

Analyse og simulering af prædiktionsfejl for vindenergiproduktion ved indmelding til NordPool

Henrik Aalborg Nielsen & Henrik Madsen
Informatik og Matematisk Modellering
Danmarks Tekniske Universitet
Lyngby

20. februar 2002

Indhold

1	Resumé	2
2	Indledning	3
3	Data	4
4	Analyse	6
4.1	Middelværdi og varians	6
4.2	Auto- og krydskorrelation	7
5	Modellering	9
5.1	Spline modellering	9
5.2	Tidsrækkemodel	13
5.3	Tidsrækkemodel med ikke-konstant varians	19
6	Simulation	22
6.1	Simulation svarende til oprindelige data.	22
6.2	Generel simulation	27
7	Konklusion	37
A	Plots	39
B	Numeriske værdier til simulering	52
	Litteratur	53

1 Resumé

I denne rapport analyseres og modelleres prædiktionsfejl for den samlede elproduktion på vindmøller i det Jysk-Fynske område. Prædiktionerne er i denne sammenhæng indmeldinger til NordPool foretaget af Eltra klokken 12:00 og gældende for det efterfølgende døgn, dvs. fra 00:00 til 24:00.

Formålet med modelleringen er at kunne simulere nye tidsrækker af prædiktionsfejl. Disse påtænkes anvendt i forbindelse med systemsimuleringer foretaget med SIVAEL. Det er i rapporten vist, at de simulerede tidsrækker har egenskaber, der svarer til de oprindelige datas egenskaber.

2 Indledning

I denne rapport analyseres prædiktionsfejl for den samlede elproduktion på vindmøller i det Jysk-Fynske område. Prædiktionerne er i denne sammenhæng indmeldinger til Nord-Pool foretaget af Eltra klokken 12:00 og gældende for det efterfølgende døgn. Datagrundlaget er den samlede elproduktion fra vindmøller placeret på Jylland og Fyn for perioden 1/1 til 31/12 år 2000. Disse prædiktioner er i stor udstrækning baseret på on-line systemet WPPT (Wind Power Prediction Tool) [11, 10, 9], men Eltra kan afvige fra disse, såfremt det skønnes relevant. Ikke alle prædiktioner fra WPPT behandles, men der fokuseres derimod på den prognose, der genereres klokken 11:00 og gælder for det følgende døgn, dvs. for tidshorisonter fra ca. 0.5 til 1.5 døgn.

Formålet med analysen er først og fremmest at muliggøre simulation af prædiktionsfejl, idet dette er interessant af hensyn til systemsimuleringer med SIVAEL [4, 12]. Sekundært kan en sådan analyse være nyttig til at påvise potentielle forbedringer af WPPT.

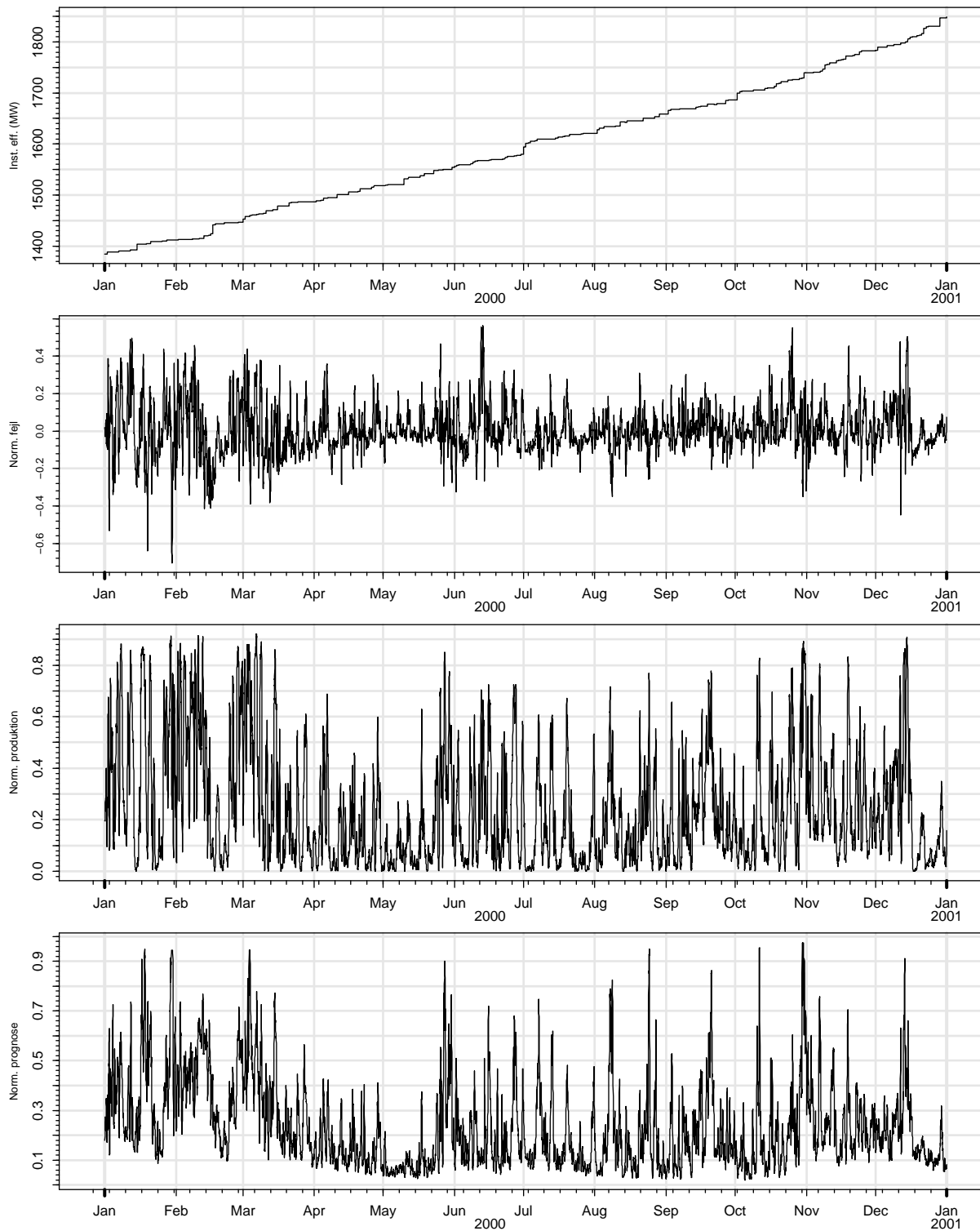
I kapitel 3 beskrives data og en indledende analyse af disse er foretaget i kapitel 4. Modeleringen beskrives i kapitel 5, og i kapitel 6 vises, det hvorledes simuleringen kan foretages, se også konklusionen i kapitel 7. Alle numeriske værdier, nødvendige for at kunne udføre simulationen, er gemt som i regnearks-format (CSV-format). Filnavne er anført i appendiks B.

3 Data

Data for vindkraften består af faktiske og prognostiserede (cf. afsnit 2) energier målt i MWh og for hvert kvarter i år 2000. Desuden er data for den installerede effekt tilgængelige.

WPPT forudsiger imidlertid elproduktionen på basis af 30 minutters intervaller (hele og halve timer). Prognoser baseret på 15 minutters intervaller er fundet ved interpolation. Disse data er omregnet til 30 minutters intervaller igen ved for faktiske produktioner at summere over halve timer. For de prognostiserede værdier multipliceres værdierne ved hele og halve timer blot med 2, derved genskabes de oprindelige prædiktioner.

Selve analysen foretages på basis af relativ produktion og prognose, hvorved forstås produktionens / prognosens relative størrelse i forhold til den installerede effekt. Inden denne normering foretages, er produktionen i *MWh* for hvert 30 min. interval først omregnet til gennemsnitlig effekt. På denne måde fås et tal mellem 0 og 1, der angiver hvor stor en del af vindmøllernes kapacitet, som udnyttes. Disse data er vist i figur 1. Endelig bemærkes, at tidspunkter relaterer til udgangen af det forløbne interval.



Figur 1: Installeret effekt, fejl, produktion og prognose. De tre sidste er normeret med den installerede effekt.

4 Analyse

Data analyseres hovedsaglig som en multivariat tidsrække, hvor der for hvert døgn haves en observation bestående af de normerede prædiktionsfejl for det aktuelle prognosetidspunkt det efterfølgende døgn. Observationerne tidsmærkes med det tidspunkt, hvor prædiktionerne blev genereret. De tre første og sidste observationer ser således ud:

Prognosetidspkt.	00:30	01:00	01:30	...	23:30	24:00
1999-12-31 11:00	0.018	-0.00039	-0.0081	...	-0.033	-0.0051
2000-01-01 11:00	-0.088	-0.08565	-0.0812	...	-0.518	-0.5320
2000-01-02 11:00	-0.044	-0.03920	-0.0388	...	0.070	0.0880
.
.
2000-12-28 11:00	0.015	0.02593	0.0143	...	0.046	0.0352
2000-12-29 11:00	0.085	0.07490	0.0722	...	0.038	0.0358
2000-12-30 11:00	-0.017	-0.01965	-0.0231	...	0.060	0.0756

Antallet af kolonner er 48, og antallet af observationer er 366. Ved skift fra vinter til sommertid findes der ikke data for 02:00 og 02:30. Dette kan være problematisk bla. ved estimation af auto- og kryds-korrelationer. I disse tilfælde dannes værdier for 02:00 og 02:30 vha. lineær interpolation i den pågældende række. Dette synes rimeligt, idet det viser sig at korrelationen inden for rækker er positiv og forholdsvis stærk.

Med mindre andet er nævnt er tidspunkter i dansk standardtid (MET), dvs. vinter/sommertid.

4.1 Middelværdi og varians

Figurene 28 og 29 (side 40 og 41) viser middelværdi og varians af prædiktionsfejlene mod klokkeslet. Der er vist plot på basis af hele år 2000 og for hvert kvartal for sig. Det bemærkes, at den højeste gennemsnitlige fejl er ca. 0.04, hvilket giver bidrag på ca. 0.002 til MSE (Mean Squared Error). Som det fremgår af figur 29 er variansens bidrag langt større; ca. 10 gange.

Gennemsnitlig prædiktionsfejl: Der ses en tydelig døgnvariation, der varierer mellem kvartalerne. Dog er variationen i den gennemsnitlige fejl markant mindre end selve fejlene, cf. figur 1.

Varians af prædiktionsfejl: Variansen af prædiktionsfejlene udviser døgnvariation, der varierer mellem kvartalerne. Den maksimale varians forekommer især i 2. og 3. kvartal

i løbet af eftermiddagen. Dette tyder på, at maksimum er relateret til solopvarmning. I 1. kvartal er variansen markant højere end i de øvrige kvartaler. Dette skyldes formentlig, at der i 1. kvartal forekommer nogle perioder med meget kraftig vind og bratte overgange fra stille til blæsende vejr. Generelt ses en tendens til at variansen stiger gennem døgnet, dvs. med prognosehorisonten.

Korrelation af prædiktionsfejl: Figur 30 (side 42) viser korrelationen af prædiktionsfejlene inden for døgn, se også [2]. Døgnet, hvor der skiftes fra vinter til sommertid, er udeladt af beregningerne. Igen ses en forskel mellem kvartaler, hvor det her specielt er 3. kvartal, der skiller sig ud. En nærmere undersøgelse viser at korrelationen er over 0.7, når klokkeslettene adskiller sig med mindre end 4 timer.

4.2 Auto- og krydskorrelation

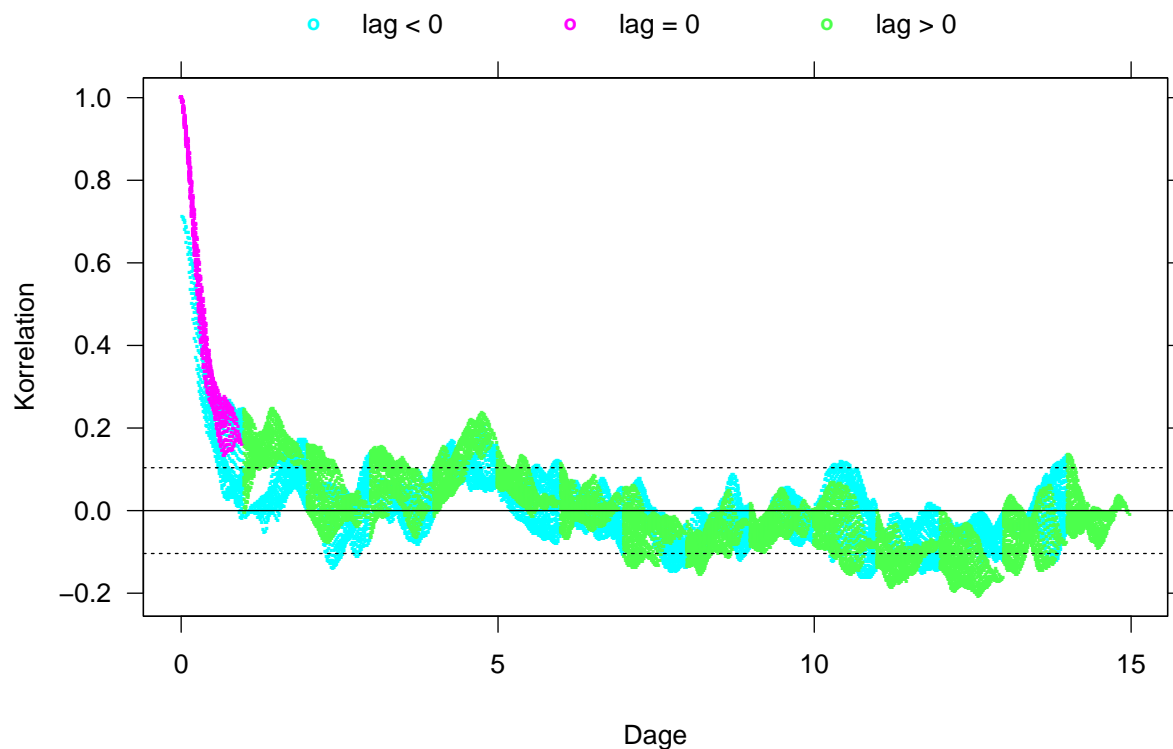
Autokorrelation: Figur 31 (side 43) viser de estimerede autokorrelationsfunktioner for de enkelte tidsrækker, se også [7]. Figurene 32–35 (side 44–47) viser tilsvarende plot for hvert kvartal. De horisontale linier indikerer $\pm 2/\sqrt{N}$, hvor N er antallet af observationer, der indgår i estimationen. Såfremt den underliggende proces er hvid støj vil ca. 95% af estimerterne for lag ≥ 1 ligge inden for disse grænser.

Det ses, at de enkelte tidsrækker er tæt på hvid støj. Dog findes der en smule korrelation i lag 1 for tidspunkter efter ca. kl. 15. Noget tilsvarende synes at gælde, når de enkelte kvartaler undersøges særskilt, dog synes korrelationen i lag 1 især at komme fra 3. og 4. kvartal.

Krydskorrelation: Undersøgelse af estimerede krydskorrelationsfunktioner [7] for de 48 tidsrækker kompliceres af, at antallet er meget stort. I det følgende undersøges krydskorrelationsfunktioner mellem tidsrækker svarende til to klokkeslet, hvor den første tidsrække svarer til det laveste klokkeslet. Dette resulterer i 1128 kombinationer. Figur 36 (side 48) viser histogrammer, af de 1128 værdier for lags fra -14 til 14 (lag 0 svarer til figur 30 (side 42)). De vertikale linier indikerer $\pm 2/\sqrt{N}$. Det ses at specielt lag -1 og 1 har krydskorrelationer signifikant forskellige fra 0. For disse lag viser figur 37 og 38 (side 49 og 50) de enkelte korrelationer. Over hvert plot er et klokkeslet anført, og plottet viser korrelationer med efterfølgende klokkeslet, f.eks. viser plottet for lag -1 med klokkeslettet 00:30, at korrelationen med fejlen 24:00 dagen før er ca. 0.7, hvor det bemærkes, at de tilsvarende prædiktioner er genereret med et døgn forskel. Plottene tyder på, at en stor del af variationen i de estimerede korrelationskoefficienter skyldes forskellen i de tidspunkter, for hvilke prædiktionerne gælder (se også afsnit 4.1). Figur 2 viser dette. Det ses, at lag 0 og -1 er langt de vigtigste. Derudover er de fleste værdier i lag 1 også over $2/\sqrt{N}$, og endelig ses der en vis positiv korrelation omkring dag 4 (lag -5 og 4) samt en negativ

omkring dag 12. For lag 0 og -1 (dag 0) ses, at korrelationen foruden den dominerende tidsforskel også afhænger af lag, dvs. af hvornår prognosen er genereret.

Figur 39 (side 51) viser tilsvarende plot for hvert kvartal. For små tidsforskelle synes forskellen mellem kvartaler at være begrænset. Dette er ikke tilfældet for højere tidsforskelle, hvor hovedparten af estimaterne dog ligger indenfor $\pm 2/\sqrt{N_q}$ (N_q er antallet af observationer i et kvartal). Sidstnævnte forskel må derfor i stor udstrækning tilskrives tilfældige variationer.



Figur 2: Estimerede auto- og krydskorrelationer (SACF og SCCF) versus tidsforskellen mellem tidspunkter for hvilke prædiktionerne gælder. De stiplede linier svare til $\pm 2/\sqrt{N}$.

5 Modellering

For at kunne simulere prædiktionsfejlene er det oplagt at modellere disse som en multivariat tidsrække, jf. datastrukturen beskrevet i starten af afsnit 4. Analysen i afsnit 4.2 viser, at hovedparten af korrelationen findes i lag -1, 0 og 1, hvilket indikerer, at en multivariat MA(1) model (kaldet en MMA(1) model i det efterfølgende) er hensigtsmæssig. Dog bemærkes, at der synes at forekomme årsvariation i middelværdi og varians (afsnit 4.1). I første omgang ses der bort fra denne årsvariation.

Til estimation i multivariate tidsrækker benyttes S-PLUS med modulet S+GARCH (se evt. www.insightful.com eller www.adeptscientific.dk). Specifikt benyttes funktionen `mgarch`. Det viser sig, at konvergenssegenskaberne for estimationen er særdeles ringe. Nærmere undersøgelser viser, at dette skyldes korrelationen mellem tidsrækker hørende til klokkeslet, der ligger forholdsvis tæt, se også afsnit 4.1 og figur 30. Eksempelvis viser det sig, at det kræver mellem 400 og 500 iterationer at opnå konvergens, når syv tidsrækker, der er jævnt fordelt over klokkeslet udvælges¹.

I afsnit 4.2 og specielt figur 2 er det vist, at auto- og krydskorrelationer hovedsaglig afhænger af tidsforskellen mellem de tidspunkter, for hvilke prædiktionerne gælder. Dette giver mulighed for med ganske få parametre at modellere disse korrelationer. Imidlertid skal modellen for korrelationerne transformeres til en model for parametrene i MMA(1) modellen, og standard software vil ikke kunne benyttes ved estimation.

Af årsagerne beskrevet ovenfor vælges en mere simpel løsning. Som beskrevet i afsnit 4 er korrelationen inden for de enkelte prædiktions-tidspunkter (rækker) positiv og i overvejende grad jævnt aftagende med afstanden mellem tidspunkter, se også figur 30. Af denne grund vil en stor del af variationen inden for prædiktions-tidspunkter kunne forklares af en regression på en spline-basis med et forholdsvis lavt antal frihedsgrader. For at undgå stor variation ved 00:30 og 24:00 anvendes en naturlig kubisk spline basis [6, kap. 5]. De tilhørende koefficienter vil herefter kunne modelleres med en MMA(1) model.

Nedenfor beskrives spline-modelleringene i afsnit 5.1 og tidsrækkemodelleringen i afsnit 5.2.

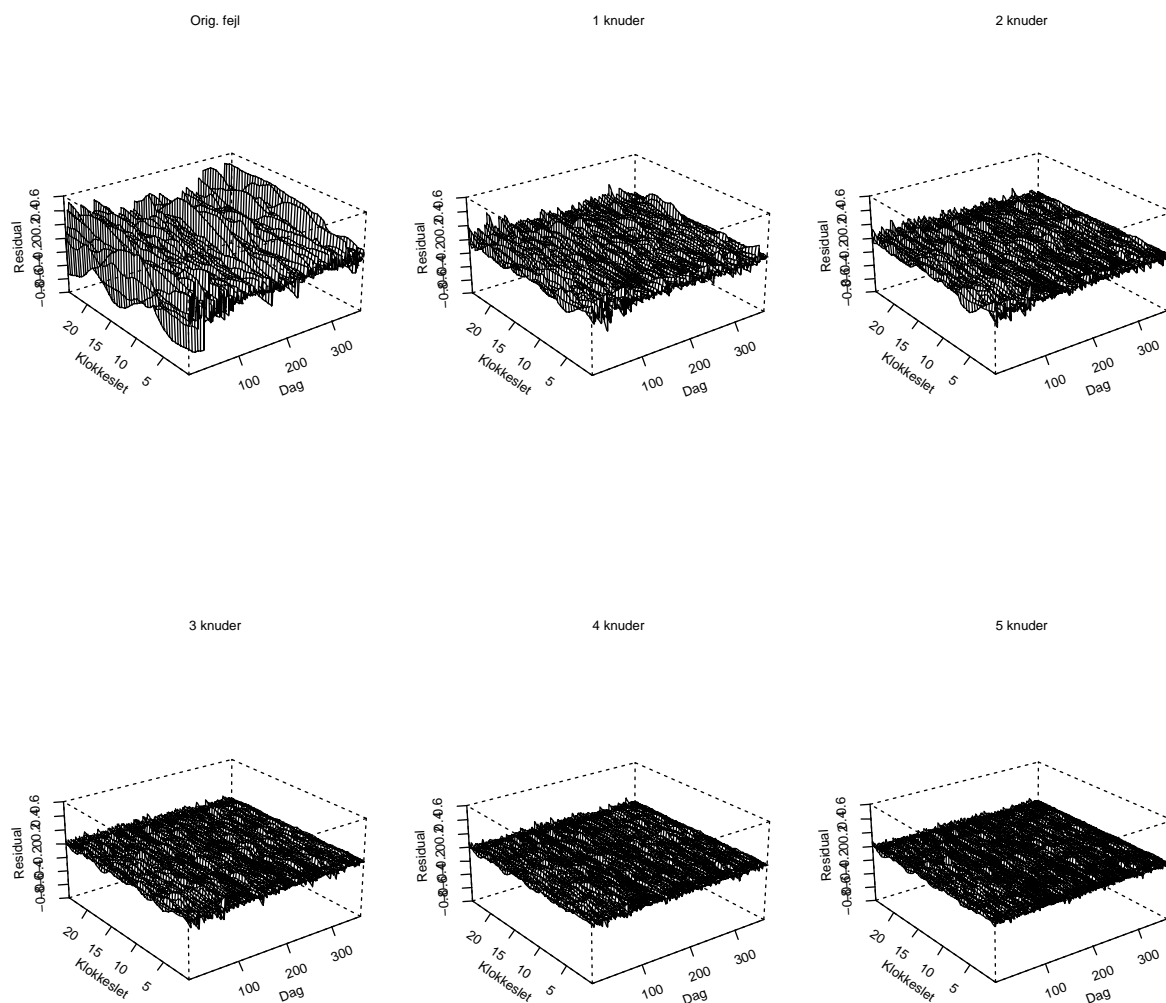
5.1 Spline modellering

Spline baser med ækvidistante knuder og med grænse-knuder ved 0.5 (00:30) og 24 (24:00) anvendes.

Figur 3 viser de originale prædiktionsfejl samt residualerne efter at en naturlig spline basis med 1 til 5 interne og ækvidistante knuder fittet til prædiktionsfejl for de enkelte dage.

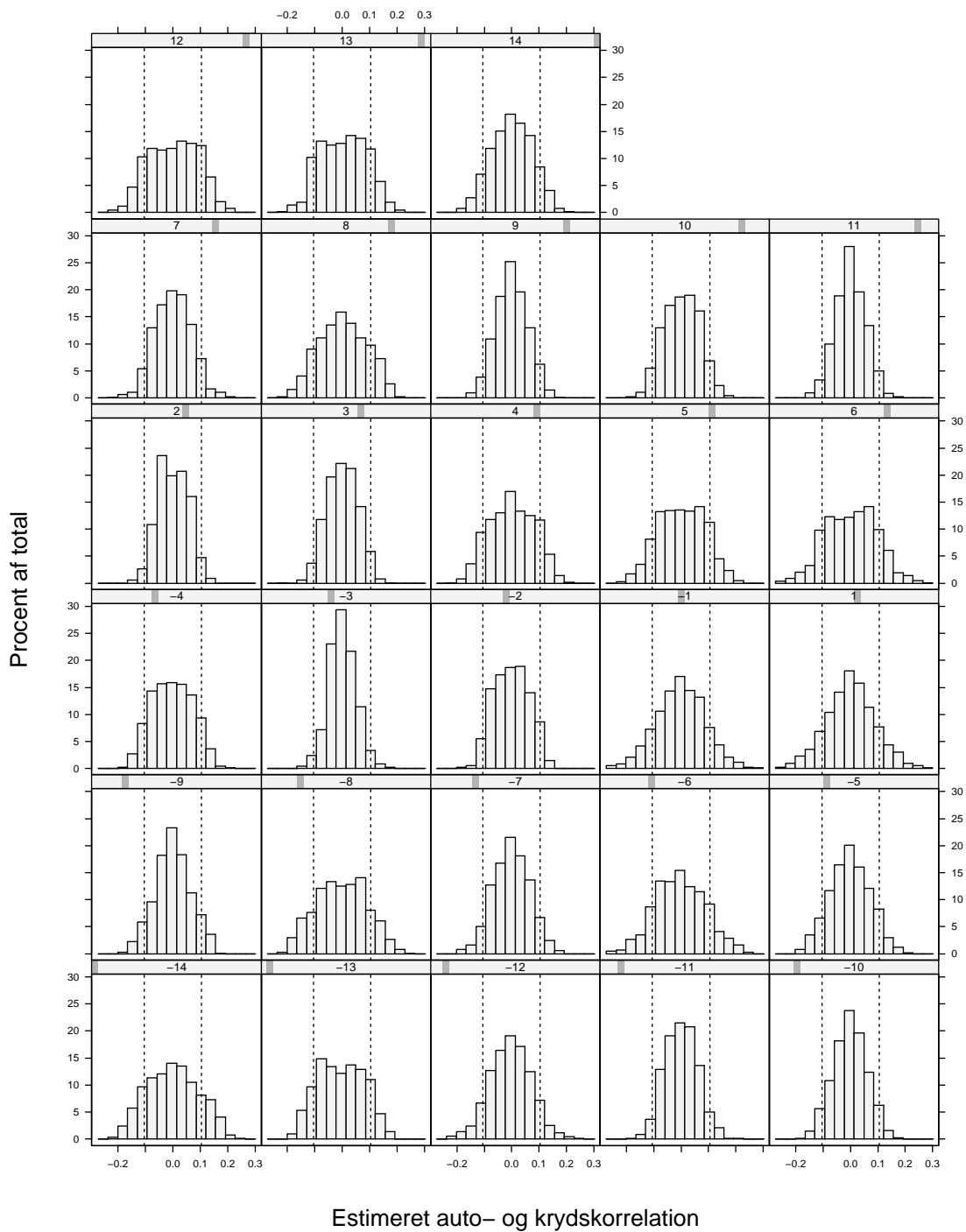
¹ Klokkeslet: 00:30, 04:30, 08:30, 12:00, 16:00, 20:00 og 24:00

Det ses, at allerede ved 3 interne knuder er en væsentlig del af variationen forklaret.

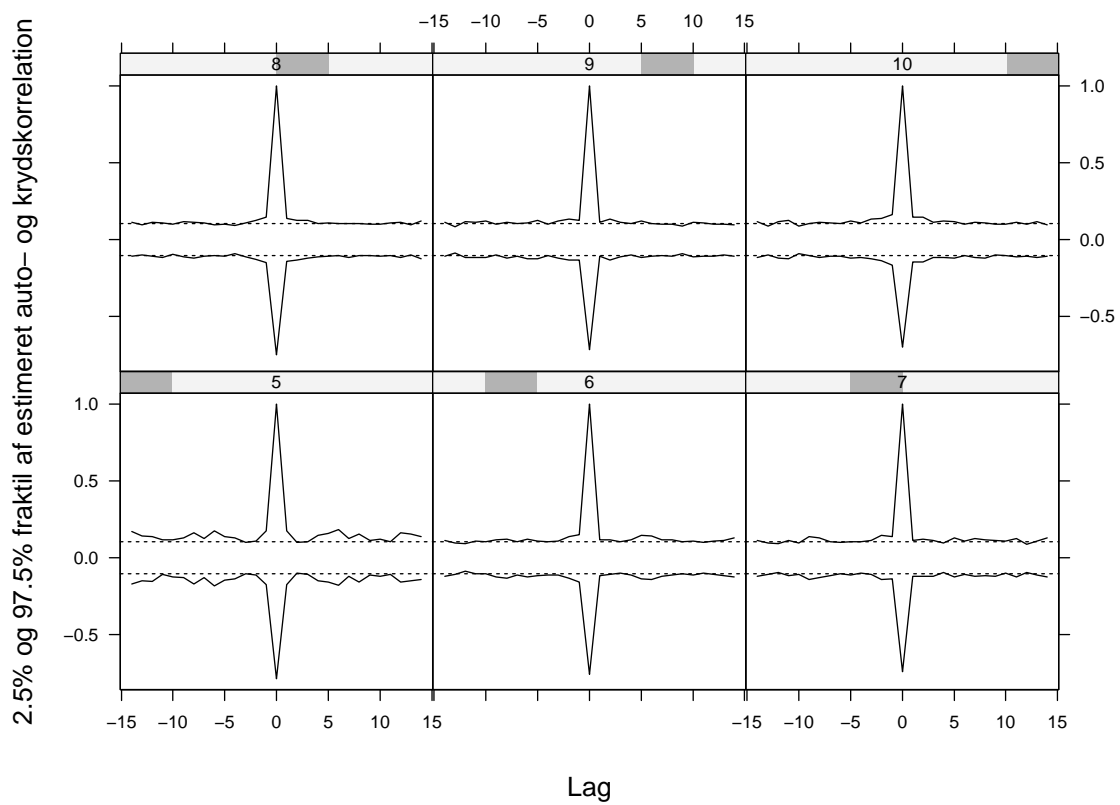


Figur 3: Originale prædiktionsfejl samt residualerne efter at en naturlig spline basis med 1 til 5 interne og ækvidistante knuder er fittet til prædiktionsfejl for de enkelte dage.

Figur 4 viser histogrammer over estimerede auto- og krydskorrelationer for residualerne, når 5 interne knuder anvendes. Det ses, at der forekommer mere end 5% uden for $\pm 2/\sqrt{N}$ ved f.eks. lag -1 og 1. I hvor stor udstrækning $\pm 2/\sqrt{N}$ overskrides er dog svært at afgøre ud fra figur 4. Dette fremgår mere tydeligt af figur 5, der viser de faktiske fraktiler sammen med $\pm 2/\sqrt{N}$. Det ses, at overensstemmelsen er langt bedre for 6 interne knuder end for 5 interne knuder, samt at overensstemmelsen forekommer at være optimal ved 8 interne knuder. For at begrænse antallet af parametre i den multivariate tidsrækkemodel vælges dog at arbejde med 6 interne knuder.



Figur 4: Histogrammer af estimerede auto- og krydskorrelationer for residualer når 5 interne knuder benyttes. De vertikale linier viser $\pm 2/\sqrt{N}$.



Figur 5: 2.5% og 97.5% fraktiler af de estimerede auto- og krydskorrelationer for residualerne når naturlige splines med 5 til 10 interne knuder fittes til data for hvert døgn for sig. De horisontale stiplede linier viser $\pm 2/\sqrt{N}$.

5.2 Tidsrækkemodel

I dette afsnit modelleres spline-koefficienterne ved 6 interne og ækvidistante knuder (se ovenfor). Som indikeret i afsnit 4.2 vil en multivariat moving average model af orden 1 (MMA(1)) formentlig være tilstrækkelig. Desuden afprøves en autoregressiv model af orden 1 (MAR(1)), samt kombinationer af de to modeller (MARMA(1,1)).

Generelt kan en MARMA(1,1) model skrives som

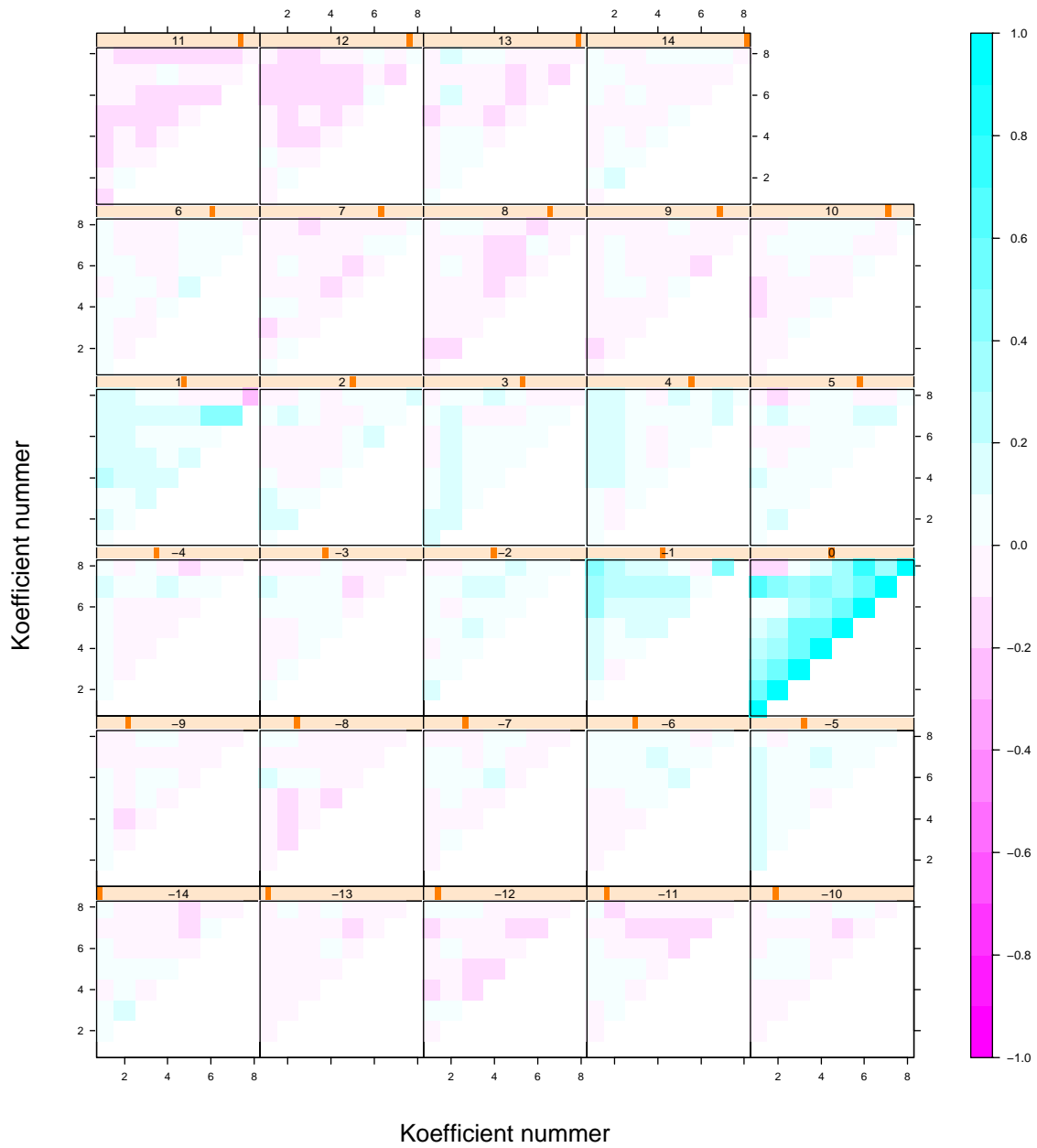
$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\Theta}_1 \boldsymbol{\epsilon}_{t-1} + \boldsymbol{\epsilon}_t, \quad (1)$$

hvor \mathbf{y}_t er en søjlevektor af dimension p indeholdende spline-koefficienterne til tid t . $\boldsymbol{\Phi}_1$ og $\boldsymbol{\Theta}_1$ er $p \times p$ matricer med parametre og $\{\boldsymbol{\epsilon}_t\}$ er en multivariat Gausisk hvid støj proces med middelværdi nul, hvis med varians $\boldsymbol{\Sigma}$. Endelig tager $\boldsymbol{\mu}$ højde for at middelværdien af $\{\mathbf{y}_t\}$ ikke nødvendigvis er $\mathbf{0}$. MAR(1) og MMA(1) modellerne fremkommer som specialtilfælde ved at udelade hhv. leddet med $\boldsymbol{\epsilon}_{t-1}$ og \mathbf{y}_{t-1} .

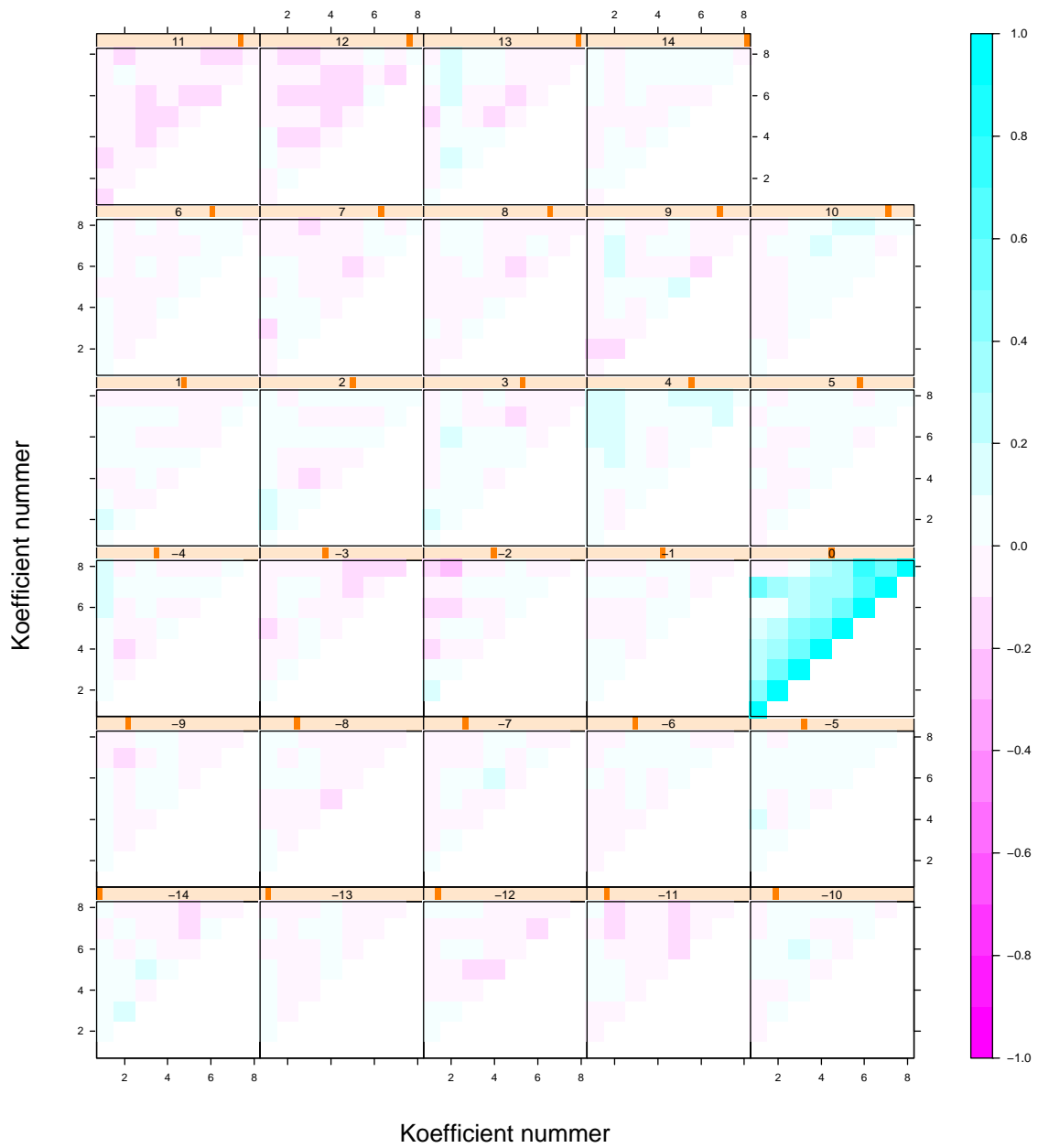
Figur 6 viser auto- og kryds-korrelationer for spline-koefficienterne og figur 7, 8, samt 9 viser tilsvarende plot for residualerne fra hhv. MAR(1), MARMA(1,1) og MMA(1) modellerne. Der er ikke nogen nævneværdig forskel på de tre sidstnævnte plot. Da MMA(1) er lettest at anvende til simulation, vælges det at arbejde videre med denne model.

Figur 10 viser tidsrækkeplot for de enkelte residual-serier. Det ses, at variansen ikke er konstant i løbet af perioden, og at der forekommer variationer, der ikke er modelleret. Som beskrevet ovenfor er de vigtigste korrelationer i lag -1 og 1 dog beskrevet.

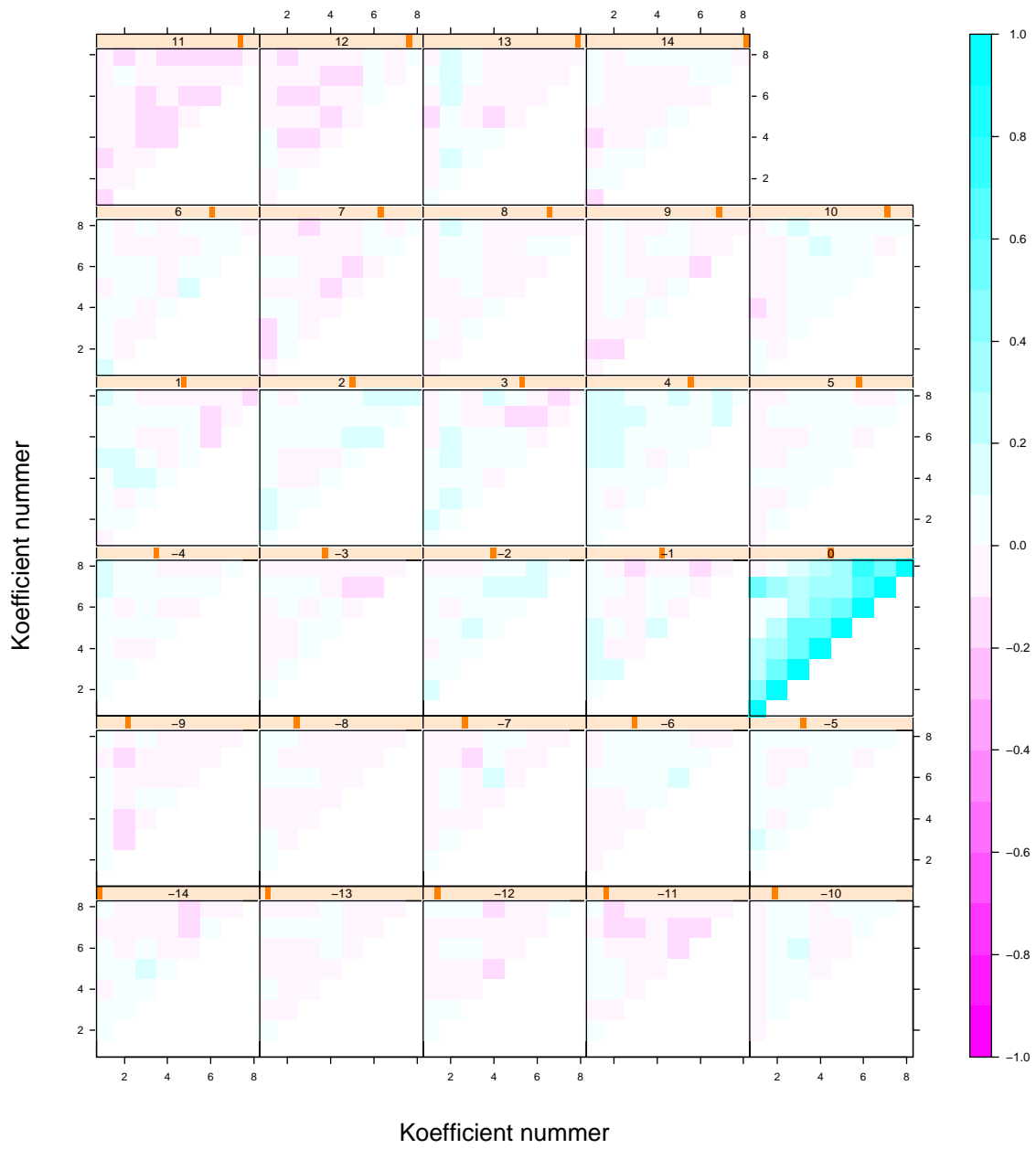
I afsnit 5.3, startende på side 19, er en simpel metode, hvorved der kan tages højde for den ikke-konstante varians, beskrevet.



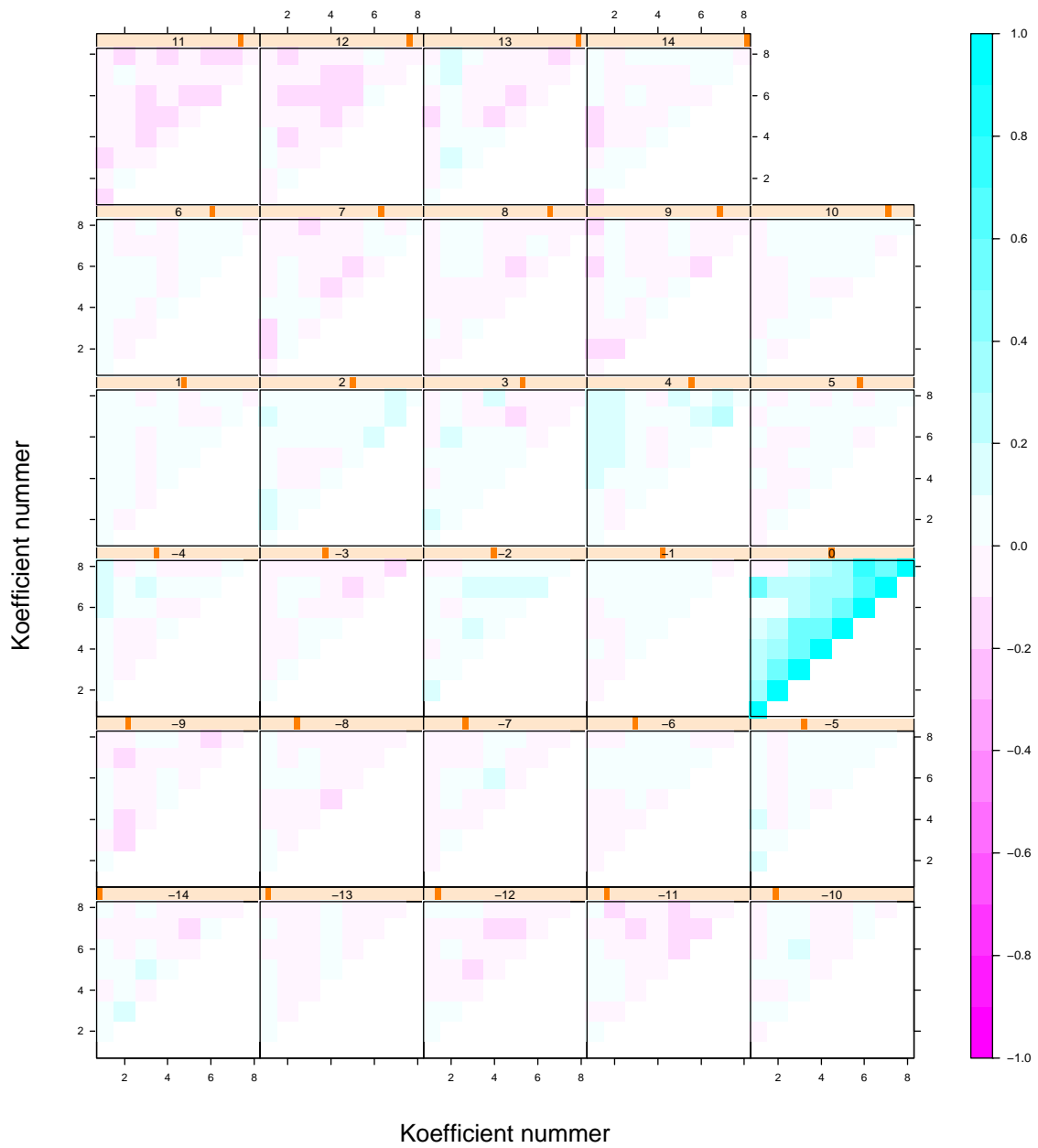
Figur 6: Estimerede auto- og krydskorrelationer for de 8 koefficienter i en naturlig spline fittet til de enkelte døgn.



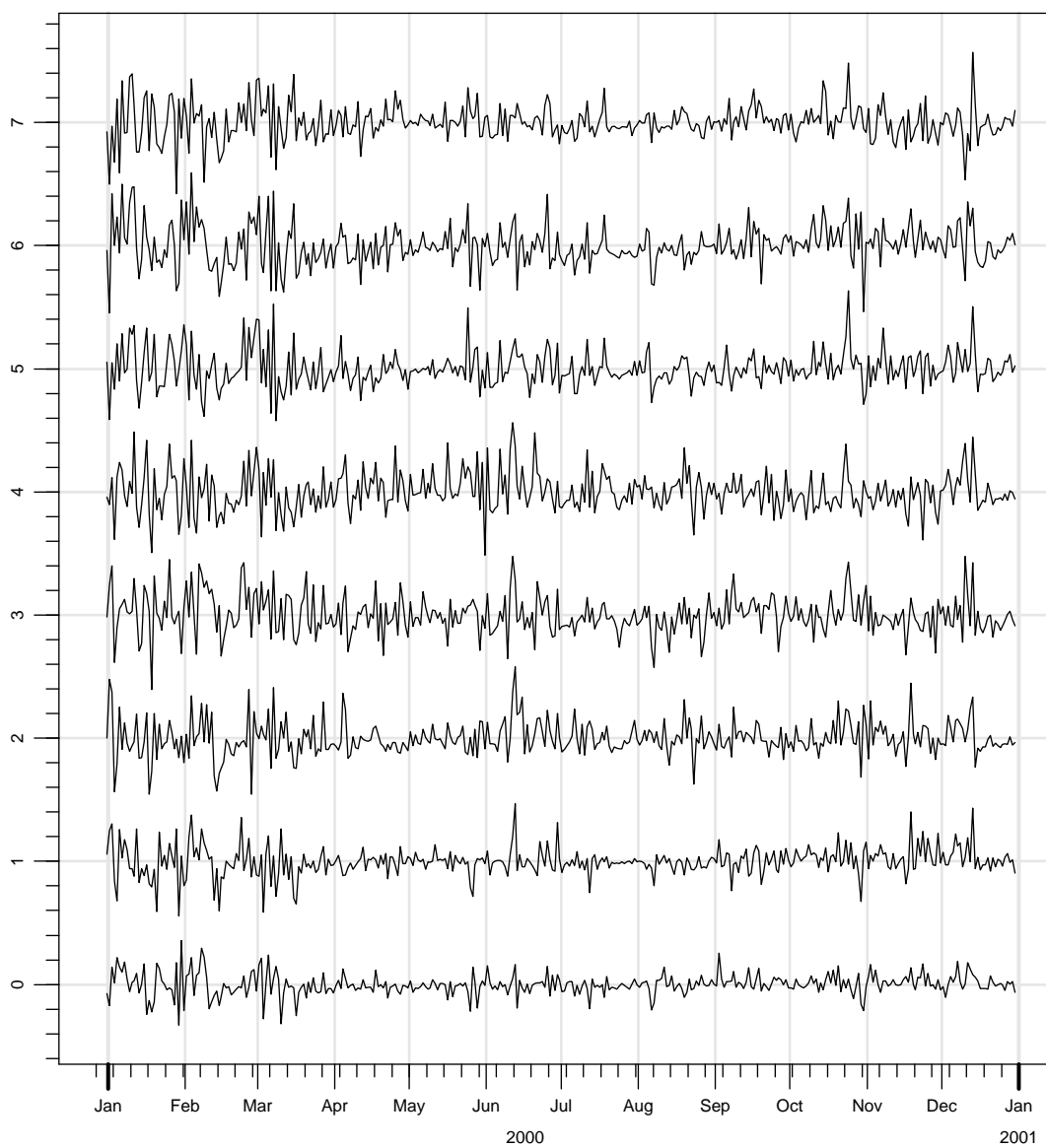
Figur 7: Estimerede auto- og krydskorrelationer for residualerne fra MAR(1)-modellen.



Figur 8: Estimerede auto- og krydskorrelationer for residualerne fra MARMA(1,1)-modellen.



Figur 9: Estimerede auto- og krydskorrelationer for residualerne fra MMA(1)-modellen.



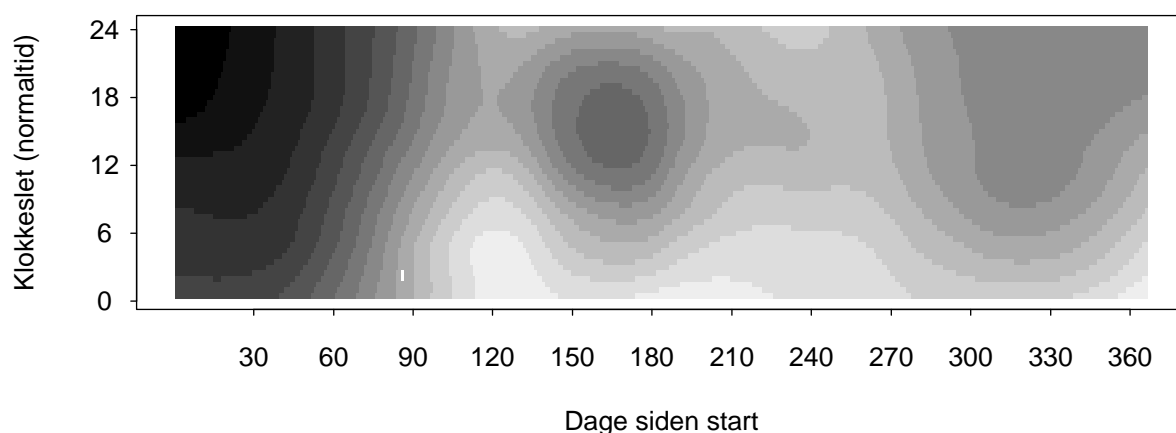
Figur 10: Residualerne for MMA(1)-modellen. Tidsrækkerne er plottet adskilt ved at addere hhv. 0, 1, ..., 7

5.3 Tidsrækkemodel med ikke-konstant varians

Som det fremgår af kapitel 4, er variansen ikke konstant over døgnet og over året. Desuden fremgår det, at middelværdiens bidrag til den forventede kvadratiske fejl er ringe. Af denne grund er det rimeligt at modellere den kvadratiske fejls afhængighed af tiden på året og døgnet samt at bruge denne model som et udtryk for, hvorledes variansen afhænger af de to variable.

Modelleringen foretages ved at estimere sammenhængen med lokal lineær regression, se [1], [5, kap. 3] eller [6, kap. 6]. Lad d_t betegne antallet af dage regnet fra starten af data sættet og h_t betegne tiden på døgnet i Dansk Normaltid (GMT + 1 time)². Udglatningen foretages nu vha. variablene $d_t/60$ og $h_t/12$, samt en nærmeste nabo båndbredde på $60/366 = 16.4\%$. Figur 11 viser standardafvigelsens (dvs. kvadratroden af variansen) variation med tid på år og døgn. Kvalitativt er plottet i overensstemmelse med figur 29.

Bemærkning: Betragtes et tidspunkt midt på dagen og året vil arealet af den cirkel, der svare til en nærmeste nabo båndbredde på $60/366$, være $24/12 \times 366/60 \times 60/366 = 2$, hvilket svare til en diameter på 1.6. Dette svare til ca. 80% i retningen bestemt ved $h_t/60$ og ca. 25% i retningen bestemt ved $d_t/12$. Sammen med en lokal lineær approksimation giver dette markant mere udglatning over døgnet end over året. Dette synes rimeligt ud fra resultaterne beskrevet i afsnit 4.1.



Figur 11: Standardafvigelsens afhængighed af tid på år og døgn. Lave værdier er vist med hvid og høje med sort (range: 0.0491 til 0.2316).

Prædiktionsfejlene standardiseres nu ved at dele de enkelte fejl med deres standardafvigelse bestemt ved figur 11. Dog justeres standardafvigelse bestemt ved figur 11 først med en proportionalitetsfaktor, således at den totale varians af de standardiserede prædiktionsfejl er 1³. Herefter approksimeres hvert døgn med en naturlig spline basis med 6

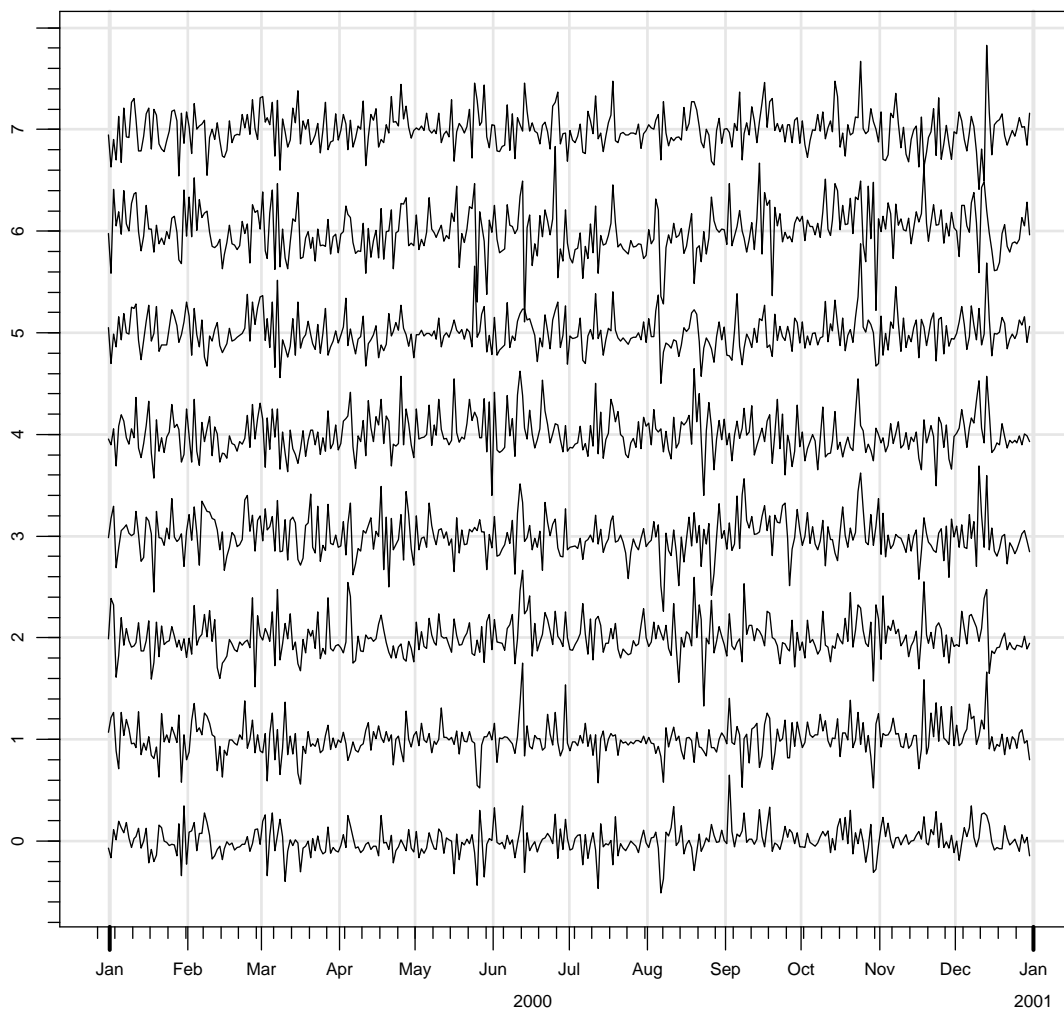
²Klokkeslettet for den længste prædiktionshorisont registreres som 24:00 (istedet for 00:00), derved bliver klokkeslettet også et udtryk for prædiktionshorisonten.

³Standardafvigelsen vist i figur 11 har, dels en lokal bias, der afhænger af den lokale krumning af

interne og ækvidistante knuder og en MMA(1)-model estimeres for spline-koefficienterne, begge dele som i afsnit 5.2. I appendiks B er der henvist til regnearksfiler indeholdende de estimerede koefficienter jævnfør model (1) på side 13. Bemærk dog at $\hat{\Sigma}$ er anført som korrelationsmatricen svarende til $\hat{\Sigma}$ og standardafvigelse svarende til diagonalen i $\hat{\Sigma}$.

Figur 12 viser residualerne for MMA(1)-modellen for de standardiserede tidsrækker. Det ses, at variansen som forventet er forholdsvis konstant over året. På dette sted vil modellens kvalitet ikke blive undersøgt nærmere. Derimod vil det i næste kapitel blive undersøgt, i hvor stor udstrækning modellen kan simulere tidsrækker, der svarer til de oprindelige data.

den sande (men ukendte) funktion, dels en overordnet bias der skyldes at de kvadrerede prædiktionsfejls afvigelse fra fladen ikke er symmetrisk fordelt.



Figur 12: Residualer fra MMA(1)-model for de standardiserede tidsrækker. De enkelte tidsrækker er delt med 6, hvorefter der er adderet 0, 1, ..., 7 til de 8 tidsrækker.

6 Simulation

I dette kapitel beskrives, hvorledes MMA(1) modellen fra afsnit 5.3 sammen med spline-approximationen fra afsnit 5.1 kan anvendes til simulering. I afsnit 6.1 undersøges, i hvor stor udstrækning modellen er i stand til at simulere tidsrækker, der svare til de oprindelige data. I år 2000 forekommer perioder med megen blæst især i de første 2-3 måneder, midt på året og i slutningen af året (figur 1). Det kan formodes, at det er dette forhold, der betyder at standardafvigelsen af fejlen afhænger af tiden på året (figur 11). I afsnit 6.2 undersøges metoder til at modellere, hvorledes standardafvigelsen af fejlen afhænger af andre forklarende variable.

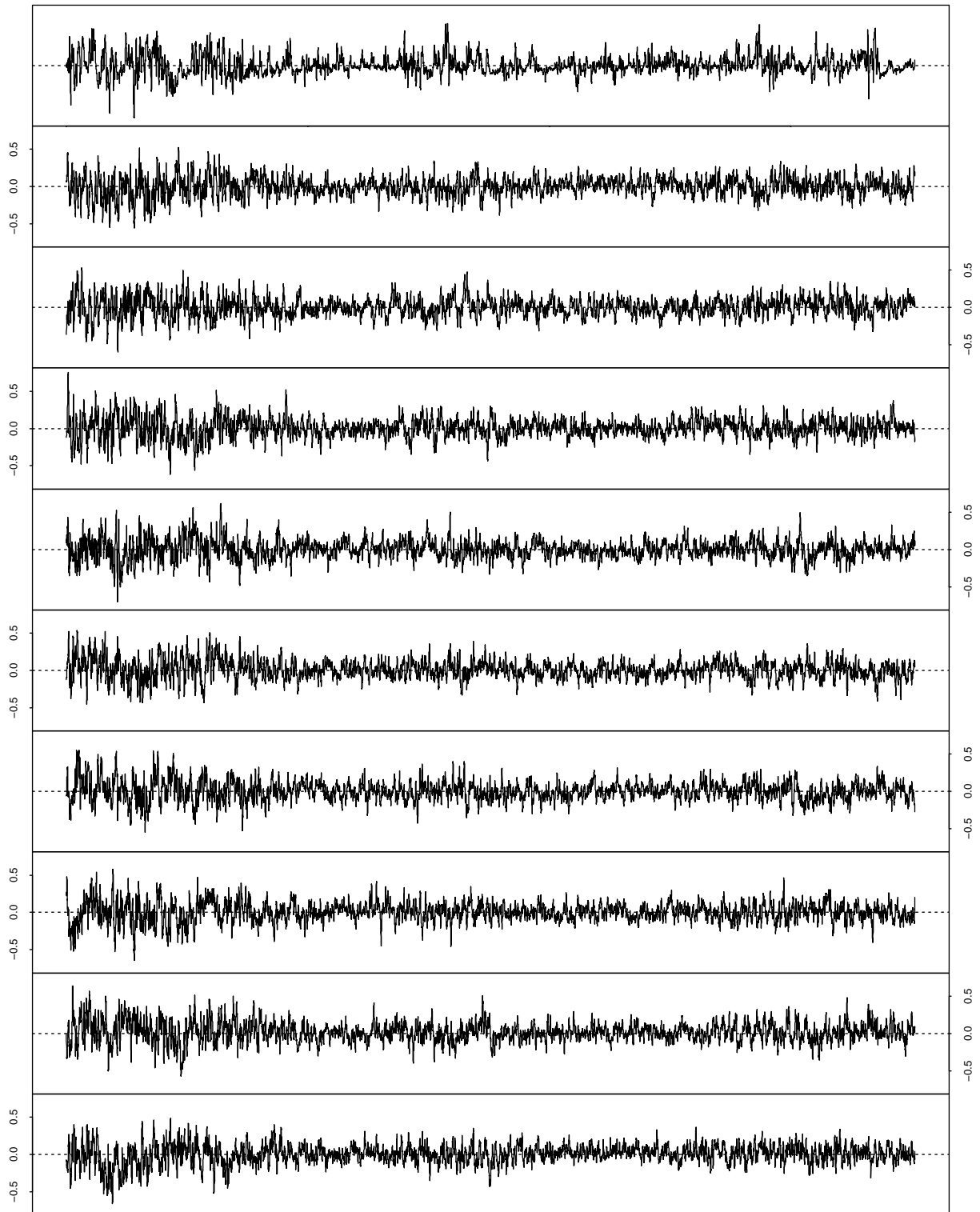
6.1 Simulation svarende til oprindelige data.

Nye tidsrækker simuleres ved at

1. Simulere 367 observationer fra en multivariat normalfordeling med korrelationsstruktur og standardafvigelser som de estimerede for ϵ_t i afsnit 5.3.
2. Bruge MMA(1)-modellen til at simulere spline-koefficienterne for dag 1 til 366 (vha. estimererne for μ og Θ_1 i afsnit 5.3).
3. Bruge spline-basen til at generere de standardiserede tidsrækker.
4. Bruge korrelationsstrukturen i spline-residualerne til at generere multivariat normalfordelt støj som adderes til tidsrækkerne fra det forrige trin.
5. Multiplicere tidsrækkerne med den standardafvigelse der blev anvendt i forbindelse med standardiseringen af de oprindelige normaliserede tidsrækker.

Figur 13 viser de oprindelige normaliserede prædiktionsfejl, samt 9 simulerede tidsrækker alle plottet mod tiden på året. Umiddelbart er overensstemmelsen god. I dette afsnit vil graden af denne overensstemmelse blive undersøgt nærmere.

For at undersøge i hvor stor udstrækning de enkelte simulerede tidsrækkers varians stemmer overens med de oprindelige data, simuleres 200 tidsrækker som beskrevet ovenfor. Herefter beregnes den empiriske varians for hver tidsrække, dvs. 48 værdier for hver simulering. På basis af disse 200×48 værdier bestemmes for hver tidsrække 5%, 50% og 95% fraktilerne. Disse er vist på figur 14 sammen med variansen af de oprindelige tidsrækker. Figuren viser også varianskurven for 10 af de 200 simulationer. Det ses, at overensstemmelsen i store træk er inden for det, man kan forvente givet længden af data-perioden (1 år). Bemærk dog at variansen er høj i starten og i slutningen af døgnet. Dette skyldes, at hvert døgn er approksimeret med en spline basis. Problemet er alment kendt [6], og det

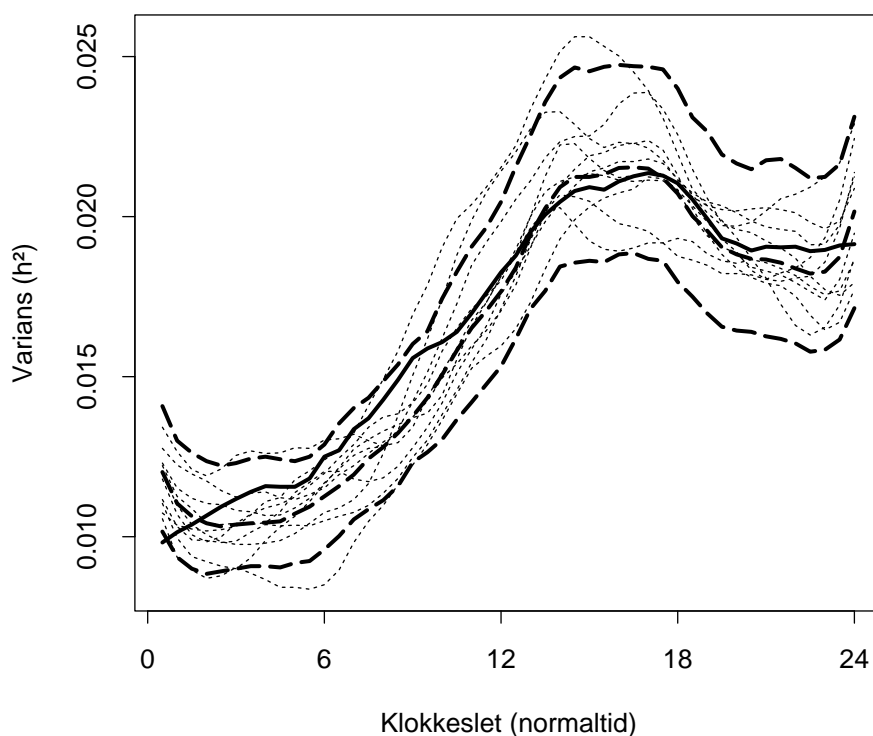


Figur 13: Oprindelige prædiktionsfejl (øverst) og 9 simulerede tidsrækker.

er forsøgt imødegået ved at anvende en naturlig kubisk spline i stedet for en kubisk spline basis. Der eksisterer følgende muligheder for at reducere problemet:

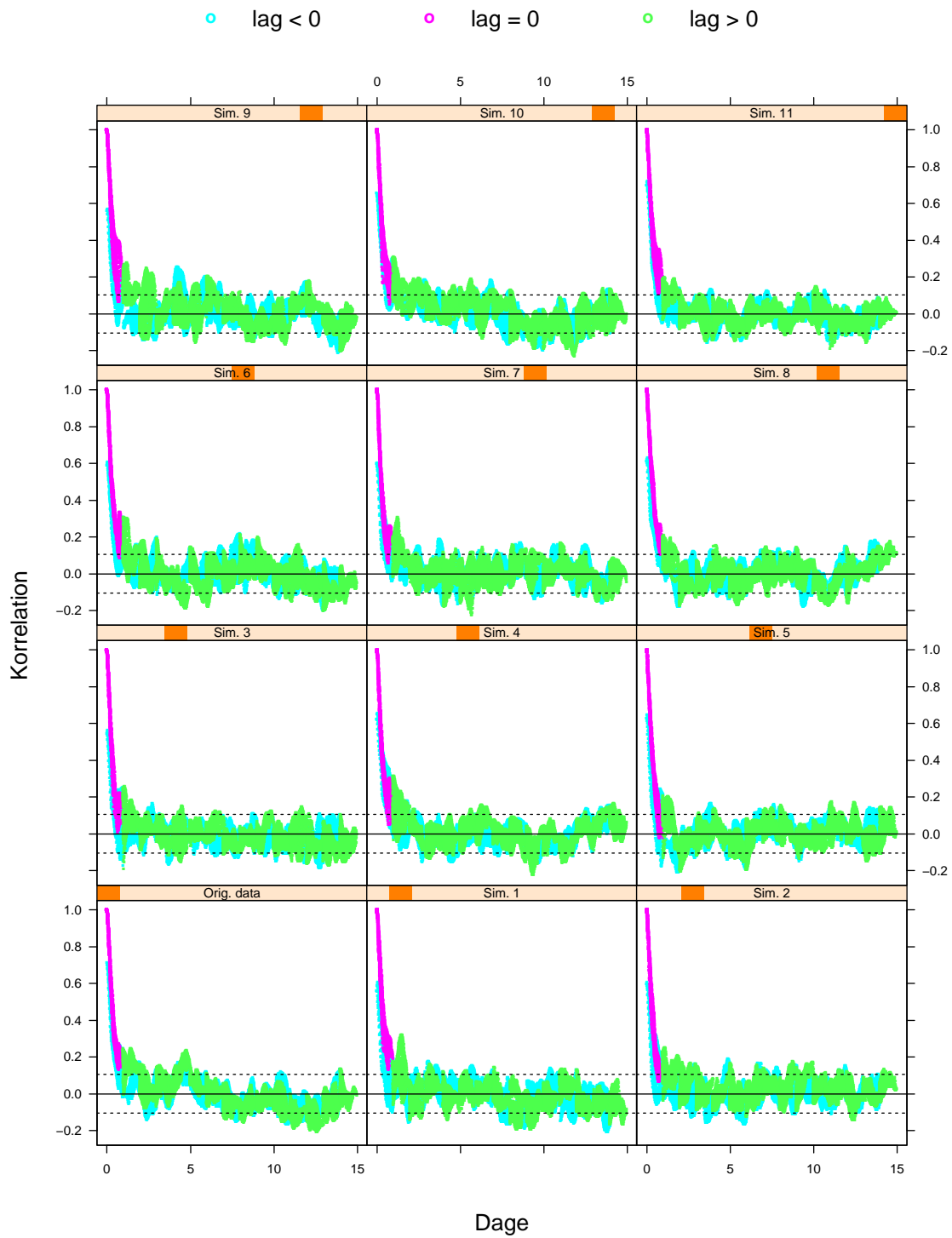
1. Anvend færre knuder i basen (dette medføre større fejl i modelleringen af døgnvariationen, se også afsnit 5.1).
2. Brug en spline af lavere orden (det viser sig, at stykkevist lineære funktioner har uheldige varians-egenskaber mellem knuderne, mens naturlige kvadratiske splines ikke umiddelbart kan genereres med tilgængelig software).

Da problemet jf. figur 14 er begrænset vil disse muligheder ikke blive undersøgt nærmere.

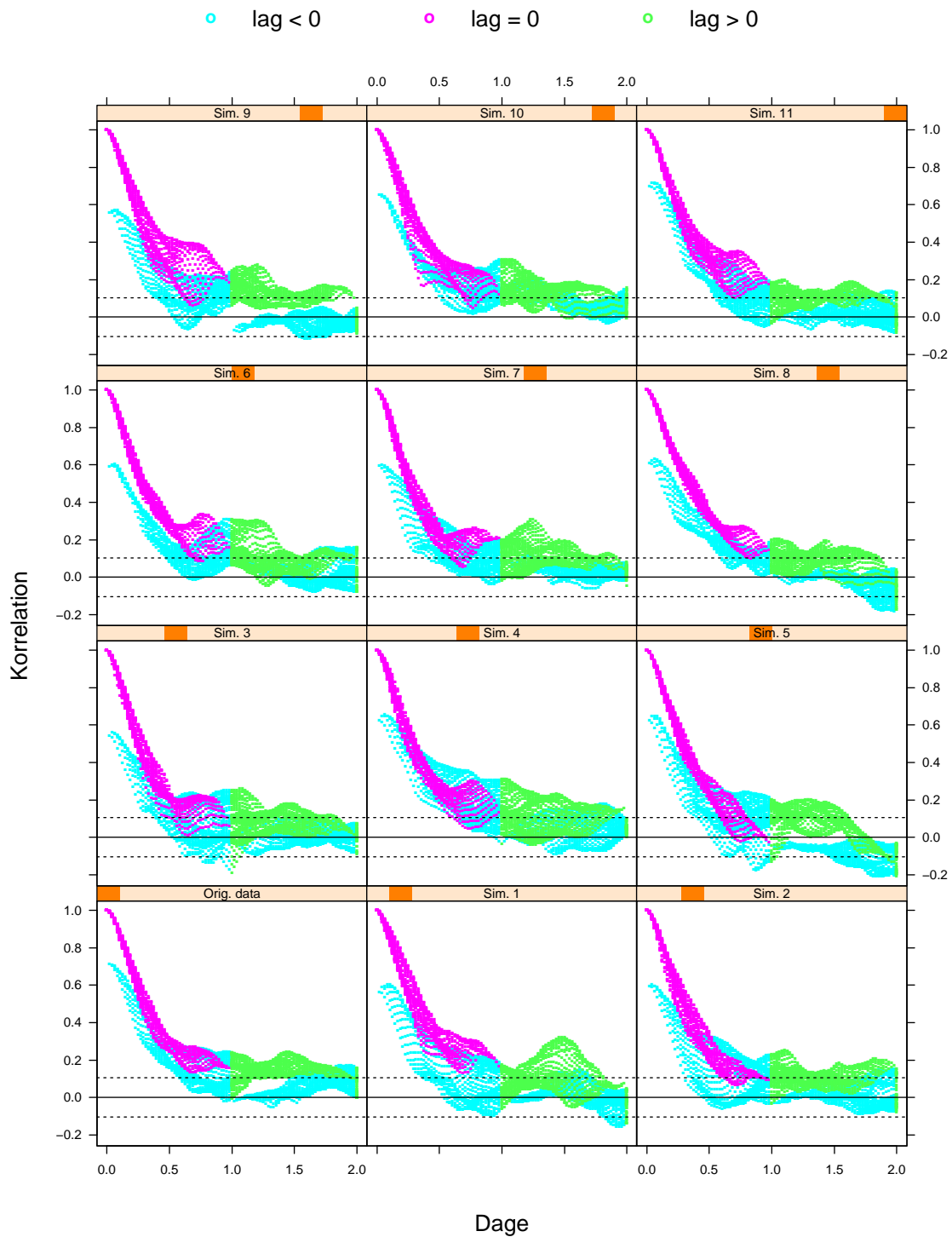


Figur 14: Den empiriske varians af de oprindelige data (fed, fuldtoptrukken), 5%, 50% og 95% fraktillerne fundet på basis af 200 simulerede tidsrækker (fed, stiptet), samt resultatet for 10 simuleringer (stiptet).

Den estimerede auto- og krydskorrelationsstruktur for de oprindelige data er vist på figur 2. Nedenfor undersøges det, i hvor stor udstrækning de simulerede tidsrækker udviser sammen struktur. Ialt 11 simuleringer er udført, og auto- og krydskorrelationer er estimeret som i afsnit 4.2, specielt figur 2 (side 8). Resultatet er vist på figur 15 og på figur 16 vises resultatet for tidsforskelle op til 2 dage. Overensstemmelsen synes tilfredsstillende, omend de simulerede værdier for negative lags og små tidsforskelle udviser nogen variation fra simulering til simulering.



Figur 15: Estimerede auto- og krydskorrelationer (SACF og SCCF) versus tidsforskellen mellem tidspunkter for hvilke prædiktionerne gælder. De stiplede linier svare til $\pm 2/\sqrt{N}$.



Figur 16: Estimerede auto- og krydskorrelationer (SACF og SCCF) versus tidsforskellen mellem tidspunkter for hvilke prædiktionerne gælder. De stiplede linier svare til $\pm 2/\sqrt{N}$.

6.2 Generel simulation

Som det fremgår af det foregående afsnit, er modellen i stand til at simulere tidsrækker, hvor de væsentlige aspekter af de oprindelige data er bevaret. Et af de væsentligste aspekter, er at de oprindelige data (normerede prædiktionsfejl) ikke har den samme varians igennem året og over døgnet, se figur 1, 11 og 29 (side 5, 19 og 41).

Disse observationer er dog ikke nødvendigvis et billede af de virkelige årsagssammenhænge. Det er plausibelt, at variansen afhænger af tiden på døgnet fordi denne er relateret til solopvarmningen, hvilket i øvrigt implikerer, at denne afhængighed ændres over året. Derimod er variansens ændringer hen over året formentlig bestemt af klimatiske forhold, der udviser en vis årsvariation, men som det ses af figur 1 (side 5), kan perioder med stærk blæst også forekomme om sommeren.

Bemærkning: En oplagt model for døgnafhængigheden er at benytte sinus til solhøjden som forklarende variabel idet denne størrelse vil være proportional med solindfaldet i klart vejr. Imidlertid skal denne størrelse lavpas-filtreres således, at maksimum forskydes fra ca. 12:00 til et tidspunkt mellem 15:00 og 17:00. Dog er det muligt, at den faktiske prædiktionshorisont også betyder noget for variansen af de normerede prædiktionsfejl. Her vil vi dog af tidsmæssige grunde ikke gå nærmere ind i denne analyse.

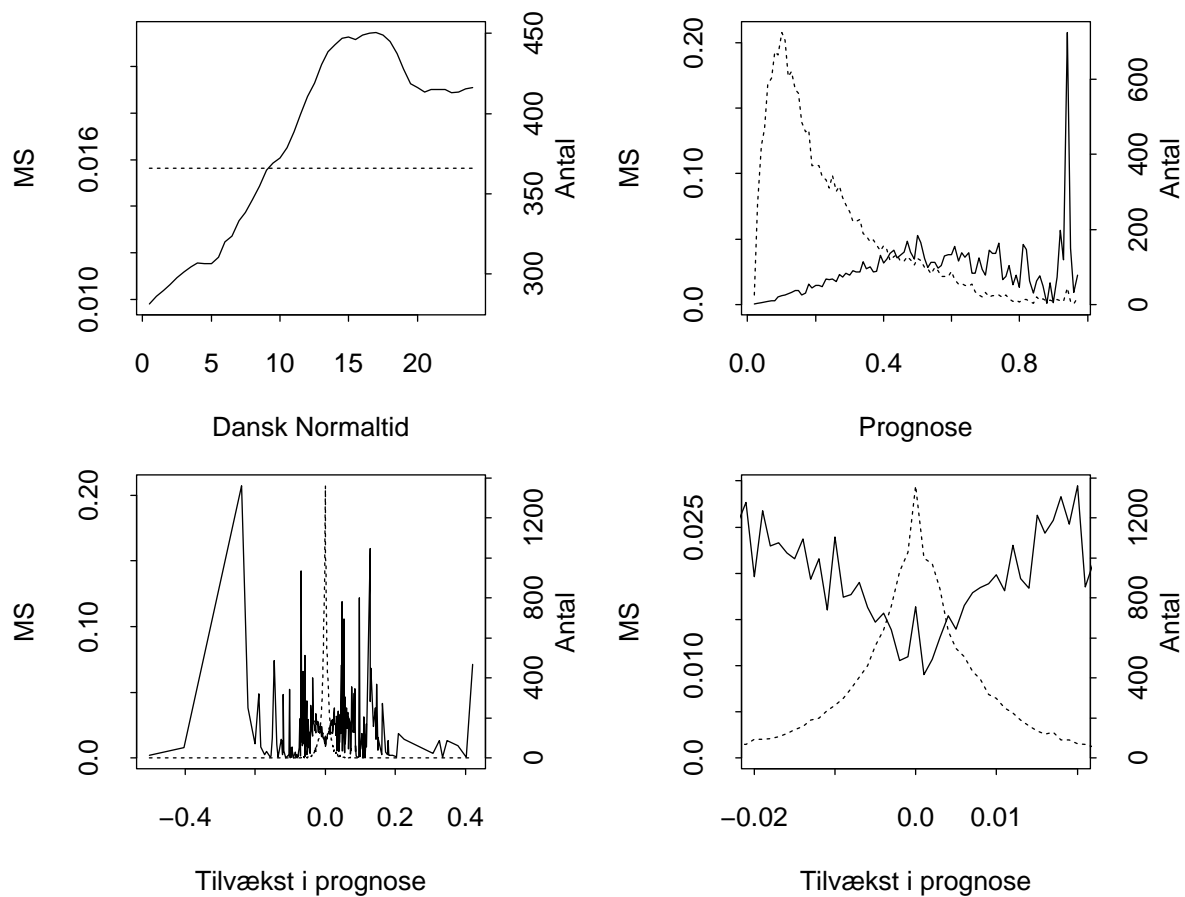
I dette afsnit vil det blive belyst, hvorledes variansen afhænger af den prognostiserede produktion og af tiden på døgnet. Bemærk at prognosen, og ikke den faktiske produktion, anvendes. Dette skyldes, at selvom prognosen skulle være perfekt således at fejlene er hvid støj, vil prognosefejlen og den faktiske produktion være afhængige.

Som det fremgår af afsnit 4.1, er middelværdiens bidrag til den forventede kvadratiske fejl ringe. Af den grund vil vi også i dette afsnit udglatte den kvadrerede normerede prædiktionsfejl. Med henblik på at analysere sammenhængene vil de forklarende variable

- den prognostiserede værdi,
- tilvæksten i den prognostiserede værdi over den forløbne time og
- klokkeslet (Dansk Normaltid som i foregående afsnit)

blive anvendt. Figur 17 viser den gennemsnitlige relative prædiktionsfejl som funktion af hver af de forklarende variable. Den høje top for høje værdier af prognosen viser sig at stamme fra den 29-30 januar, hvor prognosen er konstant høj fra 20:30 (29/1) til 13:30 (30/1)⁴, og hvor produktionen falder drastisk. Denne atypiske periode udelades af de resterende beregninger i dette afsnit.

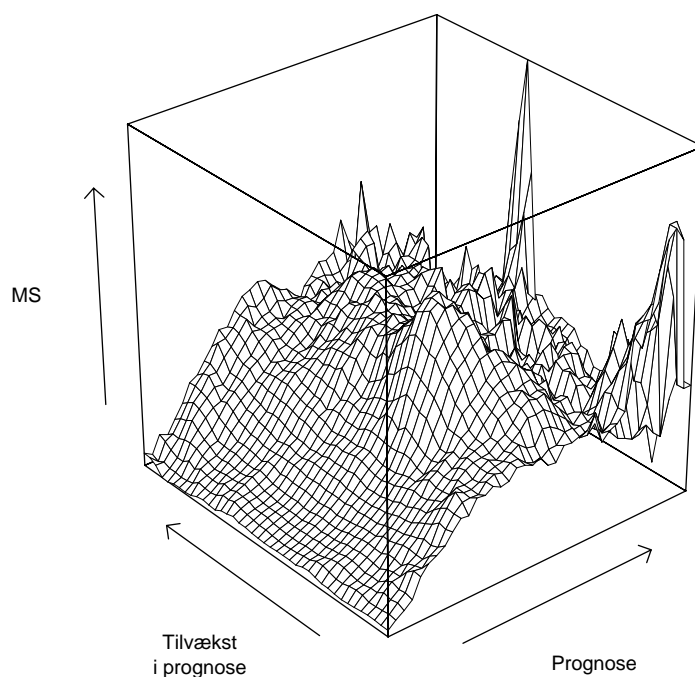
⁴Tidsangivelser i Dansk Standardtid (MET).



Figur 17: Den gennemsnitlige kvadrerede prædiktionsfejl (MS) fundet ved at beregne denne for hver af de uafhængige variable afrundet til hhv. 1, 2 og 3 decimaler. Plottet nederst til højre viser den centrale del af plottet nederst til venstre. Antallet af prædiktionsfejl er vist med stiplet og angivet på højre akse.

I det efterfølgende undersøges det, hvorledes den gennemsnitlige kvadrerede prædiktionsfejl afhænger af de nævnte variable. Undersøgelsen vil blive begrænset til ændringer i prognoser på under 0.02, idet antallet af situationer, hvor der forekommer større ændringer, er meget begrænset. Undersøgelsen udføres vha. lokal regression [1, 5, 6] og LFLM til S-PLUS [8] anvendes hertil. Den indbyggede S-PLUS funktion `loess` kunne også i stor udstrækning være anvendt, men LFLM er valgt, da den er mere fleksibel.

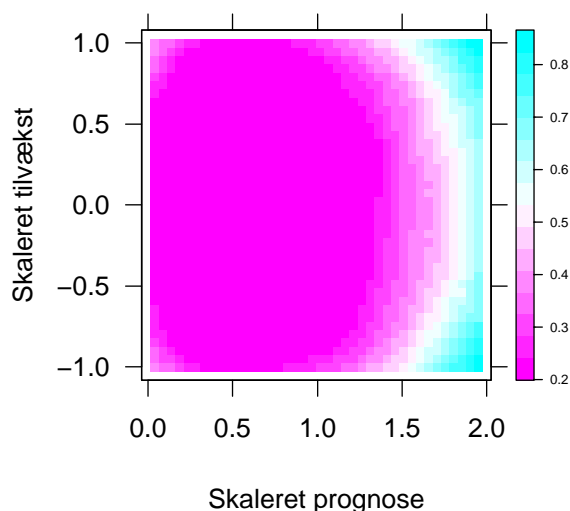
Figur 18 viser den gennemsnitlige kvadrerede prædiktionsfejl som funktion af prognosen og tilvæksten i denne. Denne sammenhæng er fundet ved kerneestimation (lokal konstant regression) med en fast båndbredde og med en skalering af de uafhængige variable, så $1/5$ af akserne ligger inden for båndbredden (hvis akserne er 1 lang er båndbredden 0.1). Endelig er en rektangulær (box) kerne anvendt. Fremgangsmåden svarer altså nøje til den ovenfor (figur 17) beskrevne. Det ses at afhængigheden af prognosen er langt mere markant end afhængigheden af tilvæksten.



Figur 18: Den gennemsnitlige kvadrerede prædiktionsfejl (MS) fundet med kerneestimation som beskrevet i teksten. MS ligger fra 0.001 til 0.087.

Ved høje prognostiserede værdier, hvor datagrundlaget er begrænset, er den estimerede flade meget ujævn. For at opnå en større udglætning ved høje prognostiserede værdier, anvendes lokal lineær regression med tricube vægtfunktion og en båndbredde defineret ud fra nærmeste nabo princippet. For at tage hensyn til de forskellige intervaller, som de forklarende variable variere indenfor, skales prognosen med 0.5 og tilvæksten med 0.02. Figur 19 viser den faktiske båndbredde, når en 1% nærmeste nabo båndbredde anvendes, og når den faktiske båndbredde ikke må blive under 0.2 (svarende til den faste båndbredde

ovenfor), 1% svarer til ca. 175 observationer. Det ses, at ved høje prognoser og ved store ændringer af disse er båndbredden over 0.8, hvilket svarer til næsten halvdelen af akserne.



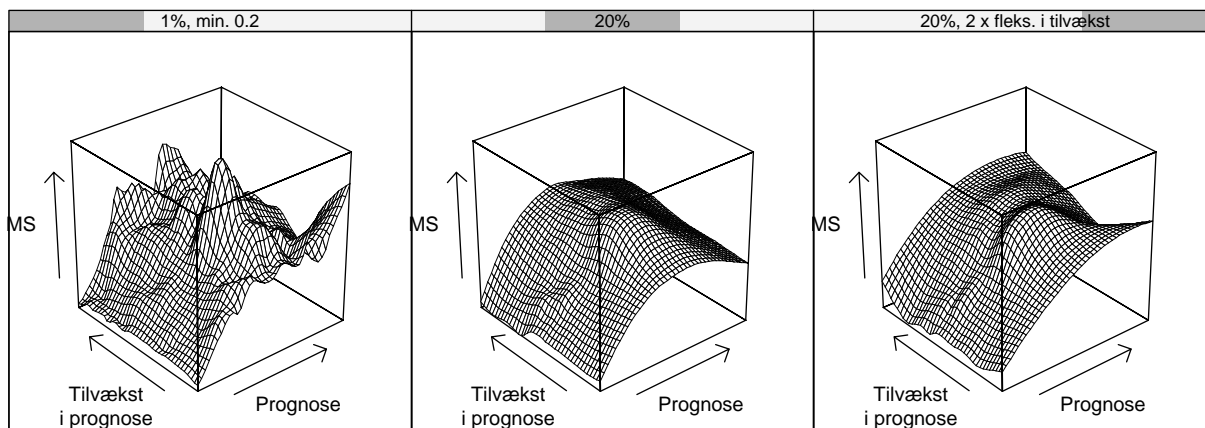
Figur 19: Faktisk båndbredde ved 1% nærmeste nabo og en minimum båndbredde på 0.2. Båndbredden er umiddelbart sammenlignelig med værdierne på akserne.

Figur 20 viser den estimerede flade sammen med estimeret opnået ved en 20% nærmeste nabo båndbredde, samt ved en 20% nærmeste nabo båndbredde når tilvæksten skaleres med 0.01 i stedet for 0.02. Sidstnævnte resulterer i mere fleksibilitet i retning af tilvæksten. I alle tilfælde ses en ringe effekt af tilvæksten i prognosen. Af de to variable synes prognosen at kunne forklare den alt overvejende del af variationen i den kvadrerede fejl.

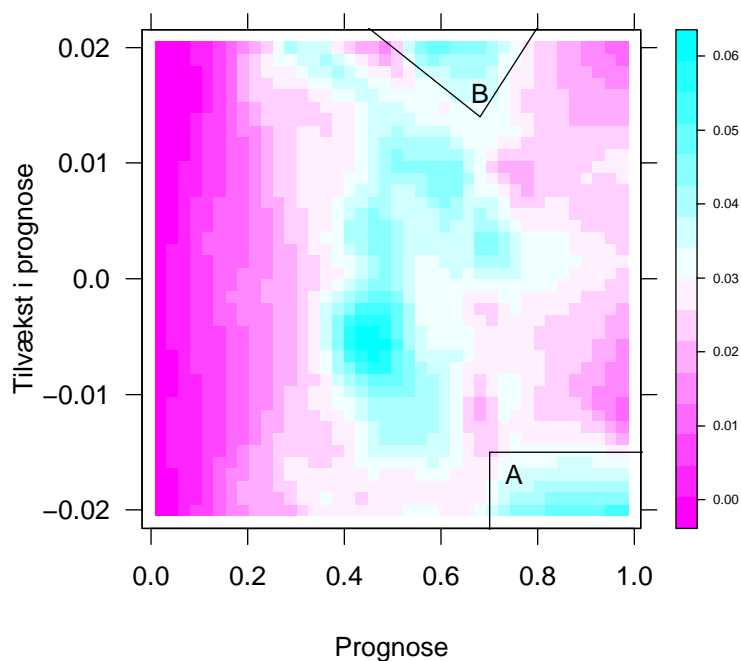
Den forventede kvadrerede fejl som funktion af prognosen synes at have et maksimum, der viser sig at forekomme ved en prognostiseret udnyttelsesgrad på omkring 0.5. Dette svarer til vindstyrker, hvor den producerede effekt er mest påvirkelig af ændringer i vindhastigheder. Ved de højeste prognostiserede udnyttelsesgrader ses en svag indikation af, at den forventede kvadrerede fejl stiger, hvis tilvæksten fjerner sig fra nul.

Figur 21 viser et niveauplot af det venstre plot i figur 20. På figur 21 er to områder „A“ og „B“ markeret. For område „A“, gælder at prognosen er meget høj, men at der forventes et fald. For område „B“ gælder at prognosen er forholdsvis høj, og at der forventes en yderligere stigning. I begge tilfælde er den forventede kvadrerede fejl høj. Begge disse forhold forekommer plausible, men vanskelige at modellere på simpel vis.

Da selve den prognostiserede værdi er langt mere betydende end tilvæksten i prognosen, vil tilvæksten i det efterfølgende blive erstattet af klokkeslettet (Dansk Normaltid) for prognosen. Det vil herefter blive undersøgt, i hvor stor udstrækning klokkeslettet kan forklare variationen i den kvadrerede fejl. Bemærk at klokkeslettet også er en indikator for længden af prognose-horisonten.

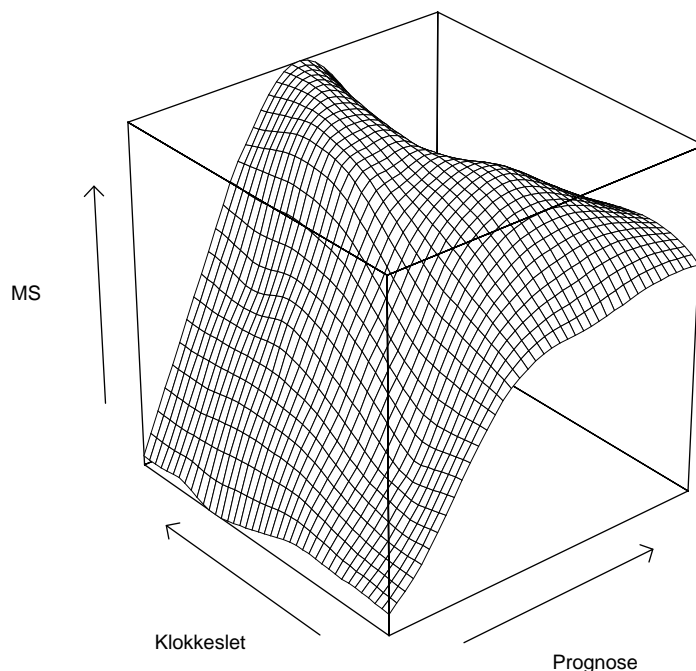


Figur 20: Den gennemsnitlige kvadrerede prædiktionsfejl (MS) fundet med lokal lineær regression som beskrevet i teksten. MS ligger fra -0.006 til 0.064.



Figur 21: Den gennemsnitlige kvadrerede prædiktionsfejl svarende til „1%, min. 0.2“ i figur 20.

Figur 22 viser den gennemsnitlige kvadrerede prædiktionsfejl (MS) som funktion af klokkeslet og prognose. Klokketslet synes at have en ringe betydning, men MS stiger dog svagt med klokkeslettet. Dette verificeres ved at variere båndbredden.



Figur 22: Den gennemsnitlige kvadrerede prædiktionsfejl (MS) som funktion af klokkeslet og prognose. Estimatet er fundet ved lokal lineær regression med en 20% nærmeste nabo båndbredde og med skalering klokkeslettet ved at dele med 24. MS ligger i intervallet -0.003 til 0.041.

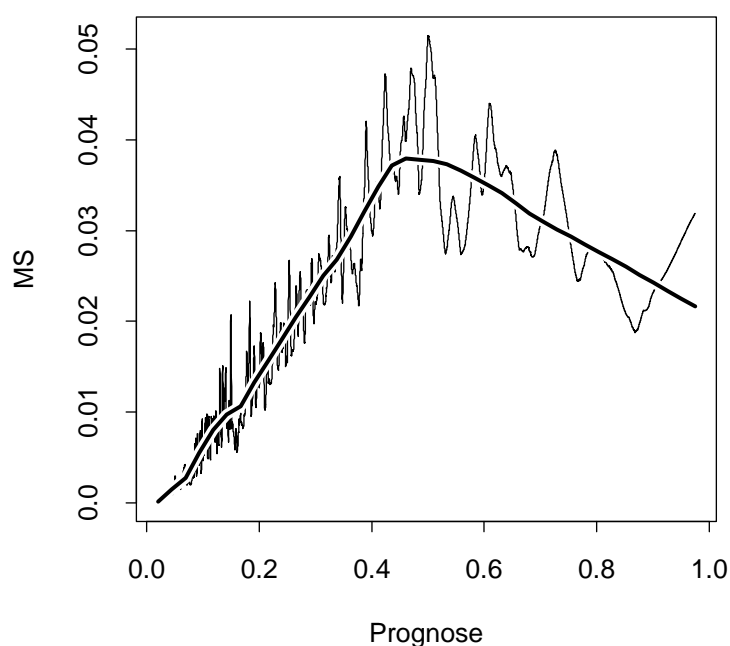
Det konkluderes, at af de undersøgte variable er det selve den prognostiserede værdi, der har størst forklarende værdi for den forventede kvadrerede fejl. Dog ses interessante aspekter vedr. ændringer i prognosen, når prognosen i forvejen er høj. I det følgende vil kun prognosen blive anvendt som forklarende variabel.

Figur 23 viser den estimerede sammenhæng for lokal lineær regression med 1% og 20% nærmeste nabo båndbredder og tricube vægt-funktion. Den meget lave båndbredde synes ikke at indikere, at der er nogen betydelig bias forbundet med estimatet opnået med en båndbredde på 20%. Sidstnævnte vil derfor blive brugt i det efterfølgende. På figur 24 ses estimatet og et 95% konfidensinterval fundet ved „bootstrap pairs“ [3, kap. 9]. Den underliggende antagelse for at anvende denne metode til at bestemme et konfidensinterval er, at de enkelte observationer er uafhængige stokastiske variable stammende fra samme (multivariate) fordeling. Dette er næppe korrekt, og de sande konfidensintervaller vil formentlig være bredere. Imidlertid giver konfidensintervallerne et udtryk for, hvor usikkert sammenhængen er bestemt for et år som år 2000.

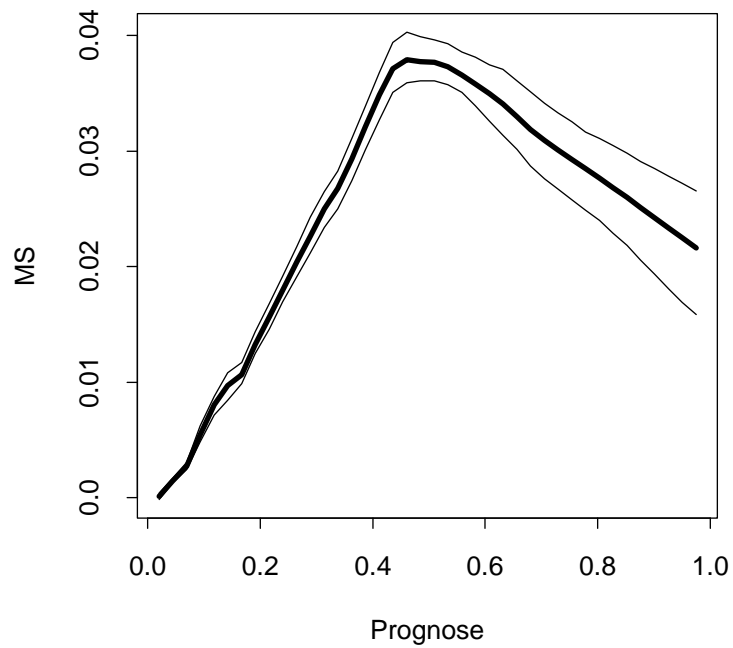
Enkelte hændelser kan resultere i stærkt afvigende resultater, specielt for høje værdier af den prognostiserede værdi. Som et eksempel viser figur 25 resultatet, når alle data fra år 2000 anvendes. Hvis man tager de store fejl den 29-30/1 som udtryk for at produktionen er faldet, fordi mange møller er standset pga. vindstyrken, indikerer resultaterne, at usikkerheden har et forløb som vist på figur 24, undtagen ved meget høje vindhastigheder, hvor store fejl kan opstå, fordi møllerne begynder at standse.

Figur 26 viser den gennemsnitlig forventede kvadrerede fejl for hver halve time i døgnet, når sammenhængen fra figur 24 anvendes for de faktiske prognostiserede værdier i år 2000. Det ses, at døgnvariationen i nogen udstrækning kan forklares ved døgnvariationer i den prognostiserede værdi.

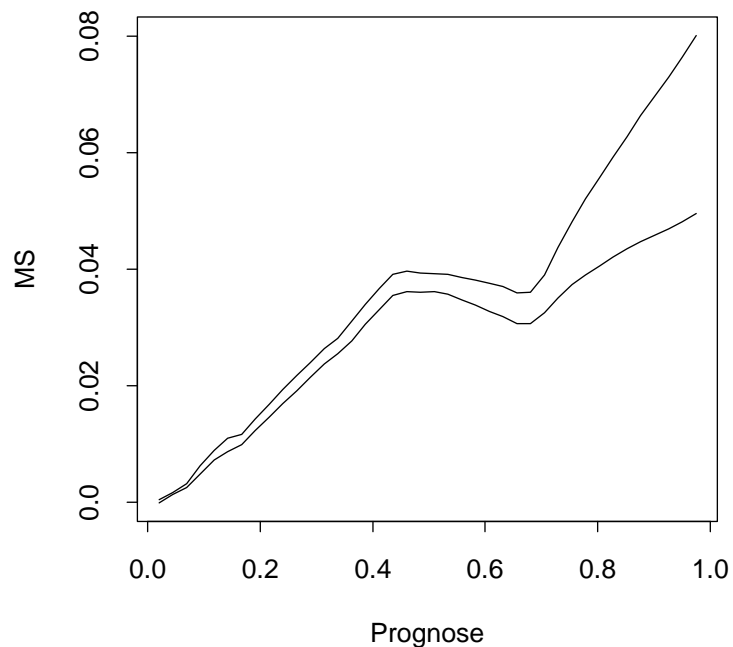
På figur 13 (side 23) ses den oprindelige tidsrække, samt 9 simulerede tidsrækker opnået bla. ved at beskrive usikkerhedens tidsmæssige afhængighed som vist i figur 11 (side 19). Tilsvarende plot er vist i figur 27 (side 36), men her er usikkerheden beskrevet ved den prognostiserede værdi via sammenhængen vist på figur 24 (side 34) (kvadratroden af værdien på kurven anvendes). Usikkerheden bestemt via prognosen synes at være velegnet i forbindelse med simulering.



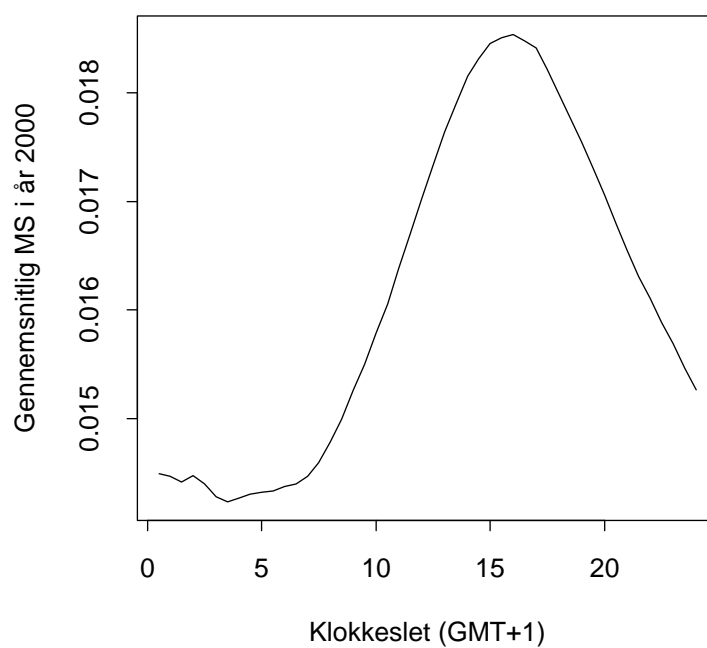
Figur 23: Den gennemsnitlige kvadrerede prædiktionsfejl fundet med lokal lineær regression med 20% (fed) og med 1% nærmeste nabo båndbredde.



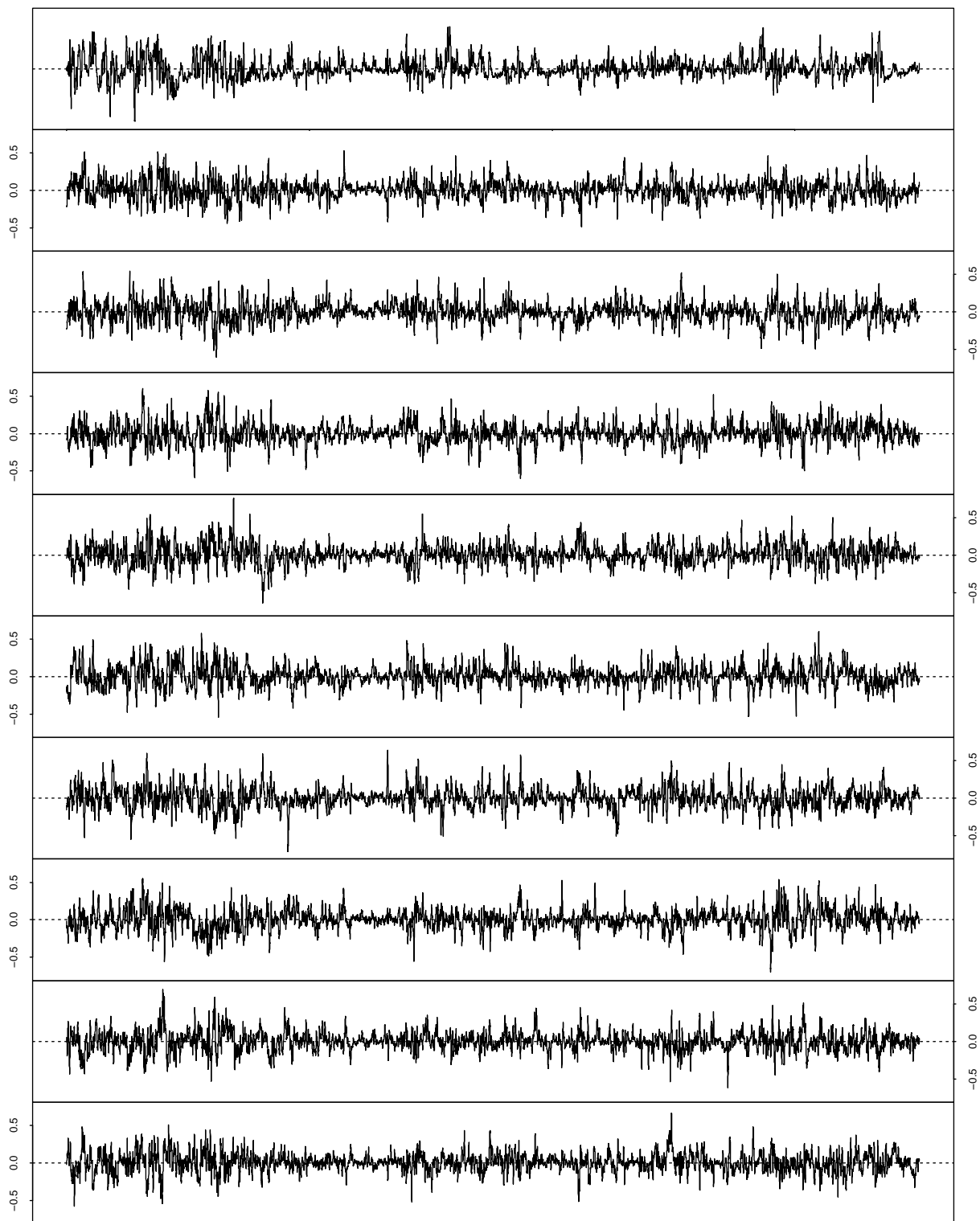
Figur 24: Den gennemsnitlige kvadrerede prædiktionsfejl fundet med lokal lineær regression med 20% (fed) nærmeste nabo båndbredde samt punktvis 2.5% og 97.5% bootstrap-fraktiler baseret på 200 gentagelser.



Figur 25: Punktvis 2.5% og 97.5% bootstrap-fraktiler baseret på 200 gentagelser når lokal lineær regression med 20% nærmeste nabo båndbredde anvendes for alle data.



Figur 26: Gennemsnitlig forventet kvadreret fejl når sammenhængen fra figur 24 anvendes for de faktiske prognostiserede værdier i år 2000.



Figur 27: Oprindelige prædiktionsfejl (øverst) og 9 simulerede tidsrækker som i figur 13, dog med variansen beskrevet ved sammenhængen i figur 24.

7 Konklusion

I denne rapport analyseres prædiktionsfejl for den samlede elproduktion på vindmøller i det Jysk-Fynske område. Prædiktionsfejlene er i denne sammenhæng indmeldinger til Nord-Pool foretaget af Eltra klokken 12:00 og gældende for det efterfølgende døgn (her aggregeret til intervaller af 30 minutter). Disse prædiktionsfejle er i stor udstrækning baseret på on-line systemet WPPT (Wind Power Prediction Tool) [11, 10, 9], men Eltra kan afvige fra disse, såfremt det skønnes relevant. Prædiktionsfejlene er normaliseret med den installerede effekt, således at de udtrykker prædiktionsfejl for udnyttelsesgraden. Data dækker år 2000.

Formålet med analysen af disse prædiktionsfejle er, ud fra en hypotetisk tidsrække for prædiktionsfejlene (normaliserede), at kunne simulere tidsrækker for fejlene. Dette vil blive anvendt i en to-trins system-simulering i SIVAEL [4, 12].

Analysen viser, ikke overraskende, at fejlene ikke har sammenlignelig varians. Variansen afhænger af årstiden og tiden på døgnet, men det viser sig, at denne ændring i stor udstrækning kan forklares ved selve den prognostiserede værdi. Det viser sig at variansen, som funktion af den prognostiserede værdi, har maksimum omkring en udnyttelsesgrad på 0.5. Denne værdi svarer til vindstyrker, hvor vindmøllernes „power curve“ har maksimal hældning. Det skal i denne sammenhæng bemærkes, at variansen stiger voldsomt i de sjældne situationer, hvor vindstyrken bliver så høj, at vindmøllerne kobles ud.

Prædiktionsfejlene mellem to på hinanden følgende døgn er korrelerede. Nærmere analyse viser, at denne auto- og krydskorrelationsstruktur i stor udstrækning kan forklares ved tidsforskellen mellem de tidspunkter, for hvilke prædiktionsfejlene gælder. Dog vil det for lave tidsforskelle (under ca. 5 timer) være vigtigt at tage hensyn til forskelle i de tidspunkter, hvor prædiktionsfejlene er genereret. Denne forskel kaldes i rapporten „lag“. Betragt eksempelvis følgende:

$e_{1,24:00}$: fejlen for prognosen genereret dag 1 og gældende for kl. 24:00 det efterfølgende døgn,

$e_{2,00:30}$: fejlen for prognosen genereret dag 2 og gældende for kl. 00:30 det efterfølgende døgn,

$e_{2,01:00}$: fejlen for prognosen genereret dag 2 og gældende for kl. 01:00 det efterfølgende døgn,

hvor der ses at være 30 minutter mellem prognoserne svarende til $e_{1,24:00}$ og $e_{2,00:30}$, samt mellem $e_{2,00:30}$ og $e_{2,01:00}$. Her estimeres korrelationen mellem $e_{1,24:00}$ og $e_{2,00:30}$ til 0.71 og korrelationen mellem $e_{2,00:30}$ og $e_{2,01:00}$ til 0.99. Se også figur 2 på side 8.

Den fundne auto- og krydskorrelationsstruktur burde udnyttes til med meget få parametre

at bygge en multivariat tidsrække model for de 48 parallelle tidsrækker (48 prædiktionsfejl pr. indmelding til NordPool). Imidlertid vil en sådan model ikke kunne håndteres med standard software, og arbejdsmængden vil derfor langt overstige, hvad der ligger inden for rammerne af dette projekt. I stedet udnyttes blot, at auto- og krydskorrelationsstrukturen kan modelleres med en multivariat moving average model af orden 1 (MMA(1)). I det konkrete tilfælde har en sådan model $48 + 48^2 = 2352$ parametre og kan derfor ikke håndteres direkte.

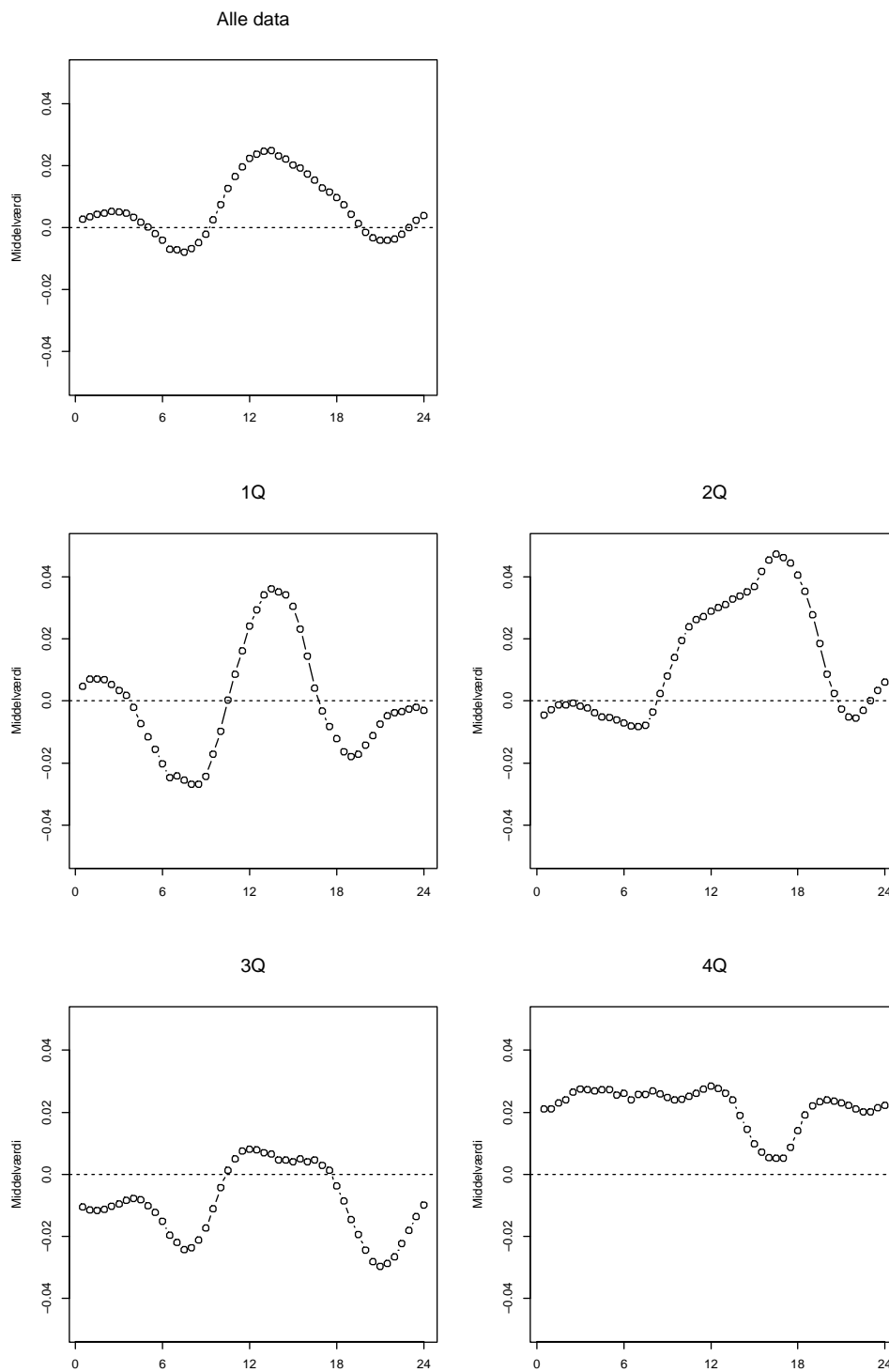
For at reducere antallet af parametre er en naturlig spline basis med seks interne knuder, placeret ækvidistant over døgnet, blevet fittet til de 48 fejl for hvert døgn. Dette er rimeligt, idet fejlene inden for et døgn er positivt korreleret. Herefter er en MMA(1) model blevet anvendt til at beskrive den tidsmæssige udvikling i de 8 spline koefficienter. Det bemærkes, at de oprindelige data blev standardiseret ved at dele med et estimat af standardafvigelsen, der tager hensyn til ændringer i denne over døgnet og over året.

Ved hjælp af de fundne modeller kan tidsrækker af prædiktionsfejl simuleres som anført i afsnit 6.1. Dog kan det være relevant at multiplicere med en standardafvigelse bestemt ved prognosen som anført i afsnit 6.2. Alle koefficienter, der er nødvendige for at kunne udføre denne simulation er gemt i regnearksfiler i CVS-format. Filnavne er anført i appendiks B (side 52). Det er i rapporten vist at de simulerede tidsrækker har egenskaber der svare til den oprindelige tidsrækkes egenskaber. Dog medfører spline-modelleringen en svagt øget varians ved lave og høje prognose-horisonter. Denne varians vil kunne mindskes ved at mindske antallet af knuder i spline basen, men herved vil der kunne optræde bias ved andre prognose-horisonter. Endelig bemærkes, at denne simulering ikke omfatter, at vindmøllerne kan standse automatisk ved meget høje vindstyrker. Noget sådant optræder meget sjældent og vil kun kunne tilføjes på ad hoc basis.

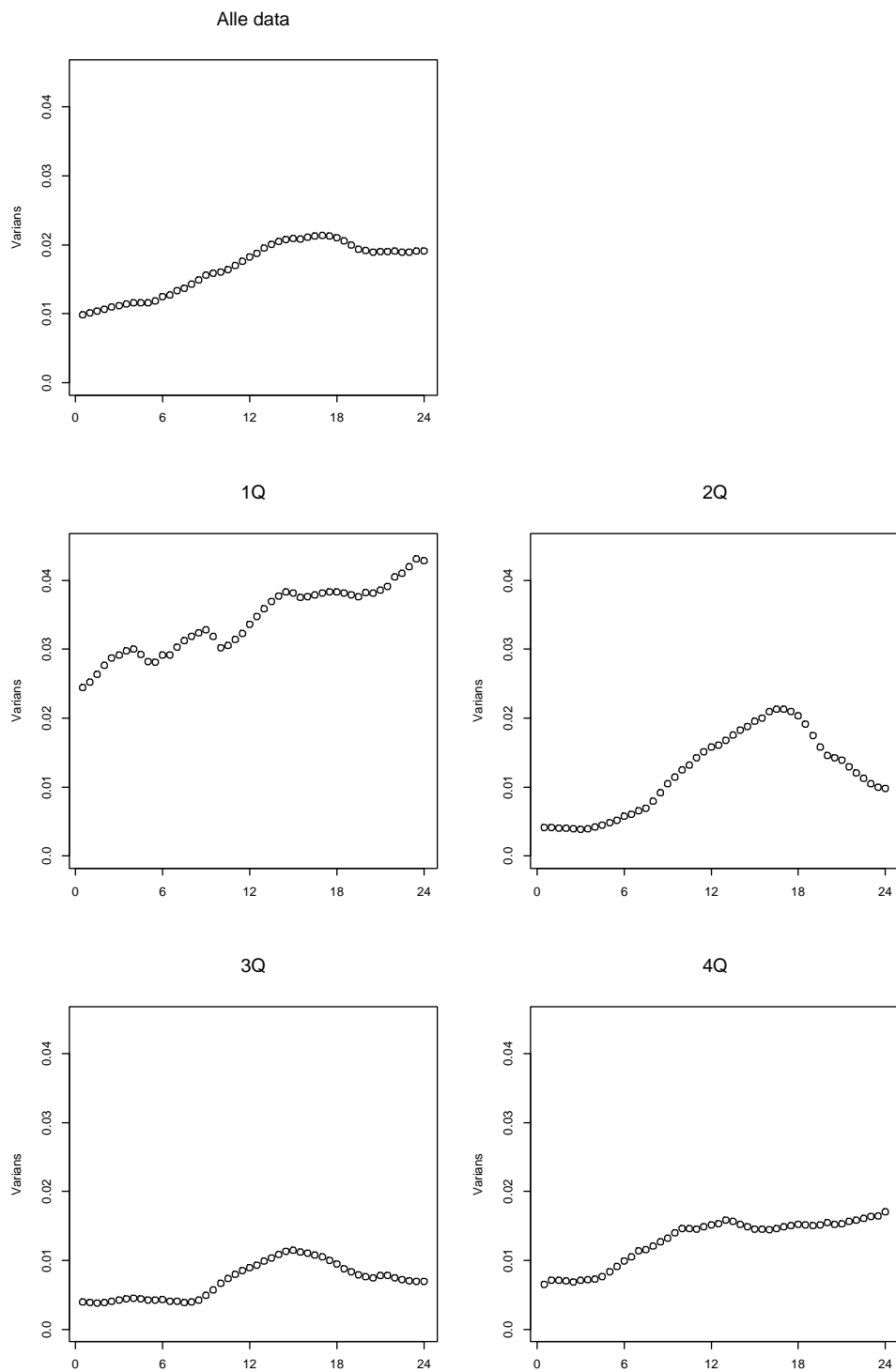
Endelig bemærkes, at der foreligger rige muligheder for at arbejde videre med mange aspekter berørt i denne rapport, heriblandt:

- Opbygning af og estimation i multivariate tidsrækkemodeller af høj dimension, der med få parametre kan modellere auto- og krydskorrelationsstrukturer som beskrevet ovenfor.
- Direkte modellering af variansinhomogenitet vha. multivariate GARCH-modeller, hvor den betingede variansen styres af prognosen.
- Analyse af situationer med markante frontpassager, for at analysere hvorledes usikkerheden ved prognosen påvirkes af fasefejl.
- Undersøgelse af i hvilket omfang den fundne auto- og krydskorrelationsstruktur kan udnyttes i en egentlig stokastisk optimering i SIVAEL.

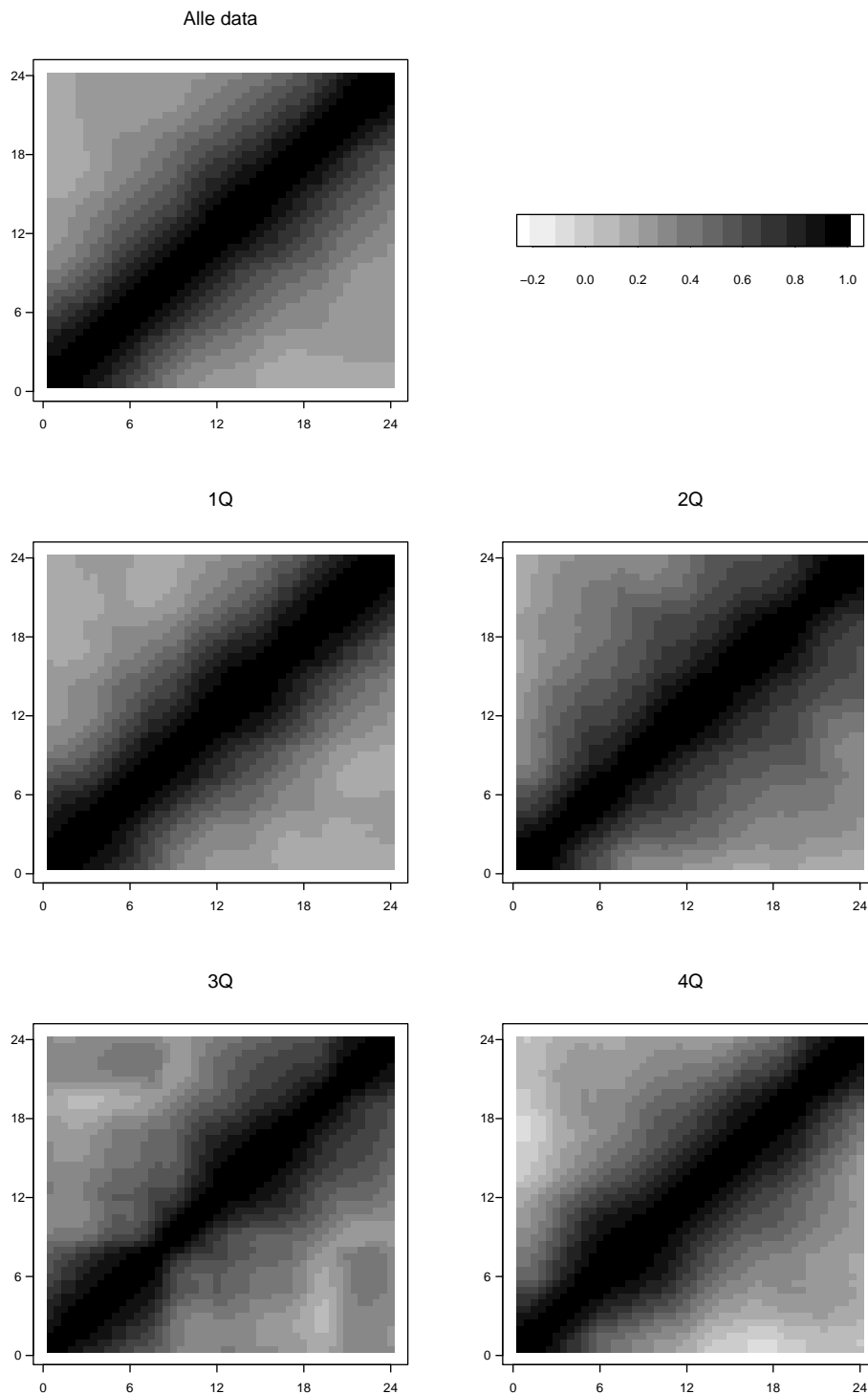
A Plots



Figur 28: Gennemsnitlig fejl mod klokkeslet for alle data og for hvert kvartal.

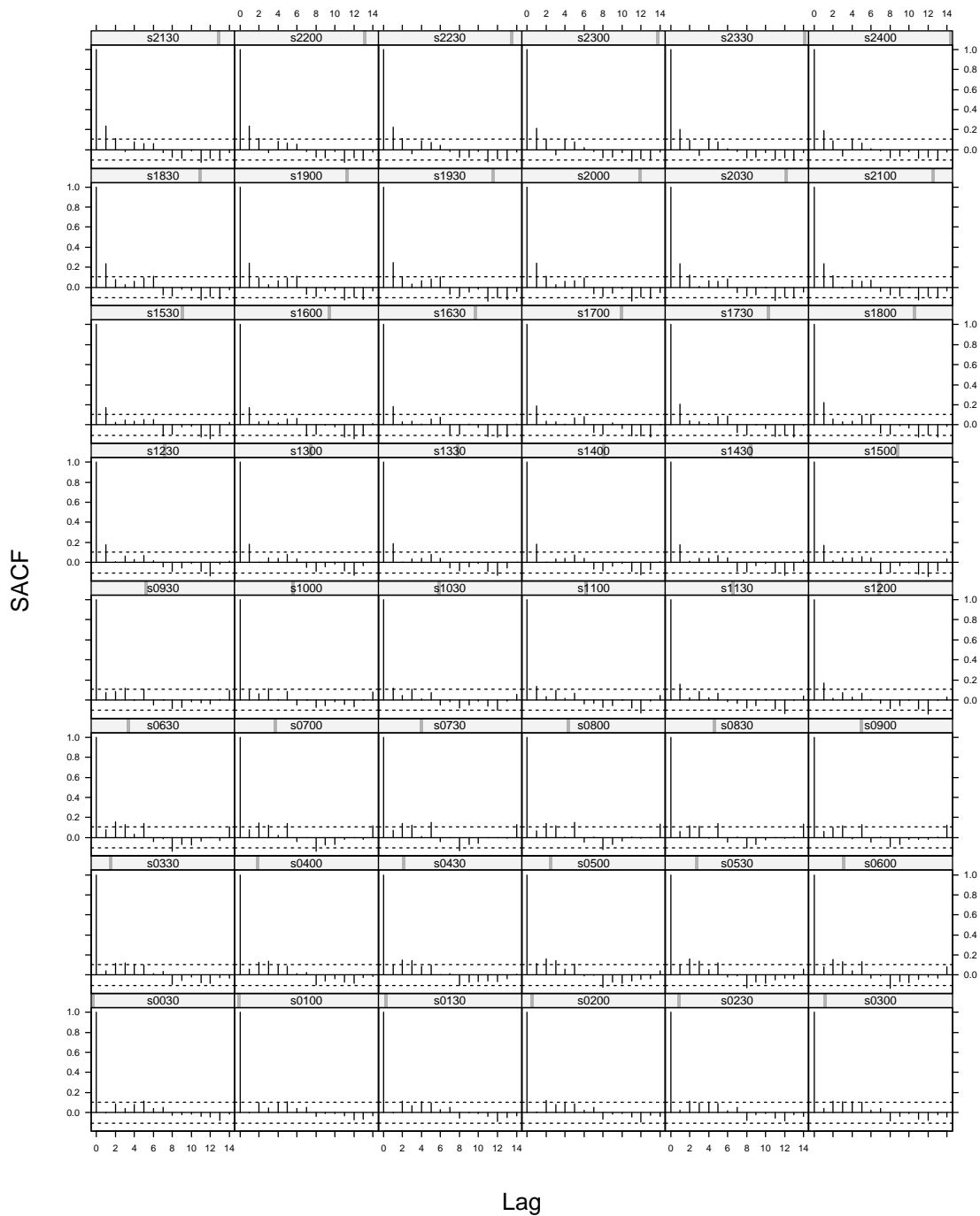


Figur 29: Estimeret varians af fejl mod klokkeslet for alle data og for hvert kvartal.



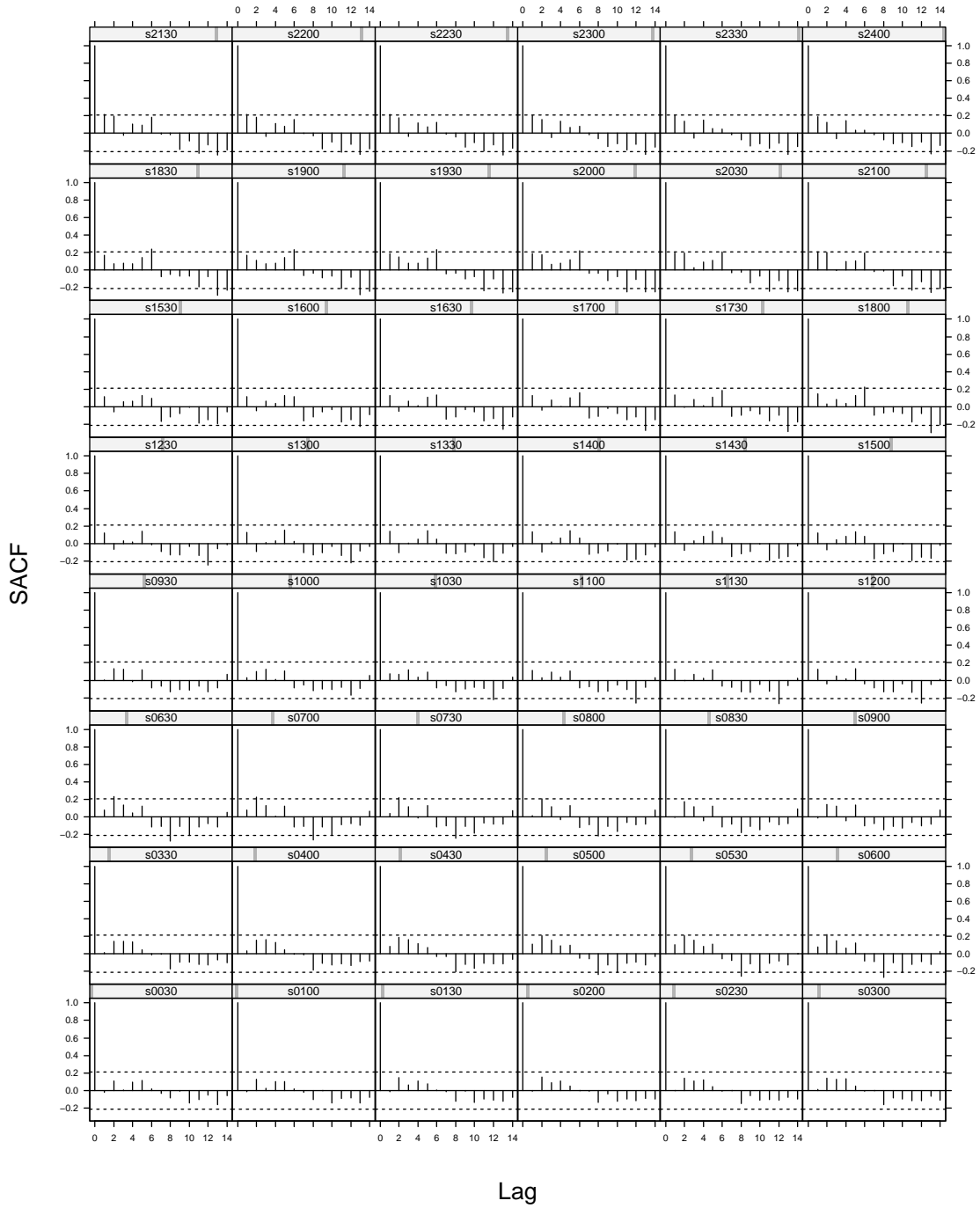
Figur 30: Estimeret korrelation af fejl til de enkelte klokkeslet for alle data og for hvert kvartal.

Alle data



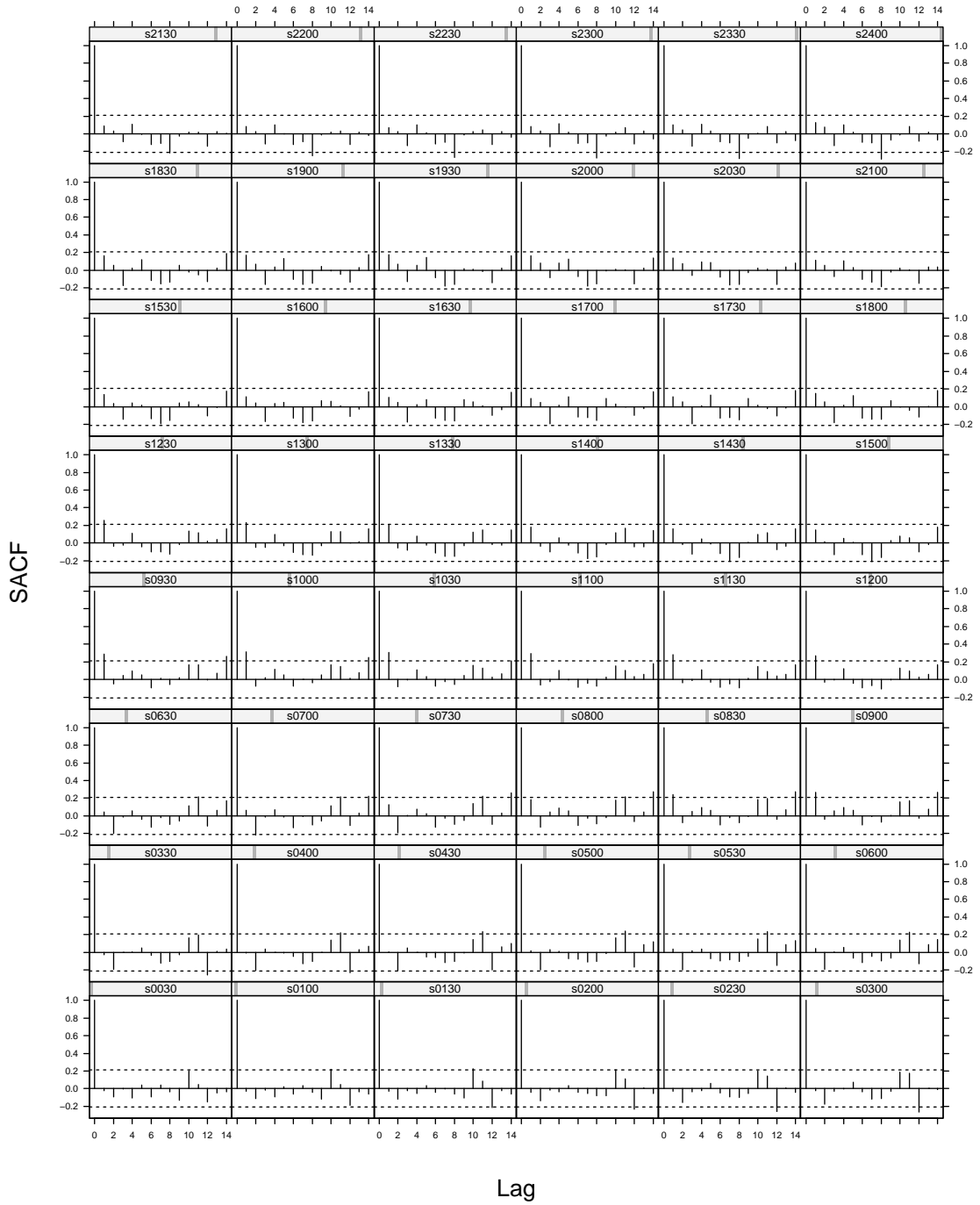
Figur 31: Estimeret autokorrelationsfunktion for lag 0 til 14.

1Q



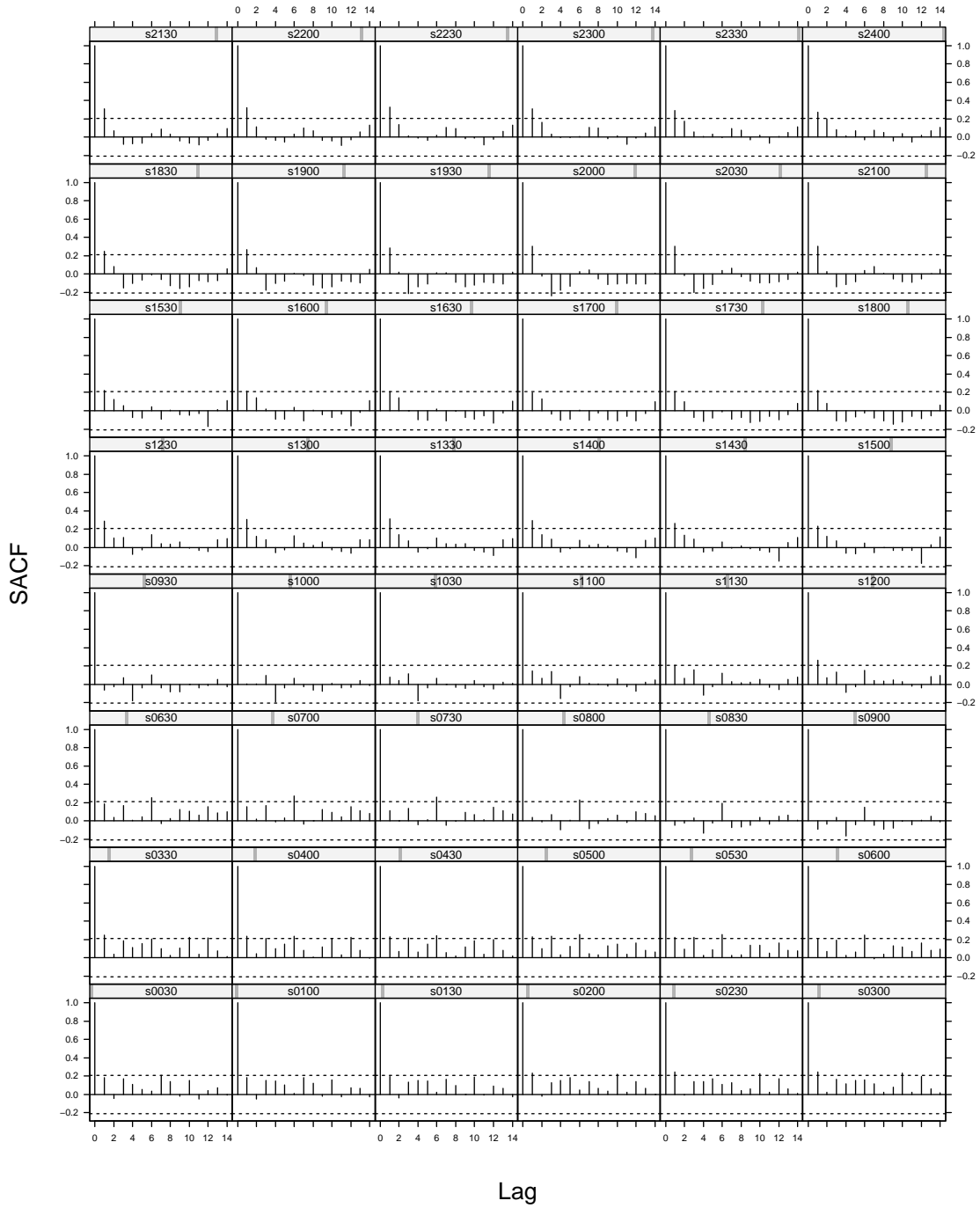
Figur 32: Estimeret autokorrelationsfunktion for lag 0 til 14; 1. kvartal

2Q



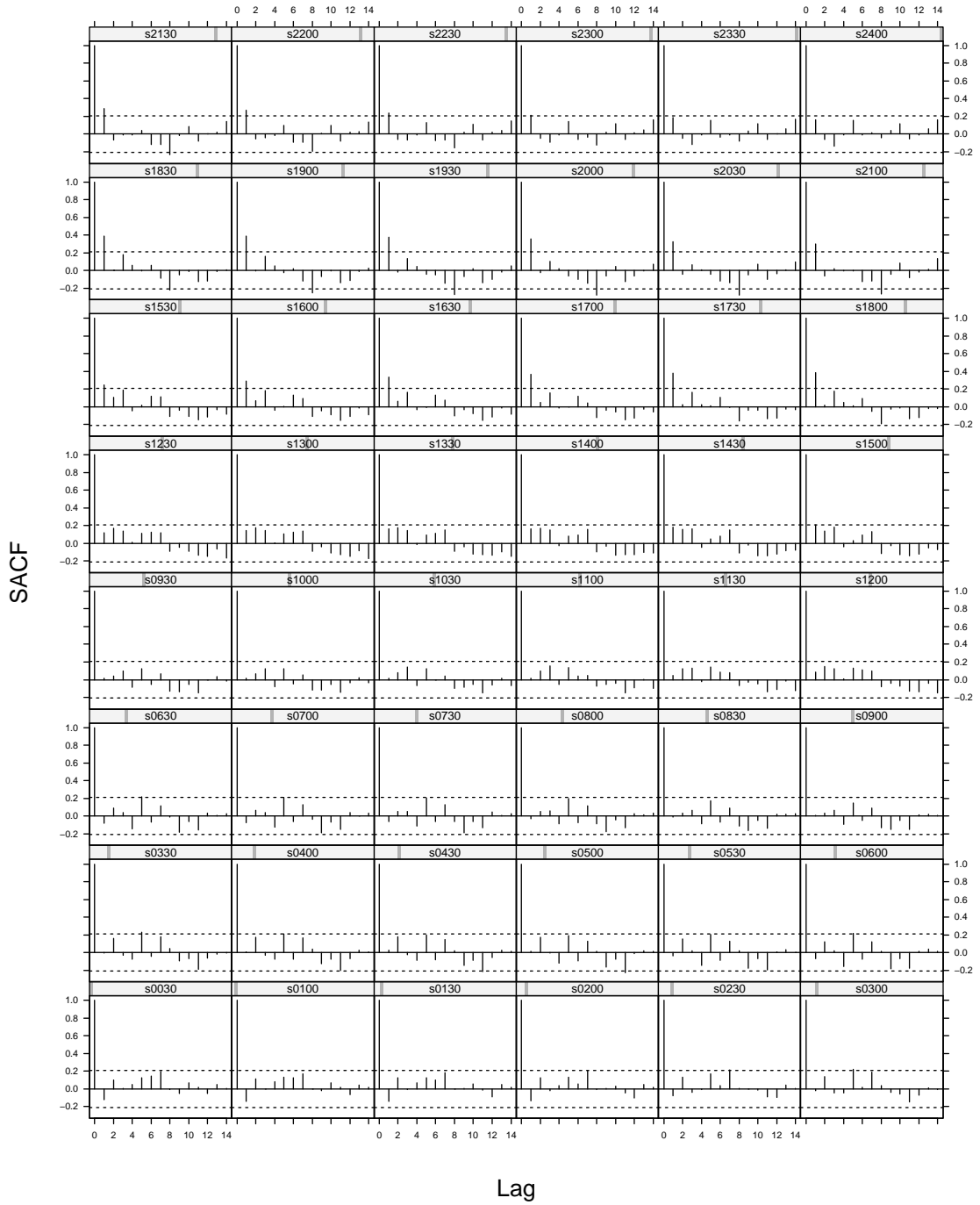
Figur 33: Estimeret autokorrelationsfunktion for lag 0 til 14; 2. kvartal

3Q

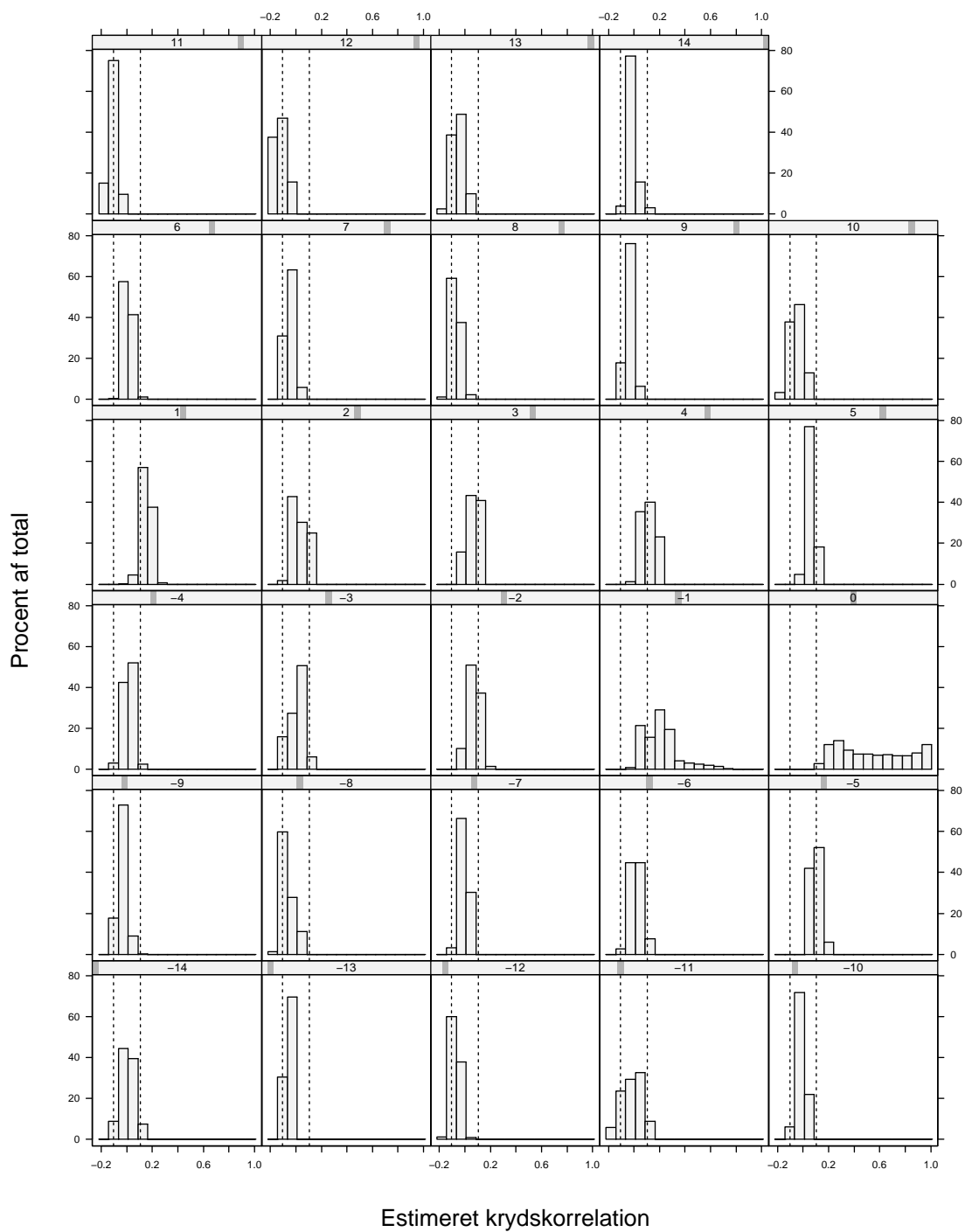


Figur 34: Estimeret autokorrelationsfunktion for lag 0 til 14; 3. kvartal

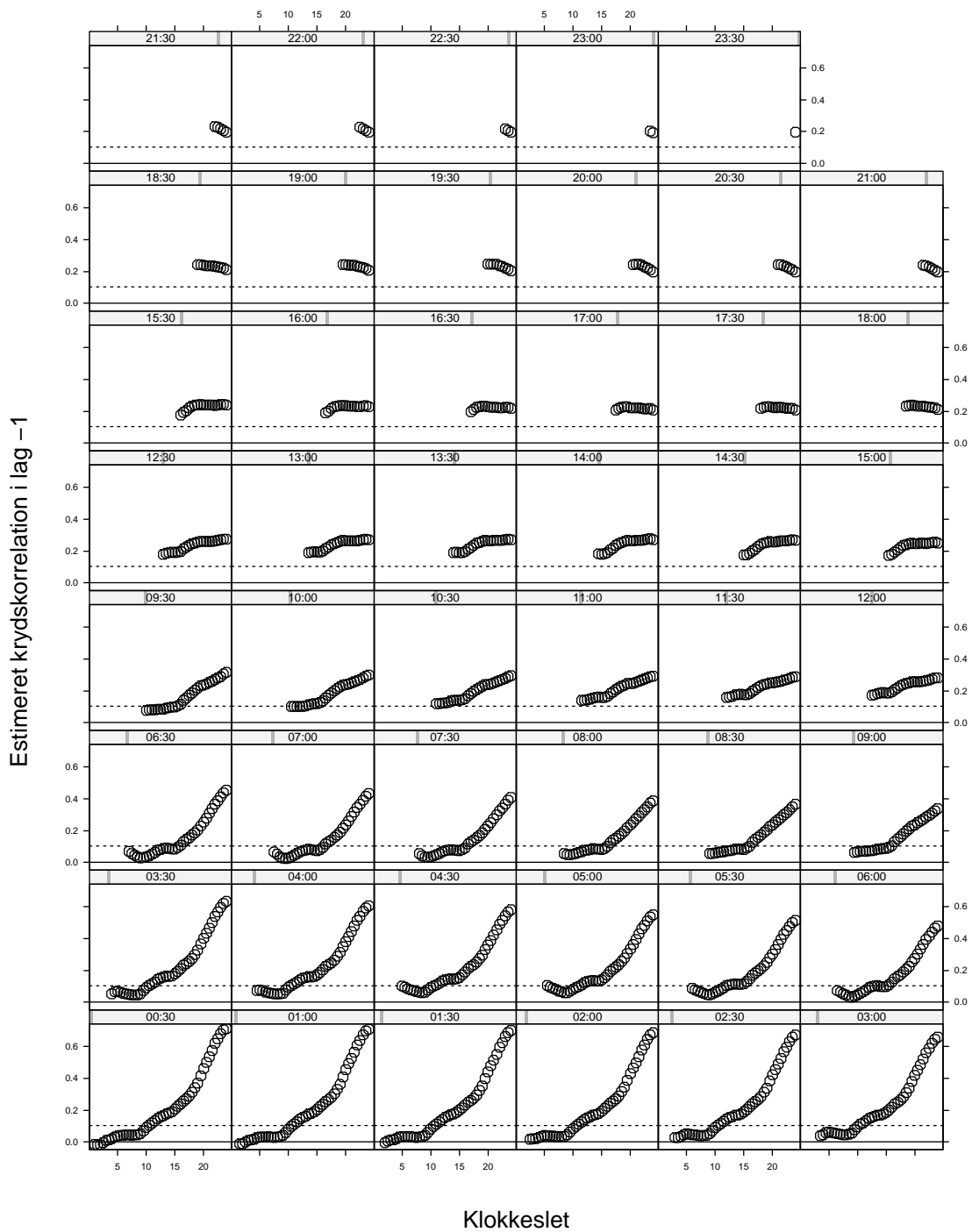
4Q



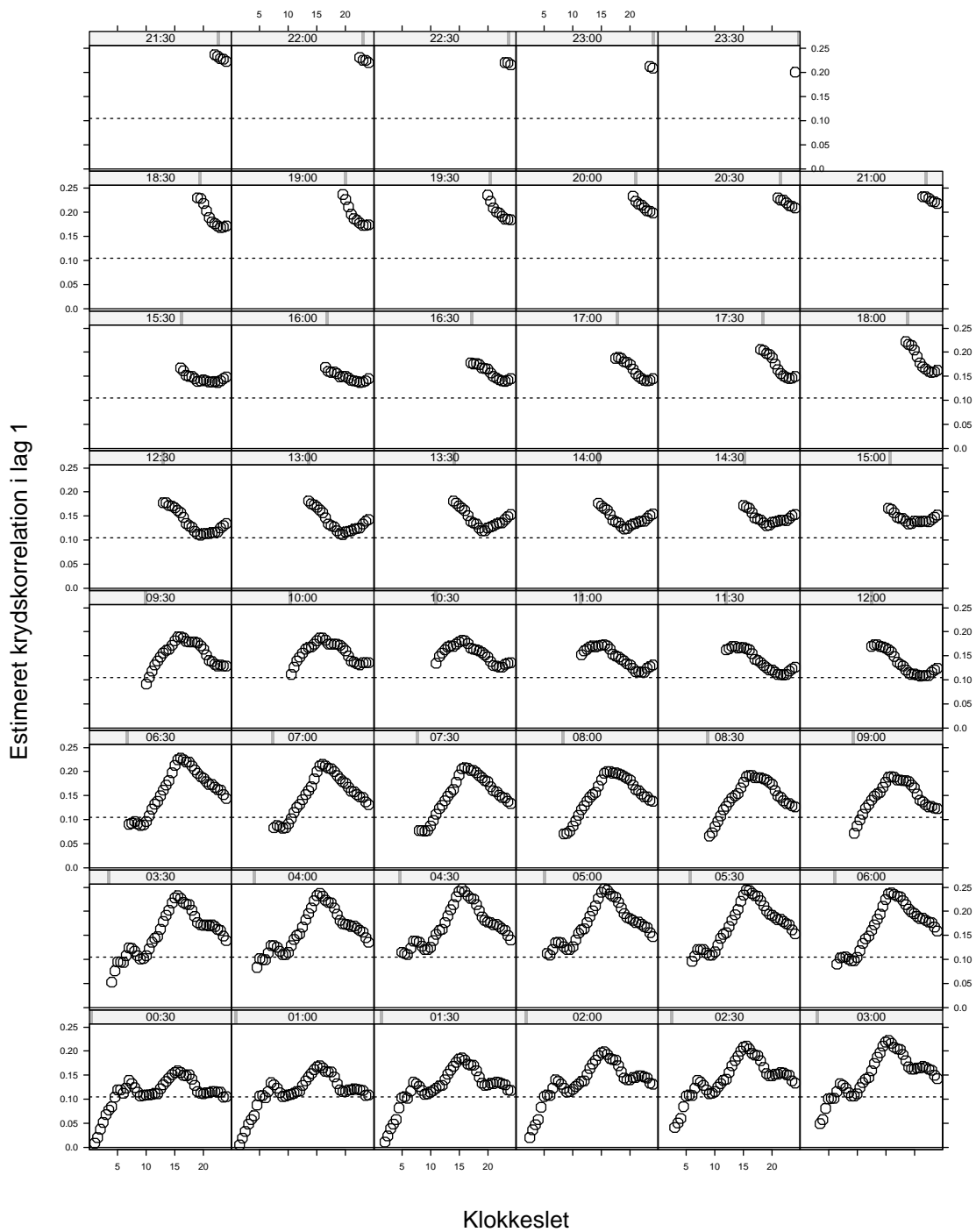
Figur 35: Estimeret autokorrelationsfunktion for lag 0 til 14; 4. kvartal



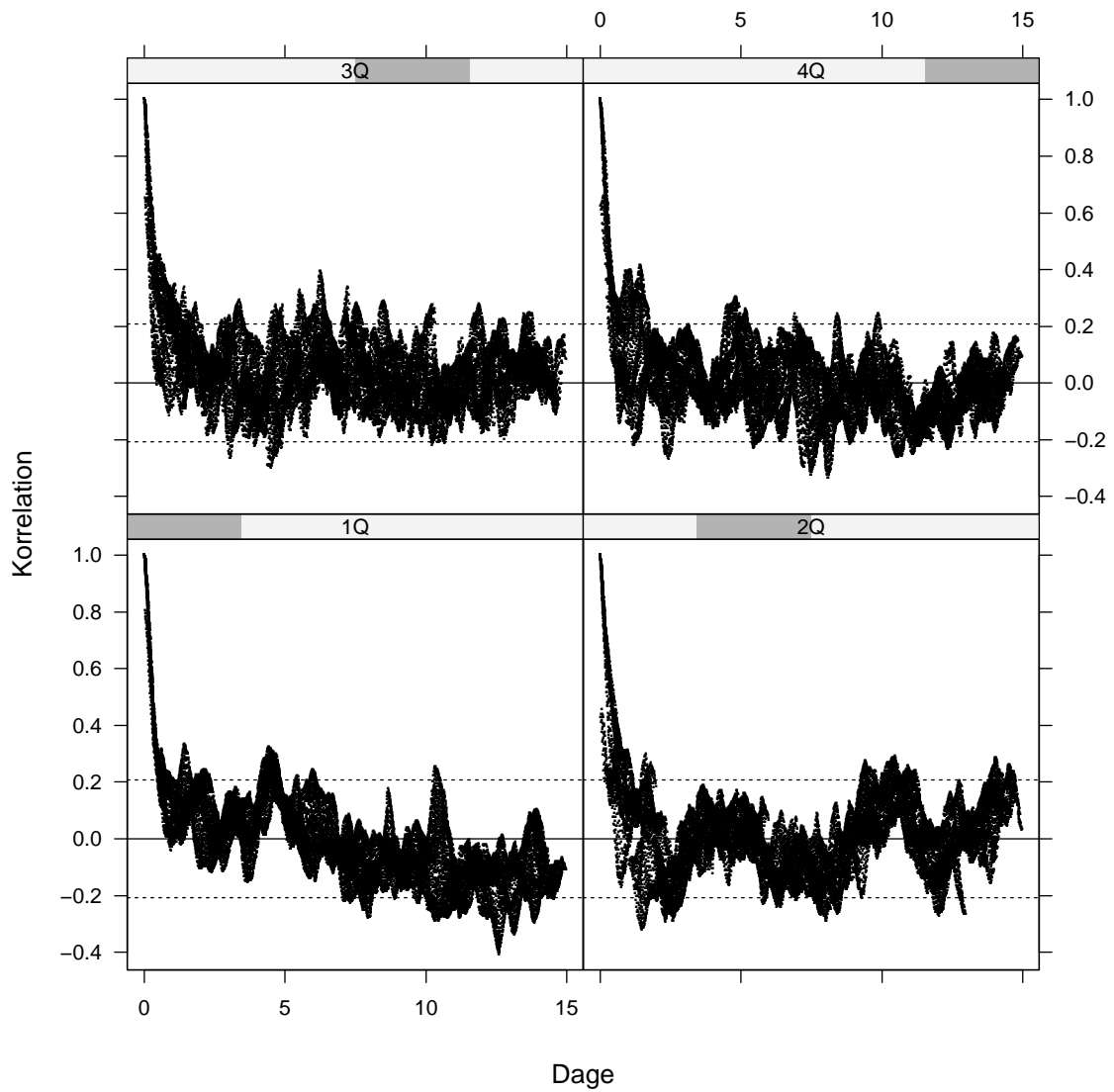
Figur 36: Histogrammer over estimerede krydskorrelationer i lag -14 til 14.



Figur 37: Estimerede krydskorrelationer i lag -1 for de enkelte klokkeslet.



Figur 38: Estimerede krydskorrelationer i lag 1 for de enkelte klokkeslet.



Figur 39: Estimerede auto- og krydskorrelationer (SACF og SCCF) versus tidsforskellen mellem tidspunkter for hvilke prædiktionerne gælder. De stiplede linier svare til $\pm 2/\sqrt{N}$. Beregningerne er foretaget for hvert kvartal for sig.

B Numeriske værdier til simulering

Variansens estimerede afhængighed af prognosen (figur 24, side 34): Se filen `progVar.csv`.

Spline basis med 6 interne ækvidistante knuder og grænseknuder ved 0.5 og 24: Se filen `basis.csv`.

Estimater svarende til MMA(1) model (model (1), side 13 uden leddet med \mathbf{y}_{t-1}):

Estimat af $\boldsymbol{\mu}$: Se filen `mma1_mu.csv`.

Estimat af $\boldsymbol{\Theta}_1$: Se filen `mma1_theta1.csv`.

Estimat af standardafvigelsen for $\boldsymbol{\epsilon}_t$: Se filen `mma1_sd.csv`.

Estimat af korrelationsmatricen for $\boldsymbol{\epsilon}_t$: Se filen `mma1_cor.csv`. Bemærk at denne er symmetrisk.

Gennemsnit af spline residualerne for hver halve time i døgnet: Se filen `spresid_avg.csv`.

Standardafvigelsen af de principale komponenter af spline residualerne: Se filen `spresid_pcsd.csv`. Bemærk at værdier under 0.1% af den maksimale værdi er sat til 0.1% af den maksimale værdi.

Loadings for de principale komponenter af spline residualerne: Se filen `spresid_pclloadings.csv`.

Idet \mathbf{L} (48×48) betegner loadings og \mathbf{R}_{pc} (366×48) eller (365×48) er simulerede principale komponenter (uafhængige) på basis af en multivariat normalfordeling med med standardafvigelser som anført i filen `spresid_pcsd.csv` fås de simulerede spline-residualer ved at løse ligningssystemet:

$$\mathbf{L}'\mathbf{R}' = \mathbf{R}'_{pc}$$

med hensyn til \mathbf{R} . Herefter adderes gennemsnittene fra filen `spresid_avg.csv` til hver kolonne.

Litteratur

- [1] William S. Cleveland and Susan J. Devlin. Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83:596–610, 1988.
- [2] Knut Conradsen. *En introduktion til Statistik*, volume 2. Informatik og Matematisk Modellering, 5 edition, 2001.
- [3] Bradley Efron and Robert J. Tibshirani. *An introduction to the bootstrap*. Chapman & Hall Ltd, 1993.
- [4] Peter Børre Eriksen. Economic and environmental dispatch of power/CHP production systems. *Electric Power Systems Research*, 57:33–39, 2001.
- [5] Trevor Hastie and Robert Tibshirani. *Generalized additive models*. Chapman & Hall Ltd, 1990.
- [6] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer, 2001.
- [7] Henrik Madsen. *Tidsrækkeanalyse*. Informatik og Matematisk Modellering, 1998.
- [8] Henrik Aalborg Nielsen. LFLM version 1.0, an S-PLUS / R library for locally weighted fitting of linear models. Technical Report 22/97, Department of Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark, 1997.
- [9] T. S. Nielsen, H. Madsen, and H. S. Christensen. WPPT – a tool for wind power prediction. In *Proceedings of the Wind Power for the 21st Century Conference*, Kassel, Germany, 2000.
- [10] T. S. Nielsen, H. Madsen, and J. Tøfting. Experiences with statistical methods for wind power prediction. In *Proceedings of the European Wind Energy Conference*, pages 1066 – 1069, Nice, France, 1999. James & James (Science Publishers).
- [11] Torben Skov Nielsen, editor. *Using Meteorological Forecasts in On-line Predictions of Wind Power*. Eltra, DK-7000 Fredericia, Denmark, 1999. ISBN: 87-90707-18-4.
- [12] Jens Pedersen. SIVAEL simulation program for combined heat and power production. In *International Conference on Applications of Power Production Simulation*, June 11–13 1990.