



Norges miljø- og
biovitenskapelige
universitet

Masteroppgave 2022 30 stp
Fakultet for realfag og teknologi

Prediksjon av elektrisitetsbruk i næringsbygg: Casestudie av Oslo Lufthavn Gardermoen

Prediction of electricity usage in commercial
buildings: Case study of Oslo Airport, Gardermoen

Kjersti Rustad Kvisberg
Miljøfysikk og fornybar energi

Sammendrag

For å tilpasse kraftsystemet til fremtidens behov, med økt elektrisitetsbruk og endret energimiks, kan forbrukerne i nettet og effektivisering av el-bruk spille en viktig rolle. Tiltak som energieffektivisering i bygg og lastrespons, som kan gi økt fleksibilitet i kraftnettet, krever prognoser for energi- og elektrisitetsbruk. Oslo lufthavn Gardermoen (OSL) har høy elektrisitetsbruk og store bygningsmasser, sammensatte laster og mange parkerte elektriske kjøretøy hvor det kan være fleksibilitet å hente. Denne oppgaven kartlegger derfor hvordan maskinlæringsalgoritmer kan brukes til å lage prognoser for elektrisitetsbruken til Oslo lufthavn én og 24 timer frem i tid.

For å predikere timeeffekt [kWh/h] for OSL i sommerhalvåret 2019 ble maskinlæringsalgoritmen *Long Short-Term Memory* (LSTM) brukt, en undertype av rekursive nevrale nettverk (RNN) som er godt egnet til prediksjon av tidsserier. Prediksjonene ble basert på meteorologiske parametere, kalenderinformasjon og passasjerantall.

38 modellkonfigurasjoner med ulike forklaringsvariabler og ulik nettverksarkitektur ble testet til prediksjon av el-bruk én time frem i tid. Modellene hadde ulike grader av treffsikkerhet, med *Mean Absolute Percentage Error* (MAPE) på 3,7-8,9 %. De fanget hovedmønstrene i el-bruken, men traff i mindre grad toppene. Med én time prediksjonshorisont fikk en enkel basismodell med ett LSTM-lag og kalenderinformasjon, meteorologiske parametere og passasjerantall som forklaringsvariabler MAPE på 6 %. Det ble vist at enkle nettverk med få forklaringsvariabler og mer komplekse nettverk med flere forklaringsvariabler ga best prediksjoner én time frem i tid. Den beste modellen hadde samme arkitektur som basismodellen, men kun time på døgnet og tidligere el-bruk som forklaringsvariabler. Modellen hadde MAPE på 3,7 %. Resultatene viser at tidligere elektrisitetsbruk var den viktigste forklaringsvariabelen.

For å kartlegge modellenes treffsikkerhet ved lengre prediksjonshorisonter, ble fire modellkonfigurasjoner testet til prediksjon et døgn frem i tid. De hadde MAPE på 4,7-5,5 %. I likhet med prediksjonene én time frem i tid, predikerte modellene mønstre bedre enn topper, men de hadde mer tydelig forverring utover i testsettet. En mer kompleks modell med tre LSTM-lag gjorde det best 24 timer frem i tid, med MAPE på 4,7 %. Modellene predikerte med tilsvarende treffsikkerhet som ved kortere prediksjonshorisont, noe som viser potensiale for ytterligere utvidelser i tid.

Undersøkelser av forklaringsvariabelenes innvirkning på prediksjonene viser at lufttemperatur bidrar til å bedre modellene, mens vindhastighet og nedbør er mindre viktige. Antallet terminalpassasjerer på OSL per uke bidro noe til prediksjonene, men det vil trolig være mer hensiktsmessig å bruke et aktivitetsmål med samme oppløsning som elektrisitetsmålingene. Riktig kombinasjon av variabelvalg og modelloppbygging ble vist å være viktig for prediksjonsresultatene.

Oppgaven viser at modeller basert på LSTM-algoritmen kan gi lastprognoser for Oslo lufthavn. For å bekrefte resultatene bør modellene testes på andre tidsrom.

Abstract

In order to adapt the power systems to the needs of the future, with increased electricity usage and changes in the energy mix, consumers in the grid and more efficient energy use can play vital roles. Improving the energy efficiency of buildings and implementing demand response to increase flexibility in the power grid require detailed forecasts of energy and electricity usage. Oslo Airport (OSL) has many buildings and a high electricity usage, but also varied loads and many parked EV's that could offer flexibility to the grid. This thesis therefore explores how machine learning (ML) algorithms can be used to forecast the electricity usage at OSL with forecast horizons of one and 24 hours.

To predict the target hourly electricity usage [kWh/h] for the airport in the summer half of 2019, the ML algorithm Recursive Neural Networks (RNN) of the subtype Long Short-Term Memory (LSTM), which is well suited for work on time series, was used. The features used in the predictions were measurements of meteorological variables, calendar variables and the number of passengers in the airport terminals.

38 model configurations using different combinations of features and network architectures were tested on forecasting electricity usage one hour ahead. The models gave varying results, with MAPE ranging from 3,7 % to 8,9 %. Most models learned the main patterns in the electricity usage, but were less well-aimed at peak prediction. Simple networks with few features or more complex networks with more features were most precise. A baseline model with one LSTM layer using time variables, meteorological variables and number of passengers as features obtained a 6 % MAPE. The best performing model had the same architecture as the baseline model, but only hour of the day and previous electricity usage as features. This model obtained a 3,7 % MAPE. The results showed the importance of including previous electricity usage as a feature.

To explore the models potential on longer prediction horizons, four model configurations were used to forecast 24 hours ahead. These models obtained MAPE scores of 4,7-5,5 %. Similarly to the predictions one hour ahead, the models were better at predicting general patterns than peaks, but the visual accuracy showed stronger deterioration over time. A more complex model with three layers and 64 units in each layer had the lowest MAPE of 4,7 %. The models were just marginally worse with a longer prediction horizon, which shows potential for forecasting further ahead.

Assessments of the impact the features had on results show that measurements of air temperature helped improve the scores, while wind speed and precipitation had less of an impact. The number of terminal passengers at OSL per week did contribute, but it would probably be more suitable to have an indicator of activity level with the same resolution in time as the target. Finding the right combination between features and network architecture was shown to be important for the final results.

This thesis shows that models based on the LSTM algorithm are suitable for forecasting electricity usage at Oslo airport. To confirm these results, the models should be tested on other time periods.

Forord

Denne oppgaven markerer slutten på min tid som student på Miljøfysikk og fornybar energi ved Norges miljø- og biovitenskapelige universitet. Da jeg begynte her i 2017 kunne jeg ikke gjettet hva disse årene skulle innebære av gleder og utfordringer eller at masteroppgaven min skulle handle om anvendelse av maskinlæring i kraftnettet, men det har vært elleve gode og lærerike semestre.

Først og fremst, tusen takk til min veileder Heidi S. Nygård for tett oppfølging og tilbakemeldinger gjennom hele prosessen, og for muligheten til å ombestemme seg og forme oppgaven som man vil. Takk også til Thomas Martinsen og resten av NeX2G-prosjektgruppa for de ukentlige møtene med nyttige innspill underveis. Takk til Åshild Grøtan, og til Ida og Rebekka, for hjelp med gjennomlesing av oppgaven.

Tusen takk til Avinor AS for tilgangen til dataene som ble brukt i oppgaven, og til Wiggo Roar Dehli for svar på spørsmål.

Jeg vil også gjerne takke venner og medstudenter som har bidratt til å gjøre disse årene så fine og interessante. Jeg har lært mye av og med dere, og gleder meg til å se hvor våre veier møtes igjen. En ekstra takk til dere jeg har bodd under samme tak som, gjengen og dere jeg har delt lesesal med denne siste høsten. Det har vært en glede å få dele hverdagen og studietiden med dere.

*A mind that is stretched by a new experience
can never go back to its old dimensions*
- Oliver Wendell Holmes, Sr.

Kjersti Rustad Kvisberg
Ås, 15. desember 2022

Innhold

1	Introduksjon	1
1.1	Bakgrunn	1
1.2	Motivasjon	2
1.3	Problemstilling og avgrensing	2
2	Teori	4
2.1	Kraftsystemet i Norge	4
2.1.1	Hovedkomponenter og overføringsnivåer	4
2.1.2	Forsyningssikkerhet	5
2.1.3	Balanse ved mer fornybar energi i nettet	6
2.1.4	Fleksibilitet	6
2.1.5	Forbrukerfleksibilitet	7
2.1.6	Prissetting av elektrisitetsbruk	8
2.1.7	Behov for lastprognoser	9
2.2	Maskinlæring	9
2.2.1	Undergrupper maskinlæring	10
2.2.2	Nevrale nettverk	10
2.2.3	Trening av nevralt nettverk	12
2.2.4	Tidsserieprediksjon	13
2.2.5	Rekursive nevralt nettverk	14
2.2.6	Long Short-Term Memory	15
2.2.7	Læringsprosessen	17
2.2.8	Evalueringsmål	18
2.3	Maskinlæring i kraftsystemet	19
3	Metode	21
3.1	Databehandling	21
3.2	Modelloppbygging og generelle valg	22
3.3	Basismodell	24
3.4	Modellkonfigurasjoner for prediksjon én time frem i tid	25
3.5	Prediksjon et døgn frem i tid	26
4	Case: Oslo lufthavn Gardermoen	28
4.1	Oslo lufthavn Gardermoen	28
4.2	Målinger av elektrisitetsbruk fra Oslo lufthavn	29
4.2.1	Mønstre i el-bruken i årene 2017 til 2021	30
4.2.2	Variasjon i el-bruk gjennom døgnet	31

4.2.3	El-bruk uke 18, 27 og 39 i 2019	33
4.3	Forklaringsvariabler og øvrige data	34
4.3.1	Forklaringsvariabler og lignende prediksjoner av el-bruk	35
4.3.2	Terminalpassasjerer	37
4.3.3	Meteorologiske målinger	39
4.3.4	El-bruk og meteorologiske målinger for uke 18 i 2019	41
4.4	Tidsperiode for prediksjon: sommerhalvåret 2019	42
5	Resultat og diskusjon	44
5.1	Mønstre i el-bruken og korrelasjonen med forklaringsvariablene	44
5.2	Evaluering og sammenligning av modeller	47
5.2.1	Scorer og evalueringsmål	47
5.2.2	Sammenligning av basismodellene fra RNN1-RNN3	49
5.3	Prediksjoner én time frem i tid	50
5.3.1	Enkel modell	50
5.3.2	Basismodell	50
5.3.3	Topp elleve beste konfigurasjoner	55
5.3.4	Innvirkning av evalueringsmål på rangering	59
5.3.5	Utvalgte prediksjonsplott	60
5.4	Prediksjoner et døgn frem i tid	65
5.5	Generell diskusjon	70
5.5.1	Overføringsverdi fra historiske til fremtidige data	70
5.5.2	Datarensing	71
5.5.3	Variabelvalg	71
5.5.4	Algoritmevalg	74
6	Konklusjon og videre arbeid	76
6.1	Konklusjon	76
6.2	Videre arbeid	77
	Referanser	79
A	Tilleggsinformasjon til kapittel 4	86
A.1	Helligdager i årene 2017 til 2019	86
A.2	El-bruk og meteorologiske målinger for uke 27 og 39 i 2019	87
B	Tilleggstabeller til kapittel 5	89
C	Tilleggsfigurer til kapittel 5	93
C.1	Oppbygging basismodell	93
C.2	Prediksjonsplott	94
C.2.1	Basismodell	94
C.2.2	Prediksjon én time frem i tid	95
C.2.3	Prediksjon et døgn frem i tid	114

Tabeller

3.1	Alternativer for egenskapene i modellkonfigurasjonene som ble testet.	25
3.2	Kombinasjoner av forklaringsvariabler som ble testet.	26
4.1	Forklaringsvariablene i datasettet.	35
5.1	RMSE-, MAE- og MAPE-scorene for den enkle gjennomsnittsmode- len på testsettet.	51
5.2	Gjennomsnittlige RMSE-, MAE- og MAPE-scorene for basismodellen for trenings- og testsettene.	51
5.3	De elleve beste modellkonfigurasjonene til prediksjon én time frem i tid.	56
5.4	Modellkonfigurasjoner testet til prediksjon et døgn frem i tid.	66
A.1	Datoene for helligdagene i helligdagsvariabelen i datasettet.	86

Figurer

2.1	Hovedkomponentene i et kraftsystem og stegene i overføring av elektrisk energi.	5
2.2	Effekttoppredusjon med lastflytting og lastklipping.	8
2.3	Oppbyggingen av Adaline-algoritmen.	11
2.4	Oppbygging av et nevralt nettverk med input-, skjult- og output-lag.	12
2.5	Sammenligning av et forovermatet og et rekursivt nevralt nettverk.	15
2.6	Oppbygging av en <i>Long Short-Term Memory</i> -celle.	16
2.7	Inndeling i trenings-, validerings- og testsett.	17
3.1	Oppbygging av basismodellen	24
4.1	Kart med plasseringen til Oslo lufthavn Gardermoen.	29
4.2	Gjennomsnittlig elektrisitetsbruk ved OSL per døgn i årene 2017 til 2021.	30
4.3	Gjennomsnittlig timebruk av elektrisitet ved OSL gjennom døgnet for ukedager og helger, vanlige dager og helligdager og de fire sesongene i årene 2017 til 2019.	32
4.4	Elektrisitetsbruken per time ved OSL tre uker i 2019.	34
4.5	Antallet terminalpassasjerer på OSL per uke i 2017 til 2021.	38
4.6	Gjennomsnittlig lufttemperatur, vindhastighet og nedbør per uke i 2017 til 2019 fra Gardermoen målestasjon.	40
4.7	Elektrisitetsbruken ved OSL i uke 18 i 2019 med lufttemperatur, vindhastighet og nedbør.	42
5.1	Pearson-korrelasjonen mellom el-bruk og forklaringsvariabler	46
5.2	Sammenligning basismodellene fra RNN1, RNN2 og RNN3.	49
5.3	Prediksjonsplott for basismodellen én time frem i tid	53
5.4	Prediksjonsplott for modell I én time frem i tid.	61
5.5	Prediksjonsplott for modell III én time frem i tid.	62
5.6	Prediksjonsplott for modell XI én time frem i tid.	64
5.7	Prediksjonsplott for basismodellen et døgn frem i tid	67
5.8	Prediksjonsplott for modell III et døgn frem i tid.	68
5.9	Prediksjonsplott for modell VI et døgn frem i tid.	69
A.1	Elektrisitetsbruken i uke 27 i 2019 med lufttemperatur, vindhastighet og nedbør.	87
A.2	Elektrisitetsbruken i uke 39 i 2019 med lufttemperatur, vindhastighet og nedbør.	88

C.1	Lagene i basismodellen med parameterantall.	93
C.2	Prediksjonsplott for basismodellen fra RNN1 én time frem i tid.	94
C.3	Prediksjonsplott for basismodellen fra RNN2 én time frem i tid.	94
C.4	Prediksjonsplott for basismodellen fra RNN3 én time frem i tid.	95
C.5	Prediksjonsplott for modell I én time frem i tid.	95
C.6	Prediksjonsplott for modell II én time frem i tid.	96
C.7	Prediksjonsplott for modell III én time frem i tid.	96
C.8	Prediksjonsplott for modell IV én time frem i tid.	97
C.9	Prediksjonsplott for modell V én time frem i tid.	97
C.10	Prediksjonsplott for modell VI én time frem i tid.	98
C.11	Prediksjonsplott for modell VII én time frem i tid.	98
C.12	Prediksjonsplott for modell VIII én time frem i tid.	99
C.13	Prediksjonsplott for modell IX én time frem i tid.	99
C.14	Prediksjonsplott for modell X én time frem i tid.	100
C.15	Prediksjonsplott for modell XI én time frem i tid.	100
C.16	Prediksjonsplott for modell XII én time frem i tid.	101
C.17	Prediksjonsplott for modell XIII én time frem i tid.	101
C.18	Prediksjonsplott for modell XIV én time frem i tid.	102
C.19	Prediksjonsplott for modell XV én time frem i tid.	102
C.20	Prediksjonsplott for modell XVI én time frem i tid.	103
C.21	Prediksjonsplott for modell XVII én time frem i tid.	103
C.22	Prediksjonsplott for modell XVIII én time frem i tid.	104
C.23	Prediksjonsplott for modell XIX én time frem i tid.	104
C.24	Prediksjonsplott for modell XX én time frem i tid.	105
C.25	Prediksjonsplott for modell XXI én time frem i tid.	105
C.26	Prediksjonsplott for modell XXII én time frem i tid.	106
C.27	Prediksjonsplott for modell XXIII én time frem i tid.	106
C.28	Prediksjonsplott for modell XXIV én time frem i tid.	107
C.29	Prediksjonsplott for modell XXVI én time frem i tid.	107
C.30	Prediksjonsplott for modell XXVII én time frem i tid.	108
C.31	Prediksjonsplott for modell XXVIII én time frem i tid.	108
C.32	Prediksjonsplott for modell XXIX én time frem i tid.	109
C.33	Prediksjonsplott for modell XXX én time frem i tid.	109
C.34	Prediksjonsplott for modell XXXI én time frem i tid.	110
C.35	Prediksjonsplott for modell XXXII én time frem i tid.	110
C.36	Prediksjonsplott for modell XXXIII én time frem i tid.	111
C.37	Prediksjonsplott for modell XXXIV én time frem i tid.	111
C.38	Prediksjonsplott for modell XXXV én time frem i tid.	112
C.39	Prediksjonsplott for modell XXXVI én time frem i tid.	112
C.40	Prediksjonsplott for modell XXXVII én time frem i tid.	113
C.41	Prediksjonsplott for modell XXXVIII én time frem i tid.	113
C.42	Prediksjonsplott for modell III 24 timer frem i tid.	114
C.43	Prediksjonsplott for modell VI 24 timer frem i tid.	114
C.44	Prediksjonsplott for modell XXV 24 timer frem i tid.	115
C.45	Prediksjonsplott for modell IV 24 timer frem i tid.	115

Ordliste

Den følgende listen beskriver ord og symboler som senere vil bli brukt i teksten.

kWh/h Timeeffekt, mål på elektrisk energi brukt per time

kWh Kilowatttime, mål på elektrisk energi

kW Kilowatt, mål på elektrisk effekt

LSTM Long Short-Term Memory, maskinlæringsalgoritme, undertype av rekursive nevrale nettverk

MAE Mean Absolute Error, evalueringsmål for regresjonsproblemer

MAPE Mean Absolute Percentage Error, evalueringsmål for regresjonsproblemer

MSE Mean Squared Error, evalueringsmål for regresjonsproblemer

RMSE Root Mean Squared Error, evalueringsmål for regresjonsproblemer

RNN Rekursivt nevral nettverk, en maskinlæringsalgoritme

Kapittel 1

Introduksjon

Dette kapittelet presenterer bakgrunnen for og motivasjonen bak denne oppgaven. Oppgavens problemstilling og avgrensninger blir også presentert.

1.1 Bakgrunn

Jorda og dens beboere står overfor seriøse, menneskeskapte endringer i planetens klima og miljø. For å begrense omfanget av endringene må utslippene av klimagasser reduseres, blant annet ved å bruke andre energikilder og utnytte energien som produseres mer effektivt [1]. To av energikildene som må øke i omfang er vindkraft og solkraft. Dette er fornybare energikilder som ikke er like enkle å regulere som vannkraft og ikke-fornybare energikilder. Produksjonen og bruken av elektrisitet må til enhver tid være lik; det må være balanse i nettet. Med større andel uregulerbare energikilder må man finne alternative måter å opprettholde balansen. Her kan forbrukerne i kraftnettet spille en viktig rolle.

Globalt står bygninger for 40 % av energibruken og 30 % av CO₂-utslippene, mye grunnet en høy andel fossile energikilder. Å forbedre energieffektiviteten i bygninger har derfor globalt blitt en sentral utfordring for å redusere mengden utslipp og å minke bruken av fossile brensel [2]. I Norge er andelen energi fra fossile energikilder lavere, men like fullt står bygninger for 40 % av energibruken i landet [3]. Den største andelen av denne energien brukes til oppvarming. I Norge estimerer Norges vassdrags- og energidirektorat at energieffektivisering i bygninger kan redusere energibruken med like mye som 10 prosent av strømbruken i landet [4]. For å redusere denne energibruken og dermed utslippene trengs en mer effektiv håndtering av energi i eksisterende bygninger.

Samtidig trenger nettet økt fleksibilitet for å opprettholde balansen. Med riktig styring av elektrisitetsbruken kan bygninger stille til rådighet lastflytting, effekttoppreduksjon, frekvensregulering og andre lastrespons-tjenester [5]. Potensialet til lastrespons fra forbrukerne i det norske kraftnettet har blitt estimert til mellom 8,5 og 24,9 % av de totale effekttoppene. Det ble estimert at husholdninger kan stå for 11,4 % av reduksjonene i effekttopper, og næringsbygg inntil 4 % [6]. Å yte denne

typen tjenester kan også ha økonomiske fordeler for forbrukerne.

Nye styringssystemer for energibruk krever bedre forståelse og prognoser av energi- og elektrisitetsbruken i bygninger. Dette kan prediksjonsmodeller gi [2]. Treffsikre prediksjoner av el-bruk i bygninger gjør at man kan planlegge og dermed redusere el-bruken. Det gjør at tilgang på elektrisitet og effekt kan planlegges og prioriteres, og dermed forebygge kraftmangel i byer og redusere sosiale kostnader ved unødvendig energiproduksjon og utbygginger av nettet [7].

Mange eksisterende løsninger for styring av energibruk i bygninger bygger på modelleringer av de fysiske energisystemene i bygningene. Maskinlæringsmodeller krever i motsetning ingen kunnskap om de fysiske systemene for å gi prognoser av fremtidig energibehov [5].

Prediksjon av laster i nettet er også sentralt for effektiv operasjon av kraftnett, og de første maskinlæringsalgoritmene ble testet til å lage lastprognoser for kraftsystemer allerede på 1980-tallet. Siden har kunstig intelligens og maskinlæring blitt testet til blant annet last- og prisprediksjon, kraftsystem-operasjoner og energistyring i bygninger [5].

1.2 Motivasjon

Som en del av forskningsprosjektet Nettbalansering fra store parkeringsanlegg og næringsbygg - NeX2G [8], undersøker denne oppgaven prediksjon av elektrisitetsbruken til Oslo lufthavn Gardermoen (OSL). Forskningsprosjektet omhandler fleksibilitetspotensialet fra parkerte elektriske kjøretøy og bygningsmassene på flyplassen, nå og i fremtiden. For å kunne estimere fleksibilitet må man også vite mønstrene og top-pene som el-bruken har i utgangspunktet. Denne oppgaven er derfor en innledende studie i bruk av maskinlæring for å predikere elektrisitetsbruken til flyplassen.

1.3 Problemstilling og avgrensing

Målet til denne oppgaven er å kartlegge hvordan maskinlæringsalgoritmer fungerer til å predikere den aggregerte el-bruken til Oslo lufthavn. Verdiene som skal estimeres er timeverdier for brukt elektrisitet, omtalt som timeeffekt [kWh/h], for hele flyplassen.

En av de største elektriske lastene til flyplassen er varmesentralen, hvor varmepumper og en el-kjel brukes til oppvarming. Varmepumpene driftes basert på temperatur, med stabil høy bruk til oppvarming om vinteren og mer varierende bruk til kjøling om sommeren. Varmepumpene har dermed et kjent bruksmønster, men egner seg ikke til fleksibel bruk. El-kjelen styres manuelt basert på både oppvarmingsbehovet og på ledig nettkapasitet, siden den er en stor last. Den anses derfor ikke som predikerbar. Det ble derfor valgt å ekskludere elektrisitetsbruken til disse to komponentene fra prediksjonene.

For å oppnå målet starter oppgaven med å undersøke hvilke mønstre som finnes i el-bruken i årene 2017 til 2021, med fokus på sommerhalvåret i 2019. Videre undersøkes det hvilke forklaringsvariabler som kan brukes i prediksjonen, på grunnlag av literatursøk og datatilgjengelighet. En dyplæring-algoritme som predikerer el-bruken én time og et døgn frem i tid blir implementeres og testes på et datasett bestående av månedene april til september 2019. Siden oppgaven er en innledende studie blir det lagt vekt på å kartlegge påvirkningen av ulike endringer i modellene, og det undersøkes derfor hvordan endringer i nettverksarkitekturen og forklaringsvariablene påvirker prediksjonene fra modellene.

Kapittel 2

Teori

Denne oppgaven er en innledende studie i anvendelsen av maskinlæringsmodeller til prediksjon av elektrisitetsbruken ved Oslo lufthavn. Dette kapittelet presenterer derfor grunnleggende teori om både det norske elektriske kraftsystemet og maskinlæring som er relevant for de øvrige kapitlene i denne oppgaven.

2.1 Kraftsystemet i Norge

I dette delkapittelet beskrives oppbyggingen til det norske kraftnettet, etterfulgt av hensyn og utfordringer i fremtidens kraftnett.

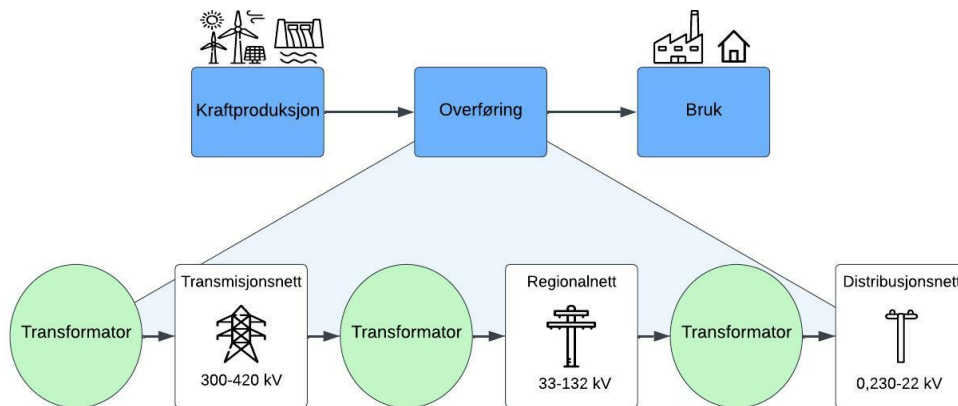
2.1.1 Hovedkomponenter og overføringsnivåer

Det norske kraftnettet har tre grunnleggende funksjoner: produksjon, overføring og bruk. I produksjonsleddet genereres elektrisiteten som så sendes ut til forbrukere gjennom overføringsleddet. Figur 2.1 viser hvordan hovedkomponentene i et kraftsystem henger sammen og hvordan den elektriske energien overføres. Et godt fungerende og sikkert kraftnett, og en pålitelig strømforsyning er essensielt for det moderne samfunnet [9].

Som figur 2.1 viser er overføringen i det norske kraftnettet delt inn i tre nivåer: transmisjonsnettet, regionalnettet og distribusjonsnettet. Transformatorstasjoner utgjør bindeleddene mellom nivåene i nettet. Transmisjonsnettet er det øverste nivået og hovedveiene i nettet. Der fraktes høyspent elektrisitet. Transmisjonsnettet kobler sammen produksjon og store forbrukere over lange avstander. Operatøren av dette nettet i Norge er Statnett SF [9].

Regionalnettet kobler transmisjonsnettet sammen med distribusjonsnettet. Regionalnettet overfører elektrisitet med mellomhøy spenning. Store forbrukere som industri og anlegg kan være koblet på dette nettnivået. Distribusjonsnettet er det laveste nivået og består av de lokale kraftnettene. Dette nettnivået har lavest spenningsnivå. Det er her de fleste forbrukere som husholdninger, småindustri og tjenester er

påkoblet. Regionalnettet og distribusjonsnettet eies og drives av lokale nettselskap som ofte er eid av kommuner og fylker [9].



Figur 2.1: Hovedkomponentene i et kraftsystem og stegene i overføring av elektrisk energi. Basert på informasjon fra Olje- og energidepartementet [9], fremstilling fra Lyslo [10, s. 5].

2.1.2 Forsyningsikkerhet

Som nevnt er en pålitelig strømforsyning svært viktig for samfunnet. Det må til enhver tid bli tatt like mye ut av strømmettet som det sendes inn; strømmettet må være i momentan balanse [9].

Forsyningsikkerheten beskriver evnen kraftsystemet har til å levere strøm av gitt kvalitet til brukerne av nettet [11]. Forsyningsikkerheten må være høy for å sikre pålitelig strømforsyning. Strømmen som leveres skal ha en viss frekvens, 50Hz med normalvariasjon på $\pm 0,1\text{Hz}$, og et visst spenningsnivå [11]. Spenningsnivået varierer ut fra hvor i nettet kundene er koblet på. Dersom det er ubalanse mellom forbruk og produksjon vil nettfrekvensen endres.

For å opprettholde forsyningsikkerheten må overføringsleddet i kraftsystemet bygges for å håndtere variasjoner i både produksjon og forbruk for både lange og korte tidshorisonter [9]. Nettet må være i stand til å sende strøm fra områder i Norge med høy produksjon og lavt forbruk, til områder med mindre produksjon. Hvor produksjonen og forbruket er høyt og lavt til enhver tid varierer.

Forsyningsikkerheten i kraftsystemet omfatter energi-, drifts- og effektsikkerhet [11]. Energisikkerhet omfatter at kraftsystemet må kunne dekke energibehovet til enhver tid. Dette krever høy nok energiproduksjon. Nettet må også driftes slik at det motstår forstyrrelser, uten at det fører til avbrudd eller avvik fra ønsket frekvens og spenning på elektrisiteten til sluttbrukere. Dette kalles driftssikkerhet [11].

I tillegg til at det må være nok energi [kWh] tilgjengelig i nettet, må det også være nok effekt [kW]. Effektsikkerheten går derfor ut på at kraftsystemet må kunne takle

den øyeblikkelige belastningen, ved å ha nok kapasitet i både installert kraftproduksjon og overføringskapasitet i kraftnettet. Mens lav energisikkerhet kan forekomme over lengre tid, er knapphet på effekt typisk et problem i enkelttimer med høyt forbruk [11].

2.1.3 Balanse ved mer fornybar energi i nettet

Elektrisitetsproduksjonen i Norge har ulike energikilder. I et normalår produseres det omlag 155 TWh kraft, hvorav nesten 90 % er fra vannkraft. Det øvrige produseres med vindkraft og termisk kraft [12]. Moderne samfunn er svært avhengige av elektrisitet, men i motsetning til mange andre samfunnskritiske varer er elektrisitet ikke enkelt å lagre. Den må brukes samtidig som den produseres. Momentan balanse krever detaljert planlegging, men også muligheter til å regulere produksjon og bruk på kort sikt.

Dagens kraftproduksjonssituasjon er i endring og behovet for elektrisitet i Norge øker. Energisystemene i hele Europa skal gå mot nullutslipp i 2050. Det innebærer netto nullutslipp i all energibruk innenfor transport, industri, byggenæring, husholdninger og kraftsystemer [13]. For å få til det må produksjonen, distribusjonen og bruken av energi endres. Bruken av elektrisitet i Norge forventes å øke fra omlag 140 TWh i dag til 180-190 TWh i 2040-50. En stor del av denne økningen skyldes elektrifisering av transport og industri [14].

Vannkraft har lenge stått for den største andelen av norsk kraftproduksjon. Vannet som utnyttes i kraftproduksjonen lagres i magasiner. Det gjør vannkraften regulerbar og pålitelig [15]. For å dekke det økende kraftbehovet uten å øke CO₂-utslippene må det bygges mer av både vannkraft og andre fornybare energiteknologier som landvind, havvind og solkraft [14]. Vindkraft og solkraft er ikke regulerbare på samme måte som vannkraft, de produseres når været tillater det [16]. Dette gir svingninger i den løpende effektbalansen [14].

Sammen kan endringene i bruken og produksjonen av elektrisitet føre til høyere effekttopper, flaskehals, manglende momentan balanse og andre problemer i driften av nettet. Deler av strømmettet der overføringskapasiteten er begrenset kalles flaskehals. Regulering av vannkraftproduksjonen kan ikke alene veie opp for disse sammensatte utfordringene, og man trenger nye verktøy for å balansere kraftnettet. Et av disse verktøyene er fleksibilitet i forbruksdelen av nettet [17].

2.1.4 Fleksibilitet

Det internasjonale energibyrået (IEA) [16] definerer fleksibilitet i kraftnettet som evnen et kraftsystem har til å pålitelig og kostnadseffektivt håndtere variabiliteten og usikkerheten til tilbud og etterspørsel over alle relevante tidsskalaer. Det har alltid vært viktig for å kunne møte variabel etterspørsel og kunne respondere på tapene fra store generatorer eller transmisjonlinjer. Når andelen uregulerbar fornybar kraft øker samtidig som elektrisitetsbruken øker blir fleksibilitet relevant for å håndtere integrasjonen av den nye produksjonen i nettet [16].

Avvik fra den momentane balansen håndteres i reservemarkeder. I Norge er det systemansvarlig Statnett som anskaffer disse. Det er ulike typer reserver for forstyrrelser av ulik størrelse og tidsskala [18]. Når det skjer endringer i kraftsystemene, som beskrevet i delkapittel 2.1.3, oppstår et behov for mer kapasitet til å regulere strømmene i nettet. Det trengs flere kilder til fleksibilitet.

Fleksibilitet kan utnyttes i alle kraftsystemets hovedkomponenter [17], og for å opprettholde balanse i fremtidens kraftnett må flere verktøy utnyttes. Kraftprodusenter som raskt og effektivt kan øke eller minke sin produksjon etter behov kan utgjøre fleksibel produksjon. På samme måte kan overføringsystemer med få flaskehals og tilstrekkelig kapasitet til å implementere balanserende ressurser, smartere styring og optimalisering av overføring utgjøre fleksibel overføring. I tillegg kommer fleksible systemoperasjoner; prosedyrer som gjør at man kan utnytte fleksibilitet i det eksisterende systemet, for eksempel å ta beslutninger oftere eller nærmere nåtid, forbedret bruk av produksjonsprognoser for sol- og vindkraft og mer samarbeid [17].

Den siste hovedkomponenten i kraftnettet er forbrukerne. I neste delkapittel beskrives måter disse kan bidra til fleksibilitet i nettet.

2.1.5 Forbrukerfleksibilitet

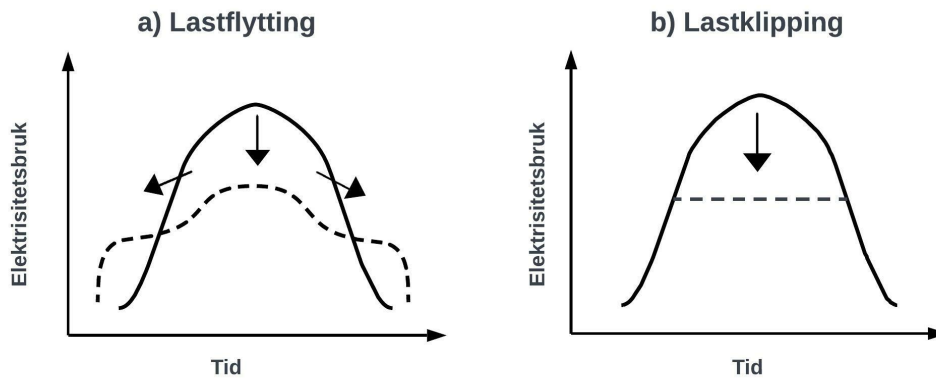
Som nevnt i delkapittel 2.1.3 kan endringene i kraftsystemet føre til effekttopper i el-bruken. Effekttopper oppstår i timer hvor det er høy sammenfallende bruk og behovet for elektrisk effekt derfor er høyere enn ellers. Disse timene er effektsikkerheten i nettet redusert, som beskrevet i delkapittel 2.1.2, ettersom det ikke er ytterligere kapasitet tilgjengelig til å dekke etterspørsel. Nettet må være dimensjonert for å håndtere disse timene med høyest momentan belastning [11]. Dette innebærer for eksempel at nettet må være laget for å dekke kraftbehovet en iskald vinterdag hvor mange forbrukere bruker mye strøm til oppvarming på samme tid.

Effekttoppene er således bestemmende for hvor mye overføringskapasitet man må ha i alle de tre nivåene av nettet som man så i figur 2.1. Den største andelen av effekttoppene i det nasjonale forbruket skyldes alminnelig forsyning, som i stor grad er energiforsyning til bygninger [19]. Effekttoppene i el-bruken utgjør typisk en liten andel av timene i året. Det er derfor ønskelig å redusere effektuttaket i de høyeste timene for å kunne sikre en bedre utnyttelse av nettet. Fra et systemperspektiv kan reduksjon av effekttopper begrense eller utsette investeringer i nettet uten å gå utover forsyningssikkerheten [19].

En last som kan endres opp og ned uten at det går ut over kvaliteten til tjenesten eller bruken av lasten kalles en fleksibel last [20]. Når en elektrisk kraftkunde endrer sin strømbruk for å bidra til å opprettholde momentan balanse i nettet kalles dette lastrespons (*demand response*), en form for forbrukerfleksibilitet som hovedsakelig brukes for å redusere effekttopper [21]. På forbrukersiden har fleksibiliteten tradisjonelt kommet fra lastrespons fra store bedrifter og laster i industriell sektor og næringsliv [22].

Tiltak for å redusere effekttopper gjennom lastrespons inkluderer lastflytting og last-

klipping (*peak shaving* eller *load clipping*) [23], som er illustrert i henholdsvis figur 2.2 a) og b). Lastflytting vil som figur 2.2 a) viser si å redusere effektbruken ved å flytte deler av den i tid til timer som ikke har like høy sammenfallende el-bruk. Lastklipping er å kutte effektbruken ved å sette et tak, som vist i figur 2.2 b). Med denne teknikken blir laster kuttet direkte, ikke erstattet eller flyttet i tid. Laster som kan brukes til lastrespons er for eksempel ventilasjonsanlegg, varmtvannsberedere og oppvarmings- og kjølesystemer, både i næringsbygg og husholdninger [20]. Planlegging av lading eller toveislading av elektriske kjøretøy på strategiske tidspunkter kan også redusere topper i effektbruken [21].



Figur 2.2: Metoder for å redusere effekttopper i elektrisitetsbruk inkluderer a) lastflytting og b) lastklipping. Basert på Hashemi [23, s. 2].

Forbrukerfleksibilitet inkluderer som nevnt implementering av lastrespons, men også utjevning av el-bruk gjennom energilagring, justering av distribuert generasjon etter behov og andre måter å la forbrukere respondere på markeds signaler. Det inkluderer også direkte lastkontroll [17].

For kunder er det to typer verdi i å ha fleksibel el-bruk. Eksplisitt fleksibilitet er å selge sin evne til å justere lastene på forespørsel fra nettselskap eller andre aktører. Implisitt fleksibilitet innebærer å selv justere sitt eget forbruk basert på prissignaler og tariffer for å få lavere strømutfgifter [24] [25]. Hvilke utgifter strømkunder har blir beskrevet i neste delkapittel.

2.1.6 Prissetting av elektrisitetsbruk

Den totale strømkostnaden til forbrukere i strømmettet avhenger av både mengden energi og mengden effekt kunden etterspør. Den består av flere deler: kraftpris, nettleie og avgifter til staten. Kraftprisen er en variabel kostand per enhet elektrisitet [kWh] som er brukt. Nettleien består av et fastledd og et variabelt ledd. Fastleddet differensieres etter hvor mye effekt [kW] kunden bruker, og skal dekke nettselskapets kostnader for å drifte strømmettet. Den variable delen kalles energiledet, og skal dekke de marginale tapskostnadene fra overføringen av strøm til kunden. Energiledet er høyere jo mer elektrisitet kunden bruker [26].

På tidspunkter hvor det er høy bruk av elektrisitet i nettet må produksjonen økes for å dekke behovet. I tillegg må det være høy nok overføringskapasitet. Dersom det er lite ytterligere kapasitet til produksjon tilgjengelig i nettet når behovet er høyt vil dette reflekteres til kundene i høyere strømpriser disse timene. Som nevnt i delkapittel 2.1.5 kan reduksjon av sammenfallende effektbehov gjøre at investeringer i nettet kan utsettes. Dersom det er lite ytterligere overføringskapasitet tilgjengelig reflekteres derfor knappheten til kunden gjennom effektledet av strømkostnaden, for å gi insentiv til å senke eget forbruk i disse timene [19]. Dermed kan kunden senke strømkostnaden sin ved å redusere en eller begge mengdene, og samtidig bidra med implisitt fleksibilitet til nettet.

2.1.7 Behov for lastprognoser

Fremtidens digitale nettsystemer omtales gjerne som *smartgrids*; smarte digitaliserte kraftnettsystemer som drives mer effektivt og dermed reduserer energi- og effektbehovet. Det innebærer blant annet styring av energilagringssystemer, styring av energibehov hos forbrukere og effekttopp-styring [7]. Flere av disse ble nevnt som former for forbrukerfleksibilitet i delkapittel 2.1.5.

Mange av denne typen styringssystemer krever estimater på fremtidig el-bruk og effekttopper til planleggingsformål. Planlegging av tiltak for å redusere effekttopper, for eksempel ved opp- og utlading av batterier, krever et tydelig bilde av når effekttoppene oppstår og hvor høye de er [27]. For å kunne justere sin egen energibruk og effektbruk, og dermed få lavere strømutgifter, må forbrukere vite på forhånd når de kan forvente å bruke mye strøm. Større energisystemer trenger også presise lastprognoser til planleggingsformål i både nettdrift og kraftproduksjon.

Effektive teknikker for lastprognoser kan dermed spille en sentral rolle på veien mot smarte nett og bygninger [28]. Til å generere lastprognoser kan det brukes flere verktøy, blant annet maskinlæringsalgoritmer. Neste delkapittel handler om maskinlæring, med fokus på undertypen rekursive nevrale nettverk. Deretter, i delkapittel 2.3 skildres anvendelser av maskinlæring i kraftnettet nærmere.

2.2 Maskinlæring

Maskinlæring er et underfelt av kunstig intelligens (AI) og går ut på at selvlærende algoritmer kan trekke sammenhenger fra data for å kunne gjøre prediksjoner, for eksempel si noe om hvordan noe vil se ut eller oppføre seg i fremtiden. En modell sitt estimat av en verdi eller klasseinndeling kalles en prediksjon. Maskinlæringsalgoritmer kan analysere store mengder data, trekke ut sammenhenger og bygge modeller mer effektivt enn om mennesker skulle gjort det manuelt [29]. Maskinlæring er et aktivt forskningsfelt, men implementering og utnytting av denne typen algoritmer har de siste årene økt i omfang i industri og næringsliv. Søkemotorer, spamfiltre og oversettelsesprogrammer bygger alle på maskinlæring. Innenfor kraftsystemer har kunstig intelligens vært brukt til planleggings- og prognoseformål, for eksempel lastprognoser, anomalideteksjon og sikkerhetsvurderinger [5].

I det neste kapittelet presenteres teori om maskinl ring. F rst presenteres undergrupper av maskinl ring. Deretter presenteres kunstige nevr le nettverk, med fokus p  rekursive nevr le nettverk. trening av disse og undertypen *Long Short-Term Memory* (LSTM) som blir anvendt i denne oppgaven. Teorien er basert p  boken *Python Machine Learning* av Raschka og Mirjalili [29], med mindre annet er spesifisert. Til sist beskrives vanlige evalueringsm l for maskinl ringsmetoder.

Oppbyggingen og treningen av maskinl ringsalgoritmer inneb rer en rekke matematiske operasjoner og ligninger. I denne oppgaven forklares teorien mer overordnet, og leseren henvendes til nevnt faglitteratur [29] for mer utdypende informasjon.

Dataene som hentes inn i en modell kalles *input*, og verdiene man f r ut kalles *output*.

2.2.1 Undergrupper maskinl ring

Maskinl ringsalgoritmer kan deles inn i tre undergrupper:

- Veiledet l ring (*supervised learning*)
- Ikke-veiledet l ring (*unsupervised learning*)
- Forsterkende l ring (*reinforcement learning*)

Veiledet l ring brukes p  merket data, som er data hvor man vet den sanne verdien. Ved at modellene i veiledet l ring kan sammenligne sine prognoser med verdiene de pr ver   predikere, kan modellene f  direkte tilbakemelding p  om de estimerer riktig. Denne typen algoritmer brukes til prediksjon av utfall eller fremtidige verdier.

Veiledet l ring kan anvendes p  to typer oppgaver. Den ene oppgaven er klassifisering, hvor man pr ver   sortere nye objekter inn i ulike grupper eller klasser basert p  tidligere objekter. Den andre oppgaven er regresjonsanalyse. I regresjon pr ver man   predikere en kontinuerlig responsvariabel (*target*) basert p  kjente forklaringsvariabler (*features*) ved   finne en sammenheng mellom de to typene variabler. Man pr ver   minske forskjellen mellom predikerte og faktiske verdier. Den gjennomsnittlige forskjellen brukes til   vurdere hvor god en prediksjon er, og dermed om prediksjonen b r justeres.

I ikke-veiledet l ring vet man ikke den sanne verdien og kan derfor ikke f  tilbakemelding underveis. Denne typen brukes til   finne skjulte strukturer i data. Forsterkende l ring brukes til   predikere valgprosesser eller serier av utfall eller handlinger, og l ringen skjer gjennom et bel nningssystem.

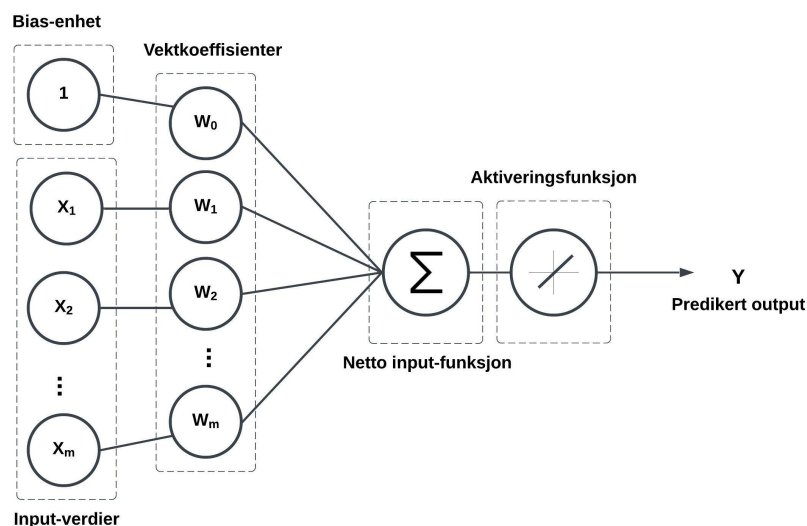
2.2.2 Nevrale nettverk

Et mye omtalt felt innenfor maskinl ring er dypl ring eller nevr le nettverk. Dypl ring kan forstås som et sett med algoritmer som ble utviklet for   trene kunstige nevr le nettverk med mange lag mest mulig effektivt [29]. Det er gjort en rekke gjennombrudd innenfor feltet de siste ti rene. I dag anvendes komplekse nevr le nettverk i en mengde digitale tjenester.

Kunstige nevrane nettverk simulerer den menneskelige hjernen ved å bruke noder, vekter og lag til å lagre informasjon i parallelle stier og hente ut informasjonen i parallell når det trengs [30]. Byggesteinene i disse nettverkene er kunstige nevroner. Nettverkene kan ha ett eller flere lag.

En representasjon av et nettverk med et enkelt lag, *ADaptive LInear NEuron* (Adaline) [31], er vist i figur 2.3. Til venstre i figuren representeres input-verdiene av x_1, x_2, \dots, x_m som sammen utgjør input-laget. Hver av disse input-verdiene, samt en *bias*-enhet 1, kobles med vekt-koeffisientene w_0, w_1, \dots, w_m og summeres. Til sammen utgjør dette netto input-funksjonen Σ . *Bias*-enheten muliggjør et skift i aktiveringsfunksjonen, analogt med å skifte krysningspunktet mellom en regresjonslinje $y = ax + b$ og y-aksen vekk fra (0,0) ved å endre b . Netto input-verdien sendes deretter til aktiveringsfunksjonen, som bestemmer hvordan den predikerte output-en skal beregnes. I Adaline i figur 2.3 er aktiveringsfunksjonen lineær, og endrer ikke verdiene fra netto input-funksjonen. Output-verdiene sendes ut til høyre i figuren.

Avviket mellom verdiene som sendes ut fra aktiveringsfunksjonen og de faktiske verdiene beregnes med en kostnadsfunksjon. Målet med treningen er å minimere denne kostnaden, noe som beskrives nærmere i delkapittel 2.2.3 om trening av nevrane nettverk. Vekt-koeffisientene w_0, w_1, \dots, w_m oppdateres basert på kostnadsfunksjonen. Algoritmen kan også brukes til klassifisering. Da settes en stegfunksjon inn mellom aktiveringsfunksjonen og output. Stegfunksjonen bestemmer hvilken klasse en output-verdi tilhører, for eksempel klasse 2 dersom verdien er over 0, og klasse 1 ellers.



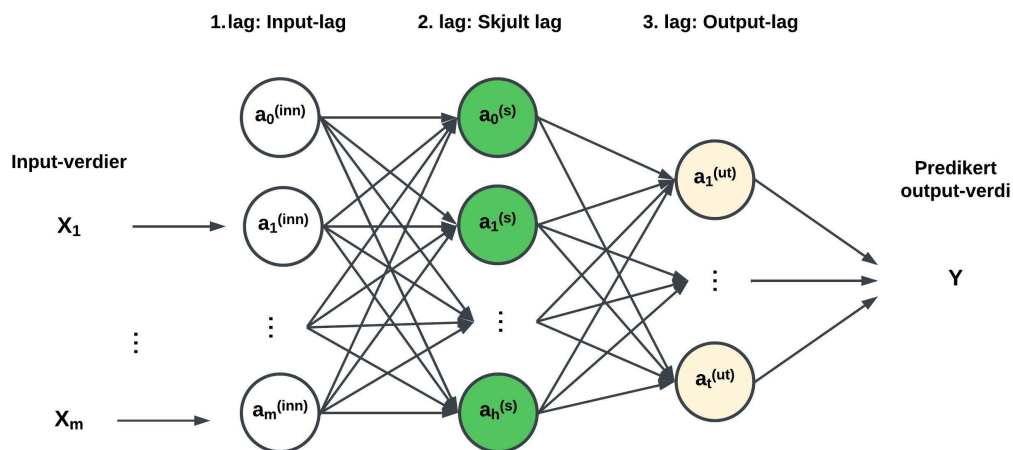
Figur 2.3: Representasjon av en Adaline-algoritme til bruk for regresjon. Nettverket tar inn m input-verdier som kobles med en vekt-koeffisient i netto input-funksjonen. Aktiveringsfunksjonen bestemmer hvordan output skal beregnes basert på netto input-verdien. Basert på Raschka og Mirjalili [29, s. 382].

Enkle lag slik som Adaline-algoritmen vist i figur 2.3 kan kobles sammen til et

større nettverk. Alle nevralt nettverk har et input-lag som tar inn input-data, og et output-lag som sender ut predikerte verdier Y . I tillegg kan det ha flere skjulte lag. Et nettverk som har flere skjulte lag kalles et dypt kunstig nevralt nettverk.

I figur 2.4 vises et forovermatet nevralt nettverk (FFNN) som har ett skjult lag, men antallet skjulte lag kan være stort. Hvert lag har et visst antall enheter eller noder som henter inn aktiveringer a fra tidligere lag og sender sine aktiveringer videre i modellen. Det er analogt med hvordan Adaline-algoritmen tok inn input og sendte ut output. Nettverket i figur 2.4 har h enheter. Hvert lag har også en *bias*-enhet a_0 . På figuren denoteres enhetene i input-laget med 1 til m , som vil si at nettverket tar inn m forklaringsvariabler. Output-laget har t enheter. Antallet output-noder avhenger av hva slags verdier man skal ha ut. For regresjonsproblemer brukes én node.

Hver enhet i et lag er koblet til hver enhet i neste lag med vektcoeffisienter. Enhetene i et lag er dermed en vektet sum av enhetene i foregående lag. Vektene oppdateres for å representere input-dataene og gi en best mulig prediksjon. Neste delkapittel forklarer hvordan vektoppdateringen foregår.



Figur 2.4: Et nevralt nettverk med et input-lag (inn) som tar inn m variabler, et skjult lag (s) med h enheter og et output-lag (ut) med t enheter. Basert på Raschka og Mirjalili [29, s. 384].

2.2.3 Trening av nevralt nettverk

Målet med opptreningen av et nettverk er å finne en best mulig representasjon av en responsvariabel, gjennom å finne sammenhenger mellom forklaringsvariabler. Avviket mellom responsvariabelens faktiske og estimerte verdier beregnes, som nevnt i delkapittel 2.2.2, med en kostnadsfunksjon.

I treningsprosessen passerer treningsdatasettet gjennom hele modellen flere ganger for å oppdatere vektene i modellen. En runde over hele treningssettet kalles en *epoch*, eller en epoke.

I hver enhet i et skjult lag beregnes en aktivering basert på netto input til enheten. Aktiveringen sendes videre til neste lag. Når informasjonen har passert gjennom hele nettverket til output-laget beregnes et avvik via en kostnadsfunksjon. Dette avviket sendes tilbake gjennom nettverket ved såkalt tilbakepropagasjon av feil. Feilen mellom output-verdiene fra det siste laget i nettverket og de målte verdiene man ønsker å predikere beregnes. Denne feilen brukes i beregningen av feilen for nest siste lag. Dermed er feilen i hvert lag avhengig av feilene fra alle foregående lag. Slik propageres avviket tilbake gjennom nettverket, og brukes til å oppdatere vektene for hvert lag. Denne prosessen gjentas for hver epoke. Til sist propageres informasjonen fra start til slutt og output beregnes.

Minimeringen av kostnadsfunksjonen og dermed oppdateringen av vektene gjøres stegvis, og hastigheten spesifiseres av læringsraten. Læringsraten sier noe om hvor mye vektene skal oppdateres for hvert steg, altså hvor store steg som skal tas på veien ned mot bunnen. For Adaline-algoritmen kan man se for seg at man leter etter bunnen på en konveks kurve. I nevralt nettverk har man et stort antall vektkoeffisienter, og kostnadsfunksjonen som skal minimeres er en multi-dimensjonell flate med mange lokale minima og maxima. Dersom stegene er for store, kan det skje at modellen passerer den optimale løsningen. Dersom de er for små, vil det ta lang tid før bunnen nås. Å velge en god læringsrate er dermed et av mange valg som må tas i byggingen av en modell.

I forkant av modellbyggingen må man velge størrelsen på nettverket, altså hvor mange lag den skal ha. Dersom nettverket er for enkelt vil det ikke klare å representere dataene godt; den vil bli undertilpasset. Dersom nettverket i motsatt fall er komplekst kan det klare å lære treningsdataene utenat. Da er sjansene store for at den fungerer dårlig på usett data, den er overtilpasset.

For å sikre at modellen finner en god balanse mellom over- og undertilpasning kan man bruke regularisering. En måte å regularisere nevralt nettverk er å bygge et relativt komplekst nettverk, og så forenkle det underveis. For nevralt nettverk kan man bruke en teknikk som kalles *dropout*, som er en form for forenkling av nettverket underveis. Ved *dropout* etter et skjult lag vil parameterne fra laget bli vilkårlig glemt med en spesifisert sannsynlighet. Siden dette er vilkårlig tvinger det modellen til å lære en overflødig representasjon av dataene, siden den ikke kan legge for mye vekt på noen av enhetene. Det gjør at modellen lærer mer generelle og robuste mønstre fra dataene.

Nettverket som ble illustrert i figur 2.4 er et forovermatet nevralt nettverk. Det finnes andre typer nevralt nettverk som er tilpasset ulike problemstillinger. Et eksempel er prediksjon på tidsserier, som målinger av elektrisitetsbruk ved Oslo lufthavn over tid er et eksempel på. I neste delkapittel presenteres hensyn som må tas ved arbeid med tidsserier, før en type nevralt nettverk som er tilpasset tidsserier presenteres.

2.2.4 Tidsserieprediksjon

Når datasett har en rekkefølge på objektene som ikke er vilkårlig kalles de sekvenser. Objektene i en sekvens er ikke uavhengige av hverandre. De fleste maskinlæringsalgo-

ritmer innenfor veiledet læring er basert på antagelser om at dataene er uavhengige og identisk distribuerte. Dette er ikke lenger gyldig når man jobber med sekvenser. Dersom rekkefølgen i dataene er avhengig av tid, dataene har en tidsdimensjon, kalles de tidsserier.

Hvis enten input- eller output-dataene til en maskinlæringsmodell har en tidsdimensjon trengs en modell som kan håndtere dette. I tilfellet med elektrisitetsdata fra Oslo lufthavn er både input- og output-dataene tidsserier. For å utføre prediksjoner på tidsserier er det en fordel å ha tilpassede modeller som kan knytte informasjon til rekkefølgen i tid. Tidsserier er ikke kontinuerlige, siden de er målinger av verdier ved bestemte tidspunkter [32], men de har en kronologi. Oppløsningen kan være høy, for eksempel ved målinger hvert sekund. Når man jobber med tidsserier er hovedregelen at testsettet bør ha nyere data enn treningssettet.

En type nevrale nettverk som er tilpasset tidsserier er rekursive nevrale nettverk (RNN). I neste delkapitel beskrives rekursive nevrale nettverk og en utfordring med dem, før undertypen *Long Short-Term Memory* (LSTM) presenteres.

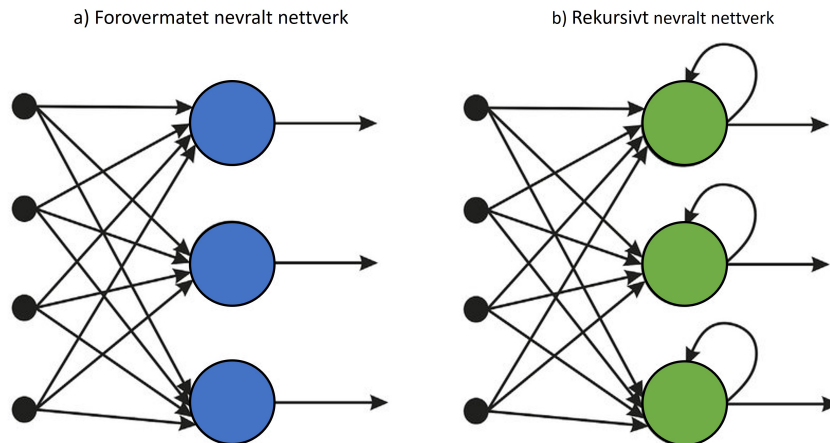
2.2.5 Rekursive nevrale nettverk

Rekursive nevrale nettverk (RNN) er laget for å kunne håndtere sekvenser, og kan huske tidligere informasjon og prosessere nye hendelser og informasjon deretter.

I vanlige nevrale nettverk sendes informasjonen fra input-variablene til input-laget til de skjulte lagene og deretter til output-laget. Informasjon om tidligere prøver som har passert laget huskes ikke fra en prøve til en annen, kun vektene i nettverket blir oppdatert.

I et RNN får derimot et skjult lag for et bestemt tidssteg input fra både tidligere lag, for eksempel input-laget, og det samme skjulte laget fra forrige tidssteg. Dette gjør at nettverket har et minne om tidligere hendelser. Flyten vises vanligvis som en loop, også kalt en rekursiv kant. Figur 2.5 illustrerer hvordan data i et forovermatet nettverk kun beveger seg mellom lagene, mens i rekursive nevrale nettverk beveger informasjonen seg også innad i samme lag på tvers av tidssteg.

Enhetene i de skjulte lagene i et RNN er i likhet med andre nevrale nettverk koblet sammen gjennom vekter. I tillegg er de enkelte enhetene i hvert lag koblet sammen med samme enhet for tidligere og senere tidssteg. Vektene i hver enhet er ikke avhengig av tid; de er felles for alle tidssteg. Beregningen av aktiveringen i hvert skjulte lag gjøres basert på vektkoeffisientene i laget og input fra både lag til lag- og tidssteg til tidssteg-koblinger.



Figur 2.5: Sammenligning av a) et forovermatet nevralt nettverk (FFNN) og b) et rekursivt nevralt nettverk (RNN). I FFNN beveger informasjonen seg i en retning, mens i RNN er det en rekursiv kant og flyt av informasjon mellom både tidssteg og lag. Modifisert fra Eliasy og Przychodzen. *The role of AI in capital structure to enhance corporate funding strategies*. 2020. URL: [sciencedirect.com/science/article/pii/S2590005620300023#fig2](https://www.sciencedirect.com/science/article/pii/S2590005620300023#fig2). Creative Commons lisens creativecommons.org/licenses/by/4.0/ (hentet 01.12.2022).

Ved tilbakepropagasjon av feil, som ble beskrevet i delkapittel 2.2.2, beregnes det overordnede tapet til hele nettverket som summen av tapsfunksjonene for hvert tidssteg. For å beregne en gradient for feilen for hvert tidssteg inngår en derivasjon av input fra tidligere lag. Denne inkluderer en multiplikasjon av gradientene til tidligere lag.

På grunn av multiplikasjonsfaktoren i beregningen av gradienten til en tapsfunksjon, oppstår et problem med eksploderende eller forsvinnende gradienter. Man ønsker at absoluttverdien av gradientene skal holde seg stabilt rundt 1. Multiplikasjonsfaktoren kan innebære mange multiplikasjoner. Dersom noe blir multiplisert med et tall mellom 0 og 1 mange ganger vil resultatet bli mindre og mindre, eller gradvis forsvinnende, for hver multiplikasjon. I motsatt fall kan resultatet bli svært stort, det kan eksplodere, dersom man multipliserer med et tall over 1 tilsvarende mange ganger. Når man skal beregne avhengigheter mellom tidssteg langt fra hverandre i tid, noe man må ved lange tidsserier, kan man derfor få gradienter som eksploderer eller forsvinner. Da vil veldig lite informasjon fra tidlige tidssteg overføres til enden av tidsserien. Det finnes flere mulig løsninger på dette problemet, men en mye brukt løsning er en utvidelse av RNN-algoritmen: LSTM-algoritmen [33].

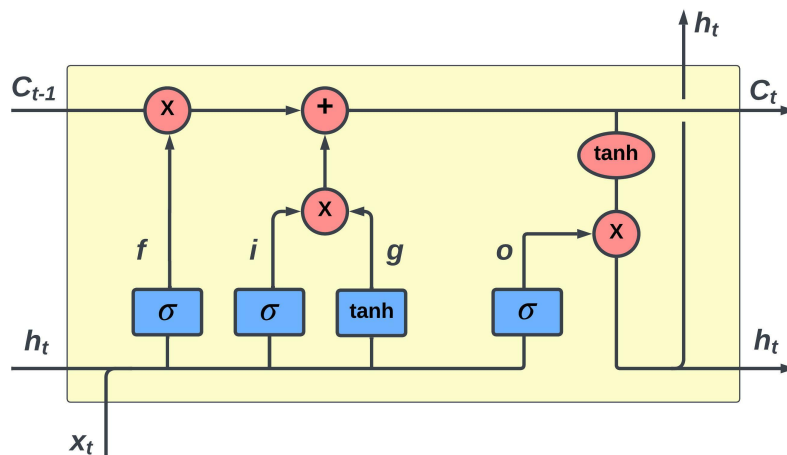
2.2.6 Long Short-Term Memory

Long Short-Term Memory (LSTM)-algoritmen ble først introdusert for å overkomme problemet med forsvinnende gradienter i læringsprosessen hos rekursive nevrale nettverk [33]. Hovedkomponentene i en LSTM er hukommelsescellene, som representerer de skjulte lagene i nettverket.

Figur 2.6 viser oppbyggingen til en LSTM-celle. Et vanlig RNN-lag vil for tidssteg t ta inn input-data x_t og verdiene h_{t-1} fra forrige tidssteg og prosessere begge input sammen i en aktiveringsfunksjon. Dette forårsaker problemet med forsvinnende eller eksploderende gradient, som beskrevet i delkapittel 2.2.5. Et RNN sender ut en output o_t , og verdiene h_t sendes videre til neste lag.

En LSTM-celle har i motsetning til et RNN også en rekursiv kant, en kobling mellom skjulte lag ved ulike tidssteg. I tillegg til input og output som et RNN har, har LSTM-cellen også en viss celletilstand C_t for hvert tidssteg. Denne tilstanden modifiseres mellom et tidssteg og det neste uten å multipliseres direkte med en vektfaktor. I stedet blir den oppdatert med verdier avledet fra disse vektene.

LSTM-cellen i figur 2.6 tar inn input-dataene x_t for tidssteg t , sammen med celletilstanden C_{t-1} og aktiveringen h_{t-1} fra forrige tidssteg. I cellen er det tre typer *port*er og en *node*, som hver har en Sigmoid- (σ) eller hyperbolsk tangent-aktiveringsfunksjon (\tanh). Porten lengst til venstre er glemselsporten, f , og har som formål å tilbakestille celletilstanden uten at den skal vokse uendelig. Altså å bestemme hvilken informasjon som får passere gjennom. Den neste porten, input-porten i , oppdaterer sammen med input-noden g celletilstanden. Den siste porten, output-porten o , bestemmer hvordan verdiene i de skjulte enhetene h_t skal oppdateres [29]. h_t sendes både til neste lag for samme tidssteg og til samme lag i neste tidssteg. Til neste tidssteg sendes også den oppdaterte celletilstanden C_t .



Figur 2.6: Strukturen innad i en LSTM-celle, med portene f , i og o og noden g som kontrollerer minnet i cellen. Basert på Raschka og Mirjalili [29, s. 548].

Det finnes mange variasjoner av LSTM beskrevet i litteratur og anvendt i praksis, blant annet toveis LSTM og konvolusjonell LSTM [34].

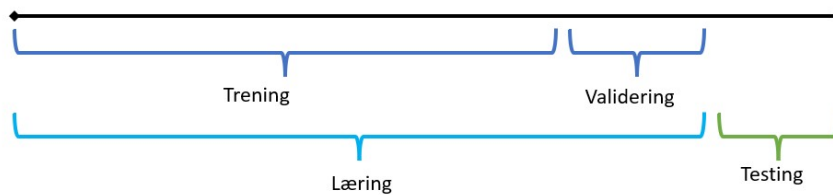
Strukturen og de underliggende beregningene i rekursive nevralt nettverk er komplekse, og matematikken er ikke forklart her. I praksis finnes implementeringer av disse algoritmene i maskinlæringsbiblioteker, for eksempel Tensorflow [35] [29]. Det gjør bygging modeller til prediksjon på praktiske problemstillinger og datasett mind-

re komplisert. I de neste delkapitlene beskrives hensyn man må ta angående oppdeling av data og evaluering av modeller i læringsprosessen av en maskinlæringsmodell.

2.2.7 Læringsprosessen

Prosessen med å trene en ny modell til å finne og bevare nye sammenhenger i data kalles læringsprosess. Etter læringen må man teste modellen på usett data. Figur 2.7 viser hvordan man kan dele dataene inn i trenings-, validerings- og testdatasett. Disse tre datasettene har ulike formål i maskinlæringsprosessen.

En modell blir først trent på å finne sammenhenger mellom dataene i treningssettet ved hjelp av en læringsalgoritme. Dette gjøres gjerne i flere omganger. Mellom omgangene undersøkes det hvor god modellen er ved å teste den på valideringsdata. Basert på resultatene fra valideringen trenes modellen igjen på treningsdatasettet. Dette gjentas til modellen er tilstrekkelig god. Valg av modellparametere skjer basert på disse stegene. Hvilke modellparametere, altså egenskaper ved modellen, som kan tilpasses varierer fra algoritme til algoritme. Deretter tester man modellen på et helt nytt datasett som kalles testsett. Dersom man bruker testsettet på nytt og på nytt for å velge beste modell i seleksjonsprosessen, vil testdataene bli en del av treningssettet og sannsynligheten for å overtilpasse modellen øker [29].



Figur 2.7: Et datasett deles inn i treningsdata, valideringsdata og testdata som brukes i lærings- og testprosessene.

Det er ønskelig å bruke den største andelen av dataene til trening, men også å ha et stort nok testsett til å få testet generaliserbarheten til modellen. For et større datasett kan størrelsen på testsettet reduseres. Det er vanlig å bruke 80 % av data til trening og 20 % til testing. Dersom man har store mengder data kan andelen treningsdata økes, og de andre delene reduseres tilsvarende [29].

Målet for justeringen av parametere til maskinlæringsalgoritmene er å finne modellparametere som gir en løsning som er så nær den beste løsningen som mulig, som omtalt i delkapittel 2.2.3. Den beste løsningen er den som gjør at det predikerte output blir så likt som mulig verdiene man ønsker å predikere. Det kan skje at modellen er for lite eller for mye trent, henholdsvis under- og overtilpasset.

En undertilpasset modell har ikke lært å ekstrahere viktige mønstre og relasjoner mellom forklaringsvariabler og responsvariabler i treningsprosessen. Undertilpassede modeller gir dermed dårlige eller upålitelige prediksjoner på både trenings- og testsett. For å forbedre et undertilpasset nevralt nettverk trengs ofte flere eller større

lag eller færre begrensninger i form av regulering. Et større datasett med andre eller flere forklaringsvariabler eller endring i pre-prosesseringen kan også forbedre en undertilpasset modell.

Overtilpasning av en modell er det motsatte. En overtilpasset modell har typisk lært mønstrene i dataene den er trent på for godt. Modellen generaliserer derfor dårlig det den har lært til nye data, og vil av den grunn gi dårlige prediksjoner. En overtilpasset modell må gjerne regulariseres mer eller ha færre forklaringsvariabler [30].

Målet er å ha en godt tilpasset modell, med en balanse mellom over- og undertilpasning. Resultatene på valideringsdatasettet underveis i treningen gir en pekepinn på hvordan modellen nærmer seg en optimal løsning. Under hver epoke oppdateres vektene i hver enhet i nettverket. Dette gjøres med mål om at modellen blir litt bedre tilpasset dataene for hver epoke [29].

2.2.8 Evalueringsmål

I maskinlæringsprosessen trengs metoder for å vurdere treffsikkerheten til modellene, både underveis i treningen og for å sammenligne modeller i etterkant. Noen vanlige metoder for å evaluere modeller innenfor tidsserieanalyse og dyplæring er *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE) og *Mean Absolute Percentage Error* (MAPE). For samtlige av disse evalueringmålene er små verdier bedre [34].

MSE er den gjennomsnittlige verdien for kvadratet av avviket mellom predikert og sann verdi. Likningen for denne feilberegningen er

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (2.1)$$

der n er antallet observasjoner i utvalget, y_i er den sanne verdien for hver av de i observasjonene i utvalget og \hat{y}_i er tilsvarende den predikerte verdien. MSE-scoren vil ha enheten til y_i opphøyd i andre [29].

For å gjøre scoren enklere å tolke kan man ta kvadratet av MSE-verdien, og få RMSE. Den vil ha samme enhet som observasjonene, og regnes ut slik [29]:

$$RMSE = \sqrt{MSE}. \quad (2.2)$$

MSE- og RMSE-verdiene straffer store feil mer enn små feil, fordi straffen øker med kvadratet av avviket for hver prediksjon som sammenlignes [36].

En annen scoringsmetode for regresjonsmodeller er MAE. For hver individuelle prediksjon som gjøres kan man beregne avviket fra den sanne verdien. Gjennomsnittet av absoluttverdiene til disse avvikene er prediksjonens MAE-verdi. Verdien kan beregnes ved [37]:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (2.3)$$

I likhet med RMSE får feilverdien samme enhet som de predikerte verdiene, men i motsetning til RMSE er endringene i MAE lineære. Med MAE vektet ikke store feil mer enn små feil, i stedet øker scoren lineært med økning i feil [36]. MAE-verdien vil være lavere enn RMSE-verdien for samme datasett.

Et annet vanlig evalueringsmål er MAPE. Den er gjennomsnittet av absoluttverdien til avviket mellom predikert og faktisk verdi, relativt til de faktiske verdiene. Den er oppgitt i prosent og kan følgelig sammenlignes mellom datasett med ulik skala [36]. MAPE beregnes ved [34]:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100. \quad (2.4)$$

I dette kapitlet ble det presentert teori om maskinlæringsmodeller og opptreningen av disse, med fokus på LSTM-algoritmen. I neste delkapittel skildres anvendelser av maskinlæringsmodeller i kraftnettsektoren.

2.3 Maskinlæring i kraftsystemet

Det finnes mange ulike maskinlæringsalgoritmer som er anvendt til prediksjon av fremtidig energibruk og elektrisitetsbruk for både husstander, næringsbygg og industri. I dette delkapitlet beskrives anvendelser av maskinlæringsalgoritmer i kraftsystemer, med fokus på prediksjonsmodeller for elektrisitetsbruk i næringsbygg.

Ifølge Shabani og Zavalani [38] kan metodene for å predikere energibruk i bygninger deles inn i to grupper: ingeniørbasert og datadrevet. Ingeniørbasert prediksjon bygger gjerne på en digital modell av en bygning basert på dens fysiske egenskaper og brukes i tidlige designfaser av bygg. De datadrevne metodene benytter statistiske verktøy og maskinlæring, og anvendes i større grad i driftsfasen av bygg. Videre skriver de at dagens *State of the Art* innen maskinlæringsbaserte prediksjonsmodeller er kunstige nevralt nettverk og *Support Vector Machine* (SVM) [38].

Hvilken maskinlæringsalgoritme man bruker for datadrevet prediksjon avhenger av den spesifikke problemstillingen. I noen tilfeller ønsker man å predikere spesifikt energibruk for ulike bygninger basert på bygningsparametre. Abbass og Hamdy [30] predikerte energibruk per areal med et dypt nevralt nettverk med 25 lag. Forklæringsvariablene var bygningskarakteristikker og energibruksdata for en rekke bygninger [30]. Dette er et regresjonsproblem.

I andre tilfeller ønsker man, som beskrevet i delkapittel 2.1.7, å generere lastprofiler eller predikere tidspunkter for høy og lav el-bruk til styring- og fleksibilitetsformål.

Fremtidig lastprofil, det vil si en fremstilling av el-bruk som funksjon av tid, kan estimeres basert på tidligere el-bruk og andre eksterne parametre. Gjøres dette med maskinlæring kalles det innenfor kraftnett for *Intelligent Load Forecasting* (ILF), eller intelligente prognoser av last. Det kan gjøres med korte eller lange tidshorisonter, fra ett minutt eller en time til år eller tiår i forveien [39]. Dette er også regresjonsproblem, men med sekvensielle data.

Lastprognoser kan genereres på ulike nivåer og ledd av kraftsystemet. I produksjonsleddet gjøres prognoser av elektrisitetsproduksjon fra fornybare energikilder og av kraftpris [40]. I overføringsleddet utarbeides lastprognoser på aggregerte nivå for hele land eller regioner, slik Rafi mfl. gjorde på kort sikt for kraftsystemet i Bangladesh [41]. I forbrukerleddet kan det også utarbeides lastprognoser for enkelte bygninger eller anlegg.

For eksempel bygde Slowik og Urban [42] en LSTM-modell for å predikere el-bruk i en fabrikk på fortløpende innkommende data med en kort tidshorison. Prediksjonene var basert på historisk el-bruk. Denne type prediksjoner gjør det mulig for sluttbrukerne å delta i et smart energimarked [42].

Hwang mfl. [43] presenterte en algoritme for optimalisering av planlegging av opp- og utladning av et energilagringssystem med mål om å redusere topplastene til en bygning. Som en del av algoritmen lagde de lastprognoser basert på en modell fra Yudiantaka mfl. [44]. Prediksjonsmodellen predikerte effekt for neste døgn basert på forrige døgn. Nøyaktigheten til lastprognosene som ble brukt i optimaliseringen var essensiell, fordi dårlige prognoser ville føre til redusert utbytte fra energilagringssystemet [43]. I sine evalueringer av modellen fokuserte de på toppene i effektbruken, og trengte følgelig modeller som presterte best ved de tidspunktene.

Hwang mfl. [7] lagde prognoser for total elektrisitetsbruk i næringsbygg med daglig og månedlig oppløsning, med mål om å bedre kunne planlegge el-bruken i bygningen. De kom frem til en to-stegsmodell hvor andre steg besto av å teste tradisjonelle teknikker for tidsserieanalyse mot maskinlæringsteknikker. De brukte fem modeller: den tradisjonelle SARIMAX, maskinlæringsalgoritmen *Support Vector Regression* og de nevralt nettverkstypene ANN, *Deep Neural Network* (DNN) og LSTM. De konkluderte med at DNN og LSTM hadde høyest presisjon [7].

Som en del av et fleksibilitetsestimeringsrammeverk predikerte Haug [32] i sin masteroppgave elektrisitetsbruken til et maskinrom ved hjelp av et rekursivt nevralt nettverk med to LSTM-lag. Basert på disse prediksjonene kunne eieren av bygget melde inn fleksibilitetspotensiale til et fleksibilitetsmarked [32].

Det finnes mange algoritmer innen statistikk og dyplæring som har vært brukt til å lage lastprognoser i kraftnett. I tillegg finnes kombinasjoner av en eller flere slike algoritmer, såkalte hybride modeller. Mange av algoritmene er komplekse og tar lang tid å bygge opp og/eller tilpasse hvert enkelt problem. Dermed kan de være mindre anvendelige i praksis enn de enklere metodene [45].

Kapittel 3

Metode

Veiledede maskinlæringsalgoritmer kan, som beskrevet i kapittel 2, brukes til å lage prognoser for elektrisitetsbruken til bygninger. I denne oppgaven undersøkes det hvordan maskinlæringsalgoritmer kan brukes til å predikere elektrisitetsbruken¹ ved Oslo lufthavn Gardermoen (OSL) én og 24 timer frem i tid.

Algoritmen som blir testet er rekursive nevralt nettverk (RNN) med skjulte lag av typen *Long Short-Term Memory* (LSTM), som ble beskrevet i delkapittel 2.2.6. Dette ble valgt fordi RNN er godt tilpasset tidsserier hvor rekkefølgen på dataene er viktig [29], som er tilfellet for elektrisitetmålinger. Videre er undertypen LSTM valgt fordi den unngår problemet med eksplodert og forminsket gradient. Dette er et problem som generelle rekursive nettverk har, som beskrevet i delkapittel 2.2.5.

Databehandlingen og maskinlæringen til denne oppgaven ble gjort i Python. De nevralt nettverkene ble bygd, trent og testet ved hjelp av grensesnittet Keras og maskinlæringsbiblioteket Tensorflow [35]. Den praktiske implementeringen av modellene er inspirert av Haug [32]. Skriptene som ble skrevet for å utføre prediksjonene er tilgjengelig fra den nettbaserte utviklerplattformen GitHub [46].

For å undersøke hvordan endringer i forklaringsvariablene og egenskapene til modellene påvirket resultatene fra prediksjonene ble det satt sammen en basismodell som ble utvidet på 37 ulike måter. Etter en innføring i databehandlingen som ble gjort, blir generelle trekk ved modellene beskrevet, etterfulgt av detaljer om konfigurasjonene som ble testet med henholdsvis én og 24 timer prediksjonshorisont.

3.1 Databehandling

Forklaringsvariablene i prediksjonene er målinger av lufttemperatur, vindhastighet og nedbør fra Gardermoen målestasjon, antallet passasjerer i flyplassterminalene og tidsvariabler. Målingene av elektrisitetsbruk presenteres i delkapittel 4.2 og forklaringsvariablene beskrives i delkapittel 4.3.

¹Aggregerte timeverdier for brukt elektrisitet [kWh/h] for hele flyplassen, unntatt elektrisiteten som brukes til oppvarming gjennom el-kjel og varmpumper i varmesentralen.

Som første steg i maskinlæringsprosessen ble datasettene lastet ned og sammenstilt for tidsrommet 2017 til 2021. Datasettene ble hentet fra sine respektive kilder Meteorologisk institutt [47] og Avinor AS [48] [49], og koblet sammen basert på tidsstempel fra elektrisitetmålingene fra flyplassen. Basert på datautforskningen som presenteres i kapittel 4 ble de kortere tidsperiode, sommerhalvåret 2019 (01.04-30.09.2019), valgt til prediksjonsformål. Dette er en periode uten manglende eller avvikende verdier.

Før dataene ble sendt inn i maskinlæringsmodellen ble de forbehandlet. Det innebar å fjerne ugyldige verdier, transformere dataene til et passende format og skalere dataene. De fleste maskinlæringsalgoritmer fungerer bedre dersom dataene som sendes inn har samme verdimengde, for eksempel verdier mellom 0 og 1 [29]. Dette ble gjort ved å først standardisere verdiene for så å skalere dem [10].

I tillegg ble Pearson-korrelasjonen mellom responsvariabelen el-bruk og forklaringsvariablene beregnet for periodene 2017 til 2019 og sommerhalvåret 2019. Den ble beregnet før skaleringen av dataene. Pearson-korrelasjonen er et kjent mål på korrelasjon mellom variabler. Positive verdier indikerer at to variabler øker eller minker i takt, mens negative verdier tilsier at en økning i en variabel medfører en minking i en annen og motsatt [50].

Tidsvariablene i det opprinnelige datasettet var kategoriske variabler. For at modellen ikke skal assosiere høyere verdier med bedre eller dårligere ble disse kategoriske variablene derfor byttet ut med indikatorvariabler. For eksempel ble ukedag i det opprinnelige datasettet markert med 0 for mandag og 6 for søndag. Det prosesserte datasettet hadde i stedet en kolonne med navn *mandag* med verdien 1 for alle timer på mandager, og 0 ellers. Det ga 24 nye time-variabler, 53 uke-variabler og sju ukedagsvariabler.

De neste delkapitlene beskriver implementeringen av de rekursive nevralt nettverkene. Først beskrives felles elementer ved alle modellkonfigurasjonene, før modellkonfigurasjonene som ble testet for å predikere el-bruken ved Oslo lufthavn én time frem i tid blir presentert. Til sist presenteres konfigurasjonene som ble testet til prediksjon 24 timer frem i tid.

3.2 Modelloppbygging og generelle valg

Det ble først bygd en enkel modell ved hjelp av biblioteket Scikit-learn [51]. Den enkle modellen predikerer gjennomsnittlig timeverdi for alle timer. Resultatene fra denne modellen blir presentert i delkapittel 5.3.1 Det er ønskelig at andre modeller skal predikere bedre enn den enkle modellen.

For å få reproduserbare resultater settes et *seed* i modellene, som er en innstilling for tilfeldig omstokking og vektinitialisering i modellene [29]. Dette kan påvirke ytelsen og robustheten til modellene [52]. Det er ønskelig å minske påvirkningen av seed-et på resultatene. Derfor ble hver av modellkonfigurasjonene trent tre ganger med ulike seed. Rundene omtales videre som RNN1, RNN2 og RNN3. For hver

modellkonfigurasjon ble det tatt gjennomsnitt av tallscorene fra de tre rundene for å minske påvirkningen av tilfeldighet.

For å kunne sammenligne visuelle resultater for modellene ble prediksjonene på trenings- og testdelene av datasettet plottet sammen med de sanne verdiene for de respektive tidsperiodene. Prediksjonsplottene som presenteres i kapittel 5 er basert på prediksjonene fra første runde, RNN1.

De nevrale nettverkene ble bygd ved å legge sammen ulike modellkomponenter fra Tensorflow-biblioteket [35]. Modellene ble kompilert med optimeringsfunksjonen Adam [53]. Videre ble *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE) og *Mean Absolute Percentage Error* (MAPE) brukt som evalueringsmål. Disse ble beskrevet i delkapittel 2.2.8. Det ble valgt å bruke alle fire evalueringsmetodene i denne oppgaven for å utvide sammenligningsgrunnlaget mot andre forskningsprosjekter.

MSE ble også brukt som tapsfunksjon, som vil si at underveis i treningen ble de beste modellene valgt basert på MSE. Siden RMSE er kvadratet av MSE og har samme enhet som den predikerte verdien, ble bare RMSE rapportert i resultatene. I etterkant av treningen og testingen ble modellene også sammenlignet basert på hvor godt de visuelt traff de faktiske verdiene for el-bruk.

Modellene ble brukt til å predikere bruken av elektrisk energi én eller 24 timer frem i tid. Et rullerende vindu med spesifisert tilbakeblikk ble brukt slik at modellen for hvert tidssteg hadde en spesifisert mengde tidligere timer som grunnlag for prediksjonen. Datasettet ble delt inn i tre deler: treningsdata, valideringsdata og testdata, slik beskrevet i delkapittel 2.2.8.

For hver epoke i treningen ble det sendt data med de tre dimensjonene [*batch*-størrelse, tilbakeblikk, antall variabler] til modellen. Batch-størrelsen anga antallet timer som skulle sendes inn i modellen for å bli predikert om gangen. Innad i hver batch var det kronologisk rekkefølge på dataene slik at hvert sett med forklaringsvariabler ble koblet til riktig responsvariabel. Batchene ble ikke sendt inn i kronologisk rekkefølge, noe som gjorde at modellen ikke lente seg for mye på verdiene fra forrige batch. Etter at modellen hadde trent ferdig på alle epokeer ble prediksjonene hentet ut med dimensjonen [antall timer], altså én predikert verdi for hver time.

For hver epoke underveis i treningen sammenlignet modellen MSE-scoren på valideringssettet med verdiene for tidligere epoke-er. Når scoren ikke forbedret seg på et visst antall epoke-er, avsluttet den treningen. Dette kalles tidlig stopping, og reduserer sjansen for at modellen overtilpasses. I tillegg kan det spare regnekapasitet og dermed tid. I denne implementasjonen var antallet epoker modellene ventet før de avbrøt treningen 20.

Når modellen nådde et platå hvor verdien til tapsfunksjonen ikke har minket på to epoker ble også læringsraten redusert til 10 % av tidligere verdi. At scoren til en modell flater ut kan indikere at man nærmer seg den optimale løsningen, som beskrevet i delkapittel 2.2.2. Å redusere læringsraten gjør at modellen har større sjanse for å finne den optimale løsningen.

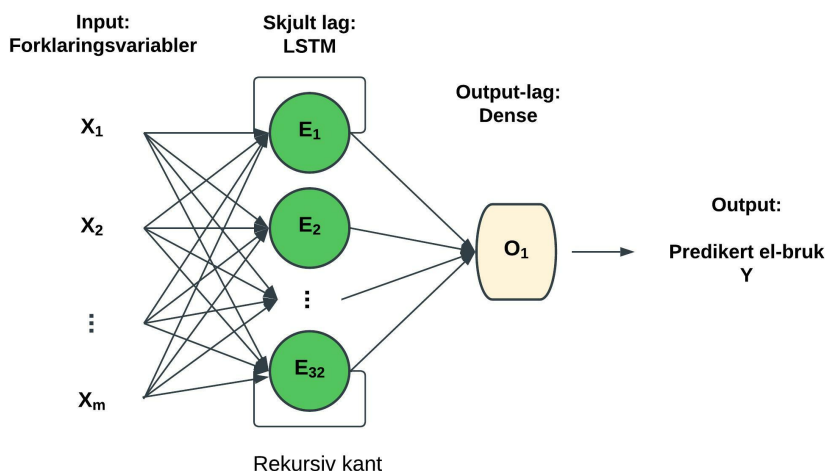
3.3 Basismodell

Det ble satt sammen en modell med parametere basert på beste praksis, heretter kalt basismodell. Videre ble ulike aspekter ved basismodellen endret for å se hvordan det påvirket treffsikkerheten til prediksjonene av el-bruk. Disse endringene blir presentert i delkapittel 3.4. Oppbyggingen til basismodellen er illustrert i figur 3.1.

Figur 3.1 viser lagene som ble satt sammen til basismodellen. I likhet med det nevralt nettverket i figur 2.4 tar modellen inn et sett med forklaringsvariabler som prosesseres i et skjult lag og output sendes ut fra et output-lag. Input-laget er i modeller i Keras integrert i LSTM-laget. Det første laget i figur 3.1 er derfor LSTM-laget med 32 enheter, E_1 til E_{32} på figuren. Hver enhet er en LSTM-celle som ble beskrevet i delkapittel 2.2.6. I figur 3.1 er output-laget et *Dense* lag med én enhet, O_1 . Dette laget får informasjon fra LSTM-laget og gir ut én prediksjon for hvert tidssteg. En oppsummering av modelloppbyggingen, inkludert antallet parametere i hvert lag, er vist i figur C.1 i vedlegg C.

I begge lag ble standard aktiveringsfunksjoner beholdt. I *Dense*-laget var det en lineær aktiveringsfunksjon som ikke endret verdiene som sendes fra det skjulte laget. Aktiveringsfunksjonene i LSTM-laget var som omtalt i 2.2.6. Modellen hadde 0 % forenkling underveis, omtalt som *dropout* i delkapittel 2.2.3.

Et tilbakeblikk på én uke, det vil si 168 timer, ble valgt for basismodellen. Dataene ble splittet inn i trenings-, validerings- og testdata etter forholdet 80:10:10 [42] [29]. Basismodellen har batch-størrelse 32.



Figur 3.1: Nettverksarkitekturen til basismodellen. For hvert tidssteg tar modellen inn n forklaringsvariabler for å predikere én output. Modellen har ett skjult LSTM-lag med 32 enheter.

Basismodellen benyttet samtlige forklaringsvariabler, unntatt tidligere timeverdier for elektrisitetsbruk, og ble brukt til å predikere el-bruk ved Oslo lufthavn både én og 24 timer frem i tid. Forklaringsvariablene og responsvariabelen timeeffekt [kWh/h]

beskrives nærmere i kapittel 4. Scorene fra prediksjonen med basismodellen én time frem i tid presenteres i delkapittel 5.3.2. Scorene fra prediksjonen 24 timer frem i tid ved hjelp av basismodellen blir presentert i delkapittel 5.4. Neste delkapittel beskriver de øvrige modellkonfigurasjonene som ble testet på datasettet.

3.4 Modellkonfigurasjoner for prediksjon én time frem i tid

I denne oppgaven blir det kartlagt hvordan rekursive nevrale nettverk med LSTM-lag kan brukes til å predikere el-bruken ved Oslo lufthavn. Siden denne oppgaven er første steg i å utvikle prediksjonsmodeller for el-bruken ved OSL ble det lagt vekt på å teste flest mulige variasjoner av LSTM-modellene. For å teste påvirkningen av ulike egenskaper og variabler på prediksjonsresultatene, ble basismodellen endret på to måter:

1. Ved å endre nettverksarkitekturen
2. Ved å endre input-dataene

Tabell 3.1 viser hvilke endringer i nettverksarkitekturen som ble testet. Et utvalg endringer ble gjort, så ikke alle mulige kombinasjoner i tabellen ble testet. I stedet ble det fokusert på å endre et aspekt om gangen. Det ble testet hvordan splitter mellom trenings-, validerings- og testsett på 70:10:20 eller 70:20:10 og tilbakeblikk på to, fem, ti og 14 dager påvirket modellen. Det ble undersøkt om forenkling av nettverket underveis gjorde resultatene bedre. I tillegg ble det undersøkt om antallet celler i hvert skjulte lag hadde en påvirkning og om et økt antall skjulte lag gjorde modellen bedre. Ved to skjulte lag ble det testet om resultatet ble best ved å ha lik eller ulik størrelse på lagene. Det ble også undersøkt hvordan modellen reagerer på økt batch-størrelse og endret aktivering i output-laget.

Tabell 3.1: Alternativene for antall skjulte lag, antall enheter i hvert lag, forenkling, splitt og tilbakeblikk som ble testet i modellene for å predikere el-bruk. Egenskaper i kursiv ble brukt i basismodellen.

Antall lag	Antall enheter	Forenkling	Tilbakeblikk	Splitt
<i>1</i>	16	<i>0 %</i>	24 · 2	70:20:10
<i>2</i>	<i>32</i>	10 %	24 · 5	70:10:20
<i>3</i>	64	20 %	<i>24 · 7</i>	<i>80:10:10</i>
	128	50 %	24 · 10	
			24 · 14	

Modeller med samme nettverksarkitektur som basismodellen ble også testet med ulike sammensetninger av forklaringsvariabler for å se hvordan dette påvirket resultatene. Kombinasjonene som ble testet vises i tabell 3.2. Basismodellens forklaringsvariabler står i kursiv. Time på døgnet ble inkludert som forklaringsvariabel i alle unntatt én konfigurasjon. For tre av kombinasjonene som inneholdt tidligere el-bruk ble også nettverk med flere lag og/eller flere enheter i hvert lag testet. Tidsvariabler

inkluderer her ukenummer, ukedag, markør for helgedag og markør for helligdag. Forklaringsvariablene blir nærmere beskrevet i kapittel 4.

Tabell 3.2: Kombinasjoner av forklaringsvariabler som ble testet til prediksjon av el-bruk én time frem i tid. Der annet ikke er spesifisert, ble time på døgnet også brukt som forklaringsvariabel. Variablene i kursiv ble brukt i basismodellen.

Variabelkombinasjoner
<i>Tidsvariabler, passasjerantall, nedbør, vindhastighet og lufttemperatur</i>
Tidsvariabler, passasjerantall, alle vær-variabler og tidligere el-bruk
Tidsvariabler, nedbør, vindhastighet og lufttemperatur
Nedbør, vindhastighet og lufttemperatur
Tidsvariabler, passasjerantall og lufttemperatur
Tidsvariabler og passasjerantall
Tidsvariabler og tidligere el-bruk
Tidligere el-bruk
Kun tidligere el-bruk, ikke inkludert time på døgnet
Ukenummer og ukedag
Passasjerantall
Tidsvariabler

Det ble i alt konstruert 38 ulike kombinasjoner av forklaringsvariabler og nettverksarkitektur, inkludert basismodellen. Disse ble brukt til å predikere el-bruk én time frem i tid i tre runder, RNN1 til RNN3, som beskrevet i introduksjonen til kapittel 3. De elleve prediksjonsmodellene som fikk best RMSE-score på testsettet for en-timesprediksjon blir presentert i delkapittel 5.3.3. Fullstendige tall-resultater og prediksjonsplott fra RNN1 finnes i henholdsvis vedlegg B og C.

Fire utvalgte konfigurasjoner som fikk god score på prediksjonen én time frem i tid, ble testet yil prediksjon 24 timer frem i tid. Hvilke modeller og hvordan de ble tilpasset en lengre prediksjonshorisont beskrives i neste delkapittel.

3.5 Prediksjon et døgn frem i tid

For å undersøke om maskinlæringsalgoritmen egnet seg til prediksjon av el-bruk lengre frem i tid ble fire modellkonfigurasjoner testet til prediksjon av timeverdier for el-bruk 24 timer frem i tid. Modellene ble valgt basert på treffsikkerhet på en-timesprediksjon, og hadde enten samme nettverksarkitektur som basismodellen eller en dypere arkitektur. Alle modellene hadde tidsvariabler, meteorologiske variabler og passasjerantall som forklaringsvariabler, og to av modellene hadde i tillegg tidligere el-bruk. De hadde ellers samme innstillinger som basismodellen, som ble beskrevet i delkapittel 3.3.

Modellene som ble testet var:

- Basismodell

- Modell med samme nettverksarkitektur som basismodellen og inkludert tidligere el-bruk som forklaringsvariabel
- Modell med tre lag og 64 enheter i hvert lag
- Modell med tre lag og 64 enheter i hvert lag og inkludert tidligere el-bruk som forklaringsvariabel

For å predikere for en lengre tidshorisont ble responsvariabelen timeeffekt [kWh/h] forskjøvet 24 timer frem. Modellene kunne dermed klokken 01:00 for et døgn predikere el-bruken klokken 01:00 neste døgn. For å predikere hele neste døgn må modellene kjøres på hele et helt døgn. Modellene hadde, i likhet med tidligere beskrevne modeller, et spesifisert tilbakeblikk med tidligere timer som grunnlag for prediksjonen, ingen forenkling og 80:10:10 splitt.

Som tidligere ble disse modellene trent i tre runder med ulike seed, og det ble for hver konfigurasjon tatt gjennomsnitt av scorene fra de tre rundene. Som før blir treffsikkerheten til modellene illustrert med plott basert på første runde, RNN1. Resultatene fra tre av modellene som ble testet til prediksjon 24 timer frem i tid blir presentert i delkapittel 5.4 med scorer og plott.

I neste kapittel beskrives casestudiet og datasettene som ble brukt i prediksjonene av el-bruk ved Oslo lufthavn Gardermoen.

Kapittel 4

Case: Oslo lufthavn Gardermoen

Denne oppgaven er en innledende studie av hvordan maskinlæringsalgoritmer kan brukes til å predikere elektrisitetsbruken¹ til flyplassen Oslo lufthavn Gardermoen (OSLO). I dette kapitlet beskrives casestudiet og datasettene som ble brukt. Dette inkluderer undersøkelser av variasjonene i elektrisitetsbruken på flyplassen gjennom året og døgnet, samt av de metrologiske målingene og passasjerstatistikken fra OSL som ble brukt forklaringsvariabler. I tillegg blir variabler brukt til lignende prediksjoner i litteratur presentert.

4.1 Oslo lufthavn Gardermoen

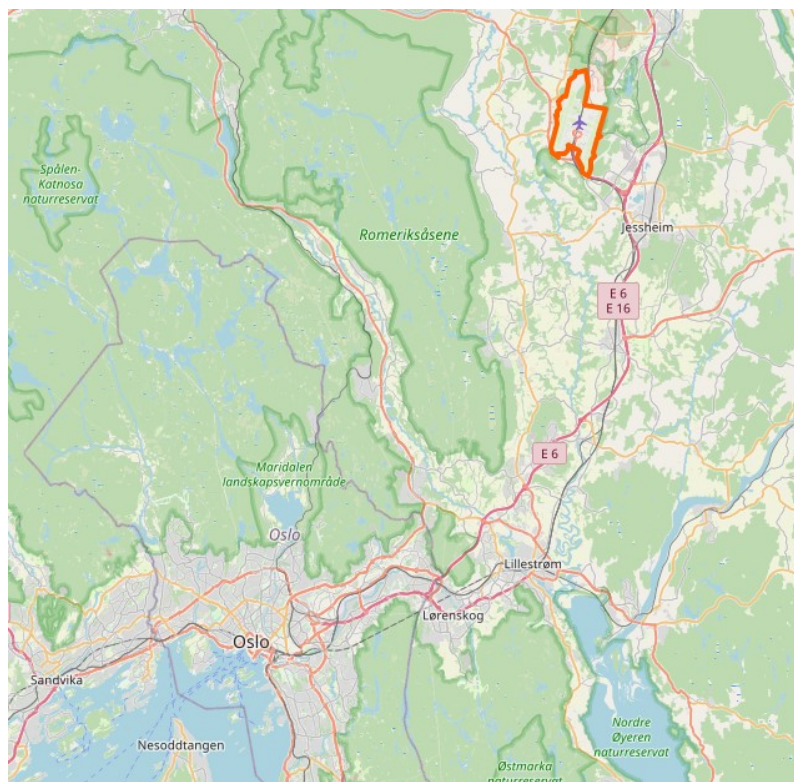
Oslo lufthavn er Norges største flyplass og ligger i Ullensaker og Nannestad kommuner i Viken fylke. I et normalår betjener flyplassen over 28 millioner reisende. Avinor AS drifter flyplassen [54]. Flyplassens plassering 40 km nord for Oslo ser man på kartet i figur 4.1, der området til OSL er markert i oransje.

I tillegg til to rullebaner og terminalbygningen, er det en rekke drifts- og administrative bygninger på OSL. Denne oppgaven presenterer prediksjon av den samlede bruken av elektrisitet til belysning, transport, varmekabler, klimaanlegg, tekniske anlegg, apparater og lignende for hele flyplassområdet. Elektrisitetsbruken i de ulike komponentene måles med høy tidsoppløsning, og er aggregert til timeverdier til bruk i denne oppgaven.

Ifølge Tveitane [55] betaler Avinor strømpris og nettleie for strømmen de bruker, hvor sistnevnte inkluderer forbruksavgift, el-sertifikater, påslag og effektledd [55]. Effektleddet bestemmes på månedsbasis av den høyeste timeeffekten per måned for den samlede el-bruken til hele anlegget på OSL [56]. OSL er tilkoblet regionalnettet. Å redusere den høyeste timeeffekten i måneden kan medføre besparelser.

I neste delkapittel presenteres målingene av timeeffekt som ble predikert og mønstre i elektrisitetsbruken ved flyplassen.

¹Aggregerte timeverdier for brukt elektrisitet [kWh/h] for hele Oslo lufthavn Gardermoen, unntatt elektrisiteten som brukes til oppvarming gjennom el-kjel og varmepumper i varmesentralen.



Figur 4.1: Oslo lufthavn Gardermoen som er markert i oransje ligger nord for Oslo, i Ullensaker og Nannestad kommuner. Skjermbilde fra OpenStreetMap [57].

4.2 Målinger av elektrisitetbruk fra Oslo lufthavn

Målet med oppgaven er å predikere mengden elektrisitet som brukes på OSL hver time, timeeffekten [kWh/h]. Målinger av timeeffekt for Oslo lufthavn Gardermoen ble lastet ned fra Avinor sitt energi- og miljøoppfølgingssystem Optima i august 2022.

Elektrisitetsbruk for flyplassen ble lastet ned for hver time i tidsperioden 01.01.2017 til 31.12.2021. Dataene var oppgitt i norsk tid, og det er tidsstempelet fra Optima som er lagt til grunn for sammenslåing med andre datasett. I Optima var døgnetts første time, fra midnatt til klokken 01:00, oppgitt som *time 0*. Dette ble konvertert til klokkeslettet 00:00:00. Dataene ble undersøkt og det ble avdekket manglende og avvikende verdier.

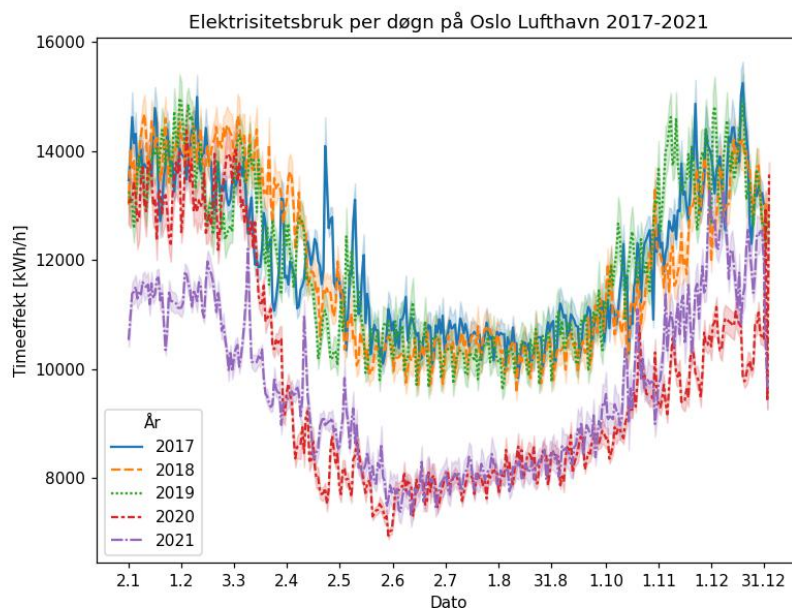
I forbindelse med stilling av klokken to ganger i året manglet målinger av timeeffekt for to timer per år. Den gjennomsnittlige timeverdien for alle målingene var 10 993 kWh/h med et standardavvik på 2 204 kWh/h. Enkelte timer i datasettet hadde avvikende verdier mer enn tre standardavvik fra gjennomsnittsverdien. Disse timene ble fjernet. Etter fjerning av timer med avvikende verdier eller manglende timeverdier var antallet timer i datasettet redusert med 0,55 %. Det resulterende datasettet blir presentert i de neste delkapitlene.

4.2.1 Mønstre i el-bruken i årene 2017 til 2021

Figur 4.2 viser den gjennomsnittlige elektrisitetsbruken per døgn for flyplassen for hele tidsperioden 2017 til 2021. Hvert år har ulik farge, og et skravert bånd representerer variasjonen i timeeffekt for timene innenfor samme døgn.

De gjennomsnittlige timeverdiene per døgn i figur 4.2 varierer fra 7 000 kWh/h sommeren 2020 til nærmere 15 300 kWh/h noen vinterdager. De faktiske målingene per time i hvert døgn varierer mer, fra 6 200 kWh/h til 16 900 kWh/h. Man ser av figur 4.2 at el-bruken er tydelig høyere om vinteren enn om sommeren, og at nivået også varierer mest fra dag til dag om vinteren. El-brukskurvene viser også at flyplassen har en høy grunnlast. I årene 2017 til 2019 faller timeeffekten ikke under 9 000 kWh/h. Dette er elektrisiteten flyplassen til enhver tid må trekke for å holde driften i gang.

Det er tydelig at årene 2017 til 2019 har årlig gjentakende mønstre, og at årene 2020 og 2021 skiller seg ut. I året 2020 kom et avvik fra den normale el-bruken i februar og deretter var timeverdiene på et lavere nivå enn tidligere år ut 2021. Dette sammenfaller med perioden Norge var inn og ut av nedstenginger i forbindelse med COVID-19-pandemien etter første nedstenging i mars 2020 [58]. På grunn av dette avviket fra den tidligere normalen ble de to årene 2020 og 2021 fjernet fra datasettet før videre arbeid for ikke å gjøre en studie i effektene COVID-19 hadde på el-bruken.



Figur 4.2: Gjennomsnittlig elektrisitetsbruk [kWh/h] for Oslo lufthavn per døgn i årene 2017 til 2021. Det skraverte beltet rundt hver linje representerer variasjonen innad i hvert døgn.

Variasjonen mellom kurvene som representerer årene 2017 til 2019 i figur 4.2 er størst på vinterstid. Variasjonen i el-bruk kan ha en sammenheng med at været varierer mer for enkeltuker fra år til år om vinteren enn sommeren. I tillegg varierer

datoene for avvikling av påskeferie og andre ferier noe mellom årene, noe som kan påvirke aktivitetsnivået og dermed el-bruken. Antallet passasjerer i terminalene og meteorologiske målinger for lufttemperatur, nedbør og vindhastighet blir undersøkt i henholdsvis delkapittel 4.3.2 og 4.3.3. Først beskrives variasjonene i el-bruken nærmere på døgn og sesongnivå.

4.2.2 Variasjon i el-bruk gjennom døgnet

Det finnes gjentakende mønstre i el-bruken til Oslo lufthavn gjennom døgnet, noe som illustreres i figur 4.3. Figuren viser gjennomsnittlige verdier for el-bruk per time i årene 2017 til 2019 separert på:

- a) Ukedager og helgedager
- b) Vanlige dager og helligdager
- c) Sesongene høst, vinter, vår og sommer

I figur 4.3 a) ser man forskjellene i gjennomsnittlig elektrisitetsbruk mellom ukedager og helgedager². El-bruken er høyere på ukedager enn helgedager, og toppene kommer ikke på samme tidspunkt. På ukedager er det en tydelig topp klokken åtte om morgenen på 13 000 kWh/h. Etter toppen flates forbruket ut på et nivå noe lavere enn toppen, og minker deretter utover ettermiddagen og kvelden. Den laveste el-bruken er hele uken om natten ca. klokken tre, på 10 700 kWh/h. På helgedager ligger el-bruken rundt sin høyeste verdi 12 000 kWh/h fra klokken 12 til 16, så det er ikke en tydelig topp slik som på ukedager.

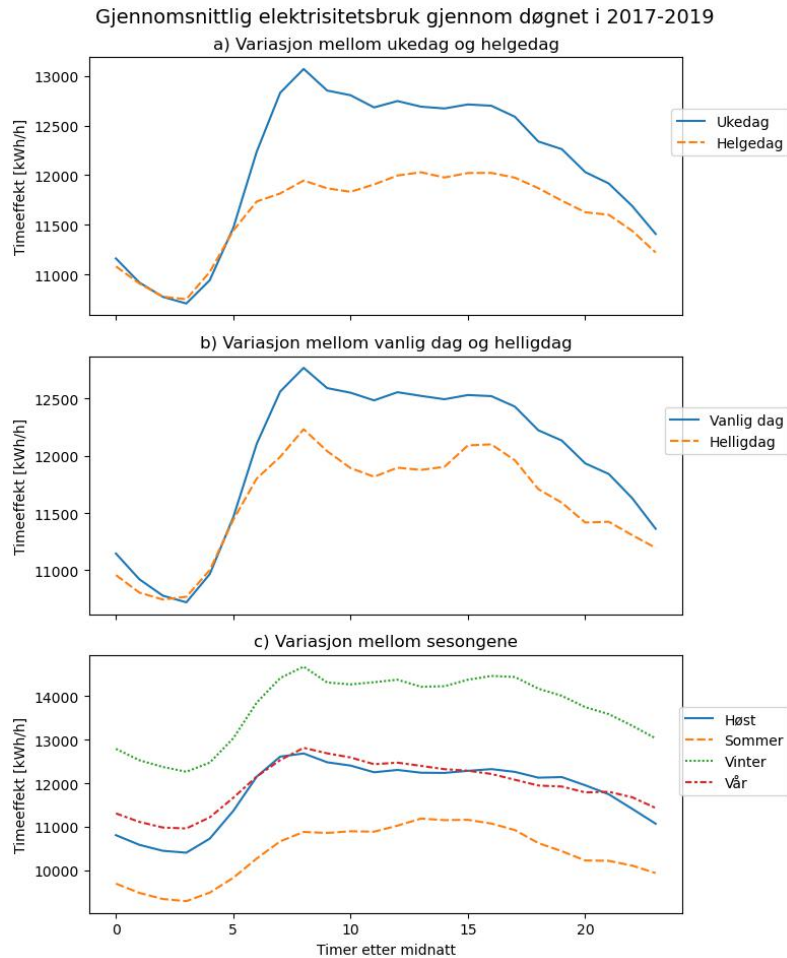
Det er også en forskjell i den gjennomsnittlige el-bruken på helligdager og vanlige dager, noe som er vist i figur 4.3 b). Vanlige dager inkluderer her både hverdager og helgedager. Helligdager inkluderer bevegelige helligdager, men ikke søndager³. På helligdager har elektrisitetsbruken to like høye topper på 12 000 kWh/h, en klokken åtte om morgenen og en fra klokken tre til fire om ettermiddagen. Den har veldig lik form som kurven for ukedager i figur 4.3 a), men noe lavere verdier.

Forskjellene mellom den gjennomsnittlige el-bruken mellom årets fire årstider er vist i figur 4.3 c). Årstidene er her delt inn i vår som mars til mai, sommer er juni til august, høst er september til november og vinter er desember til februar. El-bruken er høyest om vinteren og lavest om sommeren. Om våren, høsten og vinteren ligner formen på ukedagskurven i figur 4.3 a), mens den om sommeren ligner mer på helgedagskurven. Man ser at el-bruken i vår- og høstmånedene er nokså lik. Om vinteren er el-bruken høyere enn den gjennomsnittlige ukedagsbruken, om høsten og våren er den omtrent lik, og om sommeren er den lavere.

De viste mønstrene med lavest el-bruk om natten og høyest om morgenen sammenfaller med aktivitetsnivået på flyplassen. Enkelte funksjoner ved Oslo lufthavn er døgnåpne hele året, men servicetilbudet varierer gjennom døgnet med mest aktivitet på dagtid. Om morgenen har de første flyene avgang rundt klokken seks, mens

²Ukedager vil si mandag til fredag, og helgedager er lørdag og søndag.

³Se oversikt over inkluderte helligdager i tabell A.1 i vedlegg A.



Figur 4.3: Gjennomsnittlig timebruk av elektrisitet [kWh/h] gjennom døgnet for a) uke-dager og helger, b) vanlige dager og helligdager og c) de fire sesongene i årene 2017 til 2019.

siste avgang er ved midnatt [59]. Som man ser i figur 4.3 a)-c) og figur 4.2 trekker flyplassen hele døgnet en grunnlast for å holde grunnleggende funksjoner i gang. I timene før første avgang øker elektrisitetsbruken gradvis, grunnet økt aktivitet i terminalbygningen, administrasjon og oppstart av tekniske anlegg. Aktiviteten flates ut utover dagen og avtar mot natten, samtidig med el-bruken. Den laveste el-bruken er midt på natten, halvveis mellom siste ankomst og første avgang.

Disse mønstrene har likhetstrekk med lastprofilen til den mellomstore flyplassen Seve-Ballesteros-Santander flyplass (SDR). Ifølge Alba og Mañana [60] har flyplasser en karakteristisk lastprofil som følger lignende mønster alle dager i året, noe lastprofilene i figur 4.3 viser at Oslo lufthavn har. SDR sin lastprofil viste en høy grunnlast. Det var en skarp topp i el-bruk i morgentimene, før det stabiliserte seg på et lavere nivå utover ettermiddagen. Ukedagsprofilen til OSL i figur 4.3 a) har også en tydelig topp om morgenen og et jevnere nivå utover ettermiddagen. El-bruken ved SDR nådde en ny topp om kvelden før stenging, som skyldtes økt behov for belysning ved mørkets frembrudd [60]. Fra døgnetprofilen for vinter- og høstmånedene til OSL i 4.3 c) ser man en lignende tendens til en økning i el-bruken rundt klokken

17 som kan skyldes både økt ettermiddagstrafikk og mer behov for belysning og varme.

4.2.3 El-bruk uke 18, 27 og 39 i 2019

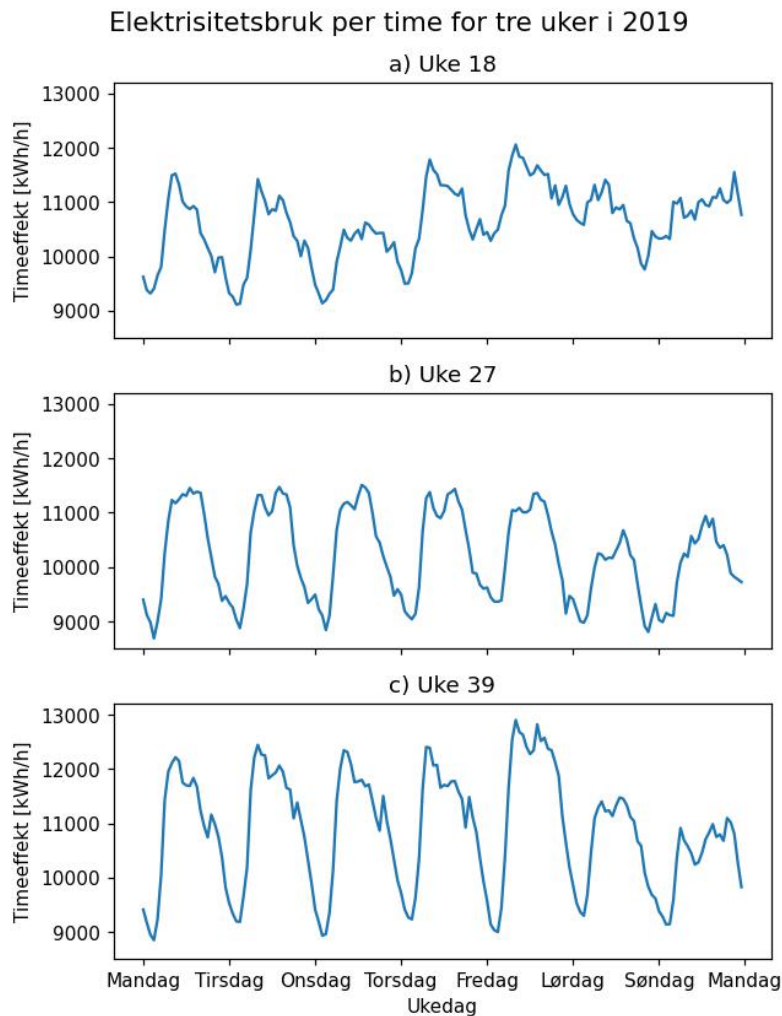
Delkapittel 4.2.2 viste hvilke mønster det er mulig å observere i el-bruken over tid. Videre blir det undersøkt hvordan disse mønstrene stemmer med el-bruken per time for uke 18, 27 og 39 i sommerhalvåret 2019, vist i figur 4.4. Disse ukene ble valgt for å vise variasjonen i el-bruk mellom uker i samme tidsperiode.

I figur 4.4 a) vises el-bruken for hver dag i uke 18 i 2019 (29.04-05.05.19). I likhet med den gjennomsnittlige ukedagsvariasjonen i figur 4.3 a) er de høyeste verdiene de fleste dager på morgenen, før el-bruken avtar utover dagen og har en lokal topp sent på kvelden. På helligdagen 1. mai og helgedagen søndag er de høyeste verdiene senere på dagen. Den høyeste timeeffekten på dagtid oppstår fredag morgen, mens den er lavest helligdagen onsdag 1. mai. I motsetning til øvrige uker var altså el-bruken lavere midt i uken enn i helgen. Verdiene på den vertikale akse begynner på 9 000 kWh/h, som var grunnlasten denne uken. Torsdag til lørdag er el-bruken over 1 000 kWh/h høyere på nattetid enn øvrige dager, og disse dagene skiller seg dermed fra øvrige uker i figuren og i figur 4.3 a).

Figur 4.4 b) viser bruk av elektrisitet i uke 27 (01.-07.07.19). Uke 27 har lignende nivå i el-bruken mandag til fredag og noe lavere i helgen. Den høyeste toppen er alle dager om ettermiddagen, men mandag, tirsdag og torsdag er den nesten like høy om morgenen. Sammenlignet med uke 18 er toppene og bunnene på samme nivå. Sammenlignet med figur 4.3 ligner døgnsprofilene på helgedagskurven i a), helligdagskurven i b) og sommerkurven i c).

Figur 4.4 c) viser elektrisitetsbruken gjennom uke 39 (23.-29.09.19). Elektrisitetsbruken er ganske lik gjennom ukedagene, med høyeste verdier om morgenen og former veldig lik mandag til tirsdag i uke 18 og ukedagskurven i 4.3 a). Den høyeste el-bruken er på fredag, etterfulgt av et lavere nivå i helgen som også har en form mer lik helgedagskurven i 4.3. Sammenlignet med uke 18 og 27 har elektrisitetsbruken i uke 39 en mer lik form fra dag til dag, høyere topper og omtrent like bunner.

Fra figur 4.4 ser man at el-bruken mange dager følger de mønstrene som man så ved å ta gjennomsnittlige timeverdier over lengre tidsperioder i figur 4.3. Det er likevel mange timer og dager som viker fra de generelle mønstrene, og som man kan trenge flere faktorer for å forklare. Datasettet med elektrisitetsdata er derfor utvidet med andre typer data, som blir presentert i neste delkapittel.



Figur 4.4: Elektrisitetsbruken per time [kWh/h] ved Oslo lufthavn i a) uke 18, b) uke 27 og c) uke 39 i 2019.

4.3 Forklaringsvariabler og øvrige data

For å kartlegge hvilke forklaringsvariabler som kan egne seg til prediksjon av el-bruk ved OSL ble forklaringsvariablene lufttemperatur, vindhastighet, nedbør, antallet passasjerer i flyplassterminalene og tidsvariabler testet. Tabell 4.1 viser variablene i datasettet, med tidsoppløsning og kilde. Siden dataene predikeres med timeoppløsning er time på døgnet tatt med som variabel. Basert på undersøkelsen av mønstre gjennom uken er i tillegg markører for ukenummer, ukedag, helligdag og helgedag lagt til. De øvrige variablene er valgt basert på tilgjengelighet og anvendelse på lignende problemstillinger. I dette delkapittelet presenteres variabler som brukes i lastprediksjonslitteratur, etterfulgt av en presentasjon av passasjer- og værdata for Oslo lufthavn Gardermoen.

Tabell 4.1: Forklaringsvariablene i datasettet.

Navn	Tidsoppløsning	Kilde
Time på døgnet	Per time	
Terminalpassasjerer	Sum per uke	Avinor AS [48]
Lufttemperatur [°C]	Gjennomsnitt per time	MET [47]
Vindhastighet [m/s]	Gjennomsnitt per time	MET [47]
Nedbør [mm]	Sum per døgn	MET [47]
Tidligere energibruk [kWh/h]	Per time	Avinor AS [49]
Helgedag (ja/nei)	Per time	
Helligdag (ja/nei)	Per time	Helligdagsfredloven [61]
Ukedag	Per time	
Ukenummer	Per time	

4.3.1 Forklaringsvariabler og lignende prediksjoner av el-bruk

Til prediksjon av el-bruk kan både eksterne og interne parametre brukes, fra tidligere el-bruk til utendørstemperatur. I delkapittelet presenteres et utvalg litteratur innenfor prediksjon av energi- og elektrisitetsbruk med maskinlæring, med fokus på hvilke variabler som er brukt.

I en sammenligningsstudie av maskinlæring til prediksjon av energibruk og energi-effektivitet i bygninger hadde Seyedzadeh mfl. [2] en oversikt over maskinlærings-teknikker og forklaringsvariabler anvendt til formålene. For å predikere elektrisitetsetterspørsel var kunstige nevralt nettverk brukt i fem av seks nevnte artikler [2].

Platon mfl. [62] predikerte timeverdier for el-bruk med en til seks timer tidshorisont ved en institusjonell bygning i Canada med kunstige nevralt nettverk. De brukte 15 måneders målinger av utendørs og innendørs lufttemperatur, utendørs relativ luftfuktighet, og temperaturen og strømmingen av luft, kaldt vann og varmt vann til bygningen [62]. Mihalakakou mfl. [63] brukte lufttemperatur og total solinnstråling for å predikere timeverdier for el-bruk i en enebolig i Hellas med kunstige nevralt nettverk [63]. Li mfl. [64] brukte både *Support Vector Machines* (SVM) og kunstige nevralt nettverk for å predikere el-bruken i 59 bolighus i Kina på årsbasis ved hjelp av en rekke bygningstekniske egenskaper, blant annet veggtykkelse og varmeoverføring mellom ulike deler av bygningen [64]. Massana mfl. [65] predikerte timeverdier for el-bruken i en kontorbygning ved et universitet, med både SVM og kunstige nevralt nettverk. De brukte innendørs- og utendørstemperaturer og luftfuktighet, innendørs belysning, solinnstråling og kalendervariabler [65]. Srivastav mfl. [66] brukte en annen modelltype, *Gaussian Mixture Models*, til å predikere el-bruk per år for en modell av en dagligvarebutikk ved hjelp av utendørstemperatur og luftfuktighet, samt solinnstråling [66].

Halvparten av artiklene som ble nevnt av Seyedzadeh mfl. [2] til å predikere elektrisitetsbruk hadde timeoppløsning på verdien de ville predikere. Over halvparten

brukte bygningsinformasjon og målinger av inneklimate som forklaringsvariabler. Flere brukte kombinasjoner av kunstige nevralt nettverk og andre algoritmer, eller brukte *Principal Component Analysis* (PCA) for variabelseleksjon. Utendørstemperatur, luftfuktighet og solinnstråling ble brukt i henholdsvis 83 %, 60 % og 50 % av artiklene [2].

Ytterligere tre artikler som ble nevnt av Seyedzadeh mfl. [2] så på prediksjon av total energibruk på timesbasis, ikke spesifikt elektrisitet. Av disse brukte to kunstige nevralt nettverk [2]. Blant dem ble temperaturer, solinnstråling, vindhastighet og kalendervariabler gjentatt som forklaringsvariabler, men også bygningskarakteristikker ble brukt. Karatsou mfl. [67] brukte et kunstig nevralt nettverk til å predikere energibruk per time for en kontorbygning i Athen. De brukte målinger av temperatur, solinnstråling, spesifikk fuktighet, vindhastighet og dag som variabler [67]. Jain mfl. [68] predikerte også energibruk per time ved hjelp av temperatur, luftfuktighet og vindhastighet, men for en boligbygning og med SVM [68].

Seyedzadeh mfl. [2] sammenstilte også 18 artikler som predikerte laster til mindre komponenter i bygninger: systemer for oppvarming, kjøling og klimaanlegg. Også her brukte over halvparten av artiklene kunstige nevralt nettverk, blant annet Paudel mfl. [69] som predikerte effektbehovet til oppvarmingssystemene i en institusjonell bygning i Frankrike. De brukte klima- og oppvarmingsdata, men også en profil som representerte tilstedeværelsen i bygningen [69]. Zhao og Magoulès [70] og Burkhart mfl. [71] brukte også målinger som representerte antallet mennesker i bygningen eller beboere i sine prediksjonsmodeller av energibehov til oppvarming eller klimakontroll [70] [71].

Hwang mfl. [7] lagde prognoser for totalt elektrisitetsforbruk i næringsbygg på daglig og månedlig oppløsning med ulike maskinlæringsteknikker, blant annet LSTM-modeller. I prediksjonen benyttet de en rekke meteorologiske parametere, deriblant temperatur, nedbør, vindhastighet, atmosfærisk trykk, relativ luftfuktighet og solinnstråling, i tillegg til flere bygningskarakteristikker. For dagoppløsning brukte de også ukedag som parameter [7].

De meteorologiske variablene som ble inkludert i modellene i denne oppgaven er lufttemperatur, vindhastighet og nedbør. Dette er variabler som blant annet Mihalakakou mfl. [63], Karatsou mfl. [67] og Jain mfl. [68] brukte til prediksjon av laster i bygninger. Luftfuktighet med timesoppløsning var ikke tilgjengelig fra Gardermoen målestasjon for det aktuelle tidsrommet, heller ikke solinnstråling. Nedbør var heller ikke tilgjengelig i timesoppløsning, men ble inkludert med døgnoppløsning.

Ingen av artiklene som ble nevnt av Seyedzadeh mfl. [2] predikerte energibruk eller elektrisitetsbruk med tidligere målinger som variabel [2]. Det gjorde Slowik og Urban [42] og Bourhane mfl. [72], som predikerte elektrisitetsbruk per time kun basert på tidligere el-bruk [42] [72].

Slowik og Urban [42] bygde en enkel LSTM-modell for å predikere energibruk i et industrianlegg på et mikrogrid. De hadde 0-4 timer tidshorisont, og predikerte kun basert på historisk el-bruk med målinger hvert sekund. De bygde, i likhet med i denne oppgaven, modellene ved hjelp av Keras og brukte 80 % treningsdata. Modellen

hadde ett LSTM-lag og ett *Dense*-lag. Med én time prediksjonshorisont fikk ett-lagsmodellen bedre MAE og MSE enn en LSTM-modell med to skjulte lag [42].

Hwang mfl. [43] presenterte en algoritme for planlegging av opp- og utladning av et energilagringssystem med mål om å redusere toppplastene til en bygning. Som en del av algoritmen lagde de lastprognoser basert på en modell fra Yudiantaka mfl. [44]. Lastprognosene bygde på temperatur- og effektmålinger og ble gjort med en kombinasjon av LSTM- og Multi-Layer Perceptron-algortimene, begge med tre lag. De foreslo å utvide maskinlæringsmodellen med for eksempel luftfuktighet [43].

Det kan skje at datasettet som blir brukt til trening av en prediksjonsmodell inneholder informasjon det ikke er forventet å ha, for eksempel variabelen man prøver å predikere. Da kan man få tilsynelatende gode resultater som viser seg å ikke være reproducerbare, eller at man overvurderer modellen mens den i virkeligheten bare hadde tilgang på mer informasjon enn ventet. Dette kalles informasjonslekkasje og er en stor utfordring i maskinlæring [73]. For å kunne inkludere tidligere verdier for variabelen som skal predikeres må man derfor passe på at verdiene er forskjøvet, så inneværende timeverdi ikke brukes som forklaringsvariabel for å predikere seg selv. Fordi tidligere elektrisitets- eller effektbruk ble brukt for å predikere fremtidig bruk av blant andre Slowik og Urban [42], Hwang mfl. [43] og Bourhnane mfl. [72] ble tidligere el-bruk vurdert som en aktuell forklaringsvariabel til prediksjonen i denne oppgaven. Det ble først forsøkt å predikere uten.

Alba og Mañana skrev at en intern faktor som påvirker el-bruken i en flyplass er bevegelsene i terminalbygningene, mens utendørstemperatur og dagslys er viktige eksterne påvirkere [60]. Elektrisitetsbehovet til Oslo lufthavn er høyere om vinteren enn om sommeren, noe som indikerer at oppvarming og klimakontroll er to av lastene som trekker mye elektrisitet også der. Siden Zhao og Magoulès [70] og Burkhart mfl. [71] brukte indikatorer på aktivitetsnivå i sine prediksjoner var det ønskelig å inkludere dette for modellene som presenteres i denne oppgaven. En tilgjengelig indikator for aktivitetsnivået på Oslo lufthavn var antallet terminalpassasjerer per uke, og dette ble derfor tatt med som variabel. Antallet passasjerer på Oslo lufthavn presenteres i neste delkapittel.

4.3.2 Terminalpassasjerer

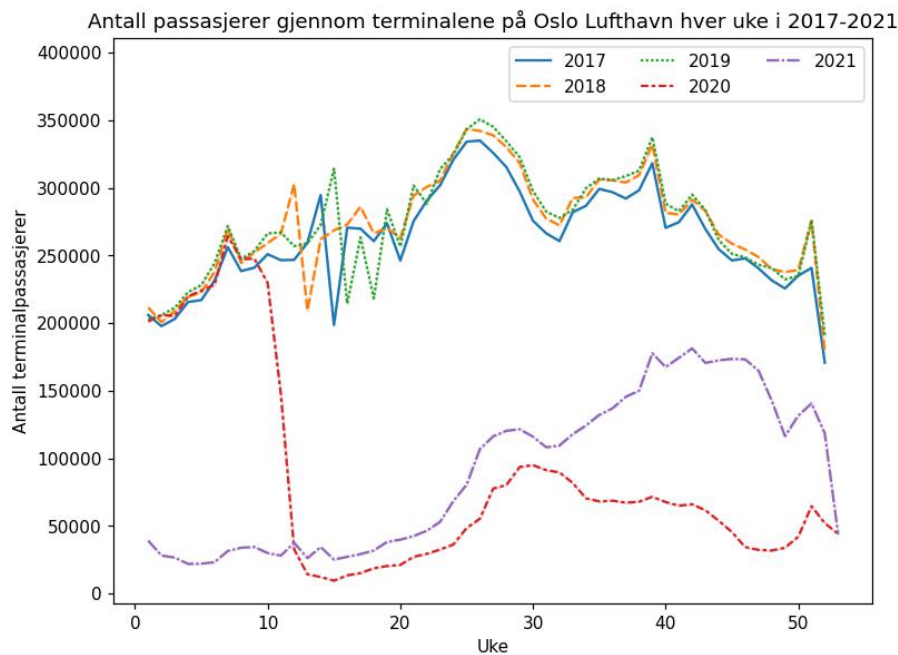
Antallet passasjerer som passerer gjennom Oslo lufthavn sine terminaler per uke, kalt *terminalpassasjerer*, ble hentet fra Avinor for årene 2017 til 2021 til bruk i oppgaven. Dette er tall på totalt avreiste passasjerer per uke som har ferdes i terminalene, både de som passerer sikkerhetskontrollen og transferpassasjerer som ikke gjør det [48]. Passasjerantallet ble inkludert for å ha et mål på aktivitetsnivået på flyplassen. For at hver time i en spesifikk uke skulle ha en verdi for denne variabelen, ble passasjerantallet per uke gjentatt for hver time i uken.

Det er et gjentakende mønster fra år til år i antallet terminalpassasjerer, som man kan se i figur 4.5. De laveste passasjertallene er alle år tidlig på året. Det bratte fallet i antallet passasjerer i årets siste uke skyldes at denne uken ikke alltid har sju dager. Uke 7 har høyere passasjerantall enn omliggende uker i 2017 til 2020. Dette

er uken før vinterferien i Oslo kommune og tidligere Akershus fylke [74]. Det høyeste passasjerantallet i årene 2017 til 2019 var i uke 25 eller 26, som er rett før fellesferien i uke 28 til 30 [75]. Deretter faller passasjerantallet disse årene før det gradvis bygger opp til en ny topp i uke 39, uken før høstferien i Oslo og tidligere Akershus fylke [74]. Det er en topp i uke 51 alle år, spesielt i 2018 og 2019.

Enkelte våruker er ikke like fra år til år i perioden 2017 til 2019. Disse ukene sammenfaller med påskeuken som er ulik uke fra år til år. I 2017 hadde uke 14 et høyt passasjerantall mens uke 15 hadde få passasjerer. Dette er henholdsvis uken før påskeuken og selve påskeuken, som var uke 15 i 2017. I 2018 hadde uke 12 et høyt antall passasjerer, mens flyplassen i uke 13 hadde få passasjerer. Dette året var påsken i uke 13. I 2019 hadde uke 15 et høyt antall passasjerer, mens uke 16 og 18 hadde et lavt antall sammenlignet med tidligere år. Dette året var uke 16 påskeuken og 1. mai var i uke 18. De øvrige årene falt ikke passasjerantallet like mye uken med 1. mai som i 2019.

Fram til uke 10 fulgte også passasjerantallet i 2020 samme mønster som i 2017 til 2019, men så sank passasjertallet kraftig og divergerte deretter fra 2017 til 2019-normalen ut året 2021. Dette sammenfaller med perioden Norge var delvis underlagt restriksjoner i forbindelse med COVID-19-pandemien [58]. Et lignende fall så man også i figur 4.2 som viste elektrisitetsbruken for årene 2017-2021. Dette bekrefter valget om å fjerne årene 2020 og 2021 fra videre arbeid.



Figur 4.5: Sum av antall terminalpassasjerer på Oslo lufthavn for hver uke i 2017 til 2021.

I tillegg til passasjerantall ble flere meteorologiske målinger inkludert som forklaringsvariabler til prediksjonen av el-bruk. Disse variablene blir presentert i neste delkapittel.

4.3.3 Meteorologiske målinger

Til prediksjon av el-bruk ved Oslo lufthavn ble det brukt nedbørs-, vind- og lufttemperaturdata fra værstasjonen Gardermoen målestasjon. Dette er en av flere meteorologiske stasjoner i tilknytning til flyplassen og drives av Meteorologisk institutt (MET). Målestasjonen ligger nord-vest på området og har id 4780 [76]. Værdata fra denne stasjonen ble hentet ut gjennom MET sin portal for nedlasting av værdata, Frost API [47], for tidsperioden 01.01.2017-31.12.2021.

Elektrisitetmålingene fra Oslo lufthavn har timesoppløsning, mens de meteorologiske dataene var tilgjengelig med ulike tidsoppløsninger. Noen timer manglet ved nedlasting en eller flere av målingene. Temperatur og vindhastighet ble lastet ned med tre målinger per time, og for å få samme tidsoppløsning som el-bruk ble det beregnet et gjennomsnitt for hver time. Nedbør hadde én måling per døgn: summen av nedbør gjennom døgnet, avlest én gang per døgn. For å få en verdi for hver time i døgnet fikk alle timene samme døgn nedbørsummen som verdi. Dette ble gjort tilsvarende som for passasjerantall per uke. For eksempel ble det i løpet av døgnet 24 timer 09.06 2019 registrert 7,4 mm nedbør ved Gardermoen målestasjon. Alle timer dette døgnet ble da oppført i datasettet med 7,4 mm nedbør.

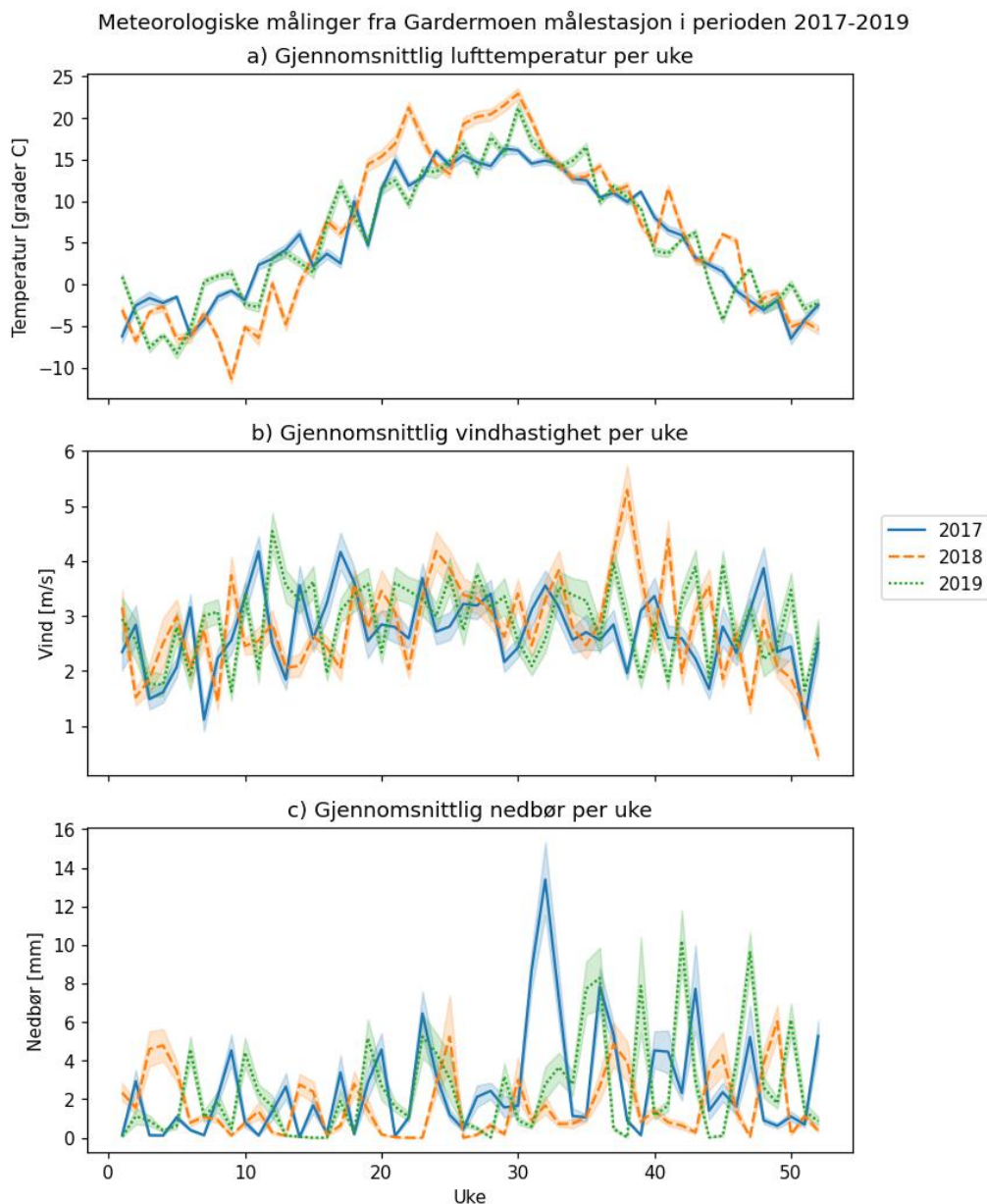
De meteorologiske dataene var logget i koordinert universaltid, og ble konvertert til norsk tid inkludert sommer- og vintertid før de ble koblet sammen med øvrige data. Det gjorde at noen timer manglet målinger av vindhastighet, nedbør og lufttemperatur, hovedsakelig ved klokkestilling vår og høst. Alle ufullstendige timene ble fjernet fra datasettet, noe som reduserte datasettet med 0,35 %.

Figur 4.6 viser for årene 2017 til 2019 gjennomsnittlige verdier per uke for

- a) Lufttemperatur [$^{\circ}\text{C}$]
- b) Vindhastighet [m/s]
- c) Nedbør [mm]

Et skravert belte rundt hver linje representerer variasjonen innad i hver uke. Grunnet opprinnelig oppløsning av dataene representerer kurvene for lufttemperatur og vindhastighet gjennomsnittlig timeverdi for de respektive målingene for hver uke. For nedbør representerer kurven gjennomsnittlig nedbør for hvert døgn den uken.

Den gjennomsnittlige lufttemperaturen ved Oslo lufthavn, vist i figur 4.6 a), viser en tydelig forskjell mellom sesongene. De høyeste temperaturene gjennom året er i juli på mellom 20°C og 30°C , og de laveste i desember til februar på ned mot -10°C . Det er en lignende variasjon gjennom året for alle tre årene, men for enkeltuker kan det være stor forskjell mellom årene og mellom påfølgende uker. Dette kan påvirke el-bruken. Det ser ut til å være minst forskjell mellom de ulike årene om høsten og tidlig vinter. I uke 25 er gjennomsnittstemperaturen rundt 14°C hvert år. I uke 9 var derimot temperaturen i 2019 2°C mens den i 2018 var -11°C . Perioden fra rundt uke 9 til uke 15 var for eksempel temperaturene lavere i 2018 enn andre år. Dette er en periode hvor man på figur 4.2 så at el-bruken i 2018 var høyere enn andre år.



Figur 4.6: Målinger av gjennomsnittlig a) lufttemperatur, b) vindhastighet og c) nedbør per uke i 2017 til 2019 fra Gardermoen målestasjon ved OSL. Det skraverte beltet rundt hver linje representerer variasjonen innad i hver uke.

Vindhastigheten ved målestasjonen, vist i figur 4.6 b), har en svak variasjon gjennom året, men den er mindre tydelig enn for lufttemperaturen. Vindhastighetene er lavere om vinteren enn om sommeren. Det er en større forskjell fra uke til uke om våren og vinteren enn om sommeren. Om sommeren er det også minst forskjell mellom årene 2017, 2018 og 2019. De høyeste gjennomsnittsvindhastighetene er opp mot 6 m/s, mens de laveste er under 0 m/s.

Som man ser i figur 4.6 c) varierer også nedbør fra uke til uke og mellom årene, men det er tydelig mer nedbør på sensommeren og høsten enn ellers i året. Den høyeste nedbøren i perioden kom i uke 32 i 2017, hvor gjennomsnittlig nedbør var nesten

14 mm. Høsten i 2019 var det også enkelte uker med gjennomsnittlig nedbør over 10 mm. Om våren har de fleste uker gjennomsnittlig nedbør på under 4 mm. Det finnes uker med nesten null gjennomsnittlig nedbør året gjennom.

Figur 4.6 a) viste at lufttemperatur har en periodisk variasjon som gjentas hvert år. Det samme gjelder i mindre grad for vindhastighet og nedbør. Man så i figur 4.2 at den gjennomsnittlige el-bruken per døgn for Oslo lufthavn også har sesongvariasjon som er motsatt som for lufttemperatur. Det er derfor interessant å se om det er en lignende sammenheng over kortere tid. I neste delkapittel studeres el-bruk for hver time en uke i 2019 sammen med målinger for lufttemperatur, nedbør og vindhastighet.

4.3.4 El-bruk og meteorologiske målinger for uke 18 i 2019

For å undersøke om det er en korrelasjon mellom el-bruk og meteorologiske variabler på timebasis blir timeeffekt for Oslo lufthavn sammenlignet med alle de tre værvarene for uke 18 i 2019. El-bruken i uke 18 ble undersøkt i figur 4.4 a). Figur 4.7 viser timeverdier for el-bruk sammen med

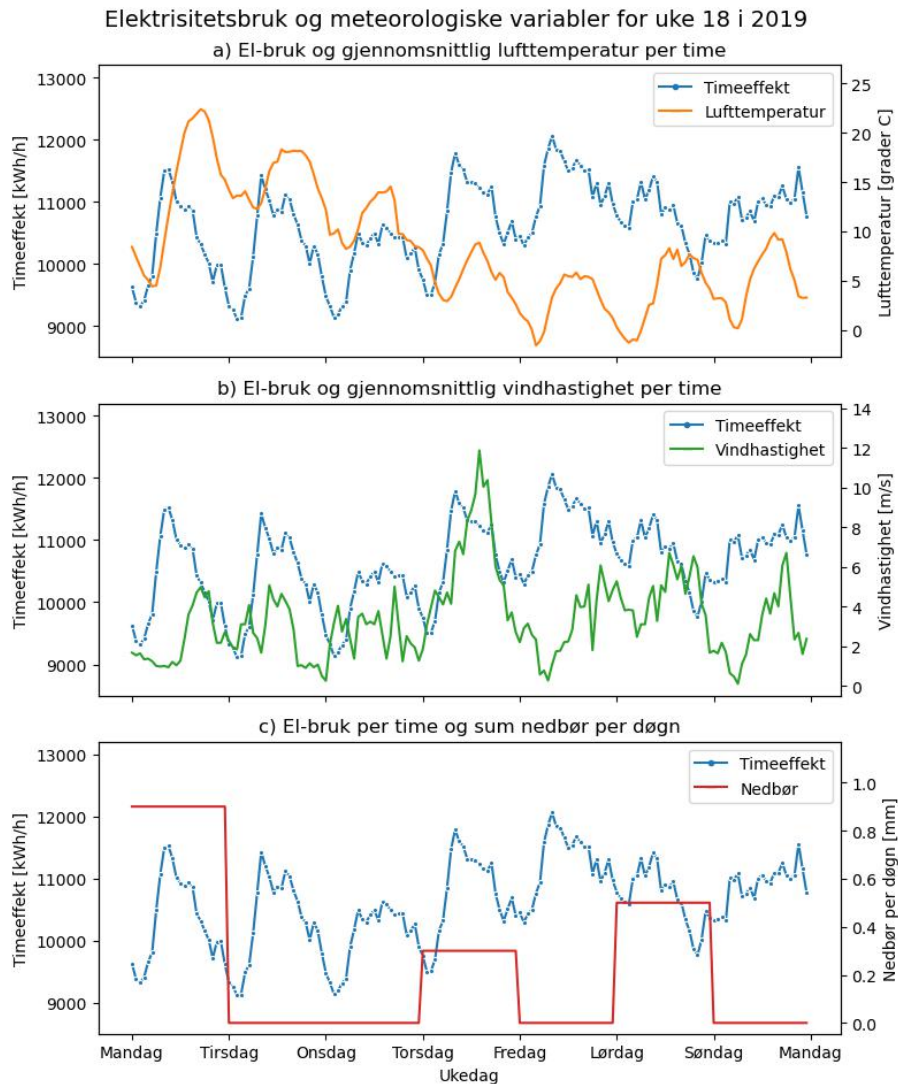
- a) Timeverdier for lufttemperatur [$^{\circ}\text{C}$]
- b) Timeverdier for vindhastighet [m/s]
- c) Døgnverdier for nedbør [mm]

Fra figur 4.7 a) ser man at det er 20°C forskjell mellom høyeste og laveste lufttemperatur denne uken. Lufttemperaturene er høyest tidlig i uken og minker mot helgen. Elektrisitetsbruken har lignende topper gjennom uken, utenom onsdag. Timeeffekten når ukens høyeste nivå på $12\,000\text{ kWh/h}$ fredag. Samme dag er lufttemperaturene på sitt laveste, med en topp på bare 5°C . Dette kan ha en sammenheng. Samtidig er el-bruken nesten like høy mandag, og da har temperaturen en topp på over 20°C . Lufttemperatur har en tydelig døgnvariasjon med høyeste verdier midt på dagen. Det så man i delkapittel 4.2 at også el-bruk har. De fleste dager er det rundt 5°C forskjell mellom døgnets høyeste og laveste lufttemperatur.

Vindhastigheten i figur 4.7 b) har også sine høyeste verdier på dagtid, men døgnprofilen er noe mindre markant enn hos lufttemperatur og el-bruk. Vindhastigheten er lavest tidlig i uken og høyest midt i uken, men dette synes ikke å ha en tydelig innvirkning på el-bruken. Forskjellen mellom døgnets kraftigste og svakeste vind varierer mellom 4 og 10 m/s .

Siden nedbøren i figur 4.7 c) kun har én måling per døgn er det ikke mulig å se døgnvariasjonene i nedbør på figuren. Det gjør det vanskelig å se en korrelasjon med el-bruk. Nedbøren er lav hele uken, med høyeste verdi på $0,8\text{ mm/døgn}$ på mandag. Det er heller ikke tydelig høyere eller lavere el-bruk de dagene det ikke regner.

Tilsvarende figur som 4.7 for uke 27 og 39 finnes i vedlegg A. I likhet med uke 18 har temperaturen og vindhastigheten i de ukene også en gjentakende døgnvariasjon,



Figur 4.7: Elektrisitetsbruken [kWh/h] i uke 18 i 2019 sammen med a) lufttemperatur, b) vindhastighet og c) nedbør.

og den er mindre tydelig enn for el-bruken. For nedbør er det ikke mulig å se en innvirkning av nedbørsmengden på el-bruken.

Basert på tilgjengeligheten til de meteorologiske dataene og anvendelsene av lignende variabler i forskning, som vist i delkapittel 4.3.1, ble lufttemperatur, vindhastighet og nedbørsmengde brukt som forklaringsvariabler i denne oppgaven.

4.4 Tidsperiode for prediksjon: sommerhalvåret 2019

El-bruk og relevante variabler har blitt presentert for årene 2017 til 2021, men siden dette er en innledende studie ble det besluttet å trene og teste modellene på data for en kortere tidsperiode. Dette ble gjort for å se hvor godt modellene lærte

sammenhenger mellom data med mindre ulikheter. Dersom man oppnår gode resultater, kan man i neste steg utvide datasettet med flere måneder eller år. Den valgte tidsperioden var sommerhalvåret 2019, bestående av månedene april til september med til sammen 4392 timer.

Basert på de sammenfallende grafene for elektrisitetsbruk i figur 4.2 ble året 2019 antatt å være representativt for normaltstanden i elektrisitetsbruken til flyplassen. Grunnet fallet i elektrisitetsbruk og passasjerantall etter våren 2020 ble årene 2020 og 2021 utelukket fra analysen.

Figur 4.2 viste at OSL har en tydelig forskjell i el-bruken sommer og vinter, med lavest el-bruk og minst variasjon fra dag til dag om sommeren. Som vist i figur 4.6 har gjennomsnittsverdiene per uke for de meteorologiske målingene større variasjon mellom årstidene enn innad i årstidene. Vindhastigheten har mindre variasjon mellom påfølgende uker i sommerhalvåret enn ellers på året. Lufttemperaturen varierer med 10 til 15 °C mellom påfølgende uker, men har over 35 °C forskjell mellom høyeste og laveste verdi gjennom året. Målingene innad i sommerhalvåret har altså mindre variasjon enn dersom flere år blir inkludert. Derfor ble dette halvåret valgt.

Alle modellkonfigurasjonene som ble beskrevet i kapittel 3.4 ble trent, validert og testet på et datasett for sommerhalvåret 2019. I neste kapittel presenteres og diskuteres resultatene fra prediksjonen av el-bruk ved Oslo lufthavn.

Kapittel 5

Resultat og diskusjon

I denne oppgaven blir det undersøkt om LSTM-modeller kan predikere timeeffekt¹ for Oslo lufthavn Gardermoen (OSL) én og 24 timer i forveien. I dette kapitlet beskrives og drøftes prediksjonene fra maskinlæringsmodellene. Først oppsummeres mønstrene i el-bruk som ble funnet i kapittel 4. Metodene for evaluering av modellene og innvirkningen av tilfeldighet på resultatene blir diskutert før modellene presenteres. Resultatene fra prediksjonene én time frem i tid fra den enkle modellen og basismodellen som ble beskrevet i kapittel 3.3 presenteres først, etterfulgt av en oversikt over de elleve beste modellkonfigurasjonene og utvalgte prediksjonsplott fra denne topplisten. Så blir resultatene fra prediksjonene av el-bruk 24 timer frem i tid presentert. Resultatene blir diskutert underveis, og så følger en generell diskusjon.

5.1 Mønstre i el-bruken og korrelasjonen med forklaringsvariablene

Undersøkelsene av variasjonene i elektrisitetsbruken ved Oslo lufthavn i kapittel 4, gjennom døgnet og for ulike uker, viser at det er tydelige mønstre i el-bruken. Det ble også vist at det over lengre tid syntes å være en kobling mellom elektrisitetsbruken og de andre variablene som ble presentert.

Mønstrene i el-bruken gjennom uken i figur 4.3 tyder på at det er en forskjell mellom ukedag og helg som kan være nyttig for modellen. Det samme gjelder mellom årstider og mellom helligdag og ukedag. Dette så man også i figur 4.4 som viste timeverdi for el-bruk for henholdsvis uke 18, 27 og 39 i 2019.

Det så også ut til at antallet terminalpassasjerer og lufttemperaturen hadde sammenhenger som var inverse av sammenhengene i elektrisitetsbruken. Fremstillingen av meteorologiske målinger fra Gardermoen målestasjon for 2017 til 2019 i figur 4.6 viste at temperaturen var høyest om sommeren, når elektrisitetsbruken var lavest.

¹Aggregerte timeverdier for brukt elektrisitet [kWh/h] for hele OSL, unntatt elektrisiteten som brukes til oppvarming gjennom el-kjel og varmepumper i varmesentralen.

Plottet av passasjerantall hver uke i 2017 til 2019 i figur 4.5 viste at også passasjerantallet var høyest om vinteren.

Samtidig viste plottet av el-bruk sammen med meteorologiske variabler gjennom uken i figur 4.7 at det var vanskelig å se sammenhenger i mønstrene ved høyere oppløsning i tid.

Barplottene i figur 5.1 viser Pearson-korrelasjonen [50] mellom elektrisitetsbruken ved flyplassen og øvrige numeriske variabler i datasettet for tidsperiodene a) 2017-2019 og b) sommerhalvåret 2019.

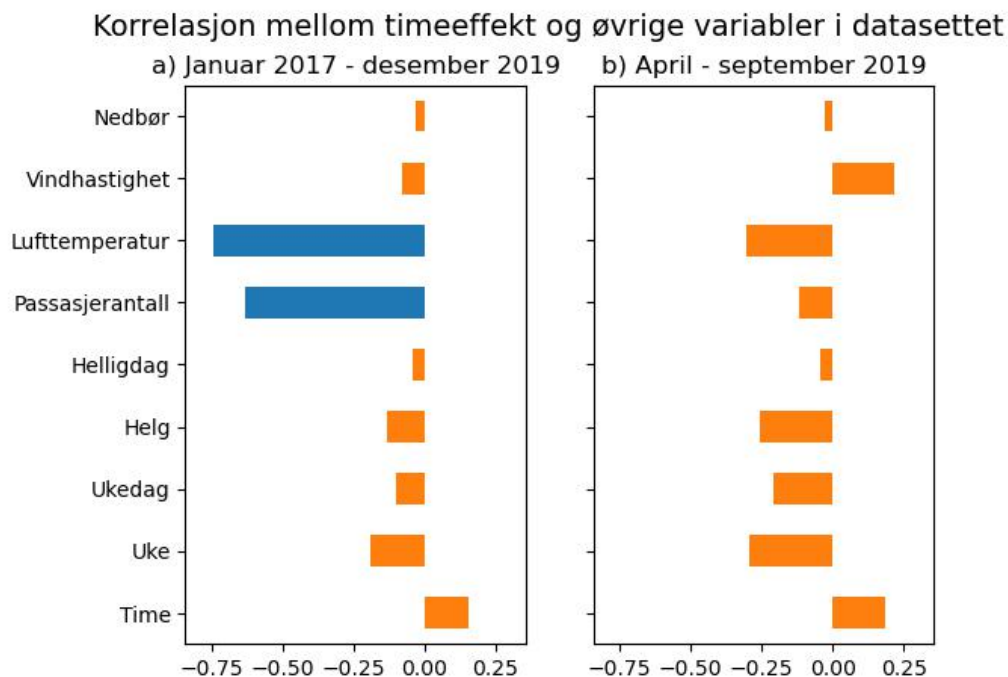
For perioden januar 2017 til desember 2019 viser figur 5.1 a) at det er en negativ Pearson-korrelasjon på over 0,60 mellom henholdsvis lufttemperatur og passasjerantall med el-bruk. Øvrige variabler har en korrelasjon på rundt null med el-bruken.

Dette bekrefter at det over lengre tid er en tydelig negativ kobling mellom lufttemperatur og el-bruk. Det meste av denne koblingen er trolig sesongbasert. Lufttemperaturen er naturligvis høyere om sommeren enn om vinteren i Norge. Samtidig fører høyere temperaturer til at mindre elektrisitet må brukes til for eksempel oppvarming. Dersom man hadde fjernet sesongvariasjonene hadde man ikke nødvendigvis sett den samme koblingen.

En negativ kobling mellom passasjerantall og el-bruk er mer overraskende. Folk reiser mer om sommeren, noe som kunne få en til å tro at el-bruken må øke om sommeren for å dekke økt aktivitet på flyplassen. Dette synes ikke å være tilfellet.

Målinger av elektrisitet som brukes til varmpumper og el-kjel i varmesentralen ved Oslo lufthavn ble ikke inkludert i verdiene for el-bruk som ble presentert. El-bruken er likevel høyest om vinteren, når passasjerantallet er lavest. Det indikerer at andre laster også bruker mer elektrisitet om vinteren, uavhengig av passasjerantall. Dette kan for eksempel være varmekabler, tekniske anlegg, ventilasjonssystemer eller lignende; laster som kan egne seg til fleksibel bruk [20].

I sommerhalvåret 2019 er det lav Pearson-korrelasjon mellom variablene i datasettet og el-bruken, som vist i figur 5.1 b) ved at ingen variabler har høyere korrelasjon med el-bruk enn +/- 0,25. Her er det lufttemperatur, helg og uke som har høyest korrelasjon med el-bruk. De to sistnevnte har litt høyere korrelasjon enn i 5.1 a). Den lave korrelasjonen bekrefter at sammenhengene mellom mønstrene er lavere ved høyere oppløsning i tid.



Figur 5.1: Pearson-korrelasjonen mellom numeriske variabler i datasettet og el-bruk for tidsperiodene a) 2017 til 2019 og b) sommerhalvåret 2019.

Pearson-korrelasjonen er et mål på lineær assosiasjon mellom variabler [50]. At denne er lav betyr at lineære regresjonsmodeller vil ha vanskeligheter med å modellere sammenhengene mellom variablene. De nevralt nettverkene som blir brukt i denne oppgaven kan modellere både lineære og ikke-lineære sammenhenger, og kan derfor gi gode prediksjoner selv om den lineære assosiasjonen er lav.

El-bruken i sommerhalvåret 2019 ble antatt å være representativ for normaltstanden i el-bruken til flyplassen. Fra år til år følger el-bruken i 2017 til 2019 den samme periodiske trenden. El-bruken er likevel ikke identisk for samme dag fra år til år. Det er blant annet en topp i el-bruken i april 2019 som skiller seg fra de to andre årene. Denne toppen er inkludert i treningssettet til prediksjonen og kan utgjøre en ekstra utfordring for modellen. Samtidig kan det i fremtidige trenings- og testsett finnes lignende avvik fra normalen. Behovet for å trene og teste modellene på andre testsett for å bekrefte generaliserbarheten til modellkonfigurasjonene blir diskutert i neste delkapittel om scorer og evaluering.

I de neste delkapitlene presenteres resultatene fra maskinlæringsmodellene i oppgaven. Først blir metodene for evaluering og sammenligning av modeller i oppgaven diskutert. Deretter blir prediksjoner av el-bruk én time frem i tid ved hjelp av basismodellen og andre utvalgte modellkonfigurasjoner presentert, etterfulgt av prediksjoner et døgn frem i tid.

5.2 Evaluering og sammenligning av modeller

Til evaluering av modellene ble både fire ulike evalueringsmål og vurdering av visuell måloppnåelse brukt. Det ble i tillegg gjort tiltak for å begrense innvirkningen av tilfeldighet på prediksjonsresultatene. Dette beskrives i dette delkapittelet.

5.2.1 Scorer og evalueringsmål

Evalueringsmålene MSE, RMSE, MAPE og MAE som ble beskrevet i delkapittel 2.2.8 ble brukt til å evaluere modellene som presenteres i denne oppgaven. Scorene fra prediksjonene kan gi et feilaktig bilde av modellenes anvendelse når det kommer til det praktiske problemet, som er å nøyaktig representere el-bruken til flyplassen. Dette vil bli illustrert med prediksjonsplottet fra modell XI i delkapittel 5.3.5. Det kan være hensiktsmessig å vurdere både score og visuelle prediksjoner sammen.

Samtidig kan det å legge for mye vekt på de visuelle resultatene gjøre de endelige modellene påvirket av testsettet. I prosessen med å trene hver modellkonfigurasjon ble et valideringssett brukt for å vurdere modellenes ytelse underveis, men i prosessen med å avgjøre hvilke modeller som overordnet gjorde det best ble alle modellene evaluert på hvor godt de estimerte el-bruken i testsettet. Modellutvelgelsen kan dermed sees på som en del av treningsprosessen. Ved å velge modell basert på egnetheten til testsettet på kun én måned kan det ha blitt valgt modeller som representerte sammenhenger som gir gode resultater på nettopp dette testsettet, men ikke nødvendigvis på andre testsett. Dette kalles testsett data-lekkasje [29]. For å minimere konsekvensene av dette vil det i neste steg være hensiktsmessig å utvide datasettet med flere verdier i både trenings- og testsettene, for eksempel ved å trene og teste de beste modellene på andre tidsperioder.

Evalueringsmålene ble valgt basert på at de er mye brukt i forskning på maskinlæring. Blant annet nevnte Abbass og Hamdy samtlige som vanlige evalueringsmetoder for prediksjonsproblemer [30]. Hwang mfl. brukte MAE som tapsfunksjon [43], Slowik og Urban brukte MAE og MSE til evaluering av modeller [42]. mens både Hwang mfl. og Haq mfl. brukte MAPE [7] [34] og Petneházi brukte blant annet MAE og RMSE [77].

Som nevnt i delkapittel 2.2.8 vektlegger MSE og RMSE store avvik mer enn små. En modell med enkelte svært høye avvik vil ha høyere RMSE enn en modell med mange, men små avvik. MAE vektlegger derimot alle avvik likt. I treningsprosessen ble MSE brukt som tapsfunksjon og i utvelgelse av beste modell ble RMSE brukt som score. Fordi RMSE ble brukt til rangering av modellene, er diskusjonen i kommende delkapitler også hovedsakelig basert på modellenes RMSE-score. Det kan ha ført til mer konservative modeller, som prefererer å undervurdere heller enn å overvurdere el-bruken.

I delkapittel 5.3.3 presenteres de elleve beste modellkonfigurasjonene til prediksjon én time frem i tid, rangert etter RMSE på testsettet. Dersom man rangerer modell-

konfigurasjonene basert på RMSE endres topplisten noe. Det er mulig at å bruke MAE som tapsfunksjon ville endret prediksjonene fra modellene.

Det kan diskuteres hvorvidt de valgte evalueringsmålene er meningsfulle når det kommer til målsetningen for modellene. Hva man ønsker at modellene skal gjøre kan variere fra situasjon til situasjon. Tidsrommet modellene predikerer er også relevant. De første timene kan være viktigere enn de øvrige timene. Scorene sier heller ingenting om variasjonen og stabiliteten til prediksjonene, men dette kan sees i plottene.

Til lastflytting vil man for eksempel at modellen predikerer døgnetts høyeste verdier bedre enn øvrige verdier, samtidig som man kan ønske et tydelig bilde av bunnene for å vite når lastene kan flyttes til. Dersom man ønsker å beregne når energilagringssystemer bør lades opp kan det også være nyttig å vite når el-bruken er lav.

For å beregne potensialet for lastklipping eller tidspunktene man bør lade ut energilagringssystemer for å redusere toppene i effektbruk, er verdiene og tidspunktene for de høyeste timeeffektene mest relevante. Hvis man bruker modellene til å estimere hvilke timer effekttoppene bør reduseres, og dermed potensialet til effekttopp-reduksjon ved hjelp av energilagringssystemer, kan et høyt avvik fra de faktiske verdiene forplante seg videre og øke usikkerhetene i estimatene av potensiale fir effekttoppproduksjon.

Et godt sett med forklaringsvariabler kan gi god informasjon om hvert eneste mønster i responsvariabelen, men en god modell må til for å lære mønstrene ordentlig og ignorere eventuelt støy.

For visuell evaluering av modellene i denne oppgaven er det lagt vekt på hvor godt modellene predikerer generelle mønstre i el-bruken. Prediksjoner som får med toppene i el-bruken er å foretrekke foran prediksjoner som råker bunnene best. Det er ønskelig å ha gode prediksjoner også frem i tid, så det legges vekt på at prediksjonen ikke har veldig store avvik i enkelte deler av testsettet.

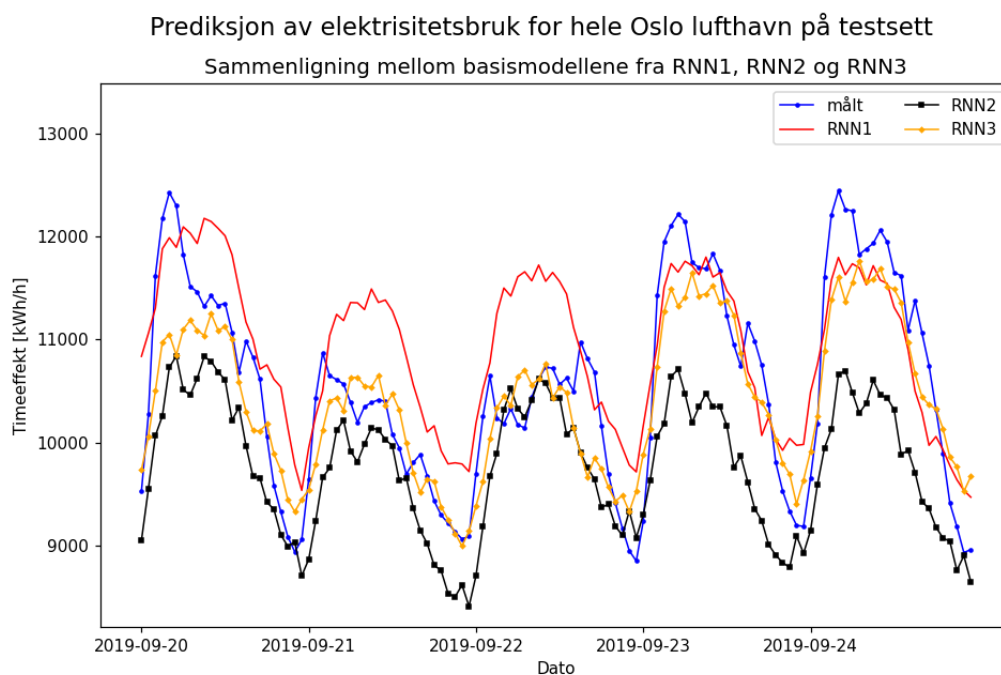
Sammenlignet med lignende modeller i forskningsartikler har de beste modellene som er presentert i denne oppgaven en lignende score. Tre av scorene som ble brukt til evaluering av modellene, MSE, RMSE og MAE, er skala-avhengige, altså kan de ikke uten videre sammenlignes mellom ulike datasett. For datasett med samme enhet på predikert variabel og kjent verdimengde ble derfor RMSE-scorene regnet om til prosenter for å kunne sammenligne.

Hver modellkonfigurasjon ble trent tre ganger med ulike innstillinger for tilfeldighet. For å vise hvilken innvirkning dette hadde på resultatene og hva som ble gjort for å minske innflytelsen på resultatene blir basismodellene fra de tre kjøringene sammenlignet i neste delkapittel.

5.2.2 Sammenligning av basismodellene fra RNN1-RNN3

Som nevnt i kapittel 3 ble alle modellkonfigurasjonene trent tre ganger med ulike *seed*, en innstilling for tilfeldighet. Tilfeldigheten hadde en innvirkning på prediksjonene fra hver modellkonfigurasjon. For å illustrere dette viser figur 5.2 prediksjonene for el-bruk én time frem i tid fra basismodellene fra de tre rundene RNN1, RNN2 og RNN3. Den siste kurven viser den faktiske el-bruken i de fem dagene midt i testperioden som er vist i figuren.

Man ser at basismodellene fra de tre rundene i figur 5.2 ikke predikerte like verdier for timeeffekt [kWh/h]. Basismodellen fra RNN3 følger de fleste dager den faktiske el-bruken godt, unntatt de øverste toppene. Sammenlignet med øvrige konfigurasjoner i RNN3 var basismodellen nest best rangert etter RMSE på testsettet. I motsetning var basismodellen i RNN2 blant de fire dårligste rangert etter RMSE på testsettet. Man ser fra figur 5.2 at den RNN1-kurven undervurderer el-bruken nesten alle timer. Basismodellen fra RNN1 var middels god. Den predikerte enkelte tidspunkter likt som RNN2 og andre tidspunkter likt som RNN3. Man ser av figur 5.2 at RNN1 enkelte dager følger faktisk el-bruk tett, mens den på nattestid overvurderer el-bruken.



Figur 5.2: Sammenligning av prediksjonsplott fra basismodellene fra RNN1, RNN2 og RNN3 sammen med faktisk el-bruk, fem dager midt i testperioden. Modellene fra de ulike rundene predikerer ikke likt for alle timer.

Forskjellen mellom basismodellen fra de tre kjøringene kom av at det er tilfeldighet involvert i initieringen av modellene og måten batch-er av data ble sendt inn til

modellen under treningsprosessen. Hver konfigurasjon i samme runde har samme innstillinger for tilfeldighet, men de ble variert mellom rundene.

Ved visuell sammenligning av prediksjonsplott fra de tre rundene så det ut til at hvordan hver konfigurasjon prestererte i forhold til basismodellen var omtrent likt fra kjøring til kjøring. Modellene som var veldig lik basismodellen i en kjøring var også det i de andre, og modellene som i gjennomsnitt presterte best var gode i alle rundene. Enkelte konfigurasjoner var likevel mye bedre eller verre i en runde enn i de to andre. For at ikke dette skulle påvirke hvilke konfigurasjoner som blir ansett som best ble det valgt å ta gjennomsnitt over alle tre rundene i de presenterte tallresultatene. Ved gjentakelse av eksperimentene enda flere ganger ville disse tallresultatene vært sikrere, men av tidshensyn ble det valgt å begrense til tre gjentakelser.

De visuelle representasjonene av modellkonfigurasjoner som blir presentert i de neste delkapitlene er hentet fra første kjøring, RNN1. Dette ble valgt fordi basismodellen predikerte middels godt den runden, mens den presterte godt i tredje runde og dårlig i andre runde. Som vist i figur 5.2 var ikke de predikerte kurvene identiske for de tre konfigurasjonene. Siden plottene som vises i de neste delkapitlene er basert på bare én kjøring er detaljene rundt for eksempel hvilke topper som ble truffet godt, og hvilke helgedager som ble overvurdert, visuelt unike for denne runden og derfor er et usikkerhetsmoment for resultatene.

5.3 Prediksjoner én time frem i tid

For prediksjon én time frem i tid, ble basismodellen og 37 andre modellkonfigurasjoner testet. Modellkonfigurasjonene hadde ulike kombinasjoner av forklaringsvariabler og nettverksarkitektur, som forklart i delkapittel 3.4. Resultatene fra konfigurasjonene blir i dette kapittelet sammenlignet med resultater fra en enkel modell og basismodellen, basert på plott fra RNN1 og gjennomsnittlige scorer fra de tre kjøringene RNN1 til RNN3. Prediksjonsplott som viser prediksjonene fra hver modell som ble testet sammenlignet med faktisk el-bruk og utfyllende tallresultater fra samtlige modeller finnes i henholdsvis vedlegg C og B. Av de 37 konfigurasjonene som ble testet, hadde 24 modeller bedre gjennomsnittlig RMSE på testsettet enn basismodellen. Kun to modeller var dårligere enn den enkle modellen.

5.3.1 Enkel modell

Tabell 5.1 viser RMSE-, MAE- og MAPE-scorene for den enkle modellen. Denne modellen predikerer at hver time har timeeffekt lik den gjennomsnittlige verdien fra treningssettet. Det er ønskelig at alle andre modeller skal predikere bedre enn denne.

5.3.2 Basismodell

Trenings- og testscorene til basismodellen, som ble beskrevet i kapittel 3.3, vises i tabell 5.2. Basismodellen hadde ett skjult LSTM-lag med 32 celler, 168 timer tilbakeblikk og ingen forenkling. Den tok inn alle forklaringsvariabler unntatt tidligere

Tabell 5.1: RMSE-, MAE- og MAPE-scorene for den enkle gjennomsnittsmodellen på testsettet.

Evalueringsmål	Testscore
RMSE [kWh/h]	974
MAE [kWh/h]	808
MAPE [%]	7,7

timeverdi for elektrisitetsbruk for å predikere timeverdi for elektrisitetsbruk én time frem i tid. I likhet med øvrige konfigurasjoner ble basismodellen trent, validert og testet på data fra sommerhalvåret 2019. Verdiene i tabellen er gjennomsnitt for de tre rundene modellene ble trent.

Sammenlignet med den enkle modellen i delkapittel 5.3.1 er basismodellen 22-26 % bedre for hver score. Dersom maskinlæringsmodellene ikke hadde predikert bedre enn den enkle modellen ville det indikert at maskinlæringsmodellene var uegnet for denne problemstillingen.

Tabell 5.2: Gjennomsnittlige RMSE-, MAE- og MAPE-scorene for basismodellen for trenings- og testsettene.

Evalueringsmål	Treningsscore	Testscore
RMSE [kWh/h]	425	760
MAE [kWh/h]	328	618
MAPE [%]	3,1	5,7

Prediksjonsplottet i figur 5.3 viser hvilke verdier basismodellen fra første treningsrunde, RNN1, predikerte for hver time i henholdsvis trenings- og testsettet. Modellen var i gjennomsnitt 25. best av modellkonfigurasjonene som ble testet, og fikk følgende navnet XXV. Grunnet tilbakeblikket på 168 timer mangler den første uken i begge datasett i plottet. Resultatene for valideringssettet er ikke vist. Sammenlignet med øvrige konfigurasjoner i RNN1 hadde basismodellen en gjennomsnittlig god RMSE-score.

Sammenligning av målte og predikerte verdier

Tidsperioden for trening som vises i øverste plott i figur 5.3 består av 20 uker. Første dato i treningsplottet er 08.04.2019, en mandag. Uke 18 og 27 som ble skildret i delkapittel 4.2 er inkludert i treningssettet.

Fra kurven med målte verdier i treningsperioden i figur 5.3 ser man at det i hele perioden er høyere el-bruk om dagen enn om natten. Man kan se skillene mellom ukene ved at en uke består av fem høye toppe i ukedagene og to lavere som er helgen. De høyeste toppene i el-bruk i perioden er tidlig i april og tidlig i mai, og på opp mot 14 000 kWh/h. I samme perioder er el-bruken høy også om natten. På dagene med lavest el-bruk er toppene mellom 10 000 og 11 000 kWh/h, og bunnene under 9 000 kWh/h. På treningssettet følger prediksjonskurven mønsteret til de faktiske

verdiene med tanke på døgnvariasjon, men den undervurderer bunnene og toppene i el-bruken.

Nederst i figur 5.3 vises kurven for faktisk el-bruk i testsettet som består av to og en halv uke og begynner med torsdag 12.09.2019. Her er skillet mellom helge- og ukedager enklere å se, ved at lørdag og søndag er to påfølgende dager med lavere verdier, etterfulgt av fem dager med høyere verdier. Mandag til lørdag har de fleste uker høyeste verdi om morgenen, mens søndag forekom høyeste verdi om ettermiddagen. Dette stemmer overens med mønstrene i el-bruken man så i delkapittel 4.2. Helgedager i perioden har høyeste verdier litt under 11 000 kWh/h, på størrelse med de laveste toppene i treningsperioden. Enkelte av ukedagene har høyeste verdier opp mot 13 000 kWh/h, som er lavere enn de høyeste toppene i treningsperioden. Laveste verdi ligger rundt 9 000 kWh/h alle døgn, på nivå med de laveste verdiene i treningsperioden.

På testsettet i figur 5.3 ser man at modellen predikerer lignende forskjeller mellom dag og natt og delvis mellom ukedag og helg som faktisk el-bruk. Alle de predikerte verdiene er innenfor samme verdimengde som de faktiske verdiene. Modellen predikerer høyere verdier om dagen enn om natten, noe som stemmer med faktisk el-bruk. Altså fanger modellen de største mønstrene i el-bruken.

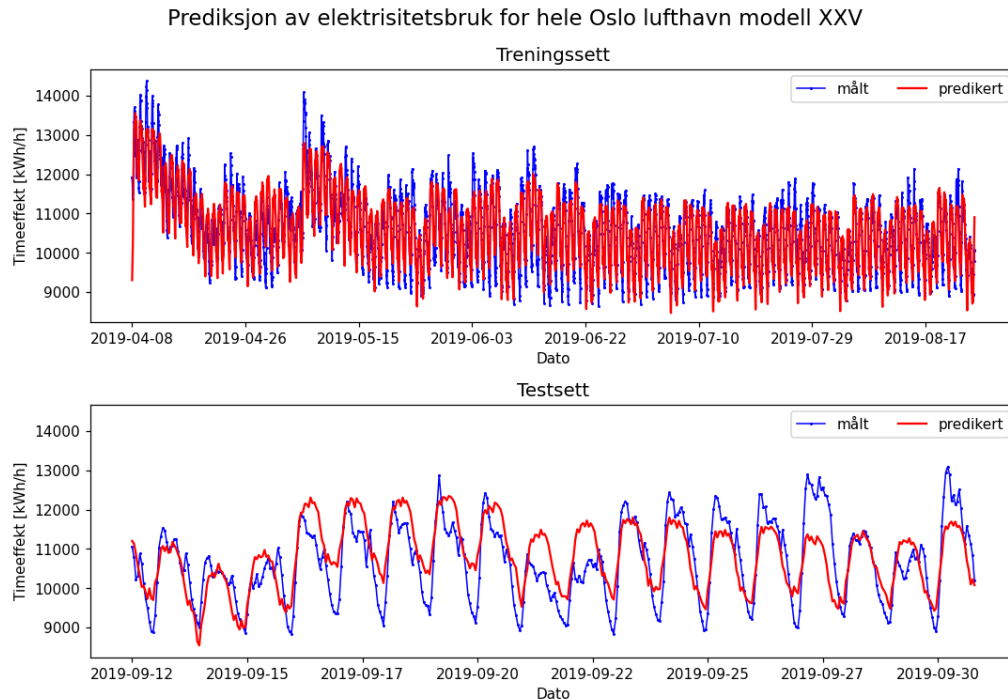
Man ser fra figur 5.3 at basismodellen visuelt har lignende treffsikkerhet på trenings- og testsettene, i den forstand at for begge datasett er basismodellen bedre på å predikere generelle mønstre enn topper og bunner. Prediksjonene på treningssettet er basert på sammenhengene modellen har lært etter å ha vært gjennom hele treningssettet i læringsprosessen, og det er derfor forventet at modellen predikerer godt i treningsperioden. Scorene i tabell 5.2 viser at den hadde 80-90 % bedre scorer på treningssettet enn testsettet. Det er vanlig at treningsscoren er høyest, men en så stor forskjell indikerer at basismodellen kan være noe overtilpasset. Det er uansett tydelig at modellen hadde de samme utfordringene i begge tidsrom; å predikere døgnets ekstremverdier godt.

Prediksjon av topper i el-bruken

Som nevnt i delkapittel 5.2.1 kan man ha som mål at modellene skal predikere topper i effekt- og energibruken, for eksempel dersom prediksjonene skal brukes til å estimere fleksibilitetspotensialet i en last eller bygning. I slike tilfeller vil ikke bunner og generelle mønstre være like viktige.

For prediksjon av effekttopper vil ikke basismodellen i figur 5.3 være helt god, fordi den enkelte dager overvurderer lengden på effekttoppene og andre dager undervurderer høyden på effekttoppene. For basismodellen er det de fleste døgn mellom null og 500 kWh/h forskjell på høyeste predikerte verdi og høyeste faktiske verdi, men enkelte dager mot slutten av testsettet er forskjellen over 1 500 kWh/h.

I en samlestudie av prediksjonsmodeller for topper i effektbruk skrev Dai mfl. [27] at for prediksjon av effekttopper er daglig, ukentlig, månedlig og årlig topplast de mest brukte tidsoppløsningene [27]. Det indikerer at timeprediksjon ikke er mest egnet for å råke topper. Modellene som ble testet i denne studien ble i treningen evaluert



Figur 5.3: Prediksjonsplott for sammenligning mellom målt og predikert el-bruk [kWh/h] for henholdsvis trenings- og testsettet for basismodellen fra RNN1. Modellen treffer i noen grad mønstrene i el-bruken.

på treffsikkerhet for hver time, ikke kun for timene med høyeste verdi, og det kan derfor forventes at de er mindre spesialisert mot topper.

Dai mfl. [27] viste til at en rekke algoritmer er forsøkt brukt til topplast-prediksjon de siste årene, blant annet rekursive nevralt nettverk, men også andre metoder som autoregresjonsmodeller og tidsseriedekomponering. Kunstige nevralt nettverk og andre dyplæring-metoder var nevnt som mest populære på feltet. Det tilsier at modellene som ble brukt i denne oppgaven kan fungere til prediksjon spesifikt av topper, men de må da trenes på andre typer data. Det er heller ikke gitt at de samme konfigurasjonene vil fungere godt for andre tidsoppløsninger. Dai mfl. foreslo også et eget evalueringsmål for effekttopp-prediksjon, *Peak Absolute Percentage Error* (PAPE) [27].

Dersom utelukkende døgnet høyeste verdi er interessant å predikere kan man bygge en annen type modell enn de som er presentert i denne oppgaven, en som har høyeste timeeffekt i løpet av døgnet som responsvariabel. Den kan være basert på målinger av døgnet høyeste verdier og andre forklaringsvariabler med samme tidsoppløsning. Det vil gjøre at input-datasettet minsker, og vil øke regneeffektiviteten. Ved lavere oppløsning på prediksjonene kan man også basere prediksjonene på færre forklaringsvariabler [27]. Med en annen tidsoppløsning vil ikke nødvendigvis de samme forklaringsvariablene være nyttige, for eksempel kan døgnet høyeste temperatur ha mer å si for toppene i effektbruk enn gjennomsnittstemperaturen.

I prediksjonsplottet fra basismodellen i figur 5.3 ser man at enkelte av ukedagene

har veldig spisse topper, og disse råker ikke modellen. De fleste dager varer døgnet høyeste verdi kun én time, mens basismodellen i figur 5.3 predikerer høye verdier i flere timer. Altså får ikke modellen med like bratt reduksjon i el-bruken utover ettermiddagen som det i virkeligheten er. Det mfl. viste at i effekttopp-prediksjon kan man predikere tidspunktet for toppene og verdien på toppene hver for seg med ulike modeller, eller predikere bare en av delene [27]. Man ser av figur 5.3 at basismodellen tidvis er bedre på å estimere verdien for effekttoppen enn det spesifikke tidspunktet, mens toppene i realiteten kommer på et nokså likt tidspunkt hver dag. Kanskje kan en kombinasjon av modeller, hvor man bruker en LSTM-modell til å predikere verdi, og en enklere modell til å predikere tidspunkt, være en måte å få en til sammen mer treffsikker prediksjon.

Som nevnt i delkapittel 5.2.1 ble det i denne oppgaven fokusert på å treffe generelle mønstre i dataene, men topper ble foretrukket foran bunner. I delkapittel 5.3.5 med utvalgte prediksjonsplott blir det vist at basismodellen ikke var den beste av modellkonfigurasjonene til å predikere verken topper, bunner eller generelle mønstre.

Basismodellen predikerer døgnet laveste verdi på samme tidspunkt som de faktiske bunnene oppsto nesten alle dager, men de fleste dager er det 500 til 1 500 kWh/h forskjell mellom faktisk laveste verdi og laveste predikerte verdi. Modellen er altså bedre på å råde riktig tidspunkt for bunner enn topper. Bunnene kommer på et mer fast tidspunkt enn toppene fra dag til dag. For disse modellene som skal treffe generelle mønstre hadde det derfor vært et dårlig tegn om modellene ikke traff riktig på bunnene i datasettene.

Flere av mønstrene i predikert el-bruk fra basismodellen kan man kjenne igjen i andre konfigurasjoner som blir presentert med scorer og plott i henholdsvis delkapittel 5.3.3 og 5.3.5.

Sammenligning mellom scorer og visuell måloppnåelse

I følge MAE-verdien til testsettet i tabell 5.2 predikerer modellen i gjennomsnitt 618 kWh/h feil for hver time. Basert på RMSE-verdien for testsettet i samme tabell er prediksjonen til basismodellen i gjennomsnitt 760 kWh/h feil. RMSE-scoren er 140 kWh/h høyere enn MAE-scoren. Som beskrevet i delkapittel 5.2.1 skyldes dette at RMSE legger større vekt på store prediksjonsfeil. Scorene er beregnet basert på basismodellene fra alle tre kjøringene RNN1-RNN3, men MAE- og RMSE-scorene ser ut til å stemme med det visuelle resultatet i figur 5.3. Formiddag og kveld, når elektrisitetsbruken øker fra bunn til topp og minker i motsatt retning, predikerer modellen de fleste dager godt. Den klarer ikke å predikere alle topper og bunner rett, og er enkelte dager over 1 000 kWh/h feil på disse tidspunktene. Denne typen avvik gjør at RMSE-verdien øker.

Tendens til forverring over tid

Plott av prediksjonene viser tendenser til at modellene predikerer bedre for dagene tidlig i testsettet. Dette kan man se av basismodellen i figur 5.3, og i flere av prediksjonsplottene som blir vist i delkapittel 5.3.5. Mange dager, spesielt mot slutten

av testperioden, predikerer basismodellen for lave verdier for el-bruken når den har sine topper. Tidlig i testsettet predikerer basismodellen bedre på døgnetts høyeste verdi, men treffer ikke samtidig døgnetts laveste. Senere i testperioden predikerer den nærmere bunnen, men ikke samtidig toppen. De første seks døgnene er den høyeste predikerte verdien hver dag mindre enn 200 kWh/h unna den faktiske toppen. Helgedagene midt i testsettet overvurderer modellen de høyeste verdiene, mens den mot slutten er over 1 5000 kWh/h feil den 27.09 og 30.09. Modellen er generelt ikke god på å treffe både bunn og topp samtidig, men er best på dette 13.09 og 14.09, helt i starten av perioden.

Avvikene antas å komme av at jo lengre ut i testsettet man kommer, jo lengre er man fra tidsrommet som modellen er trent på og dermed best tilpasset. Modellene ble trent på data fra april til august 2019, mens slutten av testsettet er sent i september. Med unntak av tidlig i treningsperioden, er de fleste toppene i treningssettet mellom 11 000 og 12 000 kWh/h. Utover i testsettet nærmer flere topper seg 13 000 kWh/h. Som nevnt vil dataene modellen er trent utgjøre på en begrensning for modellen. Modellen vil være best på å predikere verdier i et intervall den er kjent med.

Avhengig av anvendelsene til modellene kan dette være et problem. Dersom modellene primært skal anvendes på tidspunkter som er nær tidspunktene den er trent på, vil det ikke være et stort problem. Dette vil være tilfellet dersom man kontinuerlig vil trene modellen på nye data for å predikere el-bruk. I motsatt fall, hvis modellen skal trenes én gang og så anvendes på ulike tidsrom og tidspunkter langt frem i tid, bør treningssettet inkludere dager som ligner på de som skal predikeres.

Basismodellen har en minimal nettverksarkitektur. Den har to lag, et LSTM-lag og et output-lag, som er et minimum for en fungerende modell. Antallet celler i LSTM-laget ble satt til en lav verdi, og standardinnstillingene i lagene ble beholdt. Modellen er likevel i stand til å predikere el-bruk med 6 % MAPE-avvik på testsettet. Dette er vesentlig bedre enn den enkle modellen i delkapittel 5.3.1. Sammenlignet med de 37 øvrige modellkombinasjonene som ble testet til prediksjon én time frem i tid var basismodellen middels god. I neste delkapittel presenteres elleve modellkonfigurasjoner som presterte bedre enn basismodellen.

5.3.3 Topp elleve beste konfigurasjoner

Til prediksjon av el-bruk én time frem i tid ble 38 ulike modellkonfigurasjoner, inkludert basismodellen, testet og undersøkt. Totalt var treffsikkerheten til modellene 511-1168 kWh/h i test-RMSE og 3,7-8,9 % i test-MAPE. De elleve modellkonfigurasjonene I-XI som hadde lavest gjennomsnittlig RMSE-avvik på testsettet er vist i tabell 5.3. Tabellen viser hvilke forklaringsvariabler modellene hadde og nettverksarkitekturen med antall LSTM-lag og antallet enheter i hvert lag. I tillegg er RMSE-scoren til hver modell for både trenings- og testsettet inkludert. Modellene er navngitt med romertall etter hvilken rangering de har basert på RMSE-verdi for testsettet. Der det mangler en verdi i kolonnene har modellene samme egenskaper som basismodellen, som er presentert på nederste rad i tabellen og nærmere forklart i delkapittel 3.3. Samtlige modeller i tabellen har i likhet med basismodellen 168

timer tilbakeblikk og 80:10:10 splitt mellom trenings-, validerings- og testsettene.

Prediksjonsplott som viser hvordan modellene I, III og XI predikerte, sammenlignet med faktisk el-bruk, er inkludert i delkapittel 5.3.5. En fullstendig oversikt over de 38 modellkonfigurasjonene som ble testet og hvilke gjennomsnittlige RMSE-, MAE- og MAPE-scorer de fikk på trenings- og testsettene er lagt ved i vedlegg B.

Tabell med beste konfigurasjoner

Fra tabell 5.3 ser man at samtlige modeller på topp elleve-listen har færre forklaringsvariabler, inkluderer tidligere el-bruk og/eller har mer kompleks nettverksarkitektur enn basismodellen.

Tabell 5.3: Oversikt over forklaringsvariabler, nettverksarkitektur og RMSE-score for trenings- og testsettene for de elleve beste modellkonfigurasjonene I-XI og basismodellen XXV. Modellene er rangert etter test-RMSE. Tomme ruter tilsvarer samme egenskaper som basismodellen.

Navn	Forklaringsvariabler	Nettverksarkitektur	Tren.-RMSE [kWh/h]	Test-RMSE [kWh/h]
I	Tidligere el-bruk		417	511
II	Tidsvariabler, tidl. el-bruk	1 LSTM-lag m/128 celler	279	586
III		3 LSTM-lag m/64 celler	406	624
IV	Alle inkl. tidligere el-bruk	3 LSTM-lag m/64 celler	407	625
V		2 LSTM-lag m/32 celler	432	636
VI	Alle inkl. tidligere el-bruk		349	644
VII	Tidsvariabler, tidligere el-bruk		353	648
VIII		2 LSTM-lag m/32 celler 10 % forenkling	465	655
IX		1 LSTM-lag m/128 celler	337	678
X	Tidsvariabler, temperatur, passasjerdata		394	678
XI	Passasjerdata		746	686
XXV	Alle ekskl. tidligere el-bruk	1 LSTM-lag m/32 celler 0 % forenkling	425	760

Fire av modellene har færre forklaringsvariabler enn basismodellen, med el-bruk, passasjerdata og tidsvariabler nevnt to ganger hver og temperatur nevnt én gang, noe som indikerer at dette er viktige forklaringsvariabler. Nedbør og vind opptrer ikke alene, kun sammen med samtlige andre forklaringsvariabler, noe som tilsier at dette ikke er like viktige variabler. En modell er helt lik som basismodellen, men har i tillegg tidligere el-bruk som forklaringsvariabel.

Tre av modellene har en arkitektur med flere skjulte lag enn basismodellen, og to av disse har også flere celler i hvert lag. Ytterligere to modeller har 128 celler i hvert lag, i stedet for 32 som i basismodellen. Dette indikerer at flere skjulte lag eller flere celler i de skjulte lagene kan være hensiktsmessig for gode prediksjoner. Kun en av modellene har 10 % forenkling, noe som indikerer at forenkling ikke forbedret modellene særlig.

Man ser at modell I med samme arkitektur som basismodellen, og kun time på døgnet og tidligere timeeffekt som forklaringsvariabler, ga den beste prediksjonen. Marginen i RMSE ned til modell II var 75 kWh/h, nesten like stor forskjell som mellom modell II og modell XI. Sammenlignet med basismodellen har modell I 33 % forbedring i RMSE. Med en MAPE på 3,7 % har den 2 prosentpoeng forbedring i MAPE sammenlignet med basismodellen. Av modellene på topplisten har fire andre tidligere el-bruk som variabel. Det indikerer at tidligere el-bruk er en viktig forklaringsvariabel.

Fra modell II til modell XI er det 75 kWh/h kWh/h forskjell i test-RMSE, noe som indikerer at mange av modellene predikerer relativt like godt. Visuelt hadde flere av disse enten en veldig lik prediksjonskurve som basismodellen, eller noe dårligere visuell måloppnåelse. Sammenlignet med steget i score opp til modell I tyder intervallet i RMSE-score fra modell II til XI på at man kan oppnå ganske gode prediksjoner på flere måter, men for å få en virkelig god modell trengs det mer justering.

Innvirkning av nettverksarkitektur og forklaringsvariabler

Seks av modellene i tabell 5.3 har som nevnt mer kompleks arkitektur enn basismodellen, med enten tre LSTM-lag med 64 celler, to LSTM-lag med 32 celler eller ett LSTM-lag med 128 celler. Fem av disse hadde samme eller flere forklaringsvariabler enn basismodellen. Siden flere modeller med kompleks arkitektur havnet på topplisten kan det være interessant å utforske flere variasjoner av dypere nettverk til videre prediksjoner.

Når det gjelder hvor mange lag i modellen som er best spriker resultatene. Som man ser av tabell 5.3 var det på topp elleve-listen modellkonfigurasjoner med både ett, to og tre lag. Modellene med tre lag og 64 celler i hvert lag gjorde det godt, mens modellene med tre lag og 32 celler i hvert lag havnet blant modellene med dårligere scorer enn basismodellen. Det tyder på at det trengs flere celler i hvert lag i dypere nettverk for å lære sammenhenger som representerer dataene godt, men også at man ikke kan ta for gitt at større nettverk predikerer bedre.

Modell IX hadde, som man ser av tabell 5.3, samme forklaringsvariabler som basismodellen, men 128 celler i hvert lag. Det er fire ganger flere enn basismodellen sine 32 celler. Denne modellen fikk 11 % bedre test-RMSE enn basismodellen. Modellene med ett skjult lag og henholdsvis 16 og 64 celler i det skjulte laget fikk derimot henholdsvis 17 % og 5 % dårligere test-RMSE enn basismodellen. En modell med 64 celler som også hadde tidligere el-bruk som forklaringsvariabel ble også dårligere enn basismodellen. Det indikerer at veldig få LSTM-celler ikke klarer å lære de nødvendige sammenhengene i dataene, men også at man ikke ukritisk kan

øke antallet celler og forvente en økt treffsikkerhet.

I tillegg til de dypere modellene som ga gode prediksjoner og havnet i tabell 5.3, ble det testet flere modeller med to lag og forskjellige antall celler i hvert lag. Modellene med to lag og lik mengde celler gjorde modellene det bedre enn basismodellen, mens modellene med to lag og ulik mengde celler i de to lagene gjorde det dårligere. Modell V med to lag og 32 celler i hvert lag havnet for eksempel på topp elleve-listen med 16 % bedre RMSE-test enn basismodellen. En lignende modell, XXII, med henholdsvis 32 og 16 celler i hvert lag hadde 2 % bedre test-RMSE enn basismodellen og var med det kun marginalt bedre. Blant alternative versjoner av modell V og XXII med ulike prosenter forenkling presterte modellene med ulikt antall celler i lagene også dårligere enn lignende modeller med likt antall celler. Det virker som modellen predikerer bedre med lik størrelse på lagene.

Modeller med forenkling på 10, 20 og 50 %, men ellers var lik basismodellen ble også testet. Konfigurasjonene med 20 % og 50 % forenkling fikk marginalt bedre test-RMSE enn basismodellen, mens modellen med 10 % forenkling var marginalt dårligere. Visuelt basert på testsettet er det ikke mulig å se forskjell på noen av modellene med forenkling og basismodellen selv. Modeller med to lag og enten 32 celler i hvert lag eller henholdsvis 32 og 16 celler ble også testet med ulike prosenter forenkling for å se om det hadde en innvirkning i et dypere nettverk. Det ga i alle tre tilfellene en liten forverring i RMSE-avvik, men de forenklede modellene hadde veldig like predikerte kurver som de ikke-forenklede variantene.

Basert på dette virker det som forenkling har svært lite å si for modeller med bare et eller to lag og maksimalt 32 celler i hvert lag. En allerede dårlig modell, ble ikke bedre av å forenkle. En allerede ganske god modell ble heller ikke bedre. Rekursive nevrale nettverk kan dog bygges med flere enn tre LSTM-lag, og lagene kan ha titalls ganger flere celler i hvert lag enn modellene i denne oppgaven. Rafi mfl. [41] predikerte elektrisk last for hele Bangladesh' kraftsystem med en kombinasjon av konvolusjonelle og rekursive nevrale nettverk. De hadde 200 LSTM-lag i sin modell [41]. Sammenlignet med det er konfigurasjonene som er undersøkt i denne oppgaven så enkle at forenkling har mindre innvirkning.

Modellene med to, fem og ti døgn tilbakeblikk og med alternative splitter mellom trenings-, test- og valideringssettene hadde en minimal forbedring i gjennomsnittlig RMSE-avvik for testsettet sammenlignet med basismodellen. Prediksjonsplottene fra disse modellene var ikke mer treffsikre enn plottene fra basismodellen. En doblett batch-størrelse ga marginalt bedre RMSE-score på testsettet enn basismodellen, men gjorde ikke at modellen ble trent vesentlig raskere og ga ikke visuelt utslag. En endret aktivering i *Dense*-laget fra lineær aktivering til *Rectified Linear Units* (ReLU) [78] ga også bare en minimal forbedring i test-RMSE og ingen endring i prediksjonsplottet. Det indikerer at for dette datasettet og modellkonfigurasjonene som er testet har valget av variabler og nettverksarkitektur mer innvirkning på resultatet enn disse innstillingene.

En fordel med å øke størrelsen på testsettet er likevel at man får mer data å basere testscorer og refleksjoner rundt visuell måloppnåelse på. En fordel med å øke størrelsen på valideringssettet er at algoritmen under treningen har mer data å ba-

sere utvelgingen av modell på. For modellene som ble testet med henholdsvis 20 % testdata og 20 % valideringsdata virker dette ikke å ha bidratt til å bedre de endelige scorene på testsettet.

For å begrense tidsbruken til trening og prediksjon av modellene kan man begrense lengden på tilbakeblikk og kompleksiteten til nettverket. Som nevnt tidligere hadde lengden på tilbakeblikk liten innvirkning på både scorene og prediksjonsplotene, men med kortere tilbakeblikk tok modellene kortere tid å trene. Dermed kan modellene trenes eller testes oftere uten å kreve for mye regnekapasitet. I denne innledende studien var ikke tidsbruken til trening eller testing av modellene et sentralt vurderingskriterium. Slowik og Urban [42] vektla derimot i sin evaluering av en LSTM-modell for prediksjon av elektrisitetsbruk i et prosessindustri-anlegg at modellen var relativt rask med en treningstid på opptil 8 minutter, og at dette er en fordel i praktiske anvendelser og ved fortløpende prediksjoner på nye målinger som sendes direkte til modellen [42]. De mest komplekse modellene som ble testet i denne oppgaven tok også lengst tid å trene og brukte flest epoker før de konvergente, men det tilsvarte maksimalt en halv time.

5.3.4 Innvirkning av evalueringsmål på rangering

Dersom man sorterer modellkonfigurasjonene i tabell 5.3 på test-MAE, ikke test-RMSE, er de samme elleve modellene de beste, men rekkefølgen endres noe. Modell I og II vil fortsatt være best, deretter kommer modellene med mer kompleks arkitektur. Modellene med samme arkitektur som basismodellen vil falle nederst på topplisten. Dette forsterker antagelsen om at å trene modellene med MAE som tapsfunksjon kan endre prediksjonene og hvilke modeller som gjør det godt. Ved sortering på test-MAPE faller modell VIII ut av listen, og en modell med 48 timer tilbakeblikk kommer inn på 11. plass.

En sammenligning av scorene mellom trenings- og testsettene beskriver hvorvidt modellen har lært de riktige sammenhengene mellom forklaringsvariabler og responsvariabel. Hvis den har det, vil både trenings- og testscoren være lav. Testscoren vil være høyest av de to. Dette ser ut til å være tilfellet for de fleste av konfigurasjonene i tabell 5.3.

Dersom treningsscoren er veldig god og testscoren dårlig indikerer det at modellen er overtilpasset til dataene som modellen er trent på. Dersom man sorterer modellene basert på treningsscorene, for eksempel trenings-RMSE, endres topplisten vesentlig. Modeller som rangerte dårligere enn basismodellen basert på testscorene er blant de beste basert på treningsscorene. Dette gjelder for eksempel modellene XXVII og XXIX som er rangert som nummer 27 og 29 i tabellen med alle modellkonfigurasjoner i vedlegg B. Disse modellene har større differanse mellom test- og treningsscore enn modellene som er vist i tabell 5.3. Det indikerer at de er overtilpasset treningssettet.

Basert på RMSE for treningssettet er de to beste modellene II og IX fra tabell 5.3. I likhet med basismodellen har de kun ett skjult lag, men de har 128 celler i dette laget. Modell II har også kun tidsvariabler og tidligere timeeffekt som forklaringsvariabler. Disse modellene hadde en varierende treffsikkerhet fra runde til runde, og i

gjennomsnitt over dobbelt så høy testscore som treningscore på alle tre evalueringsscorene. At de har såpass mye bedre trenings-RMSE enn test-RMSE indikerer at de kan være overtilpasset treningssettet, og at ett enkelt lag med mange celler ikke er pålitelig til dette datasettet.

Dersom treningscoren er dårligere enn testscoren, indikerer det at modellen ikke har lært nok fra treningssettet og har hatt flaks med å likevel predikere godt på testsettet. Det synes å være tilfellet med modell XI som kun hadde passasjerdata og timeinformasjon som forklaringsvariabler. Som eneste modell hadde den høyere treningscore enn testscore.

Av tallresultatene for basismodellen i tabell 5.2 kan man se at scoren på testsettet er 80-90 % høyere enn på treningssettet for alle scorene. Som beskrevet over er treningscoren gjerne bedre enn testscoren, men differansen er større for basismodellen enn for modellene i tabell 5.3. Trenings-RMSE for basismodellen er på nivå med de beste modellene, men test-RMSE er vesentlig lavere. Dette indikerer at basismodellen er overtilpasset. Modellene på topplisten har enten enklere arkitektur eller færre forklaringsvariabler. Det indikerer også at basismodellen har for enkel arkitektur i forhold til mengden forklaringsvariabler, noe som kan være grunnen til at den er overtilpasset.

I tillegg til å evaluere scorene ble prediksjonene fra hver modellkonfigurasjon plottet sammen med faktiske verdier for el-bruk. I neste delkapittel presenteres et utvalgt slike prediksjonsplott.

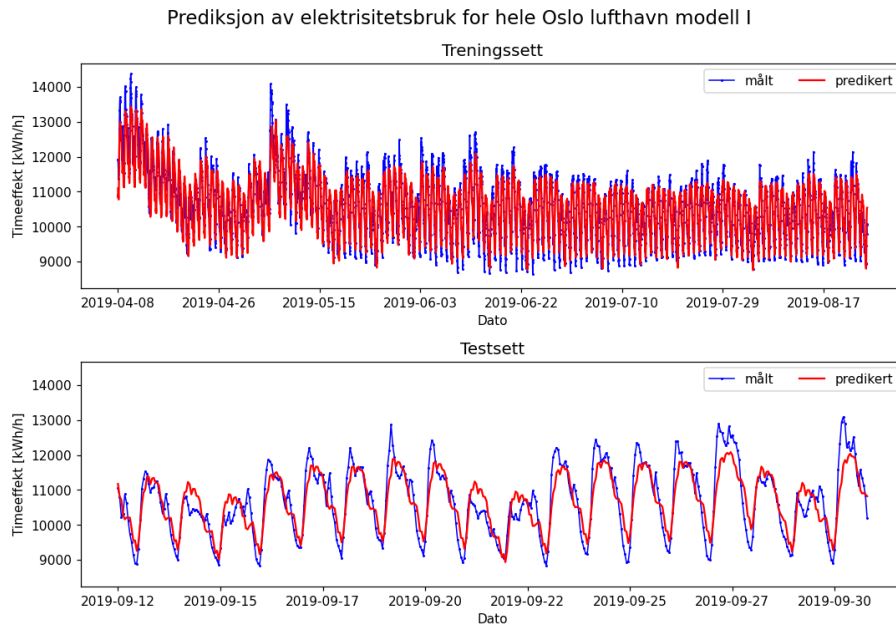
5.3.5 Utvalgte prediksjonsplott

I dette delkapittelet vises prediksjonsplott fra modell I, III og XI fra RNN1, første trening av modellene. Samtlige 37 prediksjonsplott fra RNN1, samt prediksjonsplott fra basismodellen i hver av de tre rundene, er lagt ved i vedlegg C.2.

Modell I

Prediksjonene fra modell I, som hadde laveste gjennomsnittlige RMSE-verdi på testsettet, vises i figur 5.4 for hver time i henholdsvis trenings- og testsettet. Modell I har samme arkitektur som basismodellen, men kun time på døgnet og tidligere elektrisitetsbruk som forklaringsvariabler. Prediksjonen fra modell I ligger tett inntil de målte verdiene de fleste timer, og er visuelt vesentlig bedre enn basismodellen fra RNN1. Modell I overvurderer verdien på høyeste timeeffekt i helgedagene, og undervurderer den de fleste ukedager, men prediksjonsavviket er sjelden over 1 000 kWh/h. Til sammenligning predikerte basismodellen inntil 1 500 kWh/h feil enkelte dager. Mens den faktiske el-bruken ofte har en spiss topp på ukedagene, predikerer modell I en lavere topp som flere dager ligger en time eller mer feil. Det er en mer tydelig forskjell mellom ukedag og helg enn i prediksjonsplottet fra basismodellen i figur 5.3.

Haug [32] brukte en prediksjonsmodell lignende modell I til lastprediksjoner som et ledd i estimering av kortsiktig fleksibilitet i en distribuert strømkomponent. Haug



Figur 5.4: Prediksjonsplott for henholdsvis trenings- og testsettet for modellkonfigurasjon I, som fikk beste gjennomsnittlige RMSE-verdi på testsettet. Modellen predikerer el-bruk én time frem i tid.

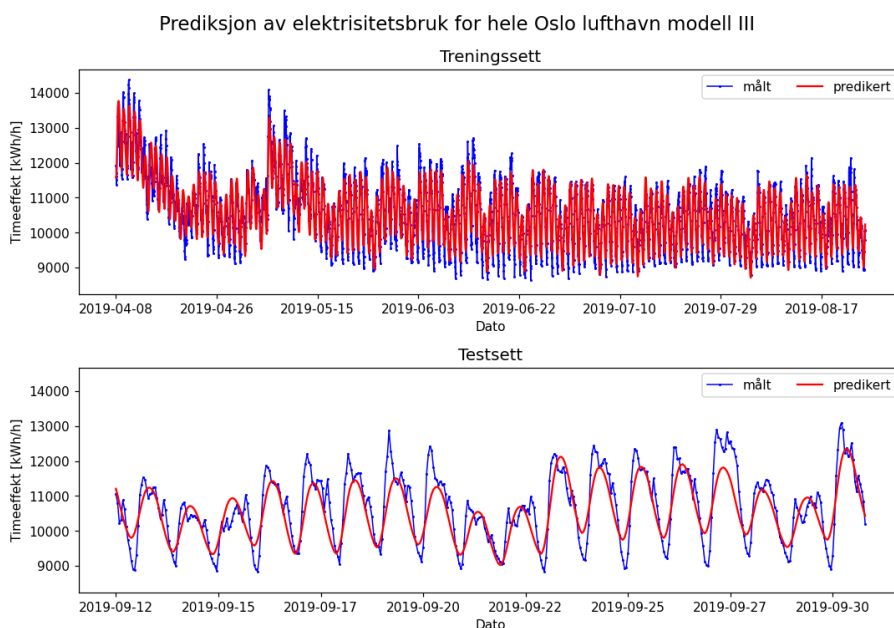
la vekt på stabile prognoser som traff topper i forbruket godt, og overestimering av el-bruken ble foretrukket framfor underestimering [32]. Til den ville det vært en ulempe at modellene som presenteres i denne oppgaven undervurderer el-bruken i døgnet mest kraftkrevende timer.

Ved prediksjon av el-bruk for neste time ved hjelp av en to-lags LSTM-modell med 100 celler i hvert lag og 10 % forenkling oppnådde Haug en RMSE på 85 kWh/h for et testsett med verdier på 400-950 kWh/h [32]. Det tilsvarer et avvik i RMSE på 9-20 % av el-bruken. Den beste modellen i denne oppgaven, modell I, oppnådde en RMSE på 511 kWh/h for et datasett med verdier på 8 600 til 14 400 kWh/h. Det tilsvarer et avvik på 4-6 % i RMSE. MAPE-scoren på testsettet fra modell I er på 3,7 %. Altså presterte modell I fra denne oppgaven noe bedre. Haug la vekt på at modellene ikke ble finjustert til formålet, men at et tilfeldig valg av egenskaper ble testet. I denne oppgaven ble det lagt vekt på å teste ulike aspekter, men de ble ikke satt sammen tilfeldig.

Til sammen sju modellkonfigurasjoner ble testet med tidligere el-bruk som forklaringsvariabel. Fire modellkonfigurasjoner var på topp elleve-listen. Disse hadde samme eller dypere arkitektur enn basismodellen. De tre siste var dårligere enn basismodellen. Av disse hadde modell XXX og XXVII var dårligere en mer kompleks arkitektur enn basismodellen. Dette tyder på at tidligere el-bruk kan være viktig å ha med i videre modeller, men at sluttresultatet likevel avhenger av nettverksarkitekturen. Den siste, modell XXXVI, hadde kun tidligere timeverdi for el-bruk som forklaringsvariabel, og ikke time på døgnet. Dette indikerer at tidligere el-bruk alene ikke gir god nok informasjon, man må ha med informasjon om hvilken time som blir predikert. Ingen andre konfigurasjoner ble derfor testet uten time-variabelen.

Modell III

I figur 5.5 ser man hvilke verdier modell III predikerer for hver time i henholdsvis trenings- og testsettet. Denne modellen har samme forklaringsvariabler som basismodellen og en dypere arkitektur med tre skjulte lag med 64 LSTM-celler i hver. Denne modellen har lært de største mønstrene i datasettet. Den hadde MAPE på 4,6 % på testsettet.



Figur 5.5: Prediksjonsplott for henholdsvis trenings- og testsettet for modellkonfigurasjon III, som fikk tredje beste gjennomsnittlige RMSE-verdi på testsettet. Modellen predikerer el-bruk én time frem i tid.

Sammenlignet med modell I i figur 5.4 er både tidspunktene og verdiene på prediksjonene fra modell III på dagtid nokså like, forskjellen er at modell III predikerer glattere kurver og treffer dårligere på lave verdier om natten. Modell III predikerer verdier som passer i skala med høyeste verdi for hvert døgn, men undervurderer de fleste dager hvor høyt toppene går. Igjen kommer vurderingen av visuell treffsikkerhet an på hva som er evalueringskriteriene for modellen. Dersom målsettingen er å predikere det generelle nivået på el-bruken er modell III god, men dersom variasjon fra time til time er sentral, kommer modellen dårligere ut. Modellen treffer verdien på effekttoppene nesten like godt som modell I. Dersom det nøyaktige tidspunktet til effekttoppen ikke er viktigst, kun verdien, kan modell III derfor gjøre nytten. I likhet med basismodellen er det enkelte dager over 1 500 kWh/h forskjell mellom predikert verdi og faktiske topper og bunner i el-bruken, men de fleste dager er forskjellen 1 000 kWh/h eller mindre. Andre tidspunkter på døgnet er forskjellen ned mot null.

Hwang mfl. [7] predikerte effektbruk per time gjennom døgnet for 28 næringsbygninger basert på meteorologiske variabler og bygningskarakteristikker. De trente og testet modeller på hver av bygningene separat. Med en LSTM-modell oppnådde de

en gjennomsnittlig MAPE på 9-11 %. Dette var en modell trent på tre år med data og med 100 celler i hvert av de tre skjulte lagene [7]. Modellene som presenteres i denne oppgaven er trent på et kortere datasett. Det kan forventes å være enklere å finne sammenhengene i korte enn i lange datasett. Modell III er den beste modellen fra denne oppgaven som kun har værdata og passasjerantall i tillegg til tidsvariabler og den har en MAPE på 4,6 %. Det er under halvparten av Hwang mfl. [7] sin beste modell. Modell III har også en lignende arkitektur som Hwang mfl. sin modell, med tre lag og 64 celler. Siden modellen trenes på et kortere datasett virker dette rimelig.

Dersom man sammenligner prediksjonsplottene fra basismodellen og modell I med modell III som har den mest komplekse arkitekturen, ser man at III predikerer mye glattere kurver enn enklere modeller. De mindre dype modellene predikerer verdier som varierer mer over kort tid. Begge typene modeller har fordeler og ulemper. De veldig jevne kurvene fra de dypere modellene råker for eksempel enkelte topper dårligere mens de varierende kurvene kan være mer utsatt for støy. Fra prediksjonsplottet ser man at med dypere nettverk kan man miste noe av variasjonen fra time til time som den faktiske el-bruken har. Samtidig treffer denne modellen bedre enn basismodellen på verdien til døgnet høyeste verdi.

Den samme tendensen til jevnere kurver ved dypere modeller kan man også se i prediksjonsplottene fra nettverkene med to lag og 32 eller 16 celler i hvert lag. Det tyder på at større nettverk skiller ut informasjonen om variasjoner gjennom døgnet i tidlige lag, og at denne informasjonen kan falle bort i dypere nettverk. Dette kan for eksempel være fallet i el-bruk mellom morgen og formiddag slik som man ser i figur 4.3 som viser elektrisitetsbruk for flyplassen gjennom døgnet. Modellene med tre skjulte lag predikerer høyeste verdier midt i døgnet, mens man ser av de faktiske el-bruuskurvene at toppene i realiteten er enten morgen i ukedager eller på ettermiddagen i helgedager.

Sammenlignet med modell I har modell III en mer tydelig tendens til forverring i tid, slik som man så i basismodellen Modell I traff bunnene i el-bruken godt nesten alle døgn. Modell III treffer bunnene generelt dårligere, er best midt i testperioden, og har avvik på 500 kWh/h eller mer siste halvdel av testsettet. Den treffer dermed få topper og bunner de siste dagene.

Modell XI

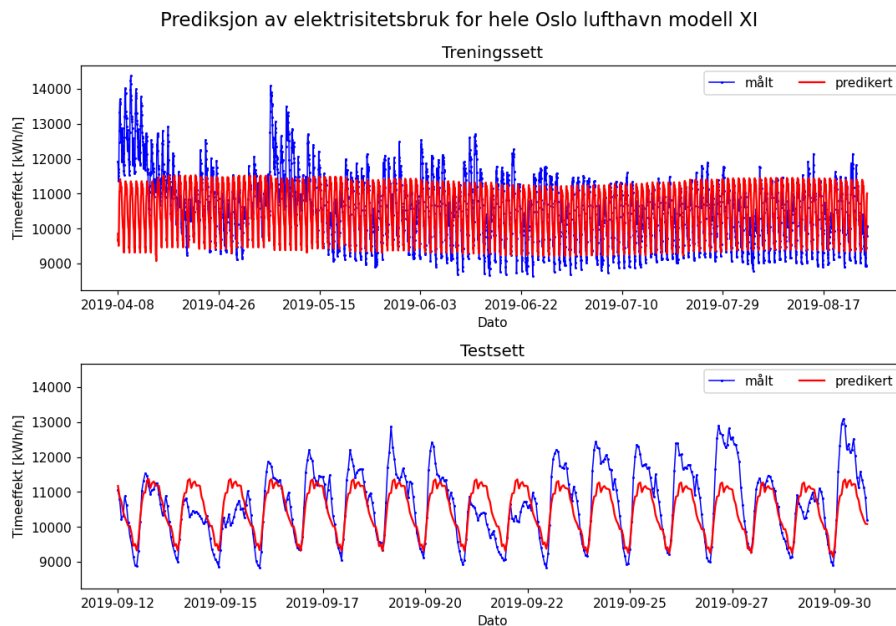
Som nevnt i delkapittel 5.2.1 er enkelte av modellene med gode scorere på trenings- og testsettet ikke visuelt bedre enn basismodellen. Dette illustreres i figur 5.6, som viser hvilke verdier modell XI predikerer for hver time i henholdsvis trenings- og datasettet. Modellen fikk en MAPE på 4,8 % på testsettet. Denne modellen har samme arkitektur som basismodellen, men kun time på døgnet og passasjerdata som forklaringsvariabler.

I figur 5.6 ser man at modell XI predikerer nesten samme verdier for samtlige døgn. Det er stor forskjell mellom faktiske verdier og predikerte verdier fra modell XI, men likevel har prediksjonen en veldig god score i forhold til basismodellen. Dette bekrefter at å se prediksjonsplott og tallscorer som en helhet kan være hensiktsmessig.

Ved å predikere et slags gjennomsnitt, er avviket mellom predikert verdi fra modell XI og faktisk verdi for el-bruk enkelte dager over 2 000 kWh/h. Andre dager er avviket nesten null. Modellen predikerer godt mange av dagene som ikke har spesielt høye topper, mens modellen predikerer for høye topper flere av dagene hvor el-bruken er lav.

Som tabell 5.3 viste har modell XI en gjennomsnittlig RMSE for testsettet på 686 kWh/h og for treningssettet på 746 kWh/h. I motsetning til mange av de andre modellene på topplisten har denne modellen liten forskjell i scorene mellom trenings- og testsettet, og den er den eneste av modellkonfigurasjonene som har en høyere RMSE på treningssettet enn testsettet. Dette kan komme av at det tidlig i treningssettet er mange høye topper som modellen ikke treffer, noe som drar opp scorene for treningssettet. Samtidig indikerer det at modellen i liten grad lærte seg mønstrene i treningssettet, og heller hadde flaks med at de mønstrene den lærte seg passet overens med testsettet. Modellen er undertilpasset.

At modell XI med kun passasjerdata med ukeoppløsning og time på døgnet predikerte dårlig visuelt tyder også på at denne kombinasjoner av forklaringsvariabler ikke kan gi nok informasjon til modellen til å gi en presis prediksjon av el-bruk. Som vist i delkapittel 4.3.2 varierer passasjerantallet i treningsperioden fra 250 000 til 350 000, med høyest passasjerantall midt i treningsperioden. Dette er ikke mulig å se igjen i prediksjonsplottet, og indikerer at lite informasjon ble overført til prediksjonene.



Figur 5.6: Prediksjonsplott for henholdsvis trenings- og testsettet for modellkonfigurasjon XI, som har dårlig visuell måloppnåelse. Modellen predikerer el-bruk én time frem i tid.

Oppsummering

Modellene som er presentert i denne oppgaven har varierende treffsikkerhet på effekt-topper. Ingen av modellene predikerer toppene i timeeffekt rett på både tidspunkt

og høyde, men flere av modellene viser potensiale. Det gjelder modell I og III, som tidlig i testsettet predikerer lavt avvik mellom faktiske topper og predikerte topper. Også små avvik kan føre til store usikkerheter dersom prediksjonene brukes videre til optimalisering. Ytterligere tilpassing av modellene forventes å gi bedre resultater.

Haq mfl. [34] kombinerte to variasjoner av LSTM-modeller, toveis og konvulusjonelle LSTM, for å predikere reaktiv og aktiv effekt i husholdninger og næringsbygg. På timeoppløsning var deres MAPE-verdier 30-40 %. Deres forklaringsvariabler var dato, time, aktiv og reaktiv effekt, spenning og effektintensitet [34]. Samtidig ser man av Haq mfl. sine prediksjonsplott at de predikerte kurvene følger den faktiske el-bruken tettere enn i denne oppgavens figurer 5.4-5.6. Haq mfl. brukte data over fem år til sine modeller og det er mulig de valgte figurer fra perioder hvor modellen presterte spesielt godt. Til tross for at modellene i denne oppgaven har en enklere arkitektur med kun standard LSTM-lag, presterer de bedre enn modellene fra Haq mfl. [34].

I dette delkapittelet ble et utvalg prediksjonsplott fra modellkonfigurasjonene som ble testet til prediksjon én time frem i tid presentert. I neste delkapittel presenteres resultatene fra prediksjon et døgn frem i tid.

5.4 Prediksjoner et døgn frem i tid

Opprinnelig ble det valgt å gjøre prediksjonene på en relativt kort tidshorisont, én time, for å se hvor godt modellene greide å lære mønstre med minimal prediksjonshorisont. Over lengre tid vil el-bruken variere mer, for eksempel fra sommer til vinter, slik man så i delkapittel 4.2. Dermed vil modellene måtte lære mer komplekse sammenhenger. I tillegg vil man til de fleste praktiske anvendelser være interessert i el-bruk lengre frem i tid enn kun én time, for eksempel dersom man skal planlegge el-bruk for en lengre periode, estimere fleksibilitetspotensiale for de neste seks timene som i Haug [32] eller utføre effekttoppreduksjon.

Det var en svakhet ved modellene at de kun predikerte én time frem i tid og det ble derfor testet om LSTM-algoritmen også kunne predikere timeeffekt [kWh/h] 24 timer frem i tid. Til dette formålet ble basismodellen og tre andre modellkonfigurasjoner testet, modellene III, IV og VI som gjorde det godt til prediksjonen én time frem i tid. I dette delkapittelet blir resultatene fra tre av disse modellene presentert. Fullstendige resultater finnes med plott i vedlegg C og med alle scorene og endringer i prosent fra prediksjonen én time frem i tid i vedlegg B.

I likhet med tidligere modeller presentert i delkapittel 5.3 ble modellene for prediksjon et døgn frem i tid trent i tre runder, og prediksjonsplottene som vises er basert på den første av disse rundene, RNN1. Responsvariabelen ble forskjøvet 24 timer. Ved et tidspunkt og basert på informasjon til og med det tidspunktet, predikerer modellene for tidspunktet et døgn frem i tid. Resultatene viser potensiale for at modellene kan utvides ytterligere i tid. En interessant utvidelse kan være å predikere alle timer i neste døgn fra et gitt tidspunkt.

En oversikt over modellkonfigurasjonene som ble testet til prediksjon av timeeffekt

24 timer frem i tid er vist i tabell 5.4. Modellene er rangert etter RMSE på testsettet, men navnene er de samme som i tabell 5.3 for prediksjon én time frem i tid.

Man ser av tabell 5.4 at modellenes test-RMSE varierer mellom 653 og 765 kWh/h. Dette er det samme intervallet som modellene som ble rangert som VII til XXVI på prediksjon én time frem i tid, og følgelig predikerer modellene i snitt litt dårligere lengre frem i tid. Det er imidlertid individuelle forskjeller mellom modellene. Basismodellen hadde en forbedring på 3 % i test-RMSE for et døgn frem i tid, sammenlignet med én time frem i tid. De øvrige modellene hadde en forverring på 5 % til 22 %, med modell IV som spesielt mye verre lengre frem i tid. Spesielt mellom ukedag og helgedag kan det være stor forskjell i nivået på el-bruk. Dette tatt i betraktning er en under 5 % endring ved prediksjon et helt døgn lengre frem i tid en relativt liten forskjell. En 22 % forverring på en av de beste modellene for kort tidshorison indikerer derimot at modellene som gjør det godt én time frem i tid ikke nødvendigvis gjør det samme et døgn eller lengre frem i tid. Dersom man i fremtiden skal utvide prediksjonsmodellene til lengre tidsrom kan man ikke ta for gitt at de samme modellene vil predikere godt uavhengig av prediksjonshorison eller tidsrom for trening og testing. Det kan gi en indikasjon, men modellene må tilpasses til hver enkelt horison.

Tabell 5.4: Oversikt over navn, forklaringsvariabler, nettverksarkitektur og RMSE-score for trenings- og testsettene for basismodellen XXV og modellkonfigurasjonene III, IV og VI som ble testet til prediksjon et døgn frem i tid. Modellene er rangert etter test-RMSE, men navnet er fra rangeringen av prediksjonsmodellene én time frem i tid. Tomme ruter tilsvarer samme egenskaper som basismodellen.

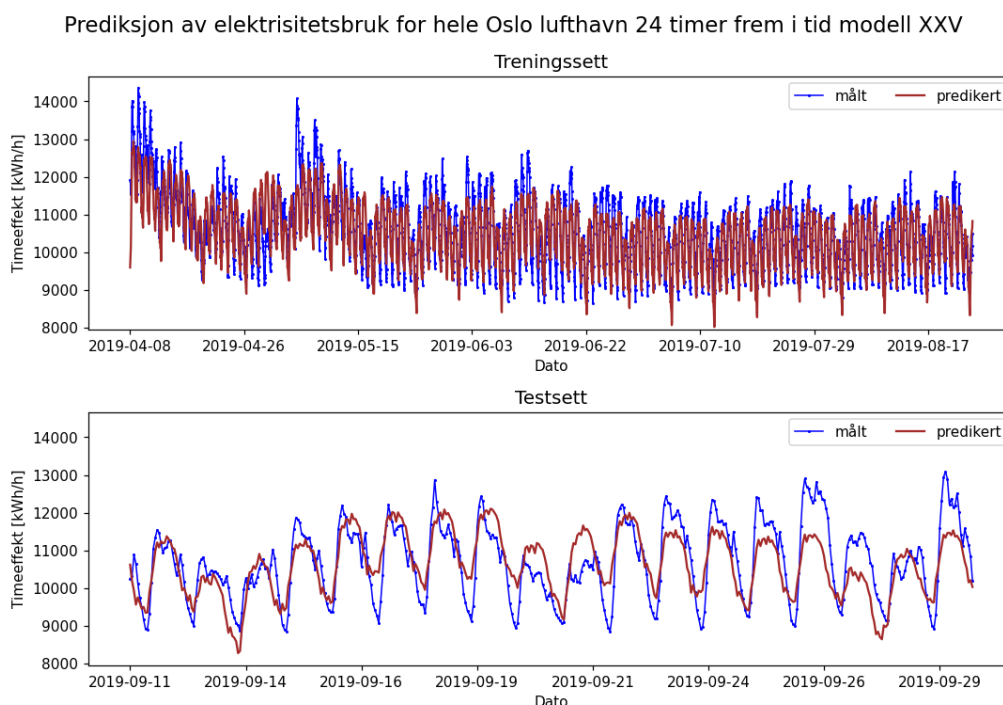
Navn	Forklaringsvariabler	Nettverksarkitektur	Tren.-RMSE [kWh/h]	Test-RMSE [kWh/h]
III		3 LSTM-lag m/64 celler	431	653
VI	Alle inkl. tidligere el-bruk		418	707
XXV	Alle ekskl. tidligere el-bruk	1 LSTM-lag m/32 celler 0 % forenkling	443	741
IV	Alle inkl. tidligere el-bruk	3 LSTM-lag m/64 celler	406	765

Modell IV presterte dårligst av modellene til prediksjon et døgn frem i tid basert på test-RMSE. Den fikk MAPE på 5,5 %, mens den én time frem i tid hadde 4,5 %. Den hadde en dypere arkitektur enn basismodellen og timeeffekt inkludert som forklaringsvariabel, og var med det den mest komplekse av modellene som ble testet med lengre prediksjonshorison. Fra treningscorene i tabell 5.4 ser man at modell IV derimot hadde beste trenings-RMSE. Det indikerer at modellen var mer overtilpasset treningssettet enn øvrige konfigurasjoner. Dette kan være grunnen til den dårlige scoren på testsettet.

Basismodellen

Et prediksjonsplott fra basismodellen fra RNN1 til prediksjon 24 timer frem i tid er vist i figur 5.7. Plottet viser samme tidsperiode som tidligere test- og treningsplott. Prediksjonen fra denne modellen er i stor grad lik som basismodellen for én time frem i tid i figur 5.3. For både trenings- og testperioden fanger basismodellen med lengst prediksjonshorisont opp forskjellene mellom dag og natt og delvis mellom ukedag og helg.

I likhet med basismodellen fra RNN1 for én time frem i tid i figur 5.7 er den ikke god til å predikere tidspunktet for høyeste topp rett, noe som kan være en ulempe i mange anvendelser. Basismodellen hadde en liten forbedring i test-RMSE sammenlignet med basismodellen som predikerer en time frem i tid. Det samme var ikke tilfellet med treningscoren, og kan derfor skyldes tilfeldigheter, ikke en faktisk forskjell.



Figur 5.7: Prediksjonsplott for sammenligning mellom målt og predikert el-bruk [kWh/h] for henholdsvis trenings- og testsettet for basismodellen fra RNN1 som predikerer el-bruk et døgn frem i tid. Modellen treffer i noen grad mønstrene i el-bruken.

I basismodellen for prediksjon 24 timer frem i tid 5.7 ser man den samme forverringen utover i tidsrommet som man så i basismodellen for prediksjon én time frem i tid i figur 5.3. Tidlig i testsettet predikerer basismodellen for prediksjon et døgn frem i tid godt på døgnetts høyeste verdi. De første dagene treffer den samtidig døgnetts laveste, men ikke for alle andre døgn i perioden. Modellen overvurderer få topper, men undervurderer spesielt toppene mot slutten av testperioden. Tidlig i testsettet har modellen et avvik på maksimalt 1 000 kWh/h fra toppene den 18.09, mens mot slutten av perioden er avviket på nærmere 2 000 kWh/h, for eksempel 26.09. Dette indikerer at ved lengre prediksjonshorisont kan forverring i tid bli et større problem.

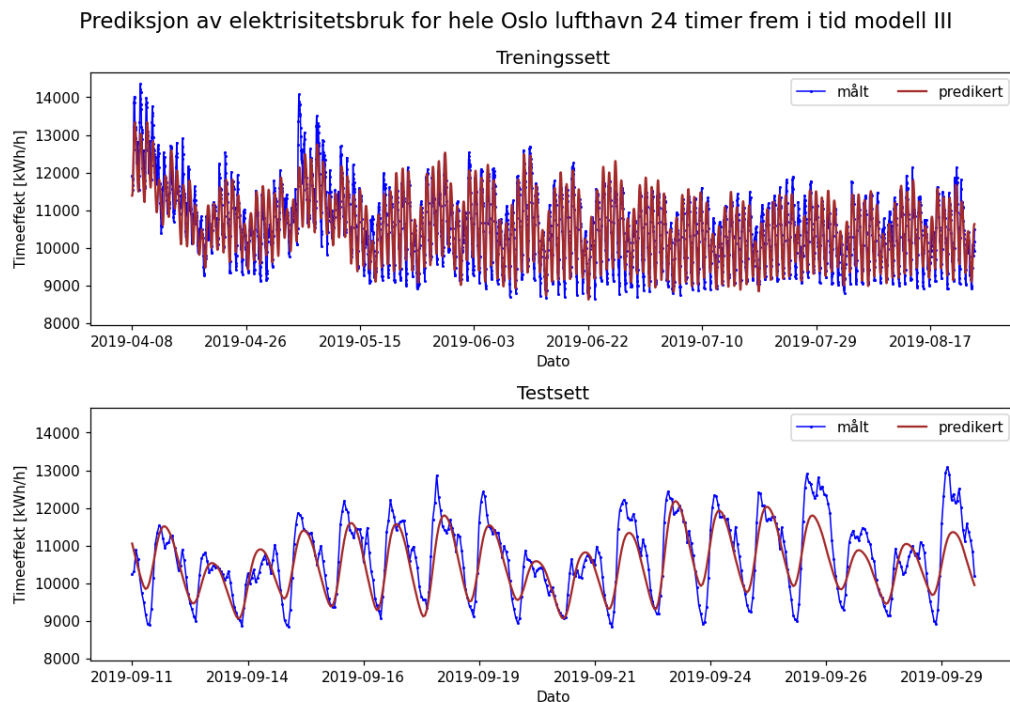
Ikke bare vil det da være en avstand i tid mellom siste oppføring i treningssettet og første oppføring i testsettet, slik det også er for prediksjon én time frem i tid. Det vil også være en lengre avstand mellom timene modellen predikerer basert på, tilbakeblikket, og timen som blir predikert. For lengre avstander mellom forklarende timer og predikert time kan avviket bli desto større.

Modell III

Et prediksjonsplott for modell III anvendt til prediksjon et døgn frem i tid er vist i figur 5.8. Denne modellen har tre skjulte lag med 64 LSTM-celler i hvert lag og samme forklaringsvariabler som basismodellen. Prediksjonen fra denne modellen skiller seg fra basismodellen ved at den predikerer glattere kurver. Dette så man også i de dypere modellene som predikerte én time frem i tid.

Modellen treffer godt i periodene hvor el-bruken øker og minker, og predikerer mange dager godt på døgnetts høyeste verdier og effekttopper. Siden den predikerer så glatt, treffer den ikke toppene på samme tidspunkt som de faktiske verdiene.

Modell III var den beste av modellene som ble testet 24 timer frem i tid, og hadde en MAPE på 4,7 %. Scorene til modell III i tabell 5.4 viser at den har 12 % bedre score enn basismodellen Dette er en mindre forskjell enn mellom basismodellen og modell III én time frem i tid. Dette gir et inntrykk av at modellene for lengre frem i tid predikerer jevnt over dårligere enn modellene med kortere horisont.

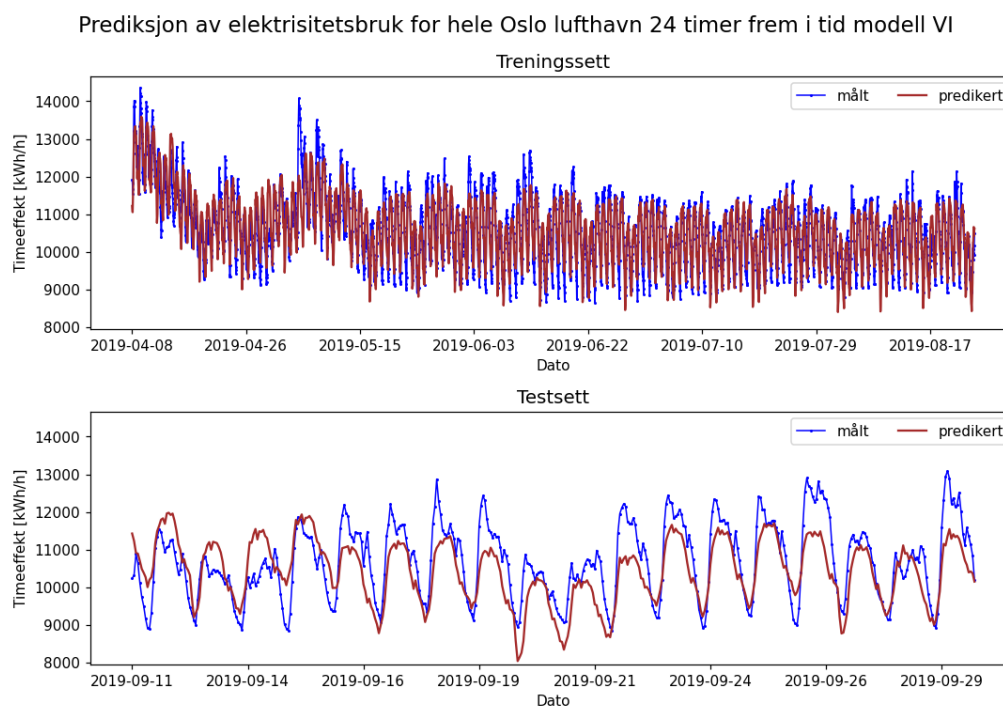


Figur 5.8: Prediksjonsplott for henholdsvis trenings- og testsettet for modellkonfigurasjon III. Modellen predikerer el-bruk et døgn frem i tid.

Modell VI

Et prediksjonsplott for modell VI anvendt til prediksjon et døgn frem i tid er vist i 5.9. Den har samme konfigurasjon som basismodellen. I tillegg til forklaringsvariablene som basismodellen har, har den også tidligere el-bruk som forklaringsvariabel.

Prediksjonen fra denne modellen er i stor grad lik som fra basismodellen vist i figur 5.7. Modell VI er noe bedre på å predikere døgnetts laveste verdier godt, men sliter like mye, om ikke mer, med toppene. Der basismodellen var merkbart bedre i starten, predikerer denne mer likt for alle døgn i testperioden. At den ikke er stort bedre enn basismodellen illustrerer at å legge til tidligere el-bruk ikke automatisk gjør modellene bedre. Dette til tross for at topp elleve-listen for prediksjon én time frem i tid indikerte at tidligere el-bruk var en viktig variabel. For å utnytte variablene til fulle må det kanskje også endringer i arkitekturen til.



Figur 5.9: Prediksjonsplott for henholdsvis trenings- og testsettet for modellkonfigurasjon VI. Modellen predikerer el-bruk et døgn frem i tid.

Oppsummering

Basert på scorene vist i tabell 5.4 var modellene noe dårligere til å predikere timeeffekt et døgn frem i tid, enn én time frem i tid. For prediksjon 24 timer frem i tid fikk basismodellen bedre scorer enn for prediksjon én time frem i tid. De øvrige modellene presterte dårligere på prediksjon 24 timer frem i tid. Modellene som ble testet hadde enten samme forklaringsvariabler som basismodellen eller samme, men inkludert tidligere el-bruk. De hadde enten samme nettverksarkitektur som basismodellen eller et dypere og større nettverk. Den mest komplekse modellen, modell III, predikerte best basert på scorer og predikerte el-bruken bedre også visuelt. Det

indikerer at en mer kompleks modell kan predikere godt lengre frem i tid, men siden den andre mer komplekse modellen ikke viste samme resultat er dette usikkert.

Samtidig viste forskjellene i scorene mellom prediksjonene én og 24 timer frem i tid at fallet i treffsikkerhet ved å predikere lengre frem i tid ikke var stort. For basismodellen og modell III var forskjellen mindre enn 5 % for hver score. For modell VI og IV var forskjellen noe større. Samtidig viste resultatene at de samme modellene ikke alltid gjør det like godt ved ulike tidshorisonter. Ved testing av flere arkitekturer og kombinasjoner av forklaringsvariabler kan det være mulig å finne modeller som predikerer enda bedre også et døgn eller mer frem i tid., men modellene må tilpasses hver enkelt horisont.

Modellene som ble presentert til prediksjon 24 timer frem i tid kan kjøres på et helt døgn. Da vil man ha prediksjoner for hele neste døgn, et døgn i forveien. Det anses også som mulig å utvide metoden og trene en modell for hver time i døgnet, slik at de ved et tidssteg til sammen predikerer hele neste døgn. Det er vist mulig med både rekursive nevrale nettverk og *Support Vector Machines* [32] [79].

Fullstendige resultater fra prediksjon både én time og et døgn frem i tid lagt ved i vedlegg B og C. Videre følger en generell diskusjon av resultatene.

5.5 Generell diskusjon

I denne oppgaven ble det undersøkt om LSTM-modeller kan predikere timeeffekt for Oslo lufthavn Gardermoen én time og et døgn frem i tid. Begge tidshorisonter kan være en del av en fleksibilitetsmetodikk som trenger prognoser av el-bruk henholdsvis én og 24 timer frem i tid. Etter en diskusjon rundt overføringsverdien fra prediksjon på historiske data til fremtidige data, følger av diskusjoner rundt valg som ble tatt i prosessen med datarensing, valg av variabler og valg av algoritme.

5.5.1 Overføringsverdi fra historiske til fremtidige data

En generell svakhet ved resultater fra maskinlæringsmodeller er at de ikke forteller hvordan modellene vil prestere på andre trenings- og testsett enn de som er anvendt. At en modell gir gode resultater på et valgt datasett garanterer ikke gode resultater på fremtidige datasett, men kan gi en indikasjon.

Det er en viss overføringsverdi fra modellene i oppgaven som ble trent på historiske data til fremtidige modeller trent på fortløpende innkommende data. Modellene begrenses dog av el-bruksintervallene de er trent på, og kan for eksempel ikke uten videre predikere mye høyere verdier. De vil trolig heller ikke gjøre det godt hvis de testes på tidsperioder med helt andre mønstre enn de er trent på.

For prediksjon på fortløpende innkommende måleverdier av el-bruk og andre variabler i fremtiden kan det være en fordel å begrense mengden forklaringsvariabler til et minimum. Dette for å begrense mengden prosessering av dataene som må til før

modellene kan kjøres. I neste delkapittel følger en drøfting av valg som ble tatt i forbindelse med pre-prosessering av dataene som ble brukt i denne oppgaven.

5.5.2 Datarensing

Valgene som ble tatt i datarensingprosessen kan ha hatt innvirkning på resultatene. Timeverdiene for el-bruk fra Avinors energi- og miljøoppfølgingsystem Optima [49] med passasjerdata fra Avinor [48] og meteorologiske data fra Meteorologisk institutt (MET) [47] ble kombinert basert på tidsstempel. Mens dataene fra Optima hadde tidsstempel i norsk tid, var dataene fra MET oppgitt i koordinert universaltid (UTC). De sistnevnte dataene ble konvertert til lokal norsk tid, noe som gjorde at enkelte målinger manglet ved tidspunktene hvor klokken ble stilt i løpet av kalenderåret. I både datasettene fra MET og fra Avinor manglet det fra før data ved visse tidspunkter når klokken stilles. Det varierte fra datasett til datasett hvilke tidspunkter som falt ut. For nedbør var det hele døgnet, siden det bare var gitt én måling per døgn. For å sikre at konverteringen ikke førte til forskyving i timeverdiene ble det gjort en manuell sammenligning av temperaturmålinger fra Optimas nettportal og MET. For utvalgte timer i 2017 til 2019 var målingene under 1 °C forskjellige, så det virker ikke som tidssonekonverteringen hadde konsekvenser.

Tidsperioden sommerhalvåret 2019 ble valgt fordi det var relativ lik elektrisitetsbruk fra uke til uke gjennom perioden, og det manglet ingen verdier for noen av forklaringsvariablene. Med algoritmene som er demonstrert i denne oppgaven kan ikke tidsrommet som skal predikeres på timeoppløsning ha manglende verdier for noen av timene i tidsrommet, heller ikke manglende tidspunkter. Da vil ikke disse timene kunne predikeres. Dersom man i fremtiden ønsker å predikere på hele tidsperioden 2017 til 2019 og oppdager manglende verdier må man bruke imputasjon, eller erstatte manglende verdier med gjennomsnitt eller andre metoder [80].

Siden videre anvendelse av de utviklede modellene er innenfor lastflytting og fleksibilitet ble det valgt å bruke aggregert elektrisitetsdata for hele flyplassen. Modellarkitekturerne er såpass generelle at de burde være i stand til å predikere på data fra enkeltmålere også, men dersom modellene skal anvendes på andre datasett, antas det at modellen bør trenes og justeres på nytt.

Til bruk i denne innledende studien av el-brukprediksjon ved Oslo lufthavn ble det som nevnt brukt variabler fra tre ulike kilder. Basert på resultatene i kapittel 5.3 ser ikke alle variablene ut til å være like viktige for en god prediksjon. I neste delkapittel drøftes variablene som ble brukt, med hensyn på nytteverdi for prediksjonene og tilgjengelighet.

5.5.3 Variabelvalg

Variablene som ble brukt i prediksjonene i denne oppgaven er tidsvariabler, tidligere timeeffekt [kWh/h] fra Oslo lufthavn, målinger av nedbør, vindhastighet og lufttemperatur fra Gardermoen målestasjon og antallet passasjerer i terminalene ved Oslo lufthavn. Variablene ble nærmere forklart i delkapittel 4.2. Dataene had-

de ulike oppløsning, og ser fra kapittel 5.3 ut til å ha bidratt ulikt til modellenes treffsikkerhet.

Nedbørsverdiene ble lastet ned med døgnoppløsning, som var høyeste tilgjengelige tidsoppløsning. For å ha timeverdier for også nedbør ble summen av nedbør gjennom døgnet brukt som en konstant for alle timer i døgnet. I virkeligheten vil det ikke være jevn nedbør hele døgnet, og i motsetning til blant annet temperatur og solinnstråling følger ikke nedbøren faste mønstre gjennom døgnet. Man kunne sett for seg at man ved hjelp av en modell kunne fordelt nedbørsmengden per døgn utover timene i døgnet. Denne typen modeller finnes for større nedbørshendelser, men ikke for normale mengder nedbør [81].

Basert på undersøkelser av mønstre i nedbør og vindhastighet sammenlignet med el-bruk i delkapittel 4.3 og Pearson-korrelasjonen som ble vist i delkapittel 5.1 virker det ikke å være en sterk sammenheng mellom disse variablene og el-bruk ved flyplassen. Modellkonfigurasjon X fra tabell 5.3 er lik som basismodellen, men vindhastighet og sum nedbør per døgn er tatt ut som forklaringsvariabler. Modell X hadde 11 % bedre test-RMSE enn basismodellen. Visuelt var det mindre forskjell. Likevel indikerer resultatet at å fjerne nedbør og vindhastighet forbedret modellen. Hwang mfl. [7] inkluderte nedbør og vindhastighet som forklaringsvariabel i sine modeller, men i kombinasjon med andre meteorologiske variabler [7]. Det samme gjelder andre prediksjoner av el-bruk nevnt i delkapittel 4.3.1. Dermed anses nedbør og vindhastighet som mindre viktige variabler å ha med videre i modeller av el-bruk ved OSL.

Som nevnt i delkapittel 4.3.1 kan det derimot være interessant å inkludere luftfuktighet i senere modeller. Det finnes målinger av gjennomsnittlig relativ luftfuktighet per døgn fra Gardermoen målestasjon for tidsrommet april 2012 til november 2020, og siden november 2020 finnes målinger med timeoppløsning. Dette ble ikke registrert i arbeidet med nedlasting av data og derfor heller ikke benyttet.

En annen forklaringsvariabel som kunne vært mer interessant å ha med er solinnstråling. Gardermoen målestasjon, som de øvrige de andre meteorologiske variablene ble hentet fra, har ikke måler for solinnstråling. Romundstad [82] estimerte i 2014 grunnlaget for solfangere på Gardermoen. Til det ble det brukt solinnstrålingsverdier med timeoppløsning fra fem målestasjoner fra Bioforsk og NM-BU, inntil 60 km fra flyplassen, og verdier fra fire solstrålingsdatabaser for PV-tjenester. Disse ble aggregert til måneds- og årsoppløsning og sammenlignet. Det ble estimert at det var 10 % forskjell mellom målingsstasjonene med størst forskjell og litt mer mellom databasene med størst forskjell. Hvordan solinnstrålingen endret seg gjennom året var sammenlignbart fra estimat til estimat [82]. Det er mulig at målinger fra hvilken som helst av disse kunne gjort nytten til el-bruksprediksjon.

Basert på undersøkelsene i kapittel 4.3 og Pearson-korrelasjonen som ble vist i delkapittel 5.1 har lufttemperatur en viss korrelasjon med el-bruk, men mye av korrelasjonen så ut til å være sesongbasert. Lufttemperatur og el-bruk har samme oppløsning. Inkludert modellkonfigurasjonene som hadde samme forklaringsvariabler som basismodellen, hadde sju av de elleve beste modellene til prediksjon én time frem i tid i tabell 5.3 med lufttemperatur som forklaringsvariabel. Det ble også testet modell-

konfigurasjoner som hadde kun værvariablene og enten kun time på døgnet eller alle tidsvariabler som forklaringsvariabler. Disse presterte ikke tydelig bedre enn basismodellen. Dette tyder på at vær-variabler alene ikke gir de beste prediksjonene, heller ikke i kombinasjon med tidsvariabler. Det er mulig at ved prediksjon av enkeltkomponenter ved flyplassen kan temperatur som forklaringsvariabel være enda viktigere, for eksempel dersom man skal predikere deler av bygningene som har all el-bruk til ventilasjon og oppvarming.

Åtte av de elleve beste modellkonfigurasjonene hadde passasjerantall som forklaringsvariabel, så variabelen må ha bidratt noe til prediksjonene. Ifølge Alba og Mañana [60] som skrev om el-bruk på flyplasser kommer toppene i el-bruken ikke av passasjerene i seg selv, men av alt utstyret som må startes opp for å yte nødvendige og ønskede tjenester til passasjerene. Fra den høye grunnlasten i el-brukfigurene i kapittel 4.2 ser man at en stor del av el-bruken er konstant på ca. 9 000 kWh/h uavhengig av tid på døgnet og passasjerantall. Forklaringsvariabelen passasjerantall hadde ikke samme oppløsning som el-bruken, og passasjerantall per uke kan ikke si noe om hvordan el-bruken endres gjennom døgnet. Man så fra prediksjonsplottet i figur 5.6 at modell XI med kun passasjerdata ikke hadde høy visuell måloppnåelse, så passasjerantallet per uke alene kan ikke ha gitt nok informasjon til å gi gode prediksjoner alene. I stedet ble passasjerantall antatt å være et mål på hvordan aktiviteten endres gjennom året. For å kunne forbedre modellene mer basert på passasjerdata må man trolig ha høyere oppløsning på målingene.

Det kunne også vært interessant å se om andre mål på aktivitet ved flyplassen påvirker prediksjonene. Gjennom døgnet, uken og året varierer antallet passasjerer på flyplassen, men også antallet som er på jobb på flyplassen og trolig hvilke elektriske laster som er i bruk. Samtidig varierer nivået på bil-, tog- og annen trafikk til og fra flyplassen, inkludert antallet flyavganger. Målinger av denne typen, eller informasjon om åpningstider for tjenester, arbeidstider for ansatte i ulike deler av flyplassen, timeplaner for tekniske anlegg eller andre typer innsikt i hvilken type aktivitet som foregår når kan derfor vurderes undersøkt og inkludert.

Basert på undersøkelsene av mønstre i el-bruk på ukedager versus helgedager og helligdager versus alle andre dager, ble markører på om en time tilhørte en helg eller helligdag inkludert som forklaringsvariabel. Man så fra undersøkelsene av passasjerdata i delkapittel 4.3.2 at passasjerantallet ofte er høyt i uken(e) før felles ferieavvikling for fylkene nær OSL. Hvilke uker dette inntreffer, i forbindelse med påsken spesielt, varierer fra år til år. Det kan tenkes at passasjerantallet i seg selv reflekterer dette godt nok, men det kunne også vært interessant å se om en markør på om en time tilhører en uke før, under eller etter en ferieuke påvirker prediksjonen. Oslo lufthavn er den viktigste flyplassen for både charter- og rutetrafikk for hovedstadsområdet [54], og det flyr betydelige mengder mennesker derfra som skal på reise i forbindelse med både jobb og fritid. Siden sommerhalvåret har så stor andel ferieuker ville det å inkludere ferieuker kanskje ikke hjulpet på dette testsettet. I større datasett som inkluderer større deler av året, med både vanlige uker, skoleferieuker og fellesferieuker, ser man for seg at ferieukene har større påvirkning på hvilke tider det brukes mye elektrisitet. Da kan en slik markør være mer relevant.

Listen over konfigurasjoner som fikk best score i tabell 5.3 viste at ved mange forklaringsvariabler trengs mer komplekse nettverk for å gi gode prediksjoner. Dersom man inkluderer flere forklaringsvariabler bør man derfor også teste både enkle og mer komplekse nettverk.

5.5.4 Algoritmevalg

Valget av algoritme falt på *Long Short-Term Memory* (LSTM) fordi rekursive nevrale nettverk er vist å være velfungerende på sekvensielle data [29], og denne oppgaven hadde som mål å predikere tidsserier av el-bruk. Videre er LSTM er en videreutvikling av RNN som fungerer bedre på sekvenser med lange tidsspekter [29] [41], noe man også hadde i dataene til oppgaven.

I forskning brukes det, som vist i kapittel 2, flere ulike algoritmer til samme formål som i denne oppgaven. Fra klustering, regresjonstrær [83] og *Support Vector Machine*-regresjon (SVR) til ulike typer nevrale nettverk, og en mengde kombinasjoner og hybrider av disse og flere [39]. Det finnes eksempler på at utvidelser av RNN-modeller fungerer bedre enn enkle versjoner. For eksempel kombinerte Rafi mfl. [41] LSTM-lag med et konvulsjonelt nevralt nettverk for å predikere elektrisitetsbehovet i Bangladesh' kraftsystem med halvtime-oppløsning. De rapporterte at den kombinerte modellen fikk en MAPE på 4-6 %, mens en vanlig LSTM fikk opptil 13 % avvik. Det samme er vist for andre maskinlæringsalgoritmer. Hong [84] presenterte for eksempel en hybridmodell som kombinerte elementer fra rekursive nevrale nettverk med SVR og *Artificial Bee Colony*-modeller (ABC) for å predikere en månedlig last i nord-østlige Kina. Hybridmodellen fikk en MAPE på 3 %, bedre enn en enklere SVRABC med MAPE på 3,5 % [84]. Selv om det er gjort mye forskning på kombinasjoner av nettverk og modelltyper er det ikke rett frem å implementere disse løsningene på egne data.

I en vurdering av forskning på feltet skrev Hong og Wang [5] at kunstige nevrale nettverk har vært den mest populære teknikken innenfor litteratur på lastprognoser de siste tre tiårene. Samtidig er de færreste av modellene fra litteraturen implementert til å gjøre prognoser av elektriske laster i praksis, for eksempel innenfor industri og kraftnett. Hong og Wang [5] skrev at mange utvikler algoritmer som er for spesifikt tilpasset visse problemstillinger og datasett til at modellene gagnar maskinlæringsfeltet som helhet og kan anvendes på andre problemer av nybegynnere og andre [5]. Abbass og Hamdy [30] skrev også at det er for få løsninger som enkelt kan implementeres av brukere som ikke er maskinlæringseksperter og forskere [30]. I tillegg fokuserer mye av litteraturen på å predikere el-bruken til store energisystemer eller brukere. Det var deler av motivasjonen Slowik og Urban [42] hadde da de implementerte en LSTM-basert modell for å predikere energibruken til en mindre sluttbruker, en fabrikk in et mikrogrid [42].

Utvalget av maskinlæringsalgoritmer er altså stort, men komplekse algoritmer kan kreve mye justering for å fungere godt. Relativt enkle LSTM-modeller ble derfor valgt i denne innledende studien. Fremfor å justere modellene til perfektjon ble det fokusert på å teste ulike konfigurasjoner for å kunne gi anbefalinger til videre arbeid

med prediksjonsmodeller for el-bruk ved OSL.

Datasettet som LSTM-modellene ble trent på i dette casestudiet er ganske kort, så andre og enklere algoritmer kunne også fungert. Sammenlignet med andre bruksområder for dyp læring, for eksempel innenfor bioinformatikk og medisinsk fysikk, har datasettet i denne oppgaven få forklaringsvariabler. Mange av studiene på lastprediksjon som er presentert brukte også flere år med data, for eksempel Hwang mfl. [7] og Haq mfl. [34], mens det i denne oppgaven ble brukt seks måneder. Dette begrenset mengden informasjon algoritmen måtte prosessere, og gjorde også treningstiden relativt kort.

Utforskingen av data i kapittel 4 viste at elektrisitetsbruken fulgte visse mønstre, og at den i tidsperioden som ble testet ikke varierte så mye. Kanskje kan man få gode resultater ved å predikere fremtidig timeeffekt som et gjennomsnitt for hver time på lignende ukedager eller ved hjelp av en *Autoregressive Integrated Moving Average*-modell (ARIMA). Sistnevnte har i flere tiår vært et mye brukt verktøy for å lage prognoser basert på historiske tidsserier [85]. ARIMA har ikke vært undersøkt i forbindelse med denne oppgaven, men det er mulig den typen metode også kunne gitt gode resultater.

Litteratur om lastprognoser prefererer enklere modeller med mindre det er god grunn til å velge en mer kompleks modell [86]. Det samme gjelder innenfor maskinlæring generelt; en mer kompleks modell er ikke nødvendigvis bedre. Det viser også resultatene i denne oppgaven, hvor en modell med få forklaringsvariabel og enkel arkitektur viste seg å være best. LSTM-algoritmen er vist å gi gode resultater på lignende problemstillinger som i denne oppgaven, den håndterer sammenhenger i tid og har mulighetet for å bli utviklet med økt kompleksitet og tilpasning til problemstillingen. Det synes derfor å ha vært et godt valg av algoritme.

Kapittel 6

Konklusjon og videre arbeid

6.1 Konklusjon

Denne oppgaven hadde som mål å kartlegge hvordan maskinlæringsalgoritmer fungerer til å predikere elektrisitetsbruken ved Oslo lufthavn Gardermoen (OSL). Basert på litteratursøk av lignende problemstillinger ble rekursive nevralt nettverk av typen *Long Short-Term Memory* (LSTM) brukt til å lage prognosene. Oppgaven viser at LSTM-algoritmen kan gi gode estimater på bruken av elektrisitet ved flyplassen.

Undersøkelser av elektrisitetsmålinger fra OSL fra årene 2017 til 2019 viste at det var gjentakende mønstre i el-bruken ved flyplassen på både døgn-, sesong- og årsbasis. Dette indikerte at elektrisitetsbruken var predikerbar.

Basert på variabelvalg i lignende prediksjoner av el-bruk ble målinger av lufttemperatur, vindhastighet, nedbør og passasjerantall undersøkt og brukt i prediksjonene. Det ble undersøkt hvordan mønstrene i el-bruken samsvarte med variablene. Sammenhengen var mest tydelig mellom lufttemperatur og el-bruk. Tidsskalaene i mønstrene i el-bruk førte til at kalendervariablene time på døgnet, ukenummer og ukedag også ble inkludert.

For å kartlegge hvordan LSTM-algoritmen fungerer til prediksjon av timeeffekt [kWh/h] for hele flyplassen ble modellkonfigurasjoner med ulike kombinasjoner av forklaringsvariabler og nettverksarkitektur testet. Modeller som predikerte el-bruk både én time og et døgn frem i tid ble undersøkt for å se hvordan en utvidet prediksjonshorisont påvirket resultatene. Alle modellene ble anvendt på data fra sommerhalvåret 2019.

De 38 modellkonfigurasjonene som ble testet til prediksjon én time frem i tid hadde MAPE-avvik på 3,7-8,9 %. Modellene traff godt på døgnvariasjon og variasjoner gjennom uken, men traff i mindre grad toppen i el-bruken. En basismodell med ett LSTM-lag med 32 enheter og tidsvariabler, meteorologiske variabler og passasjerantall som forklaringsvariabler predikerte el-bruken til flyplassen med 6 % MAPE-avvik. Mange av modellene hadde lignende resultater som basismodellen, noe som viser at mange ulike konfigurasjoner kan gi like gode prediksjoner. Det ble vist at

modeller som enten var komplekse i arkitekturen og hadde mange forklaringsvariabler, eller som var enkle i arkitekturen og hadde færre forklaringsvariabler, ga de beste prediksjonene. Dette viser at både valget av forklaringsvariabler og nettverksarkitektur er viktige.

Den beste av de 38 modellekonfigurasjonene fikk 3,7 % MAPE-avvik og hadde også visuelt de mest treffsikre prediksjonene. Modellen hadde samme arkitektur som basismodellen, men kun time på døgnet og tidligere el-bruk som forklaringsvariabler. Majoriteten av de ti beste modellene hadde tidligere el-bruk som variabel. Det viser at målinger av tidligere el-bruk var den mest informative forklaringsvariabelen.

Det ble også vist at blant de meteorologiske verdier bidro lufttemperatur mest til å bedre modellene, mens vindhastighet og nedbør var mindre viktige. Antallet passasjerer i flyplassens terminaler per uke bidro noe, men med samme oppløsning som elektrisitetsverdiene vil passasjererantallet trolig ha en større innvirkning på prediksjonene.

Til prediksjon et døgn frem i tid ble det testet fire av modellkonfigurasjonene som gjorde det best til prediksjon én time frem i tid. Disse hadde MAPE-avvik på 4,7-5,5 % og var i likhet med ved kortere tidshorisont bedre til å predikere døgn- og ukesvariasjon enn topper el-bruken. Reduksjonen i treffsikkerhet ved prediksjon 24 timer lengre frem i tid var liten, noe som viser potensiale for ytterligere utvidelser i prediksjonshorisont.

Undersøkelsene av egnetheten til LSTM-algoritmen for el-bruksprediksjon og innvirkningen av forklaringsvariablene og nettverksarkitekturen på modellene som er beskrevet i denne oppgaven legger grunnlaget for videre arbeid med prediksjonsmodeller for OSL. Disse modellene kan på sikt bidra til å estimere mulighetene for fleksibel el-bruk og reduksjon av effekttopper ved Oslo lufthavn, og dermed bidra til både økonomiske besparelser for flyplassen og minsket belastning i kraftnettet som OSL er tilkoblet.

6.2 Videre arbeid

Selv om resultatene i denne oppgaven gir indikasjoner på hvilke forklaringsvariabler og nettverksarkitekturer som fungerer til prediksjon av el-bruk ved Oslo lufthavn, trengs det ytterligere utforskning før modellene kan anvendes og inkluderes i for eksempel optimeringsrammeverk for el-bruk og fleksibilitet.

Først og fremst må modellene testes på utvidede datasett, for å se om de er generaliserbare og fungerer godt på andre tidsrom enn sommerhalvåret 2019 som ble testet i denne oppgaven.

Det ble vist at en generell svakhet ved modellene er at de ikke predikerte topper i el-bruken godt. Dette kan være grunnet manglende informasjon i forklaringsvariablene, men også at modellene ikke er i stand til å finne relasjoner i dataene. Endringer i nettverksarkitekturen kan gjøre prediksjonene mer treffsikre. Det kan undersøkes hvorvidt andre evalueringsformer er mer hensiktsmessige for måloppnåelsen, og om

større nettverk oppfører seg annerledes. I tillegg kan det undersøkes om prediksjonene blir bedre dersom man ikke prøver å predikere bunner og generelle mønstre time for time, kun daglige effekttopper. Man kan teste å kun beholde topp-tidspunktene i datasettet, eller dele toppene inn i høye, middels og lave og behandle det som et klassifiseringsproblem.

Videre ble flere forklaringsvariabler foreslått, og det kan undersøkes om disse bidrar positivt til modellenes treffsikkerhet. Det gjelder solinnstråling, luftfuktighet, informasjon om ferieuker og mer detaljerte passasjerdata. Andre mål på flyplassens aktivitetsnivå, for eksempel trafikkstatistikk eller antallet ansatte på jobb, kan også vurderes.

Referanser

- [1] Priyadarshi R. Shukla mfl. *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK New York, NY, USA: Cambridge University Press, 2022. DOI: 10.1017/9781009157926.
- [2] Saleh Seyedzadeh mfl. «Machine learning for estimation of building energy consumption and performance: a review». I: *Visualization in Engineering* 6.5 (2018), s. 20. DOI: 10.1186/s40327-018-0064-7.
- [3] Olje- og energidepartementet. *Bærekraftige bygg - Energifakta Norge*. Hentet 27.11.2022. 2019. URL: <https://energifaktanorge.no/et-baerekraftig-og-sikkert-energisystem/baerekraftige-bygg/>.
- [4] Norges vassdrags- og energidirektorat. *Energieffektivisering*. Hentet 28.11.2022. 2021. URL: <https://nve.no/energi/energisystem/energibruk-effektivisering-og-teknologier/energieffektivisering/>.
- [5] Tao Hong og Pu Wang. «Artificial Intelligence for Load Forecasting». I: *IEEE power and energy* 20.3 (2022), s. 10. DOI: 10.1109/MPE.2022.3150808.
- [6] Lennart Söder mfl. «A review of demand side flexibility potential in Northern Europe». I: *Renewable and Sustainable Energy Reviews* 91 (2018), s. 654–664. DOI: doi.org/10.1016/j.rser.2018.03.104.
- [7] Junhwa Hwang, Dongjun Suh og Marc-Oliver Otto. «Forecasting Electricity Consumption in Commercial Buildings Using a Machine Learning Approach». I: *Energies* 13.22 (2020), s. 29. DOI: 10.3390/en13225885.
- [8] Norges miljø- og biovitenskapelige universitet. *Nettbalansering fra store parkeringsanlegg og næringsbygg - NeX2G*. Hentet 12.12.2022. 2021. URL: <https://nmbu.no/forside/prosjekter/nex2g>.
- [9] Olje- og energidepartementet. *Strømnettet - Energifakta Norge*. Hentet 28.10.2022. 2019. URL: <https://energifaktanorge.no/norsk-energiforsyning/kraftnett/>.
- [10] scikit-learn-utviklere. *MinMaxScaler*. Hentet 01.10.2022. 2022. URL: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html>.
- [11] Olje- og energidepartementet. *Forsyningssikkerhet - Energifakta Norge*. Hentet 11.12.2022. 2019. URL: <https://energifaktanorge.no/norsk-energiforsyning/forsyningssikkerhet/>.

- [12] Norges vassdrags- og energidirektorat. *Kraftproduksjon*. Hentet 30.10.2022. 2019. URL: <https://nve.no/energi/energisystem/kraftproduksjon/>.
- [13] Europakommisjonen. *The European Green Deal*. COM(2019) 640 final. Brussels, 2019.
- [14] Statnett. *Langsiktig markedsanalyse - Norden og Europa 2020-2050*. Tekn. rapp. 2020.
- [15] Knut Hofstad, Jo Halvard Halleraker og Tor Haakon Bakken. *Vannkraft i Store norske leksikon*. Hentet 01.11.2022. 2022. URL: <https://http://snl.no/vannkraft>.
- [16] International Energy Agency. *Status of Power System Transformation 2018*. 2018. DOI: 10.1787/9789264302006-en.
- [17] Jaquelin Cochran mfl. *Flexibility in 21st Century Power Systems*. 2014. DOI: 10.2172/1130630.
- [18] Statnett. *Reservemarkeder*. Hentet 01.11.2022. 2020. URL: <https://statnett.no/for-aktorer-i-kraftbransjen/systemansvaret/kraftmarkedet/reservemarkeder/>.
- [19] Norges vassdrags- og energidirektorat. *Store investeringer i strømmettet*. Hentet 01.11.2022. 2020. URL: https://nve.no/energi/energisystem/energibruk-effektivisering-og-teknologier/forbrukerfleksibilitet/store-investeringer-i-stromnettet/#_ftn1.
- [20] Austin R. Coffman, Zhong Guo og Prabir Barooah. «Characterizing Capacity of Flexible Loads for Providing Grid Support». I: *IEEE Transactions on Power Systems* 36.3 (2021), s. 2428–2437. DOI: 10.1109/TPWRS.2020.3033380.
- [21] Moslem Uddin mfl. «A review on peak load shaving strategies». I: *Renewable and Sustainable Energy Reviews* 82 (2018), s. 3323–3332. DOI: 10.1016/j.rser.2017.10.056.
- [22] Norges vassdrags- og energidirektorat. *Strømmettet må alltid være i balanse*. Hentet 01.11.2022. 2020. URL: <https://nve.no/energi/energisystem/energibruk-effektivisering-og-teknologier/forbrukerfleksibilitet/stromnettet-ma-alltid-vaere-i-balanse/>.
- [23] Mohammad Hassan Hashemi. *Energy Management of Industrial Loads In a Smart Micro grid Using Particle Swarm Optimization Algorithm*. 2015.
- [24] International Energy Agency. *Demand Response*. Hentet 01.12.2022. 2022. URL: <https://iea.org/reports/demand-response>.
- [25] Norges vassdrags- og energidirektorat. *Forbrukerfleksibilitet*. Hentet 01.11.2022. 2021. URL: <https://nve.no/energi/energisystem/energibruk-effektivisering-og-teknologier/forbrukerfleksibilitet/>.
- [26] NVE - RME. *Nettleie*. Hentet 31.10.2022. 2015. URL: <https://nve.no/reguleringsmyndigheten/kunde/nett/nettleie/>.
- [27] Shuang Dai mfl. «Electrical peak demand forecasting: A review». I: *arXiv e-prints* (2021). DOI: 10.48550/arXiv.2108.01393.

- [28] Muhammad Qamar Raza og Abbas Khosravi. «A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings». I: *Renewable and Sustainable Energy Reviews* 50 (2015), s. 1352–1372. DOI: 10.1016/j.rser.2015.04.065.
- [29] Sebastian Raschka og Vahid Mirjalili. *Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-learn, and TensorFlow*. Expert insight. Packt Publishing, 2017.
- [30] Mahmoud Abdelkader Bashery Abbass og Mohamed Hamdy. «A Generic Pipeline for Machine Learning Users in Energy and Buildings Domain». I: *Energies* 14.17 (2021), s. 5410. DOI: 10.3390/en14175410.
- [31] Bernard Widrow. *An Adaptive "Adaline Neuron Using Chemical "Memistors"*. Tekn. rapp. 1553-2, B. Stanford Electron Labs, 1960.
- [32] Martin Haug. «Methodology for Assessing Short-term Flexibility in Demand-side Assets». Masteroppgave, NMBU. 2019.
- [33] Sepp Hochreiter og Jürgen Schmidhuber. «Long Short-Term Memory». I: *Neural Computation* 9.8 (1997), s. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.
- [34] Ijaz Ul Haq mfl. «Sequential Learning-Based Energy Consumption Prediction Model for Residential and Commercial Sectors». I: *Mathematics* 9.6 (2021), s. 17. DOI: 10.3390/math9060605.
- [35] Martin Abadi mfl. «TensorFlow: A system for large-scale machine learning». I: *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*. Red. av Google Brain. USENIX Association, s. 265–283. URL: usenix.org/conference/osdi16/technical-sessions/presentation/abadi.
- [36] Patrick Schneider og Fatos Xhafa. «Chapter 3 - Anomaly detection: Concepts and methods». I: *Anomaly Detection and Complex Event Processing over IoT Data Streams*. Academic Press, 2022, s. 49–66. DOI: 10.1016/B978-0-12-823818-9.00013-4.
- [37] «Mean Absolute Error». I: *Encyclopedia of Machine Learning*. Red. av Claude Sammut og Geoffrey I. Webb. Boston, MA: Springer US, 2010, s. 652–652. DOI: 10.1007/978-0-387-30164-8_525.
- [38] Aulon Shabani og Orion Zavalani. «Predicting Building Energy Consumption using Engineering and Data Driven Approaches: A Review». I: *European Journal of Engineering Research and Science* 2.5 (2017), s. 7. DOI: 10.24018/ejers.2017.2.5.352.
- [39] Seyedeh Narjes Fallah mfl. «Computational Intelligence Approaches for Energy Load Forecasting in Smart Energy Management Grids: State of the Art, Future Challenges, and Research Directions». I: *Energies* 11.3 (2018), s. 31. DOI: 10.3390/en11030596.
- [40] Carsten Croonenbroeck og Georg Stadtmann. «Renewable generation forecast studies – Review and good practice guidance». I: *Renewable and Sustainable Energy Reviews* 108 (2019), s. 312–322. DOI: 10.1016/j.rser.2019.03.029.

- [41] Shafiul Hasan Rafi mfl. «A Short-Term Load Forecasting Method Using Integrated CNN and LSTM Network». I: *IEEE Access* 9 (2021), s. 32436–32448. DOI: 10.1109/ACCESS.2021.3060654.
- [42] Maciej Slowik og Wieslaw Urban. «Machine Learning Short-Term Energy Consumption Forecasting for Microgrids in a Manufacturing Plant». I: *Energies* 15.9 (2022), s. 16. DOI: 10.3390/en15093382.
- [43] Jin S. Hwang mfl. «Optimal ESS Scheduling for Peak Shaving of Building Energy Using Accuracy-Enhanced Load Forecast». I: *Energies* 13.21 (2020). DOI: 10.3390/en13215633.
- [44] Khikmafari Yudiantaka, Jung-Su Kim og Hwachang Song. *Dual Deep Learning Networks Based Load Forecasting with Partial Real-Time Information and Its Application to System Marginal Price Prediction*. 2020. DOI: 10.3390/en13010148.
- [45] Pedro A. González og Jesús M. Zamarréno. «Prediction of hourly energy consumption in buildings based on a feedback artificial neural network». I: *Energy and Buildings* 37.6 (2005), s. 595–601. DOI: 10.1016/j.enbuild.2004.09.006.
- [46] Kjersti Rustad Kvisberg. *predicting-electricity-usage*. 2022. URL: <https://github.com/krkvis/predicting-electricity-usage>.
- [47] Meteorologisk institutt. *Frost API*. Hentet 21.09.2022. 2022. URL: <https://frost.met.no/howto.html>.
- [48] Ida Myrvold. *Terminalpassasjerer ved Oslo lufthavn*. Pers. komm. 06.09.22. 2022.
- [49] Ørn Software. *Optima gir deg full kontroll over energibruken*. Hentet 15.08.2022. 2022. URL: <https://ornsoftware.com/produkter/optima/overblikk>.
- [50] «Pearson’s Correlation Coefficient». I: *Encyclopedia of Public Health*. Red. av Wilhelm Kirch. Dordrecht: Springer Netherlands, 2008, s. 1090–1091. DOI: 10.1007/978-1-4020-5614-7_2569.
- [51] Fabian Pedregosa mfl. «Scikit-learn: Machine learning in Python». I: *The Journal of Machine Learning Research* 12 (2011), s. 2825–2830. DOI: 10.5555/1953048.2078195.
- [52] Pranava Madhyastha og Rishabh Jain. «On Model Stability as a Function of Random Seed». I: *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*. Association for Computational Linguistics, s. 929–939. DOI: 10.18653/v1/K19-1087.
- [53] Diederik P. Kingma og Jimmy Ba. «Adam: A Method for Stochastic Optimization». I: *3rd International Conference for Learning Representations*. URL: arxiv.org/abs/1412.6980v9.
- [54] Lars Engerengen, Erik Tandberg og Inge S. Kristiansen. *Oslo lufthavn Gardermoen i Store norske leksikon*. Hentet 07.11.2022. URL: https://snl.no/Oslo_lufthavn_Gardermoen.

- [55] Rasmus Tveitane. «Fleksibilitet i parkerte elbiler ved næringsbygg: En case-studie av Oslo lufthavn Gardemoen». Masteroppgave, NMBU. 2021.
- [56] Wiggo Roar Dehli. *Spørsmål om Oslo lufthavn og økonomi*. Pers. komm. 06.12.22. 2022.
- [57] OpenStreetMap-bidragstere. *Skjermdump av Oslo-regionen fra openstreet-map.org*. Hentet 28.11.2022. 2022. URL: <https://openstreetmap.org/way/425295012#map=13/60.1978/11.1206>.
- [58] Ansatte i departementene SMK og DSS. *Tidslinje: myndighetenes håndtering av koronasituasjonen på regjeringen.no*. Hentet 01.12.2022. 2022. URL: <https://regjeringen.no/no/tema/Koronasituasjonen/tidslinje-koronaviruset/id2692402/>.
- [59] Avinor AS. *Oslo lufthavn - Avgang og Ankomst*. Hentet 29.11.2022. 2022. URL: <https://avinor.no/flyplass/oslo/flyinformasjon/avganger/>.
- [60] Sergio Alba og Mario Mañana. «Characterization and Analysis of Energy Demand Patterns in Airports». I: *Energies* 10 (2017), s. 119. DOI: 10.3390/en10010119.
- [61] Helligdagsfredloven. *Lov om helligdager og helligdagsfred (LOV-1995-02-24-12)*. 1995. URL: <lovdata.no/lov/1995-02-24-12>.
- [62] Radu Platon, Vahid Raissi Dehkordi og Jacques Martel. «Hourly prediction of a building's electricity consumption using case-based reasoning, artificial neural networks and principal component analysis». I: *Energy and Buildings* 92 (2015), s. 10–18. DOI: doi.org/10.1016/j.enbuild.2015.01.047.
- [63] Giouli K. Mihalakakou, Mattheos Santamouris og Aris E. Tsangrassoulis. «On the energy consumption in residential buildings». I: *Energy and Buildings* 34.7 (2002), s. 10. DOI: 10.1016/S0378-7788(01)00137-2.
- [64] Li Qiong, Ren Peng og Meng Qinglin. «Prediction model of annual energy consumption of residential buildings». I: *2010 International Conference on Advances in Energy Engineering*, s. 223–226. DOI: 10.1109/ICAEE.2010.5557576.
- [65] Joaquim Massana mfl. «Short-term load forecasting in a non-residential building contrasting models and attributes». I: *Energy and Buildings* 92 (2015), s. 322–330. DOI: 10.1016/j.enbuild.2015.02.007.
- [66] Abhishek Srivastav, Ashutosh Tewari og Bing Dong. «Baseline building energy modeling and localized uncertainty quantification using Gaussian mixture models». I: *Energy and Buildings* 65 (2013), s. 438–447. DOI: 10.1016/j.enbuild.2013.05.037.
- [67] S. Karatasou, M. Santamouris og V. Geros. «Modeling and predicting building's energy use with artificial neural networks: Methods and results». I: *Energy and Buildings* 38.8 (2006), s. 949–958. DOI: 10.1016/j.enbuild.2005.11.005.

- [68] Rishee K. Jain mfl. «Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy». I: *Applied Energy* 123 (2014), s. 168–178. DOI: 10.1016/j.apenergy.2014.02.057.
- [69] Subodh Paudel mfl. «Pseudo dynamic transitional modeling of building heating energy demand using artificial neural network». I: *Energy and Buildings* 70 (2014), s. 81–93. DOI: 10.1016/j.enbuild.2013.11.051.
- [70] Hai-Xiang Zhao og Frédéric Magoulès. «Feature Selection for Predicting Building Energy Consumption Based on Statistical Learning Method». I: *Journal of Algorithms and Computational Technology* 6.1 (2012), s. 59–77. DOI: 10.1260/1748-3018.6.1.59.
- [71] Michael C. Burkhart, Yeonsook Heo og Victor M. Zavala. «Measurement and verification of building systems under uncertain data: A Gaussian process modeling approach». I: *Energy and Buildings* 75 (2014), s. 189–198. DOI: doi.org/10.1016/j.enbuild.2014.01.048.
- [72] Safae Bourhnane mfl. «Machine learning for energy consumption prediction and scheduling in smart buildings». I: *SN Applied Sciences* 2 (2020), s. 20. DOI: 10.1007/s42452-020-2024-9.
- [73] Qizheng Dong. «Leakage Prediction in Machine Learning Models When Using Data from Sports Wearable Sensors». I: *Computational Intelligence and Neuroscience* 2022 (2022). DOI: 10.1155/2022/5314671.
- [74] Redaksjonen hos Time og Date AS. *Skoleåret 2017-2018*. Hentet 15.11.2022. 2022. URL: <https://timeanddate.no/kalender/skolerute2017-2018>.
- [75] Fellesforbundet. *Fellesferien 2017 fastsatt*. Hentet 21.11.2022. 2016. URL: <https://fellesforbundet.no/aktuelt/nyheter/2016/fellesferien-2017-fastsatt/>.
- [76] Yr - En tjeneste fra Meteorologisk institutt og NRK. *Gardermoen*. Hentet 12.09.2022. 2022. URL: <https://yr.no/nb/v%C3%A6rvarsel/daglig-tabell/5-4780/Norge/Viken/Ullensaker/Gardermoen>.
- [77] Gábor Petneházi. «Recurrent Neural Networks for Time Series Forecasting». I: *CoRR* abs/1901.00069 (2019).
- [78] Kuniyiko Fukushima. «Cognitron: A self-organizing multilayered neural network». I: *Biological Cybernetics* 20.3 (1975), s. 121–136. DOI: 10.1007/BF00342633.
- [79] Yangyang Fu mfl. «Using Support Vector Machine to Predict Next Day Electricity Load of Public Buildings with Sub-metering Devices». I: *Procedia Engineering* 121 (2015), s. 7. DOI: 10.1016/j.proeng.2015.09.097.
- [80] Hermine N. Akouemo og Richard J. Povinelli. «Time series outlier detection and imputation». I: *2014 IEEE PES General Meeting — Conference and Exposition*, s. 1–5. DOI: 10.1109/PESGM.2014.6939802.
- [81] Omer Levend Asikoglu. «Temporal Distribution of 24-hour Maximum Rainfall». I: *International Journal of Engineering Science Invention* 6.7 (2017), s. 5.

- [82] Reidun Marie Romundstad. «Modellering av solvarmeanlegg og produksjons-simuleringer ved Hafslunds fjernvarmeanlegg». Masteroppgave, NMBU. 2014.
- [83] Baran Yildiz, Jose I. Bilbao og Alistar B. Sproul. «A review and analysis of regression and machine learning models on commercial building electricity load forecasting». I: *Renewable and Sustainable Energy Reviews* 73 (2017), s. 1104–1122. DOI: 10.1016/j.rser.2017.02.023.
- [84] Wei-Chiang Hong. «Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm». I: *Energy* 36.9 (2011), s. 5568–5578. DOI: 10.1016/j.energy.2011.07.015.
- [85] Jan G. De Gooijer og Rob J. Hyndman. «25 years of time series forecasting». I: *International Journal of Forecasting* 22.3 (2006), s. 443–473. DOI: 10.1016/j.ijforecast.2006.01.001.
- [86] Fred Collopy, Monica Adya og J. Scott Armstrong. «Research Report—Principles for Examining Predictive Validity: The Case of Information Systems Spending Forecasts». I: *Information Systems Research* 5 (1994), s. 170–179. DOI: 10.1287/isre.5.2.170.

Tillegg A

Tilleggsinformasjon til kapittel 4

Dette vedlegget inneholder utdypende informasjon om datasettene som ble beskrevet i oppgavens kapittel 4. En oversikt over helligdagene i datasettet finnes i delkapittel A.1. El-bruken ved Oslo lufthavn i uke 27 og uke 39 i 2019 er plottet sammen med meteorologiske målinger for samme uker i delkapittel A.2.

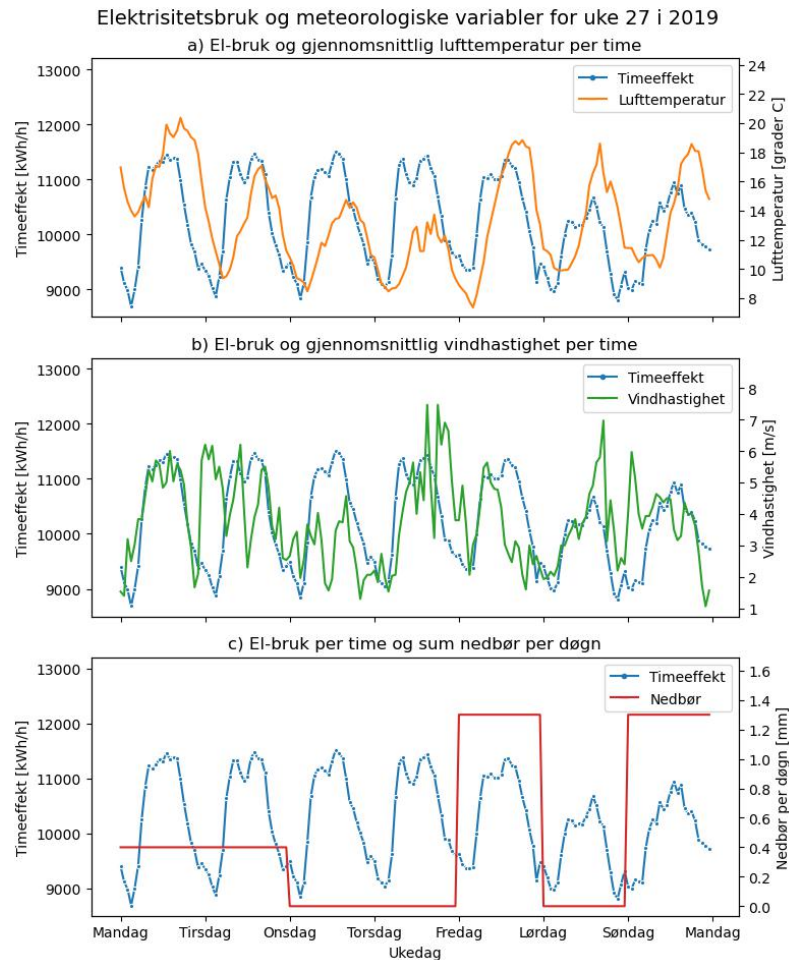
A.1 Helligdager i årene 2017 til 2019

De norske helligdagene som er inkludert i datasettet med tilhørende datoer i årene 2017 til 2019 finnes i tabell A.1 [61]. De røde dagene offentlig høytidsdag 1. mai og grunnlovsdag 17. mai er også inkludert, men ikke søndager.

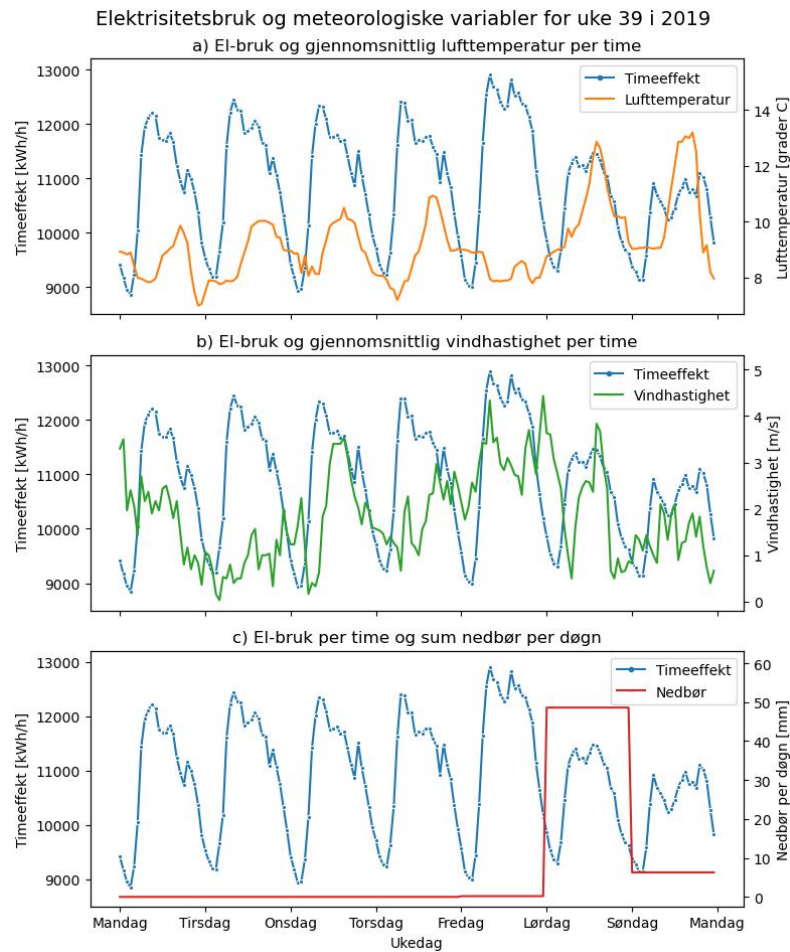
Tabell A.1: Datoene for helligdagene som inkludert i helligdagsvariabelen i datasettet [61].

Helligdag	2017	2018	2019
1. nyttårsdag	01.01	01.01	01.01
Palmesøndag	09.04	25.03	14.04
Skjærtorsdag	13.04	29.03	18.04
Langfredag	14.04	30.03	19.04
1. påskedag	16.04	01.04	21.04
2. påskedag	17.04	02.04	22.04
Offentlig høytidsdag	01.05	01.05	01.05
Grunnlovsdag	17.05	17.05	17.05
Kristi Himmelfartsdag	25.05	10.05	30.05
1. pinsedag	04.06	20.05	09.06
2. pinsedag	05.06	21.05	10.06
1. juledag	25.12	25.12	25.12
2. juledag	26.12	26.12	26.12

A.2 El-bruk og meteorologiske målinger for uke 27 og 39 i 2019



Figur A.1: Elektrisitetsbruken [kWh/h] ved Oslo lufthavn i uke 27 i 2019 sammen med a) lufttemperatur, b) vindhastighet og c) nedbør.



Figur A.2: Elektrisitetsbruken [kWh/h] ved Oslo lufthavn i uke 39 i 2019 sammen med a) lufttemperatur, b) vindhastighet og c) nedbør.

Tillegg B

Tilleggstabeller til kapittel 5

Dette vedlegget viser komplette tallresultater med RMSE-, MAE- og MAPE-scores. De to første sidene viser resultatene fra prediksjoner av timeeffekt for Oslo lufthavn Gardermoen én time frem i tid. Den siste siden viser resultatene fra tilsvarende prediksjoner 24 timer frem i tid. Begge tabeller er sortert på modellenes RMSE på testsettet, og basismodellen er markert i en annen farge.

ID	Navn	Rang	Endring					MAE test	MAPE test	RMSE test ift. basis	Kommentar visuelt		
			RMSE trening	MAE trening	MAPE trening	RMSE validering	MAE validering					MAPE validering	
26	I	1	416,5	324,6	3,0	1214,8	990,4	9,5	511,2	34,2 %	402,5	variabler: kun tidl. el-bruk 3,7 og time	> basis
42	II	2	278,5	215,8	2,0	1299,8	1054,0	9,9	586,1	24,5 %	475,5	variabler: tid og tidl. el- 4,4 bruk, # enheter:128	> basis
12	III	3	405,7	320,2	3,0	1220,8	1000,4	9,5	624,1	19,6 %	493,6	# enheter: 64 og 64 og 64, 4,6 # lag: 3	> basis
38	IV	4	406,8	322,6	3,0	1232,5	1003,7	9,5	624,6	19,6 %	491,1	variabler: alle inkl. tidl. el- bruk, # enheter: 64 og 64 4,5 og 64, # lag: 3	> basis men annerledes
10	V	5	431,5	337,6	3,2	1234,1	1003,9	9,7	636,0	18,1 %	504,4	# enheter: 32 og 32, # lag: = modell VIII, 4,7 2	= modell VIII, < basis
36	VI	6	349,2	273,9	2,6	1281,4	1039,2	10,0	644,3	17,0 %	518,9	variabler: alle inkl. tidl. el- bruk	> basis unntatt RNN3
41	VII	7	353,4	277,2	2,6	1263,9	1022,0	9,7	647,7	16,6 %	534,8	variabler: tid og tidl. el- bruk	= basis unntatt RNN3
9	VIII	8	465,1	364,3	3,4	1189,1	962,8	9,3	655,3	15,6 %	518,1	forenkling: 0.1, # enheter: 4,8 32 og 32, # lag: 2	< basis
15	IX	9	336,5	262,9	2,5	1246,3	1020,4	9,8	677,7	12,7 %	556,6	5,3 # enheter:128	< basis unntatt RNN2
40	X	10	394,4	307,4	2,9	1262,3	1034,3	9,9	678,3	12,6 %	548,8	variabler: tid, passasjerer 5,1 og temperatur	< basis unntatt RNN2
21	XI	11	745,7	536,2	4,9	1165,5	945,6	9,1	686,0	11,6 %	534,2	variabler: kun passasjerer 4,8 og time	<< basis
39	XII	12	427,5	334,6	3,1	1271,9	1037,8	10,0	701,1	9,7 %	576,1	variabler: tid og 5,4 passasjerer	< basis unntatt RNN2
30	XIII	13	401,5	314,1	3,0	1264,7	1028,5	9,8	704,2	9,3 %	571,2	5,3 variabler: tid	< basis men marginalt
25	XIV	14	629,2	491,9	4,6	1186,3	962,7	9,2	707,6	8,9 %	574,5	variabler: kun 5,3 værvariabler og time	< basis unntatt RNN2
32	XV	15	417,6	324,6	3,0	1335,3	1090,4	10,6	710,9	8,4 %	572,5	5,3 tilbakeblikk: 24*2	= basis
18	XVI	16	389,6	303,3	2,8	1261,7	1023,4	10,0	719,6	7,3 %	577,6	5,3 splitt: 70:20:10	= basis

17	XXVII	17	410,6	317,4	3,0	1147,8	936,5	9,3	723,8	6,8 %	587,1	= basis men annen tidsperiode
1	XVIII	18	388,5	299,0	2,8	1324,4	1079,0	10,5	730,5	5,9 %	583,6	= basis
33	XIX	19	428,4	330,7	3,1	1288,8	1040,2	10,1	731,0	5,9 %	588,7	= basis
4	XX	20	426,9	330,5	3,1	1256,4	1025,1	9,9	741,3	4,5 %	594,1	= basis
20	XXI	21	423,7	326,8	3,1	1280,7	1045,1	10,1	748,2	3,6 %	604,2	= basis
6	XXII	22	377,9	297,2	2,8	1276,3	1041,6	10,0	748,8	3,6 %	617,1	< basis
5	XXIII	23	472,6	362,1	3,4	1197,5	979,9	9,4	753,8	2,9 %	599,3	= basis
34	XXIV	24	418,0	322,7	3,0	1267,7	1028,0	9,9	755,0	2,8 %	606,8	= basis
2	XXV	25	425,3	327,5	3,1	1302,7	1060,9	10,3	760,2	2,1 %	617,6	BASIS
3	XXVI	26	449,3	345,3	3,2	1294,5	1056,8	10,2	766,9	1,2 %	620,6	= basis
13	XXVII	27	372,1	292,6	2,8	1356,1	1108,0	10,7	767,8	1,1 %	632,9	= basis
7	XXVIII	28	449,3	351,3	3,3	1247,2	1017,2	9,8	776,1	0,0 %	637,6	> basis
31	XXIX	29	380,3	298,2	2,8	1261,2	1024,1	9,8	776,4	0,0 %	624,8	= modell
37	XXX	30	477,7	378,9	3,6	1172,7	952,6	9,2	787,1	-1,4 %	642,7	< basis
16	XXXI	31	407,9	318,8	3,0	1298,6	1048,2	9,9	802,0	-3,3 %	653,1	< basis
11	XXXII	32	452,7	361,7	3,4	1192,1	975,6	9,3	820,1	-5,6 %	673,0	umtatt RNN2
8	XXXIII	33	605,9	479,0	4,5	1147,9	938,1	9,0	855,7	-10,2 %	696,8	> basis
35	XXXIV	34	491,4	380,8	3,6	1251,6	1018,0	9,8	875,3	-12,7 %	690,6	= basis
14	XXXV	35	573,8	448,6	4,2	1244,2	1010,4	9,5	889,3	-14,5 %	718,7	umtatt RNN1
27	XXXVI	36	759,3	634,4	6,0	1013,3	839,4	8,0	899,7	-15,9 %	724,0	<< basis
29	XXXVII	37	442,8	344,3	3,2	1306,2	1055,4	9,9	980,6	-26,3 %	789,6	> basis men
19	I	38	1107,9	886,9	8,4	1169,8	951,3	9,3	1178,3	-51,8 %	964,8	marginalt

ID	Navn	Rang.	Forskjell fra basis	Kommentar visuelt: ar visuelt: trening	RMSE	Diff. fra +1t	Diff. fra basismodel I +24t	Diff. fra +1t2	Diff. fra basismodel I +24t2	Diff. fra +1t3	Diff. fra basismodel I +24t3		
12	III		# enheter: 64 og 64, > basis, glatt kurve	430,9	-6 %	3 %	333,1	-4 %	2 %	3,1	-4 %	1 %	
36	VI		variabler: alle inkl. tidl. el-2 bruk	418,3	-20 %	6 %	324,1	-18 %	4 %	3,0	-19 %	4 %	
2	XXV		3 BASIS	442,8	-4 %		339,0	-4 %		3,2	-4 %		
38	IV		inkl. tidl. el-bruk, #4 enheter: 64	405,9	0 %	8 %	314,3	3 %	7 %	3,0	3 %	7 %	
RMSE validering													
12	III		1000,6	9,5	653,3	-5 %	12 %	509,7	-3 %	15 %	4,7	-2 %	15 %
36	VI		1011,9	9,8	707,5	-10 %	5 %	568,3	-10 %	5 %	5,3	-8 %	4 %
2	XXV		1002,7	9,7	741,2	3 %		600,7	3 %		5,5	3 %	
38	IV		1007,7	9,5	764,7	-22 %	-3 %	594,1	-21 %	1 %	5,4	-19 %	2 %

Tillegg C

Tilleggsfigurer til kapittel 5

Dette vedlegget inneholder flere figurer som illustrerer resultatene fra prediksjonene av el-bruk som ble presentert i kapittel 5.

Delkapittel C.1 viser oppbyggingen til basismodellen og antallet parametere i hvert lag.

Delkapittel C.2.1 inneholder prediksjonsplott for basismodellene fra RNN1, RNN2 og RNN3. I delkapittel C.2.2 finnes prediksjonsplott for samtlige 38 modellkonfigurasjoner som ble testet til prediksjon én time frem i tid fra RNN1, sortert etter RMSE på testsettet. Prediksjonsplott for de fire modellkonfigurasjonene som ble testet til prediksjon 24 timer frem i tid finnes i delkapittel C.2.3, sortert etter RMSE på testsettet.

C.1 Oppbygging basismodell

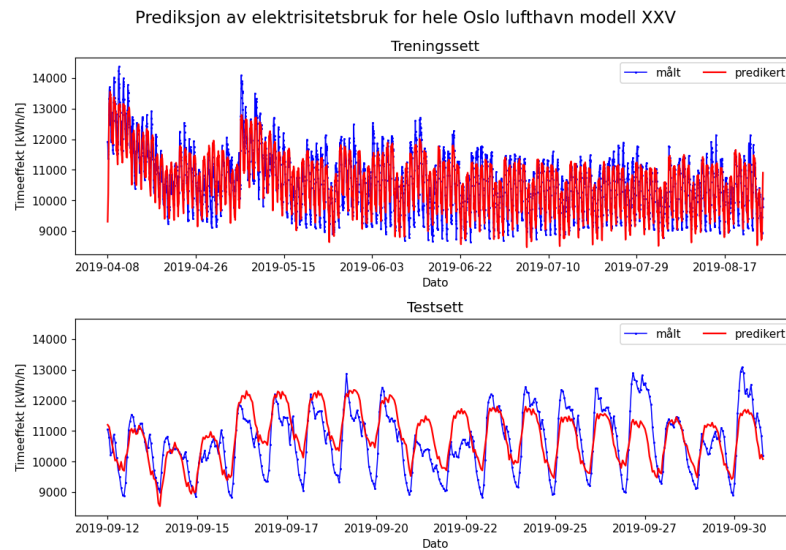
Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 32)	15744
dense (Dense)	(None, 1)	33

=====
Total params: 15,777
Trainable params: 15,777
Non-trainable params: 0
=====

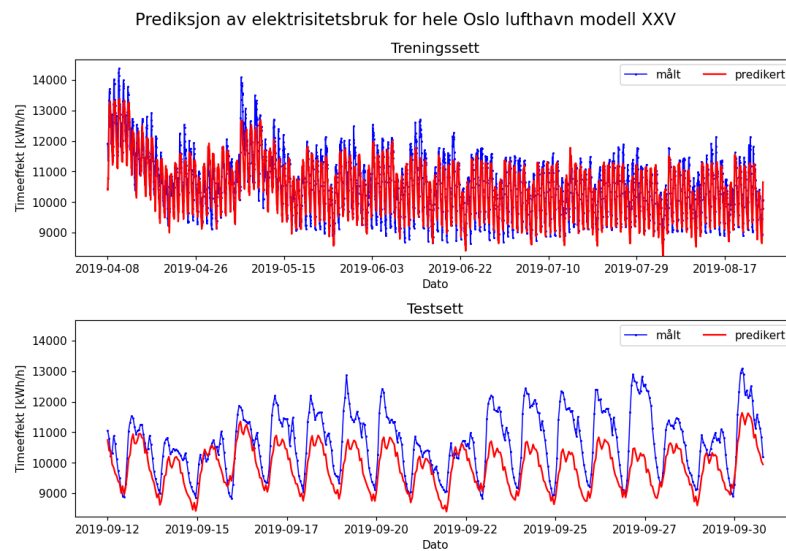
Figur C.1: Lagene i basismodellen med tilørende output-form og antall parametere. Etter omforming av kategoriske variabler hadde modellen 90 input-variabler. Skjerm bilde fra skript tilgjengelig fra GitHub [46].

C.2 Prediksjonsplott

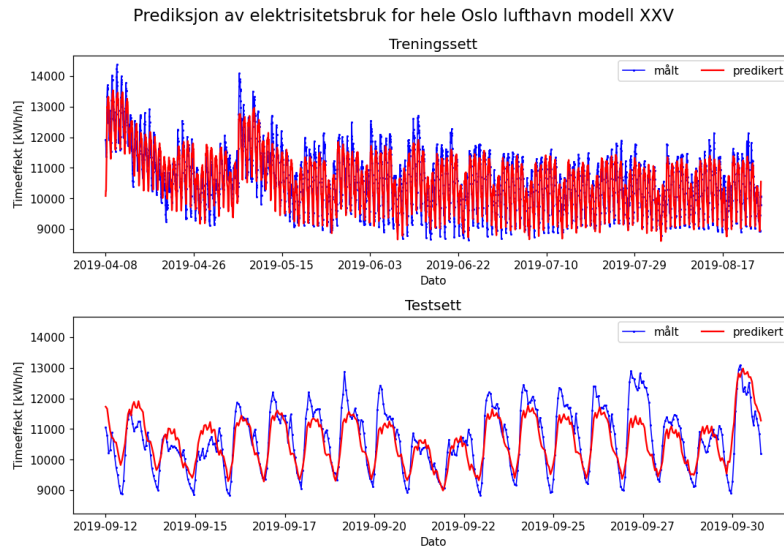
C.2.1 Basismodell



Figur C.2: Prediksjonsplott for henholdsvis trenings- og testsettet fra basismodellen fra RNN1. Modellen predikerer el-bruk én time frem i tid.

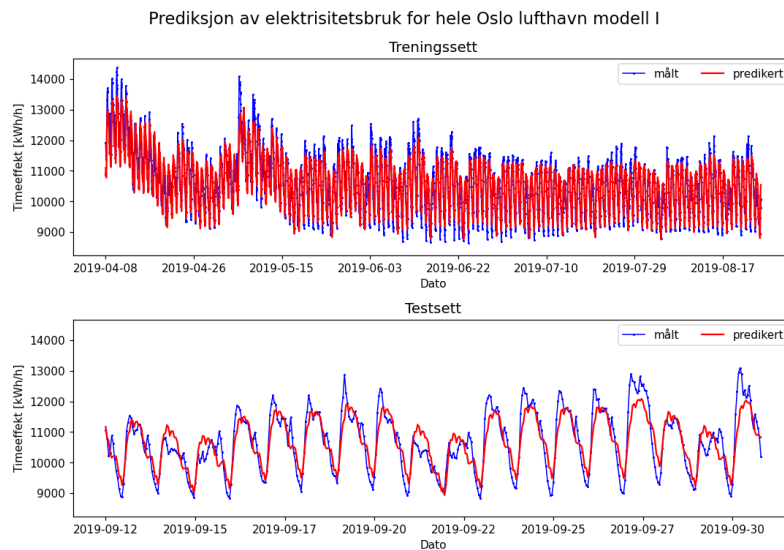


Figur C.3: Prediksjonsplott for henholdsvis trenings- og testsettet fra basismodellen fra RNN2. Modellen predikerer el-bruk én time frem i tid.

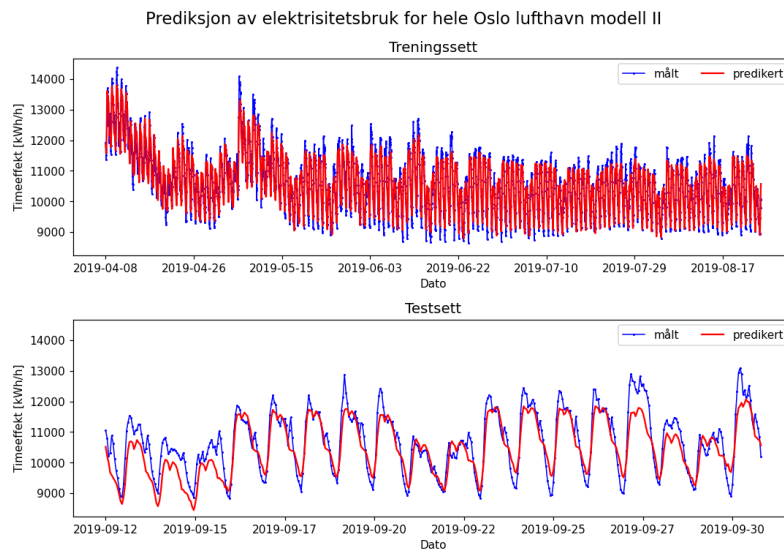


Figur C.4: Prediksjonsplott for henholdsvis trenings- og testsettet fra basismodellen fra RNN3. Modellen predikerer el-bruk én time frem i tid.

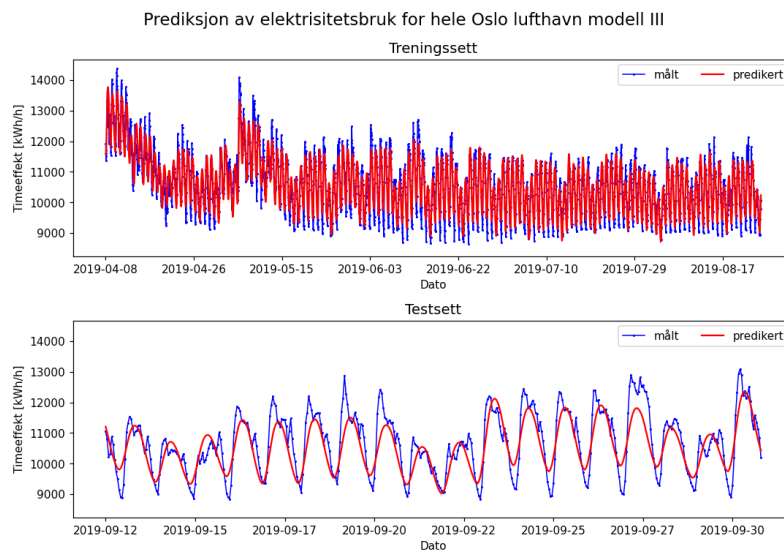
C.2.2 Prediksjon én time frem i tid



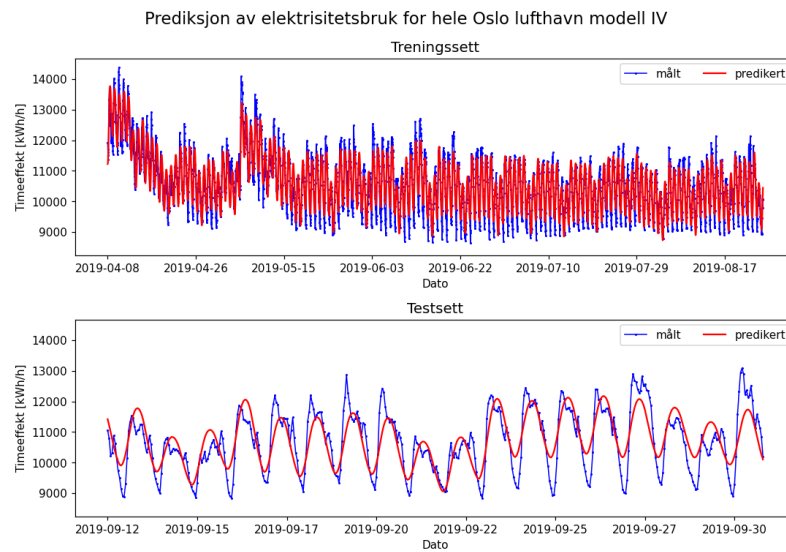
Figur C.5: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon I fra RNN1. Modellen predikerer el-bruk én time frem i tid.



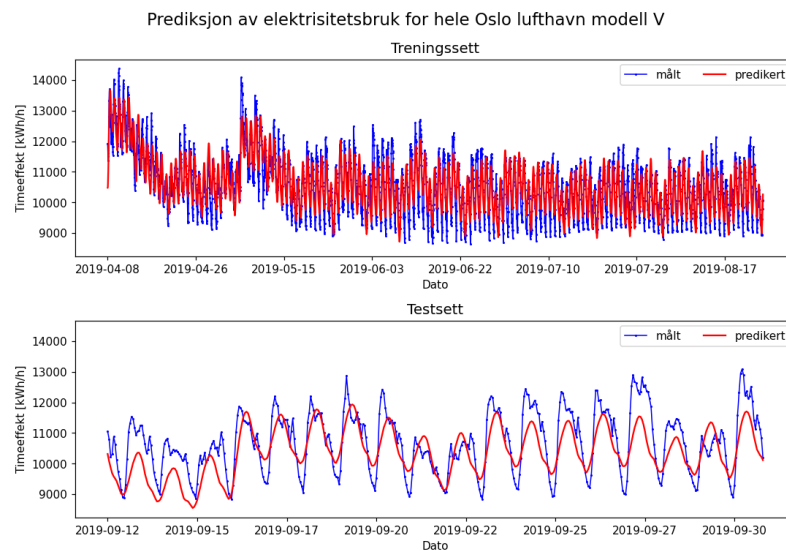
Figur C.6: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon II fra RNN1. Modellen predikerer el-bruk én time frem i tid.



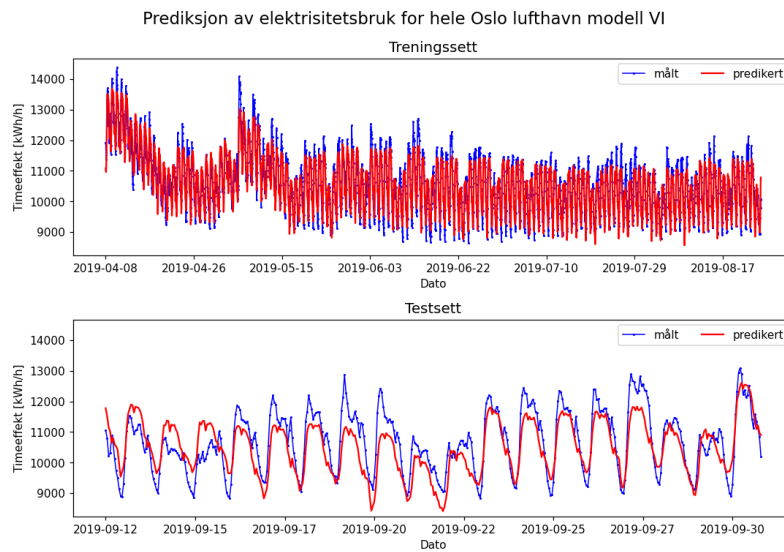
Figur C.7: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon III fra RNN1. Modellen predikerer el-bruk én time frem i tid.



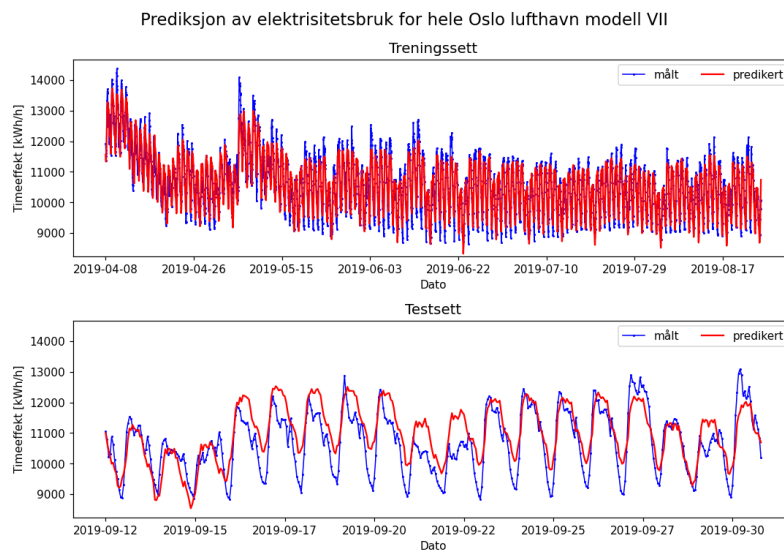
Figur C.8: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon IV fra RNN1. Modellen predikerer el-bruk én time frem i tid.



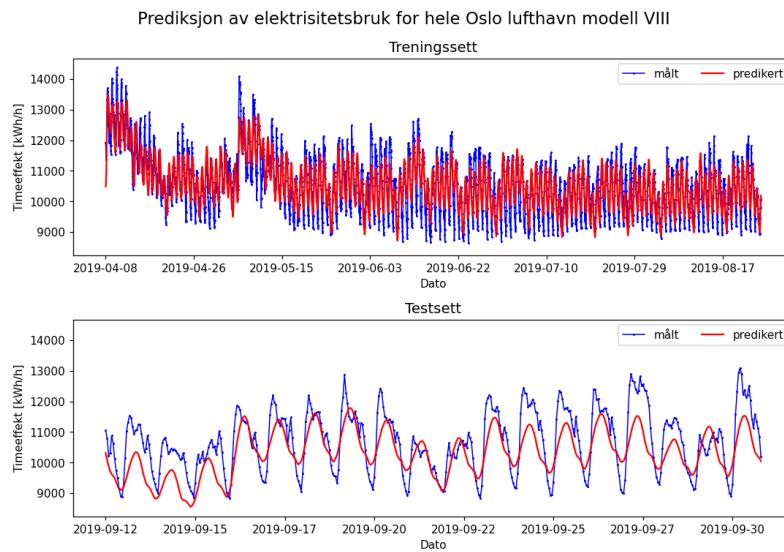
Figur C.9: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon V fra RNN1. Modellen predikerer el-bruk én time frem i tid.



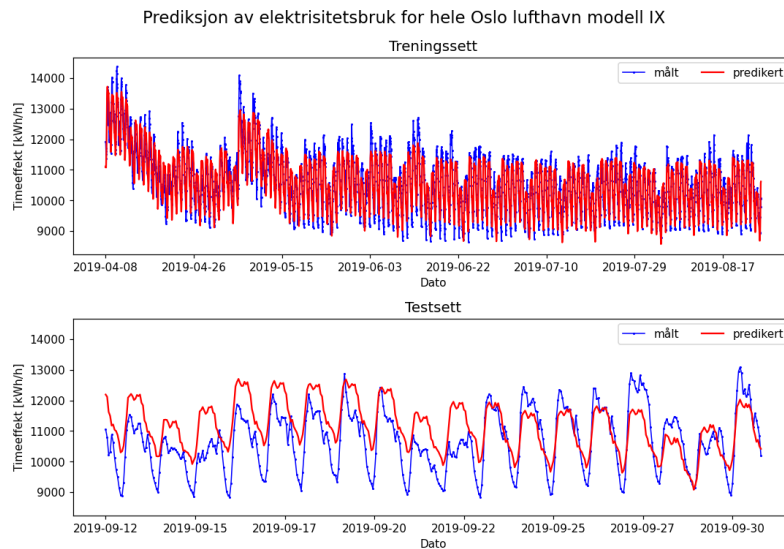
Figur C.10: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon VI fra RNN1. Modellen predikerer el-bruk én time frem i tid.



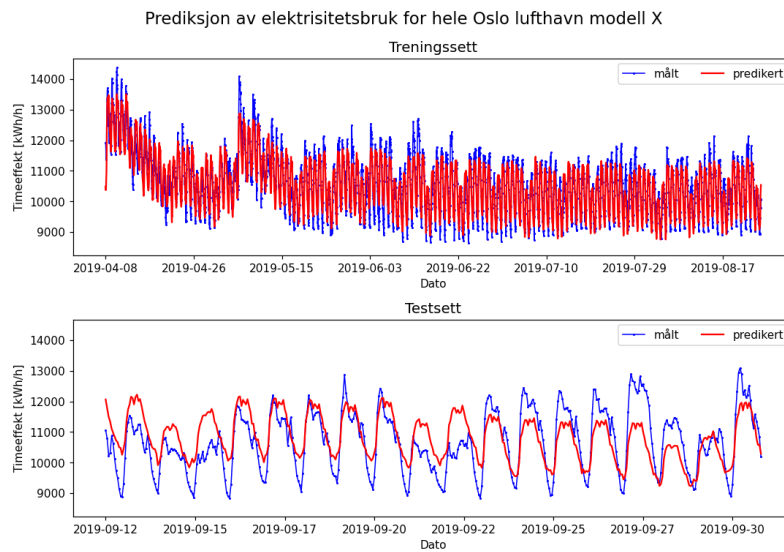
Figur C.11: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon VII fra RNN1. Modellen predikerer el-bruk én time frem i tid.



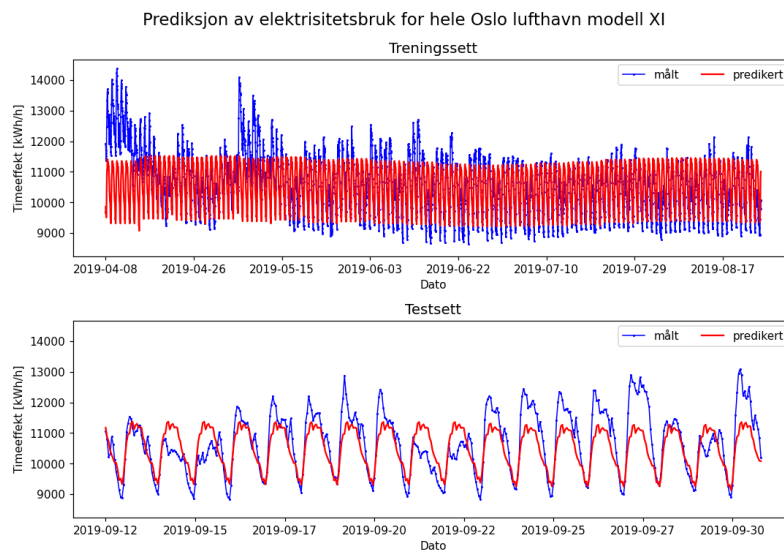
Figur C.12: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon VIII fra RNN1. Modellen predikerer el-bruk én time frem i tid.



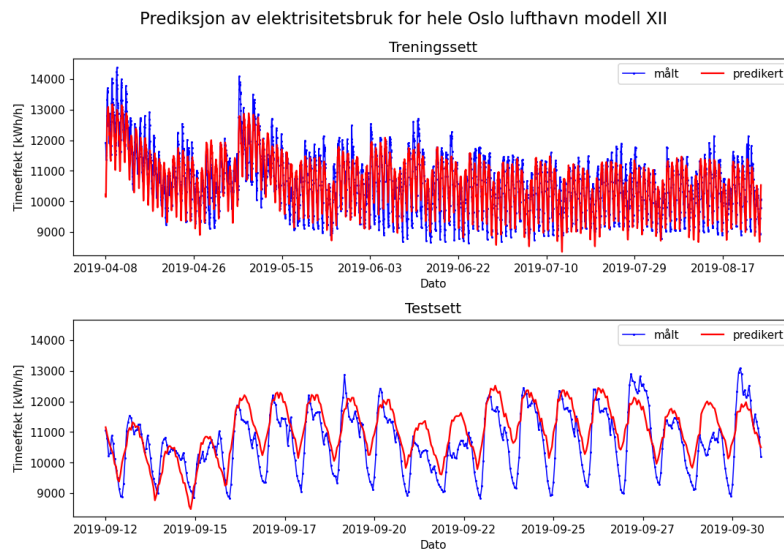
Figur C.13: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon IX fra RNN1. Modellen predikerer el-bruk én time frem i tid.



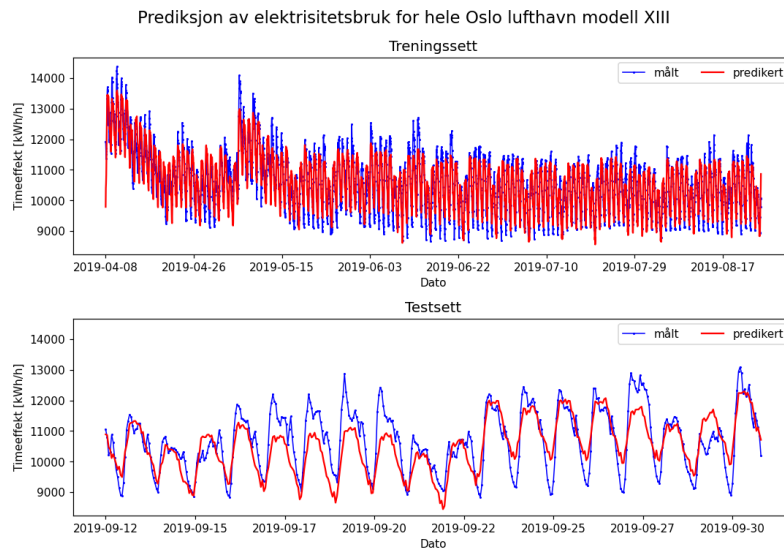
Figur C.14: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon X fra RNN1. Modellen predikerer el-bruk én time frem i tid.



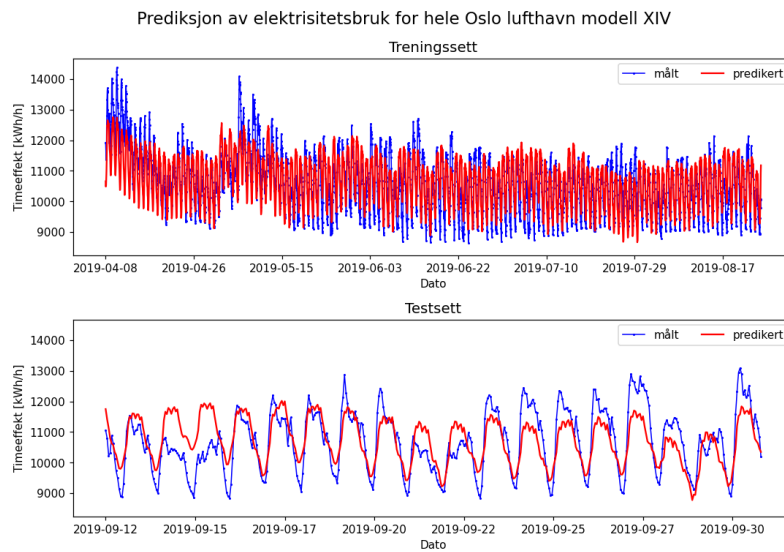
Figur C.15: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XI fra RNN1. Modellen predikerer el-bruk én time frem i tid.



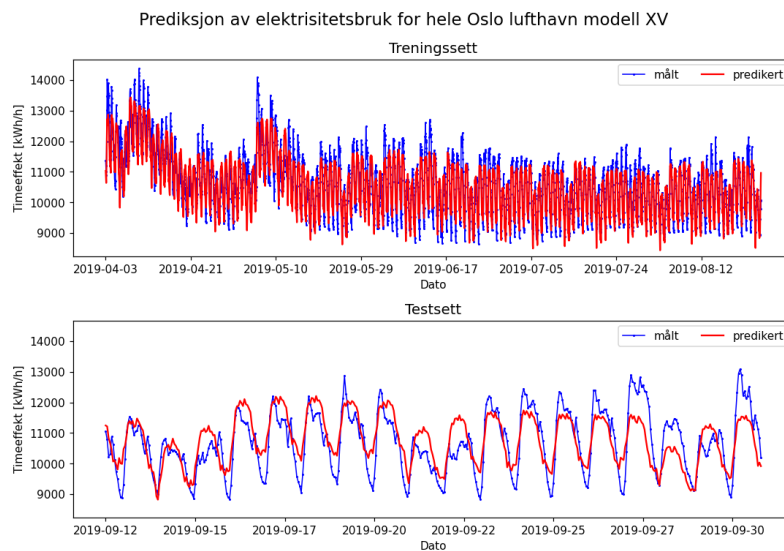
Figur C.16: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XII fra RNN1. Modellen predikerer el-bruk én time frem i tid.



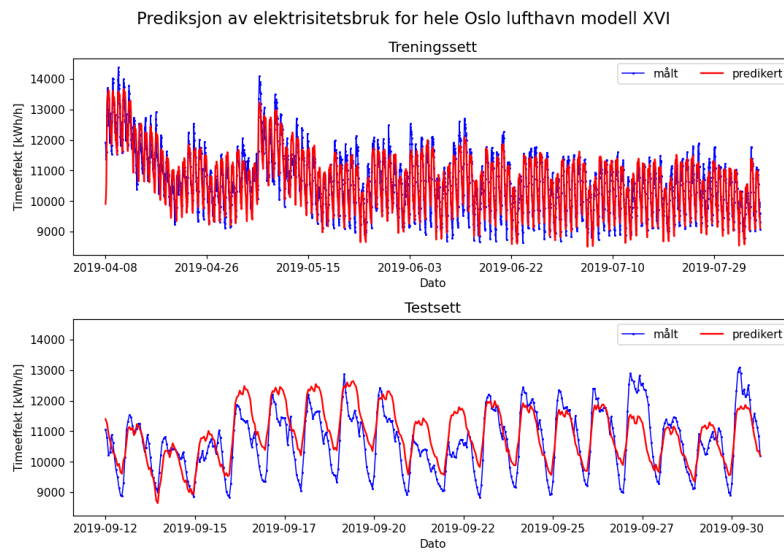
Figur C.17: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XIII fra RNN1. Modellen predikerer el-bruk én time frem i tid.



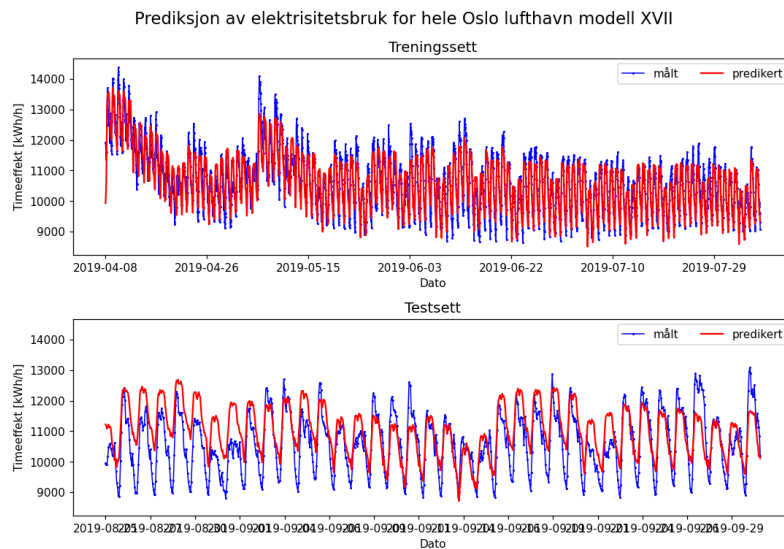
Figur C.18: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XIV fra RNN1. Modellen predikerer el-bruk én time frem i tid.



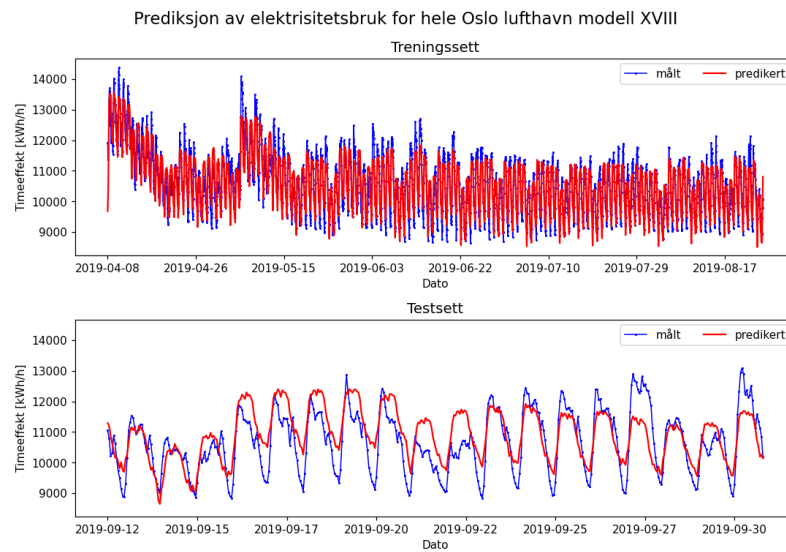
Figur C.19: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XV fra RNN1. Modellen predikerer el-bruk én time frem i tid.



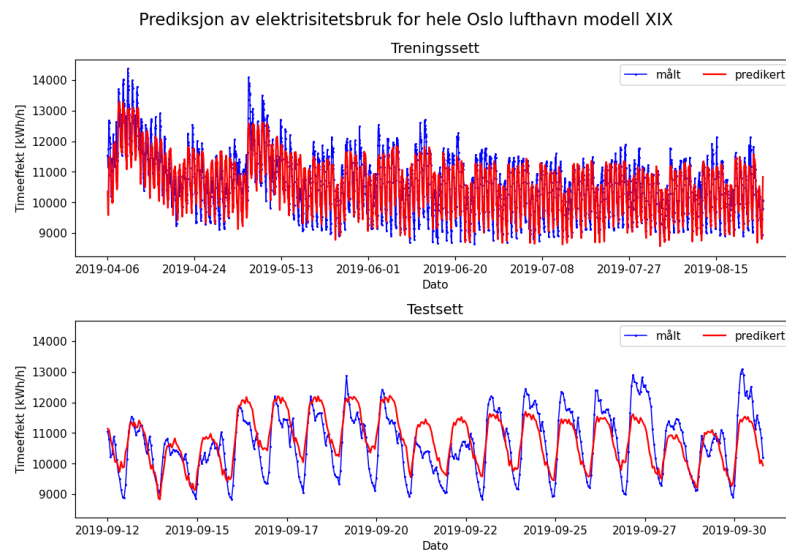
Figur C.20: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XVI fra RNN1. Modellen predikerer el-bruk én time frem i tid.



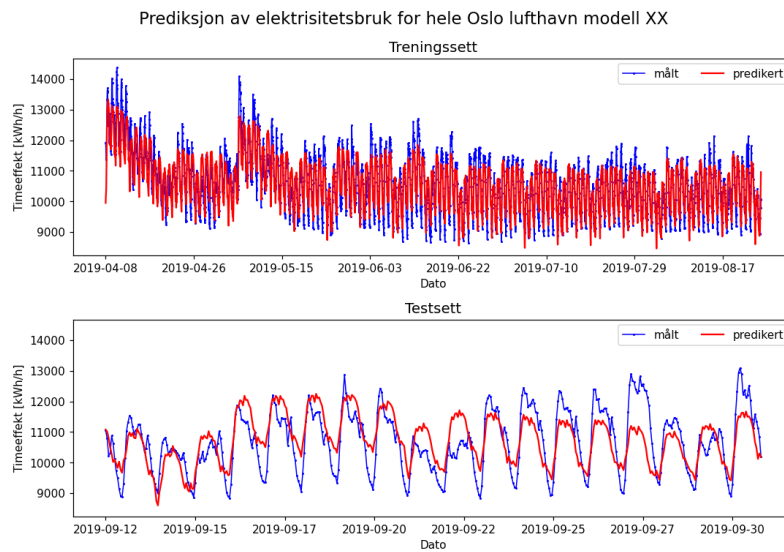
Figur C.21: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XVII fra RNN1. Modellen predikerer el-bruk én time frem i tid.



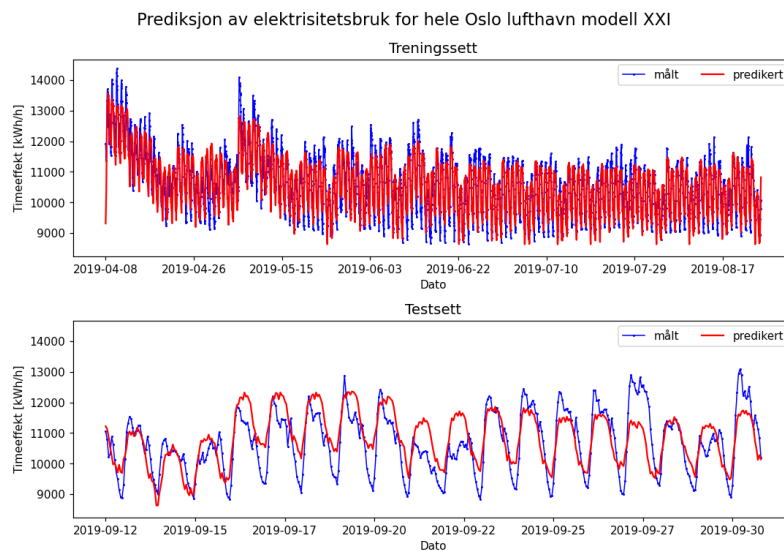
Figur C.22: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XVIII fra RNN1. Modellen predikerer el-bruk én time frem i tid.



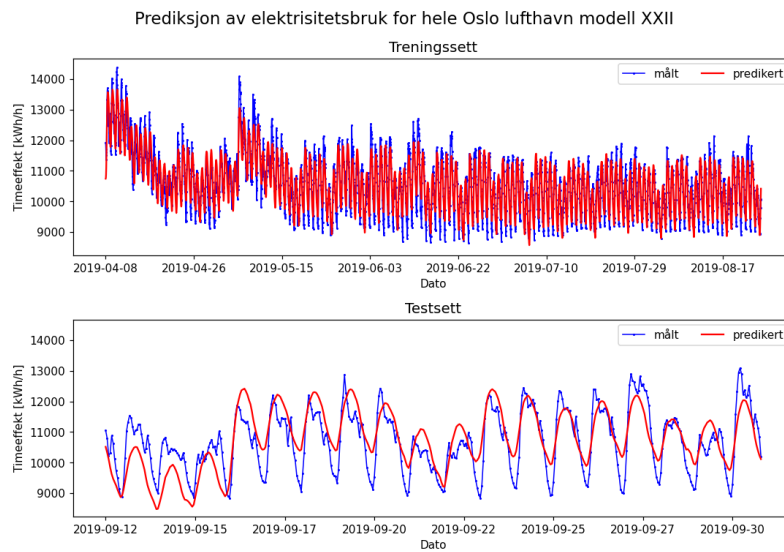
Figur C.23: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XIX fra RNN1. Modellen predikerer el-bruk én time frem i tid.



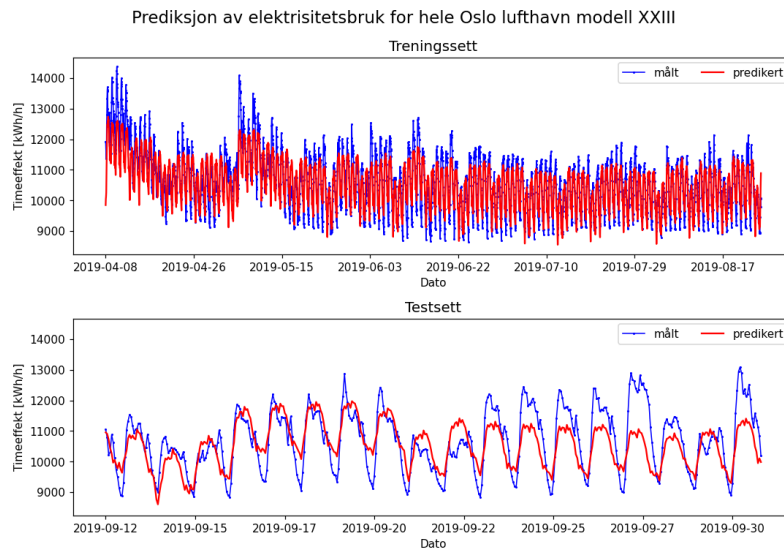
Figur C.24: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XX fra RNN1. Modellen predikerer el-bruk én time frem i tid.



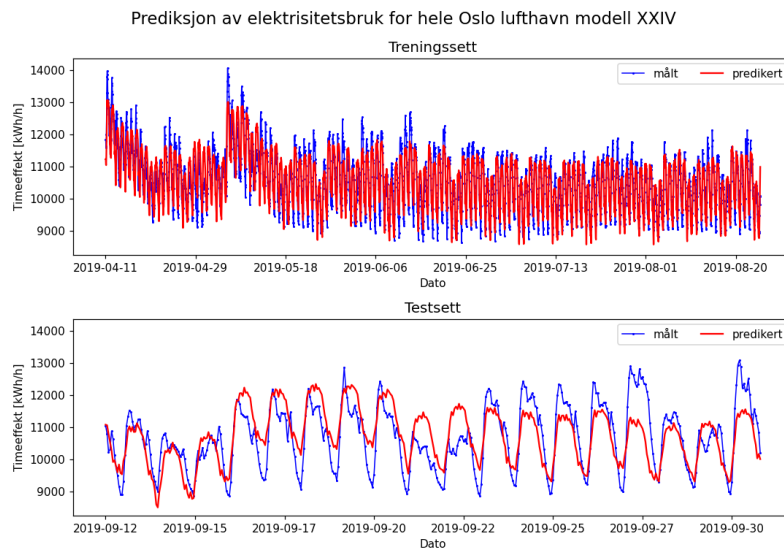
Figur C.25: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXI fra RNN1. Modellen predikerer el-bruk én time frem i tid.



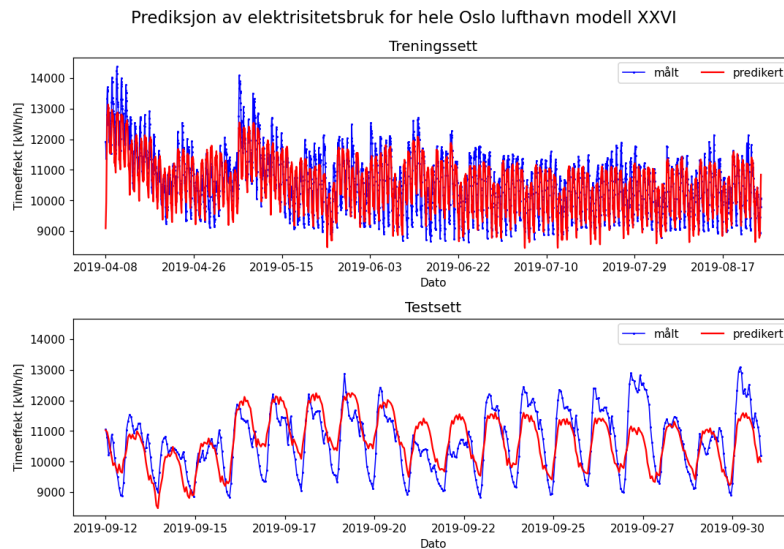
Figur C.26: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXII fra RNN1. Modellen predikerer el-bruk én time frem i tid.



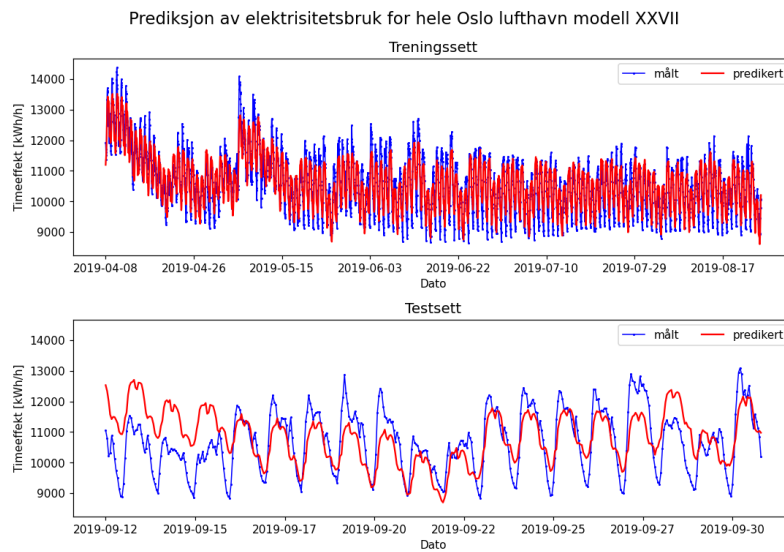
Figur C.27: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXIII fra RNN1. Modellen predikerer el-bruk én time frem i tid.



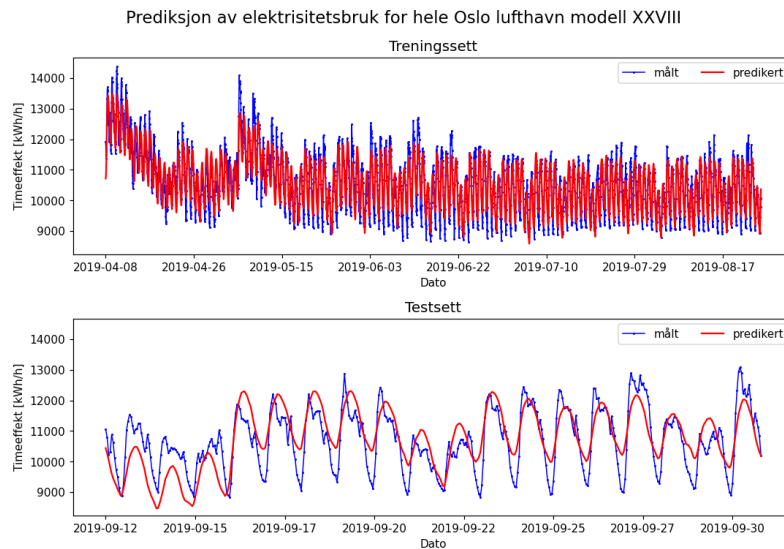
Figur C.28: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXIV fra RNN1. Modellen predikerer el-bruk én time frem i tid.



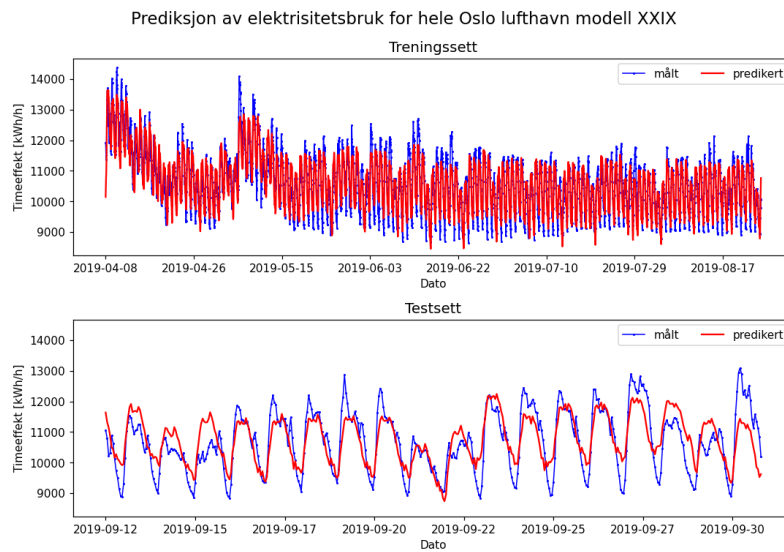
Figur C.29: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXVI fra RNN1. Modellen predikerer el-bruk én time frem i tid.



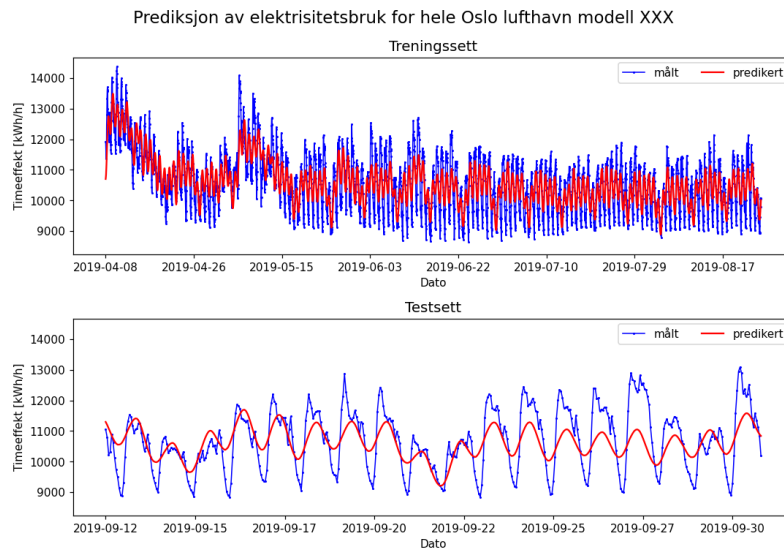
Figur C.30: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXVII fra RNN1. Modellen predikerer el-bruk én time frem i tid.



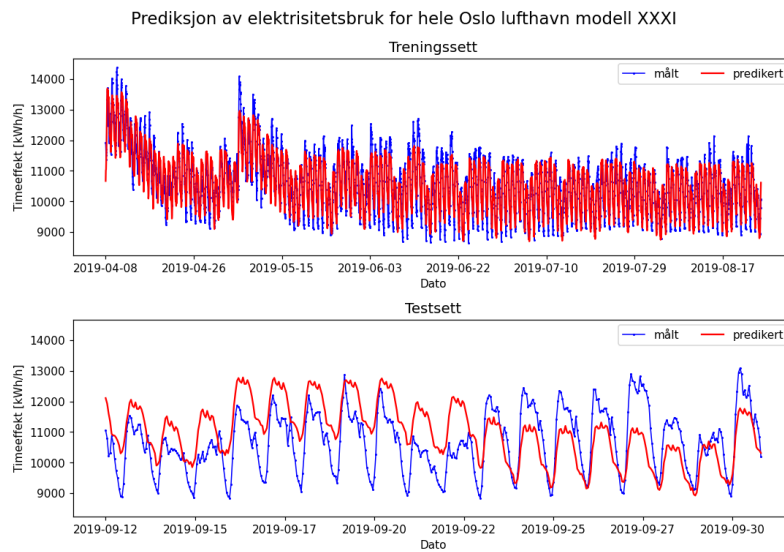
Figur C.31: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXVIII fra RNN1. Modellen predikerer el-bruk én time frem i tid.



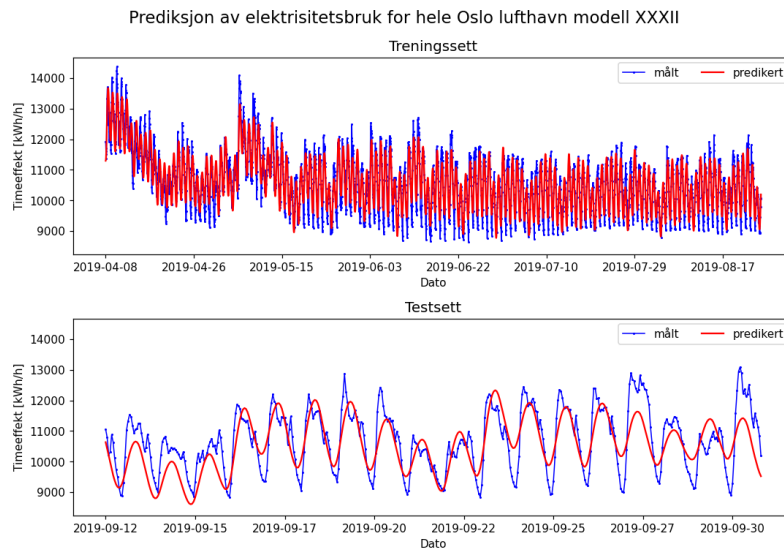
Figur C.32: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXIX fra RNN1. Modellen predikerer el-bruk én time frem i tid.



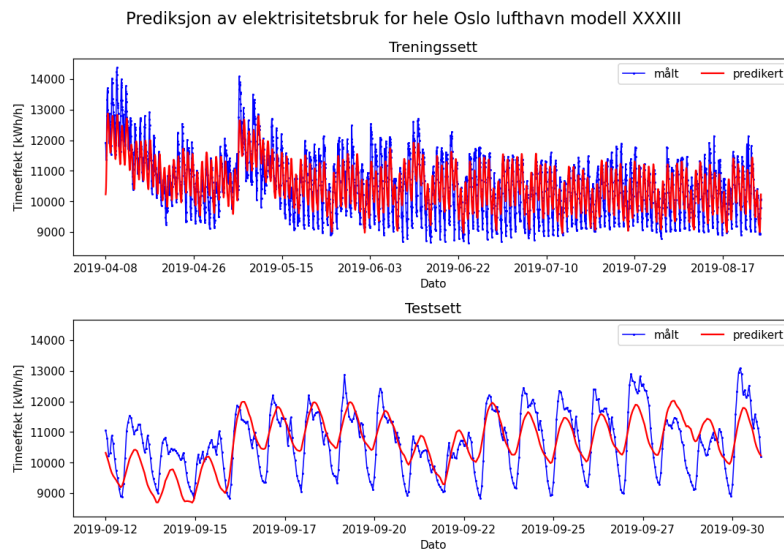
Figur C.33: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXX fra RNN1. Modellen predikerer el-bruk én time frem i tid.



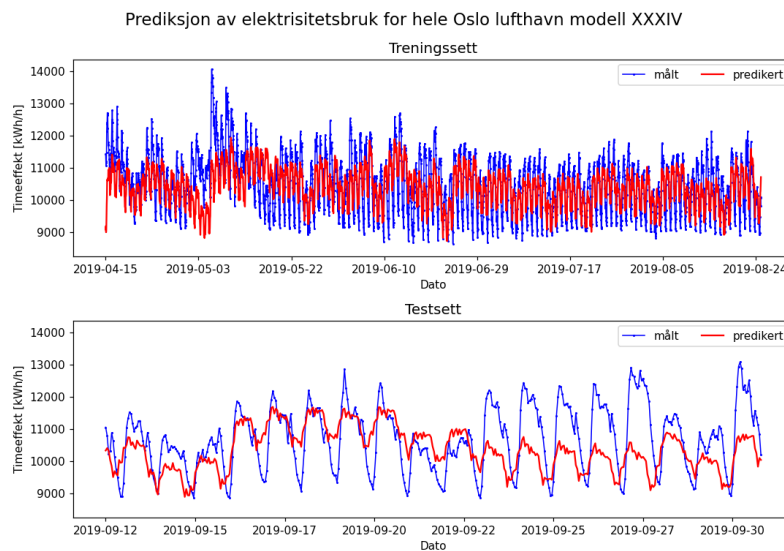
Figur C.34: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXXI fra RNN1. Modellen predikerer el-bruk én time frem i tid.



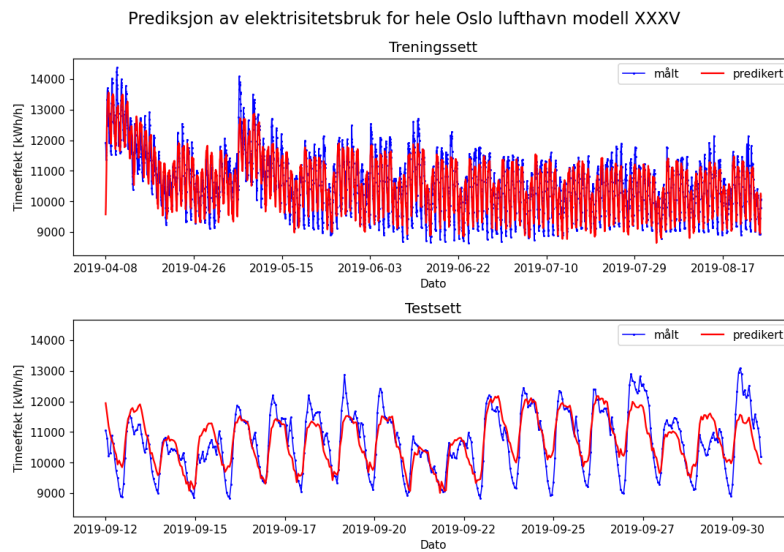
Figur C.35: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXXII fra RNN1. Modellen predikerer el-bruk én time frem i tid.



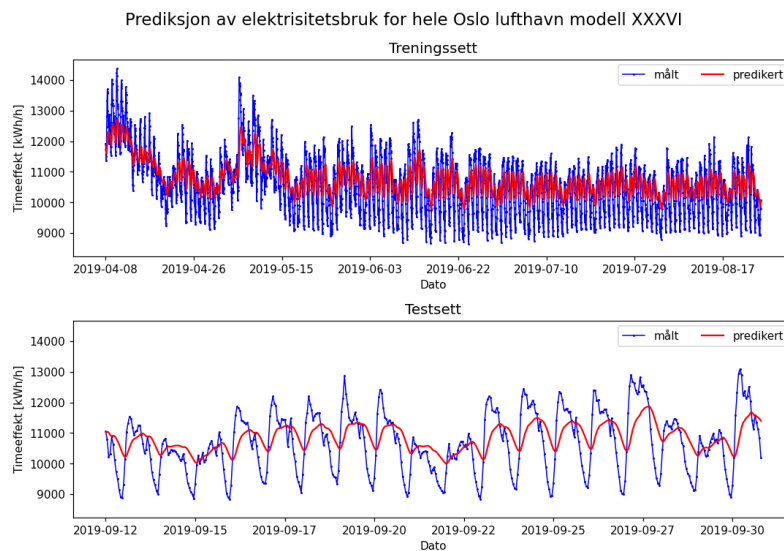
Figur C.36: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXXIII fra RNN1. Modellen predikerer el-bruk én time frem i tid.



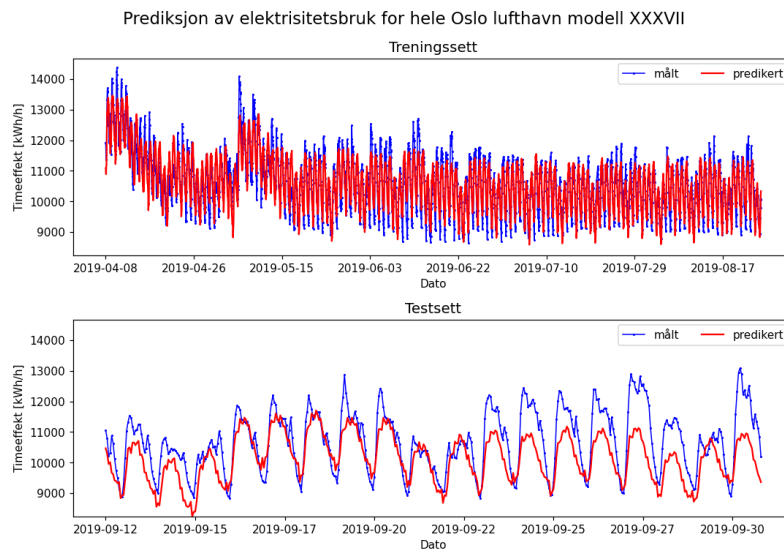
Figur C.37: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXXIV fra RNN1. Modellen predikerer el-bruk én time frem i tid.



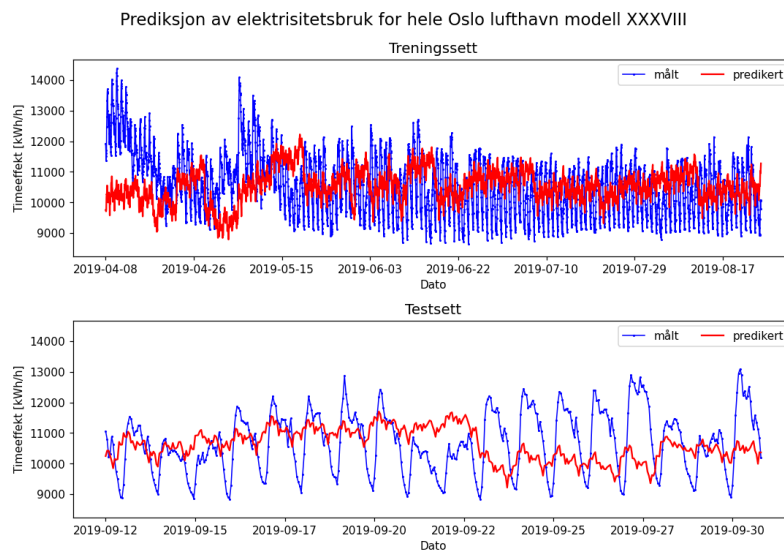
Figur C.38: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXXV fra RNN1. Modellen predikerer el-bruk én time frem i tid.



Figur C.39: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXXVI fra RNN1. Modellen predikerer el-bruk én time frem i tid.

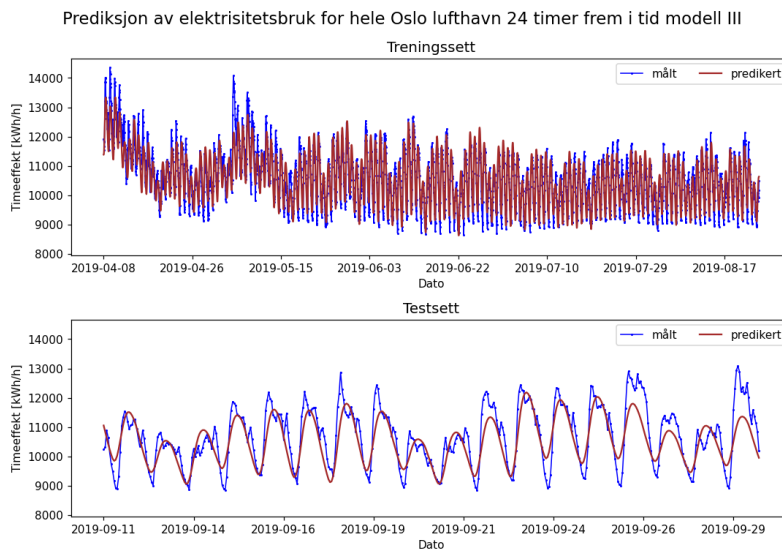


Figur C.40: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXXVII fra RNN1. Modellen predikerer el-bruk én time frem i tid.

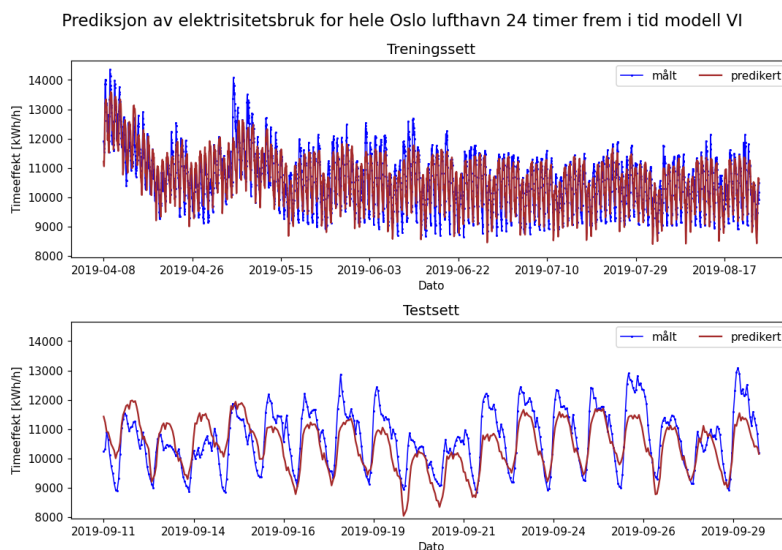


Figur C.41: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXXVIII fra RNN1. Modellen predikerer el-bruk én time frem i tid.

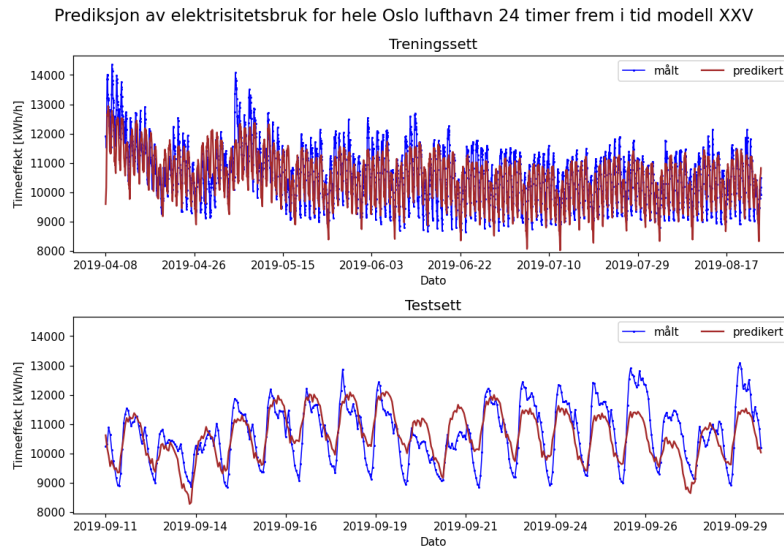
C.2.3 Prediksjon et døgn frem i tid



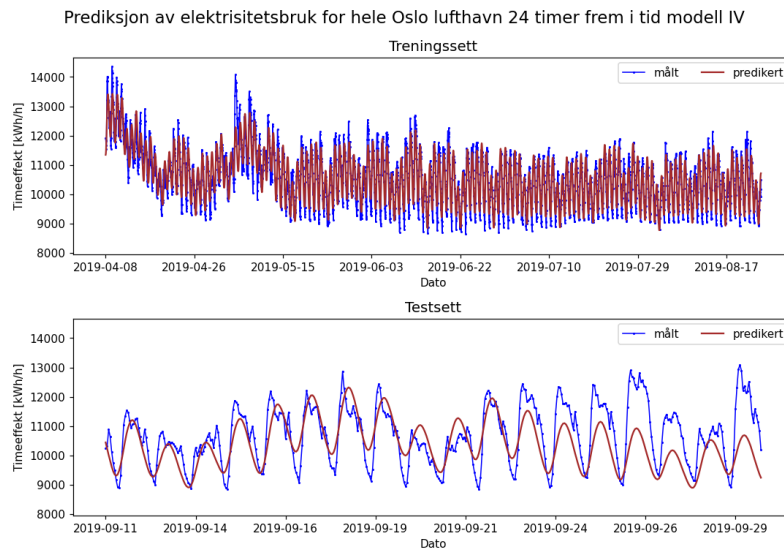
Figur C.42: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon III fra RNN1. Modellen predikerer el-bruk 24 timer frem i tid.



Figur C.43: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon VI fra RNN1. Modellen predikerer el-bruk 24 timer frem i tid.



Figur C.44: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon XXV fra RNN1. Dette er basismodellen. Modellen predikerer el-bruk 24 timer frem i tid.



Figur C.45: Prediksjonsplott for henholdsvis trenings- og testsettet fra modellkonfigurasjon IV fra RNN1. Modellen predikerer el-bruk 24 timer frem i tid.



Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitenskapelige universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway