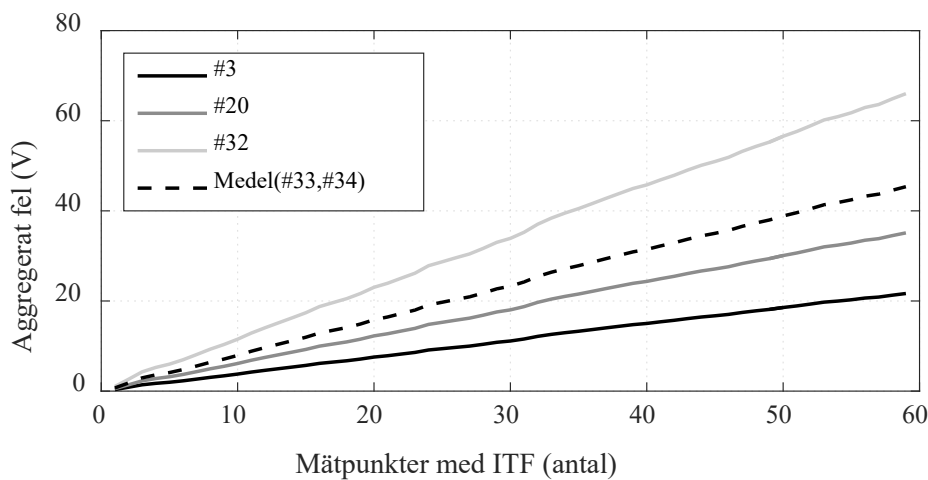
Figur 11: Förlusterna i det större exempelelnätet utan ITF.

Om sedan en ITF introduceras kommer att förlusterna att öka under vissa slumpmässiga tider.



Figur 12: Förlusterna i det större exempelnetet med ITF_2 .

Figur 12 visar liknande beteende som i det mindre exempelnetet men med lägre förlustnivåer i procent av överförd effekt. Här kommer flera tidsperioder ITF hamna under gränsen på 10%. Det är därför viktigt att veta hur mycket förluster som bör finnas i sitt distributionsnät.

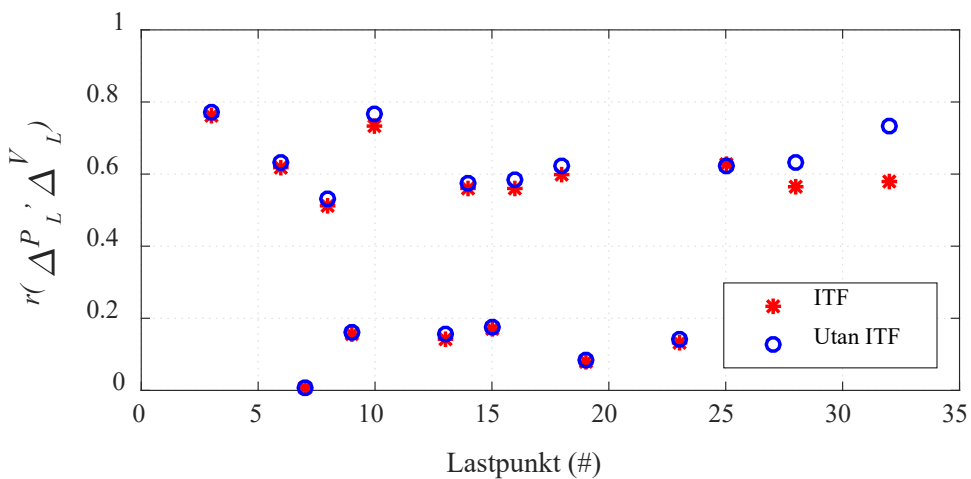


Figur 13: SiM resultatet på det större elnätet i Figur 2.

SiM resulterar i färre detekterade fall, då många perioder med mindre än 10% förluster ändå passerar obemärkt förbi. Ändå särskiljer sig #32 där den ITF är lokaliserad. Samtidigt följer närliggande lastpunkter (exempelvis #33 och #34) nära inpå. Enligt enkätsvaren från elnätsägarna kan det vara tillräckligt att lokalisera en fördelningspunkt, dvs. kopplingspunkt #29.

Korrelationsstudier – K;SE - Stort exempelnät

K;SE utfördes också på det större elnätet med syftet att se hur långt man kan komma med enbart mätvärden från energimätare så länge spänning och energi finns tillgängliga i lastpunkterna.

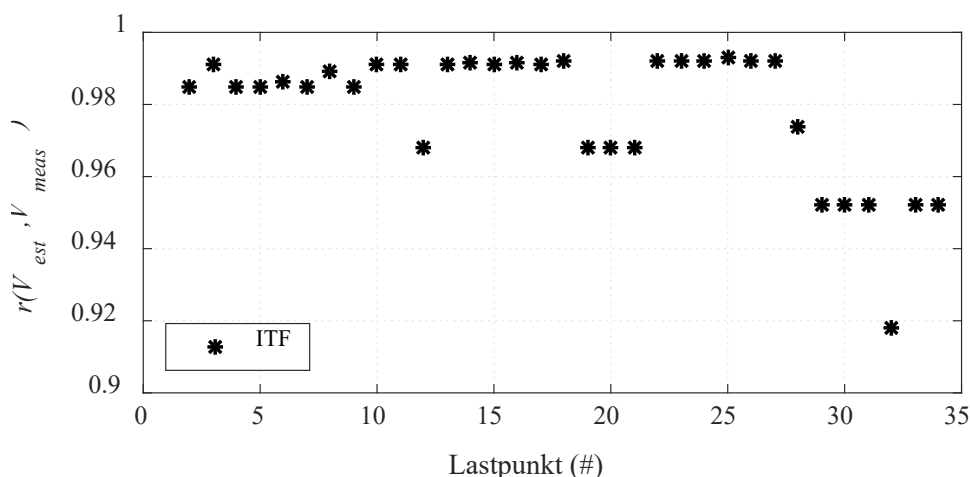


Figur 14: K;SE vid olika lastpunkter med och utan ITF.

K;SE är inte verkar lovande i det större elnätet, även om vi här har mer eller mindre ideal data. Mätbrus borde dock försvinna i korrelationsstudier, något som är en stor fördel. K;SE använder som tidigare nämnt ingen nätsstationsmätning. Men om denna finns tillgänglig kan en korrelations studie göras relaterat till uppskattad spänning och uppmätt spänning.

Korrelationsstudier – Stort exempelnät med stationsmätning och systemkunskap

Om vi antar att vi har systeminformation som i SiM metoden kan korrelationen mellan uppmätt spänning och estimerad spänning beräknas. Denna korrelation bör vara nära 1, eventuella avvikelser bör indikera ITF. Denna metod kommer härnäst att benämnas K;V.



Figur 15: Korrelation mellan uppskattad spänning och uppmätt spänning ($K:V$) i det stora exempelelnätet med en ITF lokaliserad vid lastpunkt #32.

Även om skillnaden i korrelation i Figur 15 är små pekar korrelationen ut den väg genom elnätet som kan följas mot den ITF som är lokaliserad vid lastpunkt #32. Detta då punkt #12 är kopplingspunkten ovan i elnätet (se Figur 2), och #29-#34 ligger i samma område och lasten i lastpunkt #32 särskiljer sig relativt de andra. Viktigt är dock att notera skalan på y-axeln som här går från 0.9-1, dessa små skillnader kan vara ett problem för algoritmen. Då lokaliseringen helst bör peka mot en viss punkt kan det vara bra nog med lokalisering till kabelskåp enligt enkätsvaren från elnätsbolagen. Hur detta skulle påverkas i ett verkligt elnät med större mätosäkerhet vore ytterst intressant för framtida projekt.

3.2.3 Påverkan från varierande mätkvalitet

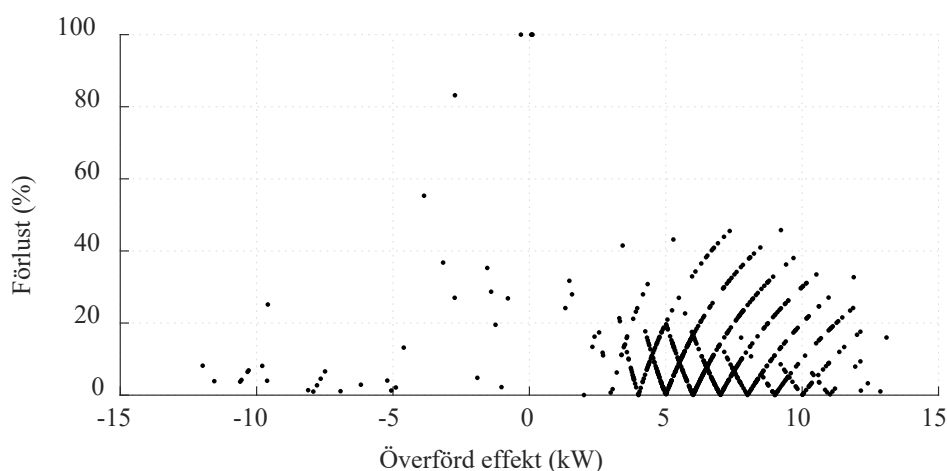
Energimätvärdena i de olika elnäten hade en upplösning på 10 Wh, detta kan dock variera stort mellan olika leverantörer. Det är därför intressant att se hur lokaliseringen i det mindre elnätet påverkades av en varierad upplösning hos energivärdena. Spannet som är av intresse är 1 Wh, 10 Wh, 100 Wh och 1kWh. Värdena för studien av 1 Wh interpolerades fram med en MATLAB funktionen *spline*. Denna funktion utför en kubisk interpolation av de ordinarie 10 Wh värdena och skapar då energivärden med flera decimaler. Dessa rundas sedan av till de olika upplösningarna i spannet (1 Wh, 10 Wh, 100 Wh och 1kWh).

Upplösning på energimätvärdena - SiM

I det lilla exemplarnätet utvärderades SiM. Detta genom att antalet timmar med höga förluster (mer än 10%, som tidigare) observerades samt avvikelsen mellan medelvärdet av lastpunkterna #3 och #5 jämfördes gentemot spänningsfelet hos #6, liknande som i Figur 10.

Resultatet visade att 1-100 Wh, gav liknande resultat. 145 detekterade timmar med ITF och liknande aggregerat spänningsfel vid timme 145 på 64,4 V. Först vid en upplösning med endast kWh noterades 241 timmar med ITF. Spänningsavvikelsen

mellan medelvärdet och lastpunkten ändrades till 65,3 V. Att skillnaden uppstår först här och inte vid exempelvis 100 Wh tros bero på triggingsgränsen på 10% (förluster av överförd effekt) som används som vid utvärderingen av SiM. Nätstationsmätningen antas ha en upplösning på 1 Wh och dessa avvikelser skapar senare vid låga överförda effekter stora procentuella skillnader som överstiger triggingsgränsen. Detta resulterar i fler registrerade timmar med ITF, även om detta inte var fallet. Detta kan ses i utspridningen av förluster gentemot överförd effekt.

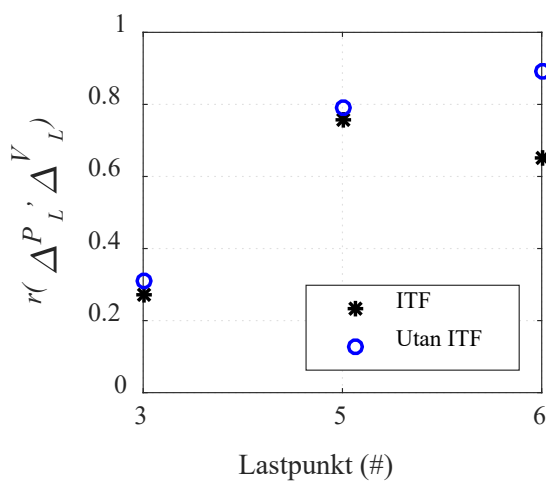


Figur 16: Uppskattade förluster i elnätet vid nätstationsmätning med upplösningen 1Wh och energimätare med upplösningen 1kWh.

Från Figur 16 ses den utökade spridningen jämfört med Figur 7, vilket indikerar vikten av upplösningen hos energiregistraren hos slutkundsmätning.

Upplösning på energimätvärdena – K:SE

K:SE utvärderades också på gentemot ett varierande spann på mätupplösning hos energidata. Här orsakades inte så stora variationer, primärt då spänning inte estimerades fram från energivärden, utan bara hur väl ändringar i spänning korrelerade mot ändringar i energi. Även om denna energiändring ibland hade låg upplösning.



Figur 17: K:SE vid 1kWh upplösning hos energimätarna.

Från Figur 17 kan observeras att en viss ökad spridning kan ses i korrelationskoefficienterna men en stor skillnad kan observeras vid lastpunkt #6. I övrigt påverkas resultaten marginellt.

4 Sammanställning: Maskininlärning för lokalisering av icke-tekniska förluster

4.1 VAD ÄR MASKININLÄRNING?

Maskininlärning (ML) är ett delområde inom datavetenskap som har utvecklats från studier inom mönsterigenkänning och artificiell intelligens. Rent förenklat kan man säga att ML uppstår i gränslandet mellan programmering/algoritmer och matematik/statistik. Statistik handlar om att lära människor om världen, men ML handlar det om att lära datorer om världen genom avancerade datadrivna algoritmer. ML kan leverera insikter och kunskap om data som kan användas till två olika saker: (i) prediktion, eller (ii) slutledning. Prediktion handlar om att göra förutsägelser om ett fenomen genom att träna en modell som har identifierat grundläggande mönster i data, som kan användas för att prediktera fenomenet med hänsyn till ny data. I prediktionstillämpningar är det givetvis centralt att identifiera modeller som har bra prediktionsförmåga på data, dvs. har låga fel. Slutledning handlar om att förstå hur variabler är kopplade och få insikter om den underliggande strukturen i datamängden. ML:s popularitet har vuxit kraftigt de senaste åren och det beror bl.a. på:

- Allmän digitalisering av samhälle och industri
- Ökad utrustning av sensorer och annan mätutrustning som genererar stora mängder data. Internet of Things (IoT) där enskilda komponenter är uppkopplade mot näten kommer accelerera den här utvecklingen ytterligare. En trend är också att data släpps öppet och gratis; exempelvis från svenska myndigheter.
- Mer data ger bättre algoritmer, eftersom dessa tränas med hjälp av data.
- Datalagring har större kapacitet och gått ner i pris vilket ger möjligheter att spara undan stora mängder data.
- Beräkningskraft har blivit billigare, bättre och mer tillgängligt. Det finns en uppsjö av molntjänster som erbjuder beräkningstjänster och datalagring.
- Kommunikationsnäten expanderar och får bättre prestanda, ex. genom utbyggnad av fiber-och mobilnät.
- Människors har ett större behov av skräddarsydda produkter och tjänster, för specifika behov och intressen. Den traditionella industriella modellen att massproducera en vara till en större kundbas är inte längre lika relevant. Netflix och Spotify är tjänster som skräddarsyr utbud och rekommendationer till användarna baserat på deras unika preferenser.

Dessa utvecklingar gör det mycket intressant att använda ML i en rad olika tillämpningsområden, såsom bilindustrin, hälsoområdet, energi, etc. Insikter om data gör det möjligt att utveckla nya tjänster som antingen kan effektivisera etablerade processer (sänka kostnader) eller identifiera helt nya marknadsområden. Som exempel kan man nämna att smarta mätare har rullats ut till alla svenska hushåll. Elförbrukningen kan mätas per timme, dessa mätvärden skickas och lagras digitalt hos nätägaren. Denna databank kan användas i ML-tillämpningar för att svara på frågor som kan ha varit blinda för nätägare tidigare,

såsom: sker energistöld i mitt nät, vad för typ av förbrukningsmönster har enskilda kunder och kan jag göra något för att ändra detta, etc.

4.2 HUR FUNKAR DET TEKNISKT?

Det första tekniska steget inom prediktion är att definiera en hypotes om hur sambandet mellan en beroende variabel (fenomenet man vill prediktera) och N stycken beroende variabler (mätbara storheter som man antar påverkar fenomenet) kan se ut. Det enklaste fallet är att sambandet är linjärt och med endast en variabel (linjär regression), som följande:

$$y = c_0 + c_1x + \varepsilon$$

Där y är den beroende variabeln, x är den oberoende variabeln, och ε är feltermen. Givet detta grundläggande samband så vill man skatta parametrarna c_0 och c_1 m.h.t. till en uppsättning data som gör anpassningen till det beroende variabeln så bra som möjligt. De optimala parametervärdena identifieras genom att minimera en fördefinierad kostnadsfunktion $J(c)$. En vetertagen kostnadsfunktionen *least squares*, där man vill minimera summan av den kvadratiske skillnaden mellan det predikterade och uppmätta värdet:

$$J(c) = \frac{1}{2m} \sum_{i=1}^m (h_c(x_i) - y_i)^2$$

Där m är antalet träningsexempel, $h_c(x_i)$ är det estimerade funktionsvärdet för träningsexemplet x_i , och y_i är det observerade värdet. Parametrarna skattas iterativt med exempelvis Newton's metod, som söker det lägsta värdet för J . Hypotesen (modellen) och de skattade parametervärdena har sedan förhoppningsvis den bästa möjliga prediktionsförmågan på ny data.

Vilka modellparametrar man inkluderar beror på: (i) vilka data man har tillgång till (variabler), och (ii) antagandet om funktionens grundläggande form, ex. linjär kontra icke-linjär. Poängen är att väga (i) och (ii) så att felet blir så lågt som möjligt med träningsdata. Här kommer dock en av huvudutmaningarna med ML; ju mer data du har (antalet variabler) och ju mer flexibla modeller du använder (icke-linjära), desto lägre blir prediktionsfelet. Detta beror på att modellen blir så flexibel att den anpassar sig till all variation som finns i data. Problemet är att denna variation kan bero på slumpmässighet (brus) och inte enbart på fundamentala samband (signal). Vad detta innebär i slutändan är att prediktionsfelet på ny uppsättning testdata kan bli högt. Felet kan t.o.m. bli högre än vad en enkel linjär modell kan uppnå. Detta betecknas som problematiken med överanpassning, och är en av de primära utmaningarna med ML; att identifiera en modell med en balans mellan flexibilitet och underliggande samband som ger det lägsta möjliga testfelet. Denna balans kan man uppnå på flera olika sätt, såsom att (i) använda den mest lämpliga modellformen för problemet, (ii) använda de viktigaste oberoende variablerna givet den expertkunskap som finns, och (iii) använda bestraffningsåtgärder på parameterskattningen i modellen (minska vikterna för variablernas påverkan). För (iii) finns olika alternativ för att hitta den bästa

modellen; såsom Aikake Information Criteria (AIC), Bayesian Information Criteria (BIC), the lasso eller ridge regression.

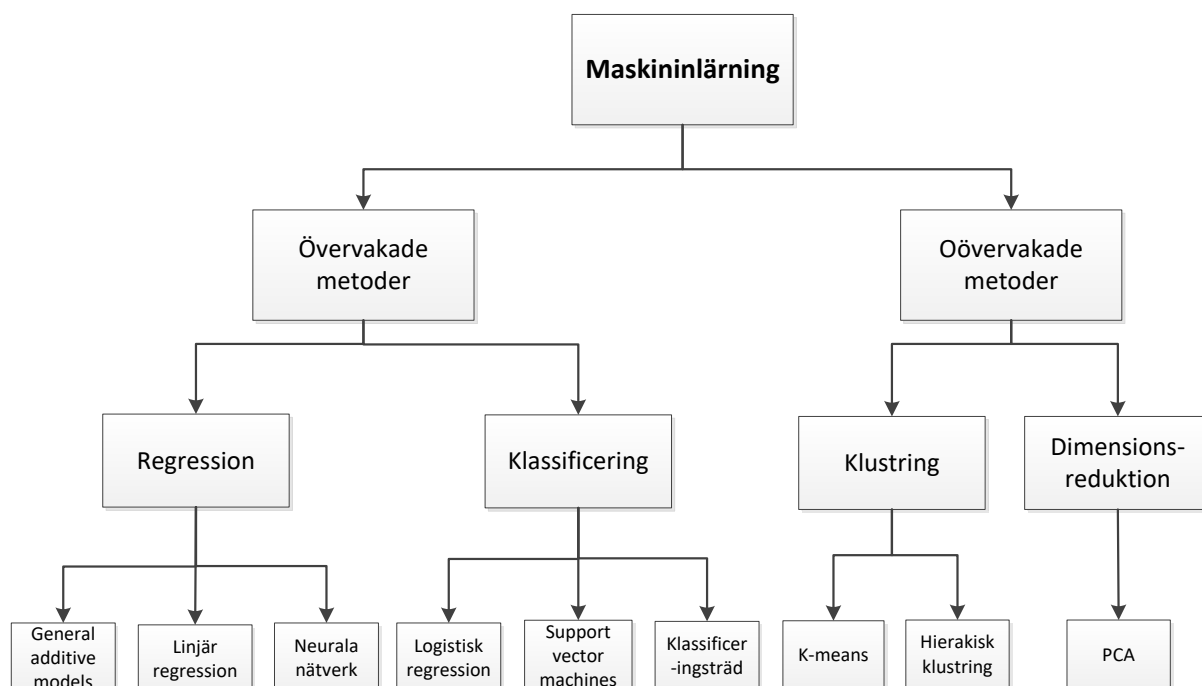
4.3 VILKA METODER INGÅR I ML:S VERKTYGSLÅDA?

I den här sektionen ska en kortare översikt om vilka typer av metoder som ingår i maskininlärningens verktygslåda. Den första nivån i indelning är övervakade resp. oövervakade metoder. Övervakade metoder innebär att man har en uppsättning oberoende variabler som ska användas för att prediktera en beroende variabel. Här ska man alltså korrelera ett antal prediktorer mot en svarsvariabel, och på så vis få ut mått på hur bra modellen är på att göra förutsägelser. I en oövervakad metod har man bara oberoende variabler, där man vill förstå hur grupper av dessa variabler är sammankopplade.

De respektive typerna har underfamiljer av metoder i nästa nivå. Övervakade metoder är uppdelade på regression och klassificering vilket skiljer sig på om det beroende variabeln är kontinuerlig eller kategorisk. Ett regressionsexempel är om man vill prediktera bostadspriser (kontinuerlig variabel) t.ex. givet bostadens storlek, månadsavgift och geografiska läge. Att t.ex. prediktera om en patient har cancer givet olika riskfaktorer är exempel på ett klassificeringsproblem då svaret är om personen har cancer eller inte, dvs. en nivå-variabel. Båda familjerna har specifika metoder som är listade i diagrammet nedan. Det som skiljer sig mellan de specifika metoderna är:

1. Komplexitetsnivå. Linjär vs. icke-linjära modeller.
2. Representerbarhet. Gråbox vs blackbox; alltså vilken typ av förståelse man kan extrahera från modellerna. Ett neuralt nätverk är exempel på en blackbox-metod då en människa inte kan ge en intuitiv förklaring till varför modellen ser ut som den gör. I en gråbox-modell kan man särskilja explicita kopplingar mellan variabler.

Vidare så består oövervakade metoder i huvudsak av klustring och dimensionsreduktion. Klustring handlar om att gruppera närliggande observationer genom att mäta avstånd mellan dessa datapunkter. Klustring är kraftfullt då man kan identifiera nyckelmönster om hur de oberoende variablerna hör ihop och förstå hur man kan dela upp data i relevanta delmängder för fortsatt dataanalysarbete. Dimensionsreduktion går ut på reducera antalet variabler i datamängden genom att hitta riktningar där data samvarierar som mest, och sedan använda dessa riktningar som approximationer till ursprunglig data. Beroende på egenskaper hos datamängden så kan den första riktningskoefficienten stå för en majoritet av den totala variationen.



Figur 18: En översikt av de metoder som finns tillgängliga i maskininlärningens verktygslåda.

4.4 HUR KAN MAN ANVÄNDA AI FÖR ATT IDENTIFIERA OCH LOKALISERA ICKE-TEKNISKA FÖRLUSTER I DISTRIBUTIONSNÄT?

I ML är identifikation av icke-teknisk förlust ett övervakat klassificeringsproblem (se diagram ovan för navigering i ML-djungeln).

4.5 VAD ÄR POTENTIALEN OCH FÖRDELEN MED ATT ANVÄNDA ML I JÄMFÖRELSE MED ANDRA METODER (EXEMPELVIS LOAD FLOW)?

Det finns andra typer av modeller och tillvägagångssätt att tillgå när man ska prediktera eller förstå något fenomen eller system. Ett av de främsta alternativen till ML är fysikaliska vita modeller, där man har grundläggande fysiska samband mellan variablerna, som är härledda från observationer av fenomenet och kunskap inom området. Om de samband man definierat är en bra representation av verkligheten, så kommer modellen vara robust, träffsäker och allmängiltig. Nackdelen med vit modellering är man vanligtvis måste ta fram modeller manuellt för varje respektive nytt problem, och kalibrera parametrarna för att uppnå tillräcklig noggrannhet. Dessa aktiviteter kostar givetvis både tid och pengar. Dock så finns det ju även fenomen som inte är fysikaliska i sin natur utan snarare handlar om information och beteenden. Exempelvis så är en klassificeringsalgoritm som kan särskilja spam-mail från vanliga mail inte baserade på fysiska samband, utan snarare samband och mönster i själva informationen i mailen som gör det möjligt att dela upp den i kategorier. Sådana algoritmer är helt datadrivna då själva informationen (mailen) sitter på mönstren, och att en självlärande algoritm kan lära sig själv att hitta de mönstren. Detta gäller även elstöld då beteenden inte baseras på fysikaliska lagar, utan snarare på irrationella och giriga beteenden hos slutanvändare.

En fördel med ML är att man inte behöver explicit programmera modellerna, utan att man givet ett flöde av indata själv kan låta programmet hitta den modell som är mest lämpligt (typiskt lägsta möjliga testfel). Då det finns mycket data tillgängligt via energimätare, gör denna approach fördelaktig i och med att den är lättare att implementeras.

Argument varför man använder AI i detektering av icke-tekniska förluster:

- Man behöver inte ha information om topologi för nätberäkningar. I synnerhet i utvecklande ekonomier så kan elnätet förändras väldigt snabbt p.g.a. en ökad efterfrågan.
- Förändringar eller omkopplingar i elnätet vilket leder till felaktiga nätberäkningar.
- Det kräver mätdata från transformater, ledningar och kunder som mäter vid samma tider.

En mer flexibel och anpassningsbar metod är Artificiell Intelligens. AI gör det möjligt att analysera kundernas lastprofiler och identifiera oregelbundna mönster. Detta kan trigga kundinspektioner för nätägaren som hjälper dem att hitta eltjuvar eller fel som orsaker ekonomiskt bortfall.

Inspektioner är dock dyra då det behöver fysiska besök från teknisk personal.

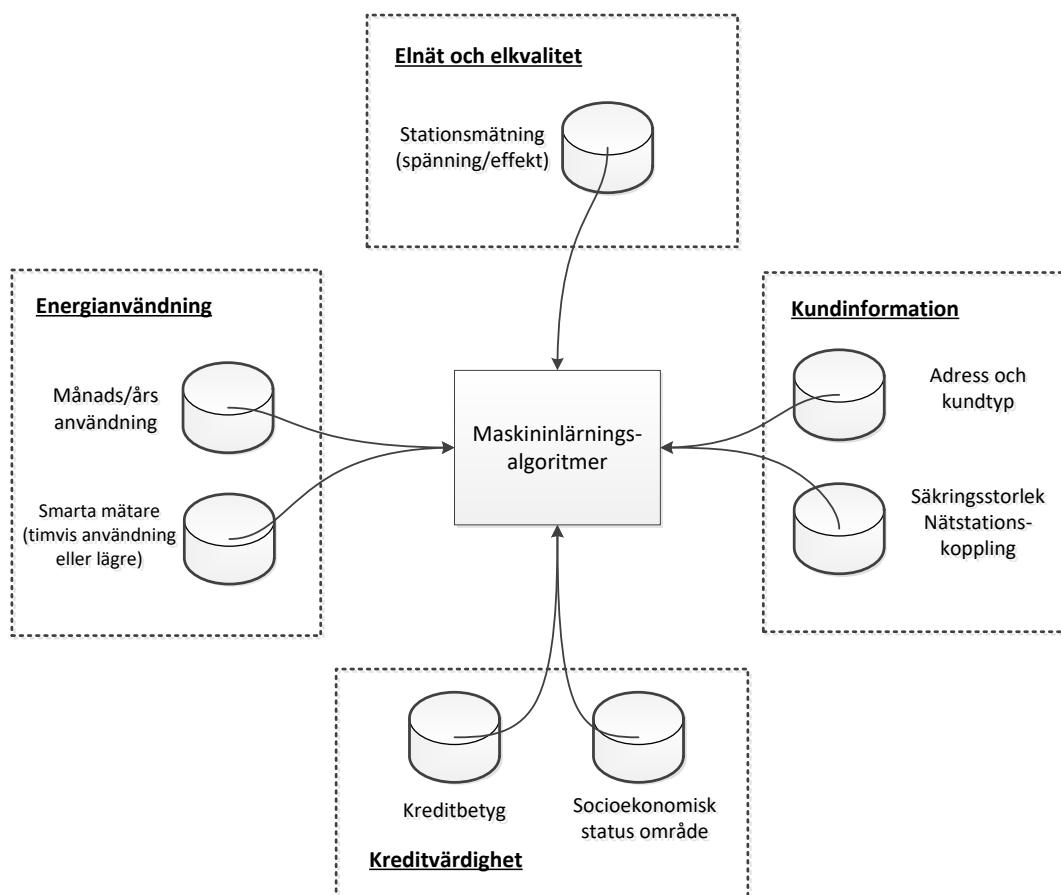
Två olika metoder finns tillgängliga;

1. *Expertsystem* - innefattar manuellt framtagna regler som används till beslutsfattande.
2. *Maskinlärning* – datorer lär sig från exempel utan att bli explicit programmerade

Historiskt sätt så har ITF-identifieringssystem baserats på domänspecifika regler.

ITF-identifiering kan räknas som ett specialfall av bedrägeriidentifiering, såsom dataintrång, kreditkortbedrägeri, etc.

Vilken typ av data som behövs för att skapa olika features till ML-algoritmerna. I diagrammet nedan finns en översikt om vilka typer av kund och elnät relaterad data som kan användas till att skapa features som går in i en maskininlärningsalgoritm.



Figur 19: De olika datakällorna som är möjliga för att skapa features som används till att träna maskininlärningsmodeller inom prediktion av icke-tekniska förluster

I figuren ser man fyra olika domäner av data som kan användas till att skapa features till ML-algoritmerna för detektion av icke-tekniska förluster:

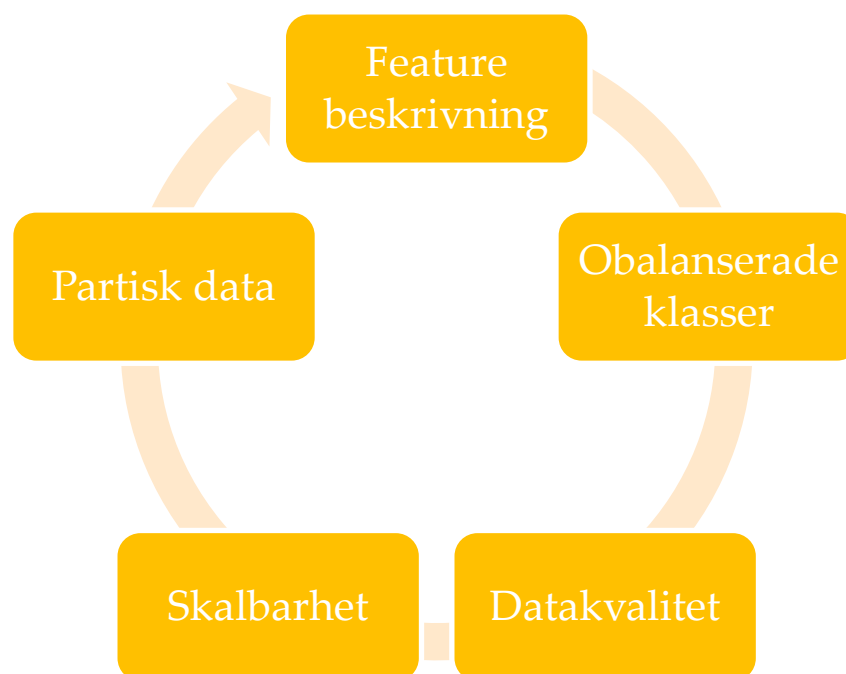
1. *Energianvändning per kund*: månads/årsförbrukning från äldre mätare eller timvis förbrukning eller högre upplösning från smarta mätare.
2. *Elnät och elkvalitet*: effekt, energi och spännings-mätningar på nätstationsnivå på sekund, minut eller kvartsupplösning. Ger info om driftstatus och möjlighet att jämföra förbrukning från stationen med den aggregerade förbrukningen från de underliggande smarta mätarna (om det finns; se datadomän ovan)
3. *Kundinformation*: information från nätbolagets kundinformationssystem; såsom adress, typ av kund, säkringsstorlek, och nätstationskoppling.
4. *Kreditvärdighet*: data om kunders betalningsvilja och hantering. Detta kan hämtas från nätbolagets debiteringssystem. Information om områdets socioekonomiska status där kunden bor kan även hämtas från exempelvis SCB, såsom medianlön, utbildningsnivåer, etc.

Notera att viss data är knepig att släppa; såsom kreditvärdighet för enskilda kunder då det kan strida mot personskyddslagar (i synnerhet extra känsligt med den nya lagen GDPR). Vidare data såsom stationsmätning eller timvis energimätning hos kund är kanske inte tillgänglig eller saknas. Denna datatillgänglighet ställer hårda krav på vilka typer av algoritmer som kan

utvecklas och där med vilken träffsäkerhet man kan uppnå i att identifiera icke-tekniska förluster. En viktig sak att nämna här är att testa olika modeller med olika typer av data och features för att se vilken träffsäkerhet man kan uppnå; då kan man kvantifiera värdet av att utöka mätning eller tillgängliggöra känslig data från olika system, då en ökad träffsäkerhet ger en potential till ökad intäkt (stoppa elstöld) och minskade kostnader (minimera fall där man skickar ut tekniker till felaktigt identifierade eltjuvar).

4.6 VAD ÄR UTMANINGARNA MED ATT ANVÄNDA ML I OMRÅDET?

I figuren nedan har de olika utmaningarna inom att använda ML för att detektera icke-tekniska förluster i distributionsnät. Varje utmaning beskrivs i separata sektioner nedan. För dessa utmaningar presenteras ett antal förslag vars syfte är att adressera dem, som sedan tas hänsyn till när det tillämpade AI-fallet utformas.



Figur 20: De olika utmaningsområdena inom ML och detektering av icke-tekniska förluster

Featurebeskrivning

De features man tar fram från rådata har stora komplikationer och inverkan på de klassificeringsalgoritmer man tar fram i nästa steg. Primärt tas features fram från två olika typer data; energianvändning och kundinformation (se Figur 19). Notera att features som kan tas fram från timdata och månadsdata är helt olika, och inte nödvändigtvis utbytliga. Om det finns nät där man både har månadsmätare och smarta mätare så kommer detta ställa krav på vilka typer av features som kan tas fram. Här skulle man vilja ha features som är så allmängiltiga och robusta som möjligt; dvs. är möjliga att ta fram i många olika kontexter och elnät. I Sverige kan dock detta vara ett mindre problem då vi har haft en bred utrullning av smarta

mätare och en stundande ny utrullning ligger nära i tiden. I andra länder och ekonomier (i synnerhet U-länder) kan detta dock vara en stor utmaning.

Obalanserade klasser

I en klassificeringstillämpningar såsom att identifiera eltjuvar kommer det finnas en utmaning kopplat till att förhållandet mellan antalet eltjuvar och totala antalet kunder är mycket lågt, dvs. det är ett relativt ovanligt fenomen. Detta ger viktiga konsekvenser och formkrav för utvärderingen av algoritmernas prestanda. Säg att man har en eltjuv i ett totalt urval om 1000 kunder. Om man antar att det inte finns en enda eltjuv så kommer algoritmen ha en träffsäkerhet på 99.9%. I verkligheten är ju algoritmen ytterst dålig på att identifiera tjuvar. Vidare om algoritmen predikterar att alla kunder är eltjuvar så kommer den med säkerhet hitta tjuven, dock så kommer detta trigga många onödiga inspektioner som är kostsamt. Därför behöver utvärderingen av algoritmen beakta två saker: (i) hur bra den är på att identifierar riktiga eltjuvar (true positives), och (ii) hur ofta den predikterar eltjuvar som i verkligheten inte är det (false positives). Denna typ av utvärdering ger en rättvis återspeglning av hur väl prediktionsmodellen kan ge bra beslutsunderlag som nätägaren kan fatta beslut på (dvs. stoppa eltjuvar och inte skicka iväg inspektörer på felaktiva uppdrag).

Partisk data

Den stora nackdelen med maskininlärning är att modellernas inte nödvändigtvis är användbara i andra kontexter eller andra uppsättning data. Exempelvis om man har tränat en modell med historisk data för att prediktera ett visst fenomen, och detta fenomen har ändrat sig så kommer modellen nödvändigtvis inte ge bra förutsägelser framöver. Då behöver modellen tränas med data från den nya situationen för att kunna göra bättre förutsägelser. Alltså; modellerna kan inte ge bättre prediktioner än de grundläggande mönster som finns i historiska data. Detta gör ju att resultaten måste analyseras med tillförsikt, varav vitboxmodeller är mer allmängiltiga, tänk typ Newton's lag.

Datakvalitet

Verklig data är ej perfekt och innehåller mycket brus. Detta gäller både för de beroende variablerna (eltjuv eller ej) och de oberoende variablerna (elanvändning). Att hantera brus i data är en stor utmaning generellt i AI, eftersom man inte vill att algoritmen ska modellera slumpmässig variation i data. Man kan hantera detta problem med regularisering eller att förenkla modellen. Oftast adresseras detta problem för de beroende variablerna. Dock så ska man vara medveten om att även de beroende variablerna kan vara utsatta för det här problemet. Exempelvis så kan en verklig eltjuv inte vara identifierad som en sådan i datat p.g.a. att kunden har mutat inspektören eller att fel data har rapporterats p.g.a. mänsklig faktor i rapportering. Detta kan ge en stor inverkan på resultat för prediktionsmodellen.

Skalbarhet

Modellerna ska helst vara skalbara för miljontals kunder; detta kan begränsas av snabbhet i algoritmerna och den processkraft man har tillgång till.

Datatillgång

Öppna databaser med stort antal exempel saknas nationellt vilket möjliggör att träna modeller.

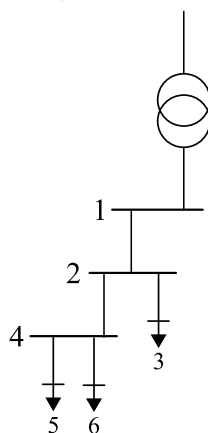
Tabell 2: Förslag på lösning för att hantera utmaningar.

Utmaning	Hanteringsförslag	Förväntad konsekvens
Featurebeskrivning	Använd <i>deep learning</i> för att automatiskt identifiera features från förbrukningstidsserier.	Deep learning kan hitta dolda korrelationer i datasetet och mer komplexa feature-hierarkier och därmed förbättra algoritmernas prestanda
Obalanserade klasser	Använd en <i>ROC-kurva</i> där man plottar TP:s mot FP. Areal under denna kurva är AUC som är ett måttetal mellan 0 och 1. Om AUC > 0,5 så är prediktionerna bättre än slumpen	Man har ett felmått som både adresserar träffsäkerheten att hitta riktiga eltjuvar, och inte peka ut falska sådana
Partisk data	I ML säger man att den som har mest data vinner, inte den bästa algoritmen. Detta är dock en sanning med modifikation, då man behöver den mest representativa data, alltså data som avspeglar den verkliga fördelningen. Man drar slumpmässig data	Förhoppningsdata som är mindre partisk, alltså eltjuvar finns enbart i vissa geografiska områden etc.
Skalbarhet	Använda map-reduce-arkitekturer vilket är en parallellberäkningsmodell där beräkningar görs på flertalet olika datorer	Man kan använda mer data utan att datorerna kraschar eller att beräkningstiden blir för lång

5 Exempel: Maskininlärning för lokalisering av icke-tekniska förluster

5.1 METOD

Vi använder en artificiell intelligens-metod med linjär regression tagen från Nikovski et al (Nikovski & al., 2013) för att identifiera och lokalisera icke-tekniska förluster i distributionsnät. I denna metod behövs två typer av data: (i) energimätning på nätstationsnivå, och (ii) energimätning hos enskild slutkund. Data kan vara på timnivå eller lägre. Se den schematiska skissen nedan för en sådan struktur.



Figur 21: En schematisk skiss över ett distributionsnät med en nätstation och tre underliggande kunder.

Den totala energiförlusten, $L(t)$, i det här lilla nätet måste vara den energi som flödar genom nätstation, minus den energin som förbrukas av de tre kunderna. Energiförlusten i varje tidssteg, t , kan i ekvationsform uttryckas som:

$$L(t) = E_{NS}(t) - \sum_{i=1}^n E_i(t)$$

Där $E_{NS,t}$ är energimätning i nätstation, och $E_{i,t}$ är energimätning hos respektive kund i . Totalt finns n stycken kunder i nätet. Förlusten kan delas in i två komponenter; tekniska förluster och icke-tekniska förluster. Tekniska förluster är oundvikligt då det är resistansförluster i ledningarna vid energiöverföring. Icke-tekniska förluster är dock aspekter man kan förbättra då det beror på mät kvaliteten (tidssynkning mät fel, avsaknad av data), omätt förbrukning eller exempelvis stöld i nätet. Den icke-tekniska förlusten O vid tidpunkt t kan beräknas genom:

$$O(t) = L(t) - L_{TF}(t)$$

Alltså, den icke-tekniska förlusten är den totala förlusten minus den tekniska förlusten $L_{TF}(t)$. De tekniska förlusterna beror bl.a. på nätets utformning, dimensionering, egenskaperna i kablagen samt kundernas lokalisering, dess förbrukningsprofiler och eventuella mikroproduktion. Den här metoden antar att varje kund är kopplad med nätstationen via en separat ledning. Detta är en grov förenkling och kommer påverka resultatets noggrannhet. En stor fördel med det

här antagandet är dock att ingen nättopologi eller nätmodellering behöver göras vilket både sparar tid i modelleringsinstatser och datainsamling. Ett av syftena med studien är att se under vilka förutsättningar träffsäkerheten i att identifiera och lokalisera icke-tekniska förlusterna fortfarande är acceptabla, exempelvis beroende på nätkomplexitet.

Vidare, metoden beräknar ledningsresistansen till varje kund. Detta görs genom att utnyttja följande samband mellan ström, resistans och energiförlust mellan två tidpunkter för kund i :

$$L_t = \frac{1}{3(I_{i,2} - I_{i,1})/(t_2 - t_1)} \int_{t_1}^{t_2} I_i(t) R_i dI_i \\ = \frac{R_i}{3(I_{i,2} - I_{i,1})/(t_2 - t_1)} [I_i(t_2)^3 - I_i(t_1)^3]$$

Eftersom de totala förlusterna är givna och strömmarna kan estimeras från energimätningarna med hjälp av antagen fasspänning (ex. 230 V) kan man lösa ut resistanserna genom Moore-Penrose pseoduinvers (UCLA, 2018):

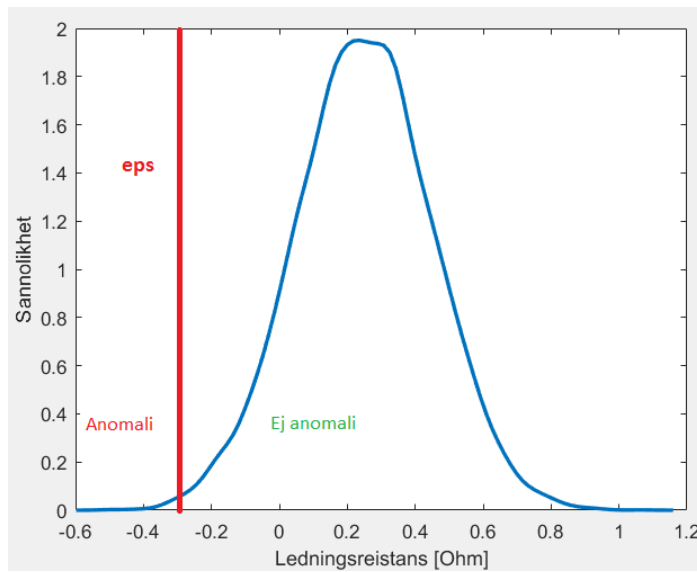
$$\hat{R} = (H^T H)^{-1} H^T L$$

L är energiförlustvektorn med alla som beräknas med (1), och H -matrisen är strömskillnaderna för varje tidpunkt och kund enligt ekvationen ovan:

$$H = \begin{bmatrix} \frac{I_1(t_2)^3 - I_1(t_1)^3}{3(I_{1,2} - I_{1,1})/(t_2 - t_1)} & \dots & \frac{I_n(t_2)^3 - I_n(t_1)^3}{3(I_{n,2} - I_{n,1})/(t_2 - t_1)} & \mathbf{1} \\ \dots & \dots & \dots & \mathbf{1} \\ \frac{I_1(t_m)^3 - I_1(t_{m-1})^3}{3(I_{1,m} - I_{1,m-1})/(t_m - t_{m-1})} & \dots & \frac{I_n(t_m)^3 - I_n(t_{m-1})^3}{3(I_{n,m} - I_{n,m-1})/(t_m - t_{m-1})} & \mathbf{1} \end{bmatrix}$$

Notera att R -värdena måste tränas med energidata där man med viss säkerhet vet att det inte skett elstöld eller något annat fel. Annars kommer R -värdena vara systematiskt underestimerade och vara dåliga på att prediktera fall av elstöld. Man kan kvalitetssäkra data genom att exempelvis beräkna medelförlusten i procent över en tidsperiod och jämföra med ett schablonvärde, typ 2 %. Om den avviker från schablonvärdet kan man ifrågasätta om det är lämplig data för modellträning. Med de estimerade ledningsresistanserna får man ut $L_{TF}(t)$ vilket tillsammans med $L(t)$ ger den icke-tekniska förlusten från ekvation

$$O(t) = L(t) - L_{TF}(t)$$



Figur 22: Visualisering av hur en anomali definieras i relation till en normalfördelning

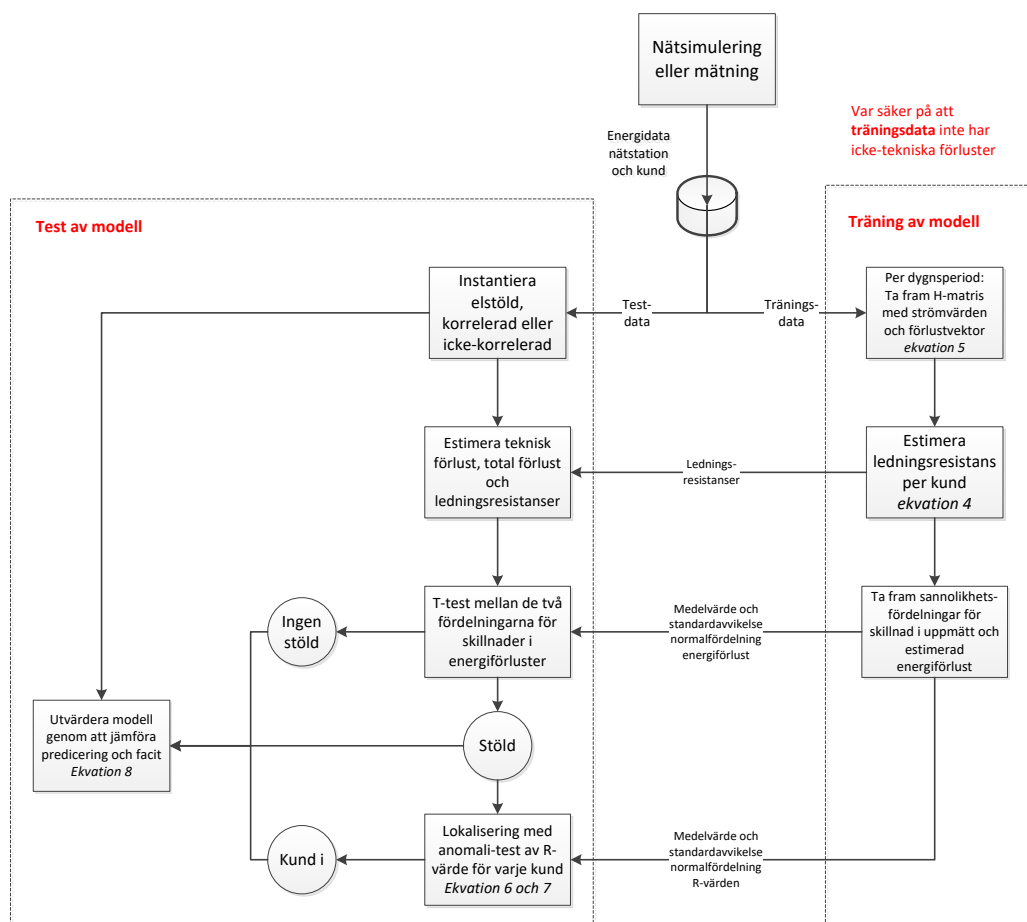
När modellen är tränad beräknar man energiförlustskillnaden mellan de uppmätta värdena (nätstation minus kundmätning) och de tekniska förlusterna (framtaget med ledningsresistanserna). Om en elstöld har skett ökar avvikelse mellan de två tidsserierna, och detta utvärderas analytiskt med ett t-test, som utvärderar om det finns en signifikant skillnad i medelvärde mellan två fördelningar. Om hypotesen inte kan förkastas tyder det på att det finns höga icke-tekniska förluster i nätet, och går vidare till att lokalisera var stölden sker genom att utvärdera skillnader i estimerade resistansvärden. Elstöld bidrar till att resistansvärdena sjunker och kan ge en indikation om vilken kund det är. Rent analytiskt definieras en anomali (dvs. kunden identifieras som eltjuv) om p -värdet är under ett visst värde ε , dvs. var den enskilda observationen ligger i förhållande till normalfördelningen (se för en visualisering av detta). p -värdet estimeras med följande funktion:

$$p_i(x_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x_i - \mu_i)^2}{2\sigma_i^2}}$$

Där μ och σ är medelvärde respektive standardavvikelsen för den skattade normalfördelningen för resistansvärden för kund i från träningsdata. Sedan klassificeras observationen x_i som en anomali om:

$$p_i(x_i) < \varepsilon$$

En översikt av AI-modellen för att identifiera och lokalisera elstöld finns i diagrammet nedan.



Figur 23: Ett översiktsdiagram av AI-metoden

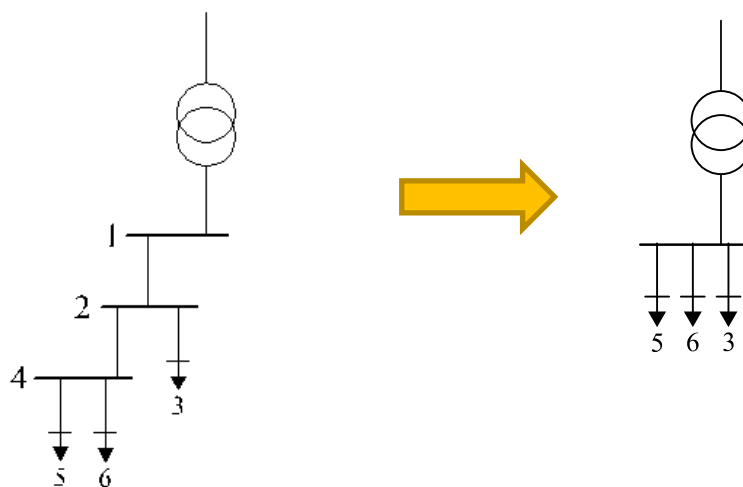
Träffsäkerheten Δ i modellen utvärderas med följande funktion:

$$\Delta = \frac{1}{n} \sum_j^n I(\hat{y}_j == y_j)$$

Där n är antalet prediktioner gjorda, $\hat{y}$ är det estimerade värdet och y är det riktiga värdet för observation j . I är identitetsoperationen vilket ger värdet 1 om villkoret stämmer, och noll om det inte stämmer.

5.2 FALLSTUDIE: ETT LITET ELNÄT MED STATIONSMÄTNING

Data från ett elnät med sex bussar i ett distributionsnät har använts som fallstudie för att testa och utvärdera modellen, se vänstra delen av Figur 24 för kopplingsschema. Systemet har tre kunder, varav två av dessa har soleanläggningar (kund 5 och 6). Data som ingår i användarfallet är timvis förbrukning för varje kund under februari månad, samt timvis simulerad energiöverföring från den överliggande nätstationen under samma tidperiod. Notera att nätet förenklats enligt det vänstra diagrammet i figuren för att kunna användas av metoden.



Figur 24: Vänstra diagrammet: kopplingsschema för elnätet i Herrljunga med tre kunder. Högra diagrammet: förenkling av nätdiagrammet med separata ledningar till varje kund

Data har delats in i två delar: (i) träningsdata vilket är energidata för de första tre veckorna, samt (ii) testdata som är energidata för den sista veckan. Modellen tränas med data-set (i), och utvärderas med data-set (ii). De riktiga ledningsresistanserna i nätverket är hämtade från nätbolagets Network Information-system och ska användas för att validera modellens förutsättningar att estimeras resistanser.

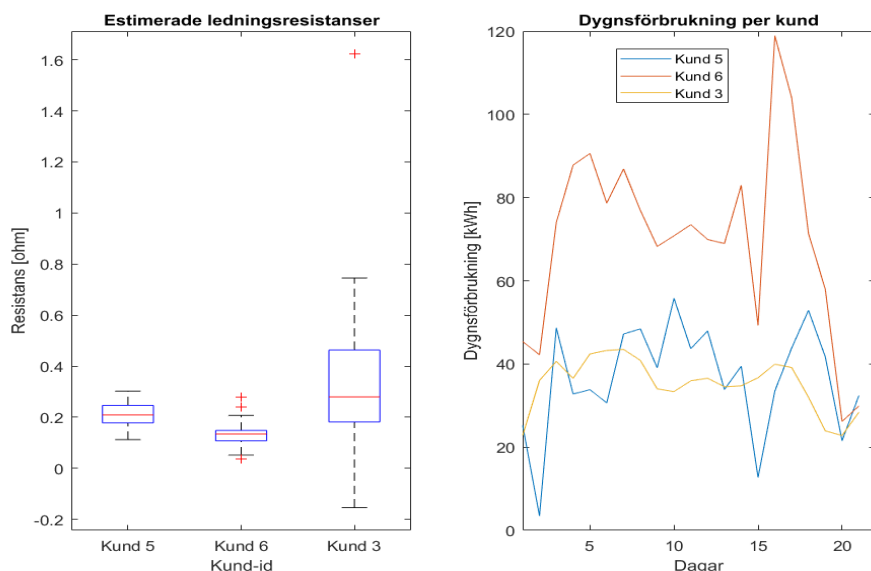
5.3 RESULTAT

Med modellen beskriven i metod-sektionen och data från användarfallet ska resultat produceras och utvärderas. Resultat genereras för ett antal olika scenarier, där ett scenario bestäms utifrån (i) nivå av elstöld i procent, (ii) vilken kund som utför stölden, samt (iv) vilka dagar stölden sker. Vi börjar dock med att utvärdera modellen utan någon elstöld först.

5.3.1 Basfallet: ingen icke-teknisk förlust

Med träningsdata (stations och kundmätning) estimeras ledningsresistanserna för varje 24-timmars perioder sekventiellt, dvs. 21 gånger totalt i hela träningsperioden (tre veckor). I figuren nedan visas de estimerade resistansvärdena i ett låddiagram, samt dygnsförbrukningen per kund för hela träningsperioden. Som synes så är variansen för de estimerade resistanserna störst för kund 3, vilket även täcker negativa värden vilket inte ska vara möjligt. För kund 5 och 6 så är variationen signifikant mindre. Notera att de riktiga resistansvärdena är cirka 0,25 ohm för de respektive tre ledningarna, vilket täcks för kund 5 och kund 3, men verkar underskattas något för kund 6. I den högra figuren syns den ackumulerade dygnsförbrukningen per kund för tidsperioden. Det verkar som att de estimerade resistansvärdenas osäkerhet (dvs. variation) är kopplade till hur mycket lasten varierar. Kund 3 har minst förbrukningsvariation men mest osäkerhet i resistansvärden, och för kund 6 är situationen tvärt om, dvs.

låg variation i resistans och hög variation i förbrukning. Notera att både kund 5 och 6 har solel.



Figur 25: Vänstra figuren: låddiagram för estimerade resistanser per kundledning. Högra figuren: dygnsförbrukningsvärden per kund för treveckorsperioden.

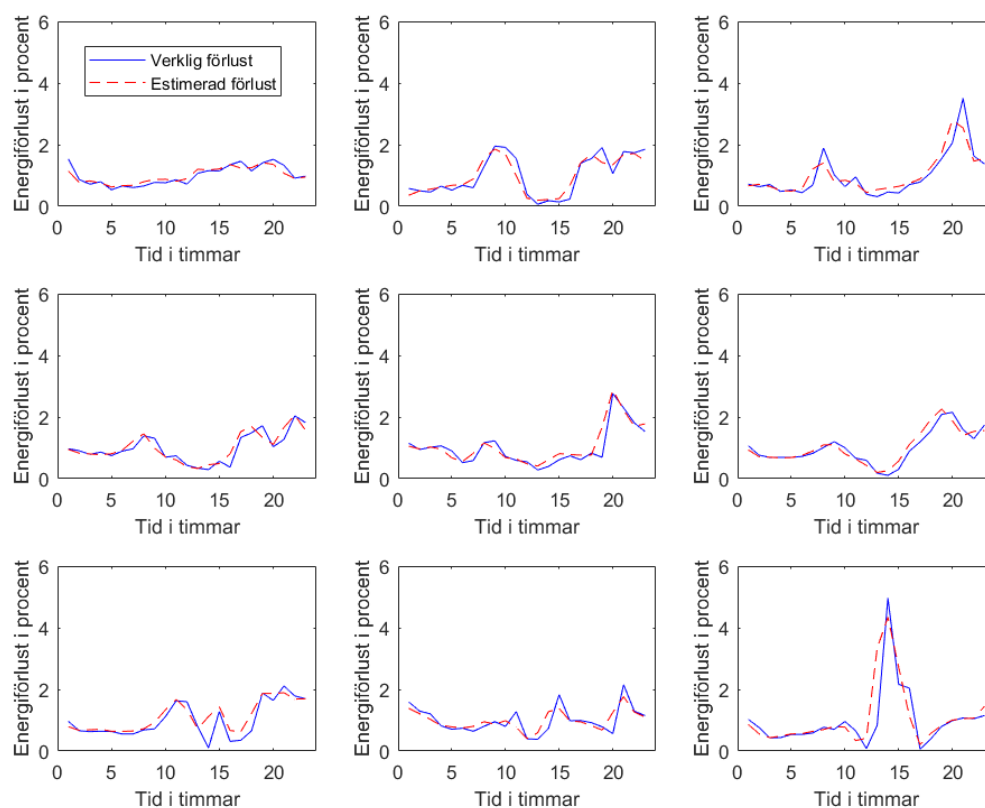
I tabellen nedan är olika statistik sammanställd för de estimerade resistanserna tillsammans med de riktiga resistansvärdena.

Tabell 3: Summerande statistik för estimerade resistansvärden per kundledning, samt det verkliga resistansvärdet. Enheten är ohm.

	Kund 5	Kund 6	Kund 3
Minimum	0,11	0,035	-0,15
Medel	0,21	0,13	0,37
Median	0,21	0,13	0,28
Max	0,30	0,28	1,62
Standardavvikelse	0,045	0,057	0,36
Riktigt resistansvärde	0,26	0,24	0,23

Genom att använda de estimerade resistansvärdena från träningsperioden så kan de tekniska energiförlusterna estimeras för testperioden. I den vänstra figuren nedan ser vi den verkliga förlusten³ och den estimerade förlusten givet medelresistanserna från tabellen ovan i en tidsserie. Den högra figuren visar skillnaden mellan de verkliga och de estimerade energiförlusterna i ett histogram. Som synes i figuren så stämmer de estimerade förlusterna väl med de verkliga, dvs. modellen kan fånga det verkliga förlustmönstret. En linjär modellanpassning mellan de två storheterna ger en variationskoefficient (R^2) på 78,8% samt ett felmått (Root Mean Squared Error) på 12,9 Wh.

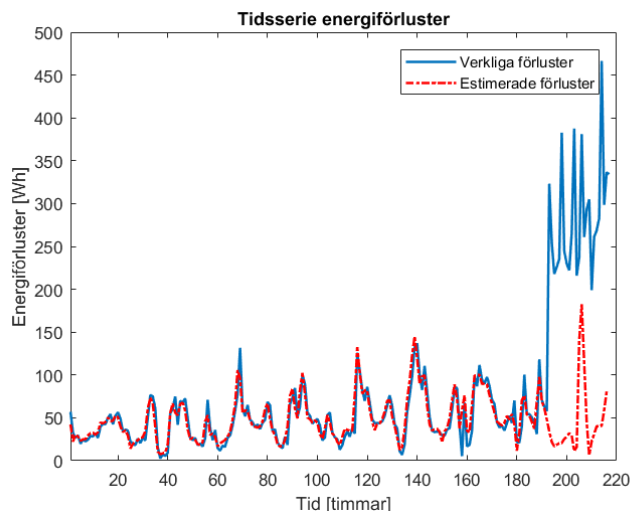
³ Stationsmätning minus kundmätning



Figur 26: Vänster figur: tidsserie över verkliga och estimerade förluster per dygnsperiod. Den blåa kurvan är de verkliga förlusterna och de röda streckade kurvorna är de estimerade.

5.3.2 Olika nivåer av elstöld

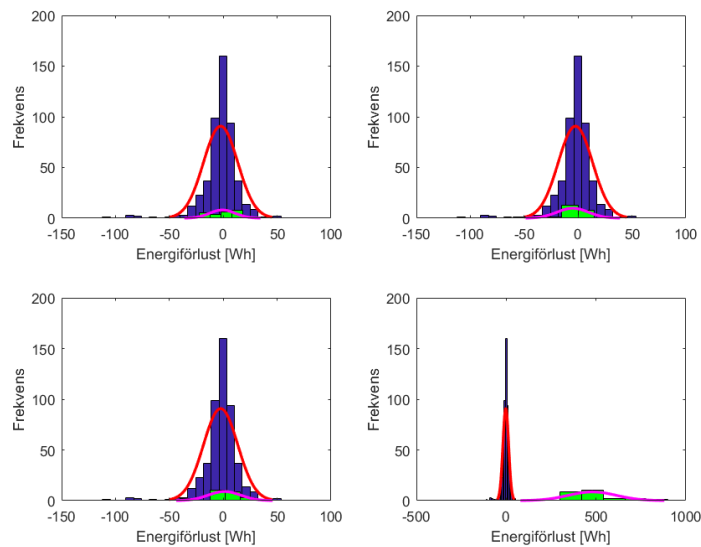
För att testa modellens förmåga att dels identifiera icke-tekniska förluster samt även lokalisera vilken kund som orsakar förlusten så ska olika nivåer av stöld utvärderas på testdata. Här testas olika nivåer av stöld genom att multiplicera den ursprungliga förbrukningen med en skalfaktor a som är ett tal mindre än 1. Är skalfaktorn 0 så stjälar kunden hela leveransen, och om den är 0,5 så är halva leveransen stulen etc. Första fallet är att kund 3 stjälar el det sista dygnet av testperioden. Vi kommer testa olika nivåer av korrelerad stöld; från 0 % till 100 % stöld för att se vid vilken nivåer algoritmen kan identifiera och lokalisera den.



Figur 27: Jämförelser mellan estimerade förluster (tekniska förluster) och verkliga förluster (stationsmätning minus kundmätning) över tid. En 25 %-ig elstöld för kund 3 simuleras de sista 24 timmarna där man ser att den blåa solida kurvan avviker kraftigt från den röda streckade kurvan

I Figur 26 ser vi de verkliga och estimerade förlusterna för testperioden där vi simulerar en elstöld på 25 % av kund 3 de sista 24 timmarna av perioden. Som vi ser avviker de verkliga förlusterna avsevärt från de estimerade tekniska förlusterna från modellen, vilket indikerar att de icke-tekniska förlusterna är höga denna period. Avvikelserna visualiseras även i Figur 27 för de fyra sista 24 timmarsperioderna i testdata i respektive histogram tillsammans med de estimerade energiförlustskillnaderna i träningsdata. De röda kurvorna är sannolikhetsfördelningsanpassning för träningsdata, och de lila kurvnorna är samma sak fast för testdata. De vi kan se här är att för de tre första dagarna så är sannolikhetskurvorna av liknande form, där topparna på kurvorna sammanfaller vid samma nivå. Dock för dag fyra ser vi att fördelningen skjuter kraftigt åt höger för testdata, och att de två fördelningarna avviker från varandra. Gör man ett t-test mellan dessa två fördelningar kan man inte förkasta hypotesen att de inte är från samma fördelning. Hypotesen förkastades för de andra dagarna, vilket sammantaget visar att algoritmen kan identifiera icke-tekniska förluster. Algoritmen kunde identifiera icke-tekniska förluster ner till 0,8 % av totalförbrukningen, vilket innebär att algoritmen troligtvis även kan hitta mindre mätfel.

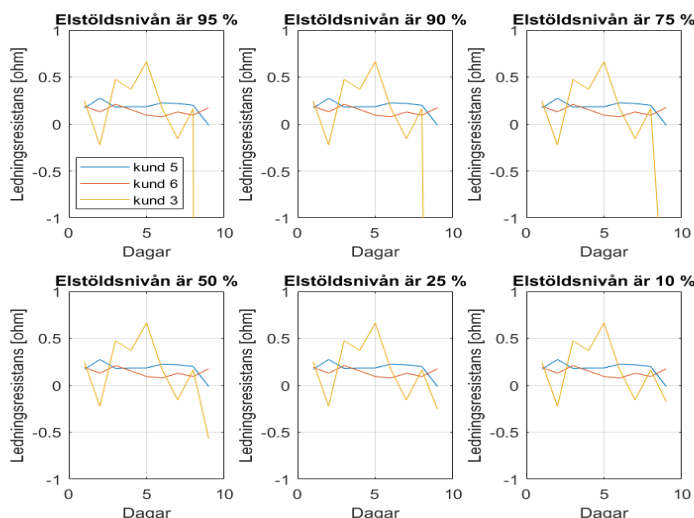
När icke-teknisk förlust har identifierats så är det relevant att se om algoritmen även kan lokalisera var stölden sker. Därför ska vi analysera hur de estimerade resistanserna förändras vid icke-tekniska förluster, och om man kan använda den informationen för lokalisering. Vi analyserar hur resistansvärdena förändras för varje 24-timmarsperiod i testdata och om eventuella avvikelser kan avslöja var icke-tekniska förlusten sker.



Figur 28: Histogram över energiförlustskillnad mellan uppmätt och estimerade energiförluster. De blåa staplarna är träningsdata med tillhörande röda anpassade fördelningskurva. De gröna staplarna och lila kurvan är samma typ av storheter men för testdata. Varje delfigur är data för en dag i testdatat. Figuren längst ner till höger är när stöld sker.

I Figur 29 ser vi de estimerade ledningsresistanser per kund och dygn under testperioden. Just i det här exemplet så är kund 3 eltjuven på olika nivåer under den sista dagen. Som man ser sjunker resistansvärdet kraftigt för kunden under sista dygnet när stölden är 75 % eller högre. Avvikelsen i resistans är mindre när stölden är 50 % eller mindre. Resistanserna för de andra kunderna är oförändrade trots stölden hos den tredje kunden. Nästa steg är att analytiskt lokalisera kunden genom att för de dagar där algoritmen har identifierat icke-teknisk förlust, jämföra de erhållna resistansvärden och se var avvikelsen är högst från resistansvärdena i träningsdata, och se om man på så vis kan lokalisera rätt kund vid olika stöldnivåer. Denna lokalisering görs enligt samma metod som för identifiering av icke-tekniska förluster, alltså stämma av sannolikhetsfördelning och var enskilda observationer ligger i förhållande till fördelningen (ekvationerna $p_i(x_i) =$

$\frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}}$ och $p_i(x_i) < \varepsilon$). Just i det här fallet kan algoritmen lokalisera stölden till kund 3 när stölden är högre än 10 %.



Figur 29: De estimerade ledningsresistanserna per dag i testdataperioden för varje kund givet olika elstödsnivåer hos kund 3.

5.3.3 Systematisk validering

I den här sektionen ska en systematisk validering göras för att få en generell träffsäkerhet på modellen, givet slumpmässig input av olika data. Validering görs för två olika typer stöder:

1. Korrelerade stölder, där stölden är proportionerlig till förbrukning
2. Icke-korrelerade stölder, där stölden är modellerad som ett vitt brus

Följande slumpmässiga input stoppas in i modellen för korrelerade stölder:

1. Slumpa ut vilken av kunderna som är eltjuv; likformig fördelning med lika stor sannolikhet att någon av kunderna är tjuven
2. Slumpa ut nivå på elstöld i procent. Följer en lognormalfördelning med ett medelvärde på noll, och olika nivåer på standardavvikelsen
3. Slumpa ut under hur många dagar stölden pågår under testperioden; likformig fördelning över totala antalet dagar i testperioden, dvs. 1 till 9 dagar.

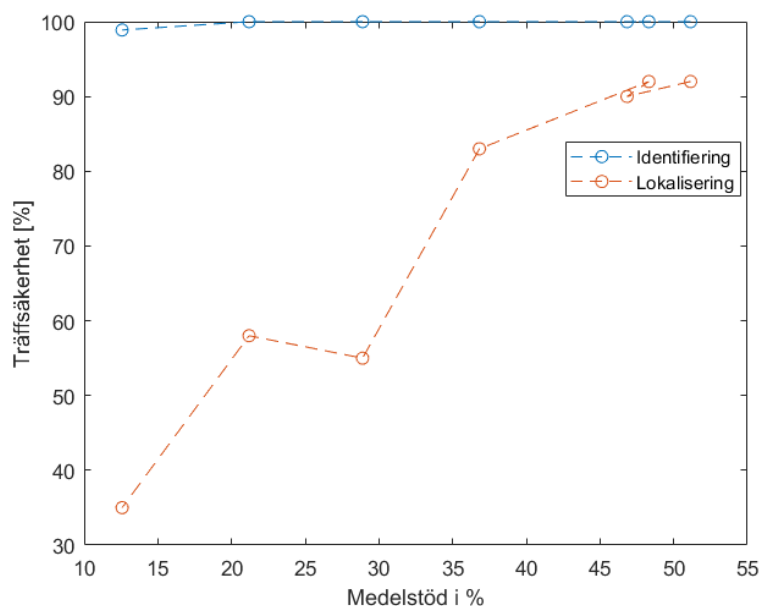
Notera att enbart en kund kan vara eltjuv i varje iteration.

Denna validering görs upprepande och iterativt 100 gånger för att få ut ett generellt resultat om hur träffsäker modellen är på att dels identifiera att stöld har skett (punkt 2 och punkt 3), för att sedan för dessa tillfällen beräkna träffsäkerheten i att lokalisera rätt kund (punkt 1). Denna prediktion görs enligt tidigare beskriven metod. Allt detta testas för respektive tjuv/fel. Vi börjar med att presentera resultaten för korrelerande stöld.

5.3.4 Korrelerade fel

I Figur 28 ser vi hur träffsäkerheten för identifiering av stöld samt lokalisering beror av nivån på medelstölden från lognormal-fördelningen i steg 2. Som synes i figuren så är träffsäkerheten för identifiering väldigt robust för storleken på stölden, där stölden måste vara på procent-nivå för att träffsäkerheten ska vara

under 100 %. Känsligheten är desto större för lokalisering vars träffsäkerhet ökar kontinuerligt som stölden ökar, tills den når ett tröskelvärde där den klingar av och inte ökar mer. En lokaliseringsträffsäkerhet på 95 % erhålls när stölden är på en nivå 45 % i medel. Slutsatsen man kan dra här är att algoritmen är dålig på att lokalisera tjuvar när stölden är liten, men är desto träffsäkrare för stora stölder.



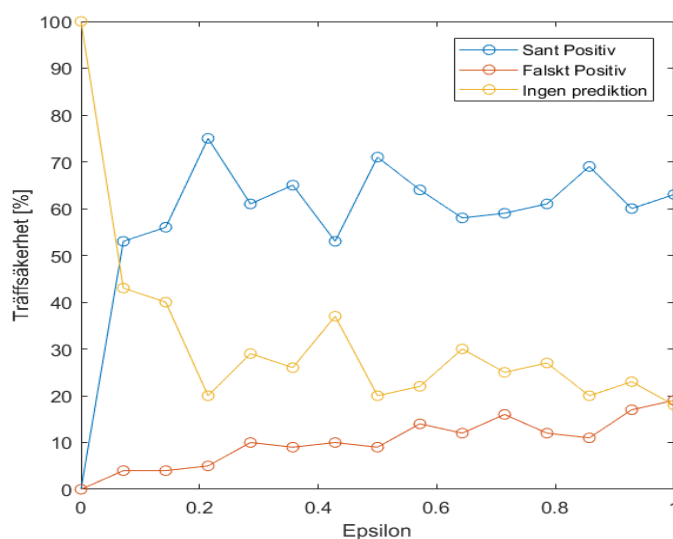
Figur 30: Träffsäkerhet för identifiering och lokalisering av elstöld som en funktion av medelstöld.

En till intressant figur att observera är hur antalet sant positiva (SP)⁴, falska positiva (FP)⁵, och ingen prediktion gjord beror på ϵ . Just detta visas i Figur 29 när medelstölden är 25 %. Som synes så görs ingen prediktion⁶ alls när ϵ är 0, eftersom tröskelvärdet för en anomali är så lågt att ingen kan pekas ut som eltjuv. Sedan ser vi hur antalet SP stiger när epsilon stiger. Dock så ser vi att andelen klingar av efter 0,2 och inte ökar mer utan snarare hoppar fram och tillbaka runt en träffsäkerhet på 65 %. Samtidigt ökar också andelen FP eftersom säkerheten i att den utpekade kunden är en riktig anomali också sjunker. Denna andel ökar kontinuerligt över ϵ . En logisk följd av det är att antalet icke-gjorda prediktioner sjunker också.

⁴ Modellen predikterar eltjuv vilket den också är i verkligheten

⁵ Modellen predikterar eltjuv men är det inte i verkligheten

⁶ Modellen kan inte peka ut någon som tjuv



Figur 31: Andel Sant Positiva, Falskt positiva och ingen prediktion som en funktion av epsilon. Notera att medelstölden är på 25 % i det här fallet

5.3.5 Slumpmässiga ITF

Samma typ av validering görs för icke-korrelerade fel och resultatet visar att träffsäkerheten på identifiering av stölder är nära 100 % för alla nivåer av stölder. Dock så kan algoritmen inte lokalisera kunderna alls när felen är icke-korrelerade, och hamnar på en träffsäkerhet mellan 0 till 5 % beroende på stöldnivå. Där bör istället metoderna i Kapitel 3 användas.

6 Slutsatser och diskussion

Rapporten visar baserat runt enkätsvar delvis designade lokaliseringsalgoritmer och detekteringslösningar för distributionselnät med nätstationsmätning. Elnäten som simulerats med verklighetstroga laster och verkliga parametrar har inte haft mer än 2% förluster. Vissa elnätsägare har beskrivit att man har nätstationsmätning och beräknar sina förluster per timma men saknar sedan utrustning eller tid att ge sig ut att lokalisera dessa. Algoritmerna som presenteras här kan vara ett steg i den jakten och minska resursåtgången vid lokalisering av icke-tekniska förluster. Då medelvärdet över årets förluster för flera distributionsnät är över 4% finns det goda incitament för elnätsägare att sätta personal på att minska sina förluster och därmed kostnader mot ovanliggande nät något som i fallet med omätt förbrukning även ger ökade intäkter men troligtvis en ökad medvetenhet om fel i databaser, rapporteringssystem och avvikande energimätare. Flera av elnätsbolagen har idag duglig upplösning på sina energivärden för att utföra ovanstående analyser (0.01 kWh) men detta förutsätter en högkvalitativ spänningsmätning i nätsstation och hos energimätarna samt god tidssynkronisering hos samtliga mätare. Om framtida krav slår igenom från Energimarknadsinspektionen riskerar däremot elnätsbolagen att överöas med data och algoritmerna som presenterats här kan ge lämpliga varningssignaler och indikationer för att förenkla driften.

Systemkunskap hos lokaliserings algoritmer skapar intressanta möjligheter i kombination med spänningsmätning. Lokalisering av stokastiska icke-tekniska förluster runt 2 kWh har påvisats med hjälp av SiM, K:SE och K:V. Samtliga algoritmer har olika fördelar och nackdelar; SiM kräver mycket data (nätstation, ledningsparametrar och energi och spänning hos slutkund) samt effektlödesanalys men ger mycket tillbaka när det gäller lokaliseringsmöjligheter och detekteringsmöjligheter. K:SE försöker tillgodose möjligheten att hitta ITF utan nätsstationsmätning om energivariationer och spänningsvariationer inte sammanfaller. K:V använder både SiM för att uppskatta spänningensvärden och ser sedan hur uppskattade och uppmätta spänningar sammanfaller med varandra. En verklig implementation av samtliga krävs för att avsluta kapitlet om lokalisering av icke-tekniska förluster. En stor fördel med korrelationsverktygen är att korrelationen inte bryr sig om eventuella mätfel (exempelvis 2 V för lågt) då dessa ändå kommer att korrelera med varandra med en konstant offset.

Maskinlärning (ML) söker efter mönster i data, då eltjuvar som här simulerats fram som slumpmässiga variationer är den föreslagna strategin oanvändbar. Enkätsvar har däremot påvisat att exempelvis strömstölden sällan är just slumpmässiga utan kan korreleras mot användarmönster av olika typer. Fördelarna med ML metoden som utvärderats är att den inte är i behov av elnätsparametrar. Kombinerat detta med nätstationsmätning kan träningsdata för algoritmen skapas då nätförlusterna är låga. Träffsäkerheten för identifiering väldigt robust för storleken på stölden, där stölden måste vara på procent-nivå för att träffsäkerheten ska vara under 100 %. Känsligheten är desto större för lokalisering vars träffsäkerhet ökar kontinuerligt som stölden ökar, tills den når ett tröskelvärde där den klingar av

och inte ökar mer. En lokaliseringsträffsäkerhet på 95 % erhålls när stölden är på en nivå 45 % i medel.

Fler typnät på distributionsnivå bör göras tillgängliga för att skapa en känsla över lämpliga förlustnivåer i elnät. Med effektlödesanalyser kan dessa förluster estimeras och påvisa på ett enkelt sätt hur stora förluster ett elnät borde ha vid vissa timmar.

Steg ett för vilken elnätsägare som är intresserad av sina förluster är ändå att med stationsmätning och god mätkvalitet uppskatta sina förluster per timma istället för per år. Detta kommer att indikera timmar med för höga förluster och troligtvis att lära elnätsägarna mer om sitt elnät, sina kunder och sina mätare. Konsekvent verkar det som att flertalet av icke-tekniska förluster som lokaliseras oavsett teknik har resulterat i minskade ekonomiska förluster även om ingen lokalisering algoritm används.

Inför nästa mätarutrustning är det viktigt att elnätsägaren själva sätter krav på sina energimätare, spänningsvärden kan vara ett klokt och relativt billigt sådant krav.

7 Citerade arbeten

- Bergh, F. v., Kadurek, P., Cobben, J., & Kling, W. (2011). Electricity theft localization based on smart metering. *Proceedings the 21st International Conference and Exhibition on Electricity Distribution (CIRED 2011)*.
- Energimarknadsinspektionen. (den 13 06 (2018-06-13)).
Sarskild_rapport_teknisk_data.xlsx. Hämtat från
<https://www.ei.se/sv/Publikationer/Arsrapporter/elstatsforetag-arsrapporter/>
- Energimarknadsinspektionen. (2015). *Ei R2015:07 Incitament för effektivt utnyttjande av elnätet*.
- Energimarknadsinspektionen. (2015). *Funktionskrav på framtidens elmätare (Ei R2015:09)*. Eskilstuna.
- Energimarknadsinspektionen. (2017). *Ei R2017:08 Funktionskrav på elmätare – Författningsförslag*.
- Hagmar, H., & Lindskog, A. (2017). *Bedömning av nätstatus baserad på dataanalys och avancerade algoritmer (2017:443)*. Energiforsk.
- Nikovski, D., & al., e. (2013). *Smart Meter Data Analysis for Power Theft Detection*. New York: Springer.
- Saadat, H. (1999). *Power system analysis*. New York: McGraw-Hill Companies.
- UCLA. (den 23 Maj 2018).
<http://www.math.ucla.edu/~laub/33a.2.12s/mppseudoinverse.pdf>.

Bilaga A: Intervjufrågor till elnätsföretagen

FRÅGOR TILL ELNÄTSBOLAG I PROJEKTET "DATAANALYS OCH AVANCERADE ALGORITMER – MÖJLIGHETER MED UTÖKAD MÄTINFRASTRUKTUR"

Introduktion

Projektet "Dataanalys och avancerade algoritmer – Möjligheter med utökad mätinfrastruktur", finansierat av Energiforsk, syftar till att öka kunskapen om hur olika distributionsägare kan använda informationen från ett elnät med en utökad mätinfrastruktur och fler mätparametrar i energimätare för att effektivisera driften. Huvudnyttorna är att:

1. påvisa nyttan med utökade mätparametrar hos energimätare i distributionsnät med nätstationsmätning.
2. Visa hur god mätkvalité kan leda till förbättrad lokalisering av icke tekniska förluster.
3. Specificera krav på nästa generations mätare för att kunna möjliggöra olika typer av analyser sett till lokalisering av omätt förbrukning.
4. Undersöka förändringar i kostnad för elnätsägaren vid ökad mätnytta, d.v.s. mätare med fler mätvärden (på grund av fler mätparametrar och ökad samplingsfrekvens).

Det är primärt med anledningen av Huvudnytta 4 (nytta och ev. kostnadsförändringar) som vi kontaktar er.

Övergripande frågor:

- Elnätsbolaget storlek? Antal kunder, geografisk storlek, GWh/år. Andra förutsättningar...
- Nätinformationssystemet. Hur noggrant? Kabellängder, kabeltyper? (resistanser vid analyser)
- Vilka är de största problemen och utmaningarna i ert elnät? Vad kostar de? På vilka sätt jobbar ni med dessa?
- Vad har ni för mätning i dagsläget, vilka storheter och hur tätt intervall?
- Har ni mätning eller övervakning i nätstationerna? Elmätare eller elkvalitetsinstrument? Använder ni det för planering och dimensionering av transformatorer och nät?
- Använder ni mätdata för att ha koll ev. förändrade effektbehov (solpaneler, elbilar, ...)?
- Hur går ni tillväga för att undersöka nätförluster i ert nät med hjälp av mätdata? Något slags verktyg och enkel analys? Effekt > säkring?
- Icke-tekniska förluster:
 - a. Vilka utmaningar och typer av problem har ni? Hur jobbar ni med dem?
Små/stora, landsbygd/tätort, områdesspecifikt
 - b. Vilket är vanligast: omätt förbrukning, elstöld och feldokumenterade nät? Exempel?
 - c. Exempel på fall där ni missat att mäta. Hur upptäcktes det?

- d. Hur identifierar ni att en mätare gått sönder?
- e. Efter identifiering kommer ju lokalisering – hur, kostnad?
- Har ni nån form av automatiska analyser som körs kontinuerligt och larmar?
- Loggas och bokförs dessa problem och åtgärder? Kartläggs nåt systematiskt, ser ni mönster? (Data kan i så fall användas i framtiden med "machine learning".)

Problemområden, utmaningar

(Skelett av frågor om de problemområden som ni arbetar med eller har på agendan)

Problemområde A (B, C, ...) - kostnad:

- Hur vanligt, ofta, ggr/år
- Kostnad per gång (eller per år)
- Var går gränsen
- Övriga svåråtkombara kostnader

Diskussion - om att leta upp och åtgärda problem av typ A med traditionell metod X:

- Enbart identifiering eller även lokalisering?
- Hur (inkl. bl.a. geografisk yta att leta på t.ex.)
- Kostnad i tid
- Kostnad systemresurser
- Tidsåtgång kalendertid
- Kostnad för åtgärd
- Hur stor andel A kan åtgärdas (per år)

Diskussion - leta upp och åtgärda problem av typ A med analysmetod baserat på elmätare med en utvecklad mätinfrastruktur:

- Vad och hur mycket ska kunna hittas?
- Grad av lokalisering (under nätstation, specifik kund, ...? Specifik kund med hur många % träffsäkerhet behövs?).
- Vilka typer av förbättringar behövs och vad kostar de:
 - f. Spänning
 - g. Effekt
 - h. Noggrannhet
 - i. tidsupplösning, tidssynkning
 - j. 15 min. avläsning
 - k. Nätstationsmätning (elkvalitetsinstrument eller elmätare)
- Särkostnad uppbyggnad av system
- Särkostnad drift
- Kostnad analys
- Kostnad för åtgärd av problem av typ A

Kompletteringsfrågor

- Nätstationsmätning: Vad skulle installation och drifttagning av mätning i en nätstation kosta er? Vi tänker oss en vanlig energimätare, inte ett elkvalitetsinstrument.
- Fler mätdata: Vi tänker oss utgångsläget att ni har eller ska ha timmätning av energi. Beskriv de kostnader och de eventuella hinder som skulle behöva övervinnas för att åstadkomma timmätning (och längre fram kvartsmätning) av tre fasspänningar, utöver energimätningen.
- Vad tror du om kostnadsförändringarna i er infrastruktur om ni skulle övergå från timmätning till kvartsmätning.
- Upplösning energi: Vilken upplösning på energimätningen har ni idagsläget? 1 kWh, 0.1 kWh, 0.01 kWh eller 0.001 kWh?
- Identifiering och lokalisering av icke-tekniska förluster / omätt förbrukning: Vi tänker oss ett sådant system byggt på utökad mätdata och avancerad, automatiserad analys. Hur bra skulle du vilja att det systemet skulle vara? Vad skulle vara "good enough" om vi tänker oss följande fyra nivåer:
 - l. Nätstationsnivå (låt oss säga att 7 kablar går ut från denna)
 - m. Kabelnivå (ett antal kopplingsskåp på varje kabel)
 - n. Kopplingsskåpsnivå (t.ex. 2 – 6 kunder)
 - o. Enskild kund-nivå

DATAANALYS OCH AVANCERADE ALGORITMER

Resultatet visar att lokaliseringsalgoritmer och detekteringslösningar för distributionsnät med nätstationsmätning. Algoritmerna kan minska resursåtgången lokalisera icke-tekniska förluster. Det finns incitament att sätta personal på att minska sina förluster och därmed kostnader mot ovanliggande nät, något som i fallet med omätt förbrukning även ger ökade intäkter men troligtvis en ökad medvetenhet om fel i databaser, rapporteringssystem och avvikande energimätare.

Grundläggande är att uppskatta sina förluster per timma istället för per år med stationsmätning och god mätkvalitet vilket visar timmar med för höga förluster. Lokaliserade icke-tekniska förluster, oavsett teknik, har resulterat i minskade ekonomiska förluster.

Enkät svar från elnätsbolag visar vilka framtida utmaningar som man ser är allt från ökade energimätarkrav, förstärkningar och att värdesäkra nätet. Enkät svaren visar att det finns stor spridning hur nätbolagen hanterar sina nätförluster, hur man i nuläget hanterar icke-tekniska förluster (ITF) och dess lokalisering och identifiering. Jämfört med kostnaden för nya generationens mätare kommer lagringskostnader för extra mätvärden och ökad mätupplösning inte vara den stora kostnadsdrivaren.

I denna rapport utvärderas korrelation, effektlödesanalys och maskininlärning för att detektera och lokalisera olika typer av ITF. Med hjälp av maskininlärning kan förluster lokaliseras utan perfekt kunskap om elnätets uppbyggnad så länge dessa förluster följer förbrukningsmönstret och perioder utan ITF finns för upplärning av algoritmen.

Ett nytt steg i energiforskningen

Energiforsk är en forsknings- och kunskapsorganisation som samlar stora delar av svensk forskning och utveckling om energi. Målet är att öka effektivitet och nyttiggörande av resultat inför framtida utmaningar inom energiområdet. Vi verkar inom ett antal forskningsområden, och tar fram kunskap om resurseffektiv energi i ett helhetsperspektiv – från källan, via omvandling och överföring till användning av energin. www.energiforsk.se