

Model for brændselsforbrug og elproduktion på Korsør kraftvarmeværk

Henrik Aalborg Nielsen & Henrik Madsen

20. marts 1995



Tidsrækkegruppen
Institut for Matematisk Modellering
Danmarks Tekniske Universitet

Indhold

1	Introduktion	7
2	Generel system- og modelbeskrivelse	7
3	Data	9
3.1	Målinger	9
3.2	Indledende databehandling	11
4	Metoder	17
4.1	Generel fremgangsmåde	17
4.2	Standardisering af variable	17
4.3	Modelspecifikation	18
4.4	Estimation	18
4.5	Modelkontrol	19
4.6	Test i modellen	20
4.7	Vurdering af modellens kvalitet	21
4.8	Anvendt program	21
5	Resultater	22
5.1	Dokumentation af beregninger	22
5.2	Standardisering af variable	22
5.3	Grafisk undersøgelse af data	23
5.4	Indledende model	27
5.5	Statistisk model for brændselsforbrug	28
5.6	Statistisk model for elproduktion	31
5.7	Vurdering af de endelige modellers kvalitet	34
6	Konklusion	35

7	Diskussion	35
A	Diagnostiske residual-plot	37
B	Styringsdiagram	44
C	Estimationsresultater	46
C.1	Indledende model for brændselsforbrug	46
C.2	Endelig model for brændselsforbrug	47
C.3	Indledende model for elproduktion	47
C.4	Endelig model for elproduktion	49
D	Test i modellerne	50
D.1	Brændselsforbrug	50
D.2	Elproduktion	52
E	Splus-funktioner	55
F	Dokumentation af beregninger	56

Figurliste

1	Grafisk fremstilling af en simpel model af motoranlægget.	8
2	Målt og beregnet varme til Korsør by og til Gasvæksvej (GV) for de timer hvor motoranlægget er i drift.	12
3	Gasforbrug, el- og varmereproduktion, når atypiske værdier betragtes som manglende.	13
4	Frem- og returtemperatur, samt differenstryk, når atypiske værdier betragtes som manglende.	14
5	Gasforbrug, el- og varmereproduktion, når atypiske værdier betragtes som manglende (motoranlægget er i drift).	15
6	Frem- og returtemperatur, samt differenstryk, når atypiske værdier betragtes som manglende (motoranlægget er i drift).	16
7	Scatterplot-matrix for alle variable (varmereproduktion i MWh (Q), fremløbstemperatur i $^{\circ}C$ (Ts), returtemperatur i $^{\circ}C$ (Tr), differenstryk i mVS (Dp), elproduktion i MWh (P), brændselsforbrug i Nm^3 (H)).	25
8	Fordelingen af de enkelte variable.	26
9	Forventet værdi af brændselsforbrug (i Nm^3) i det undersøgte område.	31
10	Forventet værdi af elproduktion (i MWh) i det undersøgte område.	34
11	Residualer.	37
12	Residualer for brændselsforbrug mod residualer for elproduktion.	37
13	Residualer. Observationer svarende til atypiske residualer udeladt.	38
14	Residualer for brændselsforbrug mod residualer for elproduktion. Observationer svarende til atypiske residualer udeladt.	38
15	Standard diagnostiske plot for residualer for brændselsforbrug. Observationer svarende til atypiske værdier udeladt.	39
16	Standard diagnostiske plot for residualer for elproduktion. Observationer svarende til atypiske værdier udeladt.	40
17	Residualer mod tiden. Observationer svarende til atypiske residualer udeladt.	41
18	Residualer mod målt og beregnet varme leveret til akkumulatortank.	42

19	Residualer for elproduktion. Observationer svarende til atypiske residualer og timer efter 20/10 kl. 21 (nr. 1222) udeladt.	43
----	---	----

Tabelliste

1	Atypiske værdier.	11
2	Gennemsnit (\bar{x}) og empirisk varians (S^2) anvendt ved standardisering af målinger.	22
3	Gennemsnit (\bar{x}) og empirisk varians (S^2) anvendt ved standardisering af energimængde til akkumulatortanken.	22
4	Test for hver enkelt forklarende variabel i model for brændselsforbrug.	29
5	Test for at de enkelte parametre svarende til leddet vist i tabellen er nul i modellen for brændselsforbrug som funktion af varmeproduktion (Q) og fremløbstemperatur (T_F). Antallet af frihedsgrader er 994.	29
6	Korrelation mellem parameterestimerne svarende til de viste led (brændselsforbrug).	30
7	Test for hver enkelt forklarende variabel i model for elproduktion.	32
8	Test for at de enkelte parametre svarende til leddet vist i tabellen er nul i modellen for elproduktion som funktion af varmeproduktion (Q) og fremløbstemperatur (T_F). Antallet af frihedsgrader er 734.	32
9	Korrelation mellem parameterestimerne svarende til de viste led (elproduktion).	33

Terminologi

Variabel:

Fysisk størrelse, der antager skiftende værdier (f.eks. temperatur).

Parameter:

Konstant der indgår i en matematisk model.

Nomenklatur

Kun symboler der anvendes gennemgående er nævnt her, øvrige symboler er beskrevet hvor de anvendes.

H Brændselsforbrug (gas) (Nm^3)

P Elproduktion (MWh)

Q Varmeproduktion (MWh)

T_F Fremløbstemperatur ($^{\circ}C$)

T_R Returtemperatur ($^{\circ}C$)

Δp Differenstryk (mVS)

De viste symboler er også anvendt for de tilsvarende standardiserede variable (se evt. afsnit 4.2). De standardiserede variable er dimensionsløse.

Tidsafledede variable angives ved en “.” over variabelen.

De variable repræsenterer timeværdier. For forbrug og produktion er der således tale om akkumulerede størrelser, mens der for de øvrige er tale om gennemsnit.

Resumé

Målinger af brændselsforbrug, el- og varmeproduktion, frem- og returtemperatur, samt differenstryk for Korsør kraftvarmeværk er blevet anvendt til bestemmelse af modeller for brændselsforbrugets og elproduktionens afhængighed af varmeproduktion, frem- og returtemperatur, samt differenstryk. Målingerne er hhv. totaler og gennemsnit pr. time.

Andengradspolynomier er tilstrækkelige til at sikre en forholdsvis høj forklaringsgrad (ca. 98%) indenfor det undersøgte område. Desuden er den tilbageværende systematik i residualerne (modelfejlen) begrænset.

Der kan ikke påvises nogen effekt af returtemperatur og differenstryk.

1 Introduktion

Formålet med arbejdet dokumenteret i denne rapport har været at bestemme modeller for det samlede brændselsforbrug og den samlede elproduktion per time, på Korsør kraftvarmeværk, potentielt som funktion af varmeproduktion, fremløbs- og returtemperatur, samt differenstryk. Modellerne søges opstillet på grundlag af målinger foretaget på systemet.

Det har kun været hensigten at undersøge stationære modeller, d.v.s. modeller hvor der ikke er afhængighed mellem timer. Dette synes rimeligt i betragtning af det anvendte tidsinterval mellem målingerne. Det er dog også blevet undersøgt hvorvidt sådanne modeller er rimelige.

Det er ikke nødvendigt at gennemlæse hele rapporten. Ønsker læseren f.eks. kun at anvende de opnåede modeller for brændselsforbrug og elproduktion kan disse findes i slutningen af hhv. afsnit 5.5 og 5.6. For at kunne anvende modellerne korrekt er det dog nødvendigt at forstå, for hvilke værdier af varmeproduktion, fremløbs- og returtemperatur, samt differenstryk, de er gyldige. Dette er beskrevet i afsnit 5.3.

Endelig bemærkes at decimaltal, grundet praktiske årsager, i denne rapport er skrevet med “.” istedet for “,”.

2 Generel system- og modelbeskrivelse

I dette afsnit gives en kort beskrivelse af produktionsenhederne for varme og kraft i Korsør. Derimod beskrives selve fjernvarmesystemet ikke. For en mere dybtgående beskrivelse af varmforsyningen i Korsør henvises til [Hellesen & Tuxen, 1994] og [Korsør Kommunale Værker, 1993].

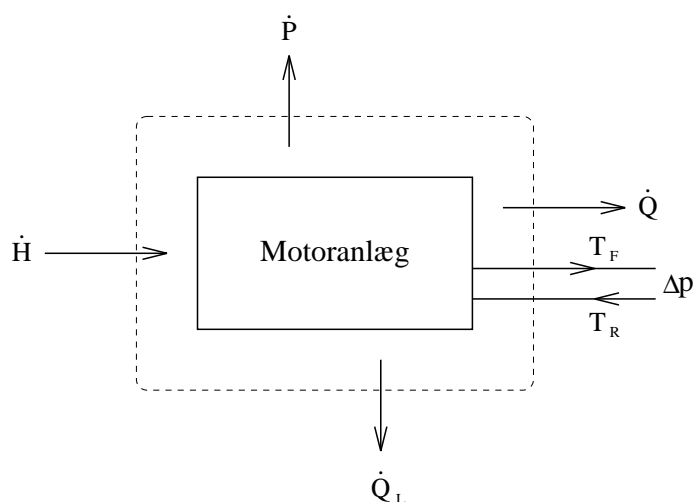
Varmeproduktionen i Korsør er mulig på to geografisk adskilte værker. På Gasværksvej findes et kedelanlæg af ældre dato, som sjældent benyttes. Størstedelen af produktionen foregår på Norbrinken, der dels består af et naturgasfyret kedelanlæg og dels af seks gasmotorer. Desuden er der på Norbrinken placeret en akkumulatortank. I appendiks B kan styrediagrammet for Norbrinken ses.

Idet der meget sjældent foregår produktion på Gasværksvej er det naturligvis ikke muligt ud fra målinger at bestemme en sammenhæng mellem brændselsforbrug og forklarende variable. I denne rapport betragtes derfor kun Norbrinken. Op- og afladning af akkumulatortanken medtages ikke i modellen, idet dette ikke vil kunne beskrives v.h.a. en stationær model. Dertil kommer at op- og afladningen ikke er et fysisk fænomen, men derimod styret af værkets personale.

I den periode hvor data er blevet opsamlet har der ikke været nogen produktion på kedlerne på Norbrinken, men kun på motoranlægget. Det er derfor kun muligt at finde en model, der beskriver brændselsforbrug og elproduktion på motoranlægget som funktion af de (potentielt) forklarende variable; varmereproduktion, frem- og returtemperatur, samt differenstryk.

Ifølge [Hellesen & Tuxen, 1994] er motoranlæggets elkapacitet på $6.3MW$, mens varmekapaciteten er på $9.7MW$.

Modelmæssigt betragtes motoranlægget som een enhed hvortil brændsel tilføres med hastigheden \dot{H} , el og varme produceres hhv. med hastighederne \dot{P} og \dot{Q} . Desuden mistes energi som varme, dette sker med hastigheden \dot{Q}_L . Varmetab og fordelingen mellem el- og varmereproduktion antages at kunne påvirkes af belastningen på værket (\dot{Q}), frem- og returtemperatur (T_F og T_R), samt af differenstrykket (Δp) over frem og retur. Modellen af motoranlægget er skitseret i figur 1.



Figur 1. Grafisk fremstilling af en simpel model af motoranlægget.

Under stationære forhold vil energien være bevaret til ethvert tidspunkt (t):

$$\dot{H}(t) = \dot{Q}(t) + \dot{P}(t) + \dot{Q}_L(t) \quad (1)$$

Under virkelige driftsforhold, hvor værket ydre driftforhold hele tiden ændres, vil relationen (1) ikke være opfyldt. Erstattes de tidsafledede størrelser derimod med gennemsnit over et tidsrum af en vis længde er det derimod klart at (1) vil gælde som en tilnærmelse, hvor præcisionen stiger med længden af det anvendte tidsrum. Tilsvarende betragtninger gælder for indflydelsen af de øvrige forklarende variable.

I denne sammenhæng blev det valgt at betragte gennemsnit for tidsrum af længden en time. Denne værdi er dels fastlagt af praktiske årsager (målingerne var direkte tilgængelige på timebasis), dels skønnes dette tidsrum at være langt sammenlignet med motoranlæggets tidskonstanter og dels skønnes tidsrummet at være kort nok til at dække værket virkelige operationsområde i den betragtede periode.

Ifølge [Korsør Kommunale Værker, 1993] anvendes varmen fra udstødning og kølevand til opvarmning af fjernvarmevand. For det betragtede motoranlæg er der derfor ikke mulighed for at påvirke forholdet mellem el- og varmeproduktion. Af denne grund vil det ikke være relevant at medtage både el- og varmeproduktion i udtrykket for brændselsforbruget. Desuden bør man ikke medtage både brændselsforbrug og varmeproduktion i udtrykket for elproduktionen.

Idet der i det efterfølgende refereres til værdier hørende til den samme time er det derfor formålet at bestemme passende udtryk for funktionerne f og g således, at der for det samlede brændselsforbrug (H), den samlede el- og varmeproduktion (P og Q), den gennemsnitlige frem- og returtemperatur (T_F og T_R), samt det gennemsnitlige differenstryk (Δp) gælder:

$$\hat{H} = f(Q, T_F, T_R, \Delta p) \quad \text{og} \quad \hat{P} = g(Q, T_F, T_R, \Delta p), \quad (2)$$

hvor det ikke på forhånd er givet at alle forklarende variable bør medtages. “ $\hat{}$ ” angiver forventningsværdien.

3 Data

3.1 Målinger

De anvendte data for motoranlægget består primært af timevise værdier for:

- Gasforbrug (Nm^3)
- Elproduktion (MWh)

- Varmeproduktion (MWh)
- Fremløbstemperatur til Korsør by ($^{\circ}C$)
- Returtemperatur fra Korsør by ($^{\circ}C$)
- Differenstryk ved Bondesbjerg (måler 6111 på styringsdiagrammet i appendiks B) (mVS)

Sekundært er følgende målinger anvendt:

- Varmeleverance til Korsør by (MWh)
- Cirkuleret vandmængde, Korsør by (m^3)
- Varmeleverance til Gasværksvej (MWh)
- Cirkuleret vandmængde, Gasværksvej (m^3)
- Fremløbstemperatur til Gasværksvej ($^{\circ}C$)
- Returtemperatur fra Gasværksvej ($^{\circ}C$)

Målingerne er indekseret således, at værdier registreret for time t hører til tidsrummet løbende fra t til $t + 1$. Målingerne dækker en periode på 1632 timer løbende fra 31. august, 1994, kl. 00:00 til 6. november, 1994, kl. 23:00.

For følgende perioder mangler målinger af en eller flere af de primære variable nævnt ovenfor:

- 6/9, kl. 00:00
- 7/9, kl. 02:00 til 8/9, kl. 08:00
- 13/9, kl. 00:00 til 23:00
- 18/10, kl. 03:00 til 06:00
- 19/10, kl. 00:00
- 21/10, kl. 00:00 til 23/10, kl. 23:00

Det samme gælder gruppen af sekundære variable.

De enkelte tidspunkter identificeres som hovedregel fortløbende således, at den første time (31/8, kl. 00:00) har nummer 1 og den sidste dermed nummer 1632.

De registrerede værdier er ikke vist, idet der forekommer meget store værdier, der betyder, at det ikke er muligt at vise de øvrige målinger med en tilstrækkelig præcision.

3.2 Indledende databehandling

Atypiske værdier

Der forekommer i data enkelte meget høje eller lave værdier. Disse blev identificeret og derefter betragtet som manglende.

For alle andre variable end differenstrykket drejer det sig om 8/9 kl. 9, 18/10 kl. 7 og 19/10 kl. 12. For differenstrykket drejer det sig om 23/9 kl. 5 og 27/9 fra kl. 8 til 11. De registrerede atypiske værdier fremgår af tabel 1.

Gasforbrug (Nm^3)	10950, 7770, 40440
Elproduktion (MWh)	36.07, 23.39, 85.35
Varmeproduktion (MWh)	68, 48, 8972
Cirkuleret vandmængde, Gasværksvej (m^3)	1070, 530, 51560
Cirkuleret vandmængde, Korsør by (m^3)	1440, 440, 35840
Leveret varme, Gasværksvej (MWh)	27, 22, 6697
Leveret varme, Korsør by (MWh)	42, 18, 8199
Differenstryk (mVS)	5.66, 2.57, 0.01, 0.01, 1.53

Tabel 1. Atypiske værdier.

Målinger

Målingerne af de primære variable er vist på figur 3 og 4¹.

Af figur 3 fremgår det at motoranlægget ofte har et brændselsforbrug på 0 Nm^3 . Dette skyldes at motoranlægget typisk slukkes om natten. De tidspunkter hvor motoranlægget er slukket bør naturligvis ikke medtages i analysen af data. Det blev besluttet at betragte motoranlægget som slukket når brændselsforbruget er nul (opløsningen på målingerne af brændselsforbruget er større end 10 Nm^3). På figur 5 og 6¹ er de i analysen anvendte data vist.

Varme leveret til Gasværksvej og Korsør by

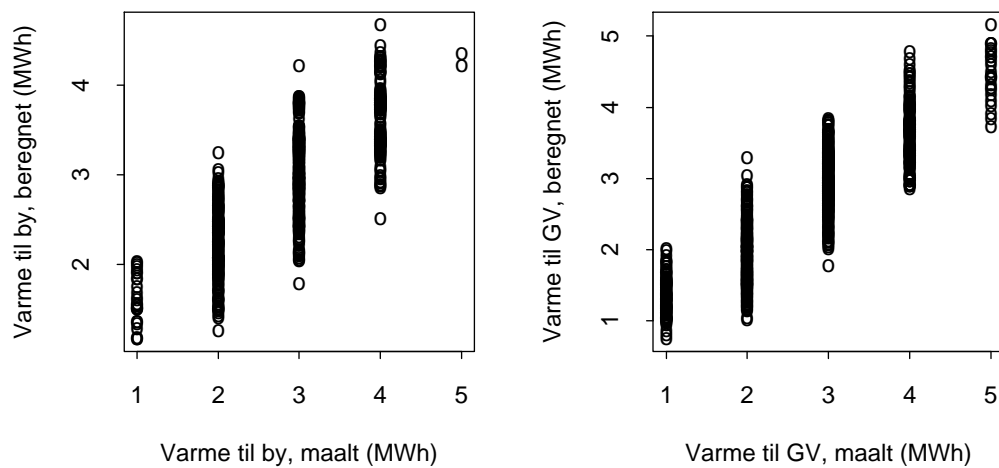
Den leverede varme er registreret som hele MWh i intervallet 0 til 5 MWh . For at få et mere nuanceret udtryk for den leverede varme (Q_{lev}) er den leverede varme i MWh beregnet som anført i ligning (3) ud fra den cirkulerede vandmængde i m^3 (V), samt frem- og returtemperaturen i $^{\circ}C$ til hhv. Gasværksvej og Korsør by.

$$Q_{lev} = \rho c_p V (T_F - T_R) / k, \quad (3)$$

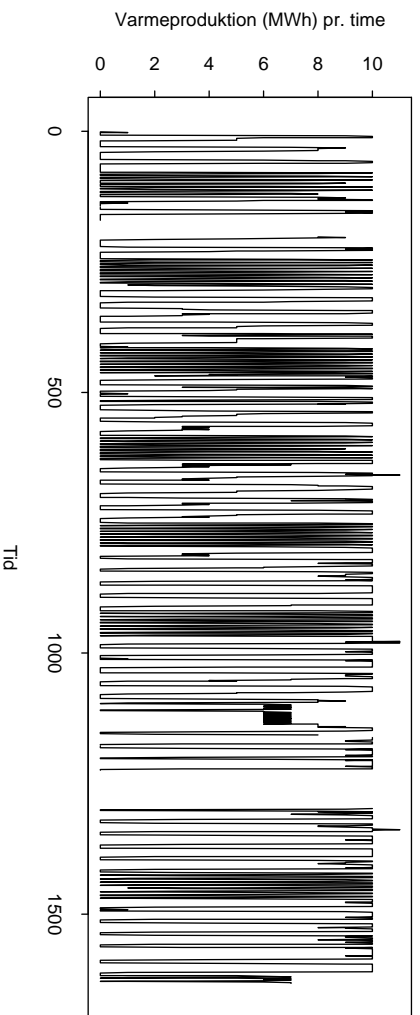
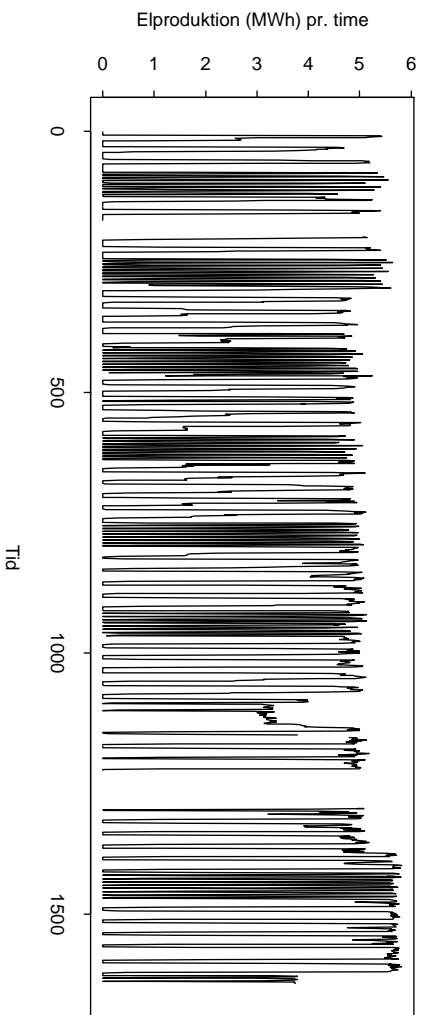
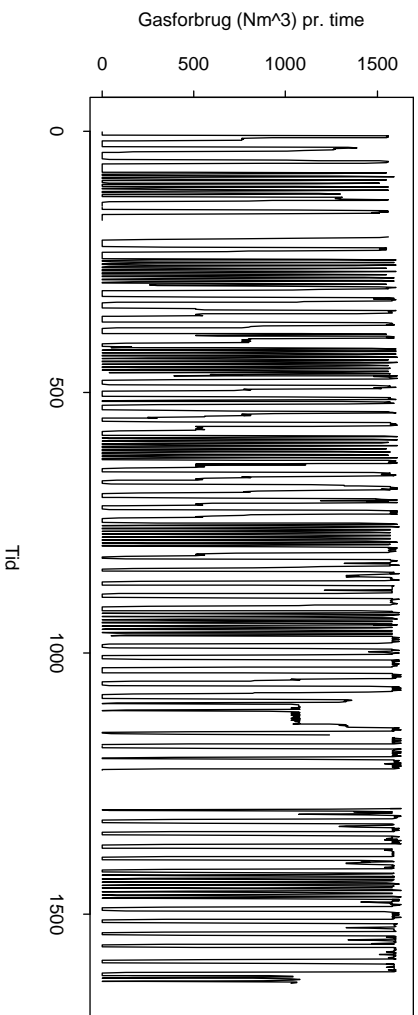
hvor $\rho = 980 kg/m^3$, $c_p = 4.187 kJ/kg^{\circ}C$ og $k = 3.6 \times 10^6 kJ/MWh$.

¹Figurene er placeret i slutningen af dette kapitel

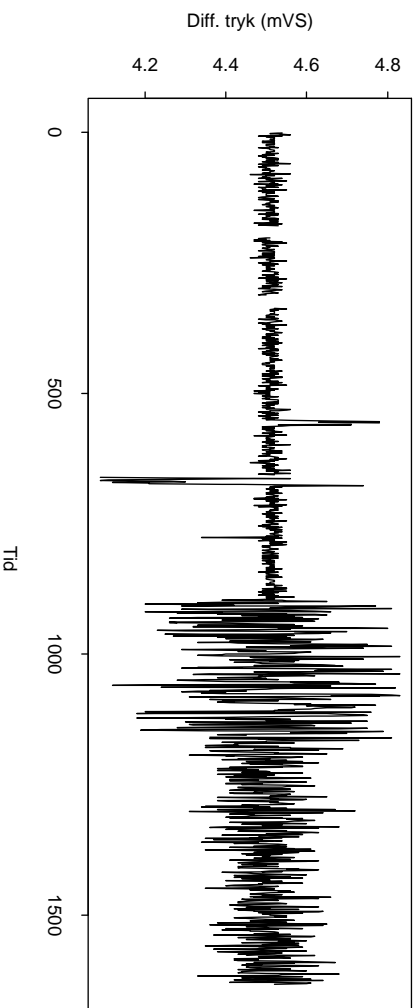
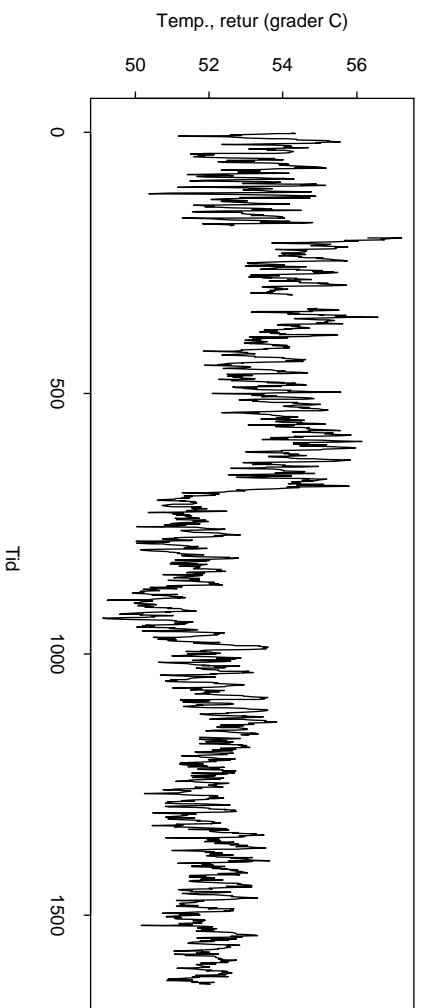
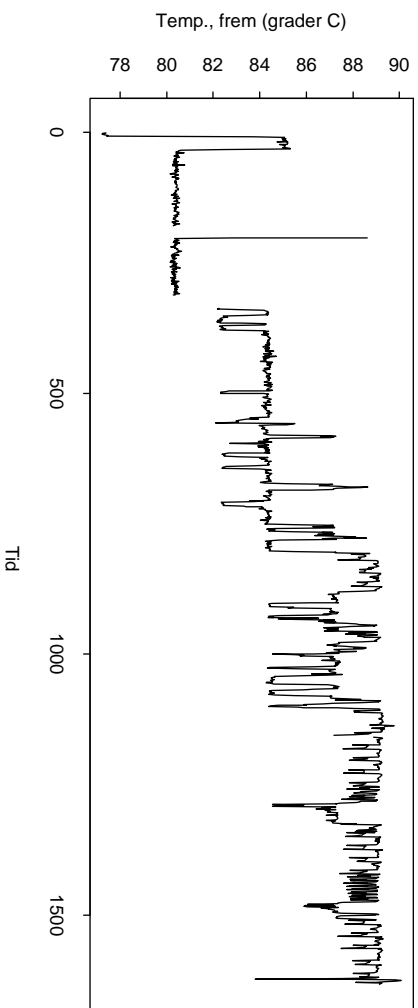
Figur 2 viser de beregnede mod de målte varmeleverancer for de timer hvor motoranlægget er i drift. Det ses at overensstemmelsen er ringe, hvilket delvis kan tilskrives den grove diskretisering af de målte størrelser.



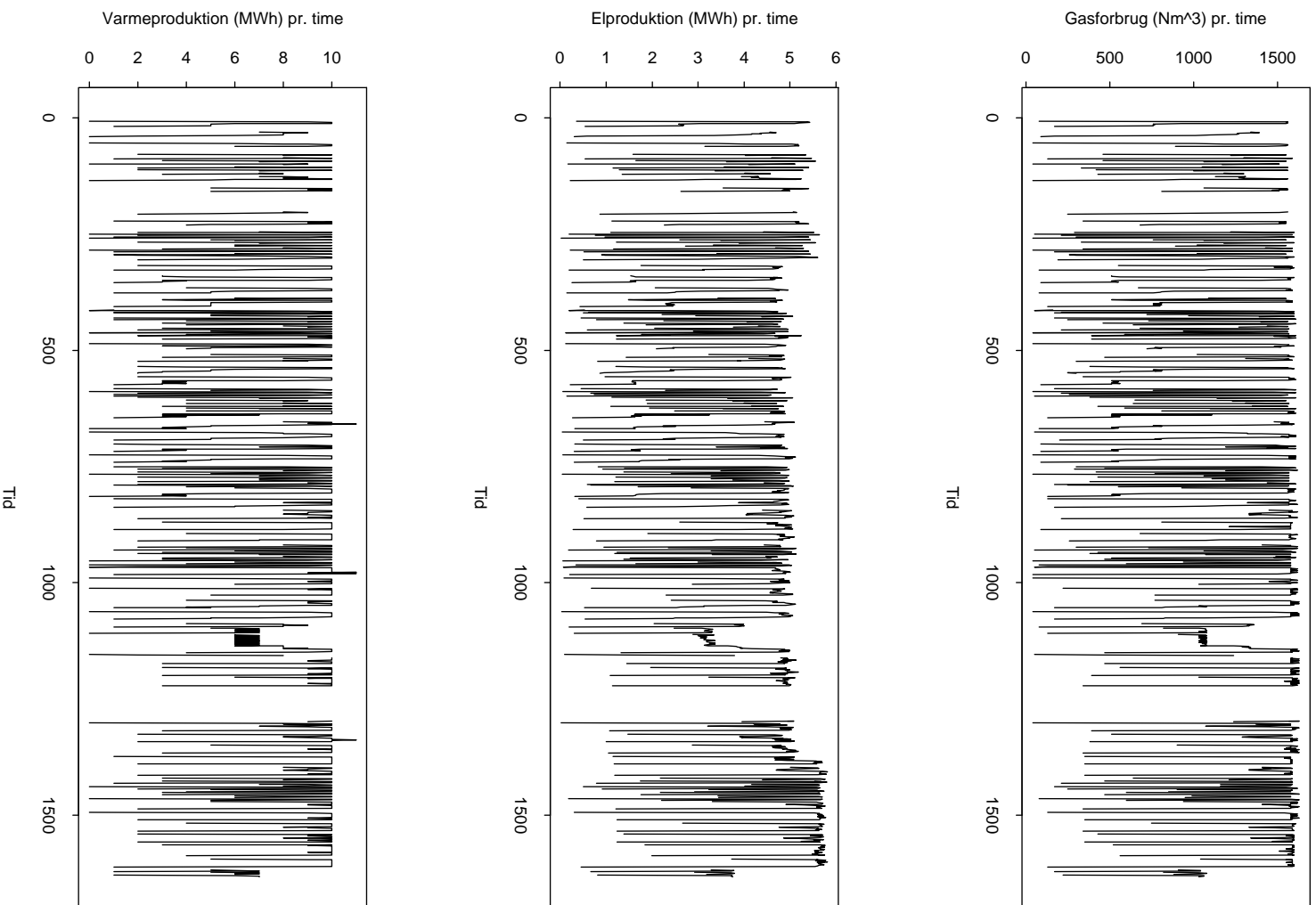
Figur 2. Målt og beregnet varme til Korsør by og til Gasvæksvej (GV) for de timer hvor motoranlægget er i drift.



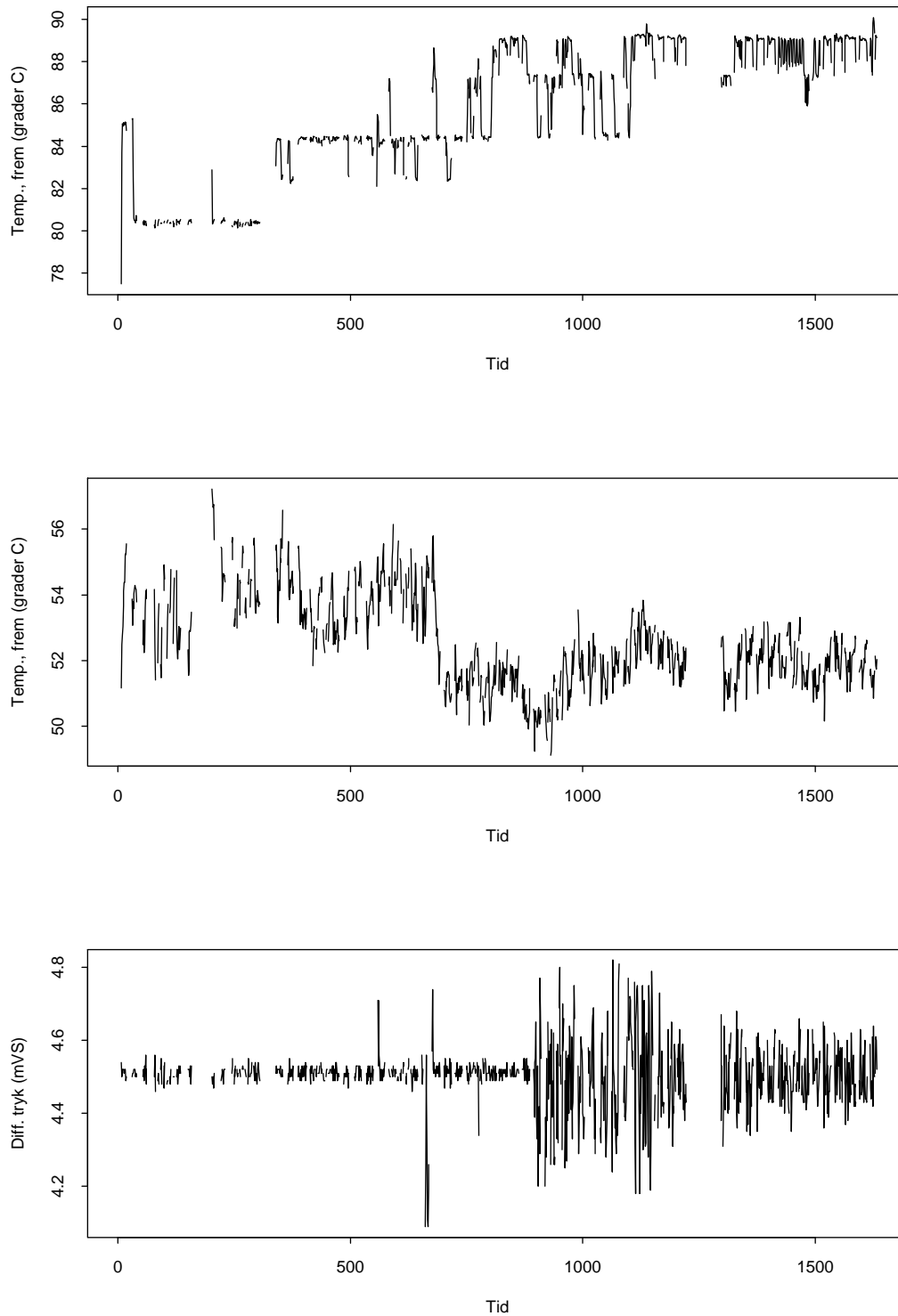
Figur 3. Gasforbrug, el- og varmeproduktion, når atypiske værdier betragtes som manglende.



Figur 4. Frem- og returtemperatur, samt differenstryk, når atypiske værdier betragtes som manglende.



Figur 5. Gasforbrug, el- og varmeproduktion, når atypiske værdier betragtes som manglende (motoranlægget er i drift).



Figur 6. Frem- og returtemperatur, samt differenstryk, når atypiske værdier betragtes som manglende (motoranlægget er i drift).

4 Metoder

4.1 Generel fremgangsmåde

I dette afsnit gøres der rede for den generelle fremgangsmåde, der er anvendt i forbindelse med opstilling af de statistiske modeller. Det grundlæggende synspunkt er, at kun de dele af modellen der på basis af data kan vises at have betydning, bør medtages. Argumentet herfor er at man herved får en lokal approksimation, der er så præcis som muligt. Approksimationen er lokal i forhold til det område der udspændes af de målte forklarende variable og bør ikke benyttes uden for dette område. Dette gælder også de af de målte forklarende variable, der ikke medtages i den endelige model.

Fremgangsmåden er følgende:

1. Standardiser samtlige variable således at de er af samme størrelsesorden.
2. Vælg en grundstruktur for modellen og estimer parametrene i denne.
3. Kontroller at modellen er tilstrækkelig.
4. Undersøg om alle forklarende variable valgt under 2 bør medtages i modellen.
5. Undersøg om yderligere modelreduktion er mulig.
6. Vurder modellens kvalitet.

Ad. 1 Dette øger den numeriske stabilitet af estimationsproceduren.

Ad. 2 og 3 Bemærk, at valideringen kan lede til den konklusion, at modellen må udvides.

4.2 Standardisering af variable

Lad x_i betegne målingerne af en variabel. Den standardiserede variabel z_i defineres ved:

$$z_i = \frac{x_i - \bar{x}}{2\sqrt{S^2}}, \quad (4)$$

hvor \bar{x} betegner gennemsnittet og S^2 den empiriske varians ($S^2 = \sum(x_i - \bar{x})^2 / (n - 1)$, hvor n er antallet af målinger). For normalfordelte data vil denne standardisering medføre at ca. 95% af de standardiserede værdier ligger i intervallet fra -1 til 1.

Gennemsnit og empirisk varians er beregnet for samtlige målinger, med undtagelse af de atypiske værdier nævnt i tabel 1.

4.3 Modelspecifikation

Idet den afhængige variabel y_i betegner enten brændselsforbrug eller elproduktion på motoranlægget for observation i er sammenhængen mellem y_i og de forklarende variable approksimeret v.h.a. et 2. gradspolynomium:

$$\begin{aligned}
 y_i = & a + b_f T_{Fi} + b_r T_{Ri} + b_q Q_i + b_p \Delta p_i \\
 & + c_f T_{Fi}^2 + c_r T_{Ri}^2 + c_q Q_i^2 + c_p \Delta p_i^2 \\
 & + c_{fr} T_{Fi} T_{Ri} + c_{fq} T_{Fi} Q_i + c_{fp} T_{Fi} \Delta p_i \\
 & + c_{rq} T_{Ri} Q_i + c_{rp} T_{Ri} \Delta p_i + c_{qp} Q_i \Delta p_i + \epsilon_i, \quad i = 1, \dots, N
 \end{aligned} \tag{5}$$

Hvor a , b og c er parametre, N er antallet af observationer og ϵ_i er normalfordelt hvid støj med middelværdi 0 og varians σ^2 .

Det forventes, at sammenhængen ikke er stærkt ulineær. Et 2. gradspolynomium vil derfor formentlig give en passende fleksibilitet. Modelkontrollen vil afsløre såfremt dette ikke er tilfældet.

4.4 Estimation

Modellen (5) er lineær i parametrene og støjen er normalfordelt, dvs. modellen er en sædvanlig lineær regressionsmodel.

Nu indføres:

$$Y = [y_1 \ y_2 \ \dots \ y_N]^T \tag{6}$$

$$x = \begin{bmatrix} 1 & T_{F1} & T_{R1} & \dots & Q_1 \Delta p_1 \\ 1 & T_{F2} & T_{R2} & \dots & Q_2 \Delta p_2 \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 1 & T_{FN} & T_{RN} & \dots & Q_N \Delta p_N \end{bmatrix} \tag{7}$$

$$\theta = [a \ b_f \ b_r \ \dots \ c_{qp}]^T \tag{8}$$

Ifølge [Conradsen, 1984], kapitel 3 er maksimumlikelihood estimatet for parametrene ($\hat{\theta}$) da givet ved:

$$\hat{\theta} = (x^T x)^{-1} x^T Y, \tag{9}$$

under antagelse af at matricen x har fuld rang. Dette estimat minimerer summen af de kvadrede afvigelser mellem modellens prediktioner og de faktiske observationer. Dette gælder selvom fordelingen af støjen ikke antages kendt. I såfald kaldes estimatet for mindste kvadraters estimatet.

Et centralt estimat for variansen af støjen er:

$$\hat{\sigma}^2 = \frac{1}{N - p} (Y - x\hat{\theta})^T (Y - x\hat{\theta}), \quad (10)$$

hvor p er antallet af parametre. For modellen (5) er $p = 15$.

Endelig estimeres kovariansmatricen ($D(\hat{\theta})$) for parameterestimerne ved:

$$\hat{D}(\hat{\theta}) = \hat{\sigma}^2 (x^T x)^{-1} \quad (11)$$

Ud fra denne bestemmes korrelationsmatricen som sædvanligt ved at dele alle elementer med kvadratroden af produktet af de tilsvarende diagonalelementer.

4.5 Modelkontrol

Idet \hat{y}_i betegner forventningsværdien af y_i givet værdier af de forklarende variable svarende til observation i , defineres residualen ved $y_i - \hat{y}_i$. Hvis modellen er tilstrækkelig bør disse residualer tilnærmelsesvis være uafhængige og have samme varians. Antagelsen om normalitet er derimod ikke så kritisk.

Modelkontrollen foretages i udpræget grad v.h.a. grafiske metoder. Det er almindeligt at plote følgende:

- Residualerne mod hver af de forklarende variable og mod tiden. Herved kan en eventuel (ulinear) sammenhæng med de forklarende variable muligvis afsløres.
- Residualerne mod de forventede værdier af den afhængige variabel, herved vil man typisk kunne afsløre om variansen er uafhængig af den forudsagte værdi.
- Fraktilerne i den empiriske fordeling af residualerne mod fraktilerne i den standardiserede normalfordeling. Hvis residualerne er normalfordelt vil dette resultere i en tilnærmelsesvis ret linie.

Yderligere kan modellens forudsætning om uafhængige residualer kontrolleres v.h.a. en række statistiske test. Da der i dette tilfælde er tale om en brudt tidsrække blev det besluttet kun at teste om sandsynligheden for fortegnsskift fra residual

til residual er lig $\frac{1}{2}$. En afvisning af denne hypotese leder til den konklusion at residualerne ikke er uafhængige.

Med N residualer er der mulighed for $N - 1$ fortegnsskift. Såfremt residualerne er uafhængige vil antallet af fortegnsskift derfor være binomialfordelt med antalsparameter $N - 1$ og sandsynlighedsparameter $\frac{1}{2}$. Testet udføres derfor ved at tælle antallet af fortegnsskift og undersøge om dette antal kan stamme fra den nævnte fordeling.

4.6 Test i modellen

Det anvendte signifikansniveau (α), dvs. sandsynligheden for at forkaste en hypotese når den er sand, er 0.05.

Test af om parameter er lig nul

For at undersøge modellens struktur kan det være af interesse at teste om en parameter er nul mod alternativet at den er forskellig fra nul. Såfremt en parameter er nul vil estimatet af denne parameter delt med standardafvigelsen på dette estimat være approksimativt t-fordelt med $N - p$ frihedsgrader. Hypotesen om at parameteren er nul afvises derfor hvis estimatet delt med standardafvigelsen ligger uden for intervallet $[t(N - p)_{\alpha/2}, t(N - p)_{1-\alpha/2}]$, se f.eks. [Madsen, 1989], side 197.

Test af om en forklarende variabel kan udelades af modellen

For at mindske antallet af test kan det være af interesse at teste hypotesen at en række parametre samtidig er nul mod alternativet at mindst en er forskellig fra nul. Vælges disse parametre således, at de svarer til at udelade en af de forklarende variable i modellen (5) fås et test for om den valgte variabel kan udelades af modellen.

Testet foretages ved at sammeligne forskellen i de forventede værdier mellem modellen (5) og en model hvor de ønskede parametre er sat til nul med residualerne for modellen (5). Hvis denne forskel er stor må det afvises, at de undersøgte parametre alle kan sættes til nul.

Lad \hat{y}_i være de forventede værdier under modellen (5) og $\hat{y}_{(red)i}$ være de forventede værdier når de parametre, der ønskes undersøgt, sættes til nul. Lad endvidere p være antallet af parametre i (5) og lad p_{red} være antallet af parametre der estimeres i den reducerede model. Teststørrelsen er da:

$$F = \frac{\sum_{i=1}^N (\hat{y}_i - \hat{y}_{(red)i})^2 / (p - p_{red})}{\sum_{i=1}^N (y_i - \hat{y}_i)^2 / (N - p)} \quad (12)$$

Under hypotesen at den reducerede model er tilstrækkelig, vil denne størrelse være F-fordelt med $(p - p_{red}, N - p)$ frihedsgrader. Hypotesen afvises derfor, såfremt $F > F(p - p_{red}, N - p)_{1-\alpha}$.

4.7 Vurdering af modellens kvalitet

Modellens kvalitet vurderes bl.a. ved inspektion af residualerne som beskrevet i afsnit 4.5. Et mere objektivt mål for modellens kvalitet er den såkaldte forklaringsgrad (R^2). Denne størrelse udtrykker hvor stor en del af den samlede variation i den afhængige variabel, der forklares af modellen:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}, \quad (13)$$

hvor \bar{y} er gennemsnittet af den afhængige variabel. Det er ikke nødvendigvis formålstjenligt at vælge den model, der resulterer i den største forklaringsgrad. Dette skyldes at, forklaringsgraden stiger med antallet af estimerede parametre alene p.g.a. tilfældige variationer. Man vil således ikke nødvendigvis få gode resultater ved anvendelse af den opnåede model på fremtidige data. Forklaringsgraden vil typisk stige voldsomt ved inddragelse af de første led. Herefter vil stigningen blive mindre, for til sidst at være marginal.

4.8 Anvendt program

Stplus version 3.0 er anvendt til de numeriske beregninger. Estimation, test m.v. er foretaget v.h.a. funktionen `lm`. Testet for antal fortegnsskift blev foretaget (eksakt) v.h.a. funktionen `binom.test`.

Resultatet af statistiske test angives ved den såkaldte p-værdi, der er den værdi af signifikansniveauet (α) som betyder at testets resultat ligger netop på grænsen mellem signifikant og insignifikant. Hvis man anvender et signifikansniveau på 5% vil man altså afvise hypotesen hvis p-værdien er under 0.05.

5 Resultater

5.1 Dokumentation af beregninger

I appendiks F ses et såkaldt “Backtrack” fra audit faciliteten i Splus. Dette “Backtrack” er udført på et objekt, der består af samtlige anvendte resultater.

I dette kapitel er der, når der henvises til resultater, også henvist til et objekt. Ved at undersøge “Backtrack” for dette objekt kan samtlige operationer, der er udført for at opnå de anførte resultater, genskabes ud fra de indlæste data. Det kan dog kræves at der udføres simple operationer på objekterne. Et typisk eksempel er beregning af residualer, der udføres med kommandoen `residuals(<objekt>)`, hvor `<objekt>` indeholder resultatet af estimationen.

5.2 Standardisering af variable

Tabel 2 viser de værdier af gennemsnit og empirisk varians, der blev anvendt ved standardiseringen af målingerne, se evt. afsnit 4.2. Værdierne er anført med stor præcision for at give mulighed for en præcis verifikation af resultaterne givet de anvendte data.

	H (MWh)	P (MWh)	Q (MWh)	T_F ($^{\circ}C$)	T_R ($^{\circ}C$)	Δp (mVS)
\bar{x}	817.9356	2.612952	5.100592	85.63312	52.6728	4.506667
S^2	490017.3	5.100949	19.26948	9.67454	1.848864	0.006999535

Tabel 2. Gennemsnit (\bar{x}) og empirisk varians (S^2) anvendt ved standardisering af målinger.

Tabel 3 viser de tilsvarende værdier for energimængden til akkumulatortanken.

	Målt (MWh)	Beregnet (MWh)
\bar{x}	1.999	2.054842
S^2	9.366365	8.87779

Tabel 3. Gennemsnit (\bar{x}) og empirisk varians (S^2) anvendt ved standardisering af energimængde til akkumulatortanken.

Med mindre andet er anført gælder parameterestimer, m.v. for de standardiserede variable.

De standardiserede variable er dimensionsløse. På figurene m.v. anføres f.eks. det standardiserede brændselsforbrug derfor som " $H (-)$ ", mens det oprindelige anføres som " $H (Nm^3)$ ".

5.3 Grafisk undersøgelse af data

De anvendte data er vist tidligere på figur 5 og 6. I dette afsnit anføres nogle grafiske undersøgelser af disse data.

Figur 7 viser en såkaldt scatterplot-matrix for alle variable i den oprindelige enhed. Hvert plot i matricen forekommer på begge sider af diagonalen, dog med akserne ombyttet. Det er derfor kun nødvendigt at undersøge plottene på den ene side af denne. I dette tilfælde omtales kun plot under diagonalen. Matricen kan umiddelbart være lidt vanskelig at forstå f.eks. er plottet øverst til venstre (under diagonalen) et plot af fremløbstemperaturen (T_s) mod varmeproduktionen (Q). Enhederne på 1. akserne er skiftevis anført overfor og under selve figuren, mens enhederne på 2. akserne skiftevis er anført til højre og venstre.

Af figuren ses:

- De fleste kombinationer af varmeproduktion og de øvrige uafhængige variable er til stede i data.
- Der synes at være en svag negativ korrelation mellem frem- og returtemperatur, således at returtemperaturen falder med stigende fremløbstemperatur. Dette øger korrelationen mellem parameterestimerne hørende til de to temperaturer.
- For lave værdier af fremløbstemperaturen og høje værdier af returtemperaturen varieres differenstrykket meget lidt (hvilket skyldes at differenstrykket kun varieres i den sidste del af perioden).
- Der er en stærk korrelation mellem el- og varmeproduktion.
- Der er en stærk korrelation mellem brændselsforbrug og varmeproduktion.
- Der er en stærk korrelation mellem brændselsforbrug og elproduktion.
- Hverken elproduktion eller brændselsforbrug synes at afhænge af frem- og returtemperatur, eller differenstryk.

Figur 8 viser fordelingen af de enkelte variable. Denne figur kan anvendes til yderligere at undersøge i hvilket område resultaterne vil være gyldige. For de forklarende variable ses:

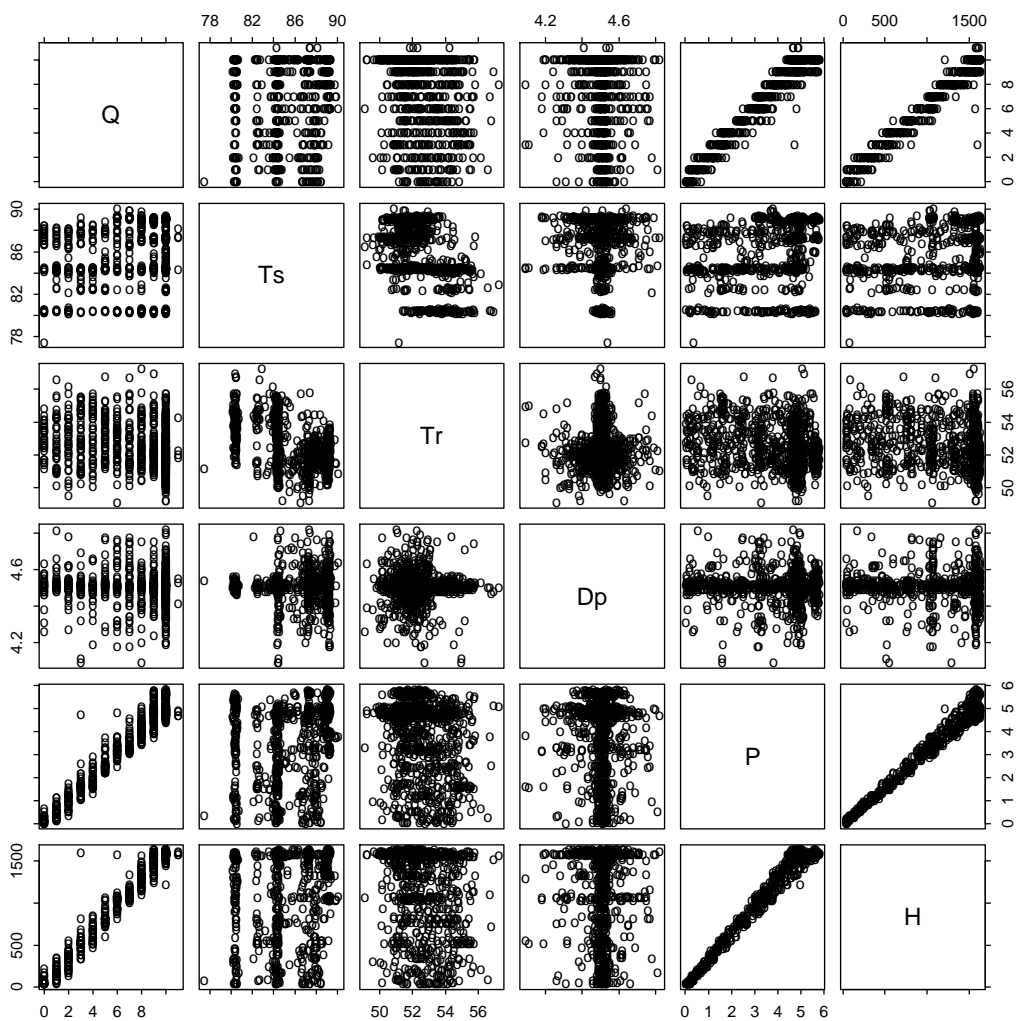
- Fremløbstemperaturen ligger hovedsagligt i intervallet 80 til 90 °C. I dette interval er især de høje værdier repræsenteret.
- Returtemperaturen ligger hovedsageligt i intervallet 50 til 55 °C, med hovedparten af data fra 51 til 53 °C.
- Langt de fleste værdier af differenstrykket ligger i intervallet fra 4.4 til 4.6 *mVS*.
- Varmeproduktionen er for det meste 10 *MWh* (per time), værdier herunder er jævnt fordelt og enkelte værdier er 0 *MWh* (per time).

Den grafiske undersøgelse viser bl.a., at korrelationen mellem de uafhængige variable til hvert tidspunkt er forholdsvis lav og at varmereproduktionen er den mest betydende af de forklarende variable. Fordelingen af varmereproduktionen viser, at hovedvægten ved estimationen vil blive lagt på den situation hvor varmereproduktionen er nær maksimum.

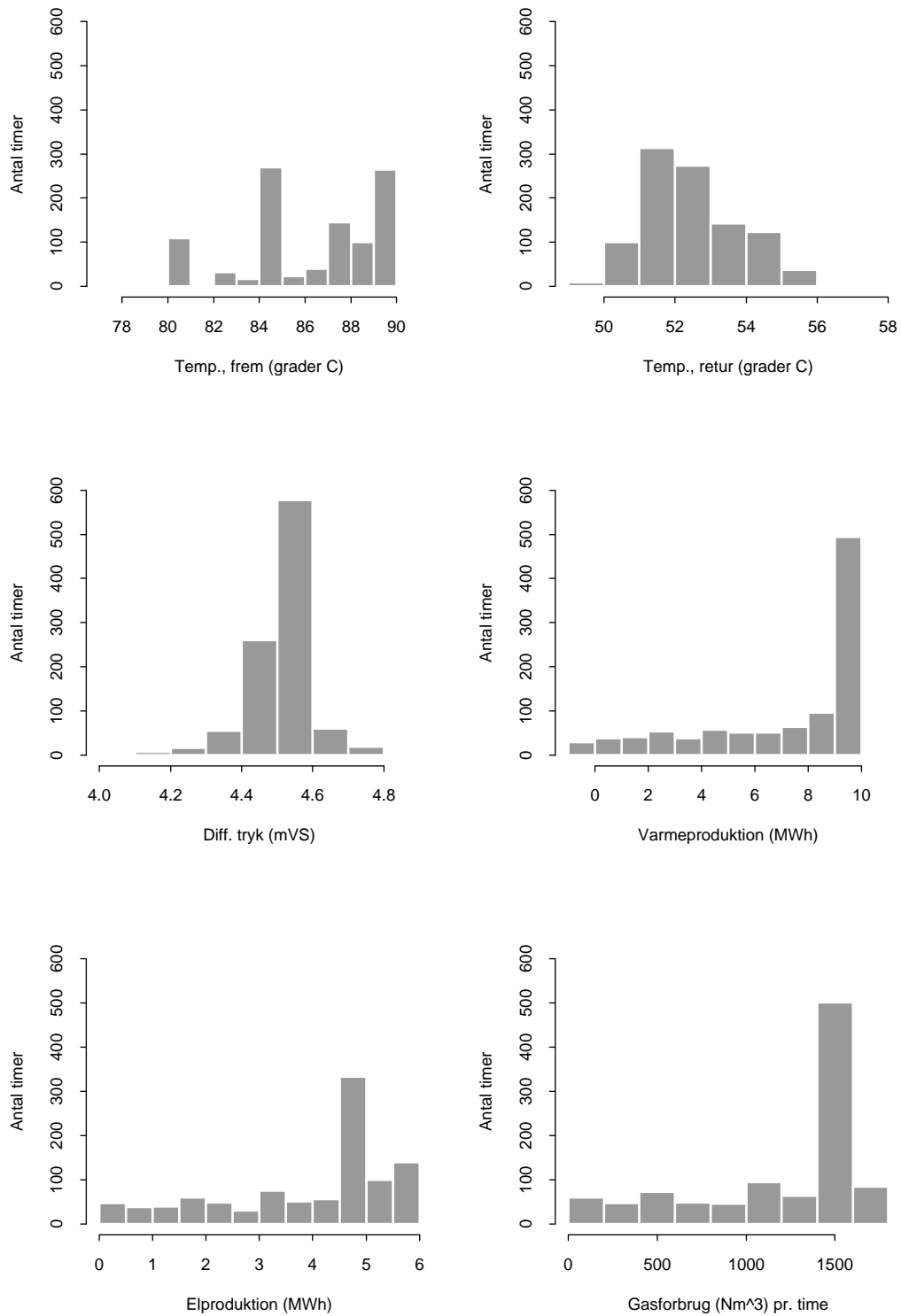
Ovenstående betyder, at der som hovedregel vil gælde, at modeller fundet på basis af de tilgængelige data vil være rimelige approksimationer såfremt:

- Fremløbstemperaturen er mellem 80 og 90 °C.
- Returtempearturen er mellem 50 og 55 °C.
- Differenstrykket er mellem 4.4 og 4.6 *mVS*.
- Varmeproduktionen er mellem 0 og 10 *MWh* (per time).

Approksimationen vil dog ikke være lige god overalt i dette område.



Figur 7. Scatterplot-matrix for alle variable (varmeproduktion i MWh (Q), fremløbstemperatur i $^{\circ}C$ (Ts), returtemperatur i $^{\circ}C$ (Tr), differenstryk i mVS (Dp), elproduktion i MWh (P), brændselsforbrug i Nm^3 (H)).



Figur 8. Fordelingen af de enkelte variable.

5.4 Indledende model

Figurer henvist til i dette afsnit er vist i appendiks A.

Parametrene i modellen (5), hvor y betegner hhv. brændselsforbrug (H) og elproduktion (P) blev estimeret på basis af data for de timer hvor motoranlægget er i drift og hvor alle variable er målt (1003 observationer). Resultaterne er samlet i objekterne `r.H.lm.full` og `r.P.lm.full`.

Figur 11 viser residualer mod det fortløbende timenummer. På figuren ses to atypiske residualer. Figur 12 viser residualer fra modellen for brændselsforbrug mod residualer fra modellen for elproduktion. Her ses tre atypiske residualer. Inspektion af data viser at disse værdier hører til timerne med nummerne 490, 491 og 879 (20/9, kl. 9-10 og 6/10, kl. 14).

For at undgå indflydelse fra disse atypiske observationer slettes disse fra data og beregningerne udføres igen. Resultaterne er samlet i objekterne `rod.H.lm.full` og `rod.P.lm.full`.

Figur 13 og 14 svarer til figurerne nævnt ovenfor. Figur 15 og 16 viser residualer for hhv. brændselsforbrug og elproduktion mod de forklarende variable, den under modellen forventede værdi, samt et fraktildiagram for residualerne (se evt. afsnit 4.5). Endelig viser figur 17 residualerne mod tiden på døgnet.

Det konkluderes at:

- Ingen af residualerne kan siges at være atypiske i størrelse, eller placering i planen bestående af de to residualekvenser (figur 13 og 14).
- Residualer fra modellen for elproduktion synes at være systematisk større end nul i den sidste del af perioden (figur 13). Hvilket betyder at modellen ikke er tilstrækkelig.
- Der synes ikke at være nogen systematik i forhold til de forklarende variable, variansen synes konstant og residualerne synes tilnærmelsesvis normalfordelte (figur 15 og 16).
- Der synes at være en svag døgnvariation (figur 17).

Da frem- og returtemperatur er målt efter akkumulatortanken og da denne delvist bestemmer fremløbstemperaturen under afladning blev det undersøgt om noget af variationen i residualerne kan forklares af hvorvidt tanken oplades eller aflades. Figur 18 viser residualerne mod standardiserede værdier af energimængden til akkumulatortanken. Af figuren ses, at der ikke er nogen synlig sammenhæng.

Idet 1000 observationer er anvendt er der mulighed for 999 fortegnsskift. Ved uafhængige residualer er det forventede antal 499.5. For brændselsforbruget observeres

535 fortegnsskift, hvilket er lidt større end det forventede antal, den tilsvarende p-værdi er 0.0267. For elproduktionen observeres 344 fortegnsskift, hvilket er meget mindre end det forventede antal, hvilket skyldes den utilstrækkelige model. Den tilsvarende p-værdi er 0.0000.

For varmeproduktionen er modellen tilfredsstillende, dog er der lidt flere fortegnsskift end forventet. Dette nedsætter gyldigheden af de test der foretages i modellen, specielt hvis resultatet er på grænsen mellem signifikant og insignifikant. Af denne grund er det vigtigt at sammenligne forklaringsgraden for den fulde model med den der opnås for den endelige model.

For elproduktionen er der overensstemmelse mellem analysen af antallet af fortegnsskift og konklusionen truffet på basis af den grafiske undersøgelse; residualerne er ikke uafhængige. Det er formentlig den sidste del af perioden der er problemet, idet residualerne her næsten alle er positive.

Under en antagelse om at målingerne af elproduktionen i denne sidste del af perioden er atypiske blev observationer efter den 20/10, kl. 21 (time 1222) udeladt og estimationen udført igen. Resultaterne er samlet i objektet `rod.subset.P.lm.full`.

Figur 19 viser residualerne fra modellen for elproduktion mod det fortløbende time-nummer, dette plot er som forventet tilfredsstillende.

Estimationen er baseret på 740 observationer, hvilket giver mulighed for 739 fortegnsskift. Ved uafhængige residualer er det forventede antal 369.5. Der observeres 344 fortegnsskift, hvilket er mindre end det forventede antal. Den tilsvarende p-værdi er 0.0658.

Det konkluderes, at modellen for elproduktionen er tilfredsstillende for de 740 første observationer. P-værdien er dog ikke væsentligt over de sædvanlige 5%. Af denne grund er det også for elproduktionen vigtigt at sammenligne forklaringsgraden for den fulde model med den der opnås for den endelige model.

5.5 Statistisk model for brændselsforbrug

Det er i afsnit 5.4 vist at modellen for brændselsforbrug giver en god approksimation. For at undersøge betydningen af de enkelte forklarende variable blev parametrene i følgende modeller estimeret, () angiver i hvilket objekt resultatet er samlet:

- Fremløbstemperatur udeladt (`rod.H.lm.Ts.excl`).
- Returtemperatur udeladt (`rod.H.lm.Tr.excl`).
- Varmeproduktion udeladt (`rod.H.lm.Q.excl`).
- Differenstryk udeladt (`rod.H.lm.Dp.excl`).

Test for at hver enkelt af disse modeller er tilstrækkelige, mod alternativet at den fulde model må anvendes, blev udført v.h.a. S-plus funktionen `f.F.test` vist i appendiks E. Resultaterne er vist i tabel 4.

	F	frihedsgrader	p-værdi
Fremløbstemperatur	2.4295	(5,985)	0.0336
Returtemperatur	0.7497	(5,985)	0.5863
Varmeproduktion	9083.328	(5,985)	0.0000
Differenstryk	0.2936	(5,985)	0.9166

Tabel 4. Test for hver enkelt forklarende variabel i model for brændselsforbrug.

Det ses, at varmereproduktionen, som forventet, har overordentlig stor betydning for brændselsforbruget. Desuden er effekten af fremløbstemperaturen svagt signifikant, mens der ikke kan påvises nogen effekt af returtemperatur og differenstryk.

Parametrene i en model hvori kun varmereproduktion og fremløbstemperatur indgår blev estimeret. Resultaterne er samlet i objektet `rod.H.1m.TrDp.excl`. Testes denne model mod modellen beskrevet i afsnit 5.4 fås $F = 0.5055$, med (9,985) frihedsgrader giver dette en p-værdi på 0.8713. Modellen hvori kun varmereproduktion og fremløbstemperatur indgår er altså tilstrækkelig. Resultaterne af testene for at de enkelte parametre i denne model er nul er vist i tabel 5.

Led	Estimat	Standardafvigelse	Teststørrelse	p-værdi
Konstant	0.0141	0.0033	4.2292	0.0000
T_F	0.0154	0.0049	3.1692	0.0016
Q	0.9797	0.0054	181.0133	0.0000
T_F^2	0.0043	0.0078	0.5522	0.5810
Q^2	-0.0416	0.0152	-2.7291	0.0065
$T_F Q$	-0.0074	0.0102	-0.7197	0.4719

Tabel 5. Test for at de enkelte parametre svarende til leddet vist i tabellen er nul i modellen for brændselsforbrug som funktion af varmereproduktion (Q) og fremløbstemperatur (T_F). Antallet af frihedsgrader er 994.

Af tabellen ses, at parametrene svarende til leddene T_F^2 og $T_F Q$ ikke er signifikante, hvorimod de øvrige parametre er signifikante. P-værdiernes størrelse indikerer, at dette resultat er særdeles tydeligt, der er derfor ikke grund til at udelade leddet med den største p-værdi for at teste leddet med den næststørste p-værdi².

²Denne p-værdi vil blive reduceret, men den vil stadig være langt over 0.05.

Parametrene i den endelige model for brændselsforbruget blev estimeret. Resultaterne er samlet i objektet `rod.H.final`. Ligning (14) viser den estimerede sammenhæng mellem det forventede brændselsforbrug (\hat{H}), fremløbstemperatur og varmeproduktion. Standardafvigelsen på parameterestimerterne er angivet i parentes efter disse.

$$\hat{H} = 0.0148(0.0030) + 0.0126(0.0034)T_F + 0.9807(0.0053)Q - 0.0432(0.0150)Q^2 \quad (14)$$

Ved anvendelse af denne formel skal fremløbstemperatur og varmeproduktion først transformeres efter formlerne:

$$T_F (-) = \frac{T_F (\text{°C}) - 85.63312\text{°C}}{2 \times \sqrt{9.67454\text{°C}^2}} \quad (15)$$

og

$$Q (-) = \frac{Q (MWh) - 5.100592MWh}{2 \times \sqrt{19.26948(MWh)^2}} \quad (16)$$

For at få brændselsforbruget udtrykt i Nm^3 skal \hat{H} fra (14) transformeres efter formlen:

$$\hat{H} (Nm^3) = 817.9356Nm^3 + 2 \times \sqrt{490017.3(Nm^3)^2} \times \hat{H} (-) \quad (17)$$

Se evt. tabel 2 og afsnit 4.2.

Korrelationen mellem parameterestimerterne er vist i tabel 6. Det bemærkes, at korrelationen mellem estimerterne svarende til Q^2 og konstantleddet er stor. Dette skyldes formentlig at motoranlægget, når det er i drift, ofte køre med maksimal effekt.

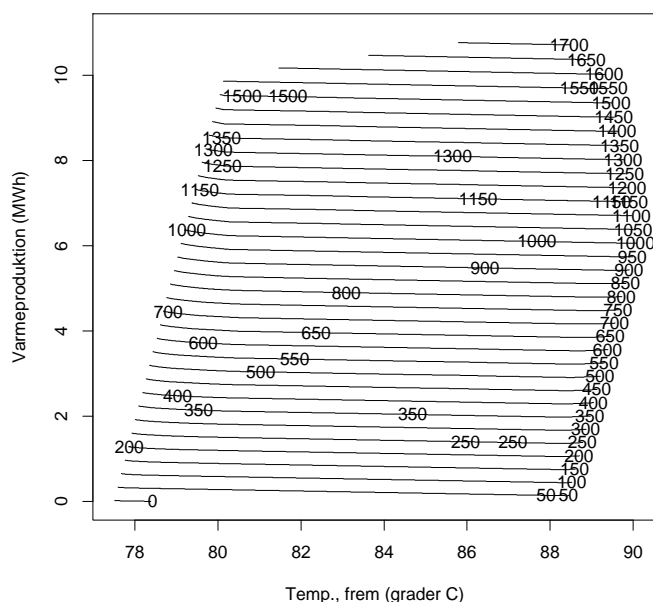
	Konstant	T_F	Q
T_F	0.1045		
Q	0.0853	-0.2003	
Q^2	-0.7631	-0.0885	-0.5455

Tabel 6. Korrelation mellem parameterestimerterne svarende til de viste led (brændselsforbrug).

Standardafvigelser på parameterestimer kan, sammen med korrelationen mellem disse, anvendes til beregning af konfidensintervaller for forudsagte værdier, se f.eks. [Conradsen, 1984].

Ovenstående resultater kan findes i appendiks C.1, C.2 og D.1, der bl.a. også viser estimer for parametrene for modellen beskrevet i afsnit 5.4.

På figur 9 ses den forventede værdi af brændselsforbruget i det undersøgte område. Bemærk at antallet af observationer hvor varmeproduktionen er over 10 *MWh* er meget begrænset.



Figur 9. Forventet værdi af brændselsforbrug (i Nm^3) i det undersøgte område.

5.6 Statistisk model for elproduktion

Det er i afsnit 5.4 vist, at modellen for elforbrug er en god approksimation for de 740 første observationer, se beskrivelsen i afsnit 5.4. For at undersøge betydningen af de enkelte forklarende variable blev parametrene i følgende modeller estimeret. () angiver i hvilket objekt resultatet er samlet:

- Fremløbstemperatur udeladt (`rod.subset.P.lm.Ts.excl`).
- Returtemperatur udeladt (`rod.subset.P.lm.Tr.excl`).

- Varmeproduktion udeladt (`rod.subset.P.lm.Q.excl`).
- Differenstryk udeladt (`rod.subset.P.lm.Dp.excl`).

Test for at hver enkelt af disse modeller er tilstrækkelige, mod alternativet, at den fulde model må anvendes, blev udført v.h.a. S-plus funktionen `f.F.test` vist i appendiks E. Resultaterne er vist i tabel 7.

	F	frihedsgrader	p-værdi
Fremløbstemperatur	25.1125	(5,725)	0.0000
Returtemperatur	0.6798	(5,725)	0.6389
Varmeproduktion	5102.3300	(5,725)	0.0000
Differenstryk	0.1948	(5,725)	0.9645

Tabel 7. Test for hver enkelt forklarende variabel i model for elproduktion.

Det ses, at varmemproduktionen, som forventet, har overordentlig stor betydning for elproduktionen. Desuden er effekten af fremløbstemperaturen signifikant, mens der ikke kan påvises nogen effekt af returtemperatur og differenstryk.

Parametrene i en model hvori kun varmemproduktion og fremløbstemperatur indgår blev estimeret. Resultaterne er samlet i objektet `rod.subset.P.lm.TrDp.excl`. Testes denne model mod modellen beskrevet i afsnit 5.4 fås $F = 0.4410$, med (9,725) frihedsgrader giver dette en p-værdi på 0.9129. Modellen hvori kun varmemproduktion og fremløbstemperatur indgår er altså tilstrækkelig. Resultaterne af test for at de enkelte parametre i denne model er nul er vist i tabel 8.

Led	Estimat	Standardafvigelse	Teststørrelse	p-værdi
Konstant	-0.0234	0.0042	-5.5505	0.0000
T_F	0.0008	0.0066	0.1152	0.9084
Q	0.9449	0.0073	129.4988	0.0000
T_F^2	0.0859	0.0098	8.7706	0.0000
Q^2	-0.0316	0.0196	-1.6166	0.1064
$T_F Q$	-0.0689	0.0137	-5.0392	0.0000

Tabel 8. Test for at de enkelte parametre svarende til leddet vist i tabellen er nul i modellen for elproduktion som funktion af varmemproduktion (Q) og fremløbstemperatur (T_F). Antallet af frihedsgrader er 734.

Af tabellen ses, at parametrene svarende til leddene T_F og Q^2 ikke er signifikante, hvorimod de øvrige parametre er stærkt signifikante. Hvis leddet T_F udelades af

modellen bliver p-værdien for leddet Q^2 lig 0.1039, se evt. appendiks D.2. Leddet Q^2 udelades derfor også af modellen.

Parametrene i den endelige model for elproduktionen blev estimeret. Resultaterne er samlet i objektet `rod.subset.P.final`. Ligning (18) viser den estimerede sammenhæng mellem den forventede elproduktion (\hat{P}), fremløbstemperatur og varmeproduktion. Standardafvigelsen på parameterestimerterne er angivet i parentes efter disse.

$$\hat{P} = -0.0280(0.0032) + 0.9392(0.0057)Q + 0.0854(0.0086)T_F^2 - 0.0701(0.0103)T_FQ \quad (18)$$

Ved anvendelse af denne formel skal fremløbstemperatur og varmeproduktion først transformeres efter formlerne (15) og (16).

For at få elproduktionen udtrykt i MWh skal \hat{P} fra (18) transformeres efter formelen:

$$\hat{P} (MWh) = 2.612952MWh + 2 \times \sqrt{5.100949(MWh)^2} \times \hat{P} (-) \quad (19)$$

Se evt. tabel 2 og afsnit 4.2.

Korrelationen mellem parameterestimerterne er vist i tabel 9.

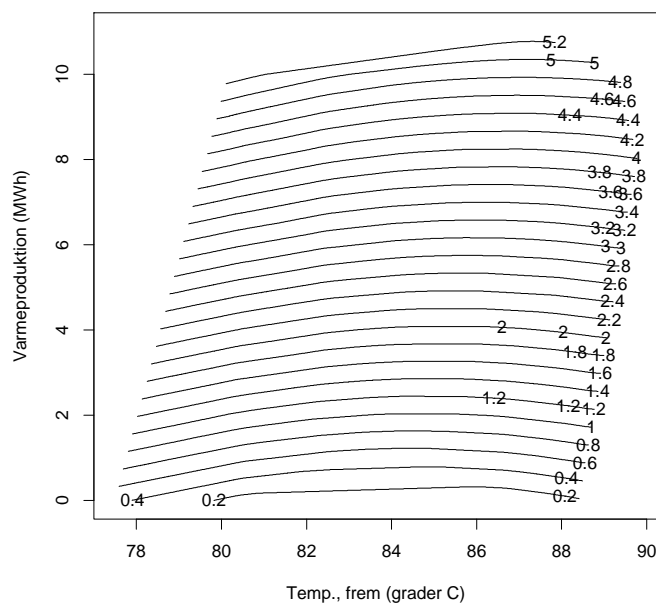
	Konstant	Q	T_F^2
Q	-0.4576		
T_F^2	-0.6077	0.0198	
T_FQ	-0.1771	0.1139	0.0806

Tabel 9. Korrelation mellem parameterestimerterne svarende til de viste led (elproduktion).

Standardafvigelser på parameterestimer kan, sammen med korrelationen mellem disse, anvendes til beregning af konfidensintervaller for forudsagte værdier, se f.eks. [Conradsen, 1984].

Ovenstående resultater kan findes i appendiks C.3, C.4 og D.2, der bl.a. også viser estimerter for parametrene for modellen beskrevet i afsnit 5.4.

På figur 10 ses den forventede værdi af brændselsforbruget i det undersøgte område. Igen skal det bemærkes, at antallet af observationer hvor varmeproduktionen er over 10 MWh er meget begrænset.



Figur 10. Forventet værdi af elproduktion (i *MWh*) i det undersøgte område.

5.7 Vurdering af de endelige modellers kvalitet

De grundlæggende modeller for brændselsforbrug og elproduktion blev undersøgt i afsnit 5.4. Disse modeller blev fundet tilnærmelsesvis tilstrækkelige for så vidt angår residualerne, men det blev også konkluderet, at forklaringsgraderne for disse modeller måtte sammenlignes med forklaringsgraderne for de endelige modeller.

Disse forklaringsgrader kan findes i appendiks C. Det viser sig at:

- For brændselsforbruget er forklaringsgraden for den grundlæggende model 98.19%, mens den for den endelige er 98.18%.
- For elproduktionen er forklaringsgraden for den grundlæggende model 97.47%, mens den for den endelige er 97.45%.

I begge tilfælde er der altså tale om meget små reduktioner, hvilket betyder, at de statistiske test ikke har ført til forkerte resultater, selvom antagelsen om uafhængig støj tilsyneladende ikke gælder.

Forklaringsgraderne for begge modeller er rimeligt store, hvilket betyder at man som en god approksimation kan anvende modellerne uden at tage hensyn til usikkerheden på modellernes prediktioner. Dette er en meget generel betragtning, hvis gyldighed naturligvis afhænger af den konkrete anvendelse.

6 Konklusion

For de data der har været tilgængelige i forbindelse med arbejdet beskrevet i denne rapport kan der ikke påvises nogen sammenhæng mellem på den ene side brændselsforbrug og elproduktion og på den anden side returtemperatur og differenstryk. Brændselsforbrug og elproduktion afhænger derimod tydeligvis af fremløbstemperatur og (naturligvis) af varmeproduktion.

Det er fundet, at sammenhængen som en rimelig approksimation kan beskrives ved 2. gradspolynomier. Disse polynomier blev reduceret således, at kun betydende led er medtaget. For disse polynomier blev det fundet, at modellen for brændselsforbruget forklarer 98.18% af den totale variation i brændselsforbruget. Modellen for elproduktionen forklarer 97.45% af den totale variation i elproduktionen.

7 Diskussion

Elproduktion

Modellen for elproduktion er kun baseret på den første del af observationerne. Dette skyldes, at den sidste del (ca. 25%) af observationerne afveg systematisk fra den første del. Hvorvidt de virkelige forhold bedst afspejles i den første eller sidste del af perioden kan dog ikke afgøres ud fra data.

Målepunkter

Som nævnt i afsnit 3.1 er frem- og returtemperatur, samt differenstryk ikke målt ved udgangen af motoranlægget.

For så vidt angår temperaturene, der er målt efter akkumulatortanken, synes valget af målepunkt ikke at være afgørende idet residualerne (modelfejlen) ikke synes at være afhængig af energimængden til/fra akkumulatortanken. Dette kan muligvis skyldes, at akkumulatortanken næsten altid oplades når motoranlægget er i drift og/eller at akkumulatortanken bestemmer fremløbstemperaturen fra motoranlægget når denne aflades.

Fordelingen af forklarende variable

Konklusionerne i denne rapport har naturligvis kun gyldighed inden for det område af de forklarende variable, som det aktuelle datagrundlag udspænder. Ønskes dette område udvidet må der enten observeres over en anden/længere periode, eller der må udføres forsøg hvor de forklarende variable varieres aktivt med det formål at kunne bestemme en model med et bredt gyldighedsområde. Et sådant forsøg er naturligvis begrænset af at fjernvarmevandet skal kunne afkøles. Desuden er returtemperaturen vanskelig at variere over et bredt interval. Endelig vil returtemperaturen udvise en dynamisk afhængighed af bl.a. fremløbstemperaturen.

I de tilgængelige data er varmeproduktionen ofte nær maksimum. Dette indebærer at de opnåede modellers præcision er størst i dette område. Ved at udføre et forholdsvist simpelt forsøg, hvor varmeproduktionen blev fordelt jævnt over det mulige område, kunne dette problem undgås.

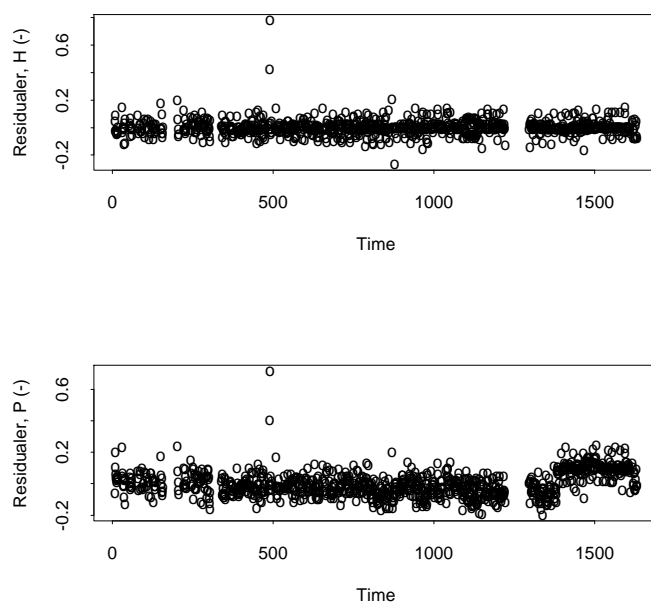
Driftudgifter

Ved hjælp af de fundne modeller kan et overslag over de variable udgifter til drift af motoranlægget bestemmes. Hertil kommer udgifter til bl.a. egetforbrug af el og varme, samt CO_2 -afgift.

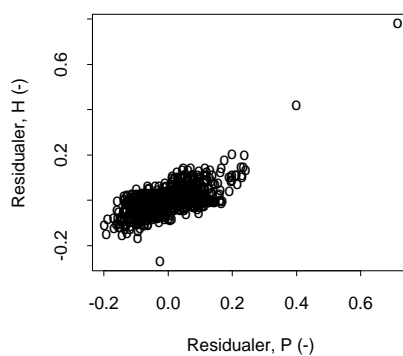
Det er ofte udgifter til drift af hele værket, der er af interesse. Man må derfor også have informationer om udgiften til drift af kedler. Disse informationer kan enten fremskaffes ud fra målinger over en periode hvori kedlerne er i drift, eller man kan anvende tilgængelig information om virkningsgrader m.v.. Såfremt kedlerne på årsbasis ikke bidrager med en væsentlig del af de samlede udgifter er det oplagt, at en præcis modellering af kedlerne ikke er nødvendig.

A Diagnostiske residual-plot

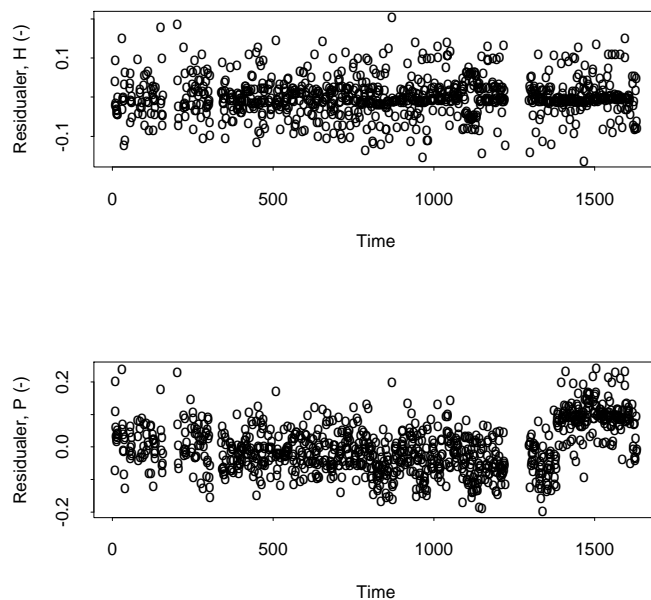
Nedenfor vises diagnostiske plot for residualerne fra modellen (5).



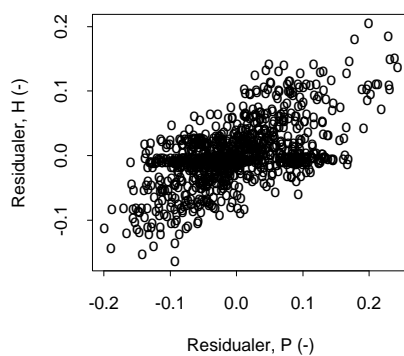
Figur 11. Residualer.



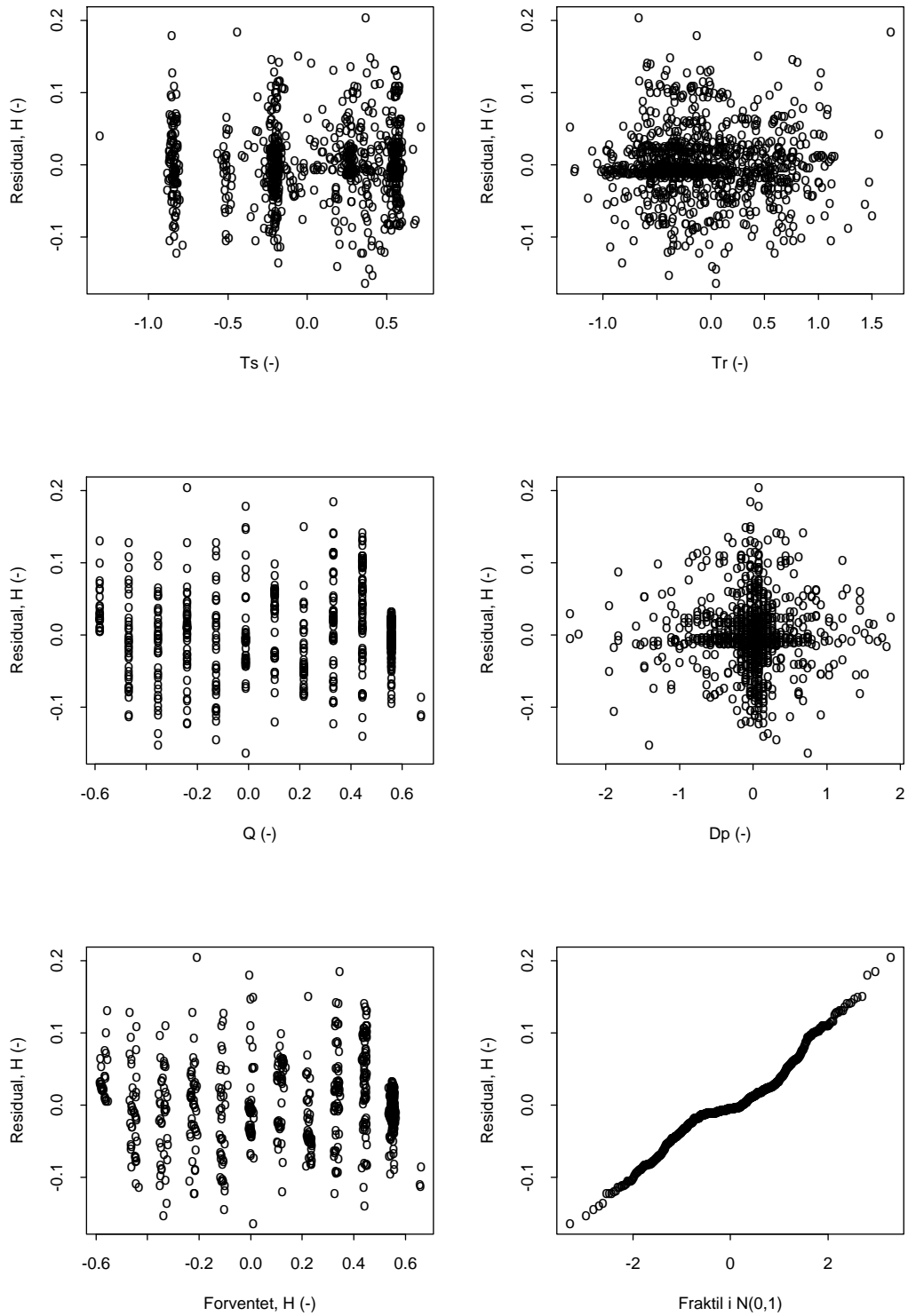
Figur 12. Residualer for brændselsforbrug mod residualer for elproduktion.



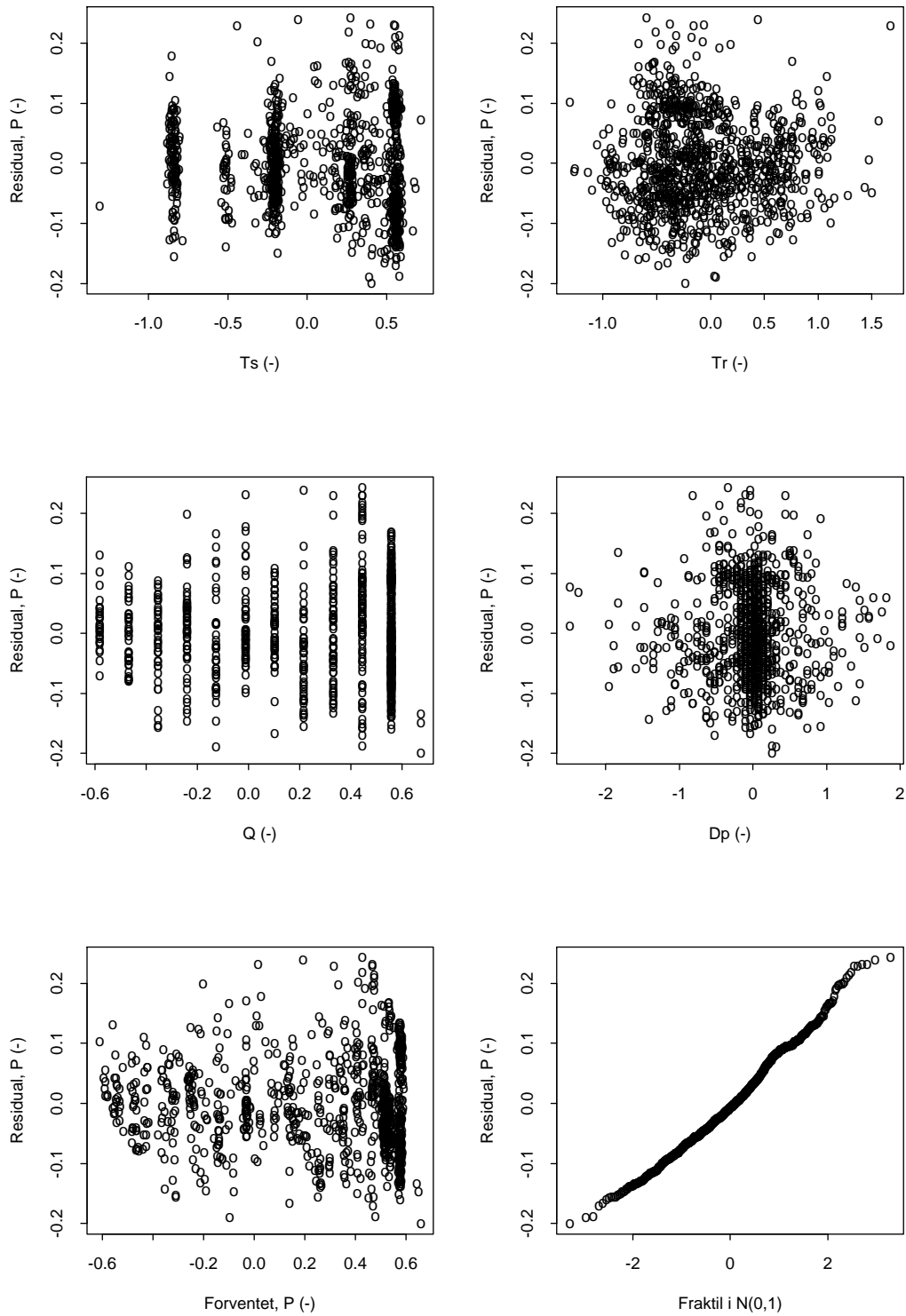
Figur 13. Residualer. Observationer svarende til atypiske residualer udeladt.



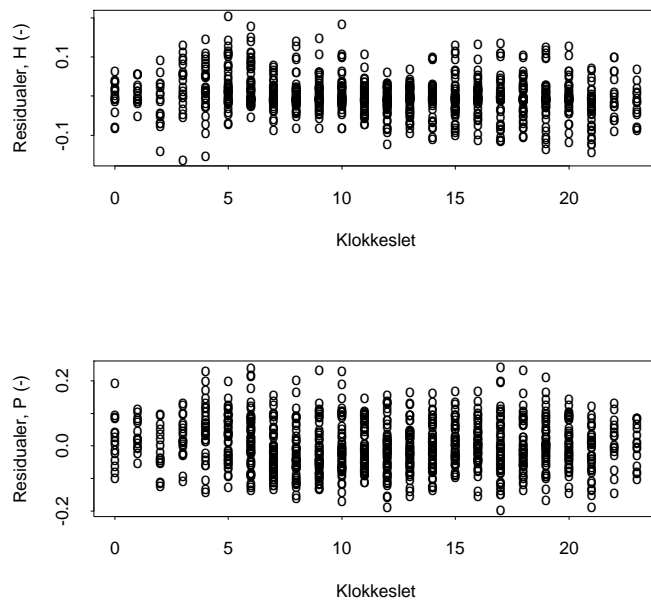
Figur 14. Residualer for brændselsforbrug mod residualer for elproduktion. Observationer svarende til atypiske residualer udeladt.



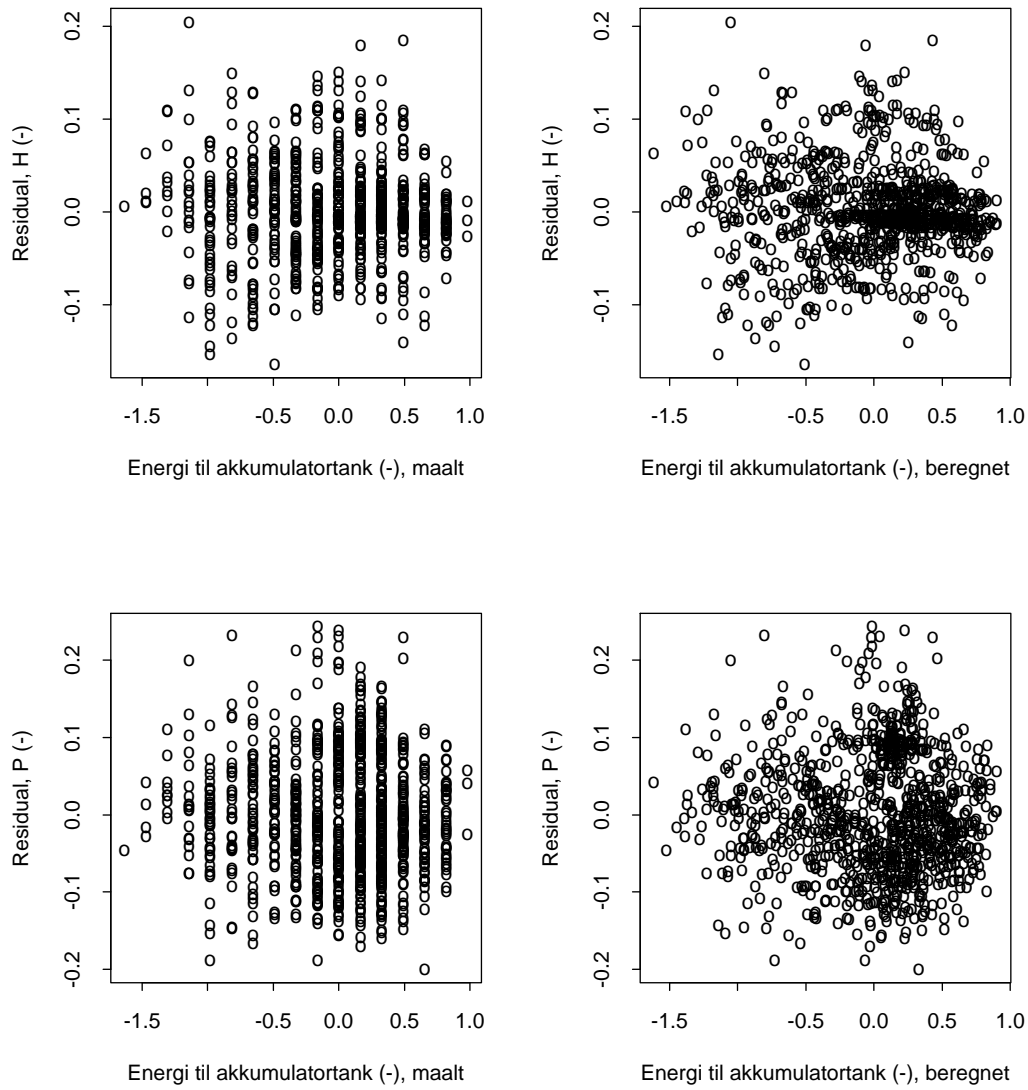
Figur 15. Standard diagnostiske plot for residualer for brændselsforbrug. Observationer svarende til atypiske værdier udeladt.



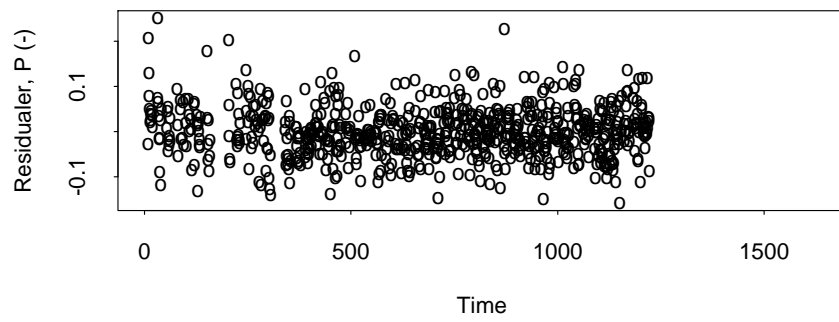
Figur 16. Standard diagnostiske plot for residualer for elproduktion. Observationer svarende til atypiske værdier udeladt.



Figur 17. Residualer mod tiden. Observationer svarende til atypiske residualer udeladt.



Figur 18. Residualer mod målt og beregnet varme leveret til akkumulatortank.



Figur 19. Residualer for elproduktion. Observationer svarende til atypiske residualer og timer efter 20/10 kl. 21 (nr. 1222) udeladt.

B Styringsdiagram

På den efterfølgende side ses en kopi af styringsdiagrammet for Norbrinken Kraftvarmeværk.

C Estimationsresultater

C.1 Indledende model for brændselsforbrug

```
> summary(rod.H.lm.full)
```

```
Call: lm(formula = H ~ 1 + Ts + Tr + Q + Dp + Ts^2 + Tr^2 + Q^2 + Dp^2 + I(Ts * Tr) +
  I(Ts * Q) + I(Ts * Dp) + I(Tr * Q) + I(Tr * Dp) + I(Q * Dp), data =
  rod.data, model = T, x = T, y = T)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.164	-0.01768	-0.004851	0.0202	0.2042

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.0124	0.0039	3.1526	0.0017
Ts	0.0144	0.0056	2.5929	0.0097
Tr	-0.0044	0.0044	-1.0003	0.3174
Q	0.9799	0.0055	179.2455	0.0000
Dp	-0.0006	0.0043	-0.1455	0.8843
I(Ts^2)	0.0060	0.0092	0.6540	0.5133
I(Tr^2)	0.0068	0.0061	1.1141	0.2655
I(Q^2)	-0.0404	0.0155	-2.5997	0.0095
I(Dp^2)	0.0000	0.0026	0.0141	0.9887
I(Ts * Tr)	0.0032	0.0108	0.2950	0.7681
I(Ts * Q)	-0.0001	0.0116	-0.0124	0.9901
I(Ts * Dp)	0.0067	0.0100	0.6719	0.5018
I(Tr * Q)	0.0160	0.0102	1.5739	0.1158
I(Tr * Dp)	-0.0052	0.0070	-0.7416	0.4585
I(Q * Dp)	-0.0104	0.0106	-0.9870	0.3239

Residual standard error: 0.0476 on 985 degrees of freedom

Multiple R-Squared: 0.9819

Correlation of Coefficients:

	(Intercept)	Ts	Tr	Q	Dp	I(Ts^2)	I(Tr^2)	I(Q^2)
Ts	-0.1764							
Tr	-0.0579	0.3886						
Q	0.1334	-0.2510	0.0417					
Dp	-0.0742	0.0143	0.0785	0.0280				
I(Ts^2)	-0.4685	0.4320	0.1180	-0.1393	-0.0163			
I(Tr^2)	-0.4285	0.2483	0.0157	-0.0179	0.0105	0.3587		
I(Q^2)	-0.6237	0.0626	0.0768	-0.5353	0.0172	0.0298	0.0549	
I(Dp^2)	-0.2400	-0.1147	0.0029	-0.0161	0.2770	-0.0160	-0.0829	0.1211
I(Ts * Tr)	-0.1673	0.1878	0.2859	0.0298	-0.0133	0.4948	0.5441	0.0246
I(Ts * Q)	0.0806	-0.6483	-0.2279	0.1909	-0.0582	-0.2757	0.0070	-0.1208
I(Ts * Dp)	-0.0113	0.0171	-0.0158	-0.0243	-0.3617	0.0665	0.0107	0.0088
I(Tr * Q)	-0.0037	-0.2344	-0.5319	0.0318	-0.1086	-0.1375	0.1203	0.0468
I(Tr * Dp)	-0.0016	-0.0263	0.0407	-0.0683	0.0788	-0.0129	-0.1073	0.0303
I(Q * Dp)	0.0443	-0.0648	-0.0655	-0.0262	-0.3599	-0.0477	-0.0792	0.0204

I(Dp^2) I(Ts * Tr) I(Ts * Q) I(Ts * Dp) I(Tr * Q) I(Tr * Dp)

```

      Ts
      Tr
      Q
      Dp
I(Ts^2)
I(Tr^2)
I(Q^2)
I(Dp^2)
I(Ts * Tr) -0.0981
I(Ts * Q)  0.0106 -0.0102
I(Ts * Dp) -0.1217  0.0086    0.0457
I(Tr * Q)  0.0124 -0.1128    0.4463    0.0818
I(Tr * Dp)  0.1683 -0.0458    0.0105    0.0264   -0.0460
I(Q * Dp)  0.0446 -0.0292    0.0342   -0.3390   -0.0009    0.3450

```

C.2 Endelig model for brændselsforbrug

```
> summary(rod.H.final)
```

```
Call: lm(formula = H ~ 1 + Ts + Q + Q^2, data = rod.data, model = T, x = T, y = T)
```

Residuals:

```

      Min       1Q   Median       3Q      Max
-0.1639 -0.01616 -0.004477  0.0199  0.212

```

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.0148	0.0030	4.9879	0.0000
Ts	0.0126	0.0034	3.7292	0.0002
Q	0.9807	0.0053	186.0635	0.0000
I(Q^2)	-0.0432	0.0150	-2.8770	0.0041

Residual standard error: 0.04747 on 996 degrees of freedom

Multiple R-Squared: 0.9818

Correlation of Coefficients:

	(Intercept)	Ts	Q
Ts	0.1045		
Q	0.0853	-0.2003	
I(Q^2)	-0.7631	-0.0885	-0.5455

C.3 Indledende model for elproduktion

```
> summary(rod.subset.P.lm.full)
```

```
Call: lm(formula = P ~ 1 + Ts + Tr + Q + Dp + Ts^2 + Tr^2 + Q^2 + Dp^2 + I(Ts * Tr) +
      I(Ts * Q) + I(Ts * Dp) + I(Tr * Q) + I(Tr * Dp) + I(Q * Dp), data =
      rod.data, subset = rod.data$time <= 1222, model = T, x = T, y = T)
```

Residuals:

```

      Min       1Q   Median       3Q      Max
-0.1588 -0.03366 -0.002819  0.03215  0.2515

```

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	-0.0271	0.0050	-5.4354	0.0000
Ts	0.0028	0.0074	0.3850	0.7004
Tr	-0.0012	0.0055	-0.2129	0.8315
Q	0.9449	0.0074	127.9709	0.0000
Dp	-0.0019	0.0053	-0.3514	0.7254
I(Ts ²)	0.0913	0.0114	7.9729	0.0000
I(Tr ²)	0.0128	0.0074	1.7424	0.0819
I(Q ²)	-0.0304	0.0201	-1.5159	0.1300
I(Dp ²)	-0.0007	0.0032	-0.2167	0.8285
I(Ts * Tr)	0.0081	0.0135	0.6011	0.5480
I(Ts * Q)	-0.0669	0.0153	-4.3595	0.0000
I(Ts * Dp)	0.0045	0.0125	0.3555	0.7223
I(Tr * Q)	0.0069	0.0124	0.5606	0.5753
I(Tr * Dp)	-0.0051	0.0086	-0.5920	0.5540
I(Q * Dp)	-0.0094	0.0141	-0.6634	0.5073

Residual standard error: 0.0561 on 725 degrees of freedom

Multiple R-Squared: 0.9747

Correlation of Coefficients:

	(Intercept)	Ts	Tr	Q	Dp	I(Ts ²)	I(Tr ²)	I(Q ²)
Ts	-0.0713							
Tr	-0.0928	0.3452						
Q	0.0690	-0.3194	0.0469					
Dp	-0.0657	0.0131	0.0656	0.0100				
I(Ts ²)	-0.4250	0.4691	0.1559	-0.1296	-0.0211			
I(Tr ²)	-0.4308	0.1821	0.0173	0.0028	0.0191	0.3367		
I(Q ²)	-0.6194	0.0793	0.1150	-0.5057	0.0062	0.0485	0.0495	
I(Dp ²)	-0.2524	-0.1542	0.0048	-0.0272	0.2839	-0.0442	-0.0659	0.1209
I(Ts * Tr)	-0.1946	0.1516	0.3525	0.0535	-0.0207	0.4865	0.5249	0.0581
I(Ts * Q)	-0.0141	-0.6069	-0.1581	0.3452	-0.0754	-0.2181	0.0243	-0.0968
I(Ts * Dp)	-0.0040	0.0456	-0.0154	-0.0323	-0.2718	0.0835	0.0156	0.0050
I(Tr * Q)	-0.0131	-0.1963	-0.4874	0.0430	-0.1058	-0.1018	0.1238	0.0469
I(Tr * Dp)	-0.0075	-0.0395	0.0585	-0.0621	0.0364	-0.0177	-0.1125	0.0390
I(Q * Dp)	0.0222	-0.0902	-0.0540	-0.0067	-0.4051	-0.0551	-0.0789	0.0341
	I(Dp ²)	I(Ts * Tr)	I(Ts * Q)	I(Ts * Dp)	I(Tr * Q)	I(Tr * Dp)		
Ts								
Tr								
Q								
Dp								
I(Ts ²)								
I(Tr ²)								
I(Q ²)								
I(Dp ²)								
I(Ts * Tr)	-0.0841							
I(Ts * Q)	-0.0046	0.0437						
I(Ts * Dp)	-0.1099	0.0111	0.0294					
I(Tr * Q)	0.0059	-0.0778	0.4296	0.0879				
I(Tr * Dp)	0.1689	-0.0233	0.0132	0.0095	-0.0486			
I(Q * Dp)	0.0477	-0.0137	0.0551	-0.3308	-0.0071	0.3766		

C.4 Endelig model for elproduktion

```
> summary(rod.subset.P.final)
```

```
Call: lm(formula = P ~ 1 + Q + Ts^2 + I(Ts * Q), data = rod.data, subset = rod.data$  
time <= 1222, model = T, x = T, y = T)
```

```
Residuals:
```

Min	1Q	Median	3Q	Max
-0.1594	-0.03444	-0.003705	0.03138	0.2547

```
Coefficients:
```

	Value	Std. Error	t value	Pr(> t)
(Intercept)	-0.0280	0.0032	-8.8951	0.0000
Q	0.9392	0.0057	165.5061	0.0000
I(Ts^2)	0.0854	0.0086	9.9480	0.0000
I(Ts * Q)	-0.0701	0.0103	-6.8066	0.0000

```
Residual standard error: 0.05593 on 736 degrees of freedom
```

```
Multiple R-Squared: 0.9745
```

```
Correlation of Coefficients:
```

	(Intercept)	Q	I(Ts^2)
Q	-0.4576		
I(Ts^2)	-0.6077	0.0198	
I(Ts * Q)	-0.1771	0.1139	0.0806

D Test i modellerne

Funktionerne `f.H.tests` og `f.P.tests` anvender blot simple kald af `f.F.test` (se appendiks E) og `summary`.

D.1 Brændselsforbrug

```
> f.H.tests()
=====
H0: Ts not in model vs. H1: Full model

      F-test of rod.H.lm.Ts.excl against rod.H.lm.full

Hypothesis (SS, df, MS): 0.02752899863267725 5 0.005505799726535449
Alternative (SS, df, MS): 2.232208714593244 985 0.002266201740703801

data: Examine objects above (eg. <obj>$call) to see which data is used
F-statistic = 2.4295, ndf = 5, ddf = 985, p-value = 0.0336

=====
H0: Tr not in model vs. H1: Full model

      F-test of rod.H.lm.Tr.excl against rod.H.lm.full

Hypothesis (SS, df, MS): 0.008495418669224066 5 0.001699083733844813
Alternative (SS, df, MS): 2.232208714593244 985 0.002266201740703801

data: Examine objects above (eg. <obj>$call) to see which data is used
F-statistic = 0.7497, ndf = 5, ddf = 985, p-value = 0.5863

=====
H0: Q not in model vs. H1: Full model

      F-test of rod.H.lm.Q.excl against rod.H.lm.full

Hypothesis (SS, df, MS): 102.9232682506972 5 20.58465365013945
Alternative (SS, df, MS): 2.232208714593244 985 0.002266201740703801

data: Examine objects above (eg. <obj>$call) to see which data is used
F-statistic = 9083.328, ndf = 5, ddf = 985, p-value = 0

=====
H0: Dp not in model vs. H1: Full model

      F-test of rod.H.lm.Dp.excl against rod.H.lm.full

Hypothesis (SS, df, MS): 0.003326328175839501 5 0.0006652656351679002
Alternative (SS, df, MS): 2.232208714593244 985 0.002266201740703801

data: Examine objects above (eg. <obj>$call) to see which data is used
F-statistic = 0.2936, ndf = 5, ddf = 985, p-value = 0.9166
```

```

=====
H0: Tr and Dp not in model vs. H1: Full model

      F-test of rod.H.lm.TrDp.excl against rod.H.lm.full

Hypothesis (SS, df, MS): 0.01030935810509765 9 0.001145484233899739
Alternative (SS, df, MS): 2.232208714593244 985 0.002266201740703801

data: Examine objects above (eg. <obj>$call) to see which data is used
F-statistic = 0.5055, ndf = 9, ddf = 985, p-value = 0.8713

=====
Test's of individual parameters when Tr and Dp is excluded

Call: lm(formula = H ~ 1 + Ts + Q + Ts^2 + Q^2 + I(Ts * Q), data = rod.data, model =
      T, x = T, y = T)
Residuals:
      Min       1Q   Median       3Q      Max
-0.1648 -0.01812 -0.004394  0.01994  0.2101

Coefficients:
              Value Std. Error  t value Pr(>|t|)
(Intercept)  0.0141   0.0033    4.2292  0.0000
           Ts  0.0154   0.0049    3.1692  0.0016
           Q  0.9797   0.0054  181.0133  0.0000
      I(Ts^2)  0.0043   0.0078    0.5522  0.5810
      I(Q^2) -0.0416   0.0152   -2.7291  0.0065
      I(Ts * Q) -0.0074   0.0102   -0.7197  0.4719

Residual standard error: 0.0475 on 994 degrees of freedom
Multiple R-Squared: 0.9818

Correlation of Coefficients:
      (Intercept)      Ts      Q I(Ts^2) I(Q^2)
Ts -0.0901
Q  0.1372      -0.2946
I(Ts^2) -0.4497      0.4247 -0.1619
I(Q^2) -0.6711      0.0426 -0.5533  0.0232
I(Ts * Q) 0.0741     -0.6775  0.1965 -0.2971 -0.1616

=====
Full model (rod.H.lm.full):
lm(formula = H ~ 1 + Ts + Tr + Q + Dp + Ts^2 + Tr^2 + Q^2 + Dp^2 + I(Ts * Tr) +
      I(Ts * Q) + I(Ts * Dp) + I(Tr * Q) + I(Tr * Dp) + I(Q * Dp), data =
      rod.data, model = T, x = T, y = T)

Ts excluded (rod.H.lm.Ts.excl):
lm(formula = H ~ 1 + Tr + Q + Dp + Tr^2 + Q^2 + Dp^2 + I(Tr * Q) + I(Tr * Dp) +
      I(Q * Dp), data = rod.data, model = T, x = T, y = T)

Tr excluded (rod.H.lm.Tr.excl):
lm(formula = H ~ 1 + Ts + Q + Dp + Ts^2 + Q^2 + Dp^2 + I(Ts * Q) + I(Ts * Dp) +
      I(Q * Dp), data = rod.data, model = T, x = T, y = T)

```

```

Q excluded (rod.H.lm.Q.excl):
lm(formula = H ~ 1 + Ts + Tr + Dp + Ts^2 + Tr^2 + Dp^2 + I(Ts * Tr) + I(Ts *
  Dp) + I(Tr * Dp), data = rod.data, model = T, x = T, y = T)

Dp excluded (rod.H.lm.Dp.excl):
lm(formula = H ~ 1 + Ts + Tr + Q + Ts^2 + Tr^2 + Q^2 + I(Ts * Tr) + I(Ts * Q) +
  I(Tr * Q), data = rod.data, model = T, x = T, y = T)

Tr and Dp excluded (rod.H.lm.TrDp.excl):
lm(formula = H ~ 1 + Ts + Q + Ts^2 + Q^2 + I(Ts * Q), data = rod.data, model =
  T, x = T, y = T)

NULL
>

```

D.2 Elproduktion

```

> f.P.tests()
=====
H0: Ts not in model vs. H1: Full model

      F-test of rod.subset.P.lm.Ts.excl against rod.subset.P.lm.full

Hypothesis (SS, df, MS): 0.3951118890731929 5 0.07902237781463858
Alternative (SS, df, MS): 2.281380349403275 725 0.00314673151641831

data: Examine objects above (eg. <obj>$call) to see which data is used
F-statistic = 25.1125, ndf = 5, ddf = 725, p-value = 0

=====
H0: Tr not in model vs. H1: Full model

      F-test of rod.subset.P.lm.Tr.excl against rod.subset.P.lm.full

Hypothesis (SS, df, MS): 0.01069519234404732 5 0.002139038468809464
Alternative (SS, df, MS): 2.281380349403275 725 0.00314673151641831

data: Examine objects above (eg. <obj>$call) to see which data is used
F-statistic = 0.6798, ndf = 5, ddf = 725, p-value = 0.6389

=====
H0: Q not in model vs. H1: Full model

      F-test of rod.subset.P.lm.Q.excl against rod.subset.P.lm.full

Hypothesis (SS, df, MS): 80.2783168960778 5 16.05566337921556
Alternative (SS, df, MS): 2.281380349403275 725 0.00314673151641831

data: Examine objects above (eg. <obj>$call) to see which data is used
F-statistic = 5102.33, ndf = 5, ddf = 725, p-value = 0

=====

```

H0: Dp not in model vs. H1: Full model

F-test of rod.subset.P.lm.Dp.excl against rod.subset.P.lm.full

Hypothesis (SS, df, MS): 0.003065319156690897 5 0.0006130638313381795

Alternative (SS, df, MS): 2.281380349403275 725 0.00314673151641831

data: Examine objects above (eg. <obj>\$call) to see which data is used

F-statistic = 0.1948, ndf = 5, ddf = 725, p-value = 0.9645

=====

H0: Tr and Dp not in model vs. H1: Full model

F-test of rod.subset.P.lm.TrDp.excl against rod.subset.P.lm.full

Hypothesis (SS, df, MS): 0.01249056373152196 9 0.001387840414613551

Alternative (SS, df, MS): 2.281380349403275 725 0.00314673151641831

data: Examine objects above (eg. <obj>\$call) to see which data is used

F-statistic = 0.441, ndf = 9, ddf = 725, p-value = 0.9129

=====

Test's of individual parameters when Tr and Dp is excluded

Call: lm(formula = P ~ 1 + Ts + Q + Ts^2 + Q^2 + I(Ts * Q), data = rod.data, subset = rod.data\$time <= 1222, model = T, x = T, y = T)

Residuals:

Min	1Q	Median	3Q	Max
-0.1606	-0.03395	-0.003338	0.03164	0.2505

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	-0.0234	0.0042	-5.5505	0.0000
Ts	0.0008	0.0066	0.1152	0.9084
Q	0.9449	0.0073	129.4988	0.0000
I(Ts^2)	0.0859	0.0098	8.7706	0.0000
I(Q^2)	-0.0316	0.0196	-1.6166	0.1064
I(Ts * Q)	-0.0689	0.0137	-5.0392	0.0000

Residual standard error: 0.0559 on 734 degrees of freedom

Multiple R-Squared: 0.9746

Correlation of Coefficients:

	(Intercept)	Ts	Q	I(Ts^2)	I(Q^2)
Ts	0.0016				
Q	0.0732	-0.3701			
I(Ts^2)	-0.3979	0.4821	-0.1663		
I(Q^2)	-0.6649	0.0639	-0.5313	0.0334	
I(Ts * Q)	-0.0319	-0.6492	0.3594	-0.2600	-0.1443

=====

Excluding only the linear term 'Ts' (the p-value for the other non-significant term is not so very large:

Call: lm(formula = P ~ 1 + Q + Ts^2 + Q^2 + I(Ts * Q), data = rod.data, subset =

```
rod.data$time <= 1222, model = T, x = T, y = T)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.1601	-0.03387	-0.003272	0.0316	0.2504

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	-0.0234	0.0042	-5.5544	0.0000
Q	0.9452	0.0068	139.5352	0.0000
I(Ts^2)	0.0854	0.0086	9.9542	0.0000
I(Q^2)	-0.0318	0.0195	-1.6283	0.1039
I(Ts * Q)	-0.0678	0.0104	-6.5309	0.0000

Residual standard error: 0.05587 on 735 degrees of freedom

Multiple R-Squared: 0.9746

Correlation of Coefficients:

	(Intercept)	Q	I(Ts^2)	I(Q^2)
Q	0.0794			
I(Ts^2)	-0.4551	0.0149		
I(Q^2)	-0.6664	-0.5476	0.0030	
I(Ts * Q)	-0.0405	0.1686	0.0795	-0.1354

Full model (rod.subset.P.lm.full):

```
lm(formula = P ~ 1 + Ts + Tr + Q + Dp + Ts^2 + Tr^2 + Q^2 + Dp^2 + I(Ts * Tr) +  
I(Ts * Q) + I(Ts * Dp) + I(Tr * Q) + I(Tr * Dp) + I(Q * Dp), data =  
rod.data, subset = rod.data$time <= 1222, model = T, x = T, y = T)
```

Ts excluded (rod.subset.P.lm.Ts.excl):

```
lm(formula = P ~ 1 + Tr + Q + Dp + Tr^2 + Q^2 + Dp^2 + I(Tr * Q) + I(Tr * Dp) +  
I(Q * Dp), data = rod.data, subset = rod.data$time <= 1222, model = T,  
x = T, y = T)
```

Tr excluded (rod.subset.P.lm.Tr.excl):

```
lm(formula = P ~ 1 + Ts + Q + Dp + Ts^2 + Q^2 + Dp^2 + I(Ts * Q) + I(Ts * Dp) +  
I(Q * Dp), data = rod.data, subset = rod.data$time <= 1222, model = T,  
x = T, y = T)
```

Q excluded (rod.subset.P.lm.Q.excl):

```
lm(formula = P ~ 1 + Ts + Tr + Dp + Ts^2 + Tr^2 + Dp^2 + I(Ts * Tr) + I(Ts *  
Dp) + I(Tr * Dp), data = rod.data, subset = rod.data$time <= 1222,  
model = T, x = T, y = T)
```

Dp excluded (rod.subset.P.lm.Dp.excl):

```
lm(formula = P ~ 1 + Ts + Tr + Q + Ts^2 + Tr^2 + Q^2 + I(Ts * Tr) + I(Ts * Q) +  
I(Tr * Q), data = rod.data, subset = rod.data$time <= 1222, model = T,  
x = T, y = T)
```

Tr and Dp excluded (rod.subset.P.lm.TrDp.excl):

```
lm(formula = P ~ 1 + Ts + Q + Ts^2 + Q^2 + I(Ts * Q), data = rod.data, subset  
= rod.data$time <= 1222, model = T, x = T, y = T)
```

NULL

>

E Splus-funktioner

I dette appendiks vises udskrifter af de funktioner der er skrevet specielt for anvendelse i dette projekt.

funktionen `f.sign.shift`

Denne funktion tester om sandsynligheden for fortegnsskift i sekvensen af residualer fra et objekt er lig $\frac{1}{2}$.

```
> f.sign.shift
function(obj, zero = "+1")
{
  ##Testing the No. of signshifts in object 'obj'...
  sign.x <- sign(as.ts(residuals(obj)))
  if(zero == "+1")
    sign.x[sign.x == 0] <- 1
  if(zero == "-1")
    sign.x[sign.x == 0] <- -1
  shift <- sign.x != lag(sign.x)
  shifts <- length(shift[shift == T])
  N <- length(residuals(obj))
  return(binom.test(x = shifts, n = N - 1, p = 0.5, alternative =
    "two.sided"))
}
```

funktionen `f.F.test`

Denne funktion tester om objektet `hypothesis` er en passende model mod alternativet at objektet `alternative` er en passende model. Modellen `hypothesis` skal altså være indeholdt i `alternative`.

```
> f.F.test
function(hypothesis, alternative)
{
  ## hypothesis = object of class 'lm' representing the hypothesis to
  ## be tested.
  ## alternative = object of class 'lm' representing the alternative
  ## to the hypothesis.
  fam.hyp <- paste(family(hypothesis)$family)
  fam.alt <- paste(family(alternative)$family)
  if(fam.hyp[1] != "Gaussian" || fam.hyp[2] != "Identity: mu" || fam.hyp[
    3] != "Constant: 1")
    stop("Hypothesis must be of Gaussian family")
  if(fam.alt[1] != "Gaussian" || fam.alt[2] != "Identity: mu" || fam.alt[
    3] != "Constant: 1")
    stop("Alternative must be of Gaussian family")
  SS.y.alt <- sum(residuals(alternative)^2)
  SS.alt.hyp <- sum((fitted.values(alternative) - fitted.values(
    hypothesis))^2)
```

```

ddf <- length(fitted.values(alternative)) - alternative$rank
ndf <- alternative$rank - hypothesis$rank
MS.y.alt <- SS.y.alt/ddf
MS.alt.hyp <- SS.alt.hyp/ndf
F.stat <- MS.alt.hyp/MS.y.alt
p.value <- 1 - pf(F.stat, ndf, ddf)
method <- paste("F-test of", deparse(substitute(hypothesis)), "against",
               deparse(substitute(alternative)),
               "\n\nHypothesis (SS, df, MS):", SS.alt.hyp, ndf, MS.alt.hyp,
               "\n\nAlternative (SS, df, MS):", SS.y.alt, ddf, MS.y.alt)
df <- c(ndf, ddf)
names(df) <- c("ndf", "ddf")
names(F.stat) <- "F-statistic"
data.name <-
  "Examine objects above (eg. <obj>$call) to see which data is used"

test.result <- list(statistic = F.stat, parameters = df, p.value =
  p.value, method = method, data.name = data.name)
attr(test.result, "class") <- "htest"
return(test.result)    ##
## The calculations above could be simplified by use of the
## generalized linear model framework (the formulation above
## corresponds to the general linear model):
##
## SS.alt.hyp = deviance(hypothesis) - deviance(alternative)
## SS.y.alt   = deviance(alternative)
##
}
```

F Dokumentation af beregninger

Nedenfor er et såkaldt Splus “Backtrack” af beregningerne vist. Dette viser de udførte beregninger i omvendt rækkefølge.

De indlæste data er placeret i objektet `korsor` som `korsor$data`.

Der anvendes nogle funktioner til håndtering af datoer (f.eks. `as.date`). Disse er ikke en del af standard S-Plus 3.0. De anvendte funktioner er skrevet af Terry Therneau (e-post: `therneau@mayo.edu`) og hentet fra “S index” under StatLib (`lib.stat.cmu.edu`).

```

snake2:/ussnk2/han/proj/efp94 $ Splus AUDIT
Reading audit file ".Data/.Audit"
audit: B to.be.documented
./S.Korsor/to.be.documented
3453: to.be.documented <- list(r.P.lm.full = r.P.lm.full, r.H.lm.full = r.H.lm.full,
  rod.P.lm.full = rod.P.lm.full, rod.H.lm.full = rod.H.lm.full,
  rod.subset.P.lm.full = rod.subset.P.lm.full, rod.H.lm.Ts.excl =
  rod.H.lm.Ts.excl, rod.H.lm.Tr.excl = rod.H.lm.Tr.excl, rod.H.lm.Q.excl
```



```

      = rod.H.lm.Q.excl, rod.H.lm.Dp.excl = rod.H.lm.Dp.excl,
      rod.H.lm.TrDp.excl = rod.H.lm.TrDp.excl, rod.H.final = rod.H.final,
      rod.subset.P.lm.Ts.excl = rod.subset.P.lm.Ts.excl,
      rod.subset.P.lm.Tr.excl = rod.subset.P.lm.Tr.excl,
      rod.subset.P.lm.Q.excl = rod.subset.P.lm.Q.excl,
      rod.subset.P.lm.Dp.excl = rod.subset.P.lm.Dp.excl,
      rod.subset.P.lm.TrDp.excl = rod.subset.P.lm.TrDp.excl,
      rod.subset.P.final = rod.subset.P.final)
3024: rod.subset.P.final <- lm(formula = P ~ 1 + Q + Ts^2 + I(Ts * Q), data =
      rod.data, subset = rod.data$time <= 1222, model = T, x = T, y = T)
2995: rod.subset.P.lm.TrDp.excl <- lm(formula = P ~ 1 + Ts + Q + Ts^2 + Q^2 + I(
      Ts * Q), data = rod.data, subset = rod.data$time <= 1222, model = T,
      x = T, y = T)
2990: rod.subset.P.lm.Dp.excl <- lm(formula = P ~ 1 + Ts + Tr + Q + Ts^2 + Tr^2 +
      Q^2 + I(Ts * Tr) + I(Ts * Q) + I(Tr * Q), data = rod.data, subset =
      rod.data$time <= 1222, model = T, x = T, y = T)
2989: rod.subset.P.lm.Q.excl <- lm(formula = P ~ 1 + Ts + Tr + Dp + Ts^2 + Tr^2 +
      Dp^2 + I(Ts * Tr) + I(Ts * Dp) + I(Tr * Dp), data = rod.data, subset =
      rod.data$time <= 1222, model = T, x = T, y = T)
2988: rod.subset.P.lm.Tr.excl <- lm(formula = P ~ 1 + Ts + Q + Dp + Ts^2 + Q^2 + Dp^
      2 + I(Ts * Q) + I(Ts * Dp) + I(Q * Dp), data = rod.data, subset =
      rod.data$time <= 1222, model = T, x = T, y = T)
2987: rod.subset.P.lm.Ts.excl <- lm(formula = P ~ 1 + Tr + Q + Dp + Tr^2 + Q^2 + Dp^
      2 + I(Tr * Q) + I(Tr * Dp) + I(Q * Dp), data = rod.data, subset =
      rod.data$time <= 1222, model = T, x = T, y = T)
2801: rod.H.final <- lm(formula = H ~ 1 + Ts + Q + Q^2, data = rod.data, model = T,
      x = T, y = T)
2795: rod.H.lm.TrDp.excl <- lm(formula = H ~ 1 + Ts + Q + Ts^2 + Q^2 + I(Ts * Q),
      data = rod.data, model = T, x = T, y = T)
2780: rod.H.lm.Dp.excl <- lm(formula = H ~ 1 + Ts + Tr + Q + Ts^2 + Tr^2 + Q^2 + I(
      Ts * Tr) + I(Ts * Q) + I(Tr * Q), data = rod.data, model = T, x = T,
      y = T)
2779: rod.H.lm.Q.excl <- lm(formula = H ~ 1 + Ts + Tr + Dp + Ts^2 + Tr^2 + Dp^2 +
      I(Ts * Tr) + I(Ts * Dp) + I(Tr * Dp), data = rod.data, model = T, x = T,
      y = T)
2776: rod.H.lm.Tr.excl <- lm(formula = H ~ 1 + Ts + Q + Dp + Ts^2 + Q^2 + Dp^2 + I(
      Ts * Q) + I(Ts * Dp) + I(Q * Dp), data = rod.data, model = T, x = T,
      y = T)
2669: rod.H.lm.Ts.excl <- lm(formula = H ~ 1 + Tr + Q + Dp + Tr^2 + Q^2 + Dp^2 + I(
      Tr * Q) + I(Tr * Dp) + I(Q * Dp), data = rod.data, model = T, x = T,
      y = T)
2638: rod.subset.P.lm.full <- lm(formula = P ~ 1 + Ts + Tr + Q + Dp + Ts^2 + Tr^
      2 + Q^2 + Dp^2 + I(Ts * Tr) + I(Ts * Q) + I(Ts * Dp) + I(Tr * Q) + I(
      Tr * Dp) + I(Q * Dp), data = rod.data, model = T, x = T, y = T, subset
      = rod.data$time <= 1222)
1841: rod.P.lm.full <- lm(formula = P ~ 1 + Ts + Tr + Q + Dp + Ts^2 + Tr^2 + Q^2 +
      Dp^2 + I(Ts * Tr) + I(Ts * Q) + I(Ts * Dp) + I(Tr * Q) + I(Tr * Dp) +
      I(Q * Dp), data = rod.data, model = T, x = T, y = T)
1840: rod.H.lm.full <- lm(formula = H ~ 1 + Ts + Tr + Q + Dp + Ts^2 + Tr^2 + Q^2 +
      Dp^2 + I(Ts * Tr) + I(Ts * Q) + I(Ts * Dp) + I(Tr * Q) + I(Tr * Dp) +
      I(Q * Dp), data = rod.data, model = T, x = T, y = T)
1826: rod.data <- as.data.frame(list(H = rod.H, P = rod.P, Ts = rod.Ts, Tr = rod.Tr,
      Q = rod.Q, Dp = rod.Dp, time = rod.time))
1819: rod.time <- r.data$time[!tmp]

```

```

1817: rod.Dp <- r.data$Dp[!tmp]
1816: rod.Q <- r.data$Q[!tmp]
1815: rod.Tr <- r.data$Tr[!tmp]
1814: rod.Ts <- r.data$Ts[!tmp]
1813: rod.P <- r.data$P[!tmp]
1812: rod.H <- r.data$H[!tmp]
1811: tmp <- r.data$time == 490 | r.data$time == 491 | r.data$time == 879
1651: r.P.lm.full <- lm(formula = P ~ 1 + Ts + Tr + Q + Dp + Ts^2 + Tr^2 + Q^2 + Dp^
  2 + I(Ts * Tr) + I(Ts * Q) + I(Ts * Dp) + I(Tr * Q) + I(Tr * Dp) + I(
  Q * Dp), data = r.data, model = T, x = T, y = T)
1649: r.H.lm.full <- lm(formula = H ~ 1 + Ts + Tr + Q + Dp + Ts^2 + Tr^2 + Q^2 + Dp^
  2 + I(Ts * Tr) + I(Ts * Q) + I(Ts * Dp) + I(Tr * Q) + I(Tr * Dp) + I(
  Q * Dp), data = r.data, model = T, x = T, y = T)
1630: r.data <- as.data.frame(list(H = r.H, P = r.P, Ts = r.Ts, Tr = r.Tr, Q = r.Q,
  Dp = r.Dp, time = r.time))
1615: r.time <- time$run[run == 1 & !is.na(H + P + Ts + Tr + Q + Dp)]
1613: r.Dp <- Dp[run == 1 & !is.na(H + P + Ts + Tr + Q + Dp)]
1612: r.Q <- Q[run == 1 & !is.na(H + P + Ts + Tr + Q + Dp)]
1611: r.Tr <- Tr[run == 1 & !is.na(H + P + Ts + Tr + Q + Dp)]
1610: r.Ts <- Ts[run == 1 & !is.na(H + P + Ts + Tr + Q + Dp)]
1609: r.P <- P[run == 1 & !is.na(H + P + Ts + Tr + Q + Dp)]
1607: r.H <- H[run == 1 & !is.na(H + P + Ts + Tr + Q + Dp)]
1579: run[run == 0 | is.na(run)] <- 0
1578: run[run > 0] <- 1
1577: run <- meas$MoGasCon
1015: Dp <- (Dp - values.std$mean$Dp)/(2 * sqrt(values.std$var$Dp))
1014: H <- (H - values.std$mean$H)/(2 * sqrt(values.std$var$H))
1013: P <- (P - values.std$mean$P)/(2 * sqrt(values.std$var$P))
1012: Q <- (Q - values.std$mean$Q)/(2 * sqrt(values.std$var$Q))
1011: Tr <- (Tr - values.std$mean$Tr)/(2 * sqrt(values.std$var$Tr))
1007: Ts <- (Ts - values.std$mean$Ts)/(2 * sqrt(values.std$var$Ts))
997: values.std <- list(description = "Mean and Varians after deletion of outliers",
  mean = means, var = vars)
990: vars <- list(Ts = var(Ts[!is.na(Ts)]), Tr = var(Tr[!is.na(Tr)]), Dp = var(
  Dp[!is.na(Dp)]), Q = var(Q[!is.na(Q)]), P = var(P[!is.na(P)]), H = var(
  H[!is.na(H)]))
988: means <- list(Ts = mean(Ts[!is.na(Ts)]), Tr = mean(Tr[!is.na(Tr)]), Dp = mean(
  Dp[!is.na(Dp)]), Q = mean(Q[!is.na(Q)]), P = mean(P[!is.na(P)]), H =
  mean(H[!is.na(H)]))
984: Q <- meas$MoHeatProd
982: H <- meas$MoGasCon
981: P <- meas$MoPowProd
980: Dp <- meas$DiffPres
979: Tr <- meas$Tr
978: Ts <- meas$Ts
958: meas$BoPowCon.NB[c(202, 1160, 1189)] <- NA
951: meas$MoPowCon[c(202, 1160, 1189)] <- NA
945: meas$BoGasCon.NB[1189] <- NA
941: meas$MoGasCon[c(202, 1160, 1189)] <- NA
937: meas$MoPowProd[c(202, 1160, 1189)] <- NA
933: meas$BoHeatProd.NB[1189] <- NA
920: meas$MoHeatProd[c(202, 1160)] <- NA
916: meas$MoHeatProd[1189] <- NA
912: meas$DiffPres[659] <- NA

```

```

909: meas$DiffPres[c(558, 657, 660, 658)] <- NA
884: meas <- as.data.frame(list(Ts = kursor$data$Ts, Tr = kursor$data$Tr, DiffPres
    = kursor$data$DiffPres, MoHeatProd = kursor$data$MoHeatProd,
    BoHeatProd.NB = kursor$data$BoHeatProd.NB, MoPowProd = kursor$data$
    MoPowProd, MoGasCon = kursor$data$MoGasCon, BoGasCon.NB = kursor$data$
    BoGasCon.NB, MoPowCon = kursor$data$MoPowCon, BoPowCon.NB = kursor$
    data$BoPowCon.NB))
838: time$run <- time$run - time$run[1] + 1
836: time <- as.data.frame(list(run = 24 * as.numeric(j.date) + kursor$data$hour +
    kursor$data$minute/60, date = date.ddmmyy(j.date)))
833: j.date <- as.date(paste(kursor$data$day, kursor$data$month, kursor$data$year,
    sep = "-"), order = "dmy")
audit:

```

Referencer

- [Conradsen, 1984] Knut Conradsen. *En introduktion til statistik*, bind 2A, 4. udgave, Institut for Matematisk Statistik og Operationsanalyse, Danmarks Tekniske Højskole, Lyngby 1984.
- [Hellesen & Tuxen, 1994] Jette K. Hellesen og Ole Tuxen. *Varme - Korsør kommunale varmforsynings 25 års jubilæum 1994*, Korsør kommunale værker, 1994.
- [Korsør Kommunale Værker, 1993] Korsør kommunale Værker. *Orientering om kraftvarmeværk i Korsør*, Korsør kommunale Værker, januar 1993.
- [Madsen, 1989] Henrik Madsen. *Tidsrækkeanalyse*, Institut for Matematisk Statistik og Operationsanalyse, Danmarks Tekniske Højskole, Lyngby 1989.