

Hydrologiens innvirkning på systemprisen i det nordiske kraftmarkedet

Kåre Olav Ramstad

Veiledere

Dennis Frestad (Universitet i Agder)

Bernt Viggo Matheussen (Agder Energi)

Tom Lassen (Universitet i Agder)

*Masteroppgaven er gjennomført som ledd i utdanningen ved
Universitetet i Agder og er godkjent som del av denne utdanningen.
Denne godkjenningen innebærer ikke at universitetet inntår for de
metoder som er anvendt og de konklusjoner som er trukket.*

Universitetet i Agder, 2012
Fakultet for teknologi og realfag
Institutt for ingeniørvitenskap

Forord

Denne rapporten er skrevet som en masteroppgave i industriell økonomi og teknologiledelse ved Universitetet i Agder. Oppgaven ble til etter at jeg kom i kontakt med Agder Energi Produksjon. Arbeidet med utredningen har vært utfordrende og svært lærerikt. Den har gitt meg god innsikt i kraftmarkedet og rekke statistiske metoder innen økonomi. Dette er kunnskap som trolig kommer godt med seinere i arbeidslivet.

Jeg ønsker å takke veileder Dennis Frestad ved UiA og Bernt Viggo Matheussen ved Agder Energi for faglig hjelp og konstruktive tilbakemeldinger som har holdt meg på rett spor.

I løpet av arbeidsprosessen har jeg også hatt kontakt med ulike personer som er bidratt med innspill og viktig informasjon som har vært nødvendig for å gjennomføre masterutredningen. Biveileder Tom Lassen har gitt meg tips om bruk av programvare, noe som har forenklet en del av oppgaven. Nord Pool Spot har gitt meg adgang på deres server, slik at jeg fikk mulighet til å hente detaljerte data.

Universitetet i Agder
Grimstad, mai 2012

Kåre Olav Ramstad

Sammendrag

Denne studien undersøker hvor sterkt nedbør, energitilsig, magasinutfylling og snømengde påvirker systemprisen i det nordiske kraftmarkedet. Måten hydrologiske faktorer påvirker systemprisen er interessant med tanke på måten kraftverk med magasin utfører sin produksjonsplanlegging og risikostyring. Studien er utført med empiriske data fra perioden 1998 til 2012. Det er laget to forskerspørsmål, et som undersøker korrelasjon mellom systempris og de hydrologiske variablene, og et som benytter regresjon til å finne ut hvor stor grad systemprisen kan forklares av de utvalgte variablene.

Korrelasjonen måles med Pearsons korrelasjonskoeffisient, hvor verdier nærme en eller minus en indikerer sterk korrelasjon, mens verdier lik null indikerer at det ikke finnes korrelasjon. Hver hydrologisk variabel er blitt analysert hver for seg for å oppnå linearitet i regresjonsmodellen. Med dette innebærer blant annet ulike dekomponeringer, filtreringer og transformasjoner. Forklaringsgraden måles med determinasjonskoeffisienten R^2 , som er et tall mellom null og en. Er verdien lav har modellen lav forklaringsgrad, er verdien høy har modellen høy forklaringsgrad. Det er blitt testet ut ulike måter å utføre regresjonen på, og to av disse er presentert og diskutert.

Av de fire hydrologiske variablene er det tilgjengelige ressurser som samvarierer mest med logaritmisk systempris med en korrelasjonsfaktor lik -0,72. Tilgjengelige ressurser defineres som magasinutfylling og snømengde. Nedbøren som er observert de 45 foregående ukene gir en korrelasjonsfaktor på -0,65 mot en logaritmisk trendjustert systempris. Energitilsiget som er observert over fem uker ser ut til å forklare best de kortsiktige svingningene, og har en korrelasjonsfaktor -0,38, mens sesongjusteres magasinutfyllingsgraden blir korrelasjonsfaktor -0,65 mot logaritmisk systempris trendjustert kraftpris. Det er knyttet problemer med å bruke magasinutfyllingsgraden og tilgjengelige ressurser i prisprognoser fordi den ser ut til å være sterkt påvirket av systemprisen.

Modellen med høyest forklaringsgrad bruker nedbøren observert 45 foregående uker, gjennomsnittlig energitilsig fem foregående uker og årstallet som forklarende variabler. Denne forklarer 75,00 % av logaritmisk systempris. En modell med tilsiget de fem siste ukene og tilgjengelige ressurser forklarer logaritmisk systempris med 59,04 %. Begge modellene bruker variabler som inneholder stor usikkerhet, men det er knyttet noe større usikkerhet til den andre modellen da tilgjengelige ressurser ser ut til å være påvirket av systemprisen. Test av regresjonene viser sterk autokorrelasjon i residualene. Dette kan være forårsaket av utelatte variabler eller at regresjonen på andre måter ikke er utført riktig. Dette er med på å redusere validiteten til modellenes resultater.

Innholdsfortegnelse

Forord	I
Sammendrag.....	II
Innholdsfortegnelse.....	III
1 Innledning.....	1
1.1 Optimeringsproblemet	1
1.2 Oppgavens struktur	2
2 Beskrivelse av kraftmarkedet.....	3
2.1 Forløpet til dagens kraftmarked	3
2.2 Kraftmarkedets virkemåte.....	3
2.3 Kraftbørsen Nord Pool.....	4
2.4 Kraftprisens priskarakteristikker.....	7
2.5 Modellering av kraftpris	9
3 Problemstilling og metode.....	11
3.1 Problemstilling og forskerspørsmål	11
3.2 Bakgrunn for problemstillingen.....	11
3.3 Forskningsdesign	12
3.4 Empiriske data	13
3.5 Korrelasjon og regresjonsanalyse	14
3.6 Transformasjoner, filtreringer og krysskorrelasjon	16
3.7 Dekomponering av tidsseriene.....	18
3.8 Innføring av ekstra ledd i modellen	20
3.9 Residualanalyse	21
3.10 Bruk av resultatene	22
4 Empirisk analyse og resultater	24
4.1 Systemprisen.....	24
4.2 Nedbør	27
4.3 Energitilsig.....	30
4.4 Magasinfyllingsgrad	31
4.5 Tilgjengelige ressurser.....	33
4.6 Multippel regresjon.....	34
5 Konklusjon	39
6 Litteraturreferanser	40

1 Innledning

Målet med denne oppgaven er å utføre en empirisk undersøkelse av i hvor stor grad hydrologiske forhold påvirker systemprisen i det nordiske kraftmarkedet. Vannkraft utgjør en stor del av det nordiske kraftmarkedet, derfor bygger i stor grad modellapparatet som mange kraftprodusenter benytter i sin produksjonsplanlegging og risikostyring på en forutsetning om at stokastiske tilsig er den viktigste driveren for utviklingen i områdeprisen. Kraftmarkedet er i stadig utvikling, noe som kan ha svekket den rollen som tilsig historisk har spilt som den viktigste driveren for utvikling i spotprisen.

Sett fra et samfunnsøkonomisk perspektiv er det mye å hente på et velfungerende modellapparat. Dette er med på å forhindre kraftkriser samtidig, som det skal behandle kraftressursene så effektivt som mulig. Derfor tar oppgaven for seg å se på hvor stor sammenhengen mellom nedbør, energitilsig, magasinifylling og snømengde mot systemprisen ved Nord Pool.

1.1 Optimeringsproblemet

På grunn av Norges vannkraftressurser har Norge en helt unik tilgang på rimelig fornybar energi som svært få andre land i verden har. Norsk vannkraft har også fordel med at produksjonen kan lett tilpasses forbruket av elektrisitet så lenge det er vann i magasinene. Likevel er det en viss risiko rundt dette da den totale kraftproduksjonen avhenger i stor grad av klimatiske forhold. Vinteren 2010/11 fikk strømkundene merke rekordhøye strømpriser som var resultat av en tørr høst og en kald vinter.

Vannkraftverkene i det nordiske kraftsystemet kan kategoriseres inn i elvekraftverk og magasin kraftverk. Elvekraftverkene må hovedsakelig produsere mens det er vann tilgjengelig, derfor kan ikke disse tilpasse seg forbruket på samme måte som magasin kraftverkene. På grunn av den unike egenskapen et magasin kraftverk med å kunne tilpasse produksjonen etter forbruket er det mye å tjene på god produksjonsplanlegging. For å få dette til er det nødvendig å ha gode prisprognoser. Spotprisen avhenger av svært mange faktorer. I det nordiske kraftmarkedet er hydrologiske forhold antydning å være en av de viktigste driverne av spotprisen. Dette er temaet denne oppgaven skal omhandle.

Opgaven løses med korrelasjon- og regresjonsanalyse av systempris, nedbør, energitilsig, magasinifyllingsgrad og tilgjengelige ressurser. Korrelasjonsanalysen gir svar på hvor sterkt systemprisen samvarierer med de ulike hydrologiske variablene. Regresjonen bygger videre på funnene i korrelasjonsanalysen og resultatet viser hvor stor grad en eller flere hydrologiske variabler forklarer systemprisen.

1.2 Oppgavens struktur

Kapittel 2 beskriver hvordan det nordiske kraftmarked er bygd opp, hvordan det fungerer og ulike karakteristikk ved kraftprisen. Det gir også et kort overblikk over oppbygningen til andre modeller som er blitt brukt til å forklare kraftpris. Kapittel 3 presenterer problemstilling, forskerspørsmål og metoden som oppgaven er utført etter. De viktigste resultatene fra undersøkelsen blir presentert og diskutert i kapittel 4.

2 Beskrivelse av kraftmarkedet

I dette kapittelet beskrives hvordan kraftmarkedet i Norden fungerer. Først presenteres litt bakgrunnsstoff som viser utviklingen til kraftmarkedet og årsaken til hvorfor kraftmarkedet er utformet som det er. Videre beskrives hvordan det fysiske markedet fungerer og hvordan dette påvirker kraftmarkedet der kraftbørsen Nord Pool er meget sentral. Til slutt er noen eksisterende modeller som er utviklet for å forklare kraftpris beskrevet.

2.1 Forløpet til dagens kraftmarked

Fram til 1990 var det hvert enkelt fylke som hadde plikt til å forsyne sitt geografiske område med elektrisitet. Kraftprisen ble satt med utgangspunkt i kostnadene ved å drive elektrisitetsforsyningen i området. I 1990 fikk Norge en ny energilov, noe som gav grunnlag for en markedsbasert omsetning av elektrisk energi i Norge. Noe av hensikten med loven var å jevne ut kostnader mellom ulike områder, at produksjon og fordeling drives effektivt og kraftprisen skulle gi signal om hvilke kraftutbygginger som kan være lønnsomme (Lindebæk-Nilsen & Strand, 2004, p. 2). Loven førte også til at Statkraftverkene, som den gangen hadde monopol på kraftoverføring og produksjon av kraft, ble splittet i to. I dag kjenner vi disse som Statkraft og Statnett, hvor Statkraft har ansvaret for den konkurransebaserte produksjonen av kraft og Statnett har ansvar for den monopolbaserte utvikling og vedlikehold av sentralnettet (Statnett, 2009).

Hovedaktørene i kraftmarkedet er produsenter, kraftleverandører, forbrukere og kraftbørsen Nord Pool. Den største andelen av omsetningen går gjennom kraftbørsen Nord Pool som leverer kraft til sluttmarkedet hvor forbrukerne befinner seg. I 2010 ble 74 % av kraften i det Nordiske markedet omsatt gjennom Nord Pool ("Striking the Balance - Annual Report 2010," 2011, p. 3). Erfaringer etter dereguleringen synes å være positive og kraftmarkedet i Norden regnes som det best fungerende kraftmarkedet i verden (Lindebæk-Nilsen & Strand, 2004, p. 3).

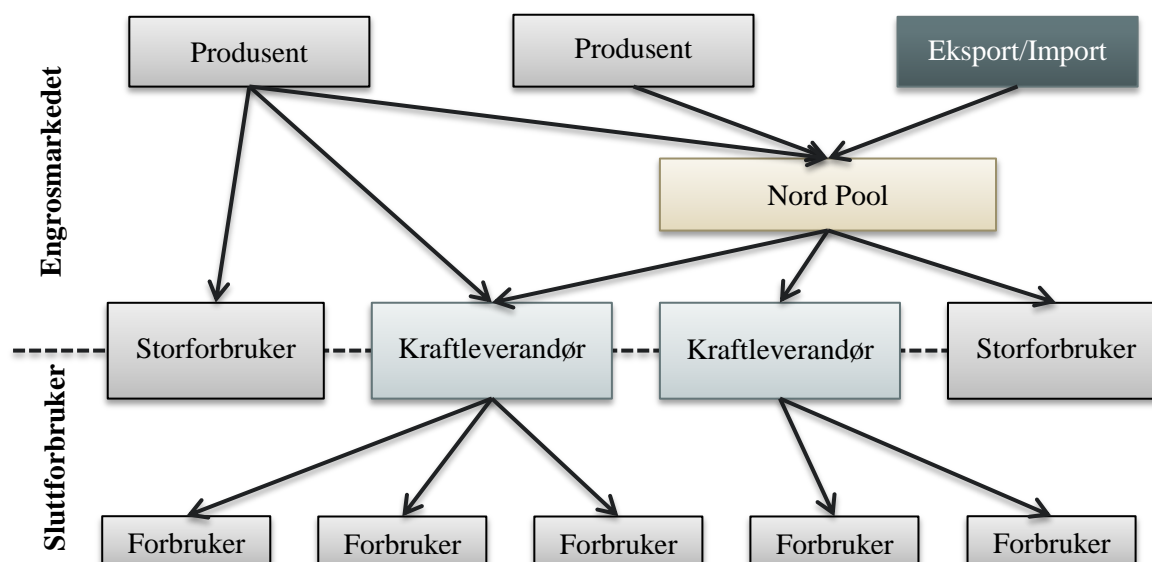
2.2 Kraftmarkedets virkemåte

I Norge sendes produsert elektrisitet ut på et felles overføringsnett. Når kraften er sendt inn på overføringsnettet er det fysisk umulig å skille ut hvor kraften kommer. Derfor måles produsert og forbrukt kraft. Produksjonen vil alltid tilsvare forbruk og nettap, og ut fra dette er det mulig å regne ut hvor mye forbrukeren skal betale sin leverandør av elektrisitet. Dette gjør det også mulig å kjøpe kraft som er produsert langt unna uten at kraften fysisk overføres over lange avstander. Ordningen som kalles punkttariff gjør markedsadgangen enkel for kundene og legger til rette for et landsomfattende marked for kraft (Spot).

Kraftlinjene i de nordiske landene er eid av staten. I Norge er det Statnett som drifter sentralnettet, en rolle som kalles transmission system operator (TSO). TSO'en er også ansvarlig for landets

forsyningssikkerhet, noe som teknisk sett innebærer å holde nettfrekvensen mellom 49,9 og 50,1 Hz. Regulerkraftmarkedet skal sørge for at det hele tiden er balanse mellom produksjon og forbruk. Her brukes tilbud for kjøp og salg fra produsenter og forbrukere som har mulighet til å endre forbruk og produksjonsplaner på kort sikt. I tilfeller hvor forbruket blir høyere enn produksjonen vil frekvensen falle, og TSO'en må sørge for at produksjonen øker (oppregulere). Hvis produksjonen blir høyere enn forbruket øker frekvensen, og da sørger TSO'en at produksjonen reduseres (nedregulere) (Spot).

I Norden omsettes mesteparten av elektrisiteten gjennom Nord Pool. Handelen kan også gjøres utenom Nord Pool ved at en storforbruker handles direkte av en produsent. Dette kalles bilaterale avtaler. Dette blir det stadig mindre av, men det er likevel vanlig måte å gjøre det på. Det skilles vanligvis også mellom engros- og sluttbrukermarked (Figur 2.1). I engrosmarkedet handles det mellom kraftprodusenter, kraftleverandører og store industrikunder. I sluttbrukermarkedet viderefremmes kraften av en kraftleverandør til vanlige forbrukere.



Figur 2.1 Aktørene i kraftmarkedets (Nome, 2010)

Hvordan kraften produseres ellers i Europa viker veldig fra de Nordiske landene. I Norge og Norden er 98 % og i overkant 50 % vannkraft (ENTSO-E, 2008), mens i Europa er det bare 1,4 % (Sæthre & Kristoffersen, 2006). Denne forskjellen betyr at prisene i Tyskland ikke avhenger av nedbør og snøsmelting på samme måte som det Nordiske.

2.3 Kraftbørsen Nord Pool

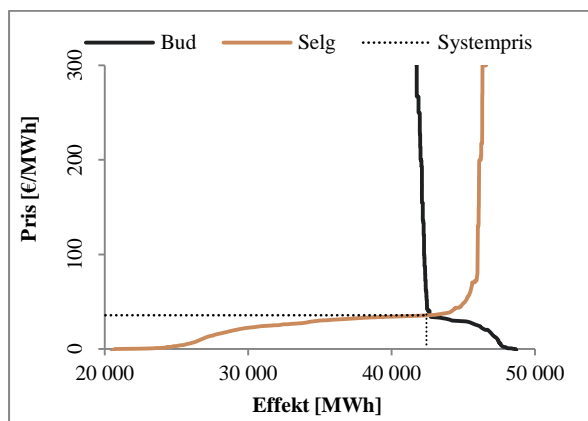
I dag består det nordiske kraftmarkedet av Norge, Sverige, Danmark, Finland. En fordel med et felles kraftmarked er at landene kan hjelpe hverandre i perioder når behovet for kraft er ekstra stort, noe som optimaliserer bruken av tilgjengelig kraft og reduserer lokale kraftunderskudd. Dette øker

samfunnsnytt og den samfunnsøkonomiske gevinsten til hvert av landene. Nord Pool-gruppen består av tre hovedkategorier; et fysisk marked, et finansielt marked og clearingvirksomhet.

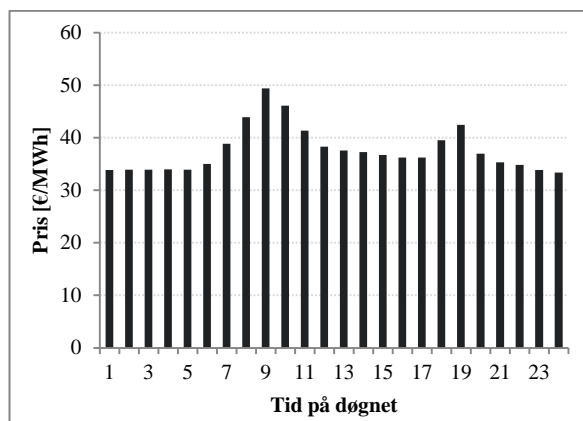
2.3.1 Det fysiske markedet

Elsport er det fysiske markedet på Nord Pool hvor kraften for neste døgn handles. Energifrisen fastsettes for hver time i døgnet, dette er gjort på grunnlag av aktørenes samlede kjøps- og salgsmeldinger. I 2010 ble 92 TWh i Norge omsatt på Elspot ("Striking the Balance - Annual Report 2010," 2011).

Systemprisen settes ut fra skjæringspunktet til aggregerte tilbud og etterspørsel, som registreres dagen før, som vist i Figur 2.2. I denne beregningen blir det ikke tatt hensyn til begrensninger i linjekapasiteten. Når det ikke er noen flaskehals i overføringslinjene vil områdepris tilsvare systempris. Spotprisen settes for hver time i døgnet (Figur 2.3).



Figur 2.2 Elspot systempriskurve (27. feb. kl. 10:00 - 11:00)



Figur 2.3 Endring i spotpris i løpet av 27. feb.

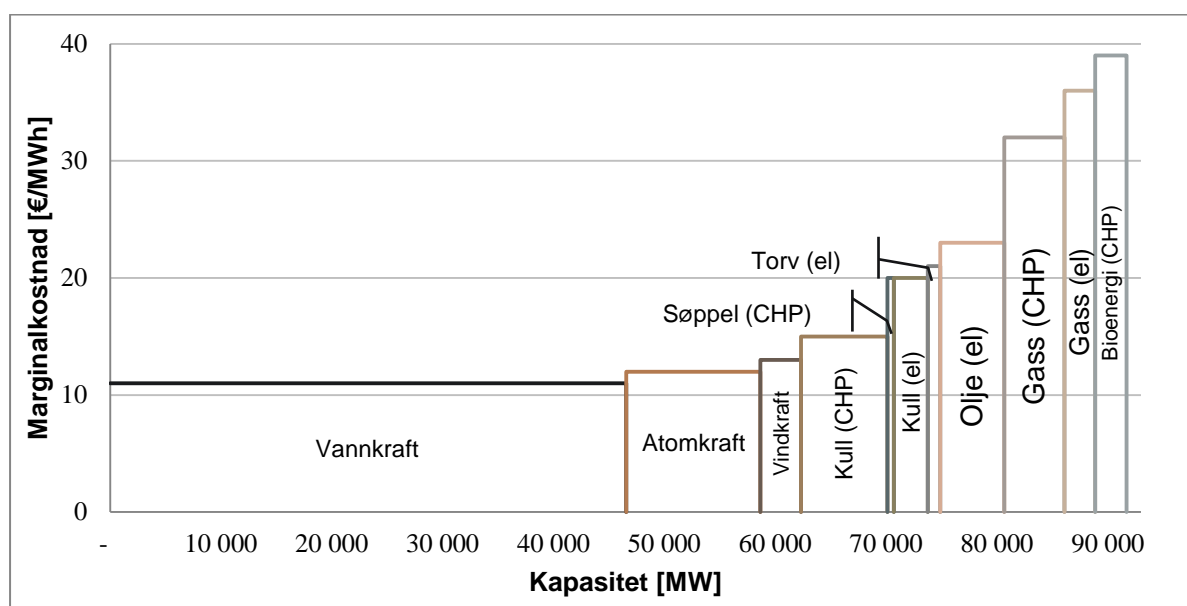
Intradagsmarkedet for handel av fysiske timekontrakter kalles *Elbas*. I dette markedet kan produsenter og forbrukere handle for å rette opp ubalanse som kan ha oppstått etter handelen i Elspot fram til en time før driftstimen I 2010 ble 88,9 GWh omsatt på Elbas ("Striking the Balance - Annual Report 2010," 2011).

Regulerkraftmarkedet brukes som et virkemiddel for TSO benytter for å opprettholde balanse mellom forbruk og produksjon i øyeblikket når kraften produseres. Dette er for å holde nettfrekvensen på 50 Hz.

Det nordiske kraftmarkedet er delt inn i ulike elspotområder, som vist i Figur 2.4. Dette skyldes begrenset overføringskapasitet i kraftnettet, da dette ikke er dimensjonert for at kraften kan produseres og forbrukes hvor som helst.

Siden markedet ble deregulert i 1991 har markedet utvidet seg geografisk og nye områder har blitt opprettet. I 1996 ble en norsk-svensk utveksling etablert. Finland ble med i 1998 og Danmark i 2000 ("Striking the Balance - Annual Report 2010," 2011). De nordiske landene er delt i flere

(Fiorenzani, 2006). Kraftproduksjonens marginalkostnad gir et godt estimat over tilbudskurven. Dette er fremstilt i Figur 2.5 i et merit order diagram, hvor installert effekt er rangert etter marginalkost.



Figur 2.5 Merit order diagram for Nord Pool (ENTSO-E, 2008)

Vannkraft dominerer det Nordiske kraftmarkedet, og i Norge er 98 % av elektrisiteten produsert fra vannkraft (Tabell 2.1). På grunn av at prisforskjellen mellom vannkraft og de andre energikildene vil prisendringen bli større jo mindre vannkraft som er tilgjengelig.

	Nordel	Danmark	Finland	Island ¹	Norway	Sverige
Totalt forbruk [TWh]	412,7	36,1	87,0	16,6	128,9	144,1
Maksimal last [GW]	61,0	6,1	12,5	1,7	18,4	22,2
Produksjon [TWh]	414,0	34,6	74,1	16,5	142,7	146,0
Fordeling av produksjon						
Vannkraft [%]	58	0	23	75	98	47
Atomkraft [%]	20	-	30	-	-	42
Annen termisk kraft [%]	19	80	47	0	1	10
Vindkraft [%]	3	20	0	-	1	1
Geotermisk energi [%]	-	-	-	25	-	-

Tabell 2.1 Produksjon og forbruk i de Nordiske landene (ENTSO-E, 2008)

2.4 Kraftprisens priskarakteristikker

En viktig egenskap til strømprisen er at den er svært volatile og har flere unike karakteristikk sammenlignet med andre finansielle markeder. På grunn av disse er strømprisen vanskelig å beskrive på samme måte som i aksjemarkedet (Knittel, Roberts, Regulation, & Institute, 2001). De største forskjellene finnes i at prisen på elektrisitet ofte er forskjellig fra marginalkostnaden for produksjon av elektrisitet. Dessuten kan ikke elektrisitet lagres, men må benyttes i øyeblikket den

¹ Er ikke på Nord Pool

produseres. Strømprisen påvirkes i utgangspunktet av faktorer knyttet til tilbud og etterspørsel, som igjen er knyttet til nedbørsmengde, temperatur, brenselpriser og flaskehals i overføringsnettet (Refvik & Djupvik, 2010).

2.4.1 Sesongvariasjoner

Spotprisen beveger seg stokastisk gjennom dager, uker, sesonger og år. Det er påvist at sesongvariasjonene gjelder spesielt for det Nordiske markedet, mye på grunn av klare skiller mellom årstidene (Clewlow & Strickland, 2000). I Norge preges vinteren av lite tilsig og høyt forbruk. Årlig brukes 45 – 50 TWh til oppvarming i bygninger, hvor 30 TWh dekkes av elektrisitet (NVE, 2009).

2.4.2 Lagringsmuligheter

Fordeler og ulemper med å ha en råvare tilgjengelig kalles ofte *convenience yield*. For å forklare dette begrepet bedre kan man si at en kraftprodusent har negativ *convenience yield* når magasinene er fulle, på grunn av at produsenten ønsker å selge i stedet for at vannet la renne utenom kraftverket. Når et kraftverk har lite vann i magasinene får det positiv *convenience yield*. Dette er på grunn av at de risikerer å tape inntekter om magasinet går tomt.

Dette påvirker kraftprisen, da produksjon og forbruk må skje samtidig og at forbrukeren ikke har mulighet til å lagre energien (Gjolberg & Johnsen, 2001). Kraftprodusenter i Norge har derimot mulighet til å lagre energi i vannmagasiner. Dette gir skjevheter i kraftmarkedet hvor produsenten stiller sterkere enn forbrukeren.

2.4.3 Volatilitet

Volatiliteten angir styrken og størrelsen på de stokastiske endringene i kraftprisen. Selv om volatiliteten ofte modelleres som konstant, kan den i seg selv være volatil og bør modelleres som en kombinasjon av tidsvarierende og stokastiske ledd (Pilipović, 2007).

At elektrisiteten må brukes umiddelbart kombinert med at etterspørselen ikke er særlig priselastisk² på kort sikt gjør at kraftprisen blir meget volatil (Fleten, Wallace, & Tomasgard, 2001). Volatilitet er estimert til å være 189 % i det nordiske kraftmarkedet i perioden 1993-1999, og det er ikke funnet bevis på en ikke-stasjonær volatilitet til spotprisen (Lucia & Schwartz, 2002). Ved større etterspørsel av kraft er volatiliteten høyere og omvendt (Knittel et al., 2001). Dette gjør også at det er meget høy risiko for ekstrempriiser i kraftmarkedet, noe aktørene velger å hedge³ ved hjelp av derivater i det finansielle markedet (Refvik & Djupvik, 2010). Forskjellen mellom volatiliteten på Nord Pool og EEX er over tre ganger, hvor den hos EEX ligger på 625 % (Sæthre & Kristoffersen,

² Liten priselastisitet medfører at endringen i prisen medfører liten endring i forbruket

³ Betyr sikring, er en investering som gjøres for å minimere eller fjerne risiko ved en annen investering

2006). Dette skyldes i at det er større kostnader ved endringer i produksjonskapasitet ved termiske anlegg enn vannkraft.

2.4.4 Mean-reversion og prissjokk

Med mean reversion mens at man vil ha svingninger på kort sikt, men på lang sikt vil prisen gå mot likevekt. Prissjokk eller rask mean reversion oppstår ved unormal etterspørsel i kraftmarkedet (Lund & Ollmar, 2002). Strømprisene viser en sterk trend til mean reversion i motsetning til det som observeres i aksjemarkedet (Knittel et al., 2001). I følge en artikkel som Schwartz ga ut i 1997 er en naturlig forklaring på men reversion på råvarer er at tilbudet av varer øker som en følge av at flere produsenter med høye kostnader kommer inn i markedet. Dette er mye omtalt av Eydeland and Wolyniec (2003) og har støtte i mange andre publiserte artikler om energipriser (Benth, Kiesel, & Nazarova, 2011), (Weron & Misiorek, 2008).

Prissjokk skjer ofte i kraftmarkedet. I perioden 1993 til 1999 er elspotprisens kurtose på 3,5 (Lucia & Schwartz, 2002), til sammenligning har en normalfordeling en kurtose på 3. Dette viser at det er høyere sannsynlighet for ekstrempriser i kraftmarkedet enn det en normalfordeling vil forklare. Kraftprisens skjevhet er positiv i samme periode. Det indikerer at sannsynligheten for høye prissjokk er større enn for lave priser.

2.5 Modellering av kraftpris

Trenden går mot et mer liberalisert kraftmarked i Europa, noe som har gjort at interessen for å forklare kraftprisen øker. Derfor er det gjort en rekke studier som forklarer prisen. Prognoser i kraftmarkedet er nødvendig for en optimal strategi for salg av kraft i et deregulert marked. Spotprisen har ulike egenskaper som sesongsvingninger, prissjokk og mean reversion (Eydeland & Wolyniec, 2003). Sesongsvingningene kan observeres daglig, ukentlig, månedlig og årlig. Prissjokkene kan utløses av blant annet uforutsette endringer i værforhold, og går tilbake til sin gjennomsnittlige verdi i ulikt hastighet.

Magasinfyllingen er en viktig variabel når spotprisen forklares (Gjolberg & Johnsen, 2001), (Botterud, Bhattacharyya, & Ilic, 2002), (Førsund, Hoel, & Mattei, 2004), (von der Fehr, Amundsen, & Bergman, 2005).

De mest nærliggende arbeidet sammenlignet med denne oppgaven som er funnet, er en masteroppgave som har testet modeller basert på fundamentalinformasjon for spot- og kontraktspriser i det nordiske kraftmarkedet (Vallevik & Øyan, 2004). Ved å bruke tidligere pris, futureskontakt, hydorbalanse og temperatur har de oppnådd en R^2 på 0,92. Denne konkluderer med at det ikke er entydig hvilken faktor som har størst innvirkning på kraftprisen.

Mange andre studier av kraftprisen baserer seg også på Box-Jenkins ARIMA modell eller andre univariate modeller. Noen er utvidet til modeller med temperatur som eksogen variabel. Det er også eksempler på modeller som bruker multippel regresjon (Huovila, 2003). Nome (2010) har i sin utredning av kraftprisen i Nord-Norge bruker ARIMA i sin prognosemodell. Her er det også gjort forsøk på dynamisk regresjon med temperatur og avvik fra normalstand som eksogene variabler. Resultatene viste seg at ingen av disse variablene var signifikante når laggede prisvariabler ble tatt med.

3 Problemstilling og metode

I dette kapittelet utledes problemstillingen og metoden blir beskrevet. Første del definerer forskerspørsmål, videre beskrives hvordan teorien i kapittel 3 kobles sammen med metoder for å løse problemstillingen. Framstillingen av metoder angående regresjon baserer seg i hovedsak på Studenmund (2011), mens modellering av tidsserie baserer seg på Chatfield (2004) og Tsay (2005).

3.1 Problemstilling og forskerspørsmål

Forskerspørsmålene i denne oppgaven retter fokus mot hvordan systemprisen påvirkes av hydrologiske faktorer. Dette er kanskje en av de viktigste driverne for spotprisen i Norden påvirker måten produksjonsplanleggingen foregår på i stor grad. Tema og problemstilling for denne oppgaven defineres:

Hvor stor innflytelse har hydrologiske forhold på systemprisen ved Nord Pool?

Dette spørsmålet er en mindre del av problemstillingen som angår prising av elektrisitet.

I denne oppgaven har jeg valgt å belyse problemstillingen ved hjelp av følgende forskerspørsmål:

- I. Hvor sterk er korrelasjonen mellom systempris og magasinifylling, nedbør, tilsig eller snømengde?
- II. Hvor stor forklaringsgrad har magasinifylling, nedbør, tilsig og snømengde på systemprisen?

3.2 Bakgrunn for problemstillingen

En av de største problemstillingene til en vannkraftprodusent er hvordan vannet i et magasin disponeres. Det vil være ønskelig for produsenten å slippe gjennom optimal mengde vann til enhver tid. For å gjøre dette på best mulig måte er planlegging av produksjonen avgjørende. Et slikt problem inneholder en del usikkerhet, derfor må denne tas hensyn til i planleggingen.

I Norden har også kraftmarkedet noen særegenheter som skiller seg fra andre markeder ved at det er deregulert og at en stor andel er produsert fra vannkraft. I 2008 ble 57 % av elektrisitetsproduksjonen i Norden (eks. Island) produsert fra vannkraft (ENTSO-E, 2008). I Norge ble over 98 % av elektrisiteten produsert fra vannkraft i samme periode.

Optimering av produksjonsplanlegging er et svært omfattende tema. Det er gjort en rekke andre studier som forsøker å forklare kraftprisen på ulike måter. Noen bruker blant annet hydrologiske faktorer til dette. Likevel er omfanget av denne forskningen begrenset.

Dereguleringen av det norske kraftmarkedet og opprettelsen av kraftbørsen Nord Pool ble produksjonsproblemet endret. Det regulerte markedet førte til at kraftprodusentene selv måtte sette

en verdi på kraften de produserte. I et velfungerende kraftmarked vil både kraftprodusent og forbruker ha nytte av at kraften som produseres prises på de riktige premissene. Dette skal gi insentiver til kraftprodusentene å disponere ressursene mest mulig effektivt. Gode prisprognoser er en avgjørende faktor i dette problemet. Svaret på denne oppgaven skal kunne være med på gjøre eksisterende prisprognoser bedre, eller reise nye spørsmål angående måten prisingen skjer på i dag.

3.3 Forskningsdesign

Metoden oppgaven bruker til å besvare forskerspørsmålene er korrelasjonsanalyse og regresjonsanalyse. Korrelasjonsfaktoren ρ beskriver styrken til korrelasjonen mellom systempris og gitt variabel, som brukes til å besvare det første forskerspørsmålet. For å oppnå linearitet mellom spotpris og de hydrologiske variablene utføres det transformasjon og filtrering av tidsseriene. Systemprisen og magasinfullingsgraden blir dekomponert, dette er også med på å øke lineariteten mellom systempris og forklarende variabler.

Regresjon er en av de mest brukte statistisk regneteknikkene til å beregne avhengigheter mellom en gitt variabel og en eller flere andre variabler. Mer spesifikt går dette ut på å forklare variasjoner i en variabel ved hjelp av en eller flere andre variabler. Bruk av regresjon i denne oppgaven vil kunne kvantifisere i hvor stor grad ulike hydrologiske faktorer påvirker systemprisen på Nord Pool.

Regresjonsanalysens determinasjonskoeffisient R^2 gir svar på hvor stor forklaringsgrad en eller flere forklarende variabler har på systemprisen. Oppgaven fokuserer på å forklare absolutt systempris. Dette gir svar på det andre forskerspørsmålet. Regresjonen som gjøres bygger videre på funnene i det første forskerspørsmålet ved at filtre, transformasjoner og dekomponering også benyttes i denne analysen. Signifikansen til regresjonen, samt test av at forutsetningene som regresjonsmetoden som benyttes er overholdt. I forbindelse med regresjonsanalysen reises også problemstillingen om årsakssammenhengen er kausal, noe som testes med krysskorrelasjon.

3.3.1 utfordringer

Problemstillingen er omfattende og kompleks, dette har gjort det krevende å kunne sette avgrensninger som er passende for oppgaven. Oppgaven fokuserer bare på en del av det som påvirker prisen, dette gir blant annet problemer med å teste residualene i regresjonen som gjøres på grunn av at det er utelatt flere forklarende variabler. Derfor er det lagt mindre fokus på analysen av residualene. Dekomponering av tidsseriene og fjerning av trend og sesonger kan være med på å forstyrre forklaringsgraden til de hydrologiske variablene da det er ukjent hva som er årsaken disse komponentene.

3.3.2 Utførelse

Mesteparten av analysene gjøres i Excel, og noe er blitt gjort med Statgraphics Centurion XVI. Statgraphics er profesjonell programvare spesialisert for statistiske analyser, og ble valgt på grunnlag av god funksjonalitet, samt et brukergrensesnitt som er forholdsvis lett å lære seg.

3.4 Empiriske data

Det ble det brukt empiriske data i tidsrommet uke 36 år 1998 fram til uke 16 2012. Systemprisen er hentet i fra Nord Pool Spots FTP server, og timesprisen er blitt regnet om til en ukentlig aritmetisk gjennomsnittsverdi. Nedbør og energitilsig er hentet fra Agder Energi. Nedbøren er oppgitt hvert døgn, derfor har denne blitt summert sammen til ukentlige verdier.

3.4.1 Hydrologiske variabler

Magasinbeholdning M_t og nedbør N_t er de eneste verdiene som er målt direkte. Energitilsiget T_t beregnes ut fra energi som er blitt produsert pluss energiekvivalenten til endringen i magasinbeholdningen. Snømengden blir beregnet ved å trekke energitilsiget fra nedbøren. Alle dataene er oppgitt som GWh. Det er knyttet noe usikkerhet til nedbørsdataene på grunn av måten disse blir målt på. Dette gjelder spesielt i perioder med snø og mye vind. Dette kan påvirke MKM regresjonen (kapitel 3.5.6), som forutsetter at de forklarende variablene er målt korrekt.

3.4.2 Tilgjengelige ressurser

Med tilgjengelige ressurser menes magasinutfyllingsgrad og snømengde. Tilgjengelige ressurser R_t ved tiden t beregnes ved

$$R_t = M_t + S_t \quad \text{Formel 3.1}$$

hvor M_t er absolutt magasinutfyllingsgrad i GWh og S_t er relativ snømengde i GWh. I analyseperioden har magasin kapasiteten variert mellom 84,1 – 84,4 TWh, og det antas at dette ikke påvirker resultatene. S_t blir beregnet ved å summere alle endringene i snømengden ΔS_t gjennom hele analyseperioden.

$$S_t = \sum_{t=1}^t \Delta S_t \quad \text{Formel 3.2}$$

På grunn av at ΔS_1 er uke 36 år 1998, blir snømengden regnet som 0 GWh på dette tidspunktet. I perioder med snømengde under dette nivået er verdiene negative, slik at $R_t < M_t$. I praksis er dette misvisende, men dette vil ikke påvirke resultatene fra korrelasjons- eller regresjonsanalysen. Endringene i snømengden regnes ut nedbør N_t og tilsig T_t :

$$\Delta S_t = N_t - T_t \quad \text{Formel 3.3}$$

3.5 Korrelasjon og regresjonsanalyse

Regresjon har noen likheter med korrelasjon. Korrelasjonen er et mål på lineær avhengighet mellom to variabler. I motsetning til regresjon vil ikke korrelasjon nødvendigvis føre til at den ene variabelen påvirker den andre. Korrelasjon indikerer bare om det finnes samvariasjon mellom to variabler, mens i en regresjon påvirkes den forklarende variabelen av en eller flere uavhengige. Dette gjør regresjon til et mer nyttig og fleksibelt verktøy enn korrelasjon.

3.5.1 Korrelasjonskoeffisienten ρ

Korrelasjonskoeffisienten, som ofte er kalt Pearsons ρ , er et mål styrken og retningen til den lineære avhengigheten mellom to variabler. Verdien til denne går fra -1 til 1, hvor -1 eller 1 er perfekt korrelasjon, mens verdi lik null sier at det ikke finnes korrelasjon. Er koeffisienten positiv betyr det at x blir større når y blir større, er den negativ blir x mindre når y blir større. Korrelasjonskoeffisienten beskrives ved

$$\rho_{x,y} = \frac{\text{cov}(x,y)}{\sqrt{\text{var}(X)} * \sqrt{\text{var}(Y)}} = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}} \quad \text{Formel 3.4}$$

hvor x og y er variabler og n er antall «par» observasjoner som er gjort. Korrelasjonsfaktoren tilsvarende den kvadrerte verdien til determinasjonskoeffisienten i en enkel regresjon. Derfor brukes ρ i stedet for R^2 når det er bare korrelasjonene som ønskes å undersøkes. En fordel med korrelasjonskoeffisienten er den er lett å forstå. Men den sier ikke noe om hvilke av variablene som påvirker hverandre, eller om den observerte sammenhengen er kausal.

3.5.2 Determinasjonskoeffisienten R^2

En av de enkleste og mest brukte metodene til å måle regresjonens presisjon på er determinasjonskoeffisienten R^2 . Denne forklarer forholdet mellom regresjonens kvadrerte forklarende verdier og kvadratet av de observerte verdiene

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{\sum \varepsilon_i^2}{\sum (Y_i - \bar{Y})^2} \quad \text{Formel 3.5}$$

hvor ESS er summen av kvadrerte feilverdier, TSS er ESS pluss regresjonens kvadrerte verdier. Regresjonens R^2 besvarer i hvor stor grad en eller flere hydrologiske faktorer påvirker spotprisen. Er verdien 0 er det ikke noe sammenheng mellom avhengig og uavhengig variabel, er verdien 1 forklares 100 % av den avhengige variabelen av de uavhengige. I prisprognoser ønskes derfor en så høy R^2 verdi som mulig. I en enkel regresjon kan R^2 også defineres ved å kvadrere verdien av korrelasjonsfaktoren ρ mellom y og x . Den største fordelingen med å bruke R^2 som mål er at den er enkel å forstå og gir en god indikasjon på hvordan modellen passer dataene.

Et problem med R^2 er at verdien aldri vil synke om det legges til ekstra uavhengig variabler. Dette resulterer i at regresjonen som inneholder flest uavhengige variabler alltid får størst forklaringsgrad, noe som ikke alltid stemmer i praksis. Derfor kan R^2 justert gi et bedre mål når ulike regresjoner skal sammenlignes da den justeres for grad av frihet. Dette medfører også at R^2 justert også kan gi negative verdier. Denne verdien brukes når antall forklarende variabler skal vurderes. Selv regresjonen bør ha så høy forklaringsgrad som mulig, er det fare for at man lager en regresjon som optimerer denne for mye. Dette vil resultere i spuriøse resultater. Derfor må regresjonen begrunnes ut fra sunn fornuft og økonomisk forståelse.

3.5.3 Regresjonens signifikans

Regresjonenes signifikans testes med prob (F) og p verdier. F-verdiene gir informasjon om hvor stor sannsynlighet det er for at regresjonen som helhet er ikke statistisk signifikant. De uavhengige variablenes p-verdi sier hvor stor sannsynlighet det er for at variabelen ikke er signifikant. Med signifikant menes at det er et lineært forhold mellom variabelen og spotprisen. Sjansen for at en koeffisient er null påvirker p-verdien, mens sjansen for at alle verdiene er null påvirker F-verdien. Signifikansnivået i oppgaven er satt til 5 %, det vil si at forklarende variabler med p-verdi over 0,05 blir forkastet.

Selv om en F- og p-verdiene gir svar på om den er statistisk signifikant, er ikke dette det samme som at sammenhengen er kausal. Virkning årsak forholdet er svært viktig å ta hensyn til i en regresjonsanalyse. Derfor er det nødvendig at en regresjonsanalyse brukes sammen med andre økonomiske teorier og sunn fornuft.

Resultatet av en regresjonsanalyse avhenger av at de riktige variablene brukes. Med dette menes at man ikke utelukker noen forklarende variabler, og at man heller ikke tar med irrelevante variabler. I denne oppgaven undersøkes fire forklarende variabler: magasinbeholdning, nedbør, tilsig og snømengde.

3.5.4 Stegvis regresjon

I en regresjonsanalyse kan man få unødvendig mange forklarende variabler å velge mellom. Stegvis regresjon er en metode som bygger en modell ved å legge til eller trekker fra variabler basert på om variablene er statistisk signifikante. R^2 -plottet viser hvor mye forklaringsgraden øker ved å legge til en variabel. Om forklaringsgraden ikke øker i særlig grad er dette tegn på at de resterende variablene ikke skal være med.

3.5.5 Systemprisfunksjonen

Regresjonsanalysen gjøres med ulike modell som forklarer spotprisen med en til fire hydrologiske variabler; magasinbeholdning, nedbør, tilsig og snømengde. Utgangspunktet for regresjonen er at systemprisen P_t kan forklares med

$$P_t = \alpha + \beta_1 M_t + \beta_2 N_t + \beta_3 T_t + \beta_4 S_t + \varepsilon_t \quad \text{Formel 3.6}$$

hvor M_t er magasininnfyllingsgraden oppgitt i prosent, N_t , T_t , S_t er nedbør, tilsig og nedbør i GWh. De forklarende variablene testes ut med ulike filtre og dekomponeringer, som er forklart i kapittel 3.6 og 3.7.

Vanligvis inneholder feilledet ε_t sanne tilfeldigheter, men på grunn av at regresjonen mangler forklarende variabler kan det ikke forventes at feilledet møter forutsetningene en MKM regresjon stiller.

3.5.6 Minste kvadraters metode (MKM)

Koeffisientene β_1 til β_4 er den deterministiske sammenhengen mellom kraftpris og hydrologiske variabler. Konstantleddet i modellen er α . Disse koeffisientene blir beregnet med minste kvadraters metode. I grove trekk går metoden ut på å finne en teoretisk sammenheng slik at forskjellen mellom empiriske hydrologiske data og spotpris blir minst mulig. De beregnede verdiene er en tilnærming til virkeligheten, og betegnes som $\hat{\alpha}$ og $\hat{\beta}_1 \dots \hat{\beta}_4$.

For at resultatene fra MKM regresjon skal bli mest mulig korrekte må et sett av forutsetninger være oppfylt:

1. Spotprisen må være en lineær funksjon av de hydrologiske variablene.
2. Residualene ε_t skal ha en forventet verdi på null og sannsynlighetsfordelingen skal være normalfordelt.
3. Residualene er ukorrelerte, dvs. ingen autokorrelasjon.
4. Konstant variasjon eller homoskedastisitet i residualene.
5. Forklaringsintervallene skal være på intervallskala eller være kategoriske. Er verdiene kategoriske (bruk av dummy variabler) skal variabelen kun ha to verdier.
6. Det må ikke være perfekt multikollinearitet, dette vil si at de hydrologiske variablene er 100 % korrelert med en av de andre forklarende variablene. Derfor ønskes lite korrelerte forklarende variabler. Ved multikollinearitet blir koeffisientene ustabile og p-verdiene høye.
7. De hydrologiske faktorene må være ukorrelerte med eksterne variabler som påvirker spotprisen. Eksempel på dette kan være at nedbøren påvirker temperaturen. Selv om ikke temperaturen er med i regresjonen, kan denne også være med å påvirke spotprisen.

3.6 Transformasjoner, filtreringer og krysskorrelasjon

Det første kriteriet i MKM regresjonen sier at spotprisen må være en lineær funksjon av de hydrologiske variablene, dessuten skal det også være homoskedastisitet i residualene. Dette

medfører at variablene som inngår i regresjonen må transformeres og filtreres for å oppnå kriteriene.

3.6.1 Logaritmisk transformasjon av spotprisen

På grunn av prissjokkene (omtalt i kapittel 2.4.4) gjøres det logaritmisk transformasjon av spotprisen. Dette gjør det mulig å få større grad av homoskedastisitet og redusere skjevheten i sannsynlighetsfordelingen til spotprisen. På denne måten passer funksjonsformen til spotprisen bedre inn i MKM regresjonen. Den logaritmiske systemprisen blir gitt ved:

$$P_{t,2} = \ln P_t \quad \text{Formel 3.7}$$

Når logaritmisk transformasjon benyttes fører dette som regel til at R^2 blir høyere enn det den i utgangspunktet er. Logaritmisk transformert spotpris er benyttet i alle analysene, derfor kan resultatene sammenlignes. Skal resultatene sammenlignes med andre modeller er det nødvendig at opprinnelig spotpris regnes ut med regresjonsmodellen, for å så måles opp mot observerte verdier.

3.6.2 Filtrering

På grunn av lagringsmulighetene (omtalt i kapittel 2.4.2) antas det at systemprisen reagerer lite på et kortvarig nedbørsperiode i forhold til en langvarig periode. For å modellere denne effekten brukes det et lineært filter på nedbørs- og tilsigsserien. I korrelasjonsanalysen og MKM regresjonen resulterer dette i økt linearitet mellom spotpris og nedbør eller tilsig.

Et lineært filter konverterer en tidsserie til en annen serie med en lineær operasjon. Dette er nyttig fordi tidsintervallet til de hydrologiske variablene må kanskje være forskjellig i fra tidsintervallet til systemprisen som skal forklares for å oppnå en god forklaringsgrad.

Dette testes ulike tidsintervaller med ulike tidsforsinkelser, og tidsintervallet som gir best forklaringsgrad benyttes. Den filtrerte nedbørs- og tilsigsserien for nedbør $N_{t,2}$ og energitilsiget $T_{t,2}$ beskrives med

$$N_{t,2} = \sum_{q=t-i}^t N_q \quad \text{Formel 3.8}$$

$$T_{t,2} = \sum_{q=t-i}^t T_q \quad \text{Formel 3.9}$$

hvor i er lengden på tidsintervallet og t er tiden observasjonen er gjort. Den minste tidsenheten som brukes er en uke. Svingninger innenfor dette intervallet antas ikke å være særlig påvirket av de hydrologiske variablene.

Det er også gjort forsøk med eksponentiell glatting. Denne metoden går ut på å vekte senere verdier sterkere enn de første verdiene i tidsintervallet. Vektfordelingen gjøres ved å justere parameteren α mellom 1 og 0.

3.6.3 Krysskorrelasjon

I en regresjonsmodell som skal brukes i prognosering er det viktig å kjenne årsakssammenhengen. Tidsforsinkelse, eller lag, brukes til å modellere eventuelle forsinkelser som kan oppstå i systemet. Denne kan være med på å gi en indikasjon på hvordan de ulike variablene påvirker hverandre, altså årsakssammenhengen. I likhet med autokorrelasjonsfunksjonen måler krysskorrelasjonsfunksjonen korrelasjonen mot observasjon gjort på et annet tidspunkt, men måler dette mot en annen variabel. Laget τ beskriver tidsforskyvningen i mellom tidsseriene. Ved tidsforskyvning kan korrelasjonskoeffisienten beskrives som $\rho(x_{t+\tau}, y_t)$.

Har ρ størst verdier når den hydrologiske variabelen har negativ tidsforsinkelse er dette med på å styrke hypotesen om at sammenhengen er kausal. Dette kalles en *ledende variabel*. Er ρ størst når den hydrologiske variabelens $\tau = 0$ er det vanskelig å bevise årsakssammenhengen. Er ρ størst når den hydrologiske variabelens $\tau > 0$ kan dette tyde på at den er et resultat av spotprisen. Dette kan bety at variabelen egner seg lite til å forklare spotprisen.

På grunn av at minste tidsintervall er en uke og at det er gjort filtreringer av tidsseriene kan dette påvirke resultatet.

3.7 Dekomponering av tidsseriene

For å møte forutsetning en, to og fire som en MKM regresjon har er det nødvendig å dekomponere tidsseriene i både den avhengige variabelen og noen av de uavhengige. I en tidsserieanalyse er antagelsen om at verdier i dataene er et resultat av tidligere verdier. Det som ønskes med en tidsserieanalyse er å identifisere fenomenet og kunne forutsi fremtidige verdier i tidsserien. Første steg i dekomponeringen av en tidsserie er å identifisere mønsteret i dataene

3.7.1 Identifikasjon av mønstre i tidsserien

Den klassiske metoden dekomponerer tidsserien ned i trend, sesong, andre sykliske variasjoner og andre irregulære svingninger. Disse irregulære svingningene gjør vanligvis at mønsteret blir vanskeligere å identifisere. Derfor er det vanlig å bruke ulike teknikker for å filtrere ut støyen slik at mønsteret blir mer tydelig. En vanlig fremgangsmåte i tidsseriemodellering er:

- Plot tidsserien og studer om serien inneholder trend, sesong eller andre sykliske variasjoner.
- Fjern funne fra tidsserien
- Velg en modell som passer residualene

I denne oppgaven er trend og sesong mest interessante. Ved dekomponering av systemprisen kan det velges mellom additiv eller multiplikativ modell. I den additive modellen er trenden og sesongen uavhengige av hverandre, mens den multiplikative er sesongen påvirket av trenden.

3.7.2 Autokorrelasjonsfunksjonen

Autokorrelasjonsfunksjonen (AKF) er et svært viktig hjelpemiddel når tidsserier skal modelleres. Denne måler korrelasjon i mellom observasjonene i med forskjellige tidsforskyvninger. AKF r_k med τ tidsforsinkelser er gitt ved $\rho(x_{t-\tau}, x_t)$

I et korrelogram er autokorrelasjonsfunksjonen r_k plottet som en funksjon av tidsforsinkelsen k . Visuell inspeksjon og tolkning av korrelogrammet kan gi nyttig informasjon om tidsserien. Dette kan være å se om serien er tilfeldig, inneholder autokorrelasjon, sesonger eller stasjonær.

3.7.3 Trend

Systemprisen inneholder trend. To vanlige måter å beregne trenden på er ved bevegelig gjennomsnitt eller med MKM. Beregnes trenden med MKM kan dette gjøres med:

$$P_{t,2} = m_t + P_{t,3} \quad \text{Formel 3.10}$$

Teknikken kan også bruke polynomuttrykk av høyere orden. Dette gjøres ved å legge til flere ledd av P_3 opphøyd en verdi ulik en.

Trenden kan også elimineres bort ved differensiering. Ved differensiering er det mulig å gjøre tidsserien stasjonær. Dette kan blant annet fjerne autokorrelasjon. På grunn av å begrense oppgavens omfang er ikke denne metoden brukt, men det kan være et godt alternativ på grunn av at kraftprisen i stor grad er bestemt av prisen som var i perioden før.

Trenden i tidsserien beskriver hvordan gjennomsnittet endrer seg i det lange løp. Denne påvirker utførelsen av undersøkelsen av det første forskerspørsmålet ved at trenden korrigeres bort. I prognoser som skal gå noen få perioder fram i tid egner metoden seg godt. Skal man prognosere mange perioder fram er det mest hensiktsmessig å fjerne trenden ved å trekke fra bevegelig gjennomsnitt eller lineær trend.

I det andre forskerspørsmålet er denne trenden vesentlig, derfor blir ikke prisen korrigeret for trend. Mer utfyllende diskusjon er presentert i kapittel 4.1.5.

3.7.4 Sesongvariasjoner

Som omtalt i kapittel 2.4.1 har kraftprisen variasjoner som skyldes forbruk og produksjon av kraft. En regresjonsanalyse med sesongjustert⁴ spotpris som uavhengig variabel vil resultere i kunstig høy forklaringsgrad, på grunn av at man i dette tilfellet også forklarer spotprisen med endring i forbruk. Utelates sesongjustering resulterer i for lav forklaringsgrad. Derfor analyseres begge tilfellene, og resultatene må gjøres i betraktning av dette. Sesongjustering av de forklarende variablene diskuteres i kapitlene hvor de analyseres.

En effektiv måte å korrigere for sesonger er å gjøre en regresjon for hver uke gjennom hele sesongen. Ulempen med dette er at det blir korte tidsserier i regresjonen, og spesielt i regresjoner med mange ledd blir resultatene mindre signifikante.

På grunn av dette er det også gjort ved å fjerne sesongkomponenten med additiv måte. Utgangspunktet i denne metoden er å lage en funksjon som beskriver sesongvariasjonen:

$$X_t = s_t + Y_t \quad \text{Formel 3.11}$$

Her er s_t sesongeffekten, mens Y_t er tidsserien uten sesongkomponent.

3.7.5 Spektral fordelingsfunksjonen

Om tidsserien inneholder en periodisk sinusformet komponent med en gitt bølgelengde kan denne forklares som

$$X_t = R \cos(\omega t + \phi) + Z_t \quad \text{Formel 3.12}$$

hvor ω er frekvensen til signalet, ϕ er faseforskyvning, R er amplituden, og Z_t er den gjennomsnittlige verdien. Både ω og ϕ er vanligvis oppgitt i radianer. I denne går sesongen over 52 uker, derfor kan den indre funksjonen skrives som $(2\pi * \text{uke}_t)/52 + \phi$. Metoden blir brukt i kapittel 4.4.2.

3.8 Innføring av ekstra ledd i modellen

For å modellere kraftprisens egenskaper er det forsøkt å legge inn flere ledd i regresjonsmodellen.

3.8.1 Kryssproduktledd

På grunn av convenience yield undersøkes det om at tilsig og nedbør påvirker prisen i større grad og resultere i prissjokk (kapittel 2.4.4) når avviket fra normalfylling er stort. Denne testen gjøres ved

$$P_{t,2} = \alpha + \beta x_t (M_{avvik,t} + 100) \quad \text{Formel 3.13}$$

⁴ Innebærer å korrigere tidsserien for variasjoner som skyldes sesonger

hvor $M_{avvik,t}$ er avviket fra normalen i prosent. Forutsetningen for testen er at prisøkningen er en lineær funksjon av magasinfyllingen.

3.8.2 Polynomregresjon

Om sammenhengen mellom systempris og forklarende variabler ikke er lineære kan polynomregresjon forklare sammenhenger som passer et polynomuttrykk av høyere grad. På grunn av at mange elvekraftverk ikke har mulighet til å lagre vannet som kommer, kan det forventes at økt tilsig i tørre perioder vil føre til større fall i systemprisen enn perioder med høyt tilsig. Dessuten er prisutviklingen i merit order diagramet (Figur 2.5) ikke lineært. Om dette er tilfellet vil samvariasjonen mellom systempris og hydrologisk tilsig ikke være lineær. Derfor gjøres det forsøk med følgende modell

$$P_{t,2} = \alpha + \beta_1 T_{t,2} + \beta_2 T_{t,2}^2 \quad \text{Formel 3.14}$$

Om β_2 verdien er negativ indikerer dette at systemprisen synker raskere jo mer nedbør som kommer. Er verdien positiv synker prisen saktere, eventuelt stiger jo mer nedbør som kommer.

3.9 Residualanalyse

For å kontrollere hvordan MKM regresjonen beskriver systemprisen undersøkes residualene for ulike karakteristikk. Residualene er feilverdien mellom regresjonens beregnede verdier og de observerte verdiene.

3.9.1 Autokorrelasjon

Autokorrelasjon, eller seriekorrelasjon, forteller oss noe om korrelasjonen mellom verdier i den samme tidsserien. Finnes dette i residualene i denne MKM regresjonen skyldes det antageligvis:

1. En forklarende variabel er utelatt fra regresjonen
2. Den regresjonsmodellen er ikke satt opp riktig. Dette kan skyldes at sammenhengen ikke er lineær, noe MKM forutsetter.
3. Tidsserien er filtrert for mye, noe som fører til at regresjonen blir for lite «følsom»

Det finnes ulike metoder å behandle modeller som inneholder autokorrelasjon. Differensiering, dvs. å bruke differansen i mellom verdiene i tidsserien, er en mye brukt metode (Box & Jenkins, 1970). Et annet alternativ er metoden foreslått av Cochrane and Orcutt (1987).

3.9.2 Test av autokorrelasjon

Den mest brukte testen for autokorrelasjon er Durbin-Watson testen. Testen tester autokorrelasjon mot verdien før og er definert som

$$DW = \frac{\sum_{t=2}^N (\hat{\epsilon}_t - \hat{\epsilon}_{t-1})^2}{\sum_{t=1}^N \hat{\epsilon}_t^2} \quad \text{Formel 3.15}$$

Verdiene av denne testen ligger i mellom null og fire. Er verdien to finnes det ikke autokorrelasjon. Er verdien null eller fire er det perfekt autokorrelasjon.

3.9.3 Normalfordelingen av residualene

Shapiro-Wilk W testen blir brukt for å teste om residualene kommer fra en normalfordeling. Hvis W statistikken er signifikant, betyr dette at residualene ikke kommer fra en normalfordeling. Konfidensintervallet settes til 0,05, dvs. at p-verdier over 0,05 er signifikante.

3.10 Bruk av resultatene

Resultatene gir skal kunne gi en indikasjon på hvor stor grad de utvalgte hydrologiske faktorene kan bestemme systemprisen. Resultatene viser at hydrologiske forhold har en stor innvirkning på systemprisen ved Nord Pool, men dette trenger ikke bety at disse variablene nødvendigvis bør være med i prisprognoser av systemprisen.

Begrepet R^2 i denne oppgaven referer til andelen av systemprisen som forklares ut fra en gitt modell. I denne sammenheng har variablene blitt behandlet på forskjellig måter for å oppnå linearitet mellom systempris og forklarende variabler. Dette kan påvirke verdien til R^2 ved at den blir høyere på grunn av tilfeldigheter.

3.10.1 Kritikk til modellen

En ulempe modellen er at den forklarer systemprisen på samme tid som målingene av de hydrologiske variablene er blitt gjort. Dette gjør det nødvendig med hydrologiske prognoser for å kunne prognosere systemprisen, og i og med at de hydrologiske prognosene i seg selv inneholder store usikkerheter vil prisprognosen være preget av mye usikkerhet. Et annet moment er at studien er basert på gjennomsnittlig systempris over en uke, og tar ikke høyde for variasjoner innenfor dette tidsintervallet.

Denne undersøkelsen har bare forsøkt å forklare absolutt spotpris. På grunn av systemprisens autokorrelasjon kan det være hensiktsmessig å løse oppgaven på er å differensiere spotprisen og forklare differansen. Alternativ til dette er å bruke Cochrane & Orcutt metoden til å fjerne autokorrelasjon. Det er også verdt å merke seg at korrelasjonsfaktoren ρ og determinasjonskoeffisienten R^2 blir lavere ved en slik modell, selv om den forklarer prisen like godt eller bedre.

På grunn av at systemprisen påvirkes av flere enn hydrologiske variabler vil regresjonen mangle viktige variabler. Dette gjør det vanskeligere å undersøke om regresjonen er utført på en

tilfredsstillende måte. I eventuelle prisprognoser må resultatene settes sammen med andre variabler som påvirker systemprisen.

4 Empirisk analyse og resultater

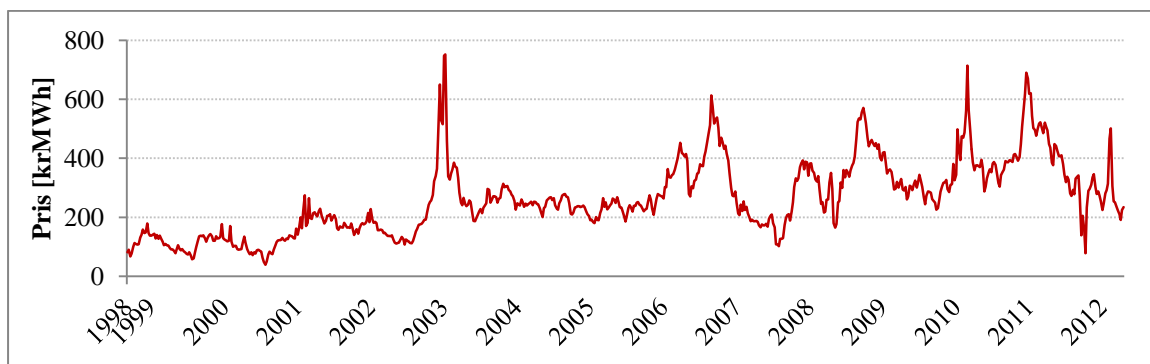
Denne delen presenteres og diskuteres de viktigste resultatene som påvirker svarene til forskerspørsmålene. I kapittel 4.1 analyseres systemprisen. Disse resultatene er grunnlaget for hvordan forskerspørsmålene besvares. I kapittel 4.2 til 4.5 undersøkes styrken på korrelasjonen i mellom systempris og de utvalgte hydrologiske variablene. Det gjøres forskjellige forsøk på å gjøre ulike filtreringer og transformasjoner for å oppnå linearitet mellom variablene. Kapittel 4.6 gjøres regresjonsanalysen for å finne svar på hvor mye de hydrologiske variablene til sammen kan forklare systemprisen.

4.1 Systemprisen

I denne delen analyseres systemprisen. Priskarakteristikkene som er beskrevet i kapittel 2.4 er grunnlaget for analysene som utføres. På grunn av dette transformeres og dekomponeres spotprisen i trend og sesongvariasjoner.

4.1.1 Plot av systemprisen

Figur 4.1 viser systemprisen som brukes i analysen. Plottet viser store svingninger i systemprisen (volatilitet), men at disse svingningene er kortvarige (prissjokk) og at de går tilbake til en gjennomsnittlig verdi (mean reversion). Prisen har i løpet av analyseperioden hatt en jevn stigning (trend). Det er også mulig å antyde sesongvariasjoner.



Figur 4.1 Plott av systemprisen P_t regnet ut til ukentlige gjennomsnittlige verdier uke 36 1998 til uke 16 2012

4.1.2 Skjevhet og kurtose

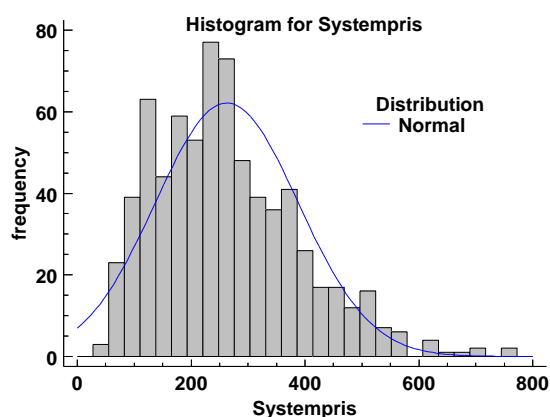
Skjevheten til systemprisen sier noe om sjansen for store prissprang (positiv skjevhet) eller store prisfall (negativ skjevhet) i forhold til om den var normalfordelt. Tabell 4.1 viser at systemprisen har positiv skjevhet, det tyder på at det er større sjanse for prishopp enn store prisnedganger. Ved en logaritmisk transformasjon endres denne skjevheten til en negativ verdi. Sammenlignes disse to fordelingene ser det ut til at logaritmen til systemprisen bedre en bedre tilnærming til normalfordelingen enn systemprisen.

Den overskytende kurtose til systemprisen er positiv. Til sammenligning har en normalfordeling overskytende kurtose på 0. I Figur 4.2 vises dette med at fordelingen er mye spissere rundt gjennomsnittet enn det en vanlig normalfordeling er. Logaritmen til systemprisen har en negativ kurtose, noe som indikerer at fordelingen ikke er like spiss (Figur 4.3).

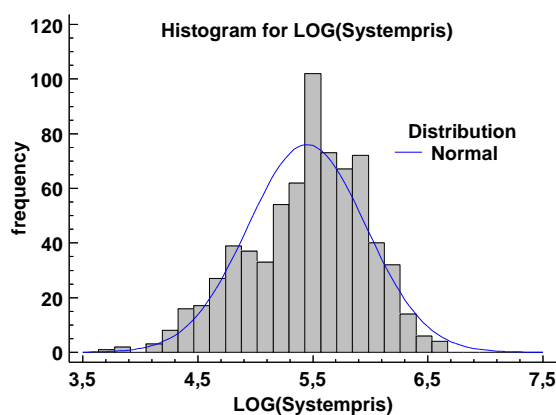
	<i>Systempris</i>	<i>Logaritmen til systemprisen</i>
Gjennomsnitt	262,59	5,4491
Standardavvik	125,52	0,513314
Skjevhet	0,784908	-0,441344
Kurtose	0,61574	-0,143815
Shapiro-Wilk W (p-value)	0,0	3,87497E-10

Tabell 4.1: Deskriptiv statistikk for systemprisen på Nord Pool

Shapiro-Wilk W testen tester sannsynligheten for at fordelingen kommer fra en normalfordeling. Resultatene forteller at det er svært liten sannsynlighet for at hverken systemprisen eller logaritmen til systemprisen kommer fra en normalfordeling. Likevel ser det ut til at det er logaritmen til systemprisen passer bedre en normalfordeling enn det systemprisen gjør.



Figur 4.2 Systemprisens histogram sammenlignet med normalfordelingen



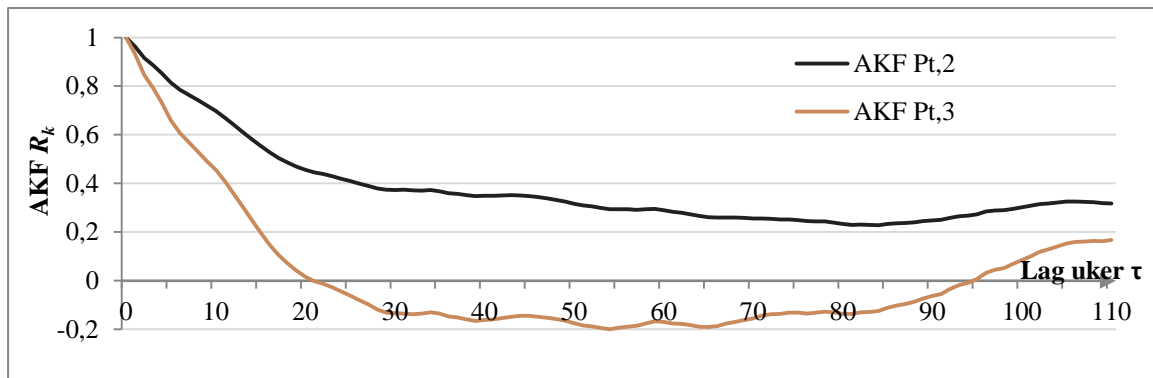
Figur 4.3 Logaritmen til systemprisen sammenlignet med normalfordelingen

Kurtosen er litt høyere enn det Lucia and Schwartz (2002) målte i perioden 1993 til 1999 på 3,5. Kurtosen på 3,5 tilsvarer en overskytende kurtose på 0,5.

En naturlig årsak til prissjokkene kan spores i markedskrysset (Figur 2.2). I perioder hvor forholdet mellom tilbud og etterspørsel er liten vil dette resultere i store utslag i systemprisen.

4.1.3 Korrelogram

Korrelogrammet (Figur 4.4) brukes til å oppdage ulike egenskaper ved tidsserien.



Figur 4.4 Korrelogram for systemprisen. Autokorrelasjonsfunksjonen til $P_{t,2}$ (svart) og $P_{t,3}$ (brun)

AKF går aldri ned til null, noe som indikerer at serien inneholder en trend. På grunn av dette bør trenden fjernes. På grunn av at denne trenden dominerer korrelogrammet fjernes trenden med lineær linje beregnet med MKM (3.7.3).

Selv om den lineære trenden fjernes, har AKF en høy verdi til å begynne med. Det vil si at sist observerte pris er en god indikator på prisen som kommer i neste periode.

Tidligere studier viser at systemprisen har sesongvariasjoner (kapitel 2.4.1). Visuell inspeksjon av korrelogrammet i Figur 4.4 viser stigende verdier ved 104 ukers lag. Etter en sesong (52 uker) er det ikke mulig å antyde noen sesongvariasjon. Dette tyder på at det finnes sesongvariasjon, men at denne er svak.

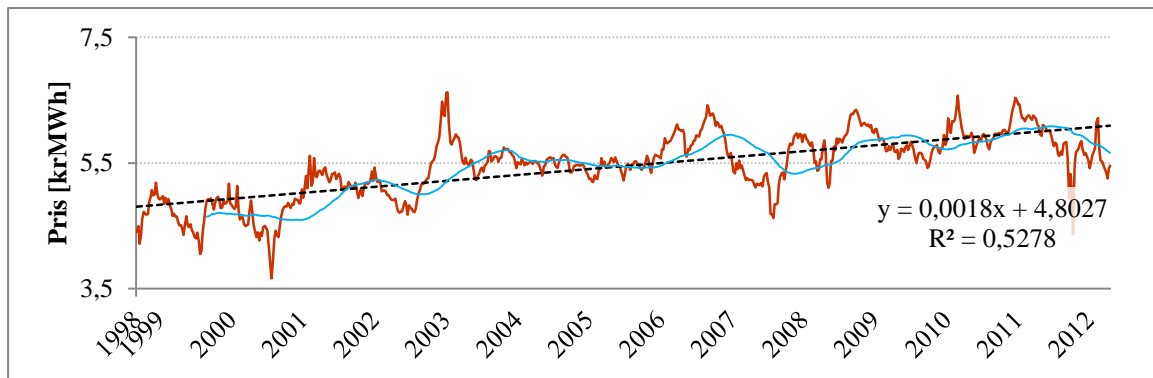
4.1.4 Logaritmen av systemprisen

Plottet av den logaritmiske systemprisen (Figur 4.5) viser at variasjonen i prisen blir mindre etter den logaritmiske transformasjonen. Dette skyldes at variasjonen i spotprisen er høyere når absolutt pris er høy.

4.1.5 Trend systempris

Etter at systemprisen er blitt logaritmisk transformert er det lettere å se at den inneholder en lineær trend. Denne vises som sort stiplet linje i Figur 4.5. Trendens R^2 forklarer over 52 % av kraftprisen, noe som vil i større grad påvirke hvor mye modellen som utvikles forklarer.

Det velges å beregne trenden etter lineær metode. En polynomtrendlinje av andre grad forklarer 59,7 %, men det er lite sannsynlig at prisøkningen flater ut som denne trenden foreslår. Grafisk inspeksjon viser at glidende gjennomsnitt over 52 uker ser ut til å ligge i nærheten av trendlinja mesteparten av tiden.



Figur 4.5 Plott av logaritmisk systempris. Trenden er illustrert med stiplet svart linje og trendens funksjon og forklaringsgrad er satt inn. Blå linje er bevegelig gjennomsnitt beregnet av de 52 foregående verdiene.

Mye av årsaken til denne trenden kan ligge i at forbruket øker raskere enn det den midlere produksjonen, og resultert i at det nesten har blitt balanse mellom forbruk og middelproduksjon. Dette påvirker det langsiktige forholdet mellom tilbud – etterspørsel, og ved å se på en mer momentan korrelasjon mellom pris og tilsig vil det være nødvendig å korrigere for denne trenden. Med et redusert kraftoverskudd kan det også tenkes at strømprisens volatilitet øker, noe som drøftes i kapittel 4.1.6. Vi ønsker derfor å fjerne denne trenden i videre analyse.

4.1.6 Sesongjustering av pris

Variasjoner som skyldes årstider elimineres bort ved at pris og tilsig sammenlignes i samme uke. I tilfeller hvor uke 53 forekommer neglisjeres denne. Dette resulterer i tilsammen 52 korrelasjonsfaktorer, og det viser seg at korrelasjonsfaktoren avhenger av årstid (kapitel 4.2).

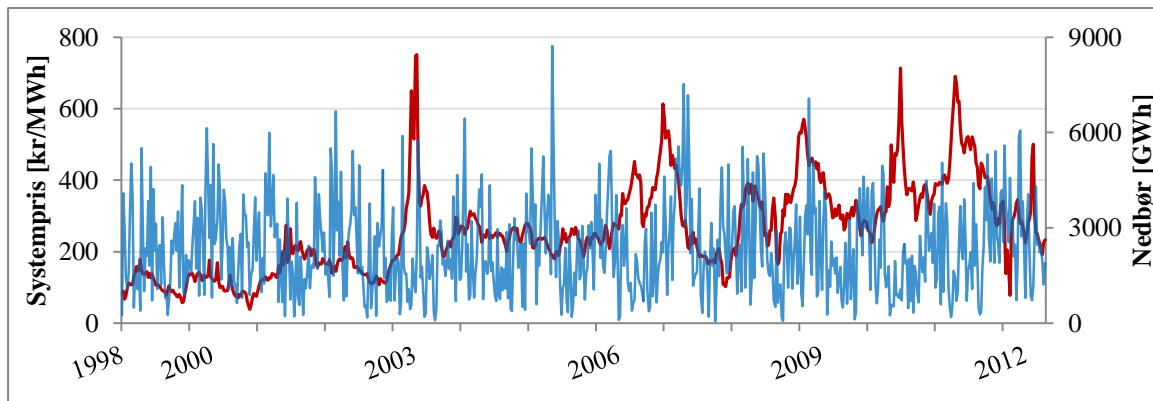
4.2 Nedbør

Å finne korrelasjon mellom spotpris og nedbørsdata er spesielt interessant fordi dette er den variabelen som på et tidligst mulig tidspunkt kan si noe om de kommende vannreservene. Dette er interessant i forbindelse med prisprognoser. Nedbørsdata er mengde vann som treffer jordens overflate, dette vil si både regn, snø, sludd mm. Rådataene er oppgitt pr. dag og har blitt lagt sammen til ukentlige verdier.

Ikke all form for nedbør er like lett å fange opp og måle. Dette gjelder spesielt tilfeller med snø og vind. Derfor er det knyttet en viss usikkerhet rundt nedbørsmålingen. Nedbør som forklarende variabel tar heller ikke hensyn til at en del kommer som snø, noe som gjør at prisen påvirkes et annet tidspunkt.

4.2.1 Beskrivelse av data

Figur 4.6 viser den observerte nedbøren i analyseperioden. Plottet viser at det er svært stor endring i nedbørsmengden i fra uke til uke. Disse variasjonene gjør det vanskelig å observere noen samvariasjon med systemprisen ut fra figuren. Variasjonene og gjennomsnittsverdien er omtrent lik gjennom hele analyseperioden. Det er også tegn til sesongvariasjoner.



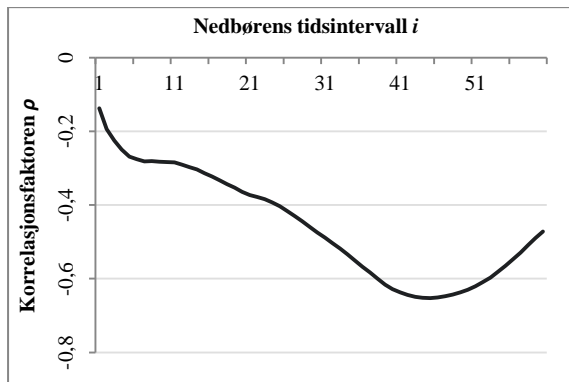
Figur 4.6 Observert systempris P_t (rød linje) og nedbør N_t (blå linje) i løpet av analyseperioden

4.2.2 Filtrering av tidsserien

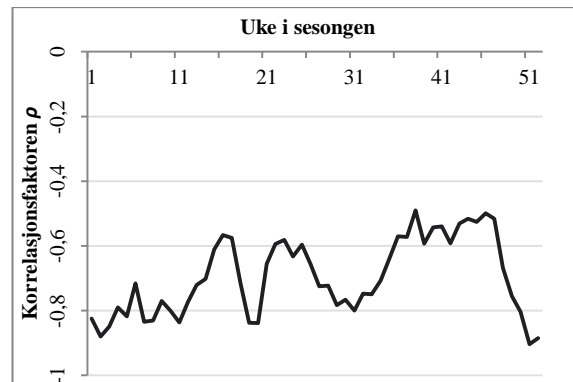
Det forventes at nedbøren har liten umiddelbar påvirkning på prisen, så det er i nedbørsdata som går over flere uker vi brukers i analysen. Derfor undersøkes korrelasjonen mellom systempris og ulike tidsintervaller nedbøren er målt på som forklarende variabel. Metoden forklart i kapittel 3.6.2.

Det gjøres en systematisk undersøkelse med i intervall fra 1 til 60 (Figur 4.7). Korrelasjonen viser seg å være sterkest om nedbøren som er målt de siste 45 ukene brukes til å forklare spotprisen, hvor $\rho = -0,652$. Ved et såpass langt tidsintervall vil nedbøren ved dette tidsintervallet trolig reflekterer tilgjengelige ressurser, og sier lite om den umiddelbare effekten nedbøren har på systemprisen.

På grunn av lengden på intervallet i gjøres analysen i fra uke 1 år 2000 til uke 16 i 2012.



Figur 4.7 Korrelasjonen ρ mellom systemprisen $P_{t,3}$ og nedbøren $N_{t,2}$ med ulik lengde på intervallet $0 < i < 60$



Figur 4.8 Korrelasjonen ρ mellom systemprisen $P_{t,3}$ og nedbøren $N_{t,2}$ i gjennom sesongen når $i = 45$

Korrelasjonskoeffisienten ρ viser seg å variere med årstiden (Figur 4.8). Dette observeres ved å utføre 52 regresjoner med data fra samme uke⁵. ρ varierer fra -0,490 i uke 38 til -0,904 i uke 51. Dette indikerer at korrelasjonen er sterkest på vinteren, og blir svakere utover sommeren / seinhøstes.

⁵ Dvs. alle data grupperes inn i ukesindeling slik at regresjonen utføres kun av verdier fra f.eks. uke 50 i perioden 2000-2012

En del av variasjonen kommer av tilfeldigheter som kommer av at det gjøres 52 ulike korrelasjonsberegninger. En annen grunn til disse variasjonene kommer av at den lengden på tidsintervallet til nedbøren i som gir sterkest korrelasjon ser ut til å variere med årstiden.

4.2.3 Krysskorrelasjon

I dette forsøket undersøkes korrelasjonen mellom spotpris, der spotprisen skal forklares ut fra nedbørsdata som er observert τ uker før. Dette er spesielt interessant i forbindelse med prisprognoser, hvor vi ønsker å vite spotprisen så langt fram i tid som mulig ut fra nedbørsdata som er gjort fram til i dag.

Tidsforsinkelsen er lagt inn som variabelen τ . På grunn at tidsforsinkelsen settes tidsintervallet i til 45, noe som er under den optimale verdien. Også i dette tilfellet blir p-verdien såpass høy ved polynomregresjon at en enkel lineær modell benyttes.

Resultatene viser at korrelasjonsfaktoren ρ svekkes lineært med 0,1449 for hver uke tidsforsinkelsen øker.

4.2.4 Endring i korrelasjonen over tid

Korrelasjonen undersøkes år for år, og resultatene viser at korrelasjonene varierer svært mye år til år. Tidsintervallet til nedbørsdataene er satt til 46 uker.

År	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
ρ	-46	40	-89	-2	-45	-29	-62	-86	-83	-37	17	-68

Tabell 4.2 Korrelasjonskoeffisienten i perioden 2000 - 2011

År 2001 og 2010 får positive korrelasjonskoeffisienter, noe som tyder på spuriøse resultater da dette betyr at spotprisen synker jo mindre nedbør som kommer. I år 2003 er det ingen korrelasjon, mens i 2002, 2007 og 2009 er det meget god korrelasjon. Årsaken til denne variasjonen er det ikke funnet noen årsak til. Finnes det en kausal relasjon til resultatene i årene 2001, 2010 og eventuelt 2003 burde disse blitt behandlet som outliers, og korrelasjonen over hele perioden hadde blitt sterkere. På grunn av stor variasjon i resultatene er det vanskelig å se om det er noen trend.

4.2.5 Korrelasjonsmatrise

Tabell 4.3 oppsummerer korrelasjonen mellom systempris og nedbør regnet på ulike måter. I dette tilfellet økte korrelasjonsfaktoren ρ ved å fjerne den lineære trenden fra systemprisen og øke tidsintervallet til energitilsiget.

	$N_{t,2}$
$P_{t,2}$	-0,532
$P_{t,3}$	-0,652

Tabell 4.3 Korrelasjonsfaktoren ρ mellom systempris og nedbør

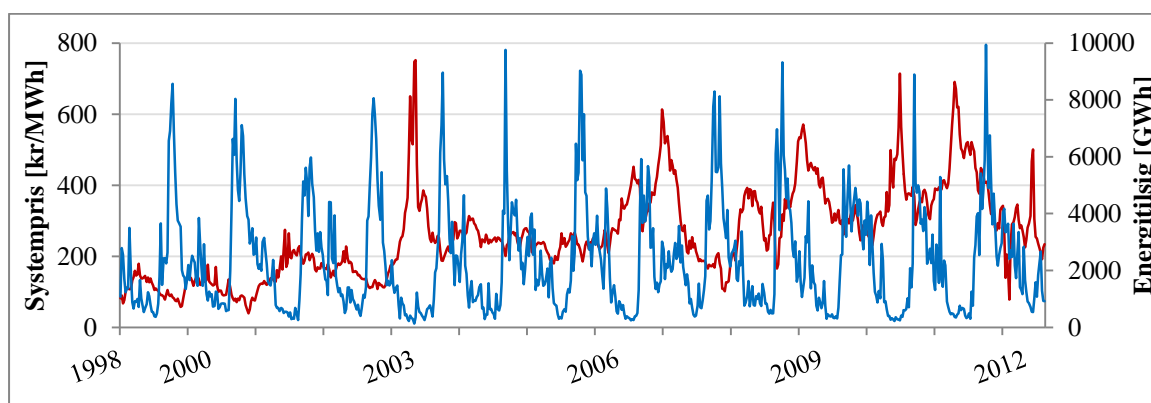
4.3 Energitilsig

Energitilsiget beskriver hvor stort tilsig magasinene har hatt, og det beregnes ut fra endring i magasinbeholdning pluss produsert effekt i den aktuelle tidsperiode.

4.3.1 Beskrivelse av data

Visuell inspeksjon av Figur 4.9 viser at tilsiget har tydelige sesongvariasjoner. De høyeste verdiene kommer i vårfloppen rundt uke 20, mens de laveste verdiene observeres om i perioden før vårfloppen kommer.

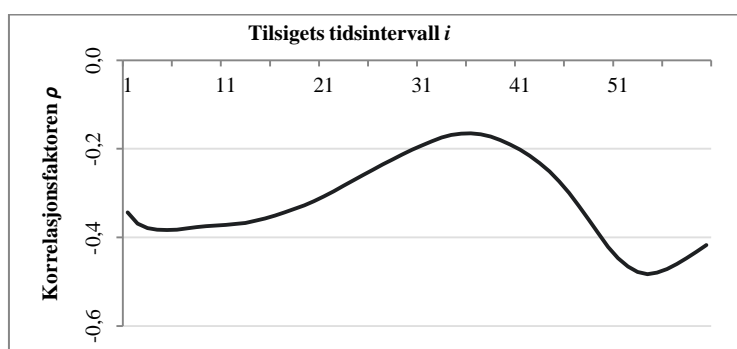
Plottet viser at pristoppene har en klar tendens til komme i perioder med lite tilsig, og at prisen ofte får et bunnpunkt ved høyt tilsig.



Figur 4.9 Observert energitilsig T_t (blå linje) og systemprisen P_t (rød linje) i løpet av analyseperioden

4.3.2 Filtrering av tidsserien

I likhet med nedbøren antas det at energitilsig i lengre perioder påvirker prisen mer enn den umiddelbare virkningen. Dette skyldes muligheter for å lagre vann i magasin, men på grunn av at det er en del kraftverk som ikke har denne muligheten blir kraftprisen påvirket vil prisen til en viss grad påvirkes noe umiddelbart.



Figur 4.10 Korrelasjonen ρ mellom systemprisen $P_{t,3}$ og nedbøren $T_{t,2}$ med ulik lengde på intervallet $0 < i < 60$

Figur 4.10 viser at korrelasjonskoeffisientens funksjon av tilsiget $\rho(i)$ har to bunnpunkter. Det ene er når $i = 5$, det andre når $i = 54$. Dette kan bety at når $i = 5$ reflekterer prisfallet som kommer av at elvekraftverk øker sin produksjon, mens når $i = 54$ reflekterer dette kraftressursene (magasin og

snø) som er tilgjengelig i Norden. På grunn av at korrelasjonen er sterkere mellom systempris og tilgjengelige ressurser (kapittel 4.4.4) er tilfellet hvor $i = 5$ det mest interessante tilfellet når systemprisen skal forklares av flere variabler (kapitel 4.6).

4.3.3 Korrelasjonsmatrise

Tabell 4.4 oppsummerer korrelasjonen mellom systempris og energitilsiget regnet på ulike måter. I dette tilfellet økte korrelasjonen ved å fjerne den lineære trenden fra systemprisen og øke tidsintervallet til energitilsiget.

	T_t	$T_{t,2},$ $i = 5$
$P_{t,2}$	-0,2261	-0,2835
$P_{t,3}$	-0,3483	-0,3834

Tabell 4.4 Korrelasjonsfaktoren ρ mellom systempris og energitilsiget

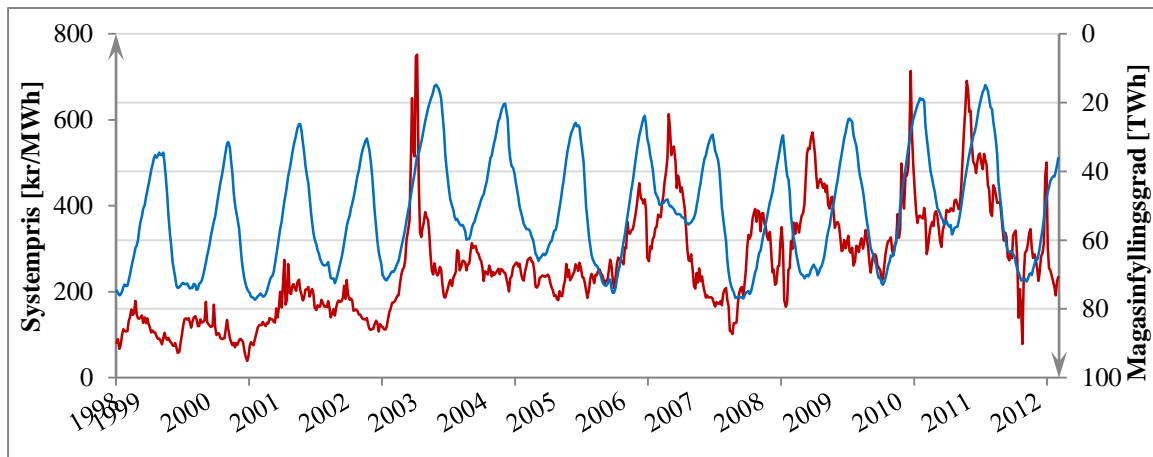
4.4 Magasinfyllingsgrad

Magasinbeholdningen og snømengden som finnes i nedbørsområdene er de beste målene på energireservene som er disponible. Derfor er det naturlig at denne påvirker spotprisen på forskjellige måter. Mengden vann produsentene har i magasinene kombinert med hydrologisk tilsig påvirker kraftprodusentenes convenience yield som er omtalt i kapitel 2.4.2.

Forholdet mellom systempris og nedbør, tilsig og snø er kausalt, det vil si at de kan påvirkesystemprisen, men ikke motsatt. Når det gjelder magasinbeholdning kan det likeså gjerne være prisen som avgjør magasinbeholdningen, som at magasinbeholdningen bestemmer prisen. Resultatet av analysen må sees i betraktning av dette. På grunn av dette vil trolig magasinbeholdningen være mindre egnet til å brukes i prognoser enn de andre variablene.

4.4.1 Beskrivelse av data

I Figur 4.13 er systemprisen plottet sammen med magasinfyllingen i synkende rekkefølge. Visuell inspeksjon viser tydelige sesongsvingninger, men at verdien til ekstremalpunktene til disse varierer fra år til år. Flere av pristoppene kommer noen uker før fyllingsgraden når sesongens laveste verdi. Prisbunnene kommer som regel litt før eller i samme tid som når magasinet har sesongens høyeste verdi. Det finnes flere unntak, dette er i spesielt i perioden 2004 - 2006.



Figur 4.11 Plot av magasinbeholdning (motsatt retning) og logaritmisk systempris uten lineær trend

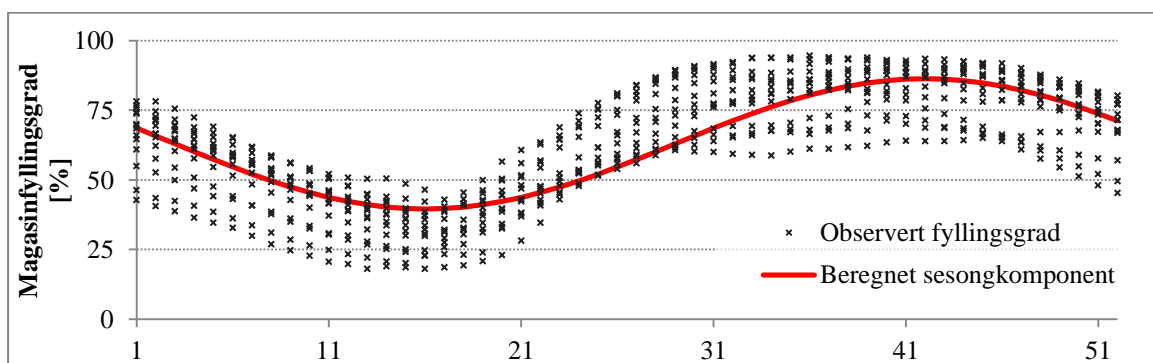
4.4.2 Sesongjustering

Magasinene ser ut til å ha tydelige sesongvariasjoner i seg. Dette vises både i plottet av magasinfyllingsgraden og autokorrelasjonsfunksjonen⁶. Det antas kraftprodusentene legger magasinivået sitt etter et normalår. Derfor fjernes denne sesongen i videre analyser.

Magasinfyllingsgraden er plottet inn i Figur 4.12 for å vise hvordan den varierer gjennom sesongen. Den ser ut til å ligne en sinuskurve, dette gjør det naturlig å forklare variasjonene med periodisk cosinusfunksjon. Normalåret (rød linje i Figur 4.12) blir regnet ut ved å tilpasse en cosinusfunksjon til magasindataene og tilpasse koeffisientene. Avviket får vi ved å trekke normalåret fra observert magasinbeholdning:

$$M_t = M_{t,2} + \alpha \cos\left(\frac{2\pi * uke_t}{52} + \phi\right) + \beta \quad \text{Formel 4.1}$$

Med MKM beregning er koeffisientene i normalåret $\hat{\alpha}$ -23,38, $\hat{\beta}$ er 62,96. Det er lagt inn en tidsforsinkelse ϕ på -16 som referer til starten på det hydrologiske året (uke 52 - 16 = 36). Dette er også verdien som gir best forklaring av magasinbeholdningen. Dette betyr at middelverdien er 62,96 %, og at den vanligvis svinger med $\pm 23,38$ %.



Figur 4.12 Magasinfyllingsgrad gjennom sesongen (kryss) og beregnet normalår (rød linje)

⁶ Appendiks

4.4.3 Krysskorrelasjon

For å gå nærmere inn på om det er magasinbeholdningen som påvirker prisen eller om det er prisen som påvirker magasinbeholdningen undersøkes krysskorrelasjonen mellom systempris og magasinbeholdning, som i regresjonen over.

Tidsforskyving τ	-2	-1	0	1	2	3	4
Korrelasjon ρ	-0,598	-0,631	-0,653	-0,661	-0,661	-0,657	-0,648

Tabell 4.5 Korrelasjonskoeffisienten ρ ved ulike tidsforskyvninger τ .

Resultatene fra Tabell 4.5 viser at korrelasjonen er sterkere i om ϕ er 1, 2 eller 3 enn om ϕ er 0. Dette indikerer at magasinfyllingsgraden er i seg selv et resultat av systemprisen. I prisprognoser vil dette være et problem, da den forklarende variabelen er påvirket av variabelen den skal forklare. Likevel kan det tenkes at magasinfyllingsgraden er med å påvirke hvor mye prisen beveger seg når noen av de andre hydrologiske variablene endrer seg. Dette undersøkes i kapittel 4.6.1.

4.4.4 Korrelasjonsmatrise

Resultatet av korrelasjonskoeffisienten ρ ved ulike forsøk er oppsummert i Tabell 4.6. Matrisen viser at korrelasjonen øker kraftig når magasinfyllingsgraden sesongjusteres. Det er derimot litt mer usikkert om det er riktig å trekke fra den lineære trenden fra kraftprisen.

	M_t	$M_{t,2}$
$P_{t,2}$	-0,3057	-0,6269
$P_{t,3}$	-0,2794	-0,6520

Tabell 4.6 Korrelasjonskoeffisienten ρ beregnet mellom systemprisen og magasinfyllingsgraden.

4.5 Tilgjengelige ressurser

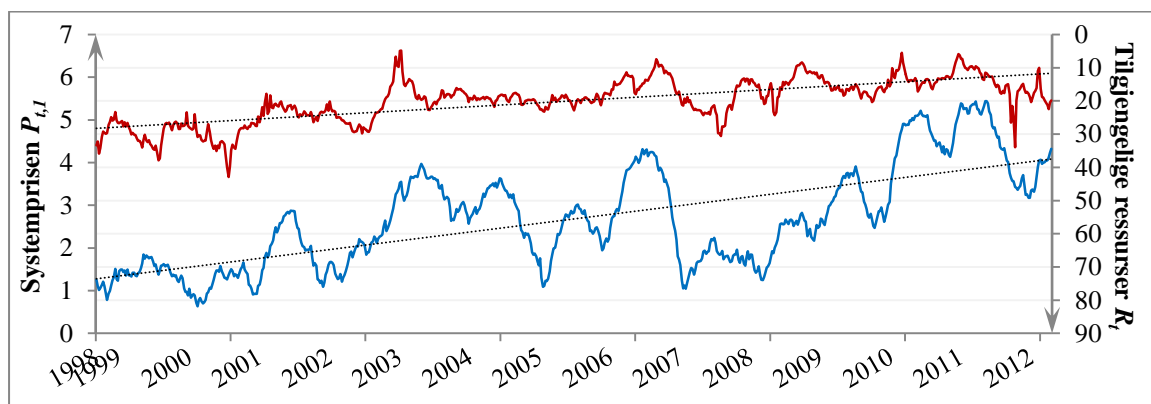
Tilgjengelige ressurser tar også hensyn til snømengden, som på et lengre perspektiv er med på å bestemme magasinfyllingsgraden.

Det er en del usikkerhet rundt hvordan snømengden påvirker systemprisen. På grunn av at snøen ofte ligger lenge før den smelter og produserer nyttbar energi vil denne kunne være med på å si noe om prisutviklingen på lang sikt. Er fyllingsgraden svært lav på starten av vinteren kan det tenkes at prisen blir høyere enn om fyllingsgraden er lav rett før snøsmeltingen.

4.5.1 Beskrivelse av data

I Figur 4.13 er systemprisen plottet sammen med tilgjengelige ressurser i synkende rekkefølge. Visuell inspeksjon viser sesongsvingninger, men at sesongens ekstremalpunkter varierer fra år til år. Ekstremalpunktene til systemprisen kommer som regel før ekstremalpunktene til tilgjengelige ressurser.

Plottet viser også at tilgjengelige ressurser har blitt kraftig redusert i løpet av analyseperioden. For å synliggjøre dette er den lineære trenden til pris og ressurser er lagt til i plottet. Dette viser også at reduksjonen i tilgjengelige ressurser har vært større enn det systemprisen har steget.



Figur 4.13 Plot av tilgjengelige ressurser R_t (blå, synkende rekkefølge) og systemprisen $P_{t,1}$ (rød). Lineær trend til tidsseriene er plottet inn med stiplede linje.

4.5.2 Korrelasjonsmatrise

Tabell 4.7 viser at det er korrelasjonen er sterkest mellom logaritmisk systempris og tilgjengelige ressurser. Ved å trekke fra den lineære trenden i systemprisen reduseres korrelasjonen kraftig. Dette indikerer at trenden i systemprisen skyldes at de gjennomsnittlige tilgjengelige kraftressursene er blitt redusert de siste årene.

Sammenlignes resultatene i Tabell 4.7 med Tabell 4.6 ser man at korrelasjonen til den opprinnelige kraftprisen P_t er mye kraftigere når snømengden medregnes.

	R_t
P_t	-0,7138
$P_{t,2}$	-0,7206
$P_{t,3}$	-0,3243

Tabell 4.7 Korrelasjon mellom systempris og tilgjengelige ressurser

4.6 Multipel regresjon

Korrelasjonsanalysen viser at energitilsiget de fem siste ukene forklarer best de umiddelbare endringene i kraftprisen, mens tilgjengelige ressurser beskriver systemprisens gjennomsnittlige prisnivå. I dette kapitlet forsøkes det å settes opp en systemprisfunksjon som forklares ut fra de hydrologiske variablene.

4.6.1 Innføring av nye ledd

Det er blitt forsøkt polynomregresjon med tilsig eller nedbør som forklarende variabel. Resultatet viser at β_2 verdien stort sett er negativ hele året. Dette sier oss at spotprisen synker raskere desto mer nedbør som kommer. Ser vi på det korte løp (med $a < 30$) blir γ positiv. Dette indikerer at relativt store nedbørsmengder på kort tid ikke reduserer spotprisen like mye relativt store nedbørsmengder over lengre tid. En naturlig forklaring på dette er at kraftstasjonene ikke er

dimensjonert for å ta kortvarige flomtopper. Likevel er økningen i forklaringsgraden såpass liten at det velges å bruke en vanlig lineær regresjon videre i oppgaven.

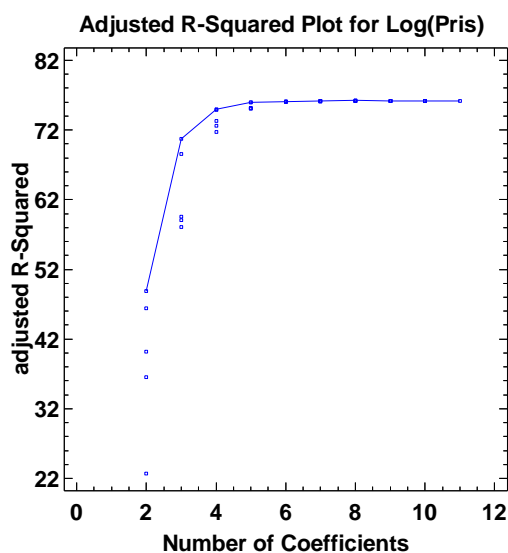
Polynomregresjon er også med sesongjustering, hvor regresjonen forklarer hver enkelt uke. Dette førte til høye p-verdier som var utenfor signifikansnivået.

Forsøk med kryssproduktledd gav ikke økt forklaringsgrad.

4.6.2 Valg av regresjonsmodell

For å finne den modellen som gir høyest R^2 -verdi utføres ulike multiple regresjon med følgende forklarende variabler til å forklare systemprisen $P_{t,2}$, $N_{t,1}$, $N_{t,2}$, $T_{t,1}$, $T_{t,2}$, $M_{t,1}$, $M_{t,2}$, $R_{t,1}$, $R_{t,2}$, uke og år⁷. Dette gir 1023 mulige modeller.

Vanligvis vil R^2 som funksjon av antall uavhengige variabler i modellen til å begynne med øke, for å så flate ut. I det området hvor funksjonen begynner å flate ut eller har en knekk i forløpet av indikerer at de viktigste variablene er tatt med. De resterende variablene vil være lite betydningsfulle eller bare ha en tilfeldig innflytelse på modellen. R^2 -plottet (Figur 4.14) viser at kurven flater ut etter tre forklaringsvariabler. Dette indikerer at det ikke er nødvendig å ha med flere enn tre forklaringsvariabler, derfor reduseres disse til tre. Dette gir 175 mulige modeller. I Tabell 4.8 er 15 av modellene som gir høyest forklaringsgrad listet opp.



Figur 4.14 R^2 plott

	R^2	R^2 -justert	Inkluderte variabler
1	75,11	75,00	$N_{t,2}$, $T_{t,2}$, år_t
2	74,94 ⁸	74,83	$N_{t,2}$, $M_{t,2}$, år_t
3	73,42	73,30	$N_{t,2}$, $T_{t,1}$, år_t
4	72,77	72,65	$M_{t,2}$, $R_{t,2}$, år_t
5	71,79 ⁹	71,67	$M_{t,1}$, $M_{t,2}$, år_t
6	70,86	70,78	$M_{t,2}$, år_t
7	68,61	68,52	$N_{t,2}$, år_t
8	59,64	59,51	$M_{t,1}$, $R_{t,2}$
9	59,16	59,04	$T_{t,2}$, $R_{t,1}$
10	58,16	58,03	$T_{t,1}$, $R_{t,1}$
11	48,96	48,88	$R_{t,1}$
12	46,50	46,42	år_t
13	40,31	40,22	$M_{t,2}$
14	36,66	36,57	$R_{t,2}$
15	22,83	22,71	$N_{t,2}$

Tabell 4.8 Modeller med høyest justert R^2 verdi.

Modellen med høyest forklaringsgrad bruker $N_{t,2}$, $T_{t,2}$ og år_t som forklaringsvariabler, og gir en forklaringsgrad på 75 %. Svakheten til denne modellen er at hverken $N_{t,2}$ eller $T_{t,2}$ forklarer den lineære utviklingen i systemprisen noe særlig bra. Variabelen år_t kan være forårsaket en rekke eksogene variabler

⁷ Verdien til årstallet brukes, dvs. observasjoner gjort i 2005 får verdien 2005.

⁸ Inneholder korrelasjon mellom de forklarende variablene ($\rho > 0,5$)

⁹ Inneholder korrelasjon mellom de forklarende variablene ($\rho > 0,5$)

som ikke skyldes hydrologiske forhold. I kapittel 4.5 er det diskutert om noe av trenden skyldes reduserte tilgjengelige ressurser. Variabelen $\hat{a}r_t$ er med på åtte av modellene, noe som indikerer at trenden i systemprisen ikke forklares særlig godt ut fra de hydrologiske variablene. De mest forklarende modellene som bare bruker hydrologiske variabler er $M_{t,1}$ og $R_{t,2}$ eller $T_{t,2}$ og $R_{t,1}$ som forklarer litt over 59 % av den logaritmiske systemprisen.

4.6.3 Korrelasjon mellom de forklarende variablene

For å unngå multikollinearitet i regresjonen undersøkes korrelasjonsfaktoren ρ i mellom de forklarende variablene. Det bør unngås å ha flere variabler som korrelerer med hverandre i regresjonen. Det er vanskelig å sette noen grense for hvor korrelasjonsfaktoren kan være, så verdier over 0,5 bør vurderes sammen med resten av regresjonen. Resultatene vises i Tabell 4.9.

	$N_{t,1}$	$N_{t,2}$	$T_{t,1}$	$T_{t,2}$	$M_{t,1}$	$M_{t,2}$	$R_{t,1}$	$R_{t,2}$
$N_{t,1}$		0,07	0,12	0,11	0,33	0,14	0,20	-0,07
$N_{t,2}$	0,07		0,05	0,05	0,19	0,62	0,56	0,61
$T_{t,1}$	0,12	0,05		0,85	0,13	0,40	-0,08	0,03
$T_{t,2}$	0,11	0,05	0,85		0,36	0,57	-0,03	-0,06
$M_{t,1}$	0,33	0,19	0,13	0,36		0,57	0,46	-0,23
$M_{t,2}$	0,14	0,62	0,40	0,57	0,57		0,55	0,44
$R_{t,1}$	0,20	0,56	-0,08	-0,03	0,46	0,55		0,69
$R_{t,2}$	-0,07	0,61	0,03	-0,06	-0,23	0,44	0,69	

Tabell 4.9 Korrelasjonsfaktoren ρ mellom de forklarende variablene. Absoluttverdi av korrelasjonsfaktoren over 0,5 vises i fet skrift

Resultatene viser blant annet at $N_{t,2}$ og $M_{t,2}$ (nummer to i Tabell 4.8) og $M_{t,1}$ og $M_{t,2}$ (nummer fem i Tabell 4.8) har en korrelasjonsfaktor på over 0,5.

Korrelasjonen mellom $N_{t,2}$ og $M_{t,2}$, $R_{t,1}$ og $R_{t,2}$ viser at $N_{t,2}$ trolig er en god indikator på tilgjengelige ressurser. Som tidligere antydnet kan magasinfyllingsgraden være påvirket av prisen, dette kan være årsaken til at $N_{t,2}$ gir bedre forklaringsgrad enn magasin eller tilgjengelige ressurser.

4.6.4 Logaritmisk systempris forklart med nedbør, tilsig og år

Resultatene fra regresjonen viser at $N_{t,2}$, $T_{t,2}$ og $\hat{a}r_t$ er signifikante med en p-verdier lik 0,000. Utrykket beskrives med:

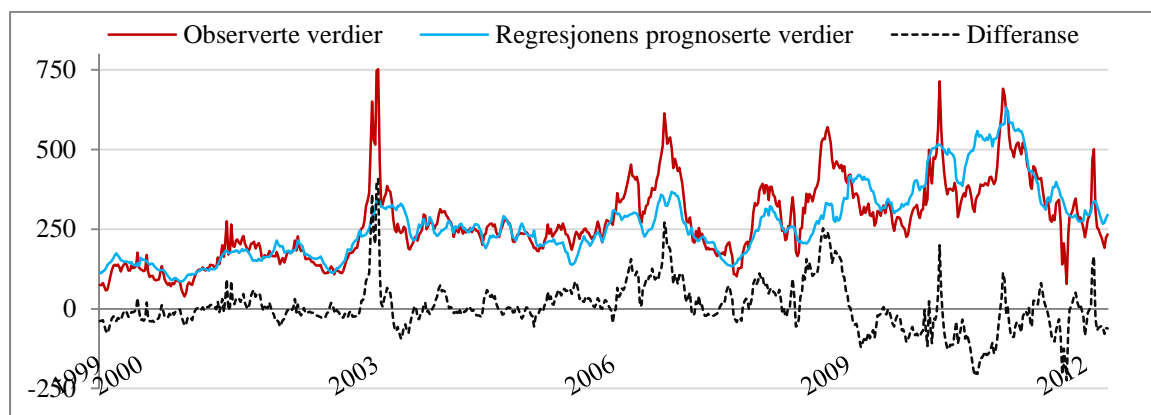
$$P_{t,2} = \ln P_t = \alpha + \beta_1 N_{t,2} + \beta_2 T_{t,2} + \beta_3 \hat{a}r_t + \varepsilon_t \quad \text{Formel 4.2}$$

Regresjonens $\hat{\alpha}$ - og $\hat{\beta}$ -verdier vises i Tabell 4.10.

$\hat{\alpha}$	-168,895
$\hat{\beta}_1$	- 0,0000162076
$\hat{\beta}_2$	- 0,0000148211
$\hat{\beta}_3$	0,0878941

Tabell 4.10 $\hat{\alpha}$ - og $\hat{\beta}$ -verdier ved regresjon med $N_{t,2}$, $T_{t,2}$ og $\hat{a}r_t$ som forklarende variabler

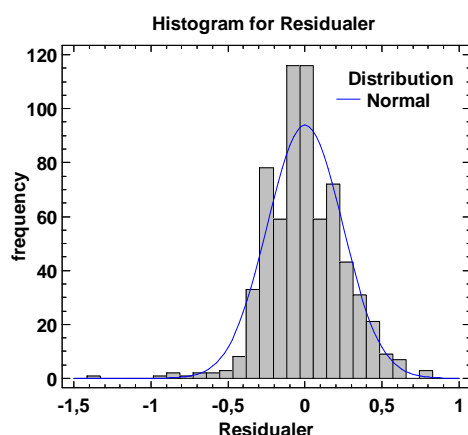
For å gjøre en visuell inspeksjon av resultatene beregnes prognosert systempris ved e opphøyd i regresjonens resultater. Dette sammenlignes med observerte systemprisen:



Figur 4.15 Systemprisen med regresjonens prognoserte verdier og differansen mellom disse

Figur 4.15 viser at differansen mellom observerte verdier og regresjonens er størst når systemprisen har topper. Dessuten endres variansen til differansen med tiden, noe som gjør seg tydelig etter år 2006. Ved dette tidspunktet kan man også observere endringer i den observerte systemprisen. Det kan også antydes at forventet verdi har sunket i tiden etter 2009.

Residualene, som er logaritmen av differansen i Figur 4.15, inneholder autokorrelasjon langs tidsaksen. Durbin–Watson statistikken gir en faktor på 0,23, noe som indikerer at det finnes korrelasjon målt mot forrige residual. Det er sannsynlig at dette skyldes utlatte variabler, siden systemprisen også påvirkes av andre faktorer enn det som er med i regresjonen.



Figur 4.16 Histogram av residualene

Shapiro-Wilk W testen gir p-verdi lik 0,0606, noe som indikerer at residualene ikke kommer fra en normalfordeling. Histogrammet i Figur 4.16 viser at fordelingen er meget spiss ($K' = 1,87$) og litt skjev i negativ retning ($S = -0,132$).

4.6.5 Logaritmisk systempris forklart med tilsig og tilgjengelige ressurser

$M_{t,1}$ og $R_{t,2}$ forklarer systemprisen $P_{t,2}$ best uten år som forklarende variabel. Selv om denne har en forklaringsgrad 0,45 % høyere enn $T_{t,2}$ og $R_{t,1}$ velges sistnevnte på grunn av mindre korrelasjon mellom de forklarende variablene, og at korrelasjonsanalysen antyder relasjonen som mer kausal. Denne gir en R^2 -verdi lik 59,04 %.

Resultatene fra regresjonen viser at $T_{t,2}$ og $R_{t,1}$ er signifikante med en p-verdi lik 0,000. Utrykket beskrives med:

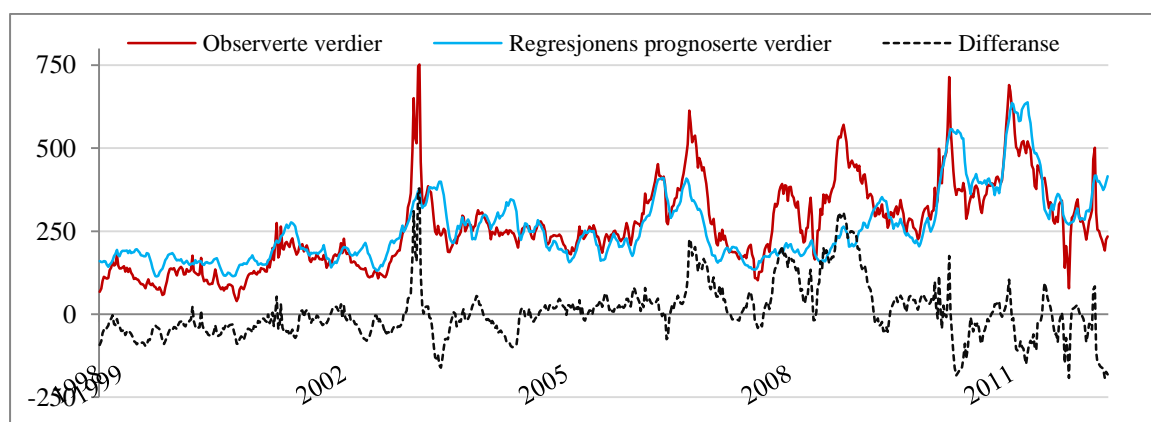
$$P_{t,2} = \ln P_t = \alpha + \beta_1 T_{t,2} + \beta_2 r_{t,1} + \varepsilon_t \quad \text{Formel 4.3}$$

Regresjonens $\hat{\alpha}$ - og $\hat{\beta}$ -verdier vises i Tabell 4.11.

$\hat{\alpha}$	6,97467
$\hat{\beta}_1$	- 0,0000185486
$\hat{\beta}_2$	- 0,0000231001

Tabell 4.11 $\hat{\alpha}$ - og $\hat{\beta}$ -verdier ved regresjon med $N_{t,2}$, $T_{t,2}$ og r_t som forklarende variabler

For å gjøre en visuell inspeksjon av resultatene beregnes prognosert systempris ved e opphøyd i regresjonens resultater. Dette sammenlignes med observerte systemprisen:



Figur 4.17 Systemprisen med regresjonens prognoserte verdier og differansen mellom disse

Figur 4.17 viser at også denne modellen, i likhet med Figur 4.15, har svakheter med at den ikke prognoserer pristoppene godt nok. Dessuten har denne modellen større avvik i forventet verdi, hvor den i starten ligger under null og stiger til over null fram til 2009. Etter 2009 har forventet verdi sunket til under null.

Residualene, som er logaritmen av differansen i Figur 4.17, inneholder autokorrelasjon langs tidsaksen.

5 Konklusjon

Opgaven har undersøkt i hvor stor grad systemprisen påvirkes av nedbør, energitilsig, magasinutfyllingsgrad og tilgjengelige ressurser i perioden 1998 til 2012. Ut fra denne problemstillingen er det laget to forskerspørsmål. Det første forskerspørsmålet bruker korrelasjonsanalyse til å undersøke hvor sterkt systemprisen korrelerer med de ulike hydrologiske variablene. Det andre forskerspørsmålet besvares ved å utføre en multipl minsket kvadraters regresjon for å finne ut hvor stor grad systemprisen forklares av de hydrologiske faktorene.

Av de fire hydrologiske variablene er det tilgjengelige ressurser som samvarierer mest med logaritmisk systempris med en korrelasjonsfaktor lik $-0,72$. De andre variablene ser ut til å ha problemer med å forklare den langsiktige økningen systemprisen har hatt i tidsrommet til analysen. Ved å trekke fra en lineær trend fra systemprisen øker korrelasjonsfaktoren til de andre variablene. Nedbøren som er observert de 45 foregående ukene gir en korrelasjonsfaktor på $-0,65$ mot en logaritmisk trendjustert systempris. Energitilsiget som er observert over fem uker ser ut til å forklare best de kortsiktige svingningene, og har en korrelasjonsfaktor $-0,38$. Sesongjusteres magasinutfyllingsgraden får denne en korrelasjonsfaktor lik $-0,65$. Det er knyttet problemer med å bruke magasinutfyllingsgraden i prisprognoser fordi den ser ut til å være sterkt påvirket av systemprisen. Dette problemet vil også gjelde tilgjengelige ressurser.

Regresjonsanalysen viser at tilsiget de fem siste ukene og tilgjengelige ressurser forklarer logaritmisk systempris med $59,04\%$. Modellen med høyest forklaringsgrad bruker nedbøren observert 45 foregående uker, gjennomsnittlig energitilsig fem foregående uker og årstallet som forklarende variabler. Denne forklarer $75,00\%$ av logaritmisk systempris. Begge modellene bruker variabler som inneholder stor usikkerhet, men det er knyttet noe større usikkerhet til den første modellen da tilgjengelige ressurser ser ut til å være påvirket av systemprisen. Test av regresjonene viser sterk autokorrelasjon i residualene. Dette kan være forårsaket av utelatte variabler eller at regresjonen på andre måter ikke er utført riktig. Dette er med på å redusere validiteten til modellenes resultater.

Modellene fra denne studien kan også suppleres med andre forklarende variabler enn hydrologiske for å kunne lage en mer komplett prognosemodell. Ellers har denne studien konsentrert seg om å forklare absolutt kraftpris. En annen måte å undersøke problemstillingen på er å forklare endring i kraftpris. Dette kan gi mer presise prognoser i en kort tidshorisont. På denne måten vil kraftprisen også bli stasjonær, noe som sannsynligvis løser problemet med autokorrelasjon.

6 Litteraturreferanser

- Benth, F. E., Kiesel, R., & Nazarova, A. (2011). A critical empirical study of three electricity spot price models. *Energy Economics*(0). doi: 10.1016/j.eneco.2011.11.012
- Botterud, A., Bhattacharyya, A. K., & Ilic, M. (2002). Futures and spot prices – an analysis of the Scandinavian electricity market.
- Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: forecasting and control*. San Francisco: Holden-Day.
- Chatfield, C. (2004). *The analysis of time series: an introduction*. Boca Raton, Fla.: Chapman & Hall/CRC.
- Clelow, L., & Strickland, C. (2000). *Energy derivatives: pricing and risk management*: Lacima Publications.
- Cochrane, D., & Orcutt, G. H. (1987). Application of Least Squares Regression to Relationships Containing Auto-correlated Error Terms. In M. L. King & D. E. A. Giles (Eds.), *Specification analysis in the linear model Donald Cochrane* (pp. 307-334): International Library of Economics series
London and New York:
Routledge and Kegan Paul.
- ENTSO-E. (2008). Annual Statistics 2008 (xls).
- ENTSO-E. (2011). Maximum NTC.
- Eydeland, A., & Wolyniec, K. (2003). *Energy and power risk management : new developments in modeling, pricing, and hedging*. Hoboken, N.J: Wiley.
- Fiorenzani, S. (2006). *Quantitative methods for electricity trading and risk management: advanced mathematical and statistical methods for energy finance*: Palgrave Macmillan.
- Fleten, S.-E., Wallace, S. W., & Tomasgard, A. (2001). Produksjonsplanlegging og risikostyring i et deregulert kraftmarked med finansielle instrumenter. Retrieved from <http://www.magma.no/produksjonsplanlegging-og-risikostyring-i-et-deregulert-kraftmarked-med-finansielle-instrumenter>
- Førsund, F. R., Hoel, M., & Mattei, F. E. E. (2004). *Properties of a non-competitive electricity market dominated by hydroelectric power*: Fondazione ENI Enrico Mattei.
- Gjolberg, O., & Johnsen, T. (2001). Electricity Futures: Inventories and Price Relationships at Nord Pool.
- Huovila, S. (2003). *Short-term forecasting of power demand in the Nord Pool market*. Master, Lappeenranta University of Technology, <http://www.doria.fi>. Retrieved from <https://www.doria.fi/handle/10024/34520>
- Knittel, C. R., Roberts, M. R., Regulation, P. o. W. E., & Institute, U. o. C. E. (2001). *An empirical examination of deregulated electricity prices*: University of California Energy Institute.
- Lindebæk-Nilsen, B., & Strand, K. (2004). *Risikohåndtering i kraftmarkedet*. B. Lindebæk-Nilsen og K. Strand, Kristiansand.
- Lucia, J. J., & Schwartz, E. S. (2002). Electricity Prices and Power Derivatives: Evidence from the Nordic Power Exchange. *Review of Derivatives Research*, 5(1), 5-50. doi: 10.1023/a:1013846631785
- Lund, A.-C., & Ollmar, F. (2002). *Analysing flexible load contracts in the energy market* (Vol. 18/2002). Bergen: Norges handelshøyskole, Institutt for foretaksøkonomi.
- Nome, S. (2010). *Prisprognoser og produksjonsplanlegging for et vannkraftverk*. Master, NHH, Bora. Retrieved from http://brage.bibsys.no/nhh/bitstream/URN:NBN:no-bibsys_brage_23986/1/Nome%202010.pdf

- NVE. (2009). Energi til oppvarming. Retrieved from Energi til oppvarming website:
- Pilipović, D. (2007). *Energy Risk: Valuing and Managing Energy Derivatives*: McGraw-Hill.
- Refvik, E., & Djupvik, H. (2010). *Opsjoner i kraftmarkedet: en beskrivelse av det nordiske kraftmarkedet med fokus påopsjoner*. [Forfatterne], Bergen.
- Spot, N. P. The Nordic Electricity Exchange and The Nordic Model for a Liberalized Electricity Market. Retrieved from <http://www.nordpoolspot.com/Global/Download%20Center/Rules-and-regulations/The-Nordic-Electricity-Exchange-and-the-Nordic-model-for-a-liberalized-electricity-market.pdf>
- Statnett. (2009). Statnetts historie.
- Striking the Balance - Annual Report 2010. (2011). Retrieved from http://www.nordpoolspot.com/Global/Download%20Center/Annual-report/annual-report-Nord-Pool%20-Spot_2010.pdf
- Studenmund, A. H. (2011). *Using econometrics : a practical guide*. Boston: Pearson.
- Sæthre, T., & Kristoffersen, Ø. S. (2006). *Elektrisitetsprisen ved EEX: Empiriske egenskaper og mulige faktormodeller*. Retrieved from http://www.iot.ntnu.no/users/fleten/students/tidligere_veiledning/KristoffersenSatre_H06.pdf
- Tsay, R. S. (2005). *Analysis of financial time series*. Hoboken, N.J.: Wiley-Interscience.
- Vallevik, I. Ø., & Øyan, S. M. (2004). *Analyse av kraftpriser*. Master, NTNU.
- von der Fehr, N. H. M., Amundsen, E. S., & Bergman, L. (2005). The Nordic Market: Signs of Stress? : Faculty of Economics, University of Cambridge.
- Weron, R., & Misiorek, A. (2008). Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *International Journal of Forecasting*, 24(4), 744-763. doi: 10.1016/j.ijforecast.2008.08.004

Appendiks

Appendiks A.1 Skewness (skjevhet)

Skjevheten er et mål på asymmetri i tidsseriens sannsynlighetsfordeling. Er fordelingen symmetrisk blir skjevheten 0, noe en normalfordeling er et eksempel på. Verdier større eller mindre enn null viser at serien er asymmetrisk. Denne verdien kan være interessant fordi den viser hvilken side av snittet et fremtidig estimat vil ligge på. Skjevheten kan beregnes med

$$S = \frac{1}{N} \sum_{i=1}^N \left(\frac{y_i - \bar{y}}{\hat{\sigma}} \right)^3 \quad \text{Formel 6.1}$$

hvor N er antall observasjoner, y_i er observert verdi, \bar{y} er beregnet verdi og $\hat{\sigma}$ er basert på den forventningsskjeve estimatoren for varians.

Appendiks A.2 Kurtose

Kurtose er et mål på hvor «spiss» sannsynlighetsfordelingen til en tidsserie er. En fordeling med høy kurtose kjennetegnes med en spiss fordeling med fete haler, lav kurtose gir avrundet topp med tynne haler. Kurtose kan beregnes med

$$K = \frac{1}{N} \sum_{i=1}^N \left(\frac{y_i - \bar{y}}{\hat{\sigma}} \right)^4 \quad \text{Formel 6.2}$$

En normalfordeling har en kurtose på 3. Begrepet overskytende kurtose kommer fra differansen fra en normalfordeling:

$$K' = K - 3 \quad \text{Formel 6.3}$$