

STATISTISKE MODELLERINGER I DET EUROPEISKE GASSMARKEDET

Anette Ullestad Frontéri



Masteroppgave i statistikk
Finansteori og forsikringsmatematikk

Matematisk Institutt
Universitetet i Bergen

1. juni 2015

Sammendrag

I denne oppgaven studerer vi naturgassmarkedet i Europa. Først ser vi på hvilke forhold som påvirker gassprisen. Deretter studerer vi statistisk teori om tidsrekker, volatilitet og copula. Til slutt blir det gjort statistiske modelleringer på naturgasspriser i tre europeiske gassmarkeder. Naturgassprisene blir omformet til log-avkastninger. GARCH-modeller blir brukt for å modellere volatiliteten i log-avkastningene. Copula blir brukt til å modellere avhengighetsstrukturen, som er mellom de marginale fordelingene i de europeiske gassmarkedene. Til slutt blir det brukt kointegrasjon for å beskrive sammenhengen mellom de europeiske gassprisene.

Forord

Naturgass er en av verdens viktigste energikilder. Den står for cirka 20 prosent av verdens energiforbruk. Selv om naturgassen er viktig, er den ofte kommet i skyggen av "hovedpersonen" av de fossile brennstoffene, nemlig råoljen. Antall studier som er skrevet om naturgasspriser i Europa er relativt få, i forhold til antall studier om oljepriser. Det har motivert meg til å skrive denne oppgaven om naturgasspriser. Under er noen artikler og oppgaver som omhandler naturgass, som har inspirert meg i denne oppgaven:

- "Using copulas to model price dependence in energy markets" av Gregoire et al. (2008)
- "Price forecasts in the european natural gas markets" av Fabini (2012)
- "Modeling and Forecasting Demand for Natural Gas of Retail Consumers" av Obžigailov (2011).

Modeller for de europeiske gassmarkedene viktig. En korrekt spesifisert modell kan brukes til mye nyttig. For eksempel kan en modell brukes for å få gode prognoser. Gode prognoser er nyttig for blant annet investorer og grunneiere som vurderer å bore nye brønner for gass. Det er av betydning for dem å vite hvor mye de kan tjene på dette i fremtiden, hvis det er noe å tjene i det hele tatt. Det kan også være nyttig for bilindustrien å ha prognoser på gasspriser, de ønsker å vite om det er det lønnsomt for dem å bygge biler som går på gass eller ikke. I tillegg er det betydningsfullt når en skal bygge kraftverk, er det mest lønnsomt å bygge gass, kull eller atomkraftverk? Dette medfører ett behov for gode prognoser innenfor gasspriser.

Takk

Først og fremst vil jeg takke min hovedveileder Hans A. Karlsen for all den hjelpen jeg har fått i arbeidet med oppgaven. Du har vært en god veileder, og gitt meg mange nyttige forslag til nye tema, og forbedringer av oppgaven. I tillegg vil jeg takke min biveileder Bård Støve for all hjelpen og nyttige innspill. De har begge vært til god hjelp, selv om siste delen av masteroppgaven ble skrevet i Oslo. Noe som har gjort at veiledningen har vært fysisk begrenset. Jeg vil igjen takke Hans for alle epostene han tatt seg tid til skrive for å hjelpe meg med oppgaven.

Jeg vil også takke mine medstudenter for den hyggelige tiden vi fikk sammen i den første delen av masteren. En ekstra takk til min medstudent Elina og hennes familie, som har latt meg for overnatte hos dem når jeg har vært i Bergen. Videre vil jeg takke min familie som har bidratt med korrekturlesning. Til slutt vil jeg takke min mann og min datter for deres støtte og tålmodighet.

Innhold

Sammendrag	iii
Forord	v
Takk	vii
1 Introduksjon	3
1.1 Innledning	3
1.2 Naturgassmarkedet	4
1.2.1 Generell fakta om naturgass	4
1.2.2 Naturgassen sin rolle opp igjennom tidene	5
1.2.3 Energi måleenheter	8
1.2.4 Naturgassmarkedet i Norge	8
1.2.5 Hvordan naturgass prises i gassmarkedet	9
1.2.6 Kjøp og salg av naturgass	12
1.2.7 Tilbud, etterspørsel og priser	12
1.3 Tilbud og etterspørsel faktorer	13
1.3.1 Vær, temperatur og årstider	13
1.3.2 Økonomisk vekst og økonomiske svingninger	13
1.3.3 Konkurransen i de ulike naturgass sektorer	14
1.3.4 Miljø og politiske og regionale reguleringer	17
1.3.5 Teknologisk utvikling	18
1.3.6 Transport, distribusjon og reserver av naturgass	19
2 Tidsrekker	21
2.1 Stasjonære tidsrekker	21
2.2 Tilfeldig gang	22
2.3 Enhetsrot test	23
2.4 Transformere dataene	25
2.5 Autokorrelasjonsfunksjon	26
2.6 Lineære prosesser	27
2.6.1 MA(q)-modell	27
2.6.2 MA(1)-modell	27

2.6.3	AR(p)-modell	28
2.6.4	AR(1)-modell	28
2.6.5	ARMA(1, 1)-modell	29
2.6.6	ARMA(p, q)-modell	30
2.7	Kointegrasjon	30
2.7.1	Engle–Granger to-steps metode	30
2.7.2	Johansen-metoden	31
3	Volatilitets modeller	35
3.1	ARCH-modeller	36
3.1.1	ARCH(q)-modell	36
3.1.2	Ordenbestemmelse i ARCH-prosessen	36
3.1.3	ARCH(1)-modell	36
3.2	GARCH-modeller	40
3.2.1	GARCH(q, p)-modell	40
3.2.2	GARCH(1, 1)-modell	41
3.2.3	Predikere volatilitet	42
3.2.4	Andre utvidelser av GARCH	43
3.3	Lagrange Multiplier ARCH test	43
3.4	Ljung-Box test	44
4	Copula	45
4.1	Definisjon	45
4.2	Fordelinger	46
4.3	Eksempel fra Copulafamilien	47
4.3.1	Normal copulafamilie	47
4.3.2	Student t-copulafamilie	48
4.3.3	Clayton copulafamilie	49
4.3.4	Gumbel copulafamilie	50
4.4	Kendalls tau og Spearmans rho	50
4.5	Estimere parametrene	52
4.6	Modelltester	52
4.7	Simuleringer med copula	53
5	Statistisk analyse av Europeiske naturgasspriser	55
5.1	Datasett av Europeiske naturgasspriser	55
5.1.1	Grafisk analyse over naturgassprisene	57
5.1.2	Omforme gasspriser til log-avkastninger	58
5.2	Fakta om log-avkastningene	60
5.2.1	Tester om log-avkastningene er stasjonære	60
5.2.2	Egenskaper til log-avkastningene	60
5.2.3	Volatilitet i log-avkastningene	64
5.3	Bygge en volatilitetsmodell	65
5.3.1	Tilpasning av lineære modeller for log-avkastningene	65

5.3.2	Teste for ARCH effekter	70
5.3.3	Tilpasning av volatilitetsmodellene	70
5.3.4	Modelltesting og valg av modell	74
5.3.5	Modeller for de marginale tidsrekkene	81
5.3.6	Predikasjon av volatilitet	83
5.4	Avhengighet	84
5.4.1	Korrelasjon	84
5.4.2	Scatter plott	85
5.5	Copula på naturgass residualer	86
5.5.1	Copula estimering	89
5.5.2	Modelltesting: Valg av copula	90
5.5.3	Prognose ved bruk av copula	93
5.6	Kointegrasjon	95
5.6.1	Engle-Grangers to-steps metode	95
5.6.2	Johansen-metode	97
6	Konklusjon	99
6.1	Resultatene	100
6.2	Sammenligne de ulike modellene	101
6.3	Problem	101
6.4	Videre arbeid	102
	Bibliography	104

Figurer

1.1	Natugassens innhold (HiST, 2015)	4
1.2	Ukonvensjonell og konvensjonell boring (NUPI, 2015)	5
1.3	Verdens naturgass forbruk målt i antall milliarder kubikkfot (Indexmundi, 2015)	6
1.4	Naturgass eksportert fra Norge (SSB)	9
1.5	Oljeprisen fra November 2013 til November 2014.	11
1.6	Gassproduksjon og forventet produksjon i USA, 1990 til 2035. Målt i trillioner kubikkfot (NUPI, 2015)	18
1.7	Rørledningsnett for naturgass (BBC, 2015)	19
2.1	Tilfeldig gang med $n = 1500$, $\mu = 0$ og $\sigma = 1$	23
4.1	Scatterplott for copula (4.2), $\rho = 0,7$	47
4.2	Scatterplott for copula (4.3), $\rho = 0,7$ og $\nu = 4$	48
4.3	Scatterplott for copula (4.4), $\rho = 5$	49
4.4	Scatterplott for copula (4.5), $\rho = 3$	50
5.1	Plott over naturgassprisene	56
5.2	Log-avkastninger for naturgasspriser	59
5.3	Empirisk tetthet til log-avkastningene	62
5.4	QQ-plot av log-avkastningene	63
5.5	De kvadrerte log-avkastningen	64
5.6	ACF og PACF (NCG)	67
5.7	ACF og PACF (TTF)	68
5.8	ACF og PACF (Zeebrugge)	69
5.9	NCG, PACF og ACF av de kvadrerte log-avkastningene	71
5.10	TTF, PACF og ACF av de kvadrerte log-avkastningene	72
5.11	Zeebrugge, PACF og ACF av de kvadrerte log-avkastningene	73
5.12	Tettheten til de standardiserte residualene mot student t-fodeling	75
5.13	QQ-plot, standardiserte residual mot student t-fodeling	76
5.14	QQ-plot, standardiserte residual mot student t-fodeling	77
5.15	ACF på de standardiserte empiriske residualene	80
5.16	GARCH(1,1) modellen og de observerte log-avkastningene	82
5.17	Scatterplott	86
5.18	Avhengighet mellom NCG og TTF	88

5.19	Avhengighet mellom TTF og Zeebrugge	89
5.20	Sammenligne avhengighet ved scatterplott	92
5.21	Sammenligne avhengighet ved scatterplott	92
5.22	Konturene til den empiriske copulaen og konturene til den tilpassete t-copulaen .	93
5.23	Prognose av de kvadrerte avkastningene ved copula- og GARCH-modell for den 9. Januar 2015.	94
5.24	NCG og TTF gasspriser	96
5.25	NCG, TTF og Zeebrugge priser i 2014	97
5.26	10 dagers prediksjon for NCG, TTF og Zeebrugge priser	98

KAPITTEL 1

Introduksjon

1.1 Innledning

Vi skal i denne oppgaven lage statistiske modeller på gasspriser i det europeiske naturgassmarkedet.

Først i oppgaven skal vi se på gassmarkedet generelt. Der ser vi på hva naturgassen består av, og hvordan vi har forholdt oss til naturgassen opp igjennom tidene. Vi skal studere hvordan naturgass prises i gassmarkedet, og hvordan naturgassen handles og selges i gassmarkedet. Videre skal vi se på ulike faktorer, som kan være med på å påvirke tilbudet og etterspørselen etter naturgass. En forandring i tilbudet eller etterspørselen etter gass vil igjen kunne påvirke naturgassprisen. Det er derfor nyttig å vite hvorfor gassprisene varierer.

Deretter skal vi studere statistisk teori som er relevant for analysene. Først skal vi introdusere tidsrekker. Vi skal bemerke oss stasjonære tidsrekker, og hvordan vi kan transformere priser om til avkastninger, ettersom priser vanligvis ikke er stasjonære. Dessuten skal vi introdusere kointegrasjon. Kointegrasjon er nyttig for å studere relasjonene som er mellom de ulike gassmarkedene. Videre ser vi på lineære modeller.

Deretter skal vi studere volatilitets modeller som ARCH- og GARCH-modeller. For å predikere og analysere volatiliteten som varierer over tid i gassmarkedet, kan vi bruke GARCH-modeller.

Det er grunn til å tro at de europeiske gassprisene er avhengige av hverandre. Copula blir brukt til å modellere avhengighetsstrukturen, som er mellom de marginale fordelinger. Vi skal i kapittel 4. studere copulateori.

I nest siste kapitlet gjennomfører vi selve analysene og modelleringene på naturgassprisene. Først studerer vi selve naturgassprisene. Deretter transformerer vi prisene om til log-avkastninger. Disse log-avkastningene blir testet for stasjonærhet. Etterpå bruker vi log-avkastningene og tilpasser volatilitetsmodeller. GARCH modellene blir brukt til å predikere fremtidig volatilitet. Videre ser vi om det er avhengighet mellom gassprisene. Copula blir brukt til å modellere avhengighetsstrukturen, som er mellom de marginale fordelingene i de europeiske gassmarkedene.

Til slutt i analysene blir det brukt kointegrasjon for å beskrive sammenhengen mellom de europeiske gassprisene.

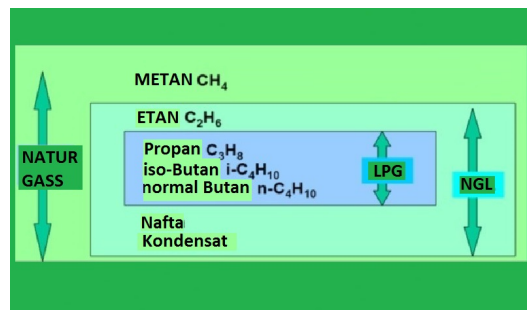
I slutten av oppgaven er selve diskusjonen, der vi først ser på hva vi har funnet ut. Diskuterer om det er noen modeller som egner seg bedre enn andre til å modellere gasspriser. Hvilke egenskaper som gasspriser har som gjør det vanskelig å predikere. Diskuterer om det finnes ting i oppgaven som kunne blitt gjort annerledes og forslag til videre arbeid.

1.2 Naturgassmarkedet

Det er viktig å bli godt kjent med gassmarkedet før vi begynner med selve modelleringen. I denne delen skal vi se på hva som kan påvirke gassprisen, og hva som gjør at gassprisene blir ustabile i perioder.

1.2.1 Generell fakta om naturgass

Naturgass er et fossilt brennstoff som består i hovedsak av metan(CH_4) og andre hydrokarboner som etan(C_2H_6), propan(C_3H_8) og butan(C_4H_{10}). I tillegg kan den også inneholde vekslende mengder av andre gasser, som nafta, nitrogen, karbondioksid og hydrogen sulfid. Mengden av de forskjellige stoffene varierer i forhold til hvilken gass det er (Lundberg, 2012).

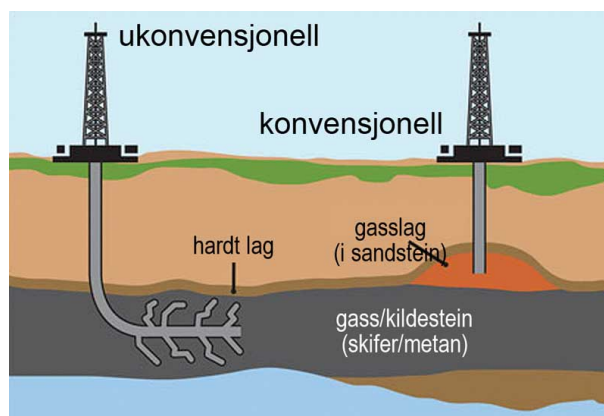


Figur 1.1: Natugassens innhold (HiST, 2015)

Naturgassen er en fargeløs og brennbar gass som er lettere enn luft. Den kan finnes i porøse bergarter i jordskorpen, både i rene gassfelt eller sammen med råolje. Naturgassen ble dannet for millioner av år siden da organisk materiale som alger, planter og rester av dyr ble begravd av sand, slam og steiner. Da bygget det seg opp et høyt trykk. I stedet for at det organiske materialet ble brutt ned, ble det liggende under dette høye trykket, og den høye temperaturen slik at det ble omdannet til olje og naturgass (Energifakta.no, 2013b).

Denne typen gass kalles for konvensjonell gass og kan bli produsert med ikke alt for mange vanskeligheter. Det finnes også ukonvensjonell gass som er mer krevende å få produsert. Den ukonvensjonelle gassproduksjonen krever bruk av fracking-utvinningsmetode, eller stimulering

for at gass kan unnslippe fra en lav gjennomtrengelige sten og strømme gjennom et borehull til overflaten. Ukonvensjonell gass er en felles betegnelse for skifer-gass, tight gas og CBM (coal bed methane) (europeunconventionalgas, 2014). I figur 1.2 kan vi se forskjellen på hvordan en borer etter konvensjonell og ukonvensjonell gass som finnes i naturen.



Figur 1.2: Ukonvensjonell og konvensjonell boring (NUPI, 2015)

Naturgass i den formen den utvinnes fra gassfeltene, inneholder som regel en blanding av tørrgass og våtgass, og går under betegnelsen rikgass. Tørrgass er navnet på naturgass etter at våtgassen er skilt ut, og består av 85 prosent metan. Resten av tørrgassen er for det meste etan. Våtgass er den delen av naturgassen som forekommer naturlig flytende, som eksempelvis kondenserbare hydrokarboner som propan og butan. Våtgass kan også inneholde etan, pentan og andre tyngre komponenter, og egner seg dårlig for rørtransport over lengre avstander (Energifakta.no, 2013a). Naturgassen som blir utvunnet fra gassfeltene blir skilt. Det vil si at gasselskapene skiller våtgassen ut i et prosesseringsanlegg som kondensat. Deretter kan tørrgassen transporteres videre til sluttbrukere i rørledninger (Energifakta.no, 2013a). Det finnes flere undergrupper av gass som LPG (Liquid Petroleum Gases) og NGL (Natural Gas Liquid). LPG er kondensert propan og butan, og det er denne gassen som er mest vanlig å bruke som drivstoff til for eksempel bil. NGL er tørrgass som er gjort flytende.

1.2.2 Naturgassen sin rolle opp igjennom tidene

Naturgass ble først sett på som et biprodukt fra oljeproduksjon, og ble dermed slippet rett ut i atmosfæren eller etterlatt igjen i grunnen. For Europa var naturgass omtrent ukjent helt frem til etter andre verdenskrig. Først ble det funnet små felt i Tyskland, Frankrike og Italia, men denne gassen ble kun konsumert lokalt. I 1959 fant de et stort felt i "Groningen" i Nederland. Det var først da produksjonen var stor nok til å bli eksportert til andre nasjoner i Europa. På 1960-tallet ble det også funnet felt i Nord-sjøen og i Storbritannia (Honore, 2010, s. 10).

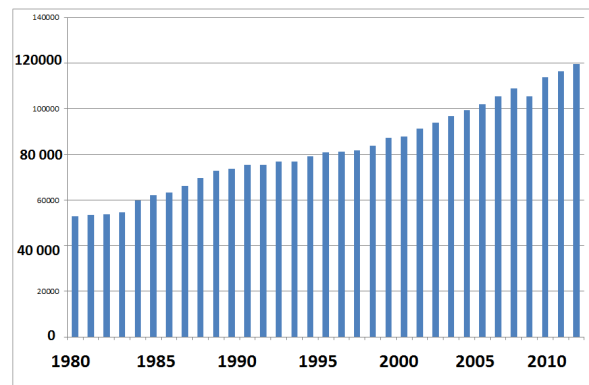
På 1970-tallet ble naturgassen ofte brukt som ett substitutt for olje. Naturgassen ble brukt til de samme energiformålene som en brukte olje til. Naturgassen ble blant annet brukt til oppvarming

og matlaging i husholdninger, industriformål og i kraftproduksjon (Honore, 2010, s. 11).

I begynnelsen av gassindustriens historie hadde flere land allerede kullkraftverk som energikilde. Å omgjøre kullkraftverkene til gasskraftverk var dyrt, derfor tok det lang tid før gass ble populært. I de nye kraftverkplanene ble gass oftere inkludert. Etter 1973 minket olje som brensel for kraftverk, samtidig fikk kull og kjernekraftverk en økning (Honore, 2010, s. 10-11). Det var først i 1990-årene at naturgass ble mer brukt i kraftverk. Naturgassens økende bruk skyldtes først og fremst teknologisk utvikling, som bedret termodynamisk effektivitet av gasskraftverk. En annen grunn til den økende bruken av naturgass var forandringer av regulerte avtaler for elektrisitetsindustrien, miljøfordeler for gass og mer stabilitet i elektrisitetsindustrien (Honore, 2010, s. 28). Gasskraftverk var lenge den viktigste sektoren for vekst i gasskonsum i den vestlige verden.

Produksjonen av ukonvensjonell gass som skifer-gass har økt kraftig siden 2000-tallet. Grunnen til dette er at det er kommet nye teknologier som fracking-utvinningsmetode og horisontal boring. Disse nye teknologimetodene har gjort at de ukonvensjonelle ressursene har blitt økonomisk lønnsomme å bruke (europeunconventionalgas, 2014).

Naturgass har opp gjennom tidene vært en stor suksess i Europa. I år 1960 stod det globale energiforbruket fra naturgass for omtrent 2,5 prosent av det totale energiforbruket, mens i år 2012 var stod naturgassen for hele 24 prosent. En forventet i 2012 at dette skulle øke ytterligere. Fra året 1990 til 2008 har det vært en årlig vekst på bruk av naturgass på 3,1 prosent (år 2000 til år 2008 en årlig vekst på 2,1 prosent) (Honore, 2010, s. 12-15). I figuren 1.3 ser vi hvordan verdens naturgass forbruk er økt de siste 30 åra.



Figur 1.3: Verdens naturgass forbruk målt i antall milliarder kubikkfot (Indexmundi, 2015)

I år 2009 ble energisektoren som alle andre sektorer påvirket av finanskrisen. Gassetterspørselen falt fra slutten av år 2008 og utover i 2009 (Honore, 2010, s. 16). Derimot i 2010 var gassetterspørselen nesten tilbake på samme nivå som før 2009. Den økende etterspørselen i 2010 var drevet av kaldt vær, men også sterk underliggende vekst i gassetterspørselen. Etter finanskrisen har kullet hatt en overraskende renesanse i Europa (Tollaksen, 2013). Kullets renesanse har

ført til at etterspørselen etter naturgass har falt. I motsetning er det ventet økende etterspørsel etter gass i Kina og India de neste årene, dette er grunnet sterk økonomisk vekst og et stort behov for energi (Oljeogenergidepartementet, 2011).

I 2011 fortsatte gassforbruket å falle, det ble registrert det laveste forbruket siden år 2009. Grunnen til dette lave forbruket kan være de milde temperaturene og den fortsatt svake økonomiske veksten i Europa (EU-kommisjonen, 2012). Den svake økonomien i Europa stimulerer til å bruke billigere energi som kull istedenfor naturgass, selv om kull er vesentlig mindre miljøvennlig.

I begynnelsen av februar i 2012 var gassprisene ekstremt volatile. Det var på grunn av en uventet økt etterspørsel etter naturgass. Volatiliteten var kortvarig, og ved midten av februar gikk prisene tilbake til mer normale nivåer (EU-kommisjonen, 2012). Den nedadgående trenden i forbruket av naturgass fortsatte i 2012. Det årlige forbruket av naturgass i 2012 var fire prosent lavere enn i 2011. Samtidig fortsatte LNG leveranser til EU å falle raskere enn forbruket. I 2012 ble det registrere en nedgang av LNG leveranse på 31 prosent i forhold til 2011 (EU-kommisjonen, 2013). Samtidig i 2012 fortsatte EU sitt kullforbruk og kullimport å stige. Europa sitt forbruk og import av kull (steinkull og brunkull) økte med henholdsvis to og ni prosent i løpet av de første elleve månedene av 2012, sammenlignet med samme periode i 2011. Kullimport fra Colombia og USA inn i EU har doblet seg siden år 2009 (EU-kommisjonen, 2013).

I de siste årene har det blitt et stadig mer fleksibelt og integrert gassmarked i Europa (EU-kommisjonen, 2013). Dette har vist seg å være svært effektivt. For eksempel ved kortvarige kuldeperioder i en del av Europa, har gassen blitt fraktet til der det er mest nødvendig. Dermed har det raskt blitt etablert en balanse mellom tilbud og etterspørsel. For eksempel har prisforskjellene mellom nederlandske og tyske handelsknutepunkter oppfordret å ta gass fra det tyske lagringsanlegget til det nederlandske markedet (EU-kommisjonen, 2013).

Kjøpe og selge gass via handelknutepunkter er stadig økende. I 2012 ble 83 prosent av det totale forbruket av naturgass levert via handelknutepunkter, sammenlignet med 70 prosent i 2011 (EU-kommisjonen, 2013). Tilsvarende synker oljeindeksregulerende naturgasskontrakter i EU (EU-kommisjonen, 2013).

Selv om naturgass har blitt mer fleksibelt de siste årene, har naturgass mistet andel i kraftproduksjonen, grunnen til dette er den økende andel av kull og fornybar-energi i elektrisitetsproduksjonen. I 2013 ble 12,4 prosent kraftproduksjonen produsert fra naturgass mot 16 prosent i 2011 (EU-kommisjonen, 2014b).

I 2014 fortsatte gassforbruket og synke i EU, det sank med cirka ti prosent i forhold til året før (EU-kommisjonen, 2014c). Etter mange år med nedgang i import av LNG, ble det i 2014 registrert en økning (EU-kommisjonen, 2014c). Denne økningen medførte at prisforskjellen mellom Asia og Europa nesten forsvant. Samtidig økte likviditeten i de europeiske knutepunktene i Europa i 2014, fordi det nederlandske TTF markedet overtok Storbritannias NBP marked (EU-kommisjonen, 2014c). Nedgangen i råoljeprisen i 2014 vil påvirke oljeindekserte kontrakter gradvis fra 2015 (EU-kommisjonen, 2014c). Fallet i råoljeprisen fører også til en forventning om at gassprisen faller ytterligere i 2015. I 2014 stod Russland og Norge for 80 prosent av den totale importen til EU, i 2010 var det 64 prosent (EU-kommisjonen, 2014c).

1.2.3 Energi måleenheter

Det finnes ganske mange måter å måle energi på, som kan virke forvirrende. Det jobbes det med å finne en standard måleenhet for energi. Vi har i denne oppgaven brukt flere ulike enheter for energimål. Om det ikke finnes nok måter og måle energi på, kan naturgass måles i både energi og i volum. Når det gjelder naturgassprisene i denne oppgaven har vi brukt pris per kWh, men ellers har vi også brukt andre måleenheter i løpet av oppgaven.

1.2.4 Naturgassmarkedet i Norge

Norge produserer omtrent 100 milliarder standard kubikkmeter oljeekvivalenter (Sm^3 o.e.)¹ naturgass i året. Gassandelen av det totale petroleumssalget målt i Sm^3 er forventet å øke fra 46 prosent i 2010 til 51 prosent i 2015.

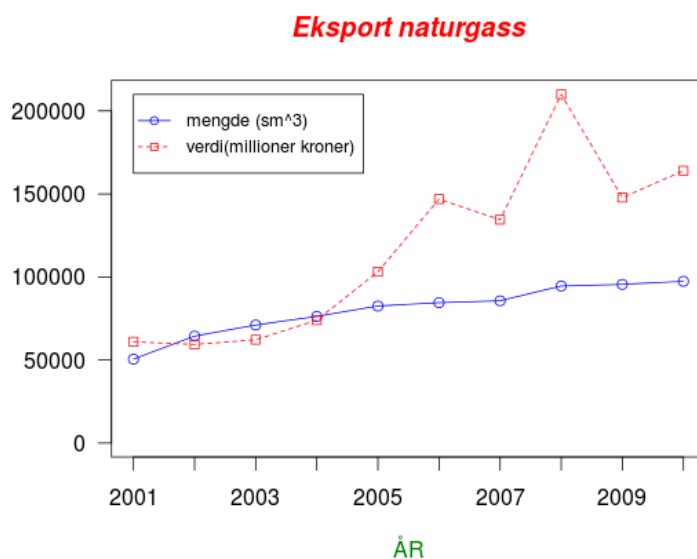
Over 95 prosent av det Norge produserer av naturgass selges til det europeiske markedet. Dette medfører at europeisk energiforbruk har stor betydning for norsk etterspørsel og norske gasspriser (Oljeogenergidepartementet, 2011). Norsk gass føres først og fremst til Europa via rørledningsnettene (Statoil, 2012). I tillegg er det i Norge prosessanlegg som omgjør gass til flytende LNG, slik at i de siste årene har det også blitt eksportert LNG fra Norge til utlandet.

Norge er et eksempel på hvor oljeproduksjonen synker, mens gassproduksjonen øker. Slik at norsk naturgasseksport utgjør en stadig økende andel av norske petroleumsinntekter. Dermed blir utviklingen av gasspriser mer og mer viktig for Norge sine fremtidige petroleumsinntekter. I første kvartal 2011 ble det eksportert naturgass for 47 547 millioner kroner, og råolje for 82 172 millioner kroner. Nedenfor i figur 1.4 kan vi se utviklingen av eksportert naturgass fra Norge, både når det gjelder mengde og verdi.

Historisk har norsk gass blitt solgt gjennom langsiktige kontrakter, der prisene har blitt justert i forhold til hvordan oljeprisen har utviklet seg med et tidsetterslep på tre til ni måneder (Statoil, 2010). I 2010 måtte Statoil reforhandle sine langsiktige kontrakter, og ble mer knyttet til europeiske gasspriser på spotmarkedet. Dette var fordi prisene i spotmarkedet var mye lavere enn oljepris relaterte langtidskontrakter (Winje, 2011).

Skifergassrevolusjonen i USA har gitt et prisdump på 3,25 mmBtu (million British thermal units) naturgass i det amerikanske markedet. På grunn av skiferrevolusjonene fikk USA et overskudd av kull som kan selges billig til Europa (Tollaksen, 2013). De norske gassprisene har vært lite påvirket av gassprisene i Amerika på grunn av den fysiske avstanden. Det at Europa kjøper billig kull fra USA, kan også svekke prisene for naturgass. I tillegg er det mulig at USA bygger ut sin gassproduksjon enda mer og begynner å eksportere LNG i stor skala. Det vil medføre at de lave prisene i USA påvirker hele verden, også Norge. Dermed kan verdien av norsk gasseksport bli lavere enn tidligere antatt (Winje, 2011).

¹ Sm^3 er forkortelsen som brukes for standard kubikkmeter, eller en kubikkmeter gass under normale forhold, definert som 1 atmosfære trykk (1,01325 bar) og ved 15 celsius (Gassco, 2012).



Figur 1.4: Naturgass eksportert fra Norge (SSB)

1.2.5 Hvordan naturgass prises i gassmarkedet

Naturgass prises forskjellig i de forskjellige regionale gassmarkedene, i motsetning til oljepris som bestemmes globalt. Dette er fordi mye av gassen blir fraktet gjennom rørledninger som medfører at markedet er svært begrenset. På grunn av geografiske begrensninger er gassen som blir solgt via rørledningssystemet i USA ikke direkte konkurrent med gass som blir solgt via rørledninger i Europa. Det finnes flere måter å prise naturgass på (Winje, 2011).

Det europeiske gassmarkedet har to ulike markeder. I Storbritannia bruker de først og fremst spotmarked for gass. Med spotmarked menes det at gassen handles på børs og har en pris som reflekterer tilbud og etterspørsel. På kontinentet blir naturgassen mye solgt via langsiktige kontrakter, men også her er vokser spotmarkedet fort. I USA bestemmes gassprisen omtrent utelukkende av spotmarkedet (Oljeogenergidepartementet, 2011), mens i Asia finnes det også en del regulerte priser.

Langtidskontrakter

Olje og gass var omtrent substitutt tidligere, slik at hvis naturgass var billigere enn olje hadde kunden mest lyst på naturgass. Dette kunne kunden endre på relativt kort sikt. Det er en av hovedgrunnene for at europeiske langtidskontrakter for naturgass er knyttet opptil oljepriser. Det er også tre andre grunner for å ha langtidskontrakter bundet opp til oljepris. Den ene grunnen er at produsentens risiko er avhengig av variasjoner i oljeprisen. Den andre grunnen er at bankene er fornøyd med å låne penger til formål som er knyttet opp til oljepris. Den

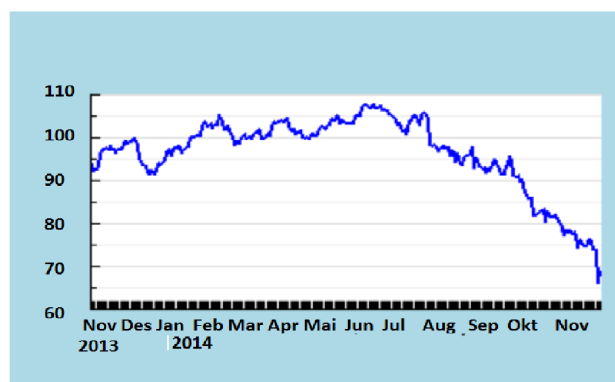
siste grunnen er at gassprodusentene som er få og sterke ikke kan påvirke prisene selv. Hvordan langtidskontrakter er bygd opp kan variere en del, men generelt er det en basispris av naturgass pluss/minus utvikling av oljeprisene (Honore, 2010, s. 50-51).

I Norge, Nederland og Russland er 85 til 92 prosent av indeksen i langtidskontrakter knyttet til oljeprisen. Disse landa står for 43,5 prosent av det totale konsumet i verden, og har dermed stor innflytelse på hva europeerne må betale for naturgassen. Storbritannia har en annen struktur, der 37 prosent av indeksen på kontrakten er knyttet opp til spotmarkedet for naturgass, 28 prosent til inflasjon, 20 prosent for oljeprodukter og 15 prosent til annet (Honore, 2010, s. 52).

I slutten av år 2000 var oljebaserte langtidskontrakter fremdeles mye brukt i fastlands Europa, men bruken av disse kontraktene var synkende. I år 2012 ble det 83 prosent av det totale forbruket av naturgass levert via handelsknutepunkter, sammenlignet med 70 prosent i 2011 (EU-kommisjonen, 2013). Tilsvarende synker oljeprisindeksregulerende naturgasskontrakter i EU (EU-kommisjonen, 2013).

Tidligere har prisene på gass fulgt oljeprisen med et visst tidsetterslep, hvor gassprisene har ligget på cirka 70 prosent av råoljeprisen. Det var dermed ikke store forskjeller mellom gassprisene i langtidskontrakter og i spotmarkedet. I 2008 medførte den internasjonale nedgangskonjunkturen til redusert etterspørsel etter gass både som innsatsfaktor i industrien og som følge av et lavere energiforbruk. Dette kombinert med økt tilbud av skifergass spesielt i USA, førte til at det ble ubalanse i det globale gassmarkedet. I tillegg førte den økte fleksibiliteten av gasstransport til at de regionale markedene ble mer påvirket av hverandre. Det førte til at prisene i spotmarkedet til gass ikke falt kun i USA, men også i Europa (Likvern, 2012a). Dette ga lave spotpriser i Europa, mens de oljebaserte langtidskontrakten fremdeles hadde høye gasspriser. I USA holder gassprisene fortsatt et lavt nivå, mens i Europa har det vært økning i gassprisen fra 2010 til 2011 (Winje, 2011). Dette er mye på grunn av den sterke utvikling i oljeprisen, som den europeiske gassprisen baserer seg på.

Langtidskontrakter basert på oljepris, har i perioder vært mye høyere enn prisen på spotmarkedet. I år 2006 og i år 2007 falt prisene i spotmarkedet i Storbritannia på grunn av det økte tilbudet fra Norge, men det påvirket ikke den tyske importprisen som var basert på langtidskontrakt (Honore, 2010, s. 58). Dermed blir gassprisene unaturlig høye. Problemet med oljebaserte gasspriser, er at gassprisene blir unaturlig høye. Dette er negativt for gassetterspørselveksten, spesielt innenfor kraftsektoren som er en av de største sektorene for naturgass. Som vi vet konkurrerer naturgass i denne sektoren i hovedsak mot kull, atom og vannkraft. Slik at høye gasspriser som oljebaserte gasspriser dermed ikke er attraktiv for kraftsektoren (Honore, 2010, s. 59). I siste halvdel av 2014 er situasjonen for oljeprisen snudd, og oljeprisen har svekket seg. I figur 1.5 er en oversikt over oljeprisen fra november 2013 til november 2014. Oljeprisen sank med 26,51 prosent fra november 2013 til november 2014.



Figur 1.5: Oljeprisen fra November 2013 til November 2014.

Langtidskontrakter basert på oljepriser er nødvendigvis ikke like logisk lenger. Før var oljebaserte kontrakter nødvendig på grunn av mangel av flytende gassmarkeder. I tillegg var olje og gass brukt til mye av det samme. I år 2010 var det omtrent ingen oljeprodukter som konkurrerte direkte mot naturgass, for eksempel brukes olje omtrent ikke lenger i kraftverk. Fremover i Europa vil langsiktige kontraktbaserte pris inngåelser reguleres av handelsknutepunkter for gass som NBP i Storbritannia og TTF i Nederland istedenfor råoljeprisen (Tollaksen, 2013).

Spotmarked

Naturgass har lenge ikke hatt et konkurransedyktig marked, men i det siste har EU ønsket at dette markedet skal bli mer konkurransedyktig. For å fremme denne konkurransen har det blitt mer fokus på at gasspris skal mer ut på spotmarkedet. Markedsbaserte gassspotpriser følger tilbud og etterspørsel sesonger. På sommeren når det er lav etterspørsel pleier Storbritannia å eksportere gass til kontinentet hvor det der er høyere oljepris relaterte gasspriser. Omvendt på vinteren (Honore, 2010, s. 217).

Den mest sannsynlige prisingen av naturgass som overtar etter de oljebaserte langtidskontraktene er priser basert på spotmarkedet for gass, slik som NBP (National Balancing Point) i Storbritannia. Det er lite trolig at alle konsumentene vil basere seg på et spotmarked i Europa. Dette er på grunn av de fysiske begrensningene som fremdeles er i gassmarkedet (Honore, 2010, s. 58).

Selv om en ikke utelukkende vil basere seg på spotmarkedet, vil det i fremtiden være langtidskontrakter som i større grad baserer seg på spotmarkedet for gass fremfor olje. En økende bruk av dette har allerede skjedd, spesielt i Norge og i Russland. De nye kontraktene ser ut til å vare i en kortere tidsperiode, med en større indeks på spotmarkedet for gasspris. Denne prisingen kan dermed bli kortvarig dersom etterspørselen etter gass tar seg opp igjen, tilbudet faller eller at oljeprisen synker under 50 dollar per fat (\$ 50/bbl) (Honore, 2010, s. 58).

1.2.6 Kjøp og salg av naturgass

I det siste har det kommet flere nye handelsknutepunkter i Europa. Et handelsknutepunkt (hub) er hvor transaksjonen skjer mellom kjøper og selger, der vi har visse betingelser som gir likviditet. Det kan være et fysisk punkt hvor flere gassrør og gasslager er knyttet sammen (f.eks. Henry Hub i USA), eller det kan være ett virtuelt punkt innenfor ett gassrør system som i NBP i Storbritannia. De dominerende handelsknutepunktene på kontinentet er Zeebrugge som er fysisk handling og TTF som er virtuell handling. Handelsknutepunkter er en markeds plass for naturgass, og det gir leverandørene mange selgere og kjøpere til gass samlet på en plass. Handelsknutepunkter gjør at leverandørene har mulighet til å tilpasse forandringer i tilbud og etterspørsel, ved enkelt å selge gass på kort og mellomlang sikt. Handelsknutepunkter blir ofte prissettingspunktet for markedet og gir dermed et åpen innsyn for gasspriser basert på gass til gass konkurranse (Honore, 2010, s. 216-217).

Det finnes to måter å handle naturgass på, den ene er ved fysisk handel og den andre er finansiell handel. Fysisk handel er når det skjer en fysisk levering eller ett fysiske mottak av naturgass som når en har et lagringsanlegg, LNG eller rørledning som tillater overføring fra en gassledning til en annen, mens finansiell handel er hvis handelen kun involverer overføring av penger. Det vil si at kjøperne driver med gjenforhandlinger. Handelsknutepunktene gjør at kjøpere kan se etter de billigste forsyningskildene, mens de tilbyr selgere med ett godt nettverk for å nå kjøpere som er villige til å betale den beste prisen (Honore, 2010, s. 217).

Storbritannia sitt NPB marked var lenge det eneste Europeiske spotmarkedet, mens nå er det flere regionale handelsknutepunkter som har dukket opp over flere år som i Nederland(TTF), Belgia(Zeebrugge), Tyskland(Gaspool og NCG) og Frankrike(PEGs). Disse varierer betydelig i forhold til deres handelsvolumer, tilgang til kapasitet, antall handlere, grad av likviditet og åpne priser og hvorvidt de er regulert (Honore, 2010, s. 218). Siden midten av 2000-tallet har likviditeten i europeiske handelsknutepunkter vokst mye. Den mest aktive handelsknutepunkten i 2008 var NPB med 960 bcm handlet, der 67 bcm var fysisk levering. På kontinentet ble 187 bcm naturgass handlet, hvor 55 bcm av disse var fysisk levering. I år 2008 ble 20 prosent av Europeisk naturgass solgt via Europeiske handelsknutepunkter for gass (Honore, 2010, s. 218), mens i 2012 ble hele 83 prosent av det totale forbruket levert via handelsknutepunkter (EU-kommisjonen, 2013).

1.2.7 Tilbud, etterspørsel og priser

I perfekt konkurranse sett fra produsentenes synspunkt er prisene satt når etterspørsel er lik tilbud (myndighetene kan ikke regulere etterspørsel, tilbud eller pris). Kun den minste profitten for at en bedrift skal overleve er den som gjelder. Hvis en aktør øker prisen for å øke profitten, kommer en annen produsent inn for å drive med minimums profitt, slik holdes prisen nede ved perfekt konkurranse. Dette er det mest effektive markedet, og det beste for kundene (Honore, 2010, s. 56). Økt etterspørsel vil i følge kjent økonomisk teori føre til økte priser, hvis tilbudet ikke endres. Stigende tilbud vil føre til lavere pris, hvis etterspørselen ikke endres. Dette forutsetter at det er ingen substitutt. Derimot har ikke gassmarkedet perfekt konkurranse og det er ofte

regulert av myndighetene. Gassmarkedet har også mange langtidsinvesteringer slik at det kan ikke reagere på prisendringer på kort sikt. Det å lagre gass er heller ikke like lett, samtidig som gass har andre substitutt som for eksempel olje og elektrisitet.

1.3 Tilbud og etterspørsel faktorer

Det er flere faktorer som kan påvirke tilbud og etterspørsel. Vi kan nevne økonomisk vekst og svingninger, teknologisk utvikling, vær og temperatur svingninger, naturkatastrofer og modenhet av markedet. I tillegg er det miljø og energipolitikk. Det kan være politiske regionale reguleringer for å begrense miljø ødeleggelser, begrense uttak av naturressurser eller for å fremme sin egen industri i sitt eget land. Ikke minst priselastisitet, og prisen på gass relativt til andre energiformer kan påvirke etterspørselen

Naturgassmarkedet er et komplisert marked. Det er fordi det er flere endogene variabler som spiller inn, som eksempelvis kan være andre energipriser, teknologi, økonomi og temperatur. Dette medfører endringer i etterspørselen og gjør at det er vanskelig å lage gode prognoser på gasspriser

1.3.1 Vær, temperatur og årstider

I spotmarkedet for gasspris som i strømpris er det sesongsvingninger i prisene som følger årstidene. I tillegg blir prisene påvirket av spesielle værforhold som temperatur og naturkatastrofer. En kan for eksempel nevne at i USA var det spesielt høye priser i vinteren i 2000/2001 og i vinteren i 2002/2003 på grunn av lave temperaturer (Likvern, 2012b). Det var også en oppgang i prisen i 2005 som skyldes orkanen Katrina (Likvern, 2012b). I Februar i 2012 var det ekstremt høye gasspriser i Europa på grunn av ekstremkulde i øst Europa og i Russland. Dette er en av grunnene til at det er mange aktører som foretrekker langsiktige oljerelaterte kontrakter (Likvern, 2012b).

1.3.2 Økonomisk vekst og økonomiske svingninger

Den økonomiske veksten påvirker både energietterspørselen og tilbudet.

Verdensøkonomien gikk inn i en lavkonjunktur i 2008. Global BNP falt med fem prosent fra 2007 til fjerde kvartal i 2008. Den økonomiske aktiviteten var svakere enn først antatt både i Tyskland, Italia, Finland, Danmark, Estland, Sverige og England ,som førte til et stort fall i gassetterspørselen i slutten av år 2008. Det vil si at gassetterspørselen falt med 5,3 prosent i år 2008 (Honore, 2010, s. 41-43).

For å finne sammenhengen mellom energiforbruk og økonomisk vekst, kan vi se på forholdet mellom antall energi konsumert per BNP. Fra år 1960 til år 1970 var energi per BNP relativ

konstant. En prosentvis økning i BNP førte til en prosentvis økning i energi (Honore, 2010, s. 41).

I den senere tid har vi hatt økonomisk vekst uten at det har vært økning i energi konsumering. Det viste seg at sammenhengen mellom økonomisk vekst og energi konsumering var mer komplisert enn først antatt. Edmonds og Reilly forklarte dette med variasjoner i andre faktorer som teknologi, reguleringer, energipris og sosiale faktorer (Honore, 2010, s. 42).

Som vi vet falt gassetterspørselen med 5,3 prosent i år 2008 på grunn av den økonomiske krisen. Det har spesielt gått utover etterspørselen etter gass i industrien. Den økonomiske tilbakegangen har medført strukturelle endringer for deler av industrisektoren. På grunn av industrinedleggelsen antas det at en del av gassetterspørselen er permanent borte (Oljeogenergidepartementet, 2011). I perioden fra oktober i år 2008 og til april i år 2009 ble det kansellert over 20 planlagte store olje- og gassprosjekter, samtidig som 35 prosjekter ble utsatt i minimum 18 måneder (Honore, 2010, s. 49).

Fra 2009 til 2013 har etterspørselen naturgass fortsatt å falle. Grunnen til dette er den økende bruken av kull og fornybar energi til elektrisitetsproduksjonen. Kullet sin rennesanse i Europa er overraskende på grunn av miljøkonsekvensene, men hovedgrunnen til denne rennesansen er den lave prisen. Flere større land i Europa har gjennom flere tiår levd på kreditt for å få økonomisk vekst. En god del av disse landene er nå inne i en fase med flere innstramningstiltak for å redusere gjelden. Disse innstramningstiltakene kan stimulere til bruk av billigere fossile brennstoff som kull (Likvern, 2012a).

1.3.3 Konkurransen i de ulike naturgass sektorer

I år 2006 stod husholdninger og andre institusjoner som for eksempel sykehus for 37 prosent av gassetterspørselen. Videre følget kraftverksektoren med 31 prosent, og industrisektoren med 21 prosent. Disse tre er hovedkonsumentene til gass utgjør hele 89 prosent av all naturgass forbruk. De siste elleve prosentene er tap (en prosent), energi(tre prosent) og annet(transport, jordbruk, o.s.v.) (Honore, 2010, s. 23).

Kraftverk

Naturgass har etterhvert blitt mer og mer vanlig i kraftsektoren. Det er denne sektoren som har stått for den største veksten i gassetterspørsel. Fra år 1990 til år 1999 var det en økning av naturgass til kraftverk på 9,3 prosent per år, og 5,4 prosent per år mellom år 2000 og år 2007 (Honore, 2010, s. 18-23).

I Europa er de tre største måtene å skape elektrisitet på ved kjerne-, kull- og gasskraftverk, der omtrent en tredjedel av kraftverka bruker naturgass. Det blir generert omtrent 1000 TWh elektrisitet årlig ved kjernekraftverk, omtrent 1100 TWh ved kullkraftverk og rundt 900 TWh ved gasskraftverk i Europa (Honore, 2010, s. 29). Alle typer kraftverk(kull, vann, kjerne o.s.v) kan i mer eller mindre grad brukes som substitutter for gasskraftverk.

Kraftsektoren er den sektoren som har høyest priselastisitet, men det kan variere fra land til land. Priselastisiteten på etterspørsel vil si hvor stor prosentvis endring det er i kvantum for salg når prisen endres med en prosent. Det som bestemmer priselastisiteten i denne sektoren er om kraftsektoren har flere kraftverk eller om kraftverket i seg selv kan veksle mellom ulike brensel (flerbrenselkraftverk). Flerbrenselkraftverk kan veksle mellom å bruke ulike kraftverk/brensel etter hva som er billigst, men de har mistet marked de siste ti åra. Uansett eier et selskap som regel flere kraftverk og er dermed veldig fleksible. Daglig eller hver time må naturgass til kraftverk generelt i Europa konkurrere med kull- og kjernekraftverk på pris (Honore, 2010, s. 36-37). Andre ting som er med å bestemme priselastisiteten er politikk, tilgjengelighet og reguleringer.

I Norge er situasjonen annerledes der det omtrent er kun vannkraftverk som skaper elektrisiteten. I følge NVE står vannkraftverk i Norge for elektrisitetsproduksjon på cirka 123 TWh i et vanlig år. Det har vært flere planlagte gasskraftverk i Norge. Hvis disse kraftverka skulle komme, ville gasskraftverka blitt mer ett tillegg for å dekke den norske befolkningen sitt behov for strøm enn en erstatning av vannkraftverk.

Husholdninger og andre institusjoner

Husholdninger og andre institusjoner som for eksempel sykehus har vært og er fremdeles hovedkonsumenten til gass. Som nevnt over står denne sektoren for 37 prosent av etterspørselen, men har hatt en minkende vekst siden år 2000.

I denne sektoren blir naturgassen brukt direkte til oppvarming og matlaging. Det er vedfyring, elektrisitet, olje og fjernvarme som er de største konkurrentene for naturgass. Samtidig har denne sektoren har ikke lett for og skifte brensel på kort sikt og dermed liten priselastisitet (Honore, 2010, s. 34-35).

Industrisektor

Industrisektoren har hatt en minkende vekst siden år 1990, og en negativ utvikling fra år 2000. Den negative utviklingen er mye på grunn av den økonomiske delen. Naturgass er dyrt og mye av industrien er flyttet til India og Kina i den senere tid. På grunn av dette tror ikke Honore (2010, s. 25-26) at etterspørselen etter gass i Europa i industrisektoren vil endre seg dramatisk frem til år 2020.

Industrisektoren bruker gass til mye forskjellig, blant annet i industri som lager papp og papir, kjemikalier, metall, petroleumsraffinering og plastikk. Små og mellomstore industrier er lite priselastiske på kortsikt, med mindre de har installert maskiner som kan gå på ulike typer drivstoff som olje og gass. I Europa er de fleste kontrakter linket mest opp mot olje, og det gjør at det er lite poeng for bedrifter å installere dyre maskiner som kan gå på ulike drivstoff. For eksempel kan papp og papir industrien bruke kull, olje eller gass for å fremstille varene sine. Dette gjør at industrien flyttes til steder hvor tilgangen til brensel er billigst. Det vil si at

det er den økonomiske aktiviteten og nettverkstilgangen som påvirker etterspørselveksten etter naturgass (Honore, 2010, s. 35-36).

Transport

En veldig liten del av naturgass blir brukt som drivstoff i fremkomstmiddel. I dag er naturgass ikke mye brukt i personbiler, men det er mer brukt for kjøretøy som kjører mange mil hver dag, som for eksempel busser (Honore, 2010, s. 11) eller ferjer. Selv om drivstoff utgjør kun en liten del av naturgass forbruket, kan den være viktig i forhold til fremtidig etterspørsel etter naturgass. Dette er fordi transport er den viktigste sektoren for olje. Omtrent 2/3 av all olje går til transport (bensin, diesel og flydrivstoff). Hvis en kun ser på bensin til biler, brukes det i underkant av halvparten av all olje (Fichman, 2012). I prinsippet kan de fleste transportmiddel bytte ut "vanlig" drivstoff som bensin/diesel til "uvanlig" drivstoff som naturgass. Det ville gjort transport til en viktig sektor for naturgass.

Utenom gass er de viktigste energiformene til transport olje og elektrisitet, der olje er den absolutt dominerende. Biler kan kjøre på forskjellige typer gass, som CNG (komprimert naturgass), LNG (flytende naturgass) og LPG (propangass). LPG er den vanligste gassen som drivstoff (Gjelstenli, 2012). I 2011 ble det brukt omtrent 1 078 000 tonn bensin som drivstoff til transportmidler, og kun rundt 2000 tonn flytende naturgass som drivstoff til transportmidler (SSB). Bensinprisen den 19.09.12 var på ca 14.50 kr per liter inkludert merverdi. LPG (flytende gass) ligger på ca 7,25 kr per liter inkludert merverdi, som er halvparten av prisen til bensin. En bruker ca 40 prosent mer gass enn bensin per kjørelengde. Den reelle "literprisen" til gass er omtrent 9,90 kr per "liter" i forhold til bensin. Dermed er det billigere å kjøre med gass.

Hvis vi nå antar at LPG og bensin er gode substitutter altså at alle bensinbiler kunne også brukt naturgass. Da ville flere velge å kjøre med LPG siden LPG er billigere enn bensin. Etterspørselen etter gass blir større og prisen forventes dermed å øke. Når mange nok velger å kjøre på gass vil den reelle gassprisen nærme seg bensinprisen. Hvis bensinprisen går opp eller ned er det naturlig og tro at den reelle LPG-prisen også vil endres i takt med bensinprisen. Nå er det slik at det ikke er mange som har den muligheten å velge å kjøre med naturgass istedenfor bensin. Slik at denne krysspriselastisiteten mellom bensin og LPG egentlig er ganske liten.

En grunn til at det ikke blir satset mer på transport som går på naturgass, er uenigheten om hvor vidt LPG er miljøvennlig i forhold til bensin og diesel. Miljøorganisasjoner som for eksempel Bellona mener at gassbiler er en kortsiktig løsning. Argumentene deres er at LPG slipper ut like mye CO₂ som vanlig drivstoff (Mæland, 2009), men hvis vi ser på de lokale utslippa er miljøgevinsten stor. Det kan dokumenteres at utslipp av andre miljøgasser og partikler lokalt er mye mindre ved bruk av LPG enn ved bensin og diesel. En vitenskapelig test gjennomført ved University of Marbor i Slovenia viste at utslipp av farlige avgasser spesielt lokalt var mindre ved LPG enn ved bensin. De sammenlignet utslipp av gasser fra samme motor drevet av LPG og bensin. De fant blant annet ut at NO_x (nitrogenoksid) som gir helseskader for mennesker ble redusert med 41 prosent lokalt og med 77 prosent ekstra lokalt ved bruke naturgass istedenfor bensin som drivstoff. Se Tasic (2011) for mer informasjon om forsøket. Lokal forurensing er

et stort verdensproblem. Lokal utslipp er en svært aktuell problemstilling idag. I Oslo har det blitt innført dieselbil stopp på kalde vinterdager på grunn av lokale utslipp. Samtidig viser tall fra Verdens helseorganisasjon (WHO) viser at sju millioner mennesker dør årlig av sykdommer knyttet til luftforurensing, der ca halvparten av disse dør på grunn av utendørs forurensing (Aftenposten, 2014). Å endret drivstoff fra olje til naturgass hadde vært ett viktig steg i riktig retning for å bedre luftkvaliteten.

1.3.4 Miljø og politiske og regionale reguleringer

Politiske og regionale reguleringer kan forekomme både ved hensyn på økonomiske og/eller miljøgevinster.

I 2007 fikk EU et miljømål som kalles 20-20-20. Fram til år 2020 skal CO₂-utslippene reduseres med 20 prosent, fornybar energi økes til 20 prosent og effektivisering skal bidra til reduksjon av etterspørsel med 20 prosent. Gjennomføringen av disse målene kan ha betydning for naturgassens fremtidige rolle i Europa. Det kan ha negativ effekt på gassetterspørselen, men det kan også påvirke etterspørselen positivt. Den 23. Oktober i 2014 ble EU-lederne enige om nye mål. Innen år 2030 skal det være en innenlandsreduksjon av drivhusgasser på minst 40 prosent i forhold til nivået i 1990. Det settes også et mål om at minst 27 prosent av all energi skal være fornybar energi innen år 2030 (EU-kommisjonen, 2014a).

I dag blir cirka 28 prosent av all elektrisitet i Europa er produsert ved kullkraftverk (Honore, 2010, s. 40). Europeiske myndigheter prøver å fornye deres energimarkeder slik at de kan redusere karbonutslippene. Naturgass er renest av alle fossile brennstoff. Dersom alle kullkraftverk erstattes med gasskraftverk reduseres 30 prosent av CO₂ utslippene fra kraftverk, som er nok til å oppfylle Europas CO₂ målsetting.

Mange av de europeiske kullkraftverka er veldig slitt, og investeringene for å bygge nye gasskraftverk er nå lavere enn å bygge nye kullkraftverk. Dermed er det mulig at det blir bygget flere gasskraftverk i fremtiden. I tillegg gir gasskraftverk fleksibel produksjon og tar kortere tid å bygge. Det kan danne betydelig vekst i gassetterspørselen i Europa de neste årene.

Det er også stor global vekst i fornybar energiproduksjon som vind, sol og biomasse energi. Spesielt solenergi satses det mye på i Europa. Bruk av fornybar energi er helt klart det mest miljøvennlige, men det vokser fra et veldig lavt nivå. Slik at fornybar energi fremdeles forventes å kun utgjøre en mindre andel av verdens energikilder i lang tid fremover. Det er derfor mulig at verdens bruk av fossile energibærere går i retning av de mer klimavennlige fossile brennstoffene som for eksempel naturgass. Som nevnt over er dette en måte å kutte CO₂-utslipp betydelig på.

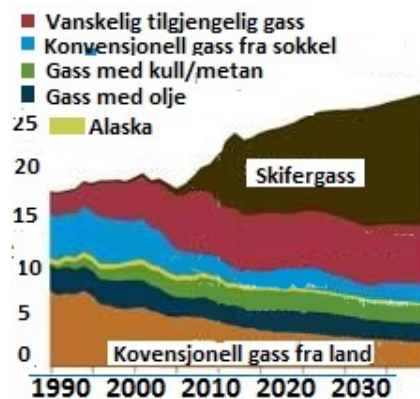
Sett fra et miljøsynspunkt er naturgass et meget godt brennstoff, ettersom svovelinnholdet er svært lavt. I tillegg oppnår man vanligvis nesten fullstendig forbrenning ved å bruke naturgass som brensel. Utslippene ved forbrenning er for det meste CO₂ og vanndamp. Energiutbyttet er høyt i forhold til utslipp av karbondioksid, sammenlignet med olje og kull. Både naturgassen i seg selv og CO₂ regnes imidlertid som klimagasser (Lundberg, 2012). Av de fossile brennstoffene er det naturgass som blir sett på som den energien som en skal bruke mer av i fremtiden. Dette

er fordi det er den mest miljøvennlige energikilden av fossile brennstoff, og fordi oljeproduksjon går nedover, mens gassproduksjonen er fremdeles i vekst.

Enkelte politiske avgjørelser spesielt i Vest-Europa er tatt for å begrense bruken av naturgass. En oppfatning av knapphet av naturgass, og at de ønsker sikkerhet av forsyninger over lengre perioder har ført til en negativ holdning for en utvidelse av gassmarkedet i Vest-Europa. For eksempel har det blitt forbud å bruke gass til elektrisitet eller som drivstoff i enkelte land. I tillegg er det en del beskyttelse av alternativ energi, som har både offentlige og private grunner. Dermed kan slike forbud begrense tilbud og etterspørsel av gass (Honore, 2010).

1.3.5 Teknologisk utvikling

I midten av 1990-tallet, ble det økt etterspørsel etter naturgass i kraftverkindustrien på grunn av ny teknologi. Den høy effektive CCGT (Combined Cycle Gas Turbine)² med lave utslipp og kostnader ble populær i kraftsektoren. Siden året 1990 har det vært en signifikant økning i etterspørsel etter naturgass, det er på grunn av at flere kullkraftverk er erstattet med gasskraftverk. Flere gasskraftverk gir økt etterspørsel etter gass, og dermed mulighet for økte gasspriser (Honore, 2010).



Figur 1.6: Gassproduksjon og forventet produksjon i USA, 1990 til 2035. Målt i trillioner kubikkfot (NUPI, 2015)

I løpet av de siste ti årene, har det blitt mer globalisering av gassmarkedet. Grunnen er at det har blitt økt tilgang på flytende gass som LNG. LNG kan transporteres over lengre avstander via skip. Det er den nye teknologien og de reduserte kostnader ved å kjøle gass ned til flytende form som LNG, som har ført til denne globaliseringen (Winje, 2011). Dette har ført til at gass har nådd nye markeder. Antall land som importerer LNG har tredoblet seg de siste ti årene, og

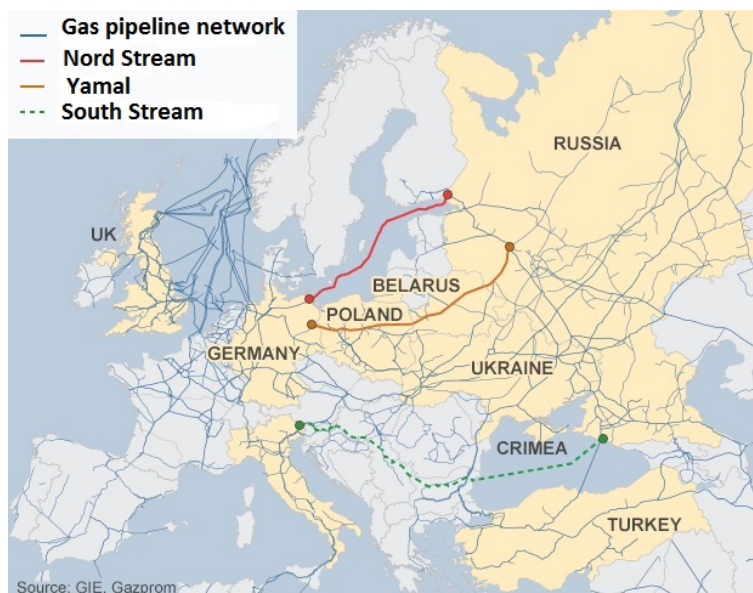
²Combined Cycle Gas Turbine (CCGT) er en form for svært effektiv energiproduksjon som kombinerer ett gasskraftverk turbin med en dampturbin. Gassturbinen skaper elektrisitet og spillvarmen lager damp, som i sin tur driver en dampturbin. Dette øker vesentlig systemets effektivitet uten noen økning i brensel (BusinessGreen, 2013)

det medfører et økt gasstilbud. Et annet forhold som har preget veksten av gassmarkedet og økt tilbudet er produksjon av skifergass. Skifergass er en ukonvensjonell gass. Denne skifergassproduksjonen har økt kraftig i USA siden 1990-tallet, og det kommer blant annet av at ny teknologi har gitt reduserte utvinningkostnader. I figur 1.6 ser vi hvordan skifergassproduksjonen i USA har økt, og hvor kraftig den forventes å øke.

Denne ukonvensjonell gassen har økt verdens gassreserver betydelig. I tillegg har høy effektive CCGT gjort at naturgass som brensel er mer effektiv, og veksten i tilbudet av LNG har gjort naturgass tilgjengelig for nye markeder. Til sammen har det styrket naturgassens posisjon som en mer sikker energikilde. Samtidig som ny teknologi fører til utvikling og større gassmarked, kan energieffektiviseringen også føre til lavere vekst i gassetterspørselen, spesielt innenfor husholdninger og industri (Oljeogenergidepartementet, 2011).

1.3.6 Transport, distribusjon og reserver av naturgass

Et rørledningsett er veldig dyrt å bygge, slik at når gass skal transporteres over lengre avstander må den først omgjøres til LNG. Koepunktet til LNG ved 1 atm er -161 grader, som gjør at det er dyrt å bygge skip som kan frakte LNG. Skipene må holde en temperatur på omkring -161 grader for å unngå overtrykk. Ved produksjon av råolje spesielt i midtøsten blir store mengder naturgass brent rett opp i atmosfæren. Naturgassen blir brent opp fordi det er høye transportkostnader knyttet til å frakte gassen. I Europa og i USA derimot er det bygget opp et stort nettverk av rørledninger. Fra Nordsjøen går det underjordiske rørledninger til både kontinentet og Storbritannia (Lundberg et al., 2009). I figur 1.7 er det en oversikt over det europeiske rørledningsnett.



Figur 1.7: Rørledningsnett for naturgass (BBC, 2015)

For å tilpasse etterspørselen etter naturgass og for å sikre forsyninger bygges det flere gasslager. I Europa har store gasskjøpere etablert flere lager. De største lagrene er ofte ferdigproduserte gassreservoarer i berggrunnen som blir fylt opp på nytt, men også andre porøse bergartslag kan fylles opp selv om de aldri har inneholdt gass. Store fjellhaller som er vasket ut i steinaltformasjoner ved saltproduksjon brukes også. Naturgass kan også lagres som LNG (Lundberg et al., 2009).

De største gassforekomster i verden ligger i Eurasia og i Midtøsten, aller mest i Russland, Iran og Qator. Disse tre landa har over halvparten av verdens kjente gassreserver (Lundberg et al., 2009). Det er ikke enighet om hvor store verdens naturgassreserver er, og det blir påstått alt fra 60 år til 220 år. Miljø-økonomene skriver 8. mars 2012 at naturgassreservene varer i ca 100 til 220 år til avhengig om vi regner med skifergassteknologien eller ikke (Gudding, 2012, s. 215).

KAPITTEL 2

Tidsrekker

En tidsrekke $\{X_t, t \in \mathbb{Z}\}$ er en rekke med historiske observasjoner av en variabel, hvor hver X_t er gitt ved en spesiell tid t . Naturgasspriser over en tidsperiode er et eksempel på en tidsrekke. Senere i oppgaven skal vi gjøre en tidsrekke analyse på de europeiske naturgassprisene. I dette kapitlet skal vi introdusere viktig teori og verktøy som vi bruker for å analysere en tidsrekke. Først i dette kapitlet skal vi se på stasjonære tidsrekker. Vi skal gjøre rede på ulike eksempler av stasjonære prosesser. Følgende skal vi betrakte tilfeldig gang. Vi skal studere ulike enhetsrot-tester som kan teste for stasjonærhet i tidsrekken. Vi skal introdusere avkastning. Deretter skal vi studere autokorrelasjonsfunksjonen (ACF) og hvordan et ACF-plott kan brukes for å tilpasse en modell. Videre i dette kapitlet skal vi se på Autoregressive Moving Average (ARMA) modeller. Til slutt skal vi iaktta kointegrasjon. Kointegrasjon kan være en nyttig for å beskrive den langsiktige sammenhengen mellom de ulike gassmarkedene.

2.1 Stasjonære tidsrekker

En tidsrekke $\{X_t, t \in \mathbb{Z}\}$ er stasjonær hvis den har like statistiske egenskaper som tidsrekken $\{X_{t+h}, t \in \mathbb{Z}\}$ for et hvert heltall h (Brockwell and Davis, 1996, s. 14).

Forventningsfunksjonen til $\{X_t\}$ er gitt som (Brockwell and Davis, 1996, s. 14):

$$\mu_X(t) = E(X_t),$$

når vi antar at $E(X_t^2) < \infty$. Videre har vi at kovariansfunksjonen til $\{X_t\}$ gitt som:

$$\gamma_x(r, s) = \text{Cov}(X_r, X_s) = E[(X_r - \mu_X(r))(X_s - \mu_X(s))]$$

for alle heltall r og s .

En tidsrekke $\{X_t\}$ er svak stasjonær dersom forventning, varians og kovarians ikke er endret ved tiden. Den presise definisjonen er gitt under.

Definisjon 2.1.1. (Brockwell and Davis, 1996, s. 15)

Tidsrekken $\{X_t\}$ er svak stasjonær hvis

- (1) $\mu_X(t)$ er uavhengig av t og
- (2) $\gamma_X(t+h, t)$ er uavhengig av t for hver h .

En tidsrekke $\{X_t\}$ er strengt stasjonær hvis $\{X_t\}$ i tillegg til egenskapene fra svak stasjonærhet også har andre kjennetegn som kvantiler, skjevhet og kurtosis som er uendret ved tiden. Den presise definisjonen er gitt under.

Definisjon 2.1.2. (Brockwell and Davis, 1996, s. 47)

Tidsrekken $\{X_t\}$ er strengt stasjonær hvis (X_1, \dots, X_n) og $(X_{(1+h)}, \dots, X_{(n+h)})$ har samme simultanfordeling for enhver $n \in \mathbb{N}$ og $h \in \mathbb{Z}$.

Eksempel om i.i.d. støy

Hvis tidsrekken $\{X_t\}$ er i.i.d. (uavhengig og identisk fordelt) støy og $E(X_t^2) < \infty$. Da er $\{X_t\}$ stasjonær fordi krav (1) i definisjon (2.1) er oppfylt, siden $E(X_t) = 0$ for alle t . Videre har vi at

$$\gamma_X(t+h, h) = \begin{cases} \sigma^2, & h = 0 \\ 0, & h \neq 0, \end{cases}$$

som ikke avhenger av t , dermed er krav (2) i definisjon (2.1) også oppfylt. Når $\{X_t\}$ innehar egenskapen i.i.d. støy bruker vi notasjonen $\{X_t\} \sim IID(0, \sigma^2)$ (Brockwell and Davis, 1996, s. 15).

Eksempel om hvit støy

Hvit støy (White Noise (WN)) er det enkleste eksempelet av en stasjonær tidsrekke. Hvis $\{X_t\}$ er en sekvens av ikke-korrelerte tilfeldige variabler, der hver sekvens har forventning null og varians σ^2 , da er det klart at $\{X_t\}$ har samme kovariansfunksjon som ved i.i.d. støy. Når $\{X_t\}$ innehar egenskapen hvit støy bruker vi notasjonen $\{X_t\} \sim WN(0, \sigma^2)$. Hver $IID(0, \sigma^2)$ er $WN(0, \sigma^2)$, men dette gjelder nødvendigvis ikke omvendt (Brockwell and Davis, 1996, s. 16).

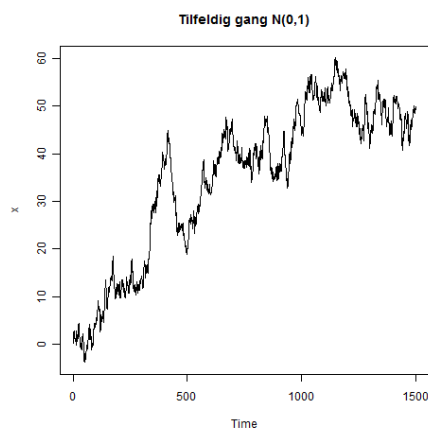
2.2 Tilfeldig gang

Proessen $\{S_t, t \geq 0\}$ kalles for tilfeldig gang og er gitt som (Brockwell and Davis, 1996, s. 8):

$$S_t = \begin{cases} S_0, & t = 0 \\ S_0 + Z_1 + \dots + Z_t, & t \geq 1. \end{cases}$$

Der S_0 er startpunktet, og Z_1, Z_2, \dots er stegene i prosessen. Z_1, Z_2, \dots er uavhengig og identisk fordelt (i.i.d.) med forventning μ og standardavvik σ . Hvis stegene er normalfordelt kalles prosessen for normal tilfeldig gang. Forventningen og variansen til S_t gitt S_0 er $E[S_t|S_0] = S_0 + \mu t$

og $\text{VAR}[S_t|S_0] = \sigma^2 t$. Parameteren μ kalles for drift og bestemmer den generelle retningen av den tilfeldige gangen. Parameteren σ er volatiliteten og bestemmer hvor mye den tilfeldige gangen svinger om den betingede forventningen. I figur 2.1 ser vi ett eksempel på en tilfeldig gang prosess, der tidsrekken er normalfordelt med forventning 0 og varians 1.



Figur 2.1: Tilfeldig gang med $n = 1500$, $\mu = 0$ og $\sigma = 1$.

2.3 Enhetsrot test

I mange modelltilpasninger er det et krav om at tidsrekken er stasjonær. Derfor er det nødvendig i en analyse å teste om tidsrekken er stasjonær. Vi kan teste for stasjonærhet ved å bruke en enhetsrotstest.

Dickey-Fuller test

En test som sjekker om vi har enhetsrøtter i tidsrekken er Dickey-Fuller test. Dickey-Fuller kan teste tre ulike modeller. Der vi har en modell med tilfeldig gang, en modell med tilfeldig gang og drift, og en modell med tilfeldig gang, drift og en lineær trend. De tre følgende modellene er gitt under:

$$\begin{aligned} X_t &= \varphi_1 X_{t-1} + \epsilon_t \\ X_t &= \varphi_0 + \varphi_1 X_{t-1} + \epsilon_t, \\ X_t &= \varphi_0 + \varphi_1 X_{t-1} + \varphi_2 t + \epsilon_t, \end{aligned} \tag{2.1}$$

der ϵ_t betegner feilledet. Nullhypotesen i Dickey-Fuller testen er $H_0 : \varphi_1 = 1$ mot den alternative hypotesen $H_1 : \varphi_1 < 1$. Testobservatoren er t-forholdet av minste kvadraters (LS) estimatet av

φ_1 . For ligning (2.1) som er modellene uten drift, gir LS metode (Tsay, 2005, 68-69):

$$\hat{\varphi}_1 = \frac{\sum_{t=1}^T X_t X_{t-1}}{\sum_{t=1}^T X_{t-1}^2}, \quad \hat{\sigma}^2 = \frac{\sum_{t=1}^T (X_t - \hat{\varphi}_1 X_{t-1})^2}{T-1},$$

der $x_0 = 0$ og T er størrelsen på utvalget. Dickey-Fuller testen er gitt som (Tsay, 2005, 68-69):

$$DF = \frac{\hat{\varphi}_1 - 1}{std(\hat{\varphi}_1)} = \frac{\sum_{t=1}^T X_{t-1} \epsilon_t}{\hat{\sigma} \sqrt{\sum_{t=1}^T X_{t-1}^2}},$$

Augmented Dickey-Fuller test

Ofte er modellen litt mer komplisert enn ligning (2.1). For eksempel for å teste om det eksisterer en enhetsrot i en $AR(p)$ -modell, må vi teste nullhypotesen, $H_0 : \beta = 1$ mot $H_1 : \beta < 1$ ved å bruke regresjonen (Tsay, 2005, 68-69):

$$X_t = c_t + \beta X_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta X_{t-1} + \epsilon_t, \quad (2.2)$$

der $\{X_t\}$ er observasjonen fra en $AR(p)$ -modell og ϵ_t er feilleddet. c_t er en deterministisk funksjon av tiden og $\Delta X_j = X_j - X_{j-1}$ er den differensierte tidsrekken til X_t . I praksis kan ligning (2.2) gi oss tre regresjonsmodeller. Den ene er hvis c_t er null, da har vi en modell med verken en konstant eller en trend. Derimot hvis c_t er en konstant har vi en modell med en konstant men ingen trend. Det tredje alternativet er at $c_t = \omega_0 + \omega_1 t$, da har vi en modell som har både en konstant og en trend (Tsay, 2005, s. 69).

Ved å trekke fra X_{t-1} på begge sider av likhetstegnet i ligning (2.2), kan ligningen skrives som:

$$\Delta X_t = c_t + \beta_c X_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta X_{t-1} + \epsilon_t,$$

der $\beta_c = 1 - \beta$. Da tester vi nullhypotesen, $H_0 = \beta_c = 0$ at det er en enhetsrot, mot den alternative hypotesen $H_1 = \beta_c < 0$ at modellen er stasjonær.

Augmented Dickey-Fuller testobservator er:

$$ADF = \frac{\hat{\beta}_c}{std(\hat{\beta}_c)},$$

der $\hat{\beta}_c$ er den minste kvadraters estimator til β_c , $std(\hat{\beta}_c) = S^2 \sum_{t=2}^n 1 / (X_{t-1} - \bar{X})^2$. S^2 er i dette tilfellet definert som $S^2 = \sum_{t=2}^n (\Delta X_t - \hat{c}_t - \hat{\beta}_c X_{t-1})^2 / (n - 3)$ og \bar{X} er gjennomsnittet.

Phillips-Perron

Phillips-Perron test er en ikke-parametrisk testobservator, der nullhypotesen er at det inneholder en enhetsrot og ϵ kan være heteroskedastisk. Phillips-Perron estimerer ikke ligning (2.2), men (Pfaff, 2008, s. 95):

$$X_t = \mu + \alpha X_{t-1} + \epsilon_t. \quad (2.3)$$

Test observatorene til ligning (2.3) er gitt som (Pfaff, 2008, s. 95):

$$\begin{aligned} Z(\hat{\alpha}) &= T(\hat{\alpha} - 1) - \frac{\hat{\lambda}}{\bar{m}_{xx}} \\ Z(\tau_{\hat{\alpha}}) &= \left(\frac{\hat{s}}{\hat{\sigma}_{Tl}} \right) t_{\hat{\alpha}} - \frac{\hat{\lambda}' \hat{\sigma}_{Tl}}{\bar{m}_{xx}^{0.5}} \\ Z(\tau_{\hat{\mu}}) &= \left(\frac{\hat{s}}{\hat{\sigma}_{Tl}} \right) t_{\hat{\mu}} - \frac{\hat{\lambda}' \hat{\sigma}_{Tl} m_x}{\bar{m}_{xx}^{0.5} m_{xx}^{0.5}}. \end{aligned}$$

Vi har at $\bar{m}_{xx} = T^{-2} \sum (x_t - \bar{x})^2$, $m_{xx} = T^{-2} \sum x_t^2$, $m_x = T^{-3/2} \sum x_t$, $\lambda = 0, 5(\hat{\sigma}_{Tl}^2 - \hat{s}^2)$, der \hat{s}^2 er utvalgs variansen for $\hat{\lambda}' = \hat{\lambda}/\hat{\sigma}_{Tl}$ og $t_{\hat{\alpha}}$ og $t_{\hat{\mu}}$ er t forholdet til $\hat{\alpha}$ og $\hat{\mu}$. Variansen i det lange løp er estimert som (Pfaff, 2008, s. 95):

$$\hat{\sigma}_{Tl} = T^{-1} \sum_{t=1}^T \hat{\epsilon}_t^2 + 2T^{-1} \sum_{s=1}^T w_{st} \sum_{t=s+1}^T \hat{\epsilon}_t \hat{\epsilon}_{t-1},$$

der $w_{st} = 1 - \frac{s}{l+1}$.

2.4 Transformere dataene

De fleste finans og økonometriske tidsrekker er ikke stasjonære, men differansen eller log-differansen i disse er ofte stasjonære (Ruppert, 2011, s. 8-10).

Endringene, differansen eller avkastningene i prisene fra en periode til en annen periode er gitt som:

$$D_t = P_t - P_{t-1},$$

der P_t er naturgassprisen gitt ved tid t .

Hvis differansen til prisene ikke er stasjonær, er det nødvendig å omforme til log-differansen som også er kalt log-avkastningen eller bare avkastningen. For å finne log-avkastningen tar vi først den naturlige logaritmen av prisene, deretter tar vi differansen. Vi bruker X som symbol på log-avkastningene.

$$X_t = \log P_t - \log P_{t-1} = \log \left(\frac{P_t}{P_{t-1}} \right), \quad (2.4)$$

der (P_t/P_{t-1}) er bruttoavkastningen.

I tillegg til at log-avkastninger ofte er stasjonære er en annen fordel at det blir å lettere regne med multiperiodiske avkastninger (Ruppert, 2011, s. 8-10). Da er log-avkastningen over k perioder bare summen av de enkelte periodene av log-avkastningene, istedenfor produktet som en måtte hatt ved bruttoavkastning. Vi ser at log-avkastningene ved periode k er:

$$\log\left(\frac{P_t}{P_{t-1}} \frac{P_t}{P_{t-2}} \cdots \frac{P_t}{P_{t-k}}\right) = \log\left(\frac{P_t}{P_{t-1}}\right) + \cdots + \log\left(\frac{P_t}{P_{t-k}}\right) = X_t + X_{t-1} + \cdots + X_{t-k+1},$$

2.5 Autokorrelasjonsfunksjon

Autokorrelasjonsfunksjonen (ACF) til en tidsrekke $\{X_t\}$ er gitt som (Brockwell and Davis, 1996, s. 15):

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \text{CORR}(X_{t+h}, X_t). \quad (2.5)$$

Der $\gamma(h)$ er autokovariansfunksjonen (ACVF) til tidsrekken $\{X_t\}$. ACVF er gitt som (Brockwell and Davis, 1996, s. 15):

$$\gamma(h) = \text{COV}(X_{t+h}, X_t),$$

der $\gamma(h) := \gamma(t+h, t)$.

Når vi derimot ikke har en modell, men starter med de observerte dataene $\{X_1, X_2, \dots, X_n\}$ må vi finne avhengighetsgraden i disse dataene og en passende modell. Dette kan gjøres ved å bruke utvalgs autokorrelasjonsfunksjon til dataene. Hvis vi tror at dataene er verdier i en stasjonær tidsrekke $\{X_t\}$, da vil utvalgs ACF'en gi oss den empiriske ACF'en til $\{X_t\}$.

Vi lar $\{X_1, X_2, \dots, X_n\}$ være observasjonene i tidsrekken. Forventningen til utvalget $\{X_1, X_2, \dots, X_n\}$ er gitt som (Brockwell and Davis, 1996, s. 17):

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Videre er utvalgs autokovariansfunksjonen gitt som (Brockwell and Davis, 1996, s. 17):

$$\hat{\gamma}(h) := \frac{1}{n} \sum_{j=1}^{n-|h|} (X_j - \bar{X})(X_{j+|h|} - \bar{X}), \quad -n < h < n.$$

Selve utvalgs autokorrelasjonsfunksjonen er gitt som (Brockwell and Davis, 1996, s. 17):

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}.$$

2.6 Lineære prosesser

En tidsrekke $\{X_t\}$ er en lineær prosess hvis den kan skrives som (Brockwell and Davis, 1996, s. 17):

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j},$$

for alle t , der $\{Z_t\} \sim WN(0, \sigma^2)$ og $\{\psi_j\}$ er en sekvens av konstanter med betingelsen $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$.

For å gjennomføre en tidsrekke analyse må vi konstruere en modell for den underliggende stokastiske prosessen. For eksempel hvis $\{X_t\}$ er en stasjonær q -korrelert tidsrekke med forventning null, kan den bli representert som en moving average (MA(q)) prosess (Brockwell and Davis, 1996, s. 48).

2.6.1 MA(q)-modell

En MA(q) modell er en lineær prosess av orden q . Modellen kan defineres slik (Brockwell and Davis, 1996, s. 48):

$$X_t = Z_t + \sum_{i=1}^q \theta_i Z_{t-i}, \quad t = 0, \pm 1, \dots,$$

der $\{Z_t\} \sim WN(0, \sigma^2)$ og $\theta_1, \dots, \theta_q$ er konstanter.

2.6.2 MA(1)-modell

Hvis $q = 1$ i modellen over har vi en første ordens moving average modell MA(1). En MA(1) modell er definert som:

$$X_t = Z_t + \theta Z_{t-1}, \quad t = 0, \pm 1, \dots,$$

der $\{Z_t\} \sim WN(0, \sigma^2)$ og θ er en konstant.

Forventningen til MA(1) prosessen er null siden $E(Z_t) = 0$. Variansen til MA(1) prosessen er:

$$\text{VAR}(Z_t + \theta Z_{t-1}) = \sigma^2 + \theta^2 \sigma^2 = \sigma^2(1 + \theta^2)$$

.

Kovariansen til MA(1) prosessen når $h = \pm 1$ er gitt som:

$$\text{COV}(X_t, X_{t+h}) = \text{COV}(Z_t + \theta Z_{t-1}, Z_{t+1} + \theta Z_t) = \theta \sigma^2,$$

mens for $h = 0$ får en bare variansen til X_t og for $h > 1$ er det ingen korrelasjon.

Generelt har en dermed at kovariansen er:

$$\text{Cov}(X_t, X_{t+h}) = \gamma(t+h, t) = \begin{cases} \sigma^2(1 + \theta^2), & \text{hvis } h = 0, \\ \sigma^2\theta, & \text{hvis } h = \pm 1, \\ 0, & \text{hvis } |h| > 1 \end{cases}. \quad (2.6)$$

Som vi ser er $E(X_t) = 0$ for alle t og dermed uavhengig av t , og $\gamma(t+h, t)$ er uavhengig av t for hver h . Dette medfører at $\{X_t\}$ er svak stasjonær.

Hvis vi setter resultatene fra ligning (2.6) inn i ligning (2.5) er autokorrelasjonfunksjonen for $\{X_t\}$ gitt som:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \begin{cases} 1, & \text{hvis } h = 0 \\ \frac{\theta}{1 - \theta^2}, & \text{hvis } h = \pm 1, \\ 0, & \text{hvis } |h| > 1 \end{cases}$$

2.6.3 AR(p)-modell

En autoregressive (AR(p))-modell av orden p er gitt som:

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + Z_t, \quad t = 0, \pm 1, \dots, \quad (2.7)$$

der $\{X_t\}$ er en stasjonær tidsrekke og $\{Z_t\} \sim WN(0, \sigma^2)$. Parametrene ϕ_1, \dots, ϕ_p representerer korrelasjonen mellom observasjonene. I AR(p)-modellen antar vi også at Z_t ikke er korrelert med X_s når $s < t$, modellen er da kausal.

2.6.4 AR(1)-modell

Hvis $p = 1$ i modellen over har vi en første ordens autoregressiv-modell (AR(1)). En AR(1) modell er definert som (Brockwell and Davis, 1996, s. 17):

$$X_t = Z_t + \phi X_{t-1}, \quad t = 0, \pm 1, \dots,$$

der $\{Z_t\} \sim WN(0, \sigma^2)$ og Z_t er ikke korrelert med X_s for hver $s < t$. Vi har at ϕ er korrelasjonen mellom observasjonene, og den er mellom -1 og 1 . Hvis $\phi = 0$ er modellen bare basert på hvit støy, $\{Z_t\}$. Det vil da si at til nærmere ϕ er 0 , til mer tilfeldig er modellen. Derimot hvis $\phi > 1$ har vi ikke en stasjonær prosess og heller ikke en AR prosess.

Forventningen til AR(1) prosessen er $E(X_t) = 0$. Kovariansen til AR(1) prosessen er:

$$\begin{aligned}\gamma(h) &= \text{Cov}(X_t, X_{t-h}) \\ &= \text{Cov}(\phi X_{t-1}, X_{t-h}) + \text{Cov}(Z_t, X_{t-h}) \\ &= \phi \text{Cov}(X_{t+h-1}, X_t) + 0 \\ &= \phi \gamma_X(h-1) \\ &= \phi^h \gamma_X(0), \quad h > 0.\end{aligned}$$

Videre har vi at $\gamma(h) = \gamma(-h)$ som medfører at $\gamma(-h) = \phi^{|h|} \gamma_X(0)$. Vi antar en stasjonær prosess, slik at $|\phi| < 1$. Under finner vi kovariansen når $h = 0$:

$$\begin{aligned}\gamma_X(0) &= \text{VAR}(X_t) \\ &= \text{VAR}(\phi X_{t-1} + Z_t) \\ &= \phi^2 \text{VAR}(X_{t-1}) + \text{VAR}(Z_t) + 0 \\ &= \phi^2 \gamma_X(0) + \sigma^2.\end{aligned}$$

Dette medfører at $\gamma(0) = \sigma^2 / (1 - \phi^2)$. Videre setter vi disse kovariansene inn i ligning (2.5), og vi får at autokorrelasjonfunksjonen er:

$$\rho_X = \frac{\gamma_X(h)}{\gamma_X(0)} = \frac{\phi^{|h|} \gamma_X(0)}{\gamma_X(0)} = \phi^{|h|}.$$

2.6.5 ARMA(1, 1)-modell

En ARMA(1, 1) (autoregressive moving average) prosess er en kombinasjon av et AR(1)-ledd og et MA(1)-ledd. En tidsrekke $\{X_t\}$ er en ARMA(1, 1) prosess hvis den er svak stasjonær og tilfredsstillende ligningen (Brockwell and Davis, 1996, s. 53):

$$X_t - \phi X_{t-1} = Z_t + \theta Z_{t-1} + \mu, \quad (2.8)$$

for hver t , der $\{Z_t\} \sim \text{WN}(0, \sigma^2)$ og $\phi + \theta \neq 0$. Størrelsen μ er trenden og kan settes lik null.

For enkelthetsskyld kan man anta at B er en "bakoverskift" operator. Det vil si at $B^j X_t = X_{t-j}$. Da kan en skrive ARMA(1, 1)-modell som

$$\phi(B)X_t = \theta(B)Z_t + \mu,$$

der $\phi(B)$ og $\theta(B)$ er henholdsvis de lineære filtrene $\phi(B) = (1 - \phi B)$ og $\theta(B) = (1 + \theta B)$.

2.6.6 ARMA(p, q)-modell

Mer generelt har vi at tidsrekken $\{X_t\}$ er en ARMA(p, q) prosess hvis den er svak stasjonær og tilfredsstillende ligningen (Brockwell and Davis, 1996, s. 81):

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q}, \quad (2.9)$$

for hver t , der $\{Z_t\} \sim \text{WN}(0, \sigma^2)$. Polynomene $\phi(z) = (1 - \phi_1 z - \dots - \phi_p z^p)$ og $\theta(z) = (1 + \theta_1 z + \dots + \theta_q z^q)$ har ingen felles røtter.

Et krav i ligning (2.9) er at $\{X_t\}$ er stasjonær. En stasjonær løsning $\{X_t\}$ av ligning (2.9) eksisterer (og er unik) hvis og bare hvis $\phi(z) = (1 - \phi_1 z - \dots - \phi_p z^p) \neq 0$ for alle $|z| = 1$ (Brockwell and Davis, 1996, s. 83).

I tillegg er en ARMA(p, q) prosess $\{X_t\}$ kausal, eller en kausal funksjon av $\{Z_t\}$, hvis det eksisterer en konstant slik at $\sum_{j=0}^{\infty} |\psi_j| < \infty$ og

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}.$$

Kausalitet er ekvivalent til betingelsen $\phi(z) = (1 - \phi_1 z - \dots - \phi_p z^p) \neq 0$ for alle $|z| \leq 1$ (Brockwell and Davis, 1996, s. 83).

I en ARMA-modell tar en ikke hensyn til tidsvarierende volatilitet. Hvis en i tillegg skal ta med tidsvariende volatilitet i modellen, må en bruke GARCH (generalized autoregressive conditional heteroscedasticity)-modeller som vi skal se på i neste kapittel.

2.7 Kointegrasjon

Kointegrasjon er interessant fordi det er mulig å oppdage stabile relasjoner mellom ikke-stasjonære tidsrekker. Gassprisene i seg selv er ikke-stasjonære tidsrekker. Et krav for å bruke kointegrasjon er at prosessene er $I(1)$. En prosess er $I(1)$ dersom $\Delta X_t = X_t - X_{t-1}$ er stasjonær. Videre er en prosess $I(d)$ dersom en prosess må differensieres d ganger før den blir stasjonær. Vi har kointegrasjon hvis forholdet mellom to $I(1)$ prosesser er stasjonært.

Engle og Granger (1987) foreslo en to-trinns estimeringsteknikk for å estimere en kointegrert vektor α (Pfaff, 2008, s. 76). Denne teknikken blir kalt for Engle-Granger to-steps metode.

2.7.1 Engle-Granger to-steps metode

Anta at vi har to tidsrekker $\{X_t\}$ og $\{Y_t\}$, der begge er $I(1)$. Det første steget er en klassisk lineær regresjonsanalyse av $I(1)$ variablene,

$$Y_t = \alpha_1 X_t + z_t, \quad (2.10)$$

Denne ligningen kalles for en kointegrasjonsligning, der z_t er restleddet og α_1 er den kointegrerte parameteren (Pfaff, 2008, s. 76). Restleddet er avviket fra langtidslikevekten. Vi kan estimere ligning (2.10) ved en klassisk lineær regresjonsanalyse. Residualene fra den statistiske regresjon (2.10) er \hat{z}_t . Hvis vi har kointegrasjon skal disse residualene være stasjonære. Hvorvidt disse residualene er stasjonære kan testes. Vanlige enhetsrotstester er Dickey-Fuller(DF) test eller augmented Dickey-Fuller(ADM) test (Pfaff, 2008, s. 76). Hvis residualene er stasjonære, følger det neste trinnet i to-steps prosedyren (Pfaff, 2008, s. 76).

I det steget er en feilkorreksjon modell (ECM) spesifisert fra Engle-Granger representasjonsteorem i 1987 (Pfaff, 2008, s. 77). En ECM modell av (2.10) er:

$$\Delta Y_t = \gamma_1 \hat{z}_{t-1} + \psi_1 \Delta X_{t-1} + \psi_2 \Delta Y_{t-1} + \epsilon_t, \quad (2.11)$$

der \hat{z}_t er residualene fra regresjonen i ligning (2.10), og ϵ_t er signifikant hvit støy prosess. ECM i ligning (2.11) sier at endringen i Y_t er forklart ved sin egen historie, laggede forandringer av X_t , og feilleddet i det lange løp fra den forrige perioden. Verdien av γ_1 bør alltid være negativ, ellers vil systemet avvike fra sin likevekt. Ut ifra ligning (2.11) og fra den statistiske regresjon i det tilfellet vi har to kointegrerte $I(1)$ variabler kan det konkluderes at Granger kausalitet må eksistere i minst en retning. Det betyr at ved å bruke en variabel kan en forutsi den andre variabelen (Pfaff, 2008, s. 77).

2.7.2 Johansen-metoden

En annen måte å teste om vi har kointegrasjon er å bruke Johansen sin metode. Johansen-metoden er basert på en vektor autoregression (VAR)-modell.

Først antar vi en VAR(1)-modell, uten trend. VAR(1)-modellen er gitt som:

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = A \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ \nu_t \end{pmatrix}, \quad (2.12)$$

der ϵ_t og ν_t er restleddenes som er hvit støy. VAR(1)-modellen kan omskrives til en vektor error-correction modell (VECM):

$$\begin{pmatrix} X_t - X_{t-1} \\ Y_t - Y_{t-1} \end{pmatrix} = (A - I) \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ \nu_t \end{pmatrix} \quad (2.13)$$

En generell k -dimensjonal VAR(p)-modell er gitt som (Tsay, 2005, s. 380):

$$\mathbf{X}_t = \boldsymbol{\mu}_t + \boldsymbol{\Phi}_1 \mathbf{X}_{t-1} + \cdots + \boldsymbol{\Phi}_p \mathbf{X}_{t-p} + \boldsymbol{\epsilon}_t, \quad (2.14)$$

der vi setter at $\boldsymbol{\epsilon}_t$ er hvit støy, og $\boldsymbol{\mu}_t$ er trenden og er en $(k \times 1)$ vektor. $\boldsymbol{\Phi}_1, \dots, \boldsymbol{\Phi}_p$ er $(k \times k)$ matriser.

VAR(p)-modellen kan omskrives til en VECM modell.

Generelt er en VECM modell gitt som (Tsay, 2005, s 381):

$$\Delta \mathbf{X}_t = \boldsymbol{\mu}_t + \boldsymbol{\Pi} \mathbf{X}_{t-1} + \boldsymbol{\Phi}_1^* \Delta \mathbf{X}_{t-1} + \cdots + \boldsymbol{\Phi}_{p-1}^* \Delta \mathbf{X}_{t-p-1} + \boldsymbol{\epsilon}_t, \quad (2.15)$$

der $\boldsymbol{\Phi}_j^* = - \sum_{i=j+1}^p \boldsymbol{\Phi}_i$ og $\boldsymbol{\Pi} = \alpha \beta'$. Leddet i ligning (2.15) som er mest interessant i forhold til å studere kointegrasjon er $\boldsymbol{\Pi} \mathbf{X}_{t-1}$.

Det er tre tilfeller i ligning (2.13) (2.15) som er interessant å finne ut av (Tsay, 2005, s 381):

- 1: (2.13): Rang ($A - I$) = 0 og i (2.15): Rang($\boldsymbol{\Pi}$) = 0. I dette tilfellet eksisterer det ingen lineær kombinasjon som gjør at $(\boldsymbol{\Pi}) \mathbf{X}_t$ er stasjonær utenom den trivielle løsningen. Dette medfører at $\boldsymbol{\Pi} = 0$ og tidsrekkene \mathbf{X}_t er ikke kointegrert. Ligning (2.15) blir da redusert til

$$\Delta \mathbf{X}_t = \boldsymbol{\mu}_t + \boldsymbol{\Phi}_1^* \Delta \mathbf{X}_{t-1} + \cdots + \boldsymbol{\Phi}_{p-1}^* \Delta \mathbf{X}_{t-p-1} + \boldsymbol{\epsilon}_t,$$

slik at $\Delta \mathbf{X}_t$ følger en VAR($p - 1$) modell med en deterministisk trend $\boldsymbol{\mu}_t$. Ligning (2.13) blir redusert til:

$$\begin{pmatrix} X_t - X_{t-1} \\ Y_t - Y_{t-1} \end{pmatrix} = \begin{pmatrix} \epsilon_t \\ \nu_t \end{pmatrix}$$

- 2: Rang($\boldsymbol{\Pi}$) = k , i ligning (2.13) er $k = 2$. Dette medfører at $|\boldsymbol{\Pi}| \neq 0$ og alle tidsrekkene i \mathbf{X}_t er $I(0)$. Da er kointegrasjon ikke relevant og det ingen poeng i å studere ECM modellen. Man kan heller studere tidsrekkene i \mathbf{X}_t direkte.
- 3: $0 < \text{Rang}(\boldsymbol{\Pi}) = r < k$, ved to tidsrekker har vi at Rang ($\boldsymbol{\Pi}$) = 1. I dette tilfellet kan vi skrive $\boldsymbol{\Pi} = \alpha \beta'$, der α og β er $k \times r$ matriser med Rang(α) = Rang(β) = k . Ligning (2.15) blir da:

$$\Delta \mathbf{X}_t = \boldsymbol{\mu}_t + \alpha \beta' \mathbf{X}_{t-1} + \boldsymbol{\Phi}_1^* \Delta \mathbf{X}_{t-1} + \cdots + \boldsymbol{\Phi}_{p-1}^* \Delta \mathbf{X}_{t-p-1} + \boldsymbol{\epsilon}_t.$$

Dette betyr at tidsrekkene i \mathbf{X}_t er kointegrert med r lineære uavhengige kointegrerte vektorer, $\omega = \beta' \mathbf{x}_t$ og har $k - r$ egenrøtter som gir at X_t har $k - r$ felles stokastiske trender. Ved $k = 2$ får vi:

$$\begin{pmatrix} X_t - X_{t-1} \\ Y_t - Y_{t-1} \end{pmatrix} = \alpha \beta' \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ \nu_t \end{pmatrix}$$

Det betyr at X og Y er to kointegrerte tidsrekker. Tilfellet som er interessant i forhold til å studere kointegrasjon er når Rang ($\boldsymbol{\Pi}$) = $r < k$.

Johansens test har to varianter, den ene er trace-testen og den andre er maksimum egenverdi-testen. Begge testene har de samme nullhypotesene, men forskjellige alternative hypoteser.

I trace-kointegrasjonstesten antar vi hypotesene:

$$H_0 : \text{rang}(\boldsymbol{\Pi}) = r \text{ mot } H_1 : \text{rang}(\boldsymbol{\Pi}) > r.$$

For å utføre testen har Johansen vist at en sannsynlighetskvoteobservator (LR) med maks r kointegrerte vektorer kan skrives som (Tsay, 2005, s. 385):

$$\text{LK}_{tr}(r) = -(T - p) \sum_{i=r+1}^k \log(1 - \hat{\lambda}_i). \quad (2.16)$$

For mer informasjon om testobservatoren i ligning (2.16) se i Tsay (2005, s. 385). De kritiske verdiene til observatoren finnes via simulering. Hvis vi forkaster nullhypotesen kan vi konkludere med at det finnes minst ett kointegrasjons forhold mellom variablene.

Utenom denne trace observatoren har Johansen og Juselius foreslått en maksimum egenverdi observator (Pfaff, 2008, s. 81). I maksimum egenverdi kointegrasjons testen antar vi hypotesene

$$H_0 : \text{rang}(\mathbf{\Pi}) = r \text{ mot } H_1 : \text{rang}(\mathbf{\Pi}) = r + 1.$$

Maksimum egenverdi observator er gitt som (Tsay, 2005, s 385):

$$\text{LK}_{max}(r) = -(T - p) \log(1 - \hat{\lambda}_{r+1}), \quad (2.17)$$

her tester vi eksistensen av r mot $r + 1$ kointegrerte forhold. Også her må de kritiske verdiene finnes via simulering.

KAPITTEL 3

Volatilitets modeller

I dette kapitlet skal vi ta utgangspunkt i Generalized Autoregressive Conditional Heteroskedasticity (GARCH)-modeller. Poenget med GARCH-modeller er å finne den betingete variansen av en prosess $\{X_t\}$. GARCH-modeller kan predikere og analysere volatiliteten i tidsrekker, når denne varierer over tid. En GARCH-modell kan ta hensyn til volatilitetsklumpingen, som vil si at store prisendringer har en tendens til å bli etterfulgt av store prisendringer.

Grunnen til at vi ønsker å studere GARCH-modeller er at mange observerte tidsrekker spesielt innenfor finans, har volatilitetsklumping (Brockwell and Davis, 1996, s. 337). I tillegg til finansmarkedet har det tidligere blitt vist at petroleumsmarkedet sine avkastninger har volatilitetsklumping. Noen som tidligere har brukt GARCH-modell for å modellere den betingede variansen i gassmarkedet er for eksempel Gregoire et al. (2008). Senere i oppgaven, skal vi modellere gassprisene. Siden gassprisene er volatile, skal vi tilpasse flere volatilitetsmodeller for gassmarkedene. I GARCH-modeller er den betingede variansen avhengig av fortiden (heteroskedastisk). Volatilitet kan ikke modelleres som en lineær prosess, fordi i en lineær prosess er den betingede variansen konstant (homoskedastisk). Lineære modeller klarer ikke å ta hensyn til mulighetene for at visse perioder av fortiden kan tillate mer nøyaktige prognoser enn andre (Brockwell and Davis, 1996, s. 337).

Ulike investeringsgrupper bruker forståelsen av volatilitet til ulike formål. For eksempel vil en risiko manager vite sannsynligheten idag for at porteføljen vil falle i fremtiden. Prisvolatilitet kan dermed oppfordre eller fraråde bedrifter mot å investere. I handel gir økende prisvolatilitet oppmuntring til en arbitrasjemulighet.

Til slutt i dette kapitlet skal vi se på noen tester som vi kan bruke for å teste om vi har volatilitet i tidsrekken.

3.1 ARCH-modeller

I dette delkapittelet skal vi se på ARCH-modeller. En ARCH-modell er et spesial tilfelle av en GARCH-modell.

3.1.1 ARCH(q)-modell

Den mest generelle ARCH-modellen, kalles for en ARCH(q)-modell, hvor q står for antall lag i modellen. Anta vi har en ARCH(q) prosess $\{X_t\}$. Løsningen til $\{X_t\}$ ble introdusert av Engle (1982) som en løsning av ligningen (Brockwell and Davis, 1996, s. 341):

$$X_t = \sigma_t Z_t. \quad (3.1)$$

Ofte er $\{Z_t\}$ standardnormalfordelt, men hvis vi ønsker å modulere med for eksempel tyngre haler kan vi anta at $\{Z_t\}$ er student t-fordelt. I ligning (3.1) er σ_t den betingede variansen. Den betingede variansen σ_t er en positiv, tidsvarierende og målbar funksjon av $\{X_s, s < t\}$, og gitt ved (Brockwell and Davis, 1996, s. 341):

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2, \quad (3.2)$$

her er $\alpha_0 > 0$ og $\alpha_i \geq 0$, der $i = 1, \dots, q$.

En ARCH(q) prosess er som hvit støy, men den betingete variansen er ikke konstant. Ligningen (3.2) fanger opp volatilitets klumpingen i prosessen. Den generelle ARCH-modellen er symmetrisk, og sier ikke noe om retningen til prisendringene.

3.1.2 Ordenbestemmelse i ARCH-prosessen

En kan bruke PACF-plott på tidsrekken $\{X_t^2\}$ for å bestemme ARCH-ordenen. Det kan vi gjøre fordi $\{X_t^2\}$ er som en AR modell (2.7). For et gitt utvalg er X_t^2 et "unbiased" estimat til σ_t^2 (Tsay, 2005, s. 106). Derfor forventer vi at X_t^2 er lineært avhengig til $X_{t-1}^2, \dots, X_{t-q}^2$ på en måte som ligner en AR modell av ordenen q . Generelt er ikke X_t^2 et effisient estimat av σ_t^2 , men det er en nyttig approksimasjon som kan brukes for å spesifisere ordenen q (Tsay, 2005, s. 106).

3.1.3 ARCH(1)-modell

Den enkleste ARCH-modellen er en ARCH(1) modell. ARCH(1)-modellen er gitt som:

$$X_t = Z_t \sigma_t \quad (3.3)$$

$$\sigma_t^2 = \alpha_0 + \alpha X_{t-1}^2. \quad (3.4)$$

I prosessen $\{X_t\}$ antar vi at $\{Z_t\}$ er i.i.d. normalfordelt hvit støy, der betinget forventning og betinget varians er gitt som:

$$\begin{aligned} E(Z_t|\mathcal{F}_{t-1}) &= 0, \\ \text{VAR}(Z_t|\mathcal{F}_{t-1}) &= 1, \end{aligned}$$

der \mathcal{F} er historien, og er gitt ved: $\mathcal{F}_{t-1} = \{X_{t-1}, X_{t-2}, \dots, X_1\}$. Variansen har nå betinget homoskedastisitet. En følgende egenskap er at $E(Z_t^2|\mathcal{F}_{t-1}) = \text{VAR}(Z_t|\mathcal{F}_{t-1}) - E(Z_t|\mathcal{F}_{t-1})^2 = 1$.

Hvis vi setter den betingede variansen i ARCH(1)-modellen (3.4) inn i prosessen $\{X_t\}$ (3.3), vil dette gi oss:

$$\begin{aligned} X_t^2 &= (\alpha_0 + \alpha_1 X_{t-1}^2) Z_t^2 \\ &= \alpha_1 X_{t-1}^2 Z_t^2 + \alpha_0 Z_t^2, \end{aligned} \quad (3.5)$$

Ligning (3.5) medfører at den betingete forventningen til X_t er $E(X_t|\mathcal{F}_{t-1}) = 0$. Videre finner vi den betingete variansen til X_t . Den betingede variansen er gitt som:

$$\begin{aligned} \text{VAR}(X_t|\mathcal{F}_{t-1}) &= E[X_t^2|\mathcal{F}_{t-1}] - E^2[X_t|\mathcal{F}_{t-1}] \\ &= E[X_t^2|\mathcal{F}_{t-1}] \\ &= \sigma_t^2 E[Z_t^2|\mathcal{F}_{t-1}] \\ &= \sigma_t^2 \end{aligned} \quad (3.6)$$

Ligning (3.6) viser oss hvordan en ARCH prosess går. Hvis X_{t-1} har en uvanlig stor verdi, vil også σ_t^2 bli uvanlig stor, slik at X_t er forventet til å ha en usedvanlig stor størrelse. Tilsvarende hvis X_{t-1} er usedvanlig liten.

I ARCH(1) tilfellet gir rekursjon av ligning (3.4) og (3.3) (Brockwell and Davis, 1996, s. 341):

$$\begin{aligned} X_t^2 &= \alpha_0 Z_t^2 + \alpha_1 Z_t^2 X_{t-1}^2 \\ &= \alpha_0 Z_t^2 + \alpha_1 Z_t^2 (\alpha_0 Z_{t-1}^2 + \alpha_1 Z_{t-1}^2 X_{t-2}^2) \\ &= \alpha_0 Z_t^2 + \alpha_0 \alpha_1 Z_t^2 Z_{t-1}^2 + \alpha_1^2 Z_t^2 Z_{t-1}^2 X_{t-2}^2 \\ &= \dots \\ &= \alpha_0 Z_t^2 + \alpha_0 \alpha_1 Z_t^2 Z_{t-1}^2 + \alpha_0 \alpha_1^2 Z_t^2 Z_{t-1}^2 Z_{t-2}^2 \\ &\quad + \alpha_0 \alpha_1^3 Z_t^2 Z_{t-1}^2 Z_{t-2}^2 + \dots + \alpha_0 \alpha_1^n (Z_t^2 \dots Z_{t-n}^2) + \alpha_1^{n+1} (Z_t^2 Z_{t-1}^2 \dots Z_{t-n}^2) X_{t-n-1}^2 \\ &= \alpha_0 \sum_{j=0}^n \alpha_1^j (Z_t^2 Z_{t-1}^2 \dots Z_{t-j}^2) + \alpha_1^{n+1} (Z_t^2 Z_{t-1}^2 \dots Z_{t-n}^2) X_{t-n-1}^2, \end{aligned} \quad (3.7)$$

Hvis $|\alpha_1| < 1$ og $\{X_t\}$ er stasjonær og avhengig, da konvergerer forventningen til det siste leddet i ligning (3.7) mot 0 og summen $\sum_{j=0}^{\infty} \alpha_1^j = \frac{1}{1 - \alpha_1}$ når $n \rightarrow \infty$. Dermed er den ubetingede

variansen gitt som (Brockwell and Davis, 1996, s. 341):

$$\sigma^2 = \text{E}(X_t^2) = \frac{\alpha_0}{1 - \alpha_1} \quad (3.8)$$

Log-likelihoodfunksjonen for ARCH(1) modellen

I denne delen skal vi finne log-likelihoodfunksjonen til ARCH(1) modellen. Fra tidligere har vi at ARCH(1) modellen er gitt som:

$$\begin{aligned} X_t &= Z_t \sigma_t \\ \sigma_t^2 &= \alpha_0 + \alpha X_{t-1}^2, \end{aligned}$$

der vi antar at $\{Z_t\} \sim N(0, 1)$. Denne modellen har en to dimensjonal vektor $\boldsymbol{\theta} = (\alpha_0, \alpha)$ som vi ønsker å estimere.

Vi betinger på $X_1|x_1$, slik at $\sigma_2^2 = \alpha_0 + \alpha x_1$ blir fast. Vi finner den betingede tettheten til X_t ved å bruke transformasjonsformelen. Transformasjonsformelen er gitt som:

$$f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{Z}}(\mathbf{z}(\mathbf{x})) \left| \frac{\partial \mathbf{z}}{\partial \mathbf{x}} \right|,$$

der $\mathbf{x}_t = (x_1, \dots, x_n)$ og $\mathbf{z}_t = (z_1, \dots, z_n)$.

Ved definisjon er tettheten til $Z_t \sim N(0, 1)$ gitt som:

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right].$$

Fra ligning (3.3) har vi at $Z_t = X_t/\sigma_t$. Dermed blir simultantettheten til X_t ved transformasjonsformelen gitt som

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{x_t^2}{2\sigma_t^2} \right] \left| \frac{1}{\sigma_t} \right|$$

Dermed blir den betingede likelihoodfunksjonen av observasjonene $\{x_2, \dots, x_n$ gitt at $X_1 = x_1\}$ gitt som:

$$\begin{aligned} L &= \prod_{t=2}^n \frac{1}{(\pi(\sigma_t^2))^{\frac{1}{2}}} \exp \left[-\frac{x_t^2}{2\sigma_t^2} \right] \\ &= (2\pi)^{-n/2} \left(\prod_{t=2}^n \sigma_t^2 \right)^{-1/2} \exp \left[-\frac{1}{2} \sum_{t=2}^n \frac{x_t^2}{\sigma_t^2} \right] \end{aligned}$$

Log-likelihoodfunksjonen blir da:

$$\log L = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^n \log(\sigma_t^2) - \frac{1}{2} \sum_{t=2}^n \frac{x_t^2}{\sigma_t^2} \quad (3.9)$$

Estimere ARCH(1)-regresjonsmodell

For å estimere en ARCH(1)-regresjonsmodell må vi estimere parametrene i θ . Dette gjøres ved å maksimere likelihoodfunksjonen med hensyn på θ . Det første vi gjør er å derivere ligning (3.9), dette gir oss likelihoodligningene:

$$\begin{aligned}\nabla_{\theta} \log L &= -\frac{1}{2} \sum_{t=2}^n \frac{1}{\sigma_t^2} \nabla_{\theta} \sigma_t^2 + \frac{1}{2} \sum_{t=2}^n \frac{x_t^2 \nabla_{\theta} \sigma_t^2}{(\sigma_t^2)^2} \\ &= \frac{1}{2} \sum_{t=2}^n \left(\frac{x_t^2 - \sigma_t^2}{\sigma_t^4} \right) \nabla_{\theta} \sigma_t^2 \\ &= \sum_{t=2}^n G_t \nabla_{\theta} \sigma_t^2,\end{aligned}$$

der $G_t = (x_t^2 - \sigma_t^2)/\sigma_t^4$. Vi definerer $Y_t = X_t^2$. For å finne maksimum setter vi

$\sum_{t=2}^n G_t \nabla_{\theta} \sigma_t^2 = 0$. I ARCH(1) modellen har vi at $\nabla_{\theta} \sigma_t^2 = \begin{bmatrix} 1 \\ y_{t-1} \end{bmatrix}$. Likelihoodligningene er dermed gitt som:

$$(1/2) \sum_{t=2}^n [(y_t - \sigma_t^2)/\sigma_t^4] \begin{bmatrix} 1 \\ y_{t-1} \end{bmatrix} = 0.$$

Dette gir oss følgende to ligninger:

$$\begin{aligned}\sum_{t=2}^n \left(\frac{Y_t - \sigma_t^2}{\sigma_t^4} \right) &= 0 \\ \sum_{t=2}^n \left(\frac{Y_t - \sigma_t^2}{\sigma_t^4} \right) Y_{t-1} &= 0.\end{aligned}$$

Vi har at $Y_0 = 0$ og setter inn $\sigma_t^2 = \alpha_0 + \alpha Y_{t-1}$,

$$\sum_{t=2}^n \frac{Y_t - \alpha_0 - \alpha Y_{t-1}}{(\alpha_0 + \alpha Y_{t-1})^2} = 0 \quad (3.10)$$

$$\sum_{t=2}^n \frac{(Y_t - \alpha_0 - \alpha Y_{t-1}) Y_{t-1}}{(\alpha_0 + \alpha Y_{t-1})^2} = 0. \quad (3.11)$$

Disse ligningene kan bli løst ved minste kvadratisk vektete iterasjoner, hvor vekten er definert ved nevneren i ligningene (3.11) og (3.10):

$$W_t = (\alpha_0 + \alpha Y_{t-1})^{-2}. \quad (3.12)$$

Setter (3.12) inn i ligning (3.11) og (3.10), og dette gir oss:

$$\sum_{t=2}^n W_t(Y_t - \alpha_0 - \alpha Y_{t-1}) = 0 \quad (3.13)$$

$$\sum_{t=2}^n W_t(Y_t Y_{t-1} - \alpha_0 Y_{t-1} - \alpha Y_{t-1} Y_{t-1}) = 0. \quad (3.14)$$

Videre definerer vi

$$S_0 = \sum_{t=2}^n W_t, S_1 = \sum_{t=2}^n W_t Y_t, S_2 = \sum_{t=2}^n W_t Y_{t-1}, S_3 = \sum_{t=2}^n W_t Y_t Y_{t-1} \text{ og } S_4 = \sum_{t=2}^n W_t Y_{t-1}^2.$$

Dermed kan vi skrive (3.13) og (3.14) som:

$$S_1 - \alpha_0 S_0 - \alpha S_2 = 0 \quad (3.15)$$

$$S_3 - \alpha_0 S_2 - \alpha S_4 = 0. \quad (3.16)$$

Ligningene (3.15) og (3.16) blir to ligninger med to ukjente og kan løses på matriseform:

$$\begin{bmatrix} S_0 & S_2 \\ S_2 & S_4 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha \end{bmatrix} = \begin{bmatrix} S_1 \\ S_3 \end{bmatrix} \quad (3.17)$$

Løsningen av (3.17) er gitt som:

$$\alpha_0 = \frac{S_1 S_4 - S_2 S_3}{S_4 S_0 - S_2^2} = (S_1 S_4 - S_2 S_3)(S_4 S_0 - S_2^2)^{-1} \quad (3.18)$$

$$\alpha = \frac{S_3 S_0 - S_1 S_2}{S_4 S_0 - S_2^2} = (S_3 S_0 - S_1 S_2)(S_4 S_0 - S_2^2)^{-1}. \quad (3.19)$$

3.2 GARCH-modeller

En ARCH(q)-modell kan utvides til en mer generell modell ved å tillate autoregresjon for σ_t^2 . Da kalles modellen for en generalisert ARCH- eller GARCH(q, p)-modell.

3.2.1 GARCH(q, p)-modell

En GARCH(q, p) modell er definert med følgende ligninger (Francq and Zakoian, 2010, s. 19):

$$X_t = \sigma_t Z_t \quad (3.20)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad (3.21)$$

der α_0 , $\alpha_i, i = 1, \dots, q$ og $\beta_j, j = 1, \dots, p$ er parametrene. Z_t er en i.i.d. tidsrekke med en fordeling F . Det er ofte antatt at fordelingen F er normalfordelt, men tidligere forskning har vist at $\{Z_t\}$ som er brukt for å tilpasse en GARCH-modell i finans og petroleumsmarkeder ofte har tykkhalet sannsynlighetsfordeling (Gregoire et al., 2008).

I ligning (3.21) er σ_t^2 spesifisert vet ett ARCH-ledd, $\alpha_0 + \sum_{i=1}^q \alpha_i X_{t-1}^2$, og ett GARCH tillegg, $\sum_{j=1}^p \beta_j \sigma_{t-1}^2$. GARCH representerer effekten som nye hendelser i forrige periode har på nåværende volatilitet slik som i ARCH-modellen, mens i motsetning til ARCH prosessen er σ_t også direkte avhengig av variansen i forrige periode. Dermed kan den betingete variansen vise mer om tidligere perioder hadde høy volatilitet, enn hva som kan sees i ARCH prosessen. Vi har at parametrene $\alpha_i \geq 0$ og $\beta_j \geq 0$. Dette er for å forsikre oss om at σ_t^2 alltid er positiv. Hvis α_i er lik 0 for alle i , må også β_j være lik 0 for alle j , ellers blir rekken σ_t^2 en konstant og β er da ikke identifiserbar. Hvis $\beta_j = 0$ for alle j , blir GARCH en ARCH-modell.

Hvis den betingede variansen σ_t er målbar fra $\mathcal{F}_{t-1} = \mathcal{F}_{t-1}^X$, og Z_t er en prosess med forventning null, varians lik en, og uavhengig av \mathcal{F}_{t-1}^X . Da har vi at forventning og varians til X_t er gitt som (Francq and Zakoian, 2010, s. 11):

$$\begin{aligned} \mathbb{E}(X_t | \mathcal{F}_{t-1}) &= \mathbb{E}(\sigma_t Z_t | \mathcal{F}_{t-1}) = \sigma_t \mathbb{E}(Z_t | \mathcal{F}_{t-1}) = 0 \\ \mathbb{E}(X_t^2 | \mathcal{F}_{t-1}) &= \sigma_t^2 \mathbb{E}(Z_t^2 | \mathcal{F}_{t-1}) = \sigma_t^2 \\ \text{VAR}(X_t | \mathcal{F}_{t-1}) &= \mathbb{E}(\sigma_t^2 Z_t^2 | \mathcal{F}_{t-1}) - [\mathbb{E}(X_t | \mathcal{F}_{t-1})]^2 = \sigma_t^2. \end{aligned}$$

3.2.2 GARCH(1,1)-modell

Den enkleste GARCH modellen får vi når $p = q = 1$. Denne GARCH-modellen kalles for en GARCH(1,1) modell, og er gitt som:

$$X_t = \sigma_t Z_t \tag{3.22}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \tag{3.23}$$

Der $\{Z_t\}$ er en i.i.d. tidsrekke med en fordeling F .

Anta at $\alpha_0 > 0$. Hvis $\alpha_1 + \beta_1 \geq 1$ eksisterer, da finnes det ikke en annen-ordens stasjonær løsning av GARCH(1,1)-modellen (Francq and Zakoian, 2010, s. 27). Derimot hvis betingelsen $\alpha_1 + \beta_1 < 1$, da er prosessen $\{X_t\}$ andre-ordens stasjonær (Francq and Zakoian, 2010, s. 27). Det vil si at prosessen $\{X_t\}$ er svak hvit støy (Francq and Zakoian, 2010, s. 27). Hvis denne stasjonære betingelsen holder, da er den ubetingete variansene gitt som:

$$\text{VAR}(X_t) = \mathbb{E}(X_t^2) = \sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1} \tag{3.24}$$

Ved hensyn på kovariansstrukturen er en GARCH(1,1) en ARMA(1,1). Ved definisjon er $U_t = \sigma_t^2(Z_t^2 - 1)$ og $Y_t = X_t^2$. Vi setter definisjonene inn i ligning (3.22):

$$Y_t = X_t^2 = \sigma_t^2 Z_t^2 = \sigma_t^2 + \sigma_t^2 Z_t^2 - \sigma_t^2 = \sigma_t^2 + \sigma_t^2(Z_t^2 - 1) = \sigma_t^2 + U_t.$$

Da har vi at $\sigma_{t-1} = Y_{t-1} - U_{t-1}$, som vi kan sette inn i ligning (3.20):

$$\begin{aligned}\sigma_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \\ &= \alpha_0 + \alpha_1 Y_{t-1} + \beta_1 (Y_{t-1} - U_{t-1}) \\ &= \alpha_0 + (\alpha_1 + \beta_1) Y_{t-1} - \beta_1 U_{t-1}\end{aligned}$$

Til slutt har vi en ARMA(1,1)-modell:

$$Y_t = \sigma_t^2 + U_t = \alpha_0 + (\alpha_1 + \beta_1) Y_{t-1} + (-\beta_1) U_{t-1} + U_t,$$

der $\theta = -\beta_1$, $\phi = (\alpha_1 + \beta_1)$ og $\mu = \alpha_0$.

3.2.3 Predikere volatilitet

Fra ligningene (3.22) og (3.23) har vi at GARCH(1,1) modellen er gitt som:

$$\begin{aligned}X_t &= \sigma_t Z_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2,\end{aligned}$$

anta at $Z_t \sim N(0,1)$. Da er de betingede forventningene til observasjonen og de kvadrerte observasjonen henholdsvis:

$$\begin{aligned}\mathbb{E}(X_t | \mathcal{F}_{t-1}) &= \mathbb{E}(\sigma_t Z_t | \mathcal{F}_{t-1}) = \sigma_t \mathbb{E}(Z_t | \mathcal{F}_{t-1}) = 0 \\ \mathbb{E}(X_t^2 | \mathcal{F}_{t-1}) &= \sigma_t^2 \mathbb{E}(Z_t^2 | \mathcal{F}_{t-1}) = \sigma_t^2.\end{aligned}$$

Anta at vi har n observasjoner. Selve prediksjonen for den betingede variansen i neste periode er gitt som (Reider, 2009, s. 7):

$$\begin{aligned}\hat{\sigma}_{n+1}^2 &= \alpha_0 + \alpha_1 \mathbb{E}(X_n^2 | \mathcal{F}_{n-1}) + \beta_1 \sigma_n^2 \\ &= \alpha_0 + \alpha_1 \sigma_n^2 + \beta_1 \sigma_n^2 \\ &= \alpha_0 + (\alpha_1 + \beta_1) \sigma_n^2 \\ &= \sigma^2 (1 - \alpha_1 - \beta_1) + \sigma_n^2 (\alpha_1 + \beta_1) \\ &= \sigma^2 + (\sigma_n^2 - \sigma^2) (\alpha_1 + \beta_1)\end{aligned}$$

Der σ^2 er den ubetingede variansen i ligning (3.24), når den stasjonære betingelsen holder. Prognosene videre for de betingede variansen i $n+l$, der $l = 2, 3, \dots$ er (Reider, 2009, s. 7):

$$\begin{aligned}
\hat{\sigma}_{n+2}^2 &= \alpha_0 + \alpha_1 E(X_{n+1}^2 | \mathcal{F}_{n-1}) + \beta_1 [\sigma_{n+1}^2 | \mathcal{F}_{n-1}] \\
&= \alpha_0 + \alpha_1 \hat{\sigma}_{n+1}^2 + \beta_1 \hat{\sigma}_{n+1}^2 \\
&= \alpha_0 + (\alpha_1 + \beta_1) \hat{\sigma}_{n+1}^2 \\
&= \sigma^2(1 - \alpha_1 - \beta_1) + \sigma_{n+1}^2(\alpha_1 + \beta_1) \\
&= \sigma^2 + (\sigma_{n+1}^2 - \sigma^2)(\alpha_1 + \beta_1) \\
&= \sigma^2 + (\sigma^2 + (\sigma_n^2 - \sigma^2)(\alpha_1 + \beta_1) - \sigma^2)(\alpha_1 + \beta_1) \\
&= \sigma^2 + (\sigma_n^2 - \sigma^2)(\alpha_1 + \beta_1)^2 \\
&\vdots \\
\hat{\sigma}_{n+l}^2 &= \sigma^2 + (\sigma_n^2 - \sigma^2)(\alpha_1 + \beta_1)^l
\end{aligned} \tag{3.25}$$

3.2.4 Andre utvidelser av GARCH

Det finnes en lang liste med ulike varianter av GARCH ligninger. Tidligere har vi sett på GARCH-modeller som er symmetriske. I tillegg eksisterer det GARCH-modeller som tillater asymmetri. Det finnes også multivariate GARCH-prosesser som CCC-GARCH og BEKK-GARCH. For mer informasjon om multivariate GARCH-modeller se kapittel 11.2 i boken GARCH models av Francq and Zakoian (2010).

Et eksempel på en utvidet GARCH-modell som er asymmetrisk er EGARCH. EGARCH er en eksponentiell GARCH-modell. EGARCH(q, p) prosessen er gitt ved (Francq and Zakoian, 2010, s. 246):

$$\begin{aligned}
X_t &= \sigma_t Z_t \\
\log \sigma_t^2 &= \alpha_0 + \sum_{i=1}^q \alpha_i g(Z_{t-1}) + \sum_{j=1}^p \beta_j \log \sigma_{t-1}^2,
\end{aligned}$$

der $g(Z_{t-1}) = \theta Z_{t-1} + \varsigma(|Z_{t-1}| - E|Z_{t-1}|)$, og $\alpha_i, \beta_j, \theta$, og ς er reelle tall.

En annen asymmetrisk GARCH-modell er GJR-GARCH-modellen. GJR-GARCH(1, 1)-prosessen er gitt ved (Bauwens et al., 2011, s. 7):

$$\begin{aligned}
X_t &= \sigma_t Z_t \\
\sigma_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \gamma X_{t-1}^2 I(X_{t-1} < 0).
\end{aligned}$$

3.3 Lagrange Multiplier ARCH test

Lagrange Multiplier (LM) ARCH test kan brukes for å teste om vi har betinget homoskedatisitet i observasjonene i en tidsrekke. Nullhypotesen er $H_0 = \alpha_1 = \dots = \alpha_q$. Anta en ARCH-modell

med $\sigma_t^2 = h(\epsilon_t \alpha)$, der h er en deriverbar funksjon som inneholder både lineære, eksponentielle og andre tilfeller (Engle, 1982, s. 1000). $\epsilon_t = (1, Z_{t-1}^2, \dots, Z_{t-q}^2)$, der Z_t er OLS residualene. Under nullhypotesen er σ_t^2 gitt som h^0 . Vi kan skrive

$$\frac{\partial \sigma_t^2}{\partial \alpha} = h' \epsilon_t',$$

der h' er den deriverte av h .

Scorefunksjonen er gitt som:

$$\frac{\partial l}{\partial \alpha} \Big|_0 = \frac{h'}{2h^0} \sum_t \epsilon_t' \left(\frac{Z_t^2}{h^0} - 1 \right).$$

Informasjonen kan skrives som:

$$I_{\alpha\alpha}^0 = \frac{1}{2} \left(\frac{(h^0)'}{h^0} \right) \mathbb{E} \epsilon' \epsilon$$

LM test-observator er:

$$\xi^* = \frac{1}{2} \left(\frac{Z_t^2}{h^0} - 1 \right)' \epsilon (\epsilon' \epsilon)^{-1} \epsilon' \left(\frac{Z_t^2}{h^0} - 1 \right)$$

Dette er en test som kan brukes for å teste om vi har heteroskedastisitet (Engle, 1982, s. 1000). Det finnes også andre forenklinger av test-observatoren.

3.4 Ljung-Box test

Ljung-Box test kan teste om det er uavhengighet mellom observasjonene i en tidsrekke. Hvis vi tar Ljung-Box test på de kvadrerte observasjonene kan vi teste om vi har heteroskedastisitet.

Ljung-Box test-observator er gitt som (Tsay, 2005, s. 27):

$$Q(m) = n(n-2) \sum_{l=1}^m \frac{\hat{\rho}_l^2}{n-l}, \quad (3.26)$$

der ρ er autokorrelasjonsfunksjonen og gitt ved ligning (2.5). $Q(m)$ er en forbedring av Portmanteau observatoren som er gitt som $Q^*(m) = n \sum_{l=1}^m \hat{\rho}_l^2$.

Hypotesene til Ljung-Box test er nullhypotesen som er gitt som $H_0 : \rho_1 = \rho_2 = \dots = \rho_m = 0$ mot den alternative hypotesen $H_1 : \rho_i \neq 0, i \in [1, \dots, m]$. Vi forkaster H_0 hvis $Q(m) > \chi_{\alpha}^2$. Der χ_{α}^2 er 100(1 - α)-persentilen av en kji-kvadrat fordeling med m frihetsgrader (Tsay, 2005, s. 27).

KAPITTEL 4

Copula

Vi har grunn til å tro at gassmarkedene i Europa er avhengige av hverandre. Grunnen til dette er blant annet fordi naturgassmarkedene er knyttet sammen via rørledningsnett. I dette kapitlet skal vi se på copulateori. Ved å bruke copula kan vi først finne fordelingene som passer hvert gassmarked best, og deretter knytte disse sammen med en avhengighetsstruktur. Først i dette kapitlet er det noen definisjoner i copulateorien og om fordelinger. Deretter studerer vi fire forskjellige copulafamilier. Til slutt i dette kapitlet ser vi på hvordan parametrene kan estimeres, modelltestes og simuleres ved bruk av copulateori.

4.1 Definisjon

Vi lar X og Y være tilfeldige variabler med kontinuerlig fordeling, med de kumulative fordelingsfunksjonene F_X og F_Y . Simultanfordelingen til de uniformt fordelte variablene $U = F_X(X)$ og $V = F_Y(Y)$ kalles da for Copulaen til (X, Y) . Sklars teorem ble introdusert av Sklar (1959). Sklars teorem sier at hvis F er den simultane kumulative fordelingsfunksjonen med marginene F_X og F_Y , eksisterer det for alle $(x, y) \in [-\infty, \infty]$ en copula C slik at:

$$F(x, y) = C(F_X(x), F_Y(y)). \quad (4.1)$$

Denne copulaen, C , er unik hvis $F_X(x), F_Y(y)$ er kontinuerlig (Nelsen, 2006, s. 18). Andre egenskaper til copula som gjelder på grunn av at $u \in [0, 1]$ og $v \in [0, 1]$, er følgende (Nelsen, 2006, s. 10):

- $C(u, v) \leq C(u', v')$, der $u \in u'$ og $v \in v'$.
- $C(u, 1) = u$ og $C(1, v) = v$
- $C(u, 0) = C(0, v) = 0$

Mer generelt sier Sklars teorem at hvis $F = F_{\mathbf{X}}$ er den simultane kumulative fordelingsfunksjon til de tilfeldige variablene X_1, \dots, X_d med marginalene $F_1 = F_{X_1}, \dots, F_d = F_{X_d}$. Da eksisterer

det en flerdimensjonal copula C slik at:

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)).$$

4.2 Fordelinger

For enkelthetsskyld begrenser vi oss videre ved å kun se på det bivariate tilfellet og antar at $C(u, v)$ er kontinuerlig tetthet. Tettheten til copulaen er dermed gitt som (Carmona, 2014, s. 149):

$$c(u, v) = \frac{\delta^2}{\delta u \delta v} C(u, v)$$

Vi har at F er den simultane kumulative fordelingsfunksjonen for parett (X, Y) , der C er deres copula. Da har vi definisjonen som er gitt i ligning (4.1). Videre kan vi finne simultantettheten f av X og Y ved å partiellderivere på begge sider av ligning (4.1).

$$\begin{aligned} \frac{\delta^2}{\delta u \delta v} F(x, y) &= \frac{\delta^2}{\delta u \delta v} C(F_X(x), F_Y(y)) \frac{\delta f_X(x)}{\delta x} \frac{\delta f_Y(y)}{\delta y} \\ f(x, y) &= c(F_X(x), F_Y(y)) f_X(x) f_Y(y) \\ &= c(u, v) f_X(x) f_Y(y) \end{aligned}$$

Nå er det mulig å spesifisere en simultanfordelingen ut fra de marginale fordelingene og en copula. De marginale fordelingene kan være fra hvilken som helst fordeling og trenger ikke å være fra samme familie eller å være lik hverandre.

4.3 Eksempel fra Copulafamilien

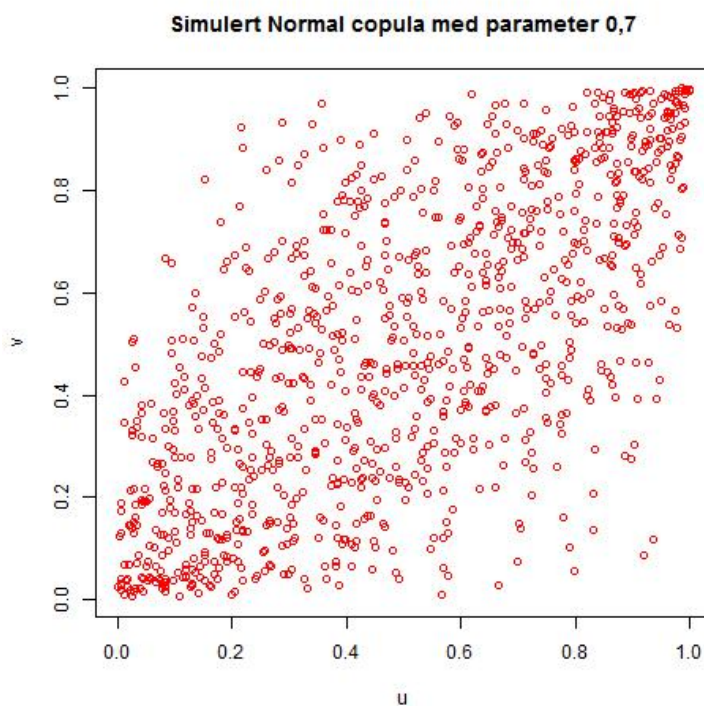
I denne delen skal vi studere fire forskjellige copulafamilier. Disse copulafamilier er henholdsvis Normal, Student t, Clayton og Gumbel.

4.3.1 Normal copulafamilie

En bivariat Normal copulafamilie med parameter $\rho \in (0, 1)$ er gitt ved (Nelsen, 2006, s. 23):

$$C_\rho(u, v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \exp\left[\frac{-(s^2 - 2\rho st + t^2)}{2(1-\rho^2)}\right] ds dt, \quad (4.2)$$

der $\Phi(\cdot)^{-1}$ er kvantilfunksjonen til standardnormalfordelingen. En Normal copulafamilie kan beskrive avhengighet hvis det er en jevn positiv og negativ avhengighet/uavhengighet. Figur 4.1 viser en simulert Normal copulafamilie med parameter, $\rho = 0,7$.



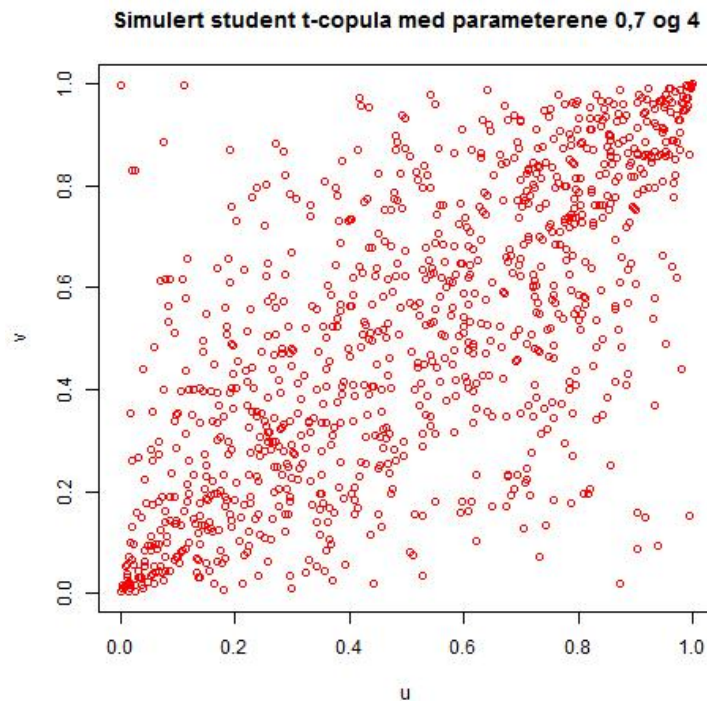
Figur 4.1: Scatterplott for copula (4.2), $\rho = 0,7$

4.3.2 Student t-copulafamilie

En Student t-copula (t-copula) gir muligheten for å inkludere tyngre haler. En t-copulafamilie tar hensyn til at det kan være en økende sannsynlighet for at vi har punkter med ekstreme hendelser. I motsetning til en Normal-copulafamilie som ikke tar med denne sannsynligheten. Student t-copulaen kan skrives som (Aas, 2004, s. 2):

$$C_{\rho,\nu}(u,v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{t_{\nu}^{-1}(u)} \int_{-\infty}^{t_{\nu}^{-1}(v)} \left[1 + \frac{s^2 - 2\rho st + t^2}{\nu(1-\rho^2)} \right]^{-(\nu+2)/2} ds dt, \quad (4.3)$$

ρ og ν er parameteren til copulaen, og $t_{\nu}(\cdot)$ er student-t-fordelingen med ν frihetsgrader, forventning lik 0, og variansen er gitt som $\nu/(\nu-2)$. I figur 4.2 ser vi at punktkonsentrasjonen er tettere i begge ender sammenlignet med den simulerte Normal copulafamilien i figur 4.1, men ellers er de ganske like i midtre del.



Figur 4.2: Scatterplott for copula (4.3), $\rho = 0,7$ og $\nu = 4$.

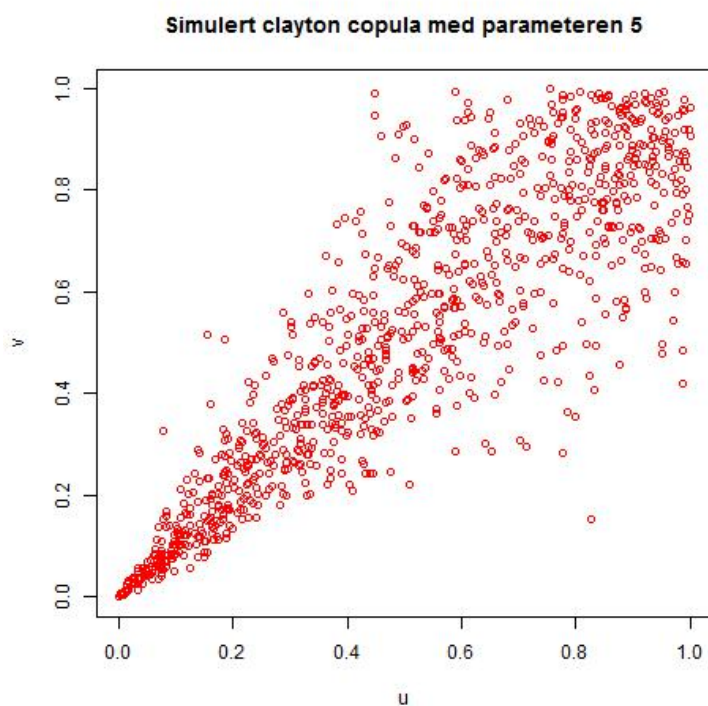
Både Normal copulafamilie og t-copulafamilie er eksempel på implisitte copulafamilier. Begge copulafamilierne er symmetriske familier. Videre skal vi se på Clayton copula og Gumbel copula som er noen eksempler på eksplisitte copulafamilier og tillater asymmetri. Clayton copula og Gumbel copula kalles for Archimedian copulaer.

4.3.3 Clayton copulafamilie

Hvis vi tror det er asymmetri er Clayton copula et bedre alternativ enn både Normal- og t-copulafamilier. Clayton copula tillater asymmetri, og har høyere avhengighet i de negative halene enn i de positive. I figur 4.3 ser vi tydelig at avhengigheten er høyest i den negative halen. Clayton copulafamilie er definert som (Nelsen, 2006, s. 116):

$$C_{\rho}(u, v) = (u^{-\rho} + v^{-\rho} - 1)^{-1/\rho}, \quad (4.4)$$

der $0 < \rho < \infty$ er en parameter som kontrollerer avhengigheten. Perfekt avhengighet er når $\rho \rightarrow \infty$, mens $\rho \rightarrow 0$ gir uavhengighet.



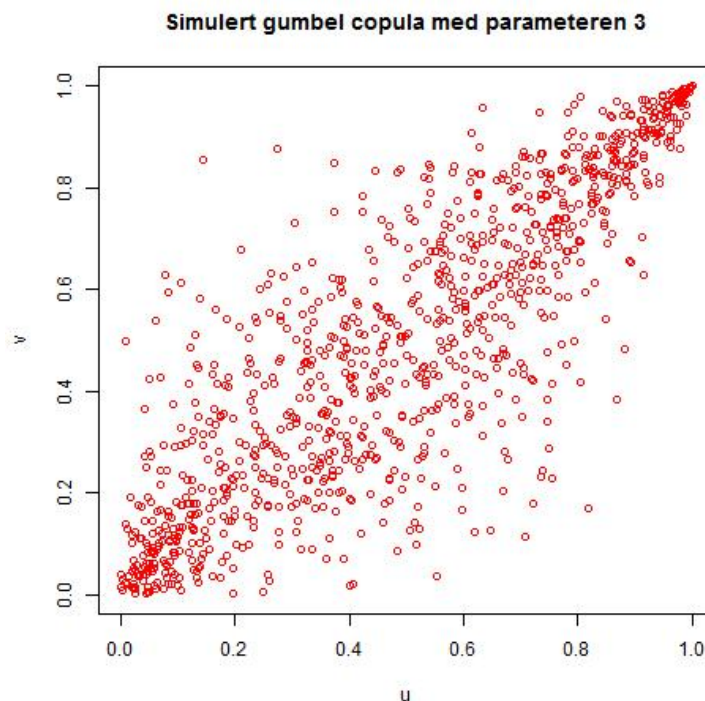
Figur 4.3: Scatterplott for copula (4.4), $\rho = 5$.

4.3.4 Gumbel copulafamilie

En Gumbel copula er også en asymmetrisk copula, men i Gumbel copulafamilien er det høyere avhengighet i den positive halen enn i den negative halen. Dette kan vi se i figur 4.4. Gumbel copulafamilien er gitt som (Nelsen, 2006, s. 116):

$$C_\rho(u, v) = \exp(-[(-\log(u))^\rho + (-\log(v))^\rho]^{1/\rho}), \quad (4.5)$$

der $\rho \geq 1$ er en parameter som kontrollerer avhengigheten. Perfekt avhengighet er når $\rho \rightarrow \infty$, mens $\rho = 1$ gir uavhengighet.



Figur 4.4: Scatterplott for copula (4.5), $\rho = 3$.

4.4 Kendalls tau og Spearmans rho

Copula beskriver avhengighetsstrukturen, dermed ønsker vi å bruke avhengighetsmål som er copula basert. Kendalls tau (τ) og Spearmans rho (S) er to avhengighetsmål som er copula basert.

Vi lar X og Y være to kontinuerlige og tilfeldige variabler, og lar C stå for copula. Da er Kendalls

tau for X og Y gitt ved (Nelsen, 2006, s. 161):

$$\tau_{X,Y} = \tau = 4 \int_0^1 \int_0^1 C(u,v) dC(u,v) - 1.$$

For Normal copula i ligning (4.2) og t copula i ligning (4.3) er forholdet mellom den lineære korrelasjonskoeffisient (CORR) og Kendalls tau gitt ved:

$$\text{CORR}(X, Y) = \sin(\pi\tau)/2. \quad (4.6)$$

Derimot for Clayton copula i ligning (4.4) er det vist i Nelson (Nelsen, 2006, s. 163) at vi har sammenhengen mellom Kendalls tau og avhengighetsvariabelen(ρ):

$$\tau = \rho/(\rho + 2), \quad (4.7)$$

der $\rho \geq -1$.

I tillegg er det vist i Nelson (Nelsen, 2006, s. 164) at for Gumbel copula er forholdet gitt som:

$$\tau = (\rho - 1)/\rho, \quad (4.8)$$

der $\rho \geq 1$.

Et annet mål på avhengighet er Spearman rho. Vi lar X og Y være kontinuerlige og tilfeldige variabler. Da er Spearmans rho for X og Y gitt ved (Nelsen, 2006, s. 167):

$$S_{X,Y} = S = 12 \int_0^1 \int_0^1 C(u,v) dudv - 3.$$

Videre antar vi at X og Y har henholdsvis fordelingsfunksjonene F og G . Da har vi følgende sammenheng mellom Spearmans rho og den lineære korrelasjonskoeffisienten:

$$S = \text{CORR}(F(X), F(Y)).$$

For den gaussiske og Student t-copulafamilien har vi følgende sammenheng mellom den lineære korrelasjons koeffisienten og Spearmans rho:

$$\text{CORR}(X, Y) = 2 \sin(\pi S)/6.$$

I følge Aas (2004, s. 4) måler τ og S graden av monoton avhengighet mellom X og Y , mens lineær korrelasjon bare måler graden av den lineære avhengigheten. I tillegg er τ og S upåvirket etter monotone transformasjoner, mens den lineære korrelasjonen er ofte endret etter monotone transformasjoner. Derfor er det bedre å bruke Kendalls tau og Spearmans rho enn den lineære korrelasjon koeffisienten.

4.5 Estimere parametrene

For hver copulafamilie en ønsker å estimere, kan vi estimere de tilhørende avhengighetsparametrene ved å bruke Kendalls tau (τ). Denne metoden kalles for den inverse til Kendalls tau metoden. Formlene for avhengighetsparametrene er hentet fra ligning (4.6), (4.7) og (4.8), og er gitt i tabell 4.1. Problemet med denne metoden er at den kun estimerer en parameter, mens t-familien faktisk inneholder to parametre. I tillegg til avhengighetsparameteren ρ , har den også frihetsgrad parameteren ν som modulerer formen til copula. Frihetsgradparameteren kan finnes ved å la ν variere for eksempel mellom 2 og 20, og deretter finne hvilken som passer best når vi kjører modelltester. Vi kan også estimere avhengighetsparametrene ved hjelp av Spearman rho.

Tabell 4.1: Formel for estimere parametre via den inverse til Kendalls tau.

Familie	Formel
Gaussisk	$\rho = \sin(\pi\tau/2)$
Student t	$\rho = \sin(\pi\tau/2)$
Clayton	$\rho = 2\tau/(1 - \tau)$ for $\tau > 1$
Gumbel	$\rho = 1/(1 - \tau)$ for $\tau > 1$

Alternativt kan man estimere parametrene ved MLE metoden. La oss si at vi vil finne en copula for utvalget $\{X\}$ og $\{Y\}$. Videre antar vi at de empiriske marginale fordelingene henholdsvis er F_X og F_Y . Derav kan det uniformt fordelte utvalget $\{U\}$ og $\{V\}$ fra copulaen konstrueres ved hjelp av de empiriske marginale fordelingene, $U = F_X(X)$ og $V = F_Y(Y)$. Etter vi har funnet U og V , kan vi bruke MLE for å estimere copula parametrene. Estimaten finner vi ved å maksimere

$$\log L(\Theta, U, V) = \sum_{i=1}^n \log c(U_i, V_i; \Theta),$$

der Θ inneholder hvilke copula parametre som skal maksimeres, og $c(u, v; \Theta)$ er tettheten til copulaen.

4.6 Modelltester

Etter at estimaten for de ulike copulafamilier er funnet, må en finne hvilken modell som passer best. Dette kan gjøres ved å kjøre tester på de ulike modellene. Eksempel på slike tester er Vuong og Clarke test, eller en test basert på empirisk funksjon. Testen basert på empirisk funksjon for bivariat copula finner Cramer-von Mises og Kolmogorov-Smirnov testobservatorer, og de tilhørende p-verdiene. Vuong og Clarke tester hvilken modell av to familier som passer best, og deretter sammenligner hvilken modell som fikk mest poeng.

Cramer-von Mises og Kolmogorov-Smirnov test basert på empirisk copula

Under nullhypotesen tester vi om $\mathbb{U} = (U_1, U_2)$ er fordelt som en copula C_θ for $\theta \in \mathcal{O}$, der \mathcal{O} inneholder spesifikasjonene til den copula familien vi vil teste om passer. Informasjonen de inneholder er summert av de tilknyttete empiriske fordelingsfunksjonene (Genest and D., 2009, s. 201):

$$C_n(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(U_{i1} \leq u_1, U_{i2} \leq u_2).$$

Der $(u_1, u_2) \in [0, 1]^2$. Denne empiriske funksjonen sammenlignes med den estimerte copulaen C_{θ_n} . Forskjellen mellom C_n og C_{θ_n} er gitt i prosessen $\mathbb{C} = \sqrt{n}(C_n - C_{\theta_n})$ (Genest and D., 2009, s. 201).

Genest and D. (2009, s. 201) foreslo blanket tester, altså rang-baserte variasjoner av Cramer-von Mises og Kolmogorov-Smirnov observator. De er henholdsvis gitt som

$$S_n = \int_{[0,1]^2} \mathbb{C}_n(\mathbf{u})^2 d\mathbb{C}_n(\mathbf{u}) \text{ og } T_n = \sup_{\mathbf{u} \in [0,1]^2} |\mathbb{C}_n(\mathbf{u})|.$$

Vuong og Clarke test

Denne testen finner hvilken copulafamilie som passer bedre enn de andre copula familiene. Både Vuong og Clarke test sammenligner to modeller med hverandre. De sammenligner modellene som er gitt i nullhypotesen med alle de andre mulige modellene. Måten de sammenlignes på er ved å se på forskjellene mellom likelihoodene mellom to modeller. Hvis modell 0 passer bedre enn en annen modell, får den +1 poeng, og -1 hvis det er den alternative modellen som passer best. Det kan gjøres for alle familiene vi ønsker å teste. For å finne den modellen som passer best, velger vi familien som får høyest poengsum. Mer detaljer om testene er gitt av Clarke (2007).

4.7 Simuleringer med copula

En kan bruke funksjonen *rcopula* fra pakken **copula**. Dette er en funksjon som kan produsere ett bivariat utvalg (U_t, V_t) fra en copula. Utvalget (U_t, V_t) har en copulaform som vi har definert, og er uniformt fordelt på $[0, 1]$. Deretter transformerer en U_t til den første marginalfordelingen ved å bruke kvantilfunksjonen til den første marginalfordelingen, og V_t ved å bruke kvantilfunksjonen til den andre marginalfordelingen. Da finner vi $X_t = F_X^{-1}(U_t)$ og $Y_t = F_Y^{-1}(V_t)$. Dermed har (X_t, Y_t) ikke bare de rette marginalene, men også den rette copulaen (Carmona, 2014, s. 153).

KAPITTEL 5

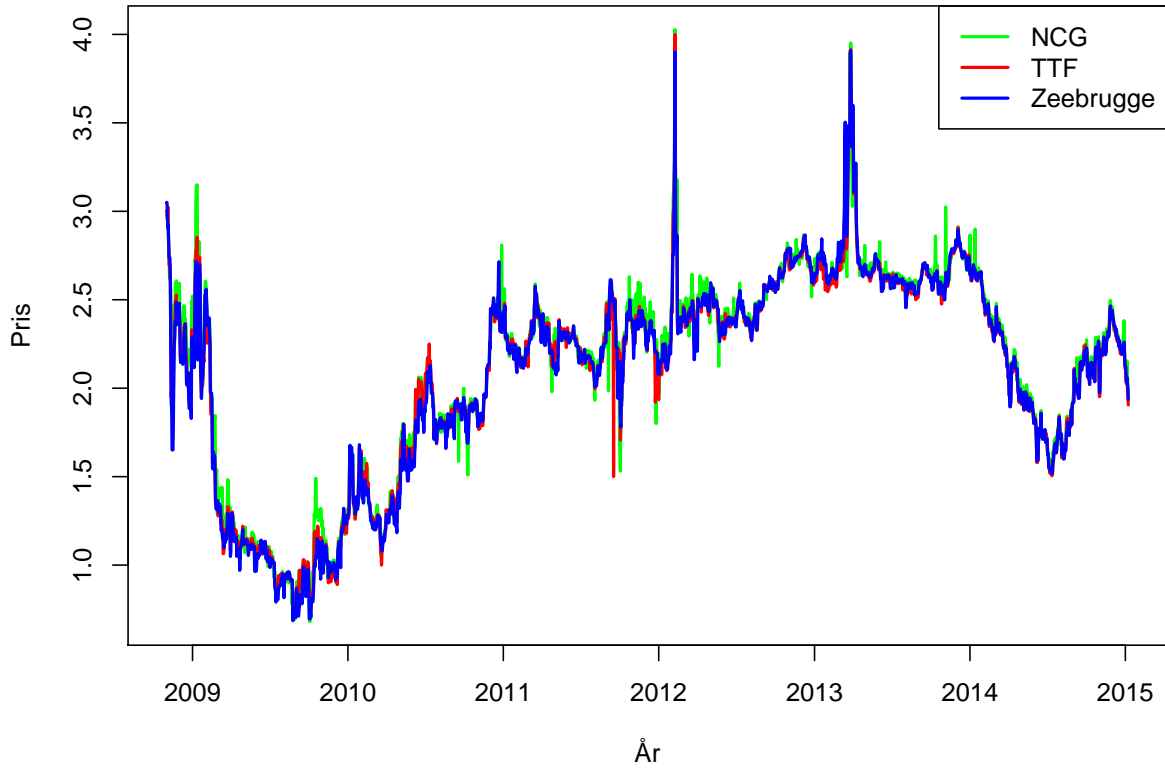
Statistisk analyse av Europeiske naturgasspriser

I dette kapitlet skal vi analysere gasspriser fra det tyske (NCG), belgiske (Zeebrugge) og nederlandske (TTF) gassmarkedet. Først skal vi ha en grafisk analyse på naturgassprisene direkte, før vi deretter omformer gassprisene til log-avkastninger. Videre studerer vi egenskapene til log-avkastningene. Log-avkastningene er tidsvarierende volatilitete, og blir tilpasset GARCH-modeller. Deretter sjekker vi for avhengighet mellom de ulike gassmarkedene. Etterpå tilpasser vi en passende copulafamilie. En copulafamilie kan knytte sammen fordelingene i de ulike gassmarkedene. Til slutt tester vi for kointegrasjon mellom gassprisene.

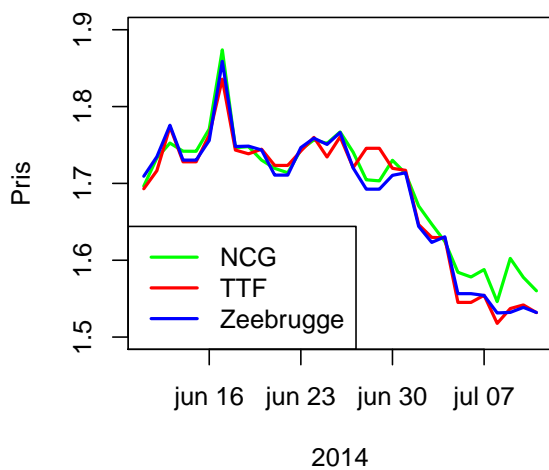
5.1 Datasett av Europeiske naturgasspriser

Hovedinteressen i denne oppgaven er å analysere og modellere prisene i gassmarkedet. Som det er nevnt tidligere finnes det mange forskjellige gassmarkeder. Det er ikke alle gassmarkedene som er tilgjengelige, og mange priser viser for eksempel kun tre måneder tilbake i tid. Datasettet som er brukt i denne oppgaven er hentet fra internettsiden: <https://www.net-connect-germany.de/en-gb/Transparency-information/Prices/Reference-Prices> (NetConnectGermany, 2015)¹. Dette datasettet inneholder daglige referansepriser for naturgasser fra handelsknutepunktene som NetConnect Germany (NCG), Title Transfer Facility (TTF) og Zeebrugge. Datasettet består av 2261 observasjoner for hver gasspris i tidsperioden 01.11.2008 til 08.01.2015. Alle prisene er gitt som Cent/kWh.

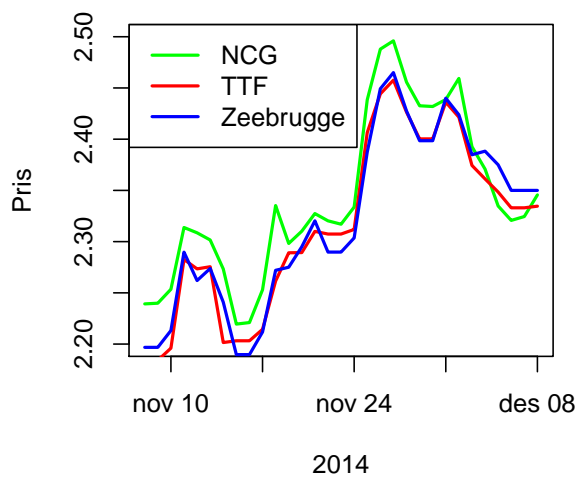
¹Sammensetningen av prisene har blitt endret på den 30.05.2011. Den tidligere referanseprisen *NCG One Day Ahead Settlement Price* har blitt erstattet av EEX priser *Daily Reference Price Natural Gas NCG*



(a) Priser fra 01.11.2008 til og med 08.01.2015



(b) Naturgasspriser i Europa fra jun/jul 2014



(c) Naturgasspriser i Europa fra nov/des 2014

Figur 5.1: Plott over naturgassprisene

5.1.1 Grafisk analyse over naturgassprisene

I figur 5.1a) ser vi alle gassprisene i hele tidsperioden under samme plott. Ut fra dette plottet ser det ut som prisene følger hverandre ganske bra, og at det ikke er store forskjeller mellom prisene i Europa. Ved å kun sammenligne en månedslengde om gangen kan vi i figur 5.1b) og 5.1c) se at det likevel er forskjeller i prisene, men at det ser ut som de følger hverandre. Denne sammenhengen skal vi analysere nærmere ved å bruke kointegrasjon senere i oppgaven.

I kapittel 1.3.1 nevnte vi at gassprisene ofte følger årstidene. Fra figur 5.1a) ser vi at gassprisene er relativt høye på vinteren i 2009, men synker mer og mer til nærmere sommeren vi kommer for deretter å øke igjen frem mot vinteren. En forklaring til denne nedgangen i gassprisen kan selvfølgelig være på grunn av de lave temperaturene som følger sesongen, men en annen grunn til dette kan være finanskrisen. Som vi nevnte i kapittel 1.3.2 ble energisektoren påvirket av finanskrisen i år 2008. Dette medførte at gassetterspørselen falt fra slutten av år 2008 og utover i år 2009. Hvis vi ser på figur 5.1a) ser vi at prisene falt fra begynnelsen av våren i 2009 og begynte ikke å øke igjen før nærmere 2010. En grunn til at gassprisen ble påvirket av finanskrisen kan være nedgangen i gassetterspørslene. Nedgangen i gassetterspørselen var mye på grunn av at industri som brukte naturgass ble lagt ned. I tillegg var det flere gasskraftverk som måtte stenges, og mange planer om å bygge nye gasskraftverk ble enten skrinlagt eller utsatt.

I år 2009 fortsatte gassprisene å synke helt frem til november. Fra metrologisk institutt sine hjemmesider kan vi lese at året 2009 var et av de varmeste årene globalt siden 1850 (Meteorologisk, 2009). Både den ekstra ordinære lave temperaturen i 2009 og ettervirkningen etter finanskrisen kan ha gjort sitt for at gassprisene falt i 2009 og holdt seg lave til langt utpå høsten. Derimot ble desember måned i 2009 en ekstra kald måned som ble etterfulgt av en kald vinter i 2010 (Meteorologisk, 2010). Det ser ut som den kalde vinteren gjorde sitt for at etterspørselen etter gass økte, og dertil økte prisene.

Etter vinteren sank prisene i 2010, se figur 5.1a). Deretter begynte naturgassprisene å øke igjen i mai/juni. Dette er ganske uvanlig siden vi forventer at gassprisene synker på sommeren. Det kan vær mange årsaker til denne uvanlige økningen i gassprisene på de europeiske knutepunktene. Blant annet var det uvanlig kaldt på en tid av året da gasssesongen skifter fra vinter til sommer sesong, samtidig som det var tilsvarende produksjonskutt fra Norge i påvente av lavere etterspørsel. Norge kunne ikke levere nok naturgass til EU og hadde flere forsyningsavbrudd i denne perioden. Virkningen av disse norske forsyningsavbruddene av gass til det europeiske markedet, ga høyere naturgasspriser i Europa (Oettinger, 2010). Naturgassprisene fortsatte å øke utover høsten i 2010.

Vinteren i 2011 hadde først høye priser, men begynte å synke for deretter å øke igjen i slutten av første kvartal i 2011. Grunnen til økningen i slutten av første kvartal 2011 kan være at var uro i midtøsten og i nord Afrika. Uroen førte til blokkeringer av Suezkanalen. Suezkanalen er en viktig forsyningsrute for den flytende naturgassen, LNG. Samtidig med uroen i midtøsten var det en tsunami i Japan den 11. mars som medførte et alvorlig kjernekraftbrudd i Japan. Dette kan være grunnen til at ble mer volatilitet i naturgassprisene (Raytchev, 2011). Selv om 2011 var det året som hadde det laveste registrerte naturgassforbruket i Europa siden år 2000, var

prisene ganske høye. Det kan også tyde på at det lave forbruket av naturgass, kan ha kommet som en effekt av de høye naturgassprisene.

Fra figur 5.1a) ser vi at i begynnelsen av februar i 2012 var prisene ekstrem volatile, men prisene var tilbake til mer normale nivåer i midten av februar. En grunn til denne volatiliteten i prisene er en uventet økt etterspørsel etter naturgass. I løpet av mars i 2013 økte gassprisene betydelig. Grunnen til prisøkningen, var et betydelig fall i temperaturen i mars, noe som var uventet i gassmarkedet. I samme periode var det lite tilgjengelig naturgass. Det var lav LNG import og avbrudd i forsyninger fra Norge (EU-kommisjonen, 2013). Dette førte igjen til en raskt tapping av naturgasslagrene. Alt dette medførte høye gasspriser perioder.

Hvis vi ser bort fra de kortvarige prisendringene som skiller seg ut, har det vært en jevn økning i prisen fra 2010 til slutten av 2013. Grunnen til denne økningen kan være den generelle økonomiske veksten, den økende oljeprisen, og nedgangen i import av LNG.

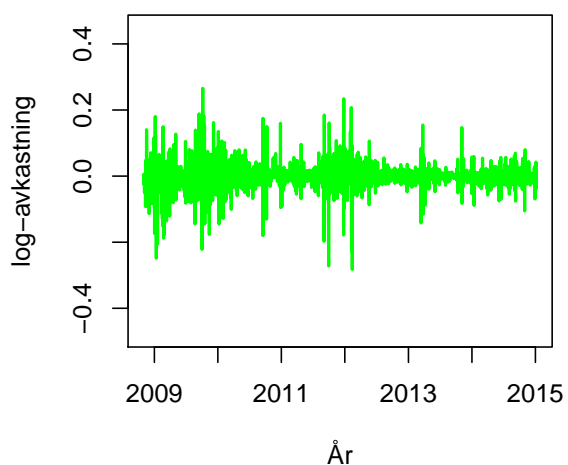
Fra figur 5.1a) ser vi at i 2014 sank gassprisen ned mot nivåer som var mer vanlige i 2011. Grunner til dette kan være at det i fjerde kvartal i 2014, var det historiske høye nivåer av naturgasslagring, samtidig som det var mildvær (EU-kommisjonen, 2014c). En annen grunn til de lave prisene kan være at det er blitt økt import av LNG noe som har gjort at prisforskjellen mellom Asia og Europa nesten har forsvunnet (EU-kommisjonen, 2014c). For det fjerde overtok det nederlandske TTF markedet, Storbritannias NBP (EU-kommisjonen, 2014c). Dette økte likviditeten i de europeiske knutepunktene i 2014. En femte grunn kan være nedgangen i råoljeprisen. Fallet i råoljeprisen i 2014 vil påvirke oljeprisindekserte kontrakter fra 2015. Videre vil nedgangen i råoljeprisen føre til en forventning om at gassprisen faller ytterligere utover i 2015 (EU-kommisjonen, 2014c).

5.1.2 Omforme gasspriser til log-avkastninger

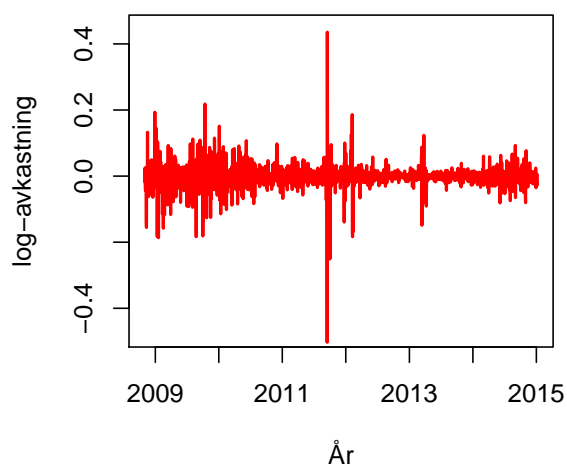
Som i de fleste finans og økonomiske tidsrekker ser vi fra figur 5.1a) at gassprisene ikke er stasjonære. Vi ønsker dermed å omforme gasspriser til log-avkastninger. For å finne log-avkastningen tar vi først den naturlige logaritmen av dataene, deretter tar vi differansen mellom dataene. Fra ligning (2.4) har vi at log-avkastningene er gitt som $X_t = \log P_t - \log P_{t-1}$.

Plottene av log-avkastningene for de ulike naturgassprisene er vist i figur 5.2. Vi ser fra figur 5.2 at det er noen forskjeller mellom log-avkastningene i de europeiske gassmarkedene. Det er tydelig at volatiliteten var høyere i gassprisene de første årene, enn den har vært etter 2012. En grunn til at svingningene i naturgassprisene var større før, kan være på grunn av den økonomiske situasjonen i Europa. Vi legger også merke til at det ser ut som log-avkastningene har variert mer på vinteren i 2012. Det var spesielt mye variasjoner i NCG markedet, samtidig som TTF markedet hadde noen ekstreme prisendringer på høsten i 2011. Som vi tidligere har nevnt var det uventete lave temperaturer og lite tilgjengelig naturgass i mars i 2013. I figur 5.2 ser man tydelig at svingningene var høyere i mars i 2013. Til slutt observerer vi at det er mer variasjon i log-avkastningene i slutten av 2014, enn det var i starten. Det kan virke som om log-avkastningene ikke er stasjonære over hele perioden, men den er delt i regimer. Det vil si at det kan se ut som vi har stasjonærhet mellom år 2009 og år 2011, og fra år 2011 til år 2015. En regime-switchende

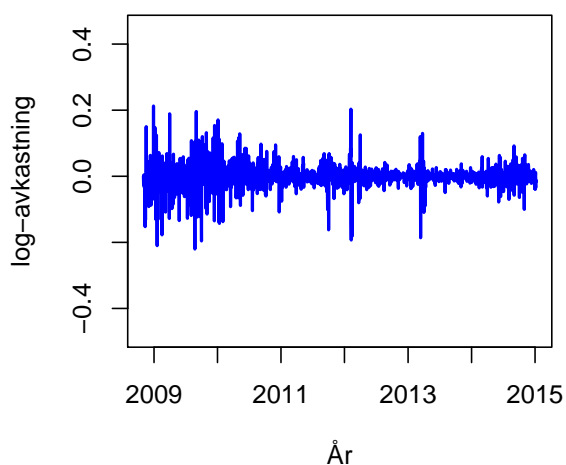
modell kunne vært brukt, for å fått med disse regimene. I denne oppgaven har vi valgt å se bort fra disse regimene. I neste delkapittel skal vi kjøre noen tester på log-avkastningene for å teste om de er stasjonære.



(a) Log-avkastinger i NCG markedet



(b) Log-avkastinger i TTF markedet



(c) Log-avkastinger i Zeebrugge markedet

Figur 5.2: Log-avkastninger for naturgasspriser

5.2 Fakta om log-avkastningene

I denne delen skal vi analysere log-avkastningene til naturgassprisene. Det første vi skal gjøre er å teste om log-avkastningene er stasjonære. Log-avkastningene i NCG, TTF og Zeebrugge markedet har følgende notasjon, X_t^{NCG} , X_t^{TTF} og X_t^{ZEE} .

5.2.1 Tester om log-avkastningene er stasjonære

En forutsetning for å bruke for eksempel en GARCH modell er at vi har en stasjonær tidsrekke. Fra figur 5.2 observerte vi at det virket litt ikke-stasjonært ut over hele perioden, men i ulike regimer virket log-avkastningene stasjonære. Samtidig ser vi at log-avkastningene svinger rundt et konstant nivå. En annen måte å teste om tidsrekkene er stasjonære på, er å teste for enhetsrøtter. Hvis vi har en enhetsrot betyr det at tidsrekken ikke er stasjonær og vi må omforme dataene. To eksempler på enhetsrottester er Phillips-Perron og ADF-test. Både i Phillips-Perron og i ADF-testen er nullhypotesen at tidsrekken inneholder en enhetsrot med en alternativ hypotese om stasjonæritet. P-verdiene er definert som: *)p-verdi $\leq 0,05$, **)p-verdi $\leq 0,01$, ***) p-verdi $\leq 0,001$, ****) p-verdi $\leq 0,0001$. Denne p-verdi notasjonen er også brukt senere i oppgaven. Fra tabell 5.1 ser vi at på et signifikansnivå på 0,05 forkaster vi hypotesen om enhetsrøtter, og dermed antar at log-avkastningene er stasjonære over hele tidsperioden.

Tabell 5.1: Test-verdiene til testene på log-avkastningene.

Marked	Test-verdi	
	Phillips-Perron	ADF
NCG	-53,20 *	-14,98 **
TTF	-49,36 *	-15,29 **
ZEE	-43,63 *	-15.62 **

5.2.2 Egenskaper til log-avkastningene

I tabell 5.2 har vi en oversikt over noen egenskaper for log-avkastningene for markedene NCG, TTF og Zeebrugge sine naturgasspriser. Vi ser at gjennomsnittet til log-avkastningene er negativ for alle tre log-avkastningene, men likevel er den nære null. Det at gjennomsnittet til log-avkastningen er nære null er ikke overraskende, men at gjennomsnittet over mange år er negativ er uventet. Den daglige gjennomsnitt volatiliteten i NCG markedet er 3,7 prosent, mens for TTF og Zeebrugge er volatiliteten henholdsvis på 3,3 og 3,2 prosent. Den høyeste prisøkningen for en dag til en annen dag har vi TTF markedet med 43 prosent, følgende kommer NCG og Zeebrugge med henholdsvis 27 og 21 prosent prisøkning.

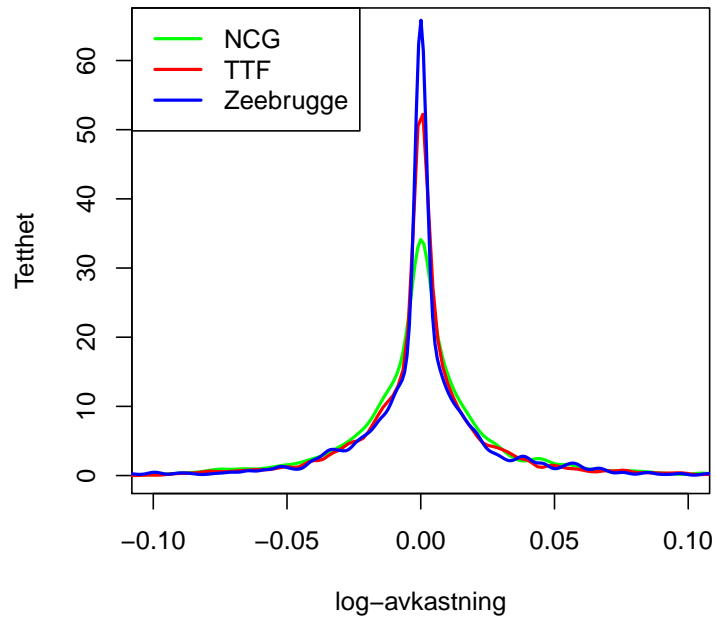
Tabell 5.2: Egenskaper til log-avkastningene

Log-avkastning	Gjennomsnitt	Standardavvik	Min	Max	Kurtosis	Skewness
NCG	$-1,83 \times 10^{-4}$	0,037	-0,28	0,27	11,97	-0,38
TTF	$-2,01 \times 10^{-4}$	0,033	-0,50	0,43	45,57	-1,00
ZEE	$-2,01 \times 10^{-4}$	0,032	-0,22	0,21	11,55	-0,43

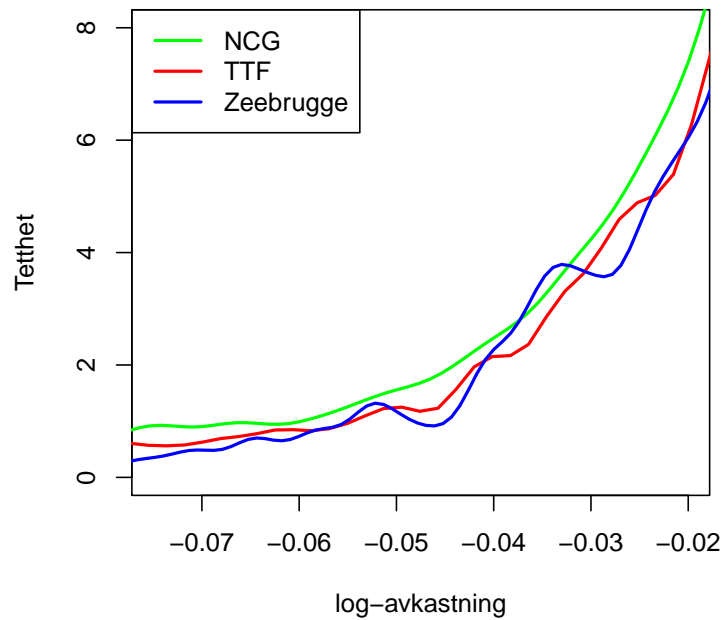
Fordelingene til log-avkastningene i markedene NCG, TTF og Zeebrugge har kurtosis på henholdsvis 11,97, 45,57 og 11,55. Kurtosis på høyere enn 3 indikerer på at fordelingene har en brattere topp og en tyngre hale enn normalfordeling. Det vil si at det er veldig lite spredning i dataene. I tillegg har log-avkastningene negativ skewness. Negativ skewness indikerer på at halen på venstre side av sannsynlighetstetthetsfunksjonen er lengre eller feitere enn høyresiden. Dermed kan vi anta fra skewness og kurtosis til log-avkastningene at vi ikke har normalfordelte data. I figur 5.4d) er det et kvantil-til-kvantil plot (QQ-plot) for log-avkastningene mot normalfordelingen, det er opplagt at log-avkastningene ikke er normalfordelt.

I figur 5.3 ser vi den empiriske tettheten til log-avkastningene, her observerer vi også at fordelingene har bratte topper og veldig lite spredning i dataene. Vi skal være litt forsiktige med å anta at disse ulike toppene har noe å si, grunnen til dette er at vi har gjort en ikke-parametrisk estimering av modellen. På grunn av lite spredning kan vi anta at vi skal bruke en student t-fordeling. I figur 5.3b fokuserer vi på den venstre halen. Siden dette er en ikke-parametrisk modell, må vi sjekke at det faktisk er nok datapunkter i enden av halen. I NCG-markedet er det 68 datapunkter som er mindre enn -0,07, samtidig som det er 289 datapunkter mellom -0,02 og -0,07. Tilsvarende i TTF-markedet er det forholdsvis 47 og 241 datapunkter. Til slutt i Zeebrugge markedet er det henholdsvis 55 og 241 datapunkter. I den venstre halen ser vi at fordelingene til de ulike log-avkastningene er ganske jevne. Slik at vi kan anta at en student t-fordeling kan beskrive alle log-avkastningene ganske godt.

Videre har vi også sammenlignet fordelingene til log-avkastningene i de ulike markedene ved bruk av QQ-plot. De tre første QQ-plottene i figur 5.4 viser de tre log-avkastningene opp mot hverandre. Vi ser i figur 5.4a at $NCG \sim TTF$ og $TTF \sim ZEE$, altså at de har lignende fordeling mot hverandre. Vanligvis ville dette ført til at $NCG \sim ZEE$, men dette er ikke tilfellet her. Det betyr at de ikke er transitive.

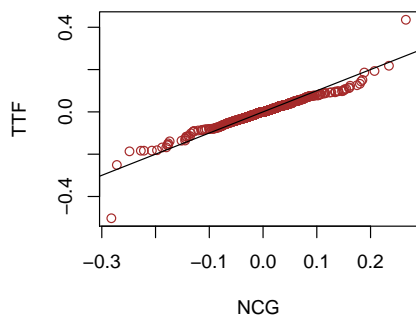


(a) Empirisk tetthet til log-avkastningene

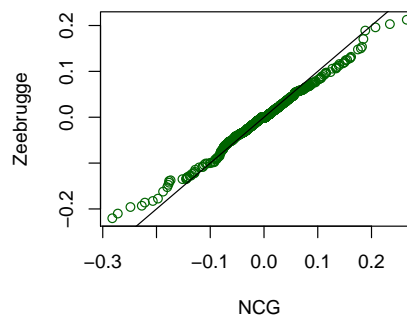


(b) Venstre hale

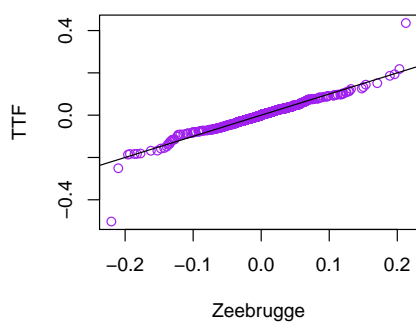
Figur 5.3: Empirisk tetthet til log-avkastningene



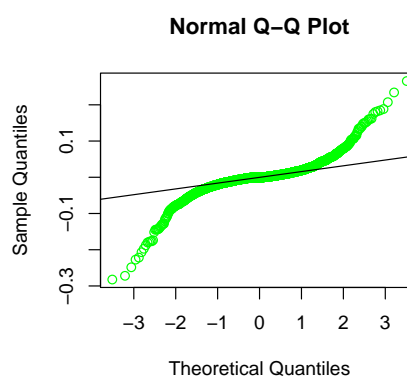
(a) NCG mot TTF



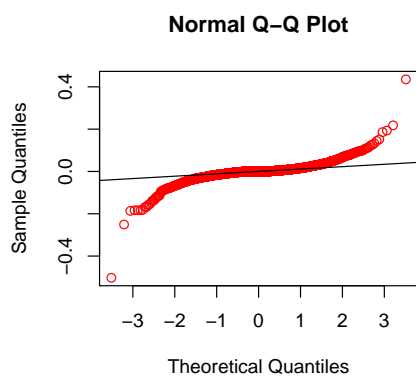
(b) NCG mot Zeebrugge



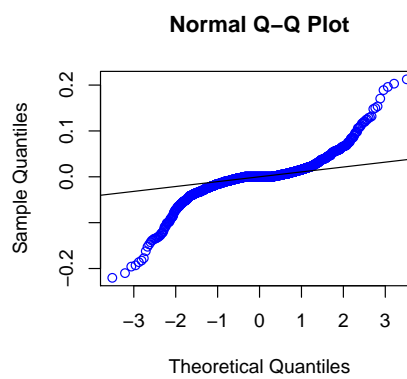
(c) Zeebrugge mot TTF



(d) NCG mot normalfordeling



(e) TTF mot normalfordelingen

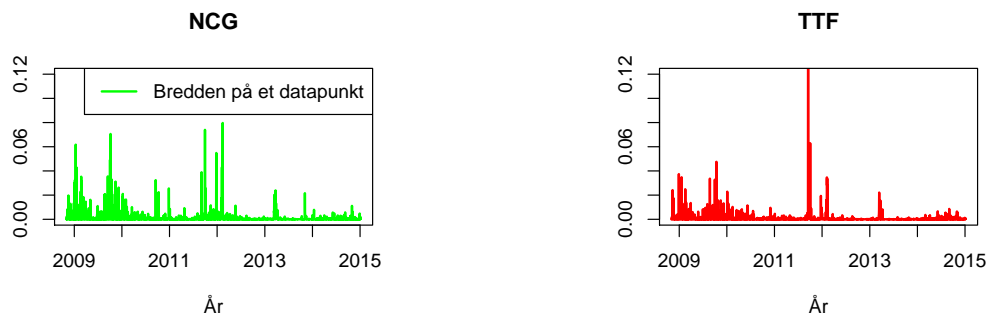


(f) ZEE mot normalfordeling

Figur 5.4: QQ-plot av log-avkastningene

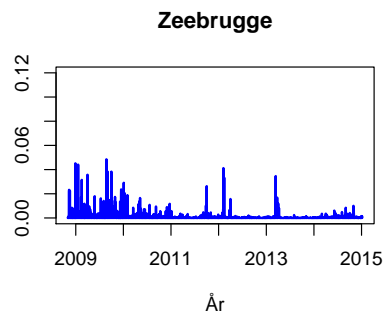
5.2.3 Volatilitet i log-avkastningene

Når vi ser på de kvadrerte log-avkastningene i figurene i 5.6, ser det ut som vi har volatilitet i dataene. Med det menes det at det kan se ut som perioder med høy absoluttverdi av avkastningene blir etterfulgt av perioder med høy avkastning. Fra figurene i 5.6 kan det virke som at volatiliteten er størst i vintersesongen. Spesielt i perioder med uventete lave temperaturer som i februar i 2012 og i mars i 2013. De største prisendringene ser vi i september TTF markedet. Da falt prisen fra 2,48 cent/kWh til 1,50 cent/kWh den 17. September til den gikk opp igjen til 2,32 cent/kWh dagen etter igjen. Det kan også være at det er en feil innskrivning i datasettet. Vi legger også merke til at prisene var mer volatile i 2009 enn i de andre årene, som igjen kan bety at det ikke er helt stasjonært. En grunn til at det var mer volatilitet i 2009, kan være fordi det var økonomisk usikkerhet i Europa som følge av finanskrisen. Det beste hadde vært å tilpasse en GARCH modell som tillater regime endringer, men dette skal ikke vi gjøre i denne oppgaven. I neste delkapittel skal vi tilpasse enkle ARCH/GARCH modeller som kan brukes for å modellere volatilitet.



(a) Den kvadrerte NCG-log-avkastningen

(b) Den kvadrerte TTF-log-avkastningen



(c) Den kvadrerte Zeebrugge-log-avkastningen

Figur 5.5: De kvadrerte log-avkastningen

5.3 Bygge en volatilitetsmodell

For å bygge en volatilitetsmodell trengs det fire steg (Tsay, 2005, s. 101):

- 1: Først skal vi tilpasse en lineær modell for log-avkastningene. Vi kan tilpasse en lineær modell ved å teste for seriell autokorrelasjon i dataene. Videre kan en finne en lineær modell som for eksempel en ARMA modell for log-avkastningsdataene. En lineær modell kan fjerne denne serielle autokorrelasjonen, hvis det er noen.
- 2: Bruk enten residualene fra den lineære modellen, eller hvis det ikke var noe autokorrelasjon kan vi bruke log-avkastningene direkte for å teste for ARCH effekter.
- 3: Hvis ARCH effektene er statistisk signifikante, tilpasser vi en volatilitetsmodell for residualene/log-avkastningen. Deretter lager vi en estimasjon av den lineære modellen og volatilitetsmodellen.
- 4: Til slutt tester vi om den tilpassede modellen passer, og avgrensner modellen om det er nødvendig.

5.3.1 Tilpasning av lineære modeller for log-avkastningene

Først må vi tilpasse en lineær modell, hvis det er seriell korrelasjon i log-avkastningene. En test som kan teste for seriell korrelasjon er Ljung-Box. Tabell 5.3 viser resultatene fra Ljung-Box testene. Vi ser at alle log-avkastningene inneholder seriell korrelasjon bortsett fra TTF.

Tabell 5.3: Log-avkastninger med testverdier og tilhørende p-verdier

Marked	Ljung-Box					
	Q1		Q5		Q10	
NCG	21,31	$3,9 \times 10^{-6}$	36,02	$9,4 \times 10^{-7}$	58,33	$7,5 \times 10^{-9}$
TTF	2,53	0,11	11,90	0,036	20,81	0,022
ZEE	15,00	$1,0 \times 10^{-4}$	39,90	$3,9 \times 10^{-6}$	55,92	$2,1 \times 10^{-8}$

Siden log-avkastningene er stasjonære kan vi anta at den lineære modellen kan bli modulert som en ARMA-modell. For å finne ut om vi skal ha en AR-modell, MA-modell eller en ARMA-modell og hvor mange ledd vi skal ha med i modellen, må vi se på den serielle autokorrelasjonen mellom log-avkastningene. Forskjellen er at en AR-modell inkluderer laggede ledd av tidsrekken selv, mens en MA-modell inkluderer laggede ledd av støyen eller residualene.

Tilpasning av en lineær modell for NCG markedet

Vi fant i tabell 5.3 at vi hadde seriell korrelasjon i log-avkastningene i NCG markedet. I tillegg kan vi bruke et ACF-bånd for å teste om det er noe signifikant seriell autokorrelasjon på fem og ti prosentnivå. Ut fra ACF-testen i figur 5.6a) har vi signifikant seriell autokorrelasjon på fem prosentnivå ved flere lag, men denne serielle korrelasjonen er under ti prosentnivå. Dette kan medføre at det kanskje ikke er nødvendig å tilpasse en ARMA modell. Likevel ønsker vi først å prøve å tilpasse en lineær modell.

For å få en ide om hvor mange ledd som skal være med i den lineære modellen kan vi bruke ACF og PACF på log-avkastningene. Ved fem prosentnivå i figur 5.6a) har vi signifikant ACF ved lag 1,4,6,7,10 og 11. For å ikke gjøre modellen for komplisert velger vi å se bort fra lag-ene som er fire eller mer. En MA-modell tilpasses ved å se på hvor vi har ACF kutt (Tsay, 2005, s. 32-56), dette vil gi oss en MA(1)-modell. I tillegg skal PACF være synkende, der den er høyest ved lag 1.

Nederst i figur 5.6a) er det et plott over partial ACF, der har vi PACF-kutt etter første lag i tillegg som den også er signifikant ved lag 4,6,10,11. For å tilpasse antall ledd i en AR-modell ser vi på hvilken lag vi har PACF kutt (Tsay, 2005, s. 32), det gir oss en AR(1)-modell. Samtidig skal en AR-modell ha synkende ACFer.

Fra figur 5.6a ser det hverken ut som log-avkastningene kan moduleres som en ren AR-modell eller en ren MA-modell, fordi hverken ACF eller PACF er synkende. Oppførselen til ACF og PACF minner derfor mer om en ARMA-modell. Siden både ACF og PACF kutter ved første lag kan vi tilpasse en ARMA (1, 1)-modell.

Vi har fra ligning (2.8) at ARMA(1, 1) prosessen er gitt som:

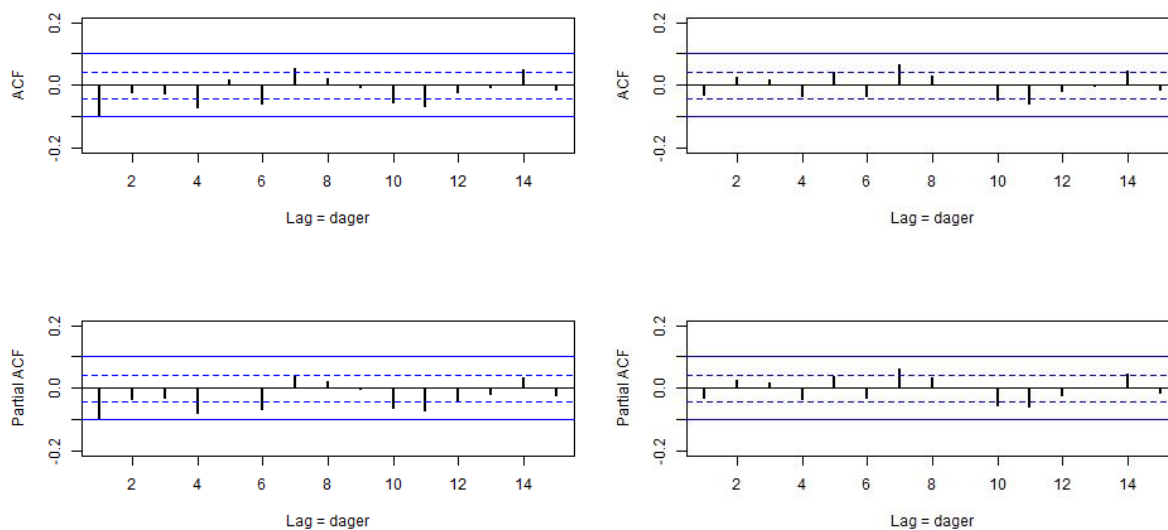
$$X_t = Z_t + \theta Z_{t-1} + \phi X_{t-1}, \quad (5.1)$$

der X_t er log-avkastningen, Z_t er hvitstøy og ϕ og θ er henholdsvis koeffisientene i AR-leddet og i MA-leddet.

For å tilpasse en ARMA(1, 1) modell til NCG log-avkastningene bruker vi funksjonen *arma* fra pakken *tseries* i programmet R. Vi får at koeffisientene i ligning (5.1) er $\hat{\phi} = 0,79$ og $\hat{\theta} = -0,87$, der begge koeffisientene er signifikant forskjellig fra null. Dette trenger nødvendigvis ikke å bety at modellen er bra, fordi når n er stor nok kan vi få en signifikant ARMA modell uten særlig prediksjonskraft. Selv om koeffisientene er statistisk signifikante, kan forklaringskraften være veldig liten. ARMA(1, 1) modellen er nært på å ha en felles rot i AR og MA delen slik at

$$\text{ARMA11}(0, 79, -0, 87) \approx \text{ARMA}(0, 0),$$

det medfører at modellen ikke er langt fra hvit støy. Det kan dermed være at det ikke er nødvendig å tilpasse en ARMA-modell.



(a) ACF og PACF av NCG log-avkastninger (b) ACF og PACF av NCG ARMA(1,1)-residualene

Figur 5.6: ACF og PACF (NCG)

Samtidig viser figur 5.6b) ACF og PACF til ARMA(1,1)-residualene. I figur 5.6b) ser vi at vi får noe mindre korrelasjon ved å bruke ARMA(1,1) residualene istedenfor NCG log-avkastningene direkte. Det er spesielt det første laget som har fått mindre autokorrelasjon, men det er fremdeles noe avhengighet ved fem prosentnivå ellers. Forskjellene mellom ARMA-residualene og log-avkastningene er minimale. Ljung-Box test på residualene, se tabell 5.4, viser at den serielle korrelasjonen på de første lag-ene er fjernet.

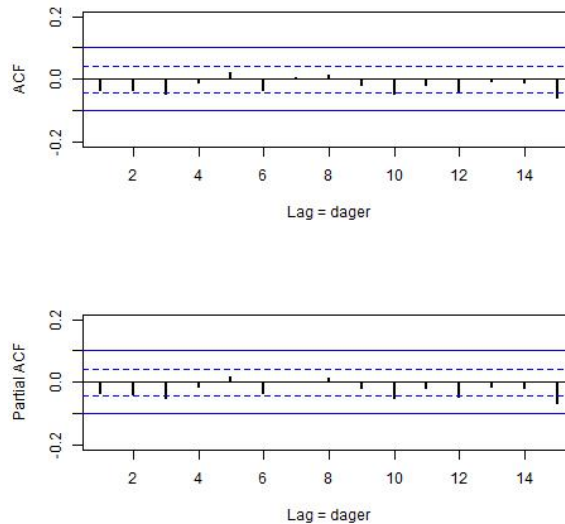
Tabell 5.4: ARMA(1,1)-residualer med testverdier og tilhørende p-verdier

Marked	Ljung-Box					
	Q1		Q5		Q10	
NCG	1,96	0,16	10,85	0,05	30,63	$6,7 \times 10^{-5}$

Siden den serielle korrelasjonen for log-avkastningene fra før av var relativt liten, og vi ønsker å gjøre modellen enklest mulig velger vi å bruke NCG log-avkastningene direkte i den videre analysen. Dermed ser vi bort fra ARMA-residualene.

Tilpasning av en lineær modell for TTF markedet

En ACF og PACF test over TTF sine log-avkastninger er vist i figur 5.7a). I figur 5.7a) og i tabell 5.3 ser vi at den serielle korrelasjon til TTF sine log-avkastninger er svært liten, og vi velger dermed å ikke gå videre med en ARMA tilpasning.



(a) ACF og PACF av TTF log-avkastninger

Figur 5.7: ACF og PACF (TTF)

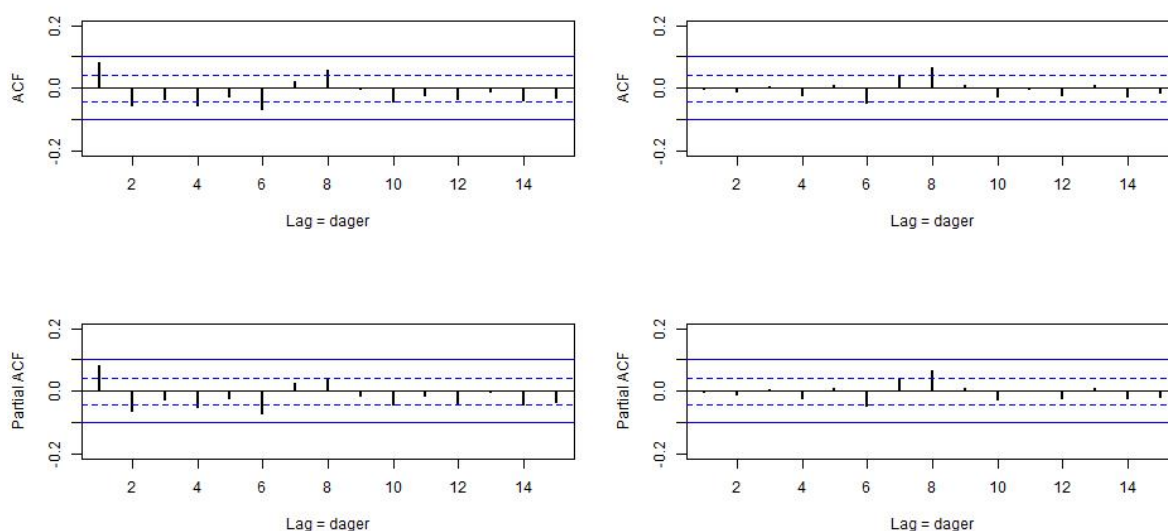
Tilpasning av en lineær modell for Zeebrugge markedet

Det samme er gjort for Zeebrugge sine log-avkastninger. Figur 5.8a) viser at det er litt mer seriell autokorrelasjon på fem prosentnivå enn det som var i de andre log-avkastningene, men heller ikke her er det noe på ti prosentnivå. Ut ifra ACF og PACF av log-avkastningene i figur 5.8a) ser det verken ut er en ren AR-modell eller en ren MA-modell. Som vi har sett i de forrige log-avkastningene minner også oppførselen til ACF og PACF over Zeebrugge sine log-avkastninger mest om en ARMA-modell. Siden at både ACF og PACF kutter ved andre lag kan vi bruke en ARMA (2, 2)-modell. En ARMA(2, 2) prosess er gitt som:

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \phi_1 X_{t-1} + \phi_2 X_{t-2}. \quad (5.2)$$

Vi får at koeffisientene i ligning (5.2) er $\hat{\phi}_1 = 0,86$, $\hat{\phi}_2 = 0,01$, $\hat{\theta}_1 = -0,79$ og $\hat{\theta}_2 = -0,13$, der $\hat{\phi}_1$ og $\hat{\theta}_1$ er signifikant forskjellig fra null. Resten av koeffisientene er ikke signifikant ulik null. Vi tilpasser derfor en ARMA(1, 1)-modell. Ved en ARMA(1, 1) tilpasning får at koeffisientene i ligning (5.1) er $\hat{\phi} = -0,53$ og $\hat{\theta} = 0,61$, der begge koeffisientene er signifikant forskjellig fra null.

Også her er ARMA modellen nært på å ha felles rot med motsatt fortegn i AR og MA delen. Dermed er ARMA modellen nært på hvit støy.



(a) ACF og PACF av Zeebrugge log-avkastninger (b) ACF og PACF av Zeebrugge residualene

Figur 5.8: ACF og PACF (Zeebrugge)

Figur 5.8b) viser ACF og PACF til ARMA(1, 1)-residualene. Etter å ha tilpasset en ARMA(1, 1) modell for Zeebrugge sine log-avkastninger fikk vi fjernet den serielle korrelasjonen for de to første lagene. I tabell 5.5 viser resultatene fra Ljung-Box test på arma-residualene. Tabellen viser at vi har fått fjernet en del av den serielle korrelasjonen som var. Samtidig kan det se ut som det ikke er nødvendig å tilpasse en ARMA-modell, siden den serielle korrelasjonen i utgangspunktet var veldig liten.

Tabell 5.5: ARMA(1, 1)-residualer med testverdier og tilhørende p-verdier

Marked	Ljung-Box					
	Q1		Q5		Q10	
ZEE	0,02	0,89	11,99	0,04	31,71	$4,5 \times 10^{-5}$

5.3.2 Teste for ARCH effekter

Selv om Ljung-Box test på log-avkastningene viser at vi ikke har noe seriell korrelasjon, betyr ikke det at vi ikke har noe volatilitet. For at vi ikke skal ha volatilitet må vi beholde hypotesen til Ljung-Box test om ingen seriell korrelasjon på de kvadrerte log-avkastningene eller eventuelt de kvadrerte residualane. I tillegg kan en ARCH-LM test brukes for å teste om vi har heteroskedastisitet i log-avkastningene eller residualene.

Vi bruker Ljung-Box testen på de kvadrerte log-avkastningene $\{(X_t^{\text{NCG}})^2\}$ i NCG markedet, de kvadrerte log-avkastningene $\{(X_t^{\text{TTF}})^2\}$ i TTF markedet og de kvadrerte log-avkastningene $\{(X_t^{\text{ZEE}})^2\}$ i Zeebrugge markedet. I tillegg bruker vi LM ARCH test på $\{X_t^{\text{NCG}}\}$, $\{X_t^{\text{TTF}}\}$ og $\{X_t^{\text{ZEE}}\}$ for å sjekke om vi har ARCH effekter. Resultatene er vist i tabell 5.6. Både Ljung-Box testen og ARCH-LM testen viser sterke ARCH effekter for samtlige log-avkastninger med p-verdier som er nære null. Dermed må vi tilpasse ARCH/GARCH-modeller for de videre analysene.

Tabell 5.6: Kvadrerte log-avkastninger med testverdier. Alle p-verdiene er under 0.5%.

	Ljung-Box				ARCH-LM			
Marked	Q1	Q5	Q10	Q15	LM1	LM5	LM10	LM15
NCG	60,64	267,83	413,52	435,91	60,5	220,78	289,46	291,96
TTF	396,75	401,51	408,82	430,05	396,05	483,31	484,20	497,65
ZEE	66,76	162,37	367,44	439,54	66,05	127,34	228,34	250,72

5.3.3 Tilpasning av volatilitetsmodellene

Fra ligningene (3.1) og (3.2) har vi at ARCH-modellen er gitt som:

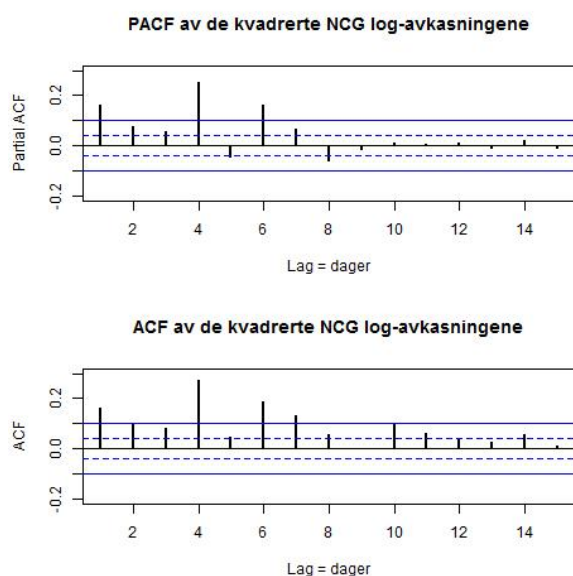
$$X_t = \sigma_t Z_t.$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2.$$

GARCH-modell for NCG sine log-avkastninger

Først tilpasser vi en volatilitetsmodell for NCG sine log-avkastninger $\{X_t^{\text{NCG}}\}$. Vi kan bruke PACF på X_t^2 for å finne ut hvor mange ledd som skal være med i ARCH modellen (Tsay, 2005,

s. 106). Figur 5.9 indikerer på at vi skal ha med de 8 første leddene i en ARCH modell. Derimot gir ordensbestemmelsen i ARCH(q) modellen at $q = 3$, på grunn av at estimatene $\hat{\alpha}_4, \dots, \hat{\alpha}_8$ ikke er statistisk signifikante. Tidligere i oppgaven observerte vi at log-avkastningene kunne være student t-fordelt. Vi tilpasser dermed en ARCH(3) modell for NCG-log-avkastningene og antar at Z_t er standardisert student-t-fordelt med $\nu = 3$ frihetsgrader.



Figur 5.9: NCG, PACF og ACF av de kvadrerte log-avkastningene

De estimerte koeffisientene i ARCH(3)-modellen er gitt i tabell 5.7.

Tabell 5.7: Estimaten til NCG-ARCH(3)-modellen

	Estimat	Signifikant
α_0	$3,32 \times 10^{-4}$	***
α_1	0,56	***
α_2	0,18	***
α_3	0,16	***

Hvis vi kun hadde hatt en ARCH effekt ville vi sett at prosessen $\{(X_t^{\text{NCG}})^2\}$ kun hadde hatt seriell avhengighet (PACF) og ingen seriell korrelasjon (ACF), men i $\{(X_t^{\text{NCG}})^2\}$ ser vi at vi har både seriell avhengighet og seriell korrelasjon. Slik at en ARCH-modell generelt sett ikke er god nok, vi må derfor bruke en GARCH-modell. Fra ligning (3.21) har vi at GARCH-modellen er gitt som:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-1}^2 + \sum_{j=1}^p \beta_j \sigma_{t-1}^2.$$

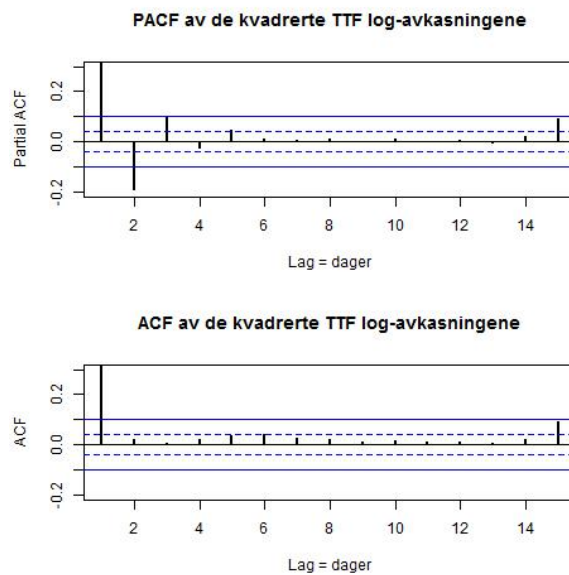
Estimatene for en GARCH(1,1) modell der Z_t er student-t-fordelt med frihetsgradsparameter, $\nu = 3$ er gitt i tabell 5.8.

Tabell 5.8: Estimatene til NCG-GARCH(1,1)-modellen

	Estimat	Signifikant
α_0	$3,57 \times 10^{-5}$	***
α_1	0,24	***
β_1	0,77	***

GARCH-modell for TTF sine log-avkastninger

På tilsvarende måte som vi gjorde for å tilpasse en ARCH-modell for NCG, indikerer figur 5.10 at log-avkastningene for TTF sine gasspriser kan modelleres som en ARCH(3) modell, hvor Z_t er standardisert student-t-fordelt med $\nu = 3$.



Figur 5.10: TTF, PACF og ACF av de kvadrerte log-avkastningene

Følgelig er estimatene i modellen gitt i tabell 5.9.

Tabell 5.9: Estimaten til TTF-ARCH(3)-modellen.

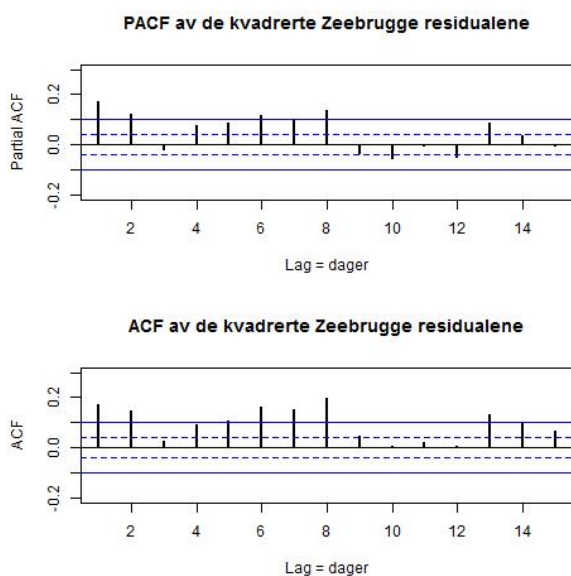
	Estimat	Signifikant
α_0	$2,38 \times 10^{-4}$	***
α_1	0,61	***
α_2	0,07	**
α_3	0,09	**

Deretter tilpasser vi en GARCH(1,1) modell, der Z_t er antatt å være student-t-fordelt med $\nu = 3$ for log-avkastningene i TTF markedet. Dette gir følgende estimater til modellen:

Tabell 5.10: Estimaten til TTF-GARCH(1,1)-modellen

	Estimat	Signifikant
α_0	$1,54 \times 10^{-5}$	***
α_1	0,16	***
β_1	0,83	***

GARCH-modell for Zeebrugge sine log-avkastningene



Figur 5.11: Zeebrugge, PACF og ACF av de kvadrerte log-avkastningene

Figur 5.11 på log-avkastningene i Zeebrugge markedet indikerer på at log-avkastningene kan modelleres som en ARCH(8)-modell. Vi antar nå at Z_t er student-t-fordelt med $\nu = 2,5$ fri-

hetsgrader og tilpasser en ARCH(8). I ARCH(8)-modellen er det kun ARCH(1)-leddet som er statistisk signifikant. Estimaten til ARCH(1)-modellen er gitt under:

Tabell 5.11: Estimaten til Zeebrugge-ARCH(1)-modellen

	Estimat	Signifikant
α_0	$4,14 \times 10^{-4}$	***
α_1	0,87	***

Videre tilpasser vi en GARCH(1, 1)-modell der vi fremdeles antar at Z_t er student-t-fordelt med $\nu = 2, 5$. Koeffisientene i GARCH(1, 1)-modellen er gitt som:

Tabell 5.12: Estimaten til Zeebrugge-GARCH(1, 1)-modellen

	Estimat	Signifikant
α_0	$1,58 \times 10^{-5}$	***
α_1	0,15	***
β_1	0,87	***

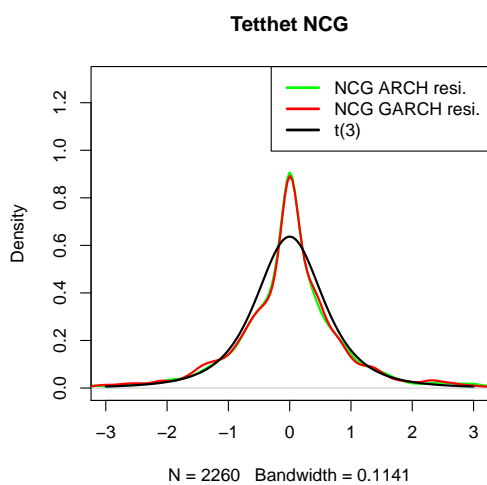
5.3.4 Modelltesting og valg av modell

I ARCH(3)-modellen og i GARCH(1, 1) modellen for NCG og TTF markedet, var alle parametrene statistisk signifikante. En test-observator som kan brukes for å velge hvilken modell som passer best er informasjonskriteriet AIC (Brockwell and Davis, 1996, s. 171). Den modellen som har minst AIC-verdi, er den modellen som passer best. Det er nødvendigvis ikke en god modell, men det er den modellen som er den beste av de som blir testet. For NCG markedet var AIC = -4,51 for ARCH(3)-modellen, og i GARCH(1, 1) var AIC = -4,55. AIC er minst i GARCH(1, 1). Samtidig for TTF var AIC = -4,94 for ARCH(3)-modellen, og i GARCH(1, 1) var AIC = -5,00. I ARCH(1)-modellen og i GARCH(1, 1) modellen for Zeebrugge markedet, var alle parametrene statistisk signifikante. I ARCH(1) modellen var AIC = -4,88, og i GARCH(1, 1)-modellen var AIC = -4,97. AIC er minst i GARCH(1, 1)-modellene for samtlige av naturgassmarkedene.

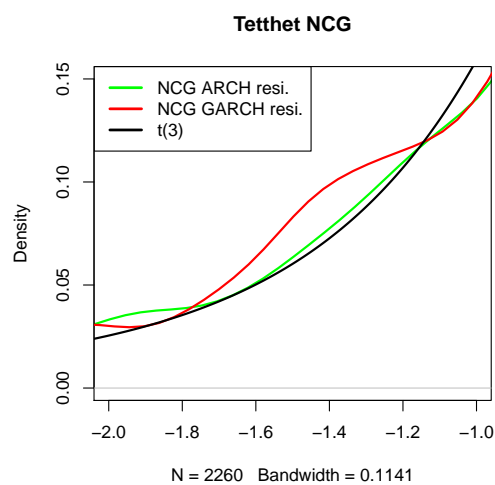
Tabell 5.13: Beste modell ved AIC

Marked	Modell	AIC
NCG	ARCH(3)	-4,51
	GARCH(1,1)	-4,55
TTF	ARCH(3)	-4,94
	GARCH(1,1)	-5,00
ZEE	ARCH(1)	-4,88
	GARCH(1,1)	-4,97

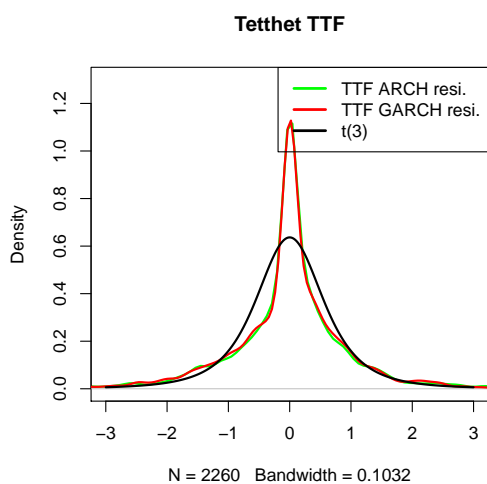
For en riktig spesifisert GARCH-modell skal de empiriske residualene, $\hat{Z}_t = X_t/\sigma_t$, ha samme fordeling som den valgte standardiserte t-fordelingen.



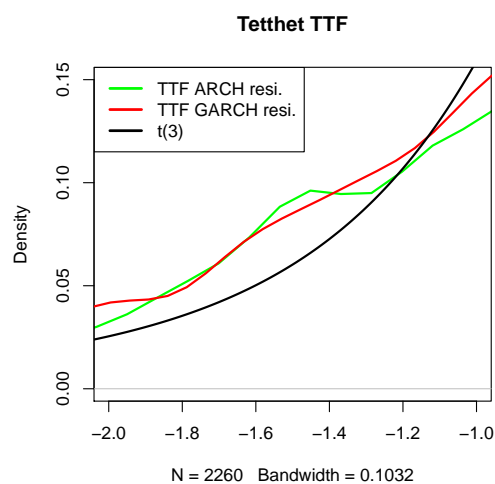
(a) NCG markedet



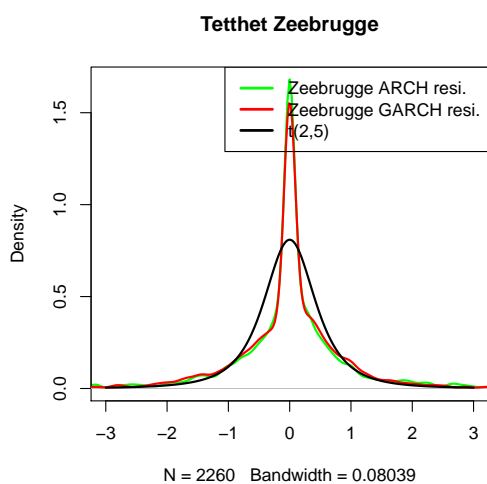
(b) Venstre hale i NCG



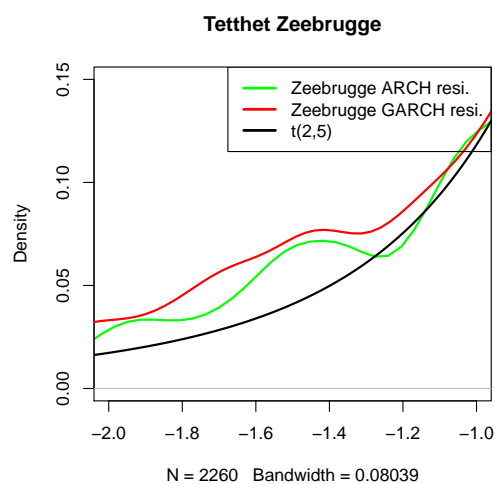
(c) TTF markedet



(d) Venstre hale i TTF



(e) Zeebrugge markedet

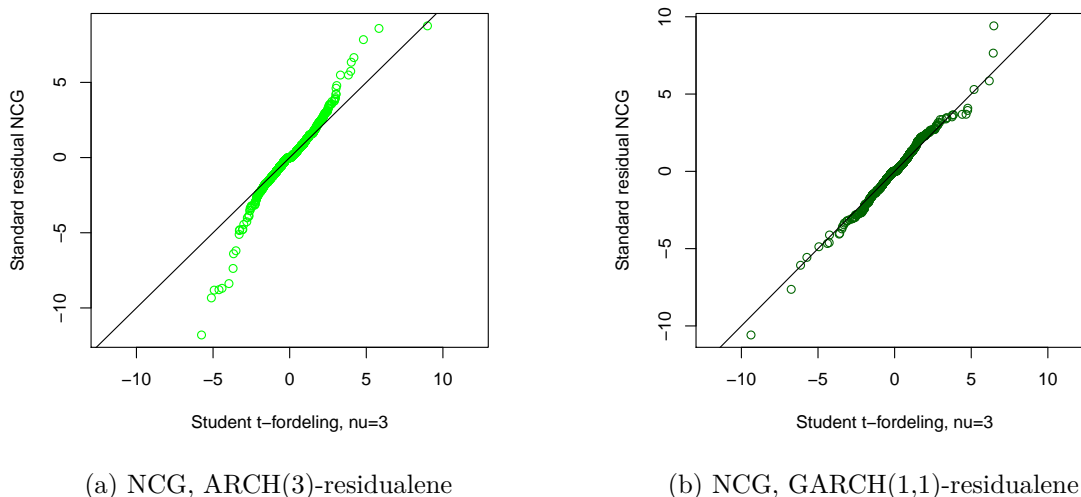


(f) Venstre hale

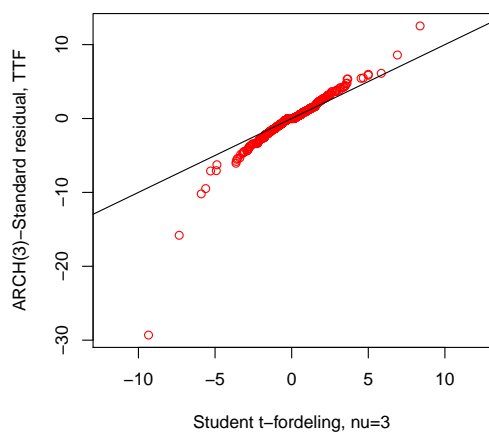
Figur 5.12: Tettheten til de standardiserte residualene mot student t-fodeling

I figur 5.12 er det flere plott over tetthetene til \hat{Z}_t , sammen med den antatte student t-fordelingen. Vi legger merke til at de empiriske residualene har en brattere kurve enn den antatte t-fordelingen, men båndviddene er ganske like. Siden det er den empiriske tettheten over residualene, er det vanskelig å si noe om de ulike toppene. Samtidig ser vi ikke noe merkbar forskjell mellom tettheten fra residualene til ARCH-modellene og tettheten fra residualene fra GARCH modellene, med mindre vi ser nærmere på halene. Halene er til høyre i figur 5.12. Det er 50 eller mer punkter etter -2 . Det er vanskelig å konkludere fra figur 5.12, hvilken av ARCH eller GARCH som passer best, i figur 5.12b ser det faktisk ut som ARCH(3) modellen gir en bedre tilpasning.

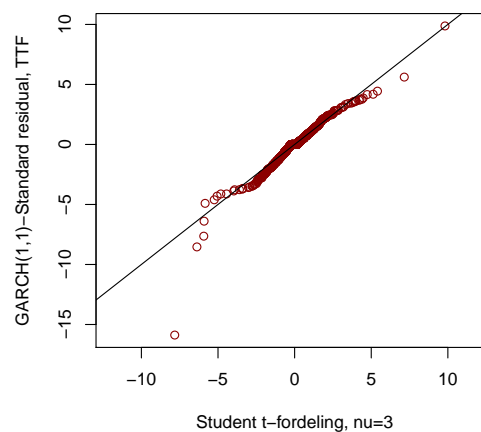
Likeledes studerer vi QQ-plottene som er vist i figur 5.13 og i figur 5.14, her ser vi også residualene, \hat{Z}_t , mot den antatte student t-fordelingen. Fra QQ-plottene kan det se ut som det er en del avvik. I motsetning av hva vi kunne se i figur 5.12, er det i figur 5.13 og 5.14 tydelig at residualene fra GARCH-modellene er mer like t-fordelingen, enn residualene fra ARCH-modellene. GARCH(1,1) residualene stemmer bra overens med de antatte t-fordelingene. Selv om det er noe avvik i QQ-plottene, betyr det nødvendigvis ikke at vi skal forkaste antakelsen om at residualene er student t-fordelt.



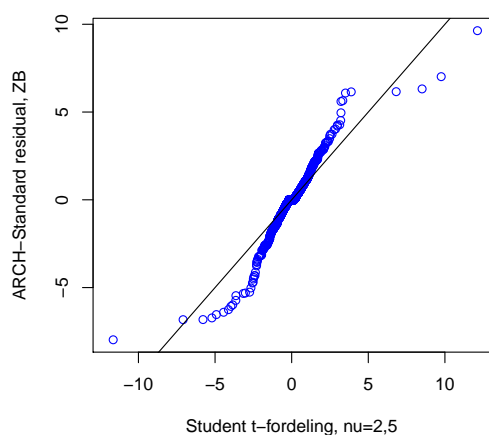
Figur 5.13: QQ-plot, standardiserte residual mot student t-fordeling



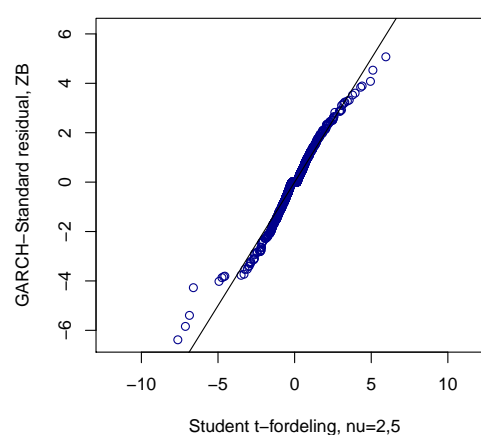
(a) TTF, ARCH(3)-residualene



(b) TTF, GARCH(1,1)-residualene



(c) Zeebrugge, ARCH(1)-residualene



(d) Zeebrugge, GARCH(1,1)-residualene

Figur 5.14: QQ-plot, standardiserte residual mot student t-fordeling

Hvis modellen er tilstrekkelig skal de standardiserte empiriske residualene oppføre som i.i.d. variabler. Dermed er en annen måte å sjekke om vi har en riktig spesifisert GARCH-modell, å teste om de standardiserte estimerte residualene, $\hat{Z}_t = X_t/\sigma_t$, er uavhengige. Vi har brukt to tester som kan teste for uavhengighet. Den ene testen er Ljung-Box testen, der vi ser på de kvadrerte residualene. Ljung-Box observator er gitt i ligning (3.26). Den andre testen er ARCH-LM testen, som tester om ARCH-effektene er fjernet. I tabell 5.14 er det en oversikt over de ulike testene som er gjort på de standardiserte residualene.

Tabell 5.14: Standardiserte empiriske residualene med testverdier og tilhørende p-verdier

Test	Modell	Marked		Testverdi	P-verdi
Ljung-Box	ARCH(3)	NCG	Q(1)	1,45	0,23
			Q(5)	12,69	0,03
			Q(10)	68,67	$8,01 \times 10^{-11}$
			Q(15)	91,99	$4,21 \times 10^{-13}$
ARCH-LM	ARCH(3)	NCG	LM(1)	1,45	0,23
			LM(5)	12,33	0,03
			LM(10)	70,67	$3,28 \times 10^{-11}$
			LM(15)	85,19	$7,76 \times 10^{-12}$
Ljung-Box	GARCH(1,1)	NCG	Q(1)	0,34	0,56
			Q(5)	3,67	0,60
			Q(10)	5,83	0,83
			Q(15)	8,49	0,90
ARCH-LM	GARCH(1,1)	NCG	LM(1)	0,34	0,56
			LM(5)	3,71	0,60
			LM(10)	5,73	0,84
			LM(15)	8,50	0,90
Ljung-Box	ARCH(3)	TTF	Q(1)	0,06	0,80
			Q(5)	1,22	0,94
			Q(10)	1,45	0,99
			Q(15)	1,86	1
ARCH-LM	ARCH(3)	TTF	LM(1)	0,06	0,80
			LM(5)	1,22	0,94
			LM(10)	1,45	0,99
			LM(15)	1,85	1
Ljung-Box	GARCH(1,1)	TTF	Q(1)	0,20	0,66
			Q(5)	1,47	0,92
			Q(10)	1,94	0,99
			Q(15)	4,03	0,99
ARCH-LM	GARCH(1,1)	TTF	LM(1)	0,20	0,66
			LM(5)	1,46	0,92
			LM(10)	1,90	0,99
			LM(15)	3,99	0,99
Ljung-Box	ARCH(1)	ZEE	Q(1)	2,84	0,09
			Q(5)	33,80	$2,61 \times 10^{-6}$
			Q(10)	88,44	$1,08 \times 10^{-14}$
			Q(15)	122,32	0,00
ARCH-LM	ARCH(1)	ZEE	LM(1)	2,84	0,09
			LM(5)	33,05	$3,68 \times 10^{-6}$
			LM(10)	85,22	$4,72 \times 10^{-14}$
			LM(15)	107,10	$5,55 \times 10^{-16}$
Ljung-Box	GARCH(1,1)	ZEE	Q(1)	1,65	0,20
			Q(5)	9,26	0,10
			Q(10)	24,84	0,01
			Q(15)	36,45	$6,34 \times 10^{-4}$
ARCH-LM	GARCH(1,1)	ZEE	LM(1)	1,65	0,20
			LM(5)	8,82	0,12
			LM(10)	23,03	0,01
			LM(15)	31,44	0,01

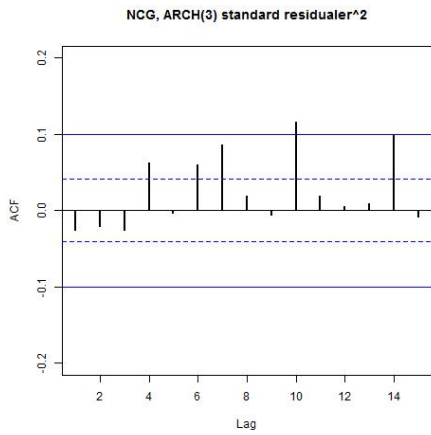
Resultatene fra tabell 5.14 viser at vi har uavhengighet ved lave lag etter å ha tilpasset en ARCH(3)-modell for log-avkastningene i NCG markedet. Derimot er residualene uavhengige på alle lag etter å ha tilpasset en GARCH(1, 1)-modell til NCG sine log-avkastninger. Dermed kan vi beholde nullhypotesen om at det ikke er seriellkorrelasjon i de standardiserte residualene ved å bruke GARCH(1, 1)-modell for log-avkastningene i NCG markedet. Det vil si at GARCH(1, 1) modell er en god modell for log-avkastningene i NCG markedet.

For TTF sine log-avkastninger fikk vi fjernet volatiliteten både ved å bruke en ARCH(3)-modell og ved å bruke en GARCH(1, 1)-modell. Dermed kan vi også her beholde nullhypotesen om ingen seriellkorrelasjon i de standardiserte residualene. Både ARCH(3) modellen og GARCH(1, 1) modellen er gode modeller for log-avkastningene i TTF markedet.

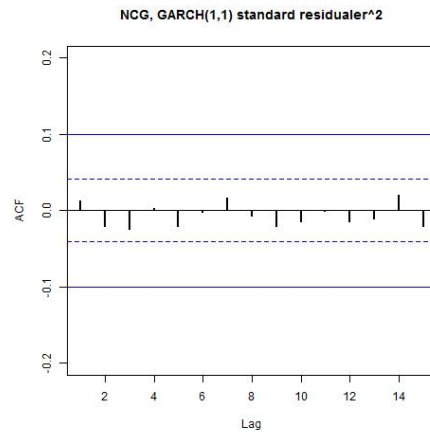
Derimot ved modelltilpasningene for log-avkastningene i Zeebrugge markedet, var det ikke en fullt så passende modell. De standardiserte residualene fra Zeebrugge modellene var ikke uavhengige på alle lag. ARCH-effektene vi hadde i log-avkastningen til Zeebrugger er ikke fjernet ved lag 10 og 15, men på lavere lag fikk vi fjernet volatiliteten ved å tilpasse en GARCH(1,1)-modell.

Til slutt kan vi sjekke for uavhengighet ved å studere et ACF-plott av de kvadrerte avkastningene. Et ACF-plott til NCG sine kvadrerte standardiserte residualer $\{\hat{Z}_{NCG,t}^2\}$, etter vi tilpasset en ARCH(3)-modell er vist i figur 5.15a). Her ser vi at det fremdeles er noe seriell korrelasjon i ARCH-residualene, mens figur 5.15b) viser at det er ingen seriell korrelasjon i de kvadrerte residualene når vi brukte GARCH(1, 1)-modell. Det vil si at GARCH(1, 1) modellen gjør en bedre jobb med å fjerne volatiliteten, og er dermed en bedre tilpasning en ARCH(3)-modellen. Vi ser at ARCH-effektene til TTF log-avkastningene er fjernet i både figur 5.15c) og figur 5.15d). Fra figur 5.15e) og figur 5.15f) er det tydelig at det er GARCH(1,1) som er den mest passende modellen til Zeebrugge sine log-avkastninger.

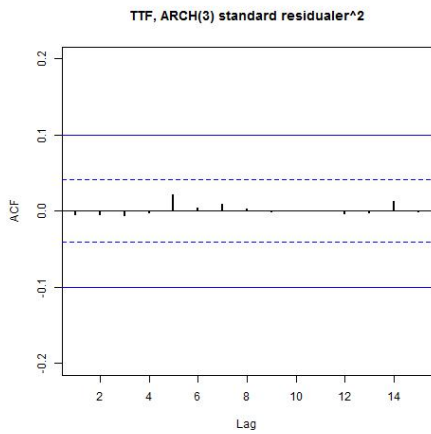
Ut ifra disse testene og plottene kan vi si at det er GARCH(1, 1)-modellen som er den modellen som passer best for modellering i alle tre gassmarkedene. Vi kan dermed anta at GARCH(1, 1)-modellen er tilstrekkelig for å beskrive den betingede heteroskedastisiteten av dataene.



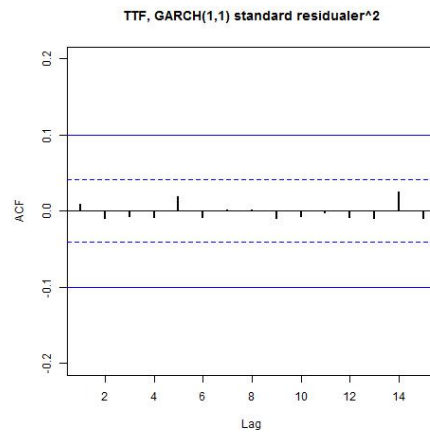
(a) NCG ARCH(3)-residualene



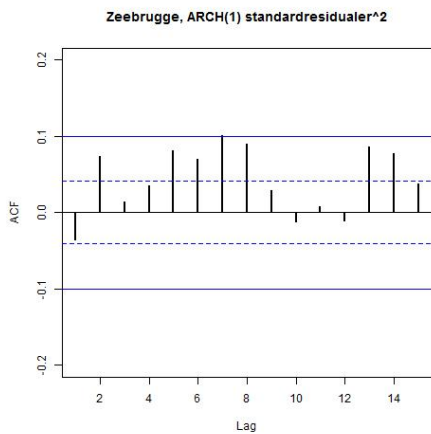
(b) NCG GARCH(1,1)-residualene



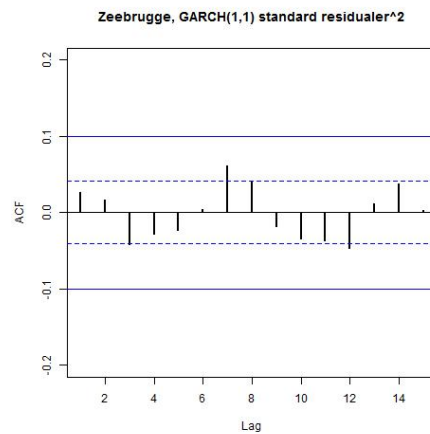
(c) TTF ARCH(3)-residualene



(d) TTF GARCH(1,1)-residualene



(e) Zeebrugge ARCH(1)-residualene



(f) Zeebrugge GARCH(1,1)-residualene

Figur 5.15: ACF på de standardiserte empiriske residualene

5.3.5 Modeller for de marginale tidsrekkene

Etter vi tok modelltestene på de ulike modellene fant vi at en GARCH(1,1)-modell er den mest passende beskrivende modellen for både NCG sine log-avkastninger (X_t^{NCG}), TTF sine log-avkastninger (X_t^{TTF}) og Zeebrugge sine log-avkastninger (X_t^{ZEE}). De respektive modellene for NCG, TTF og Zeebrugge sine log-avkastninger er gitt under.

$$X_t^{\text{NCG}} = \sigma_t^{\text{NCG}} Z_t \quad (5.3)$$

$$\begin{aligned} (\sigma_t^{\text{NCG}})^2 &= 3,57 \times 10^{-5} + 0,24(X_{t-1}^{\text{NCG}})^2 + 0,77(\sigma_{t-1}^{\text{NCG}})^2 \\ X_t^{\text{TTF}} &= \sigma_t^{\text{TTF}} Z_t \end{aligned} \quad (5.4)$$

$$\begin{aligned} (\sigma_t^{\text{TTF}})^2 &= 1,54 \times 10^{-5} + 0,16(X_{t-1}^{\text{TTF}})^2 + 0,83(\sigma_{t-1}^{\text{TTF}})^2 \\ X_t^{\text{ZEE}} &= \sigma_t^{\text{ZEE}} Z_t \end{aligned} \quad (5.5)$$

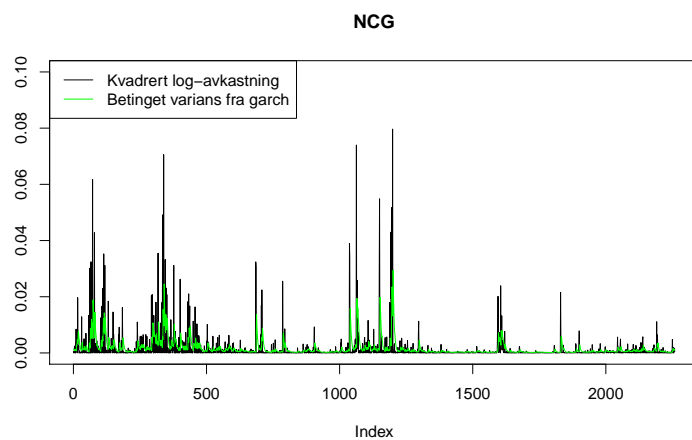
$$(\sigma_t^{\text{ZEE}})^2 = 1,58 \times 10^{-5} + 0,15(X_{t-1}^{\text{ZEE}})^2 + 0,87(\sigma_{t-1}^{\text{ZEE}})^2$$

I ligning (5.3) er $\{Z_t\}$ fra et tilfeldig utvalg fra en fordeling F . Videre i ligning (5.4) er $\{Z_t\}$ fra et tilfeldig utvalg fra en fordeling G . Til slutt i den siste ligningen (5.5) er $\{Z_t\}$ fra et tilfeldig utvalg fra en fordeling H . Der F og G er standardisert student t-fordeling med $\nu = 3$ frihetsgrader, varians lik en og forventning lik null. H er standardisert student t-fordeling med $\nu = 2,5$ frihetsgrader.

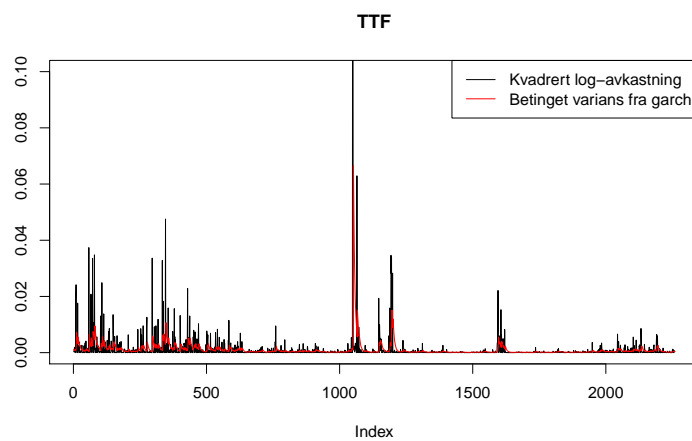
Ligningene (5.3), (5.4) og (5.5) har noen interessante egenskaper. Volatilitetsmodellen til TTF er den eneste modellen som er kovarians stasjonær, fordi betingelsen $\alpha_1 + \beta_1 < 1$ er tilfredsstillt. Dette er den eneste modellen hvor vi kan finne den ubetingede variansen i GARCH-modellen. Den ubetingede variansen er gitt som $\sigma^2 = \alpha_0 / (1 - \alpha_1 - \beta_1) = 1,54 \times 10^{-3}$. Derimot i de andre modellene er $\alpha_1 + \beta_1 > 1$. Det betyr at volatiliteten vil bli større for hver dag den blir modellert, i forhold til hva volatiliteten var den forrige dagen.

Størrelsen α_1 er størst i ligning (5.3), slik at NCG prisene reagerer kraftigere på prisendringer enn i de to andre modellene. β_1 er større enn α_1 for alle modellene. For NCG, TTF og Zeebrugge er β_1 henholdsvis 0,77, 0,83 og 0,87. Dermed vil volatiliteten den forrige dagen påvirke volatiliteten neste dag mer enn endringen av log-avkastningen den forrige dagen.

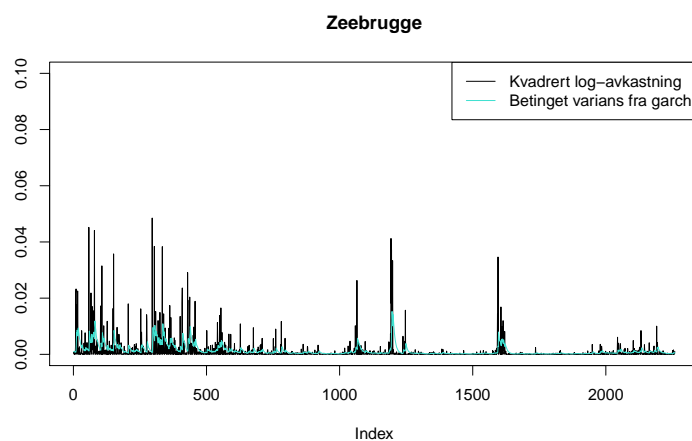
I figur 5.16 som er vist på neste side, er de betingede variansene fra de antatte GARCH(1,1) modellene, sammenlignet med de kvadrerte avkastningene. Den betingede variansen er funnet fra $t = 2$ til $t = n + 1$, dermed starter også de kvadrerte avkastningene når $t = 2$ i figur 5.16. Vi observerer at GARCH(1,1)-modellene følger volatilitetene som er i de kvadrerte log-avkastningene.



(a) GARCH(1,1) i NCG markedet



(b) GARCH(1,1) i TTF markedet



(c) GARCH(1,1) i Zeebrugge markedet

Figur 5.16: GARCH(1,1) modellen og de observerte log-avkastningene

5.3.6 Predikasjon av volatilitet

GARCH-modellene for den betingede variansen av de observerte avkastningene $\{X_t\}$ er funnet over. Den ubetingede variansen i $n + 1$ er gitt som følgende:

$$\left(\sigma_{n+1}^{\text{NCG}}\right)^2 = 3,57 \times 10^{-5} + 0,24\left(X_n^{\text{NCG}}\right)^2 + 0,77\left(\sigma_n^{\text{NCG}}\right)^2 \quad (5.6)$$

$$\left(\sigma_{n+1}^{\text{TTF}}\right)^2 = 1,54 \times 10^{-5} + 0,16\left(X_n^{\text{TTF}}\right)^2 + 0,83\left(\sigma_n^{\text{TTF}}\right)^2 \quad (5.7)$$

$$\left(\sigma_{n+1}^{\text{ZEE}}\right)^2 = 1,58 \times 10^{-5} + 0,15\left(X_n^{\text{ZEE}}\right)^2 + 0,87\left(\sigma_n^{\text{ZEE}}\right)^2 \quad (5.8)$$

(5.6), (5.7) og (5.8) kan forutsi de daglige volatilitetene av log-avkastningene en dag frem. Hvis vi skal predikere den betingede variansen for de neste periodene, kan vi bruke formlene som vi fant i ligning (3.25):

$$\begin{aligned} \hat{\sigma}_{n+1} &= \sigma^2 + (\sigma_n^2 - \sigma^2)(\alpha_1 + \beta_1), \\ \hat{\sigma}_{n+2} &= \sigma^2 + (\sigma_n^2 - \sigma^2)(\alpha_1 + \beta_1)^2, \\ &\vdots \\ \hat{\sigma}_{n+l} &= \sigma^2 + (\sigma_n^2 - \sigma^2)(\alpha_1 + \beta_1)^l, \end{aligned}$$

der σ^2 er den ubetingede variansen og gitt som $\sigma^2 = \alpha_0/(1 - \alpha_1 - \beta_1)$. Initialverdiene er satt til 8. Januar 2015 som er den siste observasjonsdagen i datasettet. De betingede variansene i $n + 1$, og de observerte kvadrerte log-avkastningene er gitt i tabell 5.15:

Tabell 5.15: Predikasjon av volatilitet

Marked	σ_{n+1}^2	$\hat{\sigma}_{n+1}^2$	X_{n+1}^2
NCG	$1,29 \times 10^{-3}$	—	$1,74 \times 10^{-6}$
TTF	$4,54 \times 10^{-4}$	$4,06 \times 10^{-4}$	$1,32 \times 10^{-4}$
ZEE	$4,95 \times 10^{-4}$	—	$2,30 \times 10^{-5}$

Den betingede variansen i NCG markedet for en dag frem i tid er 0,00129. Samtidig er den observerte prisen i NCG markedet den 9. Januar på 1,9711 Cent/kWh, mens prisen dagen før var på 1,9685 Cent/kWh. Dermed er den kvadrerte log-avkastningen på $1,74 \times 10^{-6}$, som er mindre enn den betingede variansen.

Den betingede variansen er noe mindre i modellen for TTF markedet. Den observerte gassprisen i TTF markedet den 9. Januar 2015 var 1,9068 Cent/kWh, mens prisen dagen før var 1,9051 Cent/kWh. Det medfører at den kvadrerte log-avkastning er på $1,32 \times 10^{-4}$. I GARCH-modellen

for TTF markedet har vi at $\alpha_1 + \beta_1 < 1$, dermed har vi også funnet den empiriske betingede variansen i denne modellen. Det hadde vært mulig å predikert den empiriske betingede variansen l dager frem i tid.

Observert gasspris for Zeebrugge markedet den 9. Januar 2015 var 1,9453 Cent/kWh, mens prisen dagen før var 1,9360 Cent/kWh. Dette ga en kvadrert log-avkastning på $2,30 \times 10^{-5}$. Den betingede variansen var $4,95 \times 10^{-4}$.

Vi legger merke til at variansen er størst for NCG, og minst for TTF. Samtidig som de observerte kvadrerte log-avkastningene er størst for NCG, og minst for TTF.

Videre har vi grunn til å tro at naturgassprisene i Europa er avhengig av hverandre. Zeebrugge, TTF og NCG er alle tre sentrale knutepunkter i Europa. Den 8. Januar var TTF-prisen for naturgass 1,9051 Cent/kWh, som var under NCG og Zeebrugge sine naturgasspriser på henholdsvis 1,9685 Cent/kWh og 1,9360 Cent/kWh. Slik at hvis det er noen sammenheng, og vi kunne tatt med denne sammenhengen i modellen. Da kunne modellene kanskje gitt oss bedre prediksjoner. I neste delkapittel skal vi se nærmere på sammenhengen mellom de ulike naturgassprisene, og videre modellere sammenhengen. Copula kan knytte sammen fordelingene i de ulike gassmarkedene, og kointegrasjon kan finne om det er en langtidslikevekt mellom naturgassmarkedene.

5.4 Avhengighet

GARCH-modellene over gir tilstrekkelig beskrivelse av de volatiliteten i log-avkastningene hver for seg. Imidlertid er det grunn til å tro at prisene er avhengig av hverandre. Å ta med denne avhengigheten i modellene, vil kanskje gi oss bedre prediksjoner for både NCG, TTF og Zeebrugge sine naturgasspriser.

5.4.1 Korrelasjon

En måte å betrakte avhengigheten på er måle korrelasjonen mellom tidsrekkene. Verdiene av korrelasjons koeffisientene er kalkulert ved Pearsons (P), Kendall's (K) tau og Spreaman's (S) rho. Der den vanligste å bruke er Pearsons. Pearson er en lineær korrelasjonskoeffisient. I tabell 5.16 er det en oversikt over korrelasjonskoeffisientene til log-avkastningene. Vi observerer at det er rimelig høy kryss-korrelasjon mellom log-avkastningene i de ulike markedene.

Tabell 5.16: Korrelasjons koeffisientene til log-avkastningene

Variabel	NCG-log-avkastinger			TTF-log-avkastinger			ZEE-log-avkastinger		
	P	(K)	S	P	(K)	S	P	(K)	S
NCG-log-avkastinger	1			0,50	(0,42)	0,54	0,51	(0,37)	0,48
TTF-log-avkastinger	—			1			0,69	(0,59)	0,72
ZEE-log-avkastinger	—			—			1		

Videre i den neste tabellen, tabell 5.17 er det en oversikt over korrelasjonskoeffisientene fra GARCH(1,1)-residualene. For samtlige korrelasjonskoeffisienter er avhengigheten statistisk signifikant, og også her ganske høye. Vi kan si at vi har avhengighet mellom variablene.

Tabell 5.17: Korrelasjons koeffisientene til GARCH(1,1)-residualene

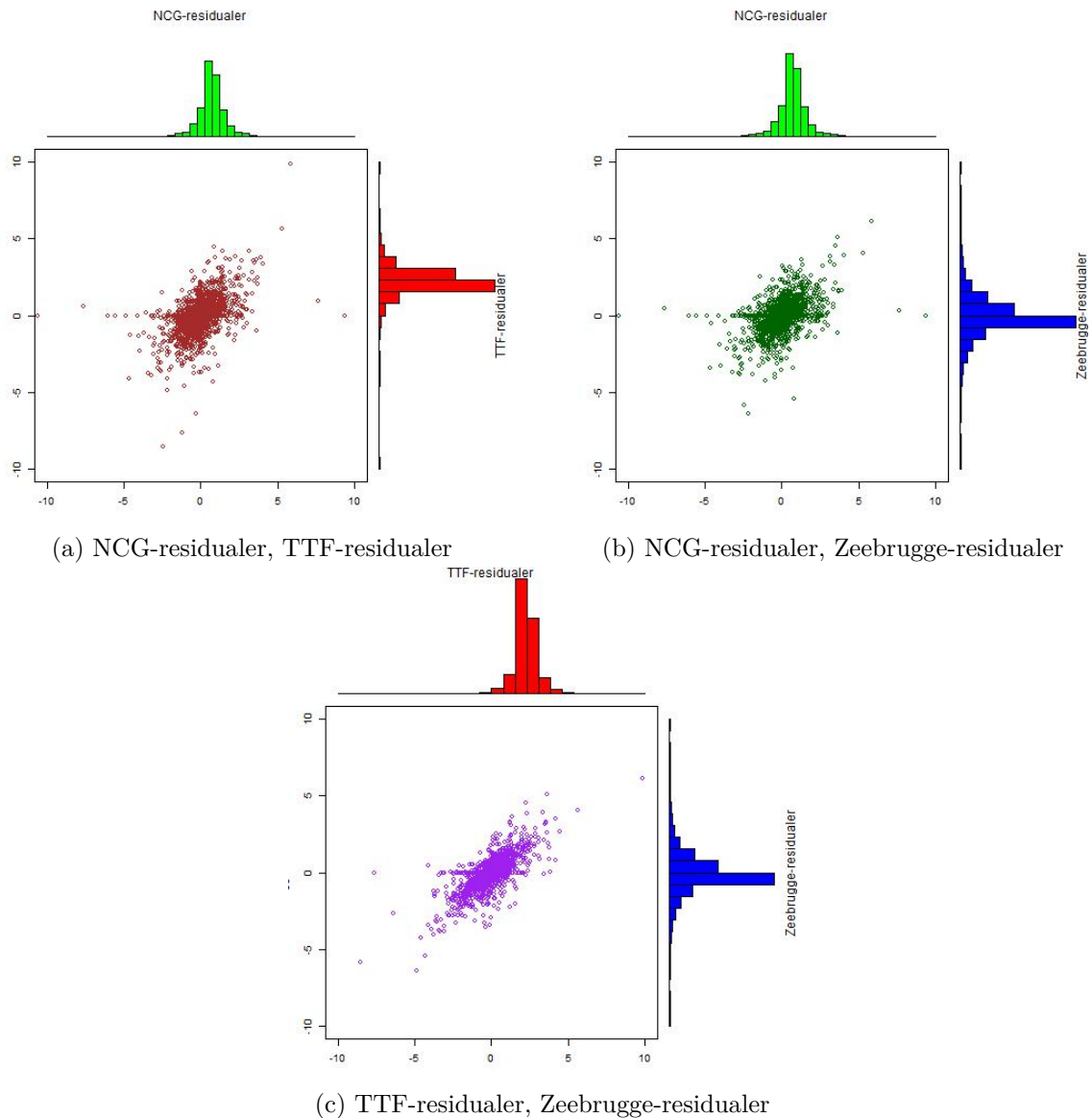
Variabel	NCG-residualer		TTF-residualer		ZEE-residualer	
	P	(K) S	P	(K) S	P	(K) S
NCG-residualer	1		0,49	(0,42) 0,56	0,47	(0,37) 0,50
TTF-residualer	—		1		0,71	(0,59) 0,73
ZEE-residualer	—		—		1	

5.4.2 Scatter plott

En annen måte å teste om det er avhengighet mellom variablene er ved å studere de parvise scatterplottene. Figur 5.17 viser en oversikt over de parvise scatterplottene til GARCH-residualene, og histogrammene av de marginale fordelingene til residualene fra GARCH-modellene. I figur 5.17a) er det tydelig at det er avhengighet. Vi ser at det er en positiv sammenheng mellom NCG-residualene og TTF-residualene, slik at høyere NCG-residualer gir også høyere TTF-residualer.

I figur 5.17b) er det derimot noe mindre avhengighet, men også her er det avhengighet. Til slutt i figur 5.17c) er det den tydeligste avhengigheten. Det er tydelig at når prisene går opp i Zeebrugge-markedet, vil dette medføre at prisene i TTF-markedet også øker. At det er mest avhengighet mellom TTF og Zeebrugger er ikke veldig overraskende, siden disse to knutepunktene har kort fysisk avstand mellom seg.

En nærmere sjekk på scatterplottene i figur 5.17 viser at en del punkt er samlet. Denne samlingen betyr at det ganske ofte ikke er en prisendring fra den ene dagen til den andre. Tilstedeværelsen til alle disse nullene indikerer på at sannsynlighetsfordelingen er singular, i den forstand at den kumulative fordelingen må hoppe ved 0. Disse hoppene kan bli ett hinder i den videre analysen. Dermed velger vi å fjerne disse nullene for den videre analysen (Carmona, 2014).



Figur 5.17: Scatterplott

5.5 Copula på naturgass residualer

I forrige delkapittel observerte vi at de standardiserte residualene fra GARCH-modellene var avhengig av hverandre. I denne delen skal vi finne en copula C som kan koble de marginale fordelingene fra GARCH-modellene sammen. Mye av fremgangsmåten i dette delkapittelet er hentet fra Gregoire et al. (2008) og fra Carmona (2014).

I kapittel 5.3 observerte vi at vi hadde volatilitet i log-avkastningene. GARCH-metodologi

ble brukt for å modellere denne volatiliteten. Ved å tilpasse GARCH(1,1)-modeller for log-avkastningene i kapittel 5.3 fikk vi følgende tre modeller:

$$\left(X_t^{\text{NCG}}\right)^2 = \left(\sigma_t^{\text{NCG}}\right)^2 Z_t^2 \quad (5.9)$$

$$\left(\sigma_t^{\text{NCG}}\right)^2 = 3,57 \times 10^{-5} + 0,24\left(X_{t-1}^{\text{NCG}}\right)^2 + 0,77\left(\sigma_{t-1}^{\text{NCG}}\right)^2$$

$$\left(X_t^{\text{TTF}}\right)^2 = \left(\sigma_t^{\text{TTF}}\right)^2 Z_t^2 \quad (5.10)$$

$$\left(\sigma_t^{\text{TTF}}\right)^2 = 1,54 \times 10^{-5} + 0,16\left(X_{t-1}^{\text{TTF}}\right)^2 + 0,83\left(\sigma_{t-1}^{\text{TTF}}\right)^2$$

$$\left(X_t^{\text{ZEE}}\right)^2 = \left(\sigma_t^{\text{ZEE}}\right)^2 Z_t^2 \quad (5.11)$$

$$\left(\sigma_t^{\text{ZEE}}\right)^2 = 1,58 \times 10^{-5} + 0,15\left(X_{t-1}^{\text{ZEE}}\right)^2 + 0,87\left(\sigma_{t-1}^{\text{ZEE}}\right)^2$$

I ligning (5.9) er $\{Z_t\}$ fra et tilfeldig utvalg fra en fordeling F , i ligning (5.10) er $\{Z_t\}$ fra et tilfeldig utvalg fra en fordeling G . I den siste ligningen (5.11) er $\{Z_t\}$ fra et tilfeldig utvalg fra en fordeling H . Der F og G er fra en standardisert student t-fordeling med 3 frihetsgrader, og H er en standardisert student t-fordeling med 2,5 frihetsgrader.

Vi kan bruke en copula, C , til å koble de marginale fordelingene F og G , G og H , eller H og F sammen. I copula analysen i denne oppgaven har vi kun fokusert på en to-dimensjonal copula. Siden vi har tre ulike gasspriser, kobler vi derfor kun sammen to og to. Vi har kun sett på de to parene som har høyest korrelasjon. De to parene med høyest korrelasjon er paret TTF og Zeebrugge, og paret NCG og TTF. I figur 5.4 oppdaget vi at gassmarkedene $\text{NCG} \sim \text{TTF}$ og $\text{TTF} \sim \text{ZEE}$. Dermed har vi ikke sett på paret NCG og Zeebrugge.

NCG markedet og TTF markedet i copulaskala

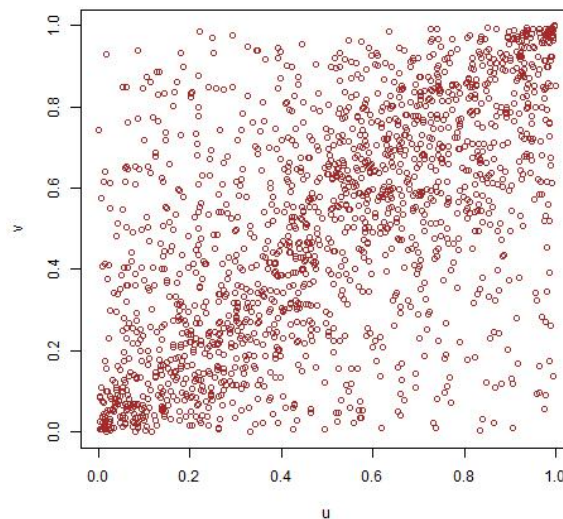
Først ser vi på sammenhengen mellom NCG og TTF sine naturgasspriser. Siden vi har volatilitet i log-avkastningene bruker vi de standardiserte residualene fra GARCH-modellene, $\hat{Z}_T^{\text{NCG}} = X_t^{\text{NCG}}/\sigma_t$ og $\hat{Z}_t^{\text{TTF}} = X_t^{\text{TTF}}/\sigma_t$, for å modellere copulaen.

For å bruke copula er det nødvendig å omforme GARCH-residualene, \hat{Z}_t^{NCG} og \hat{Z}_t^{TTF} , til henholdsvis u_t og v_t slik at vi har et bivariat utvalg på $[0, 1]$. Det bivariate utvalget kan vi kalle for copulaskalaen. Før transformasjonen gjennomføres fjerner vi de dataene der det ikke har vært prisendringer for begge variablene. Å transformere de standardiserte residualene til en copula skala kan gjøres ved å bruke den antatte fordelingen i GARCH-modellene. Alternativt kan de standardiserte residualene omformes til en copulaskala ved å bruke kernel estimator for den kumulative fordelingsfunksjonen.

For transformasjonen i denne oppgaven har vi valgt å bruke de antatte fordelingsfunksjonene. Funksjonen *pstd* fra pakken **fGarch** i programmet **R**, er den vi har brukt for å transformere

til copulaskalaen. Parametrene i funksjonen $pstd$ er spesifisert i forhold til fordelingene F og G som vi antok når vi tilpasset GARCH-modellene.

Etter transformasjonen er den marginale fordelingen uniform. Det er et tegn på at innflytelsen av de originale marginale fordelingene er blitt fjernet fra dataene. Slik at det eneste som er igjen er hvordan u_t og v_t er koblet sammen. Avhengigheten mellom residualene fanges opp på måten disse koblingene er gjort.

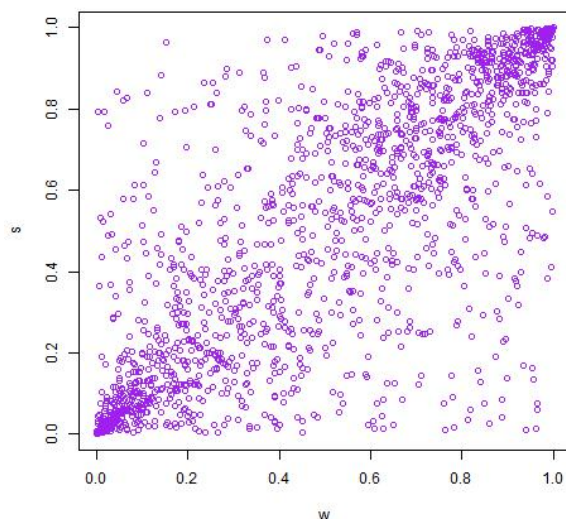


Figur 5.18: Avhengighet mellom NCG og TTF

Figur 5.18 viser en grafisk presentasjon over avhengigheten som er mellom variablene u_t og v_t i copulaskalaen. Vi ser at det er en tettere punkt konsentrasjon rundt diagonalen. Det tyder på at det er en positiv sammenheng mellom u_t og v_t . Vi ser også at det er mest klumping ut mot sidene, noe som kan bety at det er en t-copulafamilie som passer best til denne modellen. Hvis det hadde vært en Clayton eller en Gumbel copulafamilie, ville de vært mer avhengige i den ene halen i forhold til den andre. Ved Normal copulafamilie ville det vært endel klumping i midten. For å bekrefte avhengigheten mellom variablene, kan vi ta korrelasjonen mellom u_t og v_t . Når vi bruker Spearmans rho korrelasjon får vi at $\text{CORR}_{S,u_t,v_t} = 0,60$ og ved Kendalls tau får vi at $\text{CORR}_{\tau,u_t,v_t} = 0,44$. Begge korrelasjonskoeffisientene er signifikant forskjellig fra null.

TTF markedet og Zeebrugge markedet i copula skala

Videre gjør vi samme transformasjon for TTF-residualene $\hat{Z}_{TTF,t}$ og Zeebrugge-residualene $\hat{Z}_{ZEE,t}$, ved å bruke de respektive fordelingene H og G som ble definert i GARCH tilpassningene. GARCH-residualene blir transformert til henholdsvis w_t og s_t .



Figur 5.19: Avhengighet mellom TTF og Zeebrugge

Figur 5.19 viser en grafisk presentasjon over avhengigheten som er mellom variablene w_t og s_t . Vi ser i figur 5.19 som i figur 5.18 at det er en tettere punkt konsentrasjon rundt diagonalen. Det betyr at også her er det en positiv sammenheng mellom w_t og s_t . Samtidig ser vi at det er mest klumping ut mot sidene, noe som kan tyde på at det er en t-copula modell som passer best. Korrelasjonen mellom w_t og s_t er bekreftet ved å bruke Spearmans rho og Kendalls tau. Når vi bruker Spearmans rho får vi at $\text{CORR}_{S,w_t,s_t} = 0,78$ og ved Kendalls tau får vi at $\text{CORR}_{\tau,w_t,s_t} = 0,60$ begge er signifikant forskjellig fra null.

5.5.1 Copula estimering

I figur 5.18 er det en tydelig kobling mellom u_t og v_t , og vi må finne en passende copula C som kan beskrive denne sammenhengen. Tilsvarende kobling er det mellom w_t og s_t . Vi ønsker å finne den copulafamilien av Clayton, Gumbel, Normal og Student t-copulafamilien som passer best til paret u_t og v_t og paret w_t og s_t . Der henholdsvis de tre første familiene har en parameter som styrer avhengigheten, mens student t-copulafamilien i tillegg har en frihetsgradsparameter.

Vi bruker estimeringsmetoden inversion of Kendall's tau for å estimer parameteren ρ i de fire ulike familiene. Resultatene fra copula estimatene mellom paret u_t og v_t og paret w_t og s_t er gitt i tabell 5.18.

Ved inversion of Kendalls tau metoden finner man ikke frihetsgradsparameteren ν til student t-copula familien. Hvis en bruker den inverse av Kendalls tau metoden, går det ann å la frihetsgradsparameteren variere og velge den som passer best (Gregoire et al., 2008). Til å finne

frihetsgradsparameteren ν har vi brukt maksimum likelihoods estimering for å tilpasse en t-copulafamilie. Når man bruker maksimum likelihood estimering for å tilpasse en t-copulafamilie, finner man begge parametrene. Da er ν ofte ikke et heltall. Siden vi har brukt inversion of Kendalls tau har vi latt parameteren ν variere mellom de nærmeste heltallene som ble estimert ved mle. Deretter finner man hvilken ν som gir den mest passende t-copulafamilien. Alternativt kunne vi brukt estimerings resultatene fra mle, istedenfor resultatene fra den inverse til Kendalls tau.

Tabell 5.18: Estimaten til avhengighetsparameteren ρ for de forskjellige copulafamiliene

Copulafamilie	Markeder		Formel for metode
	NCG(u_t) og TTF(v_t)	TTF(w_t) og ZEE(s_t)	
	Estimat	Estimat	
Normal	$\hat{\rho} = 0,64$	$\hat{\rho} = 0,81$	$\hat{\rho} = \sin(\pi \text{CORR}_\tau / 2)$
t	$\hat{\rho} = 0,64$	$\hat{\rho} = 0,81$	$\hat{\rho} = \sin(\pi \text{CORR}_\tau / 2)$
Clayton	$\hat{\rho} = 1,60$	$\hat{\rho} = 3,05$	$\hat{\rho} = 2\text{CORR}_\tau / (1 - \text{CORR}_\tau)$
Gumbel	$\hat{\rho} = 1,80$	$\hat{\rho} = 2,52$	$\hat{\rho} = 1 / (1 - \text{CORR}_\tau)$

5.5.2 Modelltesting: Valg av copula

Ved copula tilpasningen for NCG og TTF oppnår alle modellene minimumskravet på $\text{CORR}_{\tau, u_t, v_t} = 0,44$, men det betyr nødvendigvis ikke at alle familiene er passende modeller. Minimumskravet til TTF og Zeebrugge var $\text{CORR}_{\tau, w_t, s_t} = 0,60$, også her møter alle modellene dette kravet. For å finne den copulafamilien i tabell 5.18 som passer best med de standardiserte residualene kan vi bruke modelltester. Hypotesen er at copulaen, C , tilhører en copulafamilie C_0 , slik at nullhypotesen er $H_0 : C \in C_0$.

En modelltest kan gjøres ved å bruke funksjonen *BiCopGofKendall* fra pakken **CDVine**. Dette er en modelltest basert på Kendalls prosessen for bivariate copula data. Den finner *Cramer-von Mises* og *Kolmogorov-Smirnov* testobservator med tilhørende p-verdier. Testresultatene fra modellene til NCG og TTF er gitt i tabell 5.19, der p-verdiene er i parentesene. Tilsvarende testresultater for modellene til TTF og Zeebrugge er gitt i tabell 5.20.

Tabell 5.19: Testverdier med tilhørende p-verdier for de estimerte copulamodellene til markedene NCG og TTF

Copulafamilie	Cramer-von Mises	Kolmogorov-Smirnov
Normal	1,53(0,00)	2,43 (0,00)
t(3)	0,67(0,00)	1,83 (0,00)
t(4)	0,66 (0,00)	1,82(0,00)
Clayton	3,90(0,00)	3,38(0,00)
Gumbel	2,00(0,00)	2,48(0,00)

Tabell 5.20: Testverdier med tilhørende p-verdier for de estimerte copulamodellene til markedene TTF og Zeebrugge

Copula	Cramer-von Mises	Kolmogorov-Smirnov
Normal	1,46(0,00)	2,16(0,00)
t(3)	0,99 (0,00)	2,17 (0,00)
t(4)	0,91 (0,00)	2,03 (0,00)
Clayton	4,12(0,00)	3,45(0,00)
Gumbel	1,89(0,00)	2,32(0,00)

Tabellene 5.19 og 5.20 over viser at ingen av copulafamiliene er passende. Den copulafamilien som hadde lavest testobservator var student t-familien.

Noen andre modelltester som kan velge den mest passende copulafamilien av de utvalgt copulafamiliene, er Vuong- og Clarketestene, disse testene er gitt i funksjonen *BiCopVuongClarke* fra pakken **CDVine**. Denne funksjonen sammenligner to og to familier, den copulafamilien som passer best får +1 poeng, og den som passer minst for -1. Resultatene fra Vuong- og Clarketesten for markedene NCG og TTF er gitt i tabell 5.21.

Tabell 5.21: Vuong og Clarke test for de estimerte copulamodellene til markedene $NCG(u_t)$ og $TTF(v_t)$

Test	Gaussian	t	Clayton	Gumbel
Vuong	-1	3	-3	1
Clarke	-1	3	-3	1

Vuong- og Clarketest favoriserer t-copulafamilien. Dermed kan vi anta fra de fire forskjellige testene i tabell 5.19 og 5.21 at det er t-familien som er den parametriske copulafamilien som passer best.

Vuong- og Clarketesten for TTF og Zeebrugge ga samme resultat som for NCG og TTF, og er gitt i tabell 5.22. Slik at t-copulaen gir den beste tilpasningen. Vuong og Clarke sier ikke noe om antall frihetsgrader i t-copula familien, slik at det kan være at det finnes en t-copula som hadde blitt bedre med en annen frihetsgradparameter.

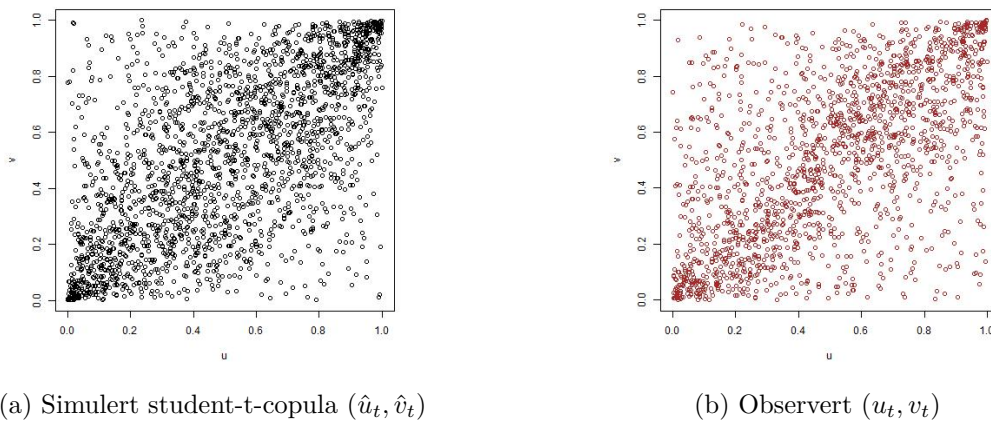
Tabell 5.22: Vuong og Clarke test for de estimerte copulamodellene til markedene $TTF(w_t)$ og $Zeebrugge(s_t)$

Test	Gaussian	t	Clayton	Gumbel
Vuong	-1	3	-3	1
Clarke	-1	3	-3	1

Siden målet var å finne den mest passende copulafamiliene av de fire forskjellige copulafamiliene, velger vi å bruke t-copulafamilien. Et alternativ kunne vært å tilpasse modellen, ved hjelp av

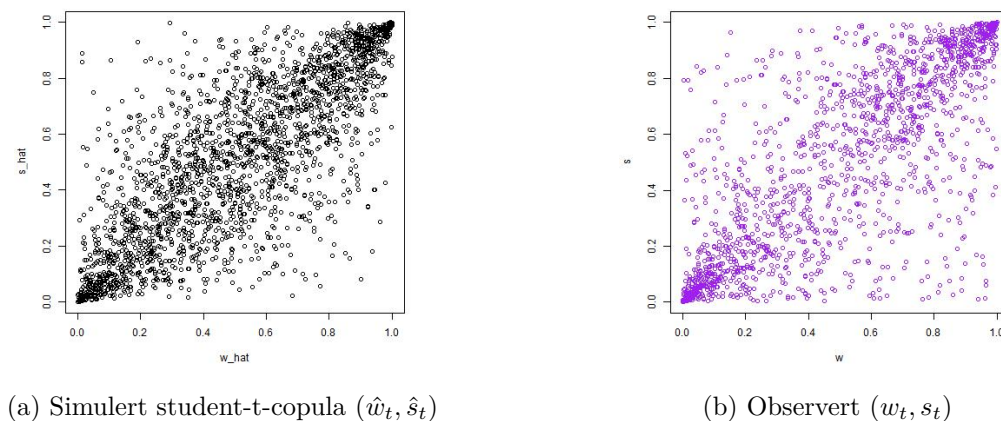
den empiriske kumulative fordelingsfunksjonen. Ved å bruke den empiriske kumulative fordelingsfunksjonen kunne vi fått en fordeling som passet dataene bedre enn en antatt parametrisk fordeling.

En måte å sjekke om modellen passer bra, er å sammenligne den simulerte student t-copulaen (\hat{u}_t, \hat{v}_t) i figur 5.21a) med figur 5.20b). Figur 5.20b) er den samme som er gitt i figur 5.18, denne viser den faktiske avhengigheten mellom u_t og v_t . I den simulerte student t-copulaen har vi brukt estimatene vi har fra tabell 5.18. Når vi sammenligner de simulerte verdiene mot de observerte verdiene, legger vi merke til at de ser ganske like ut i figur 5.20.



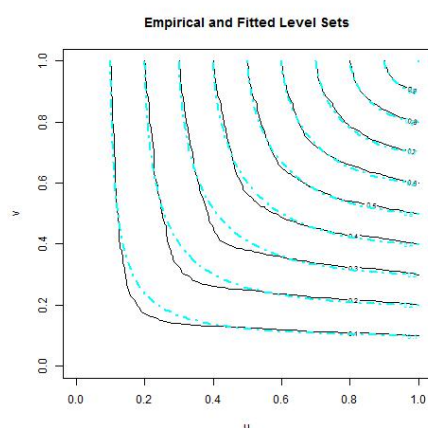
Figur 5.20: Sammenligne avhengighet ved scatterplott

Tilsvarende sammenligning er gjort i figur 5.21. Her ser det også ganske bra ut.



Figur 5.21: Sammenligne avhengighet ved scatterplott

En annen måte å sjekke at t-copula faktisk er en passende modell, er å sammenligne den empiriske copulaen med den modulerte t-copulafamilien (Carmona, 2014, s. 151). Den empiriske copulaen er funnet ved å bruke funksjonen `empirical.copula` fra pakken **Rsafd**. I figur 5.22 ser vi hvordan den empiriske copulaen er sammenlignet med estimerte parametriske t-copulaen. Vi ser at den empiriske copulaen er veldig lik den estimerte copulaen, men dette trenger ikke å bety at de passer bra sammen. Fordelingsfunksjoner har en tendens til å se like ut, selv om de nødvendigvis ikke er det.



Figur 5.22: Konturene til den empiriske copulaen og konturene til den tilpassete t-copulaen

5.5.3 Prognose ved bruk av copula

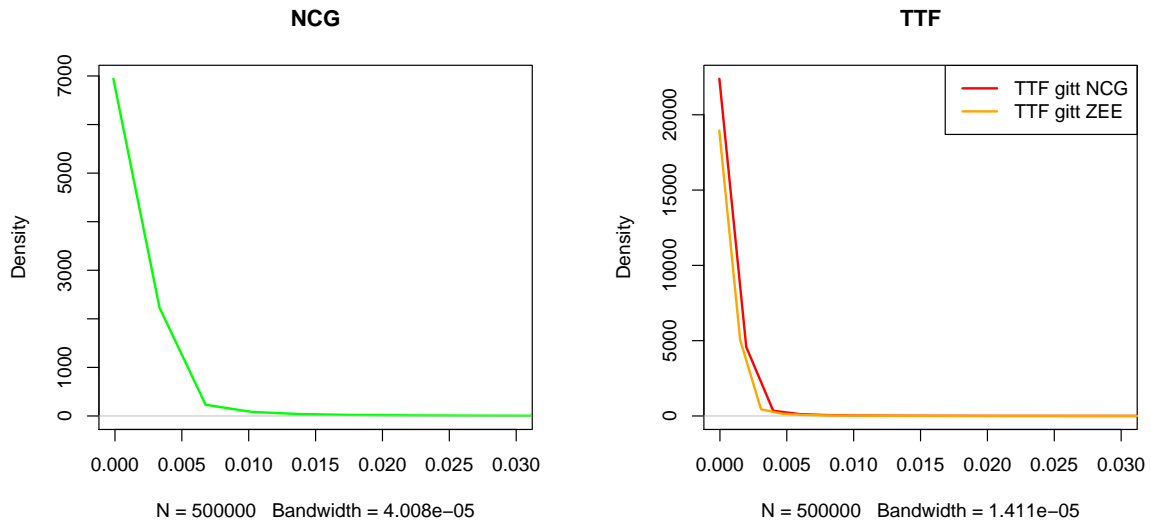
I denne delen skal vi bruke den tilpassede t-copulafamilien for å forutsi de kvadrerte log-avkastningene. For å predikere ved bruk av en copulamodell må man først generere et par u_t, v_t fra student t-copulafamilien (4.3) med parametrene som er gitt i tabell 5.18, der $\rho = 0,64$ og frihetsgradsparameteren $\nu = 4$. Etterpå omformer vi u_t og v_t slik at vi kan bruke det i GARCH-modellen. Omformingen gjøres ved å sette $\zeta = F^{-1}(u_t)$ og $\xi = G^{-1}(v_t)$, der F og G er en student t-fordeling med $\nu = 3$.

På tilsvarende måte genererer vi et par w_t, s_t fra student t-copulafamilien med parametrene som er gitt i tabell 5.18, der $\rho = 0,81$ og frihetsgradsparameteren $\nu = 4$. Etterpå omformer vi w_t og s_t slik at vi kan bruke det i GARCH-modellen, ved å sette $\varepsilon = G^{-1}(w_t)$ og $\gamma = H^{-1}(s_t)$, der G er en student t-fordeling med $\nu = 3$ og H er fra en t-fordeling med $\nu = 2,5$.

De kvadrerte log-avkastningene for en dag frem i tid finnes ved å bruke $(X_{t+1}^{\text{NCG}})^2 = \zeta^2(\sigma_{t+1}^{\text{NCG}})^2$, der $\sigma_{t+1}^{\text{NCG}}$ er den betingede variansen i GARCH-modellen. Tilsvarende framgangsmåte brukes for å finne de kvadrerte log-avkastningene $((X_{t+1}^{\text{TTF}})^2 = \xi^2(\sigma_{t+1}^{\text{TTF}})^2)$ i TTF markedet. TTF markedet er med i to forskjellige copula modelleringer, slik at vi har generert både v_t og w_t og omformet begge med fordelingen G . Det gjør at vi kan predikere de kvadrerte log-avkastningene, ved å ta utgangspunkt i v_t eller w_t . Begge prediksjonene er gitt i tabell 5.23. Til slutt har vi regnet ut

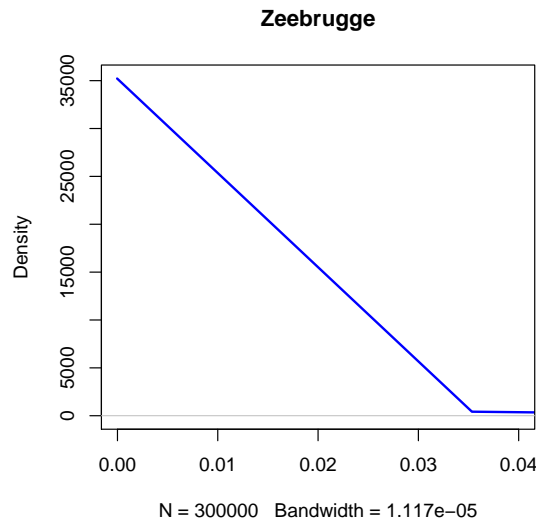
$(X_{t+1}^{ZEE})^2 = \gamma^2(\sigma_{t+1}^{ZEE})^2$. Initialverdiene er satt til 8. Januar 2015 som er den siste observasjon dagen i datasettet. For å fjerne usikkerheten ved denne prosedyren har vi gjentatt prosessen flere ganger.

De empiriske fordelingene av prognosene av de kvadrerte log-avkastningene er gitt i figur 5.23. Medianene til fordelingene er gitt i tabell 5.23. I tabell 5.23 finner vi også de observerte kvadrerte log-avkastningene.



(a) Tettheten til NCG

(b) Tettheten til TTF



(c) Tettheten til Zeebrugge

Figur 5.23: Prognose av de kvadrerte avkastningene ved copula- og GARCH-modell for den 9. Januar 2015.

Tabell 5.23: Kvadrerte log-avkastninger for $(n + 1)$

Marked	Median	Observert
NCG	$2,25 \times 10^{-4}$	$1,74 \times 10^{-6}$
TTF	$8,84 \times 10^{-5}$	$1,32 \times 10^{-4}$
TTF	$8,86 \times 10^{-5}$	$1,32 \times 10^{-4}$
ZEE	$5,92 \times 10^{-5}$	$2,30 \times 10^{-5}$

5.6 Kointegrasjon

En annen måte å ta med avhengighet mellom tidsrekkene, er ved å bruke kointegrasjon. Hvis vi har kointegrasjon innebærer det at det er en statistisk sammenheng mellom prisene. I denne delen av oppgaven har vi brukt pakken **egcm** i programmet R for å gjøre Engle-Granger to steps prosedyre. Videre har vi brukt pakken **urca** i programmet R for å gjøre Johansen-metoden. I tillegg har vi brukt referansen Pfaff (2008).

For å sjekke om vi har kointegrasjon mellom tidsrekkene, velger vi å se på gassprisene i 2014 til og med 8. Januar 2015. Naturgassprisene er brukt direkte, alternativt kunne man ført ha gjennomført en marginal transformasjon til normalfordeling.

Først sjekker vi om de ulike tidsrekkene er $I(1)$, det vil si om de inneholder enhetsrøtter. For å finne ut av dette kan vi bruke Augmented-Dickey-Fuller (ADF) test. En ADF test har nullhypotesen om at tidsrekkene inneholder en enhetsrot. Funksjonen `adfTest()` fra R-pakken: **fUnitRoots** kan brukes for å sjekke om man har enhetsrøtter. P-verdiene fra ADF testen er 0,26, 0,21 og 0,23 for henholdsvis NCG, TTF og Zeebrugge sine gasspriser. Basert på disse p-verdiene kan vi beholde nullhypotesen om enhetsrøtter. Dermed har vi funnet at alle gassprisene inneholder enhetsrøtter. For at de skal være $I(1)$ må ΔX_t være stasjonær, der X_t er gasspris. Vi finner at alle gassprisene er $I(1)$. En forutsetning for å bruke kointegrasjon er at tidsrekkene er $I(1)$.

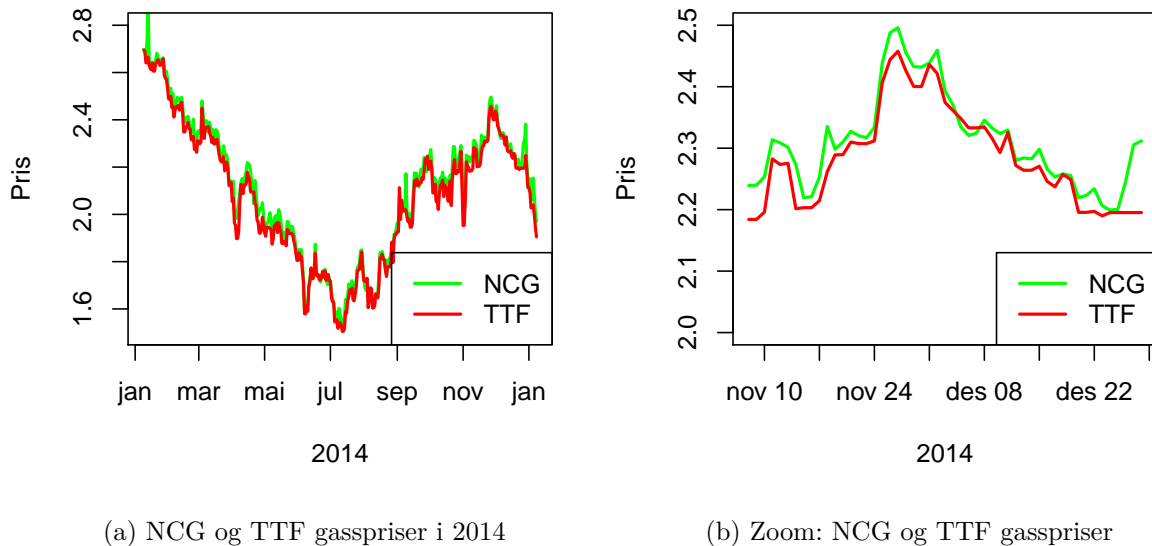
5.6.1 Engle-Grangers to-steps metode

Naturgassprisene i NCG- og TTF-markedet er vist i figur 5.24. Fra plottet kan det se ut som gassprisene i NCG markedet ligger generelt over naturgassprisene i TTF markedet. Over fant vi at alle gassmarkedene var $I(1)$. Vi lar naturgassprisen i NCG og TTF markedet i tid t være definert som henholdsvis X_t og Y_t .

I Engle-Grangers to-steps metode estimerer vi modellen:

$$X_t = \beta Y_t + r_t, \quad (5.12)$$

ved å bruke en lineær tilpasning av X_t med hensyn på Y_t .



Figur 5.24: NCG og TTF gasspriser

Regresjonsmodellen (5.12) ved å bruke gassprisene, gir oss følgende kointegrasjonsligning:

$$X_t = 1,01Y_t + r_t,$$

der $\hat{\beta} = 1,01$ er signifikant. Den lineære kombinasjon av variablene er residualene \hat{r}_t . For at vi skal ha kointegrasjon må residualene være stasjonære. En ADF-test på residualene gir oss en p-verdi på mindre enn 0,01, slik at vi kan forkaste hypotesen om enhetsrot. Dermed antar vi at residualene er stasjonære, og at det er kointegrasjon mellom tidsrekkene. Vi kan finne $r_t = \rho r_{t-1} + \epsilon_t$ ved en lineær tilpasning, der $\epsilon_t \sim N(0, \sigma^2)$. Da er $\rho = 0,455$ som er signifikant på fem prosentnivå. Siden $\rho < 1$ kan vi konkludere med at X_t og Y_t kointegrerte.

Til slutt lager vi en ECM regresjon. En ECM-modell av (5.12) kan skrives som $\Delta X_t = \alpha_0 \hat{r}_t + \alpha_1 \Delta Y_{t-1} + \alpha_2 \Delta X_{t-1} + \nu_t$. Resultatene fra ECM regresjonen er vist i tabell (5.24). Likevektsfeilen av siste periode er signifikant og har en verdi på $-0,24$ og en t-test på ΔY_{t-1} og ΔX_{t-1} viser at vi kan forkaste nullhypotesen om at $\alpha_1 = 1$ og $\alpha_2 = 1$. Vi legger derimot merke til $\alpha_1 + \alpha_2 \approx 0$. Den estimerte ECM er dermed gitt som $\Delta X_t = -0,24 \hat{r}_t + 0,38 \Delta Y_{t-1} - 0,34 \Delta X_{t-1} + \nu_t$.

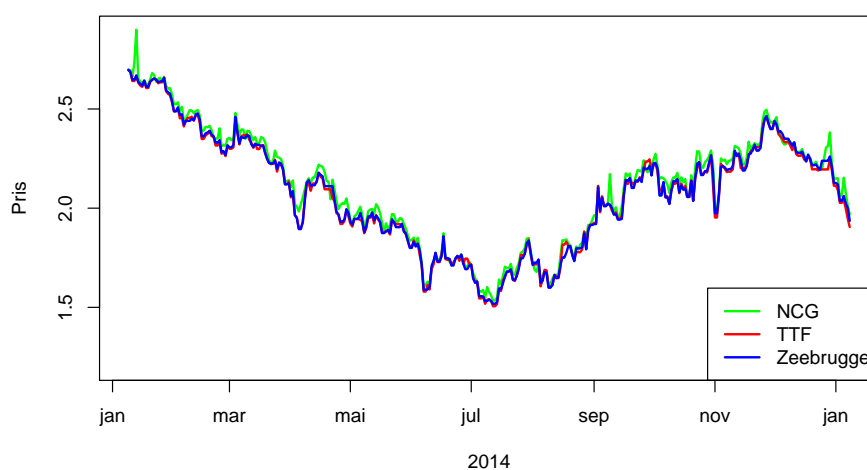
Tabell 5.24: Resultat fra Engle-Granger prosedyre med NCG og TTF sine gasspriser

Variabel	Estimat	std.error	t-verdi	P-verdi
ΔY_{t-1}	0,38	0,031	12,08	***
ΔX_{t-1}	-0,34	0,028	-12,11	***
\hat{r}_t	-0,24	0,026	-09,45	***

Problemet med å bruke Engle-Granger to-steps metode er å vite hvilken variabel vi skal bruke som den avhengige variabelen. I tillegg har vi til nå kun brukt bivariat tilfelle og vi får dermed bare en kointegrasjonsvektor. En viktig test for kointegrasjon som ikke er påvirket hvordan variablene er satt opp i forhold til hverandre, og som i tillegg kan takle multippelt tilfelle er Johansen testen.

5.6.2 Johansen-metode

Vi bruker alle tre gassprisene, forholdsvis NCG, TTF og Zeebrugge. Tidligere har vi sett at alle disse tidsrekkene er $I(1)$. Tidsrekkene er vist i figur 5.25.



Figur 5.25: NCG, TTF og Zeebrugge priser i 2014

Ved johansens prosedyre må vi først estimere VECM-modellen. VECM-modellen er gitt som $\Delta \mathbf{X}_t = \boldsymbol{\mu}_t + \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{X}_{t-1} + \boldsymbol{\epsilon}_t$, der $\boldsymbol{\epsilon}_t$ er hvit støy. For å estimere VECM-modellen og for å teste kointegrasjonsrangen i en VECM-modellen, har vi brukt R-pakken **urca** med funksjonen `ca.jo()`. Vi har satt standardverdiene til testobservatoren til å være maksimum egenverdi test, og vi setter at $\boldsymbol{\mu}_t$ er konstant. Resultatene er vist i tabell (5.25).

Tabell 5.25: Verdien til testobservatoren og kritiske verdier av maksimum egenverdi test

	Testverdi	10pct	5pct	1pct
$r \leq 2$	5,78	7,52	9,24	12,97
$r \leq 1$	81,51	13,75	15,67	20,20
$r = 0$	91,21	19,77	22,00	26,81

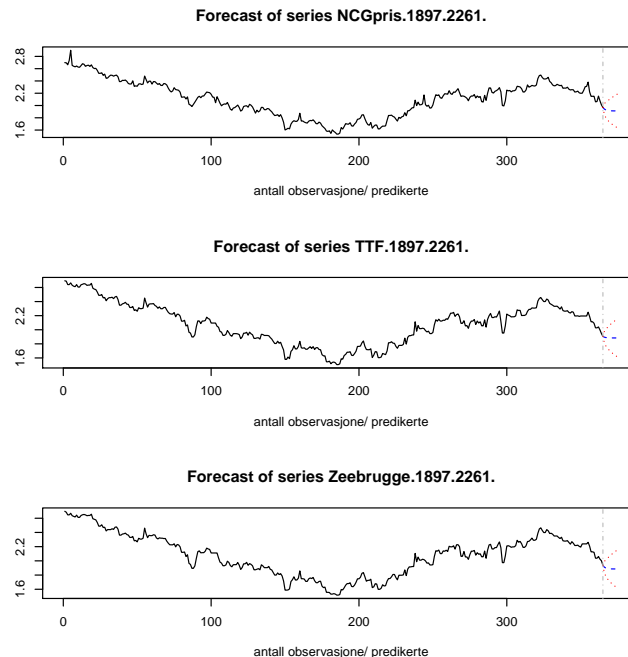
Testobservatoren for $r = 0$ og $r \leq 1$ er veldig høy, høyere enn de kritiske verdiene. Dermed forkaster vi nullhypotesen og antar at vi har kointegrasjon mellom dataene. Konintegrasjonsrangen er to. Vi får da at de estimerte verdiene for β og α med hensyn på NCGs gasspris er som

følger: $\beta = \begin{pmatrix} 1,00 \\ 0,88 \\ -1,89 \end{pmatrix}$ og $\alpha = \begin{pmatrix} 0,14 \\ -0,23 \\ -0,01 \end{pmatrix}$. Ut fra de estimerte verdiene til β og α ser det ut som

Zeebrugge med α lik $-0,01$ og β lik $-1,81$ vil nå likevekten senere enn de to andre gassprisene.

Det har blitt vist at en VECM-modell lages fra en VAR-modell (Pfaff, 2008, s. 84). Det er derfor mulig å uttrykke den estimerte VECM-modell som en VAR-modell. Dermed blir de estimerte koeffisientene av VECM-modellen transformert til VAR-ligningene. Denne transformasjonen er tilgjengelig i R-pakken **vars** ved funksjon `vec2var()`. Der vi setter at alle prisene har kointegrasjonrelasjoner som var utfallet av `ca.jo()`.

Vi bruker VAR-modellen for å predikere naturgassprisen. Nedenfor i figur 5.26 er det predikert prisen for ti dager frem ved å bruke VAR-modellen. Prognosene for NCG, TTF og Zeebrugges gasspriser den 9. Januar 2015 er på henholdsvis 1,9420 (1,9711) Cent/kWh, 1,8949 (1,9272) Cent/kWh og 1,9114 (1,9453) Cent/kWh. Der tallene i parentesene er de faktiske prisene. Vi legger merke til at prognosene er et stykke under de observerte prisene den dagen. Det kan være flere grunner til dette. En grunn kan være at vi har antatt normalfordelte tidsrekker, noe som de egentlig ikke er. Et alternativ kunne vært å gjennomført en marginal transformasjon til normalfordeling.



Figur 5.26: 10 dagers prediksjon for NCG, TTF og Zeebrugge priser

KAPITTEL 6

Konklusjon

I det første kapittelet studerte vi gassmarkedet generelt, slik at vi kunne finne mer ut om hva som påvirker naturgassprisene og gjør dem volatile. Til slutt i kapittel 5 analyserte vi gassprisene og tilpasset modeller, ut fra teorien som vi studerte i kapittel 2-4.

Det første vi gjorde i analysen var å studere hvordan naturgassprisene hadde oppført seg i den observerte tidsperioden. Vi observerte blant annet at temperatur hadde innvirkning på naturgassprisen. Spesielt hadde ekstreme kuldeperioder innvirkning på naturgassprisen. Andre ting som hadde innvirkning var den økonomiske situasjonen. Samt de politiske urolighetene i midtøsten og i Ukraina førte til mer svingninger i gassmarkedet.

Naturgassprisene i seg selv var ikke stasjonære, det medførte til at vi omformet prisene til log-avkastninger. Til disse log-avkastningene tilpasset vi GARCH-modeller. For hvert marked endte vi opp med at en GARCH(1,1)-modell var den beste tilpasningen for log-avkastningene. GARCH-modellene ble videre brukt for å predikere volatiliteten.

Vi hadde grunn til å tro at naturgassprisene var avhengige av hverandre. En grunn til at vi kan tro dette er fordi det er gassrør som kobler de forskjellige knutepunktene i Europa sammen. Vi fant det var avhengighet mellom de europeiske log-avkastningene. Videre brukte vi copula for å knytte sammen de marginale fordelingene fra GARCH-residualene sammen. Student t-copulafamilien viste seg å være den mest passende familien. Deretter fant vi prognosene for dagen etterpå.

Vi observerte at vi hadde kointegrasjon i naturgassprisene. Dette er den eneste modellen hvor vi bruker gassprisene direkte, og det er den eneste modellen der alle de tre variablene blir modulert samtidig. Først brukte vi Engle-Granger to stegs prosedyre og tilpasset en ECM modell. Videre brukte vi Johansens prosedyre og estimerte en VECM-modell. VECM-modellen ble uttrykt som en VAR-modell. Denne VAR-modellen ble brukt til prediksjon.

6.1 Resultatene

GARCH-modellene vi tilpasset viste at tidligere volatilitet hadde større innvirkning på volatilitetsmodellen, enn de tidligere endringer i log-avkastningene hadde. Vi fant at det kun var for det ene markedet, at GARCH-modellen oppfylte kravet om at $\alpha_1 + \beta_1 < 1$. Videre fant vi at det var en student t-copulafamilie, som best beskrev sammenhengen mellom de marginale fordelingene i GARCH-modellene. Til slutt oppdaget vi at vi hadde kointegrasjon mellom gassprisene.

Tabell 6.1: Predikerte kvadrerte avkastningen for den 9. Januar 2015

Marked	GARCH	Copula	VAR	Observert
NCG	$1,29 \times 10^{-3}$	$2,25 \times 10^{-4}$	$1,83 \times 10^{-4}$	$1,74 \times 10^{-6}$
TTF	$4,54 \times 10^{-4}$	$8,84 \times 10^{-5}$	$3,35 \times 10^{-5}$	$1,32 \times 10^{-4}$
ZEE	$4,95 \times 10^{-4}$	$8,86 \times 10^{-5}$	$1,16 \times 10^{-4}$	$2,30 \times 10^{-5}$

I tabell 6.1 er det en oversikt over alle prediksjonene. Der alle prediksjonene er gitt som kvadrerte avkastninger, i tillegg er det satt inn for de observerte kvadrerte log-avkastningene. Siden vi kun har sett på en dags prediksjon, er det vanskelig trekke noen bestemt konklusjon ut fra prediksjonene. Spesielt vanskelig er det å si hvilken modell som gir det beste resultatet.

Det vi kan gjøre er å sammenligne de predikerte verdiene fra modellene, med de trivielle prediksjonene. TTF markedet er det eneste markedet som oppfyller kravet om $\alpha_1 + \beta_1 < 1$, slik at vi kan finne den trivielle prediksjonen som er den ubetingede variansen. Den ubetingede variansen er gitt i ligning (3.24), og er i TTF markedet gitt som:

$$\sigma_t^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1} = 0,00154$$

Vi får at trivielle prediksjon ($1,54 \times 10^{-3}$) > GARCH prediksjon ($4,54 \times 10^{-4}$) > copula prediksjon ($8,84 \times 10^{-5}$) > VAR prediksjon ($3,35 \times 10^{-5}$).

I TTF markedet kan prediksjonfeilens størrelse oppgis i prosent av den ubetingede variansen.

$$\text{Pred.feil} = \frac{|\text{pred} - \text{obs}|}{\sigma^2} \quad (6.1)$$

Prediksjonfeilens størrelse (6.1) i TTF markedet er gitt i tabell 6.2:

Tabell 6.2: Prediksjonsfeil i TTF markedet

GARCH	Copula	VAR
20,97 %	2,77 %	6,33 %

Ut fra tabell 6.2 er det ved copula tilpasningen at vi får den minste prediksjonfeilen. Samtidig er det umulig å si om copula faktisk gir den beste prediksjonen, siden vi kun har sammenlignet med en predikert verdi.

6.2 Sammenligne de ulike modellene

Hovedforskjellen mellom GARCH- og copula-tilpasningen, er at vi i GARCH har en marginal modell, mens copula binder de marginale fordelingene sammen og finner en familie som kan beskrive denne sammenhengen. I GARCH(1,1)-modellen har vi tilpasset en GARCH(1, 1)-modell for hver tidsrekke, uten å ta hensyn til de andre tidsrekkene. Derimot ved copula-tilpasningen ble to og to markeder koblet sammen ved bruk av en copulafamilie. Et alternativ kunne vært å brukt en multivariat GARCH modell på to eller tre tidsrekker.

Vi observerte kointegrasjon i de naturgassprisene i Europa. Ved kointegrasjon har vi sett på naturgassprisene og ikke avkastningene, som vi har gjort ved GARCH-modellene. I GARCH og copula studerte vi hele tidsperioden fra 2008 til 2015, mens vi har kun testet for kointegrasjonen for det siste året. En annen forskjell er at vi har antatt normalfordeling ved kointegrasjon. I de andre modellene har vi antatt t-fordeling.

6.3 Problem

Når vi tilpasset GARCH-modellene, så vi bort fra regime endringene. Selv om vi observerte at volatiliteten var høyest i de to første årene. Datasettet er over en relativ lang tidsperiode som inneholder mye variert og forskjellig informasjon. Dette er både oppturer og nedturen som vi igjen kan finne forklaring på via andre situasjoner som har skjedd i den tidsperioden. Dette kan være et problem, fordi vi ikke får nøyaktig nok modell. Vi observerte at det kun var et marked som oppfylte kravet om andre ordens stasjonæritet ($\alpha_1 + \beta_1 > 1$). Et annet problem med GARCH er at vi antok fordelinger når vi tilpasset modellene, som kanskje ikke er de mest passende fordelingene.

Problemene fra GARCH-modellen tok vi med oss videre til copula tilpasningen. Når vi valgte ut copulafamilien, valgte vi kun den mest passende av de fire vi testet. Det kan være at en annen copulafamilie ville passet bedre. Når vi transformerte residualene til en copula skala brukte vi de antatte fordelingene, istedenfor en kernel estimator for den kumulative fordelingsfunksjonen. Dette kan ha gitt oss problemer. Et annet problem var måten vi fant frihetsgradsparameteren i t-copulafamilien. Det kunne vært at det finnes en t-copulafamilie som passer med en annen frihetsgradparameter. Et annet dilemma var at selv om Z_t^{TTF} og Z_t^{NCG} var sterkt avhengige, var ikke X_t^{TTF} avhengig til Z_{t+1}^{NCG} . Det betyr at X^{TTF} prosessen greier ikke å titte fram i X^{NCG} prosessen. Det kan gjøre at predikerte verdiene ved å bruke copula, nødvendigvis ikke gir noen forbedring.

Et problem med kointegrasjon var at vi antok at prosessene var normalfordelte, selv om dette ikke var tilfelle.

6.4 Videre arbeid

Problemene som er nevnt over kunne vært løst i videre studier. Ved å ta med GARCH-regime switching inn i modellen, kunne vi fått en mer nøyaktig modell. Da kunne vi hatt stasjonæritet i perioder, istedenfor å anta at vi har stasjonæritet i hele perioden. GARCH modellene for NCC og Zeebrugge gassmarkedene var ikke andre ordens stasjonære GARCH modeller. Det var i disse to markedene vi observerte noe seriell autokorrelasjon i log-avkastningene direkte. Denne autokorrelasjonen ble ikke fjernet, siden den var veldig liten. I et videre arbeid av oppgaven kunne vi fått fjernet denne autokorrelasjonen i log-avkastningene ved hjelp av en lineær modell. Det kunne vært at dette hadde gitt bedre GARCH tilpasninger etterpå. Vi kunne også forbedret modellen, ved å funnet en mer passende fordeling i GARCH-modellene. For eksempel kunne vi brukt QML estimering som ikke trenger noen forutsetninger om den observerte prosessen (Francq and Zakoian, 2010, s. 141). Istedenfor å ha brukt de antatte marginale fordelingene når vi transformerte variablene til copulaskalaen, kunne vi heller brukt den empiriske kumulative fordelingsfunksjonen i et videre studie av oppgaven. Vi kunne også ha testet for flere frihetsgradsparametre. I kointegrasjonen kunne vi brukt marginal transformasjon på prisene.

En annen ting vi ikke har gjort i denne oppgaven er å ta hensyn til andre variabler, som vi kan ha grunn til å tro at påvirker gassprisen. Dette er variabler som for eksempel temperatur, økonomi og andre energikilder. I et videre studie kunne vi testet om noen av disse variablene var signifikante. Å regne med denne avhengigheten, om det er noen, kunne ført til mer realistiske prediksjoner for prisene. Samtidig er det ikke sikkert at en mer komplisert modell ville blitt særlig mye bedre. Ideelt sett skal markedet i seg selv sluke andre avhengigheter som vær-avhengighet slik at denne ikke har en predikerende effekt.

Multivariat GARCH-modell kunne blitt brukt i et videre prosjekt. Da kunne vi sett på hvor stor effekt for eksempel volatiliteten til TTF-markedet har på volatiliteten NCG-markedet. Vi kunne også sammenlignet multivariat GARCH-modell, med copula modelleringen som er gjort i denne oppgaven. I denne oppgaven har vi kun funnet prognosen for en dag frem i tid. I en videre analyse kunne vi predikert prisen for flere dager frem i tid. Konklusjonene kunne blitt bedre, ved at vi kunne sett på ulike mål for feil predikering, som for eksempel gjennomsnittlig kvadratfeil (RSE).

Et annet tema som kunne som kunne blitt med i et videre studium, er om vi kunne brukt modellene i denne oppgaven til å tjene penger. Vi observerte at det var kointegrasjon mellom de ulike gassprisene. Denne kointegrasjonen kunne blitt brukt for å tjene penger. Anta at gassprisen i NCG markedet avviker en del positivt i forhold til de andre gassprisene. Siden vi har kointegrasjon vil denne avvikingen over tid normalisere seg igjen. Da vil det lønne seg å selge gass i NCG marked nå, og kjøpe gass i NCG markedet igjen når gassprisen har kommet tilbake til det forventete nivået i forhold til de andre prisene. Da vil man sitte igjen med et overskudd. Samtidig må vi huske på at det ofte er transaksjonskostnader og forskjellige kjøps og salgspriser knyttet til denne typen spekulering.

Det finnes flere måte å tjene penger i gassmarkedet. Volatilitets modeller er mye brukt i praksis for å tjene penger. Ulike investeringsgrupper bruker forståelsen av volatilitet til ulike formål. For

eksempel vil en risiko manager vite sannsynligheten idag for at porteføljen vil falle i fremtiden. Prisvolatilitet kan dermed oppfordre eller fraråde bedrifter mot å investere. I handel gir økende prisvolatilitet oppmuntring til en arbitrasjemulighet. Det kunne også vært mulig å studert om langtidskontrakter for gasspris har en sammenheng med spotprisen. Hvis dette hadde vært tilfelle, kunne vi stilt spørsmålet om det er mulig å tjene penger på å vite denne sammenhengen?

Bibliografi

- Aas, K. (2004). Modelling the dependence structure of financial assets: A survey of four copulas. *Norwegian Computing Center*, 1:s. 1–19. 48, 51
- Aftenposten (2014). Who: Luftforurensing tar sju millioner liv årlig. <http://www.aftenposten.no/nyheter/uriks/Luftforurensning-tar-syv-millioner-liv-arlig-7514544.html>. Lest: 28.11.2014. 17
- Bauwens, L., Hafner, C., and Laurent, S. (2011). Volatility models. *Core discussion paper*, 1(58):1–56. 43
- BBC (2015). <http://www.bbc.com/news/world-europe-27308526>. Lest 28.11.2014. xiii, 19
- Brockwell, P. J. and Davis, R. A. (1996). *Introduction to Time Series and Forecasting*. Springer texts in statistics. Springer-Verlag New York Inc., 1 edition. 21, 22, 26, 27, 28, 29, 30, 35, 36, 37, 38, 74
- BusinessGreen (2013). Combined cycle gas turbine (ccgt). <http://www.businessgreen.com/bg/glossary/1807455/combined-cycle-gas-turbine-ccgt>. Lest: 2013. 18
- Carmona, R. (2014). *Statistical Analysis of Financial Data in R*. Springer, 2 edition. 46, 53, 85, 86, 93
- Clarke, K. (2007). A simple distribution-free for nonnested model selection. *Political Analyses*, 15. 53
- Energifakta.no (2013a). Naturgass. <http://www.energifakta.no/documents/Energi/Ressurser/Naturgass.htm>. Lest: 2013. 5
- Energifakta.no (2013b). Råolje og naturgass. http://www.energifakta.no/documents/enkel20innforing/energi/Ressurser/forbr_petroleum.htm. Lest: 2013. 4
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50:987–1007. 36, 44
- EU-kommisjonen (2012). Quarterly report on european gas markets. *Market observatory for energy*, 4. 7
- EU-kommisjonen (2013). Quarterly report on european gas markets. *Market observatory for energy*, 6. 7, 10, 12, 58

- EU-kommisjonen (2014a). 2030 framework for climate and energy policies. http://ec.europa.eu/clima/policies/2030/index_en.htm. Lest 2014. 17
- EU-kommisjonen (2014b). Quarterly report on european gas markets. *Market observatory for energy*, 6. 7
- EU-kommisjonen (2014c). Quarterly report on european gas markets. *Market observatory for energy*, 7. 7, 58
- europeunconventionalgas (2014). About unconventional gas. <http://www.europeunconventionalgas.org/unconventional-gas>. Lest: 2014. 5, 6
- Fabini, C. (2012). Price forecasts in the european natural gas markets. *Masters Thesis*. v
- Fichman, B. T. (2012). August 2012 monthly energy review. <http://www.eia.gov/totalenergy/data/monthly/pdf/mer.pdf>. The U.S. Energy Information Administration's, (s. 38). 16
- Francq, C. and Zakoian, J. M. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley. John Wiley and Sons Ltd, 1 edition. 40, 41, 43, 102
- Gassco (2012). Ord og uttrykk. <http://www.gassco.no/wps/wcm/connect/gassco-no/Gassco/Home/presse/ord-og-uttrykk/>. 8
- Genest, C. R. B. and D., B. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Mathematics and Economics*, 44. 53
- Gjelstenli (2012). Fyllestasjoner norge. <http://www.gjelstenli.no/lpg/FyllestasjonerNorge.htm>. 16
- Gregoire, V., Genest, C., and Gendron, M. (2008). Using copulas to model price dependence in energy markets. *Energy Risk*. v, 35, 41, 86, 89
- Gudding, P. (2012). Verdens naturgass reserver. <http://miljookonomene.blogspot.no/2012/03/verdens-naturgass-reserver.html>. Miljøøkonomene. 20
- HiST (2015). Produktkrav gass. <http://hist-materialer-og-energi.wikispaces.com/Produktkrav+gass,HiST>. Lest 16.05.2015. xiii, 4
- Honore, A. (2010). *European natural gas demand, supply, and pricing*. Oxford University Press, 1 edition. 5, 6, 10, 11, 12, 13, 14, 15, 16, 17, 18
- Indexmundi (2015). World dry natural gas consumption by year. <http://www.indexmundi.com/energy.aspx/?product=gas&graph=consumption>. Lest 16.05.2015. xiii, 6
- Likvern, R. (2012a). Bp statistical review 2012 og naturgass. <http://www.fractionalflow.wordpress.com/2012/07/12/bp-statical-review-2012-og-natugass/>. 10, 14
- Likvern, R. (2012b). Skifergass. <http://www.fractionalflow.wordpress.com/2012/08/08/skifergass/>. 13

- Lundberg, Nesse, and Hagland (2009). Naturgass. <https://snl.no/naturgass>. lest 15.12.2014. 19, 20
- Lundberg, N. H. (2012). Naturgass. <http://snl.no/naturgass>. Store Norske leksikon. 4, 17
- Metreologisk, i. (2009). 2009 blant de varmeste ti globalt. http://met.no/2009+blant+de+varmeste+ti+globalt.b7C_wlvK5A.ips. 57
- Metreologisk, i. (2010). Desember 2009:. http://met.no/Desember+20093A.b7C_wlv00r.ips. 57
- Mæland, K. (2009). Har gitt opp norske gassbiler. <http://www.nettavisen.no/nyheter/article2770211.ece>. Nettavisen. 16
- Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer Science & Business Media, 2 edition. 45, 47, 49, 50, 51
- NetConnectGermany (2015). Reference prices. <https://www.net-connect-germany.de/en-gb/Transparency-information/Prices/Reference-Prices>. 55
- NUPI (2015). Skiferrevolusjonen hva betyr den for verden. <http://hvorhenderdet.nupi.no/Artikler/2013-20142/Skiferrevolusjonen-hva-betyr-den-for-verden/Arbeidsoppgaver>. Lest 16.05.2015. xiii, 5, 18
- Obžigailov, S. (2011). Modeling and forecasting demand for natural gas of retail consumers. *Master's thesis*. v
- Oettinger, G. H. (2010). Quarterly report on european gas markets. *Market observatory for energy*, 3. 57
- Oljeogenergidepartementet (2011). En næring for framtida – om petroleumsvirksomheten. <http://www.regjeringen.no/nb/dep/oed/dok/regpubl/stmeld/2010-2011/meld-st-28-2010-2011/3/2.html?id=649738>. 7, 8, 9, 14, 19
- Pfaff, B. (2008). *Analyses of Integrated an Cointegrated Time Series with R*. Springer texts in statistcs. Springer-Verlag New York Inc., 1 edition. 25, 30, 31, 33, 95, 98
- Raytchev, D. (2011). Quarterly report on european gas markets. *Market observatory for energy*, 4. 57
- Reider, R. (2009). Volatility forecasting i: Garch models. http://cims.nyu.edu/~almgren/timeseries/Vol_Forecast1.pdf. lest: 23.05.2015. 42
- Ruppert, D. (2011). *Statistics and Data Analysis for Financial Engineering*. Springer Science Business Media, LLC, 1 edition. 25, 26
- Statoil (2010). Gas sales agreements. <http://www.statoil.com/annualreport2010/en/ouoperations/businessareas/naturalgas/pages/gassalesagreements.aspx>. 8
- Statoil (2012). Kollsnes prosessanlegg. <http://www.statoil.com/no/ouoperations/terminalsrefining/processcomplexkollsnes/pages/default.aspx1>. lest 04.09.2012. 8

- Tasic, T. (2011). http://maja.uni-mb.si/files/APEM/APEM6-2_087-094.pdf. 16
- Tollaksen, T. G. (2013). Norge kan tape 20 mrd. kroner. <http://www.aftenbladet.no/energi/Norge-kan-tape-20-mrd-kroner-3147844.html>. 6, 8, 11
- Tsay, R. S. (2005). *Analysis of Financial Time Series*. Wiley Series in Probability and Statistics, 2 edition. 24, 31, 32, 33, 36, 44, 65, 66, 70
- Winje, P. (2011). Økt gasseskport, men hva med prisen. http://www.norges-bank.no/pages/87141/AK_4_2011.pdf. 8, 9, 10, 18