

Förord:

Denna uppsats har tillkommit sedan uppsatsförfattarna blivit kontaktade av Electrolux med en förfrågan om att undersöka statistikmodulen i deras nyimplementerade affärssystem. Vi vill därför tacka vår handledare på Electrolux, Magnus Cohen, för att ha gett oss möjligheten att arbeta med deras tidsserier, samt för att ha tillhandahållit datorer och ordnat nödvändig utbildning i SAP-APO-DP. Det har varit spännande att ges en roll i analysen av vilka möjligheter som statistiska metoder kan ha i att förbättra prognoser på Electrolux. Mats Nielsen, Yoon Kong Lee och Klas Lindgren är annan personal på Electrolux som har gett oss omedelbart och nödvändigt stöd. Vi tackar även Pär Stockhammar som, utan att ha varit anlitad för detta, har bidragit med viktig vägledning i metodologiska frågor.

Sammanfattning:

Statistiska hjälpmedel blir allt vanligare i tillverkningsföretags planeringsprocess. Syftet med den här uppsatsen är att testa ett antal statistiska metoders förmåga att prognostisera försäljning och därmed efterfrågan på Electrolux produkter. Studien utfördes på månatlig försäljningsdata mellan januari 2004 och juni 2010 för fem olika produktgrupper, vars försäljningsmönster skiljdes sig åt. De valdes ut som testdata för att ge ett så brett underlag som möjligt. På dessa serier testades sedan fyra olika modeller och utvärderades både med ex-post-prognos och med halveringsmetod på de 24 sista observationerna. Olika modeller presterade bäst beroende på tidsavstånd och serie. Mest robusta resultat påvisade Holt-Winters modeller. Resultaten pekar på att noggrannhet i anpassandet och val av modell är avgörande för prognosresultaten.

Abstract:

Statistical tools become more and more common in manufacturing companies planning processes. The aim of this study is to test a number of statistical methods ability to forecast sales and therefore demand on Electrolux products. The study was conducted on monthly sales data between January 2004 and June 2010 for five different product groups, with different sales patterns. They were chosen in order to maximize the scope of the tests. On the series four different models were tested and evaluated on ex-post forecasts as well as pseudo out-of-sample forecasts on the last 24 months. Different models performed best depending on series and distance in time. The most robust results were obtained by the Holt-Winters models. The results imply that choosing and fitting the appropriate model is crucial for good forecasts.

1 Inledning	3
1.1 Bakgrund.....	3
1.2 Syfte.....	4
1.3 Avgränsningar.....	4
1.4 Metod	4
1.5 Uppsatsens avsnitt	5
2. Metoder inom tidsserieanalys	5
2.1 Centrala begrepp i tidsserieanalys.	5
2.2 Holts metod (dubbel exponentiell utjämnning).....	7
2.3 Holt-Winters metoder.....	8
2.3.1 Holt-Winters additiva metod.....	8
2.3.2 Holt-Winters multiplikativa modell.....	9
2.4 ARIMA.....	10
2.5 Utvärderingsmetoder	13
2.5.1 Halveringsmetod	13
2.5.2 Jämförelsetal	14
2.5.3 Informationskriterier.	15
2.5.4 Diebold-Marianos test.....	16
3. Datamaterial	18
3.1 Val av tidsserier	18
3.2 Jämförelse med ARIMA	19
3.3 Val av mjukvara	19
4. Resultat	19
4.1 Tillämpning av exponentiella utjämningsmetoder.....	20
4.2 ARIMA resultat och exempel	20
4.3 Sammanfattande jämförelse.....	21
4.3.1 Jämförelsetal från ex-post-prognos	21
4.3.2 Diebold-Marianos test.....	22
4.3.3 Jämförelsetal med halveringsmetod.....	23
5. Slutsatser och rekommendationer	24
5.1 Slutsatser.....	24
6. Diskussion	25
6.1 Analys.....	25
6.1.1 Parameterskattningarna för de exponentiella utjämningsmetoderna.	25
6.1.2 Skillnad mellan anpassningsförmåga och prognostisering.	26
6.1.3 Teoretiska förväntningar och utfall.....	27
6.2 Brister i studien.....	28
6.3 Vidare analysmöjligheter – diskussion	28
7. Referenser	30
Appendix 1	32
Appendix 2	32
Appendix 3	33
Appendix 4	33

1 Inledning

1.1 Bakgrund

Det finns ett naturligt behov av tillförlitliga prognoser för alla företag som agerar på en marknad. En god prognos kan effektivisera lagerhantering och även vara ett hjälpmedel i strategisk planering av kampanjer. På Electrolux sker för närvarande denna prognostisering helt utan hjälp av statistiska metoder. På enheten Demand Planning arbetar ett team av väl insatta experter (så kallade demand planners) som, i samråd med säljare med kundkontakt, gör en subjektiv bedömning över kommande försäljning. Denna prognos justeras sedan i samrådsmöten med ledning och formaliseras i en prognos kallad Final Agreed Demand Plan. Prognosen utgörs av punkttestimationer över kommande försäljning. Dessa prognoser görs både för aggregerad försäljning och för produktgrupper, samt subgrupper inom produktgrupperna. De uppdateras och utvärderas löpande. Eftersom säljorganisationerna är olika från land till land, så skiljer rågången för demand planners sig åt. Det innebär att systemet är mycket personberoende och svårt att standardisera.

I april 2010 implementerades ett nytt affärssystem i Electrolux Norden, SAP. Detta ingår i en större process där Electrolux avser att börja nyttja SAP globalt. I detta program finns en modul för statistiska prognosberäkningar, Advanced Planner and Optimizer - Demand Planning (APO-DP). Där kan användaren producera olika prognoser och det finns vissa möjligheter till utvärdering. Styrkan och svagheten hos APO-DP är att en oerfaren användare enkelt kan producera prognoser oavsett om denne inte förstår eller kan utvärdera dem. Utan statistiska kunskaper blir det alltså svårt att avgöra vilken modell som bör användas. Det blir svårt att förstå hur programmet gör sina automatiska modellval samt konsekvenserna av detta.

Uppsatsförfattarna blev kontaktade under sommaren av den person som är ansvarig för den globala implementeringen av SAP mot bakgrund av att Electrolux i dagsläget inte kunnat utnyttja de tillgängliga statistiska metoderna på ett tillfredsställande sätt. Uppsatsförfattarna har, med tillgång till de data som Electrolux har lagrat, formulerat ett syfte som kombinerar Electrolux behov med de akademiska kraven för en kandidatuppsats i statistik.

I Electrolux modell så är tanken att statistiska metoder skall användas för att uppskatta marknadens obegränsade efterfrågan (Electrolux term är 'unconstrained demand') på Electrolux produkter, för att mer långsiktigt planera tillverkningen utan hänsyn till vad Electrolux kan leverera och opåverkat av eventuella kampanjer. Därför ska helst prognoserna utföras på data som korrigerats för försäljning som resulterat från kampanjer och försäljning som uteblivit på grund av produktions- och logistikproblem (så kallade stock-outs). Detta eftersom kampanjer styrs av Electrolux själva och att historiska stock-outs inte kan ligga till grund för framtida produktionsbeslut.

1.2 Syfte

Syftet med uppsatsen är att beskriva fyra olika modeller för tidsserieanalys och prognoser samt, med hjälp av olika lämpliga statistiska utvärderingsinstrument, analysera dessa. De fyra modellerna som skall studeras är Holts metod, Holts-Winters additiva och multiplikativa metoder samt ARIMA-metoder.

Med hjälp av vår analys vill vi ge Electrolux rekommendationer om hur de ska tillämpa de statistiska metoder som finns tillgängliga för dem via SAP.

1.3 Avgränsningar

Uppsatsens avgränsningar är aktuella för tre områden; val av datamaterial, val av metoder samt val av programvara. Datamaterialet kommer att väljas på subgruppsnivå i Electrolux produkthierarki. Tidsserierna väljs efter principen om spridning. Detta innebär att tidsserier som liknar varandra vad det gäller de tre faktorerna trend, säsong och varians inte kommer att analyseras var för sig. Istället kommer tidsserier som skiljer sig åt på ovan punkter att väljas. Detta kommer att ge en god grund för att utvärdera styrkor och svagheter med de valda metoderna för olika datamaterial.

All analys kommer vara univariat då vi utgår från det data som Electrolux för närvarande har tillgängligt för SAP APO-DP. Vår utgångspunkt är de metoder som är tillgängliga i SAP och vi lade därefter till ARIMA-modeller för att öka förståelsen av datamaterialet. Ett annat skäl till att vi valde att analysera även ARIMA-modellers prestationer är att denna grupp av modeller är väldigt vanliga och populära men saknas i SAP APO-DP. Detta kan alltså visa om de tillgängliga metoderna är tillräckliga eller om Electrolux bör arbeta utanför sin modul.

1.4 Metod

Prestationen av de utvalda metoderna kommer att utvärderas med hjälp av lämpliga jämförelsetal. Vi undersöker ex-post-prognoser, det vill säga prognoser på tidsavstånd ett under anpassningstiden, samt med halveringsmetod, där datamaterialet delas upp och anpassas på den ena delen och prognostiserar och prövas mot den andra delen. Dessa kommer sedan att styra rekommendationer av vilka analysmetoder som ska väljas för vilka tidsserier.

1.5 Uppsatsens avsnitt

I avsnitt 2 går vi igenom de metoder som kommer att tillämpas i uppsatsen samt ett antal grundläggande begrepp som är nödvändiga att känna till inom tidsserieanalys. Vi redovisar även de jämförelsetal vi kommer att använda för att utvärdera de olika metoderna samt Diebold-Marianos test.

I avsnitt 3 redovisar vi hur vi valt vilka tidsserier vi använde för att analysera metoderna samt information angående datamaterialet som varit viktigt i vårt arbete. Vi berättar även hur vi använt olika programvara samt vilken betydelse detta haft för uppsatsen.

I avsnitt 4 presenteras hur genomförandet av studien gått till samt vilket utfallet blev. Det görs i form av beskrivningar av hur vi applicerat de metoder som vi studerar och resultaten redovisas i tre tabeller.

I avsnitt 5 preciserar vi våra slutsatser och utifrån dessa ger vi sedan rekommendationer till Electrolux.

I avsnitt 6 diskuterar vi olika aspekter av studien, resultatet och genomförandet. Vi identifierar brister i studien samt olika intressanta analysmöjligheter samt vidare utredning som skulle bygga vidare från denna uppsats.

2. Metoder inom tidsserieanalys

2.1 Centrala begrepp i tidsserieanalys.

Själva idén med tidsserieanalys är att identifiera och extrapolera mönster i en uppsättning observationer i en tidsserie. Den del av tidsserien som inte kan uttryckas i någon form av återkommande mönster kan anses bero på slumpen. Det brukar oftast kallas *vitt brus* (Montgomery m.fl., 2008, s 171) och kan betecknas som en slumpmässig felterm, ε_t . De modeller som beskrivs längre fram försöker att beskriva mönstret i tidsserien så väl att då mönstret har subtraherats från datamaterialet så återstår vitt brus. Modellerna handlar alltså om allt som placeras innan feltermen ε_t .

Feltermen ε_t antas ha konstant varians och konstant medelvärde, och den förväntade kovariansen för alla tidsavstånd är noll.

Förutom det sistnämnda kriteriet är detta även uttrycket för en tidsserie som uppvisar *stationaritet*. Detta begrepp används återkommande för att beskriva egenskaper hos en tidsserie. Stationaritet innebär att egenskaperna för en tidsserie är konstanta över tid. Med egenskaper menas här medelvärdet och variansen (Montgomery m.fl. 2008, s 25-

26) och kovariansen (Kirchgässner & Wolters, 2009, s 13-14.) I korthet gäller för en stationär tidsserie följande:

$$E[X_t] = \mu_t = \mu$$
$$Var[X_t] = E[(X_t - \mu_t)^2] = \sigma^2$$

Vidare gäller att kovariansen för en stationär tidsserie endast är en funktion av tidsavståndet mellan två observationer och ej av var i tiden som observationerna befinner sig. (Kirchgässner & Wolters, 2009, s 13-14). Vitt brus är alltså en stationär tidsserie, men en stationär tidsserie behöver inte vara vitt brus.

Kovariansen i en tidsserie delat med variansen ger oss korrelationen. I univariata tidsserier benämns korrelationen mellan olika tidsavstånd för *autokorrelation*. Det tal som uttrycker detta kallas följaktligen för autokorrelationskoefficienten. Om man kontrollerar för alla mellanliggande korrelationer får man den partiella autokorrelationen.

Ett av de första stegen i försöken att avslöja underliggande trender i en tidsserie är att tillämpa ett *glidande medelvärde*. Detta är en av de enklaste modellerna inom tidsserieanalys och innebär att varje observation byts ut mot ett medelvärde av några närliggande observationer (inklusive den aktuella observationen). Ett glidande medelvärde kan vara centrerat eller inte. Närliggande det glidande medelvärdet är glidande medianer som kan användas om det finns en del kraftigt avvikande observationer eller andra skäl att tillämpa denna mer robusta beskrivning. Glidande medelvärden kan inkorporeras i mer komplicerade modeller.

Enkel exponentiell utjämning (EEU) är en vanligt förekommande metod inom tidsserieanalys. Den utgör även den konceptuella basen för flera mer avancerade metoder. EEU kan sägas vara en linjärkombination av den senaste observationen och den utjämnade föregående observationen. Vikten mellan dessa bestäms av en konstant θ . Denna konstant kan benämnas utjämningsfaktorn. (Montgomery m.fl. 2008, s 176).

$$(2.1) \quad \tilde{Y}_t = (1 - \theta)Y_t + \theta\tilde{Y}_{t-1}$$

Konstruktionen av (2.1) leder till slutsatsen att man i användandet av EEU behöver veta vad det första utjämnade värdet är. Detta värde kan kallas initialvärde och betecknas \tilde{Y}_0 . Montgomery et al (2008, s 178) nämner två sätt att skatta \tilde{Y}_0 :

a) $\tilde{Y}_0 = Y_1$, och b) $\tilde{Y}_0 = \bar{Y}$

Medelvärdet \bar{Y} kan beräknas på en undergrupp av observationerna. Till exempel kan medelvärdet av de första observationerna vara en bra grund, särskilt om de är lokalt stationära. I EEU ges observationer längre bak i tiden en geometrisk minskande vikt.

Därför kommer betydelsen av hur initialvärdet skattas att minska allteftersom antalet observationer blir fler. Ju lägre θ , desto mer vikt kommer att läggas på den senaste observationen och utjämningen kommer att vara känsligare för ändringar i datamaterialet. Prognosen i enkel exponentiell utjämning är samma som det förra utjämnade värdet. Alltså:

$$(2.2) \quad \tilde{Y}_{t+1} = \tilde{Y}_t$$

Alla prognoser för h steg framåt i tiden är desamma för (2.2). (Montgomery m.fl. 2008, s 194).

2.2 Holts metod (dubbel exponentiell utjämning)

EEU är olämplig att använda på serier som innehåller en trend. I de fallen kommer EEU att antingen underskatta de faktiska värdena (vid positiv trend) eller överskatta dem (vid negativ trend) (Montgomery m.fl., 2008, s 184). Ett sätt att inkludera en trendfaktor i exponentiell utjämning är den så kallade Holts metod, en form av dubbel exponentiell utjämning. Grundtanken är att tidsserien beror på två komponenter, en nivå (L_t) och en trend (T_t) som båda förändras över tid:

$$\hat{Y}_t = L_t + T_t$$

Med trend menas skillnaden med aktuellt och föregående värde i en tidsserie. Namnet dubbel exponentiell utjämning kommer av att både nivån och trendkomponenten utjämnas över tid med var sin utjämningskonstant. De skattas med följande ekvationer:

$$(2.3) \quad \hat{L}_t = \alpha Y_t + (1 - \alpha)(\hat{L}_{t-1} + \hat{T}_{t-1})$$

$$(2.4) \quad \hat{T}_t = \beta(\hat{L}_t + \hat{L}_{t-1}) + (1 - \beta)\hat{T}_{t-1}$$

Utjämningsparametrarna α och β i (2.3) och (2.4) ligger mellan 0 och 1. Värdena på α och β avgör vilken tyngd vi lägger på de senaste observationerna. Ju lägre värde på respektive konstant, desto trögrörligare blir den skattade serien. En generell tumregel då dessa bestäms är att sätta dem så att prognosfelel minimeras (Montgomery m.fl. 2008, sid 194). L_t blir en dubbel exponentiell utjämningskattning av datamaterialet i period t . Med hjälp av T_t , adderad till L_t får vi en prognos för perioden $t + \tau$:

$$(2.5) \quad \hat{Y}_{t+\tau} = \hat{L}_t + \tau \hat{T}_t$$

För skattning av både L_t och T_t behövs ett initialvärde, som kan beräknas på olika sätt. Montgomery m.fl. (2008, s 191) föreslår linjär regression. Interceptet blir \hat{L}_0 och lutningen blir \hat{T}_0 .

2.3 Holt-Winters metoder

Vissa tidsserier, inte minst sådana som beskriver försäljning, uppvisar det som kallas för säsongsvariation. Att glassförsäljningen ökar på sommaren medan skridskor säljs mer på vintern är ganska självklart, men detta gäller även de flesta varor som tillverkas av Electrolux. Detta mönster kan uppträda med eller utan trend. Det är möjligt att i termer av exponentiell utjämning bygga modeller som tar hänsyn till den här typen av variationer. Den modell som finns tillgänglig i SAP är Winters metod, som eftersom den kan reduceras ner till Holts metod också kallas Holt-Winters metod. Beroende på säsongsvariationens natur kan man använda två olika metoder, en additiv eller en multiplikativ.

2.3.1 Holt-Winters additiva metod

Holt-Winters utökar Holts metod genom att addera en säsonganpassning:

$$\hat{Y}_t = \hat{L}_t + \hat{S}_t$$

L_t överensstämmer med motsvarande term i Holts metod och S_t är säsonganpassningen för period t . För S_t gäller:

$$(2.6) \quad S_t = S_{t+is} ; i = 1, 2, \dots, n ; \sum_{t=1}^s S_t = 0$$

I (2.6) är S antalet perioder som utgör en cykel och n är antalet cykler. Detta innebär att den teoretiska säsongsvariationen är densamma för varje period under olika cykler. Om vi har månadslånga perioder över årlånga cykler har alltså en given månad samma säsonganpassning år från år. Vi får tre skattningar, trend, nivå och säsong, som skattas löpande enligt följande:

$$(2.7) \quad \hat{L}_t = \alpha(Y_t - \hat{S}_{t-s}) + (1 - \alpha)(\hat{L}_{t-1} + \hat{T}_{t-1})$$

$$(2.8) \quad \hat{T}_t = \beta(\hat{L}_t - \hat{L}_{t-1}) + (1 - \beta)\hat{T}_{t-1}$$

$$(2.9) \quad \hat{S}_t = \gamma(Y_t - \hat{L}_t) + (1 - \gamma)\hat{S}_{t-s}$$

I (2.7), (2.8) och (2.9) gäller att parametrarna α, β, γ ligger mellan 0 och 1. Som tidigare behövs initialvärden. Dessa fås genom att anpassa en linjär regression med $s-1$ dummyvariabler till materialet (Montgomery m.fl., 2008, s 212). Interceptet ger oss \hat{L}_0 , lutningen \hat{T}_0 och dummyvariablernas parametrar, standardiserade så de summerar till noll, ger \hat{S}_i för den första säsongen (SAS, 2010a).

En prognos för period $(t + \tau)$ ges av:

$$\hat{Y}_{t+\tau} = \hat{L}_t + \tau \hat{T}_t + \hat{S}_t (\tau - s)$$

Eftersom storleken av säsongsvariationen är oberoende av storleken på nivåskattningen, så passar den här modellen bäst när säsongsvariationen inte är proportionell mot nivån på observationen. Detta innebär att om t.ex. en produkt i Electrolux sortiment ökar i försäljning så kommer skillnaden mellan olika månader under året vara densamma i absoluta tal.

2.3.2 Holt-Winters multiplikativa modell

Om säsongsvariationen är proportionell mot storleken på datamaterialet, så är en multiplikativ modell att föredra. Detta är alltså lämpligt om man tror att t.ex. en ökad försäljning av en viss produktgrupp kommer öka skillnaden i försäljning mellan olika månader under året i absoluta tal men i relativa tal vara densamma. I Holt-Winters multiplikativa metod multipliceras säsonganpassningen med nivåskattningen:

$$\hat{Y}_t = \hat{L}_t \hat{S}_t$$

För S_t gäller nu istället:

$$\sum_{t=1}^s S_t = s$$

Vi har återigen tre skattningar, trend, nivå och säsong, att uppdatera enligt:

$$(2.10) \quad \hat{L}_t = \alpha \left(\frac{Y_t}{\hat{S}_{t-s}} \right) + (1 - \alpha) (\hat{L}_{t-1} + \hat{T}_{t-1})$$

$$(2.11) \quad \hat{T}_t = \beta (\hat{L}_t - \hat{L}_{t-1}) + (1 - \beta) \hat{T}_{t-1}$$

$$(2.12) \quad \hat{S}_t = \gamma \left(\frac{Y_t}{\hat{L}_t} \right) + (1 - \gamma) \hat{S}_{t-s}$$

Som tidigare gäller det att parametrarna i (2.10), (2.11) och (2.12) ligger mellan 0 och 1.

SAS skattar initialvärdena på samma sätt som för den multiplikativa modellen, men dummyvariablernas parametrar standardiseras till medelvärde 1. En prognos för period $(t + \tau)$ ges av:

$$\hat{y}_{t+\tau} = (\hat{L}_t + \tau \hat{T}_t) \hat{S}_t (\tau - s)$$

2.4 ARIMA

ARIMA modeller är en kombination av s.k. autoregressiva modeller (AR-modeller) och glidande medelvärdesmodeller (Moving Average – MA-modeller). Både glidande medelvärdesmodeller och autoregressiva modeller utgår från att en stationär tidsserie kan beskrivas som summan av medelvärdet och en viktad summa av slumpmässiga avvikelser. En MA-modell kan skrivas (Montgomery m.fl. 2008, s 240):

$$(2.13) \quad y_t = \mu + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$$

Avvikelsena antas vara identiskt och oberoende fördelade. Ett rimligt antagande man kan göra om observationernas påverkan på varandra är att de minskar med tidsavståndet. Då kan vi tillskriva det oändliga antalet avvikelser minskande vikter för att beskriva det minskande inflytande de har på den aktuella observationen. Ofta antas dessa vikter vara exponentiellt avtagande och modelleras genom att sätta $\psi_i = \phi^i$ med $|\phi| < 1$. Då kan ekvation 2.13 skrivas:

$$(2.14) \quad y_t = \mu + \varepsilon_t + \phi \varepsilon_{t-1} + \phi^2 \varepsilon_{t-2} + \dots = \mu + \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i}$$

Detta innebär att:

$$(2.15) \quad y_{t-1} = \mu + \varepsilon_{t-1} + \phi^1 \varepsilon_{t-2} + \phi^2 \varepsilon_{t-3} + \dots$$

Eftersom alla avvikelser i ekvation (2.15) finns i ekvation (2.14) kan dessa kombineras till:

$$(2.16) \quad y_t = \mu + \varepsilon_t + \phi(y_{t-1} - \mu) = \delta + \phi y_{t-1} + \varepsilon_t$$

I (2.16) är $\delta = \mu - \phi\mu$. Ovan transformation har resulterat i en autoregressiv modell av första ordningen. Både (2.14) och (2.16) kan ses som specialfall av (2.13). Det kan vara så att påverkan från tidigare observationer är mer komplicerad än att autokorrelationen är exponentiellt avtagande. I sådana fall kan det vara nödvändigt att anpassa en AR-modell av högre ordning. En sådan AR (p)-modell kan skrivas:

$$(2.17) \quad y_t = \delta + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + e_t$$

Ändliga MA-modeller utgår även de från ekvation (2.13). För dem gäller att:

$$\psi_i = \begin{cases} 1, & i = 0 \\ -\theta_i, & 0 < i \leq q \\ 0, & i > q \end{cases}$$

En MA (q) modell kan alltså betraktas som ytterligare ett specialfall av (2.13) och kan skrivas (Montgomery m.fl. 2008, sid 235):

$$(2.18) \quad y_t = \mu + e_t - \theta e_{t-1} - \dots - \theta_q e_{t-q}$$

Dessa modeller utgår alltså från att en observation kan beskrivas som summan av medelvärdet plus aktuell observations avvikelse minus summan av q föregående avvikelser i avtagande betydelse.

Modellerna presenterade i (2.17) samt (2.18) har den viktiga egenskapen gemensamt att deras parametrar är skattningsbara.

En ARMA-modell är en kombination av en AR- och en MA-modell. En ARMA-modell kan generellt skrivas (Montgomery m.fl., 2008, s 253):

$$(2.19) \quad y_t = \delta + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + e_t - \theta e_{t-1} - \dots - \theta_q e_{t-q}$$

Det vanligaste sättet att avgöra hur en modell skall byggas är att se på hur autokorrelationen och den partiella autokorrelationen beter sig på olika tidsavstånd. Kirschgässner & Wolters tillhandahåller en konceptuell matris för detta och så även Montgomery m.fl. (se appendix 3). För att modellen ska vara praktiskt möjlig att bygga måste AR- och MA-processerna båda två vara finita, det vill säga:

$$p < \infty; q < \infty$$

Som tidigare nämnts måste tidsserien, för att kunna beskrivas av ARMA- processer, vara stationär. Problemet är att ganska många tidsserier inte är det, till exempel sådana som innehåller en trend. Lösningen på detta är att tillämpa en lämplig transformation, med syfte att åstadkomma stationaritet. En av de vanligaste transformationerna är så kallad differentiering. Detta kan skrivas (Montgomery m.fl., 2008, s 256):

$$(2.20) \quad w_t = y_t - y_{t-1} = (1 - B)y_t$$

$$(2.21) \quad w_t = (1 - B)^d y_t; \quad 1 < d < \infty$$

Faktorn $(1 - B)$ är den så kallade bakåtoperatoren som enkelt uttrycker vilken ordning d vi har differentierat vår serie på. Ett hjälpmedel för att avgöra om en tidsserie är stationär är Dickey-Fullers utökade test. Det kontrollerar om en serie är stationär vid existensen av autokorrelerade feltermen. Nollhypotesen är att tidsserien är icke-stationär och mothypotesen är att serien är stationär (Se appendix 1). Tillämpas AR, MA och integrering på en gång kallar vi y_t för en autoregressiv integrerad glidande medelvärdesprocess (ARIMA) av ordningarna p , d och q . (Integrerad eftersom y_t kan beskriva som summan, eller integralen, av w_s , $s = 1, 2, \dots, t$, med ett startvärde y_0). En ARIMA(p, d, q)-process kan alltså skrivas:

$$(2.22) \quad (1 - B)^d y_t = \delta + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + e_t - \theta e_{t-1} - \dots - \theta_q e_{t-q}$$

En anledning till att ARIMA-modeller fick ett snabbt genombrott är att de, jämfört med stora ekonometriska modeller, har hög pålitlighet i prognoserna och att prognostiseringen är relativt enkel (Kirchgässner & Wolters, 2009, s 77). En ARIMA(1,1,1)-prognos för tidpunkt $t + 1$ kan skrivas (Montgomery m.fl., 2009 s 277):

$$(2.23) \quad \hat{y}_{t+1} = (1 + \phi)y_t - \phi y_{t-1} - \theta e_t(1)$$

Uttrycket $(1 + \phi)y_t - \phi y_{t-1}$ kan skrivas om till $y_t + \phi y_t - \phi y_{t-1}$ och följaktligen till $y_t + \phi(y_t - y_{t-1})$. Enligt (2.4.8) får vi då $y_t + \phi(1 - B)y_t$. Detta är alltså innevarande observation plus en vikt gånger första differensen på innevarande och föregående observation. Uttrycket $e_t(1)$ skall läsas som den beräknade feltermen baserat på föregående tidsperiods prognos för innevarande tidpunkt: $y_t - \hat{y}_t(T - 1)$. Därmed finns en MA(1)-del och en AR(1)-del gjord på första differensen.

MA-delen, som utgår från de siffror som gavs i $e_t(1)$, kommer att påverka prognosen endast i den första prognosen och om det är en MA (2) lösning så påverkar den prognosen i två prognoser och så vidare (Kirschgässner & Wolters, 2009, s 83). ARIMA behandlar säsong på ungefär samma sätt som andra processer. Vi föreställer oss en säsongprocess såsom en summa av två delar; en del består av säsongvariationen S_t , och den andra delen är en stokastisk process N_t som kan modelleras som en ARMA-process:

$$(2.24) \quad y_t = S_t + N_t$$

Eftersom säsongsvariationen innebär att $S_t = S_{t-s}$ är det möjligt att differentiera bort den:

$$(2.25) \quad \begin{aligned} y_t - y_{t-s} &= S_t - S_{t-s} + N_t - N_{t-s}; & S_t - S_{t-s} &= 0 \\ (1 - B^s)y_t &= (1 - B^s)N_t \\ w_t &= (1 - B^s)N_t \end{aligned}$$

I (2.25) är $(1 - B^s)$ en differentiering på avstånd s och alltså inte av ordningen s . Resultatet w_t är säsongstationär och möjlig att beskriva med ARMA-modeller.

Om vi istället för att se S_t som enbart en deterministisk process även tillskriver den stokastiska egenskaper så kan betydande autokorrelation på säsongstidsavståndet kvarstå även efter differentieringen. Denna variation kan även den beskrivas som en ARMA-process. En generell ARIMA-modell som innehåller en stokastisk säsongskomponent kan då vara av ordning $(p,d,q)(P,D,Q)$. Versalerna representerar den ARMA och den differentiering som genomförts på säsongsavståndet. Montgomery m.fl. (2008, s 283) skriver att i de flesta fall förväntas inte ARIMA på säsong överstiga $(1,1,1)$.

Prognoser för säsongmodeller erhålls på samma sätt som i ickesäsongmodeller.

2.5 Utvärderingsmetoder

Vägledande för all form av modellval är vilken utvärderingsmetod som väljs. Det går att dela upp dessa i två typer. Den ena kan benämnas jämförelsetal och den andra test (Clements & Hendry, 1998, s 312-313). Denna konceptuella uppdelning hjälper oss att förstå grunderna för våra modellval. När vi väljer modell baserat enbart på jämförelsetal finns det en risk att slumpen påverkar vårt beslut. Därför behöver vi utföra test för att pröva om de skillnader vi kan se är statistiskt signifikanta.

I detta avsnitt beskriver vi styrkor och svagheter i ett flertal jämförelsetal samt ett test för skillnaden mellan prognosmodeller.

2.5.1 Halveringsmetod

När styrkan med en prognosmetod skall utvärderas är det viktigt att komma ihåg att den modell som bäst beskriver det historiska datamaterialet inte behöver vara den modell som lämpar sig bäst för prognoser. Enligt Ashley & Patterson (2010, sid 2-3) kan en bra modell Anpassning till ett visst datamaterial snarare ses som ett mått på modellens flexibilitet. Därför är det bra att utvärdera en modells prestation på ett datamaterial som inte använts för att bygga detsamma. Detta gäller för både jämförelsetal och test. Det vanligaste sättet att göra detta på är att dela upp tidsserien i två delar. En del används för att anpassa modellen och en del för att testa modellen. Montgomery m.fl. (2008, s 25) föreslår att minst 20-25 observationer bör vara en del i testdatat. Ashley & Patterson

(2010) lyfter också fram betydelsen av att inkludera observationer för längre perioder så att utvärderingen inte blir alltför beroende av en specifik tidsperiod.

För vårt test och för våra jämförelsetal tillämpar vi en uppdelning där 24 observationer är en del av testdatat. Detta ger oss 54 observationer att anpassa våra modeller utifrån. Denna uppdelning valde vi för att 24 observationer i våra serier representerar två år (eller två säsongcykler). Vårt testdata som utgörs av observationerna mellan juli 2008 till och med juni 2010 täcker en finansiellt orolig period. Våra tidsserier har synbarligen inte ändrat mönster under denna period (se appendix 3). Därmed utgår vi från att den period testperioden täcker inte är sådan att den är alltför speciell för att inte kunna användas till att testa och utvärdera våra modeller.

2.5.2 Jämförelsetal

Som grund för att bygga jämförelsetal används prognosfel. Ett prognosfel är skillnaden mellan prognosen och det verkliga värdet och kan för en prognos gjord för en tidsperiod framåt i tiden skrivas:

$$e_t(1) = y_t - \hat{y}_t(t-1).$$

Från denna felterm, residualen på tidsavstånd ett, kan olika jämförelsetal beräknas. *Genomsnittligt fel* är summan av residualerna delat med antalet. Väntevärdet för denna är 0 såtillvida den är fri från systematiskt fel. Om man kvadrerar feltermen får man ett storleksmått på feltermen. Medelkvadratsumman, given i (2.26) används ofta, men har den nackdelen att extrema observationer har en stor påverkan. Så är också fallet för tal som är närliggande medelkvadratsumman, som till exempel kvadratroten ur medelkvadratsumman definierad i (2.27), (vi kommer referera till detta tal enligt dess engelska förkortning RMSE). I de fall man vill begränsa effekten av avvikande observationer kan man använda absolutvärden.

$$(2.26) \text{ Medelkvadratsumman} = \frac{\sum_{i=1}^n e_i(1)^2}{n}$$

$$(2.27) \text{ Roten ur medelkvadratsumman} = \sqrt{\frac{\sum_{i=1}^n e_i(1)^2}{n}}$$

Det absoluta medelfelet är summan av absolutfelet delat med antalet observationer.

$$\frac{\sum_{i=1}^n |e_i(1)|}{n}$$

Det absoluta medelfelet storlek är skalberoende och det lämpar sig därför inte att använda i jämförelser mellan tidsserier. Därför kan man använda ett mått som på engelska heter Mean Absolute Percentage Error (MAPE). Vi kan översätta det till det procentuella absoluta medelfelet men kommer i texten benämna det med dess engelska förkortning.

MAPE beräknas på följande sätt för residualen på tidsavstånd ett:

$$\frac{\sum_{i=1}^n \left(\frac{|e_i(1)|}{x_i} \right)}{n} \cdot 100$$

2.5.3 Informationskriterier.

Det kan argumenteras för att det generellt föreligger ett positivt samband mellan hur komplicerad en modell är och hur den presterar. En modell med fler parametrar klarar oftast att förklara större del av variationen i datamaterialet. Ett bra exempel på detta är multipel regressionsanalys där fler förklarande variabler ger en ökad förklaringsgrad och ett lägre MAPE. Men en förklarande variabel kan enbart genom slumpen öka förklaringsgraden och även om den individuella förklaringsförmågan hos en variabel är låg, kan modellen som helhet enligt vissa mått prestera bättre. Frågan är då om vår förståelse av datamaterialet ökar för att vi kopplat en del av den slumpmässiga variationen till en variabel. Det är alltså inte entydigt hur en komplex modell ska bedömas (Gujarati m.fl., 2008, s 206).

Som ett svar på denna problematik är det möjligt att använda olika former av informationskriterier. De är jämförelsetal som påverkas av modellens komplexitet.

Akaiikes informationskriterium (AIC) och Schwarz informationskriterium (SIC) är två av dessa. De ges enligt följande formler:

$$(2.28) \quad AIC = \ln \left(\frac{\sum_{t=1}^T e_t^2}{n} \right) + \frac{2p}{n}$$

$$(2.29) \quad SIC = \ln \left(\frac{\sum_{t=1}^T e_t^2}{n} \right) + \frac{p \ln(n)}{n}$$

I (2.28) och (2.29) är p = antal parametrar i modellen. Som kan ses så ökar AIC och SIC då antalet parametrar ökar. Enklare modeller får därför lägre tal och bedöms som mer informativa. Modellerna är lika i den första delen av formeln. Denna uttrycker logaritmen av det kvadrerade medelfelet. Skillnaden mellan AIC och SIC ovan, som uttrycks i den senare delen av formlerna, innebär att AIC straffar modellen med två gånger modellparametrarna och SIC straffar modellen med parametrarna multiplicerat med logaritmen på urvalet. Skillnaden mellan formlerna innebär att då n ökar så kommer bestraffningen i AIC minska snabbare än SIC. SIC kommer att ge en högre bestraffning för komplicerade modeller på alla urval som är större än sju. Under detta gränsvärde bestraffar AIC modellkomplexitet mer än SIC. (Vid gränsvärdet 7,39 gäller att $2p = p \cdot \ln(n)$.)

2.5.4 Diebold-Marianos test

Många test utgår från ett visst jämförelsetal, ofta den kvadrerade medelavvikelsen, och antar även att residualerna har noll i medelvärde (dvs. väntevärdesriktig), att de är normalfördelade och att de är okorrelerade över olika tidsavstånd (Clements & Hendry 1998, s 321). Diebold och Mariano har föreslagit ett test av skillnaden mellan konkurrerande prognosmodeller som inte kräver att ovan angivna antaganden är uppfyllda. (Clements & Hendry 1998, s 322). Detta innebär att Diebold-Marianos test har ett brett tillämpningsområde samt är relativt robust.

Man kan anta att en modells kvalité kan uttryckas i en funktion av residualen på olika tidsavstånd, $g(e_i)$. När två modeller presterar lika bra antar vi att:

$$E[g(e_{1t}) - g(e_{2t})] = 0$$

Om vi definierar skillnaden mellan modell 1 och modell 2 såsom:

$$d_t = g(e_{1t}) - g(e_{2t}); t = 1, \dots, n$$

så blir det skattade medelvärdet för skillnaden mellan modellerna:

$$(2.30) \quad \bar{d} = n^{-1} \sum_{t=1}^n d_t$$

Utgångspunkten för Diebold-Marianos teststatistiska är (2.30).

När vi jämför prognoser på tidsavstånd större än 1 är prognosen för period h till sin natur korrelerad till prognosen för period $h-1$, $h-2$ och så vidare. Därför kan man inte enbart använda den skattade variansen i det statistiska testet.

Diebold-Mariano inkluderar därför autokorrelationer mellan feltermerna upp till tidsavståndet h , för en prognos h steg framåt, i sin skattning av variansen för \bar{d} som asymptotiskt är lika med:

$$(2.31) \quad V(d) \approx n^{-1} \left[\gamma_0 + 2 \sum_{k=1}^{h-1} \gamma_k \right],$$

I (2.31) kan autokorrelationen γ_k skattas med:

$$\hat{\gamma}_k = n^{-1} \sum_{t=k+1}^n (d_t - \bar{d})(d_{t-k} - \bar{d})$$

För prognoser på tidsavstånd ett får man följande uttryck för variansen (Harvey m.fl., 1997, s 283).

$$V(d) \approx n^{-1} [\gamma_0] \text{ eller } V(d) \approx n^{-1} [\text{var}_{\bar{d}_t}]$$

Diebold- Marianos teststatistika beräknas enligt följande:

$$(2.32) \quad S_1 = \left[\hat{V}(\bar{d}) \right]^{-1/2} \bar{d}$$

Noll- och mothypotes för testet kan skrivas:

$$H_0 : E[d_t] = 0$$

$$H_a : E[d_t] \neq 0$$

Under nollhypotesen är teststatistikan i (2.32) normalfördelad med medelvärde 0 och varians 1.

Nollhypotesen förkastas enligt ovan med $\alpha = 0,05$ för värden som är större/mindre än $\pm 1,96$.

Det har dock visat sig att detta test är överkänsligt i små och medelstora stickprov samt att detta problem förvärras ju större h -steps-prognos som görs (Harvey m.fl., 1997, s. 290, s 282). Därför har Harvey m.fl. föreslagit ett modifierad Diebold-Mariano-test.

Deras modifikation innebär att skattningen av variansen på \bar{d} förändras något. Testet presterar med förändringen bättre, men är fortfarande något överkänsligt.

Teststatistikan för modifierade Diebold-Mariano kan skrivas:

$$S_1^* = \left[\frac{n+1-2h+n^{-1}h(h-1)}{n} \right]^{1/2} S_1$$

För prognoser som utförts på tidsavstånd ett får man följande uttryck:

$$S_1^* = \left[\frac{(n-1)}{n} \right]^{1/2} S_1$$

Harvey m.fl. (1997) motiverade även användandet av t-fördelningen istället för normalfördelningen som jämförelse till teststatistikan (s 283).

3. Datamaterial

3.1 Val av tidsserier

På den svenska marknaden har Electrolux ca 1700 olika produkter indelade på ca 120 subgrupper och ca 27 produktgrupper. Prognostiseringar sker som regel på subgruppsnivå, då enskilda produkter kan ha kort livslängd, medan produktgrupperna aggregerar produkter med olika konsumtionsmönster. Att genomföra en djupare analys på de 120 subgrupperna vore förstås önskvärt, men omöjligt inom den här uppsatsens ramar. Det nödvändiga urvalet har gjorts enligt principen om spridning. Det innebär att vi helt enkelt valt tidsserier som skiljer sig när det gäller trend, säsong och varians.

I SAP APO-DP lagras Electrolux 2 års historik av försäljningssiffror, som går att dela upp på månader och veckor. Då detta material är aningen begränsat, framförallt för utvärdering, har Electrolux ställt ett annat datamaterial till förfogande. I denna databas, EFS, finns tillgång till data från januari 2004 till juni 2010. Eftersom detta material ger större möjlighet till analys och utvärdering använder vi denna källa istället för att exportera från SAP-APO-DP. Detta innebär att vårt datamaterial inte helt täcker samma period som det datamaterial som för närvarande är tillgängligt för analyser i den dagliga produktionsmiljön på Electrolux.

Siffrorna i SAP APO-DP och EFS skiljer sig åt i viss mån. Detta beror på att man tillämpat något olika affärslogik i systematiseringen av dessa. Enligt experter på Electrolux skall detta inte innebära att det påverkar tidsseriernas natur. Därmed är det möjligt för oss att använda det större datamaterialet, analysera detta och sen ge modellvalsrekommendationer som är aktuella för datamaterialet i SAP-APO-DP.

Det är också värt att nämna att datamaterialet i SAP-APO-DP i vissa fall är påverkat även av en manuell rensningsprocess som inte tillämpats på datamaterialet i EFS. Detta är dock så pass sällsynt och i sådan pass begränsad utsträckning att det inte kan antas påverka våra analyser i stort.

3.2 Jämförelse med ARIMA

SAP APO- DP tillåter användare att importera externa prognoser och dessutom tillhandahåller SAS en modul som kan skapa tidsserieanalys till SAP, därför ansåg vi det var intressant, även för Electrolux, att analysera metoder utanför SAPs räckvidd. Genom att göra detta så går det även att göra en bedömning över kvalitén på de i SAP-APO-DP direkt tillgängliga metoderna. Om det visar sig att en annan uppsättning modeller presterar mycket bättre så bör Electrolux överväga att komplettera med annan programvara. Vi valde att jämföra de i SAP-APO-DP tillgängliga exponentiella utjämningsmetoderna med ARIMA- modeller. De är med tanke på deras popularitet och flexibilitet ett naturligt val för en jämförelse.

3.3 Val av mjukvara

Ett flertal begränsningar i SAP-APO-DP gjorde att det var naturligt för oss att genomföra våra analyser i SAS. För det första, som ovan nämnt, har SAP-APO-DP enbart 24 månaders lagrad data. För det andra har SAP-APO-DP mycket begränsade möjligheter till analys av residualer, autokorrelationer etc. Slutligen är SAS enklare att använda för att exportera data till externa program. Detta argument var aktuellt då vi kompletterade användningen av SAS med olika beräkningar i Microsoft Excel. Detta har varit fallet i utförandet av Diebold-Marianos test och även för beräkning av vissa jämförelsetal.

4. Resultat

Detta avsnitt ger först en genomgång av hur analysen genomfördes för de metoder vi valde och därefter en sammanställning av resultaten för metoderna avseende jämförelsetalen och det modifierade Diebold-Marianos test.

Utifrån tidigare redogjorda urvalsprinciper så kom vi fram till att följande fem tidsserier tillfredsställer de krav vi har för att testa metoderna. Dessa finns grafiskt presenterade i appendix 3.

Tabell 1.
Undersökta tidsserier och deras karaktäristiska

<i>Serie</i>	<i>Produkttyp</i>	<i>Karaktäristiska</i>
3503	Frysack	Säsong
4022	Tillbehör till spisar	Oregelbunden säsong
2224	Fristående spisar	Trend och säsong – multiplikativ
920	Handhållna golvdammsugare	Trend och oregelbunden säsong
4033	Tillbehör till köksfläktar	Trend och säsong – additiv

Som man kan se så innehåller alla serierna någon form av säsongskomponent. Dessa är däremot olika för de olika tidsserierna För 3503 innebär karaktäriseringen att den inte innehåller trend. Uppdelningen har gjorts utifrån en visuell betraktelse.

4.1 Tillämpning av exponentiella utjämningsmetoder

För Holts metod samt Holt-Winters additiva och Holt-Winters multiplikativa metoder använde vi proceduren ESM i SAS. Inom denna lät vi programmet automatiskt bestämma vilka värden som skulle ges på utjämningskonstanterna. SAS sätter dessa utefter vilket som ger lägst RMSE vid ex post prognos (SAS, 2010b).

Det går även att låta programmet använda en mindre datamängd i sin skattning av initialvärdena men då vi inte har något specifikt teoretiskt skäl att göra detta lät vi grundinställningarna stå orörda.

4.2 ARIMA resultat och exempel

Vi har följt gängse metodik för anpassning och analys av ARIMA-modeller. Arbetsgången är som följer. Det första som vi försökt avgöra är om tidsserien kan betraktas som stationär. Ett tydligt tecken på att den inte är det är om autokorrelationsfunktionen (ACF) avtar långsamt. I de fall en tidsserie uppvisade detta mönster gjorde vi en differentiering och studerade sen ånyo ACF. I samtliga fall utom 2 valde vi att göra detta, i ett fall även på säsong.

Om den differentierade tidsserien uppvisade ett stationärt mönster kompletterade vi den uppfattningen med att genomföra Dickey-Fullers utökade test, vilket finns tillgängligt i SAS. I samtliga fall bekräftade testet att serien var stationär.

Då stationaritet uppnåtts undersöktes ACF och partiell autokorrelation (PACF) för den bearbetade tidsserien. Utifrån beslutsregler presenterade av Montgomery samt Kirschgässner bestämmer man sedan vilken kombination av autoregressiva- och glidande medelvärdesprocesser som bör tillämpas (se appendix 2). En sammanfattning av våra bedömningar finns i tabell 2 och underlaget finns i appendix A4.

Tabell 2
Skattning av ARIMA modeller.

<i>Tidsserie</i>	<i>ARIMA</i> <i>(p,d,q)(P,D,Q)</i>
3503	(1,0,1)(0,1,0)
2224	(0,1,0)(0,0,1)
4033	(0,1,1)(0,0,1)
4022	(1,0,1)(1,0,1)
920	(0,2,2) (0,0,0)

I tidsserie 920 gjorde vi även den ovan beskrivna analysen på logaritmerat data, då förändringen i tidsserien i viss mån kan tänkas följa ett exponentiellt mönster. Då resultatet av detta gav ARIMA modeller med sämre resultat valde vi att inte logaritmera datat.

Som ett komplement till vår analys använde vi även en scan-funktion i SAS under proceduren ARIMA. Denna ger förslag på vilka ARIMA modeller som bör tillämpas. I samtliga fall fanns de modeller vi valt bland de olika modeller som SAS förslög.

4.3 Sammanfattande jämförelse

4.3.1 Jämförelsetal från expost-prognos

Tabell 3.
Jämförelsetal för 24 månaders expost prognoser.

	<i>RMSE</i>	<i>MAPE</i>	<i>AIC</i> ¹	<i>SIC</i>
3503				
H	428,18	23,90	294,85	297,21
AW	291,52	16,62	276,41	278,76
MW	334,78	18,68	283,05	285,40
ARIMA	374,16	20,36	288,38	290,74
2224				
H	79,99	18,49	214,32	216,69
AW	116,62	29,80	232,43	234,78
MW	81,73	20,42	215,36	217,72
ARIMA	83,03	20,36	216,12	218,48
4033				
H	258,35	20,72	270,67	272,96
AW	193,98	14,45	256,85	259,21
MW	214,40	15,65	261,66	264,01
ARIMA	240,68	18,48	267,21	269,56
4022-				
H	436,02	28,58	295,73	298,08
AW	300,25	21,22	277,82	280,12
MW	338,80	24,34	283,62	285,98
ARIMA	344,31	24,36	284,39	286,75
920-				
H	4599,53	46,40	408,82	411,17
AW	3401,27	33,58	394,33	396,69
MW	3136,53	32,57	390,44	392,80
ARIMA	4647,06	60,21	409,31	411,67

H= Holt, AW=Additiv Winter, MW=Multiplikativ Winter, ARIMA=den anpassade ARIMA modellen.

I tabell 3 presenteras resultatet för de fem olika tidsserierna för de fyra utvalda jämförelsetalen. Dessa siffror är framtagna på expost-prognoser och för de 24 sista observationerna i datamaterialet.

¹ Den metod som SAS använder för att räkna ut AIC och SIC är densamma som vi redovisat för tidigare med den skillnaden att de multiplicerar talet med urvalet. Därmed var vi tvungna att dividera deras tal med 24 för att få en samstämmighet med vår redovisade modell.

För att bättre kunna jämföra resultaten från expost-prognoserna och halveringsmetod har vi valt att göra sammanställningen på de sista 24 observationerna. På så sätt jämförs båda prognosavstånden mot samma observationer.

Vi har markerat de tal som är lägst med fet stil. Man kan se att additiv Holt-Winters har lägst tal i tre tidsserier, multiplikativ Holt-Winters i en och Holts metod i en.

4.3.2 Diebold-Marianos test

För att se om skillnaderna mellan expost-prognoser är statistiskt signifikanta har vi tillämpat det tidigare nämnda modifierade Diebold-Marianos test. I tabell 4 redovisas resultaten.

Tabell 4
Metoder jämförda med modifierat Diebold-Mariano test² på kvadrerat fel

<i>Serie</i>	<i>MW/AW</i>	<i>MW/H</i>	<i>MW/ARIMA</i>	<i>AW/H</i>	<i>AW/ARIMA</i>	<i>ARIMA/H</i>
3503	0,44	-0,26	-0,17	-0,41	-0,38	0,15
4022	0,49	-0,34	-0,04	-0,39	-0,33	0,29
2224	-0,42	0,03	-0,02	0,37	0,40	-0,06
920	-0,17	-0,47	-0,63	-0,64	-0,65	-0,03
4033	0,40	-0,30	-0,32	-0,44	-0,49	0,12

Inget av dessa värden är större eller mindre än de kritiska värdena, hämtade ur t-tabellen (2,0 för 95 % konfidensgrad och 1,7 för 90 % konfidensgrad). Som synes är inte en enda jämförelse signifikant.

² Utfört på de sista 24 observationerna och avrundat till två decimaler.

4.3.3 Jämförelsetal med halveringsmetod

Tabell 5
Jämförelsetal från prognos med halveringsmetod på de 24 sista observationerna.

	<i>RMSE</i>	<i>MAPE</i>	<i>AIC</i>	<i>SIC</i>
3503				
H	675,96	44,40	316,77	319,13
AW	385,86	21,15	289,86	295,40
MW	388,55	21,22	292,20	295,73
ARIMA	393,11	25,57	292,76	296,29
2224				
H	255,17	72,53	270,01	272,37
AW	261,44	70,75	271,18	276,71
MW	216,44	61,75	264,11	267,65
ARIMA	123,02	31,50	234,99	237,34
4033				
H	256,47	22,37	270,26	272,61
AW	238,04	17,98	266,68	272,21
MW	221,93	16,41	265,31	268,85
ARIMA	227,20	19,11	264,44	266,80
4022				
H	432,89	32,52	295,38	297,74
AW	343,54	26,20	284,29	289,82
MW	349,38	25,06	287,10	290,63
ARIMA	337,57	28,02	289,45	295,34
920				
H	4851,23	60,63	411,38	413,73
AW	4540,45	56,13	408,20	413,73
MW	4314,08	51,28	407,74	411,28
ARIMA	4889,11	72,44	411,75	414,10

Tabell 5 visar samma tal som i tabell 3 men för en prognos med halveringsmetod. Vi har alltså gjort en prognos för 24 månader framåt och därefter jämfört prognosen med utfallet. Resultatet visar att det nu är någon av de tre metoder som inkluderar en säsongskomponent som presterar bäst. Denna tabell visar, liksom den förra, att en modell oftast presterar bäst på samtliga jämförelsetal.

En avvikelse från detta finns för subgrupp 4022 där multiplikativ Holt-Winters presterar bättre än additiv Holt-Winters avseende MAPE. Här är det också ARIMA modellen som får lägst RMSE. Skillnaderna mellan modellerna är dock marginell och detta gäller även för AIC och SIC. En trolig orsak till detta är att extrema observationer påverkar jämförelsetalen på olika sätt.

5. Slutsatser och rekommendationer

5.1 Slutsatser

Den tydligaste slutsatsen är att ingen metod var den bästa för samtliga undersökta tidsserier. Olika modeller måste alltså användas på olika typer av serier. Vidare kan vi notera att för flera serier presterade samma modeller mycket olika på de olika prognosavstånden. Intressant är att Holts metod och den additiva Holt-Winters metod presterar relativt sett bättre på ex-post prognoser än på halveringsmetod, med motsatsen gäller för de övriga. Det bästa resultatet på halveringsmetod kan tillskrivas Holt-Winters multiplikativa modell men i det fall de presterade sämre än ARIMA var skillnaden betydande. Det är alltså inte möjligt att rekommendera en enda modell för alla tidsserier utan det är viktigt att identifiera och tillämpa olika modeller. Vi kan också konstatera att de skillnader som uppvisas i jämförelsetal inte bekräftas av det modifierade Diebold-Marianos test. Enligt det testet kan alltså modellerna inte sägas skilja sig åt signifikant på ex-post-prognoser.

5.2 Rekommendationer

Electrolux, begränsat av möjligheterna i SAP APO-DP, kan endast basera sina prognosutvärderingar på de jämförelsetal som systemet genererar. De använder för närvarande en utvärderingsperiod på 12 månader efter en anpassningsperiod på 12. Utvärderingsmått beräknas endast på ex-post prognoser. En jämförelse mellan ex-post-prognos och halveringsmetod visar att goda jämförelsetal under anpassningsperioden inte nödvändigtvis resulterar i goda prognoser flera steg fram. Med tanke på att utvärderingsperioden enbart är 12 månader ökar det dessutom risken för att slumpen skall påverka modellvalet. Vi har också sett en tydlig risk för överanpassning när man låter programmet självt generera optimala prognoser. Mot bakgrund av att ARIMA fungerade bättre än samtliga exponentiella utjämningsmetoder endast i ett fall, dessutom på en svårprognostiserad serie, anser vi att SAP APO-DP i princip täcker Electrolux behov.

Våra resultat visar alltså att det är riskabelt att blint tillämpa olika modeller utifrån en generell rekommendation. Även fast det visserligen går att rekommendera Holt-Winters multiplikativa metod före andra val är detta som vi sett inte entydigt. Den bästa rekommendationen är därför att demand planners, ibland i samråd med statistiker, väljer en modell som inte enbart presterar väl på historiskt material, men som också ger en prognos som stämmer överens med demand planners förväntningar. Tillämpningen av statistiska metoder bör ses som en process, där en regelbunden utvärdering av tillämpade metoder är en självklarhet. Makradakis & Hibon (2000) konstaterar i sin stora studie då de jämförde ett antal modeller på 3003 stycken serier att metoderna presterar olika bra på olika tidshorisonter. Det kan alltså vara lämpligt att ta hänsyn till vilken tidshorisont man ska prognostisera på när man väljer modell.

6. Diskussion

I detta avsnitt analyserar vi det presenterade resultatet och avslutar med uppslag för ytterligare intressanta analyser och teoretiska frågeställningar.

6.1 Analys

6.1.1 Parameterskattningarna för de exponentiella utjämningsmetoderna.

Tabell 6
Skattningstabell för utjämningsvikter för de exponentiella utjämningsmetoderna då SAS optimeringsfunktion används.

<i>Tidsserie</i>	<i>Konstanter</i>	<i>H</i>	<i>AW</i>	<i>MW</i>
3503	Nivå	0,946	0,532	0,389
3503	Trend	0,001	0,001	0,001
3503	Season		0,001	0,376
2224	Nivå	0,999	0,028	0,019
2224	Trend	0,001	0,001	0,001
2224	Season		0,001	0,105
4033	Nivå	0,024	0,029	0,009
4033	Trend	0,001	0,001	0,001
4033	Season		0,001	0,285
4022	Nivå	0,288	0,333	0,304
4022	Trend	0,001	0,002	0,004
4022	Season		0,001	0,419
920	Nivå	0,02	0,026	0,213
920	Trend	0,001	0,001	0,005
920	Season		0,001	0,339

Som vi nämnde i föregående avsnitt så har vi låtit SAS göra optimerade skattningar av parametrar till våra exponentiella utjämningsmetoder. Detta gav oss ett antal skattningar som var nära noll och två skattningar som var nära ett.

Dessa situationer beskrivs av SAS. De vikter som är nära noll ger stor betydelse till föregående observation, som för exponentiella utjämningsmetoder i sin tur är beroende av den förr- förra observationen. Alltså antyder dessa vikter att modellerna har konstanta nivåer, trender eller säsonger. I motsvarande situation, med vikter nära ett, kan det innebära att en faktor kanske inte borde vara en del av modellen. T.ex. om säsongsvikten är nära ett är det kanske mer rimligt att exkludera denna (SAS, 2010c). Detta innebär vidare att initialskattningarna har stor betydelse när utjämningskonstanterna är nära noll.

Då vårt syfte var att skapa så träffsäkra modeller som möjligt har vi inte valt att exkludera olika faktorer utifrån det resultat som framkommit från SAS optimeringsprocess.

För serien 3503 är alla säsongsskattningar för additiv Holt-Winter nära noll. Detta stämmer bra överens med själva grundtanken för denna typ av modell som förutsätter att säsongsvariationen inte skiljer sig åt beroende på vilken nivå som serien ligger på. Man kan också se att säsongsskattningarna för multiplikativ Holt-Winter här ligger mellan 0,11 och 0,42, vilket bekräftar funktionen av en multiplikativ modell.

Det är värt att notera att parameterskattningarna för alla de exponentiella utjämningsserierna avseende trendfaktorn ligger nära noll. Detta innebär att den trends-kattning som minimerar feltermen nästintill kan betraktas som en linjär regression. Trendförändringen är med andra ord en konstant förändring. Ser man grafiskt på tidsserierna i appendix 5 är det lätt att föreställa sig att detta är korrekt för serierna 3503, 4022 och 4033. Det är inte lika klart i de två resterande tidsserierna.

Avseende Holts metod så gör den en parameterskattning på nivå för serie 3503 som är 0,95 och på serie 2224 som är 0,99. Detta innebär att modellen i denna tillämpning inte skiljer sig mycket från en random walk process, eller en process med s.k. naiva skattningar. Detta innebär att nästkommande observation prognostiseras till samma värde som den föregående.

6.1.2 Skillnad mellan anpassningsförmåga och prognostisering.

Om man jämför utfallet på ex-post-prognos och halveringsmetod så kan man konstatera att det är skillnad på en modells anpassningsförmåga och en modells prognoskapacitet. Allra tydligast framträder detta för Holts metod på serie 2224. För prognoser på ett tidsavstånd får den ett MAPE på drygt 18 procent, vilket är bättre än alla de tre andra konkurrerande modellerna. För prognoser på 24 månader däremot presterar den sämst, med en MAPE på 72 procent.

Vår generella bedömning är att då det råder stora skillnader mellan prognoser på dessa olika tidsavstånd så är det ett tecken på att modellen är överanpassad. Detta innebär att den är alltför beroende av närliggande observationer och därmed avviker från den teoretiska funktion som kan anses ligga bakom utfallen i tidsserien.

För att vända på resonemanget kan man uppmärksamma multiplikativ Holt-Winters metod på serie 4033 som har en MAPE på 15,7 procent på ex-post-prognoser och 16,4 procent på halveringsmetoden. Även ARIMA uppvisar ett likvärdigt resultat på denna serie. Detta indikerar alltså att modellerna ligger nära den teoretiskt "sanna" modellen.

6.1.3 Teoretiska förväntningar och utfall

Utifrån den beskrivning som finns i litteraturen av de olika metoderna för exponentiell utjämning och ARIMA kan man förvänta sig att modellerna presterar bättre eller sämre i olika situationer. De tidsserier vi utgick från presenterar olika utmaningar för modellerna. Tidsserierna verkar uppvisa additiv eller multiplikativ säsong, trender som är starkare eller svagare och även ett variansmönster som är oregelbundet.

Utifrån vår teoretiska förståelse förväntade vi oss att additiv Holt-Winter skulle prestera bättre än multiplikativ Holt Winter på serie 4033 men sämre på serie 2224. Vi förväntade oss ett likvärdigt resultat dem emellan på serie 3503 och ett visst övertag för den multiplikativa varianten på serie 920. Vi förväntade oss att Holts metod skulle prestera dåligt på alla serier, och vara väsentligt mycket sämre än de andra modellerna. ARIMA modellerna förväntade vi oss som högpresterande i samtliga fall med tanke på dess flexibilitet och i ungefärlig paritet med Holt-Winters båda metoder.

Vi kan konstatera att våra uppfattningar angående skillnaden på multiplikativ Holt-Winter och additiv stämde överens grovt. Den multiplikativa presterade märkvärt bättre på serie 2224 och var förövrigt likvärdig den additiva. Detta beror troligen på att seriernas karaktär inte entydigt följt ett additivt eller multiplikativt mönster.

Resultaten från ARIMA-modellen på serie 2224 är det som visar den största skillnaden mellan modellerna. Den anpassade ARIMA-modellen är här väldigt enkel och beroende enbart av säsongsförskjutningen. Anledningen till att den multiplikativa och den additiva Holt Winters inte får bättre resultat kan vara att säsongsskillnaden är kvadratisk eller att den minskande säsongsvariationen inte beror på seriens nivå utan snarare sjunker med tiden. Därmed kan en modell som inte gör antaganden om säsongsvariationens karaktär, utan istället gör prognoser utifrån enbart observerade värden, prestera mindre dåligt.

Multiplikativ Holt-Winter presterar minst dåligt på serie 920, som vi betraktar som den serie som är svårast att prognostisera. Att multiplikativ Holt-Winter här överträffar även den anpassade ARIMA modellen kan eventuellt förklaras på samma sätt som ovan. I ARIMA modellen beror prognosen på de föregående observationerna medan multiplikativ Holt-Winter teoretiskt förutsätter en multiplikativ funktion. Den underskattar därmed inte rörelsen.

Att de relativt enkla modellerna, det vill säga de exponentiella utjämningsmetoderna sammantaget, kan säga prestera bäst är inte helt förvånade med tanke på att Makradakis & Hibon (2000) konstaterar att mer sofistikerade modeller ofta presterar sämre än enklare modeller (s.460).

Vi förväntade oss att vi i alla fall skulle få några signifikanta resultat på det modifierade Diebold-Marianos test. Tvärtom var ingen av skillnaderna i kvadrerat fel i närheten av

att vara statistiskt signifikant. Huvudorsaken till detta är troligen att variansuttrycket för \bar{d} i samtliga fall var relativt stora.

6.2 Brister i studien

En uppenbar brist i uppsatsen är att vi inte kunnat genomföra det modifierade Diebold-Marianos test på feltermerna från halveringsmetodsprognoserna. Detta hade varit mycket intressant då resultatet visar relativt närliggande prestationer för ex-post-prognoserna. Denna brist kan delvis härledas till begränsningar i tillgänglig programvara. SAS har inte modifierade eller omodifierade Diebold-Marianos test som funktion. Uppsatsförfattarna har räknat ut teststatistikan på ex-post-prognoserna med hjälp av Microsoft Excel men bedömde det som alltför tidskrävande att göra detta för halveringsmetodsprognoserna. Inte heller SAP-APO-DP har tillgång till detta test.

Avseende datamaterialet finns den självklara invändningen att den var något begränsad. Visserligen kan man kalla datamängden för varje serie medelstor, om man jämför med underlaget till t.ex. Harvey m.fl. (1997), men det hade givetvis gått att göra en mer komplett analys med fler observationer tillgängliga.

Vår analys av ARIMA modellerna är behäftade med en viss osäkerhet. Vi har använt riktlinjer från litteraturen och även använt oss av olika funktioner i SAS för att definiera vilken modell vi tillämpar. Vi vill dock inte hävda att vi varit säkra på våra bedömningar, utan tvärtom. Modellvalsprocessen skulle kunna förfinas och även kanske breddas till annan programvara. En uppsats framlagd vid Statistiska institutionen vid Stockholms universitet använde till exempel programvaran *R-2.10.1* för att för att få vägledning till att välja ARIMA modell (Banefelt & Dehdari, 2010, s 9). Detta är något vi skulle ha kunnat testa.

En annan brist är den begränsade litteraturstudie vi har gjort. Liknande undersökningar som denna har gjorts förut, inte minst på Statistiska Institutionen vid Stockholm Universitet och om vi i ett tidigare skede tillgodogjort oss mer av detta hade vi kunnat forma en undersökning som i större utsträckning kompletterar, och bygger vidare på, dessa tidigare verk. Det hade också hjälpt uppsatsen om vi tillgodogjort oss mer litteratur angående ytterligare test för att avgöra vilken av konkurrerande prognosmodeller som presterar bäst.

6.3 Vidare analysmöjligheter – diskussion

Om målet med prognostiseringen är att göra punktskattningar är de modeller vi valt att studera funktionella. De ger relativt goda skattningar och de hanterar de viktigaste två faktorerna: trend och säsong. Däremot tillvaratar de inte information som kan finnas i fler variabler. Det finns ibland så kallad kannibalism, det ville säga att ökad försäljning av en Electrolux-produkt kan resultera i minskad försäljning av en annan. Därför kan det vara av intresse för Electrolux att skapa modeller som beror på även andra tidsserier, så kallade VARIMA modeller. Dessutom kan man föreställas sig påverkan från externa

faktorer, till exempel konjunkturläge. Då kan till exempel Konjunkturinstitutets bedömningar av BNP utveckling användas till att förbättra träffsäkerheten.

En annan aspekt som skulle tillföra information är skattning av konfidensintervall. För att göra dessa på ett optimalt sätt så är tillämpningen av olika ARCH/GARCH modeller intressant (Kirchgässner & Wolters, 2009, s 261).

Ett tredje spår för utvecklad analys av prognostisering av försäljning inom Electrolux produktgrupper är effekten av olika kombinationer av subjektiva prognoser och statistiska prognoser. Det finns forskning som visar att kombinationer av dessa kan förbättra träffsäkerheten, samt även under vilka förutsättningar som manuella justeringar av statistiska prognoser ger ett positivt utfall (se bl.a. Fildes, R. m.fl., 2009, eller Lawrence, M., 2006).

Slutligen skulle en fördjupad analys av hur bra metoderna prognostiserar på olika tidsavstånd kunna bringa klarhet ifall man ska variera modellvalet beroende på detta.

7. Referenser

Ashley, R. A. & Patterson, D. M., 2010. A Test Of The Garch(1, 1) Specification For Daily Stock Returns, *Macroeconomic Dynamics*, Cambridge University Press, 14, 137-144.

Banefelt, J. Dehdari, S (2010). *En komparativ studie av olika prognosmetoder tillämpad på förhållandet mellan urbanisering och tillväxt*, Statistiska Institutionen, 15 - högskolepoängsuppsats inom Statistik III, Stockholm Universitet

Clements, M. P., Hendry, D.F. (1998), *Forecasting economic time series*, Cambridge University Press, Cambridge, United Kingdom

Diebold, F.X. och Mariano, R. (1995). Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, 13, 253-265.

Fildes, R. Goodwin, P. och Lawrence, M. (2009). Effective forecasting and judgemental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting*, 25, 3-23.

Gujarati, D. N., Porter, D. C., (2009). *Basic Econometrics*, 5th ed. New York: McGraw-Hill

Harvey, D. Leybourne, S. och Newbold, Paul. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13, 281-291.

Kirchgässner, G. och Wolters, J. (2008) *Introduction to modern time series analysis*, Berlin: Springer

Lawrence, M. Goodwin, P. O'Connor, M. och Önkal, D. (2006). Judgmental forecasting: A review of progress over the last 25 years. *International Journal of Forecasting*, 25, 493-518.

Makridakis, S., Hibon, M, (2000), The M3-Competition: results, conclusions and implications, *International Journal of Forecasting*, 16, 451-476.

Montgomery, D. C., Jennings, C. L. and Kulachi, M. (2008) *Introduction to Time Series Analysis and forecasting*. New Jersey: John Wiley & Sons

SAS (2010a) hämtad från 16 december 2010 från
<http://support.sas.com/documentation/cdl/en/etsug/63348/HTML/default/viewer.htm#etsug_tffordet_sect010.htm>

SAS (2010b) hämtad från 16 december 2010 från
<http://support.sas.com/documentation/cdl/en/etsug/63348/HTML/default/viewer.htm#etsug_esm_sect013.htm>

SAS (2010c) hämtad från 16 december 2010 från
<http://support.sas.com/documentation/cdl/en/etsug/60372/HTML/default/etsug_tffordet_sect013.htm>

Appendix 1: Dickey-Fullers utökade test (ADF)

ADF testar existensen av en enhetsrot. En tidserie kan beskrivas som:

$$(A.1.1) \quad y_t = \rho y_{t-1} + \varepsilon_t; \quad -1 < \rho < 1$$

där ε_t är vitt brus. Om $\rho = 1$, så har vi en enhetsrot och tidsserien är en slumpvandring, en icke-stationär stokastisk process. Ett flertal enhetsrotttest utgår ifrån att om vi i närvaro av enhetsrot borde kunna anpassa en regressionlinje mellan y_t och y_{t-1} och se ifall lutningskoefficienten för y_{t-1} är lika med 1. Om ja, är processen icke-stationär. För att undvika vissa problem utförs regression på första differensen av y_t . Dickey-Fullers utökade test tar hänsyn till möjlig autokorrelation hos feltermerna. Det utgår från följande regression:

$$(A.1.2) \quad (1 - B)y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \sum_{i=1}^m \alpha_i (1 - B)y_{t-i} + \varepsilon_t$$

De m stycken laggade differenserna åsyftar att plocka upp autokorrelation för att säkerställa att feltermen är vitt brus. Testet har följande hypoteser:

$$H_0 : \quad \delta = 0$$

$$H_1 : \quad \delta < 0$$

Nollhypotesen innebär att serien ska vara icke-stationär eller har en stokastisk trend. När den förkastas är den istället stationär, möjligtvis kring en deterministisk trend. När nollhypotesen förkastas kan man använda ett ensidigt t-test (Gujurati och Porter, 2009, s. 755-758).

Appendix 2: Teoretiskt uppträdande av autokorrelationsfunktioner och partiella autokorrelationsfunktioner

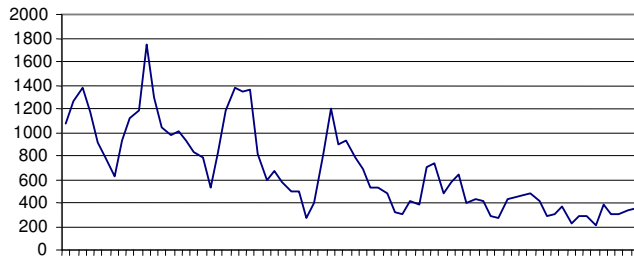
Tabell 7.

Enligt Montgomery m.fl. (2008) sid 256 och Kirchgässner & Wolters (2009) sid 67

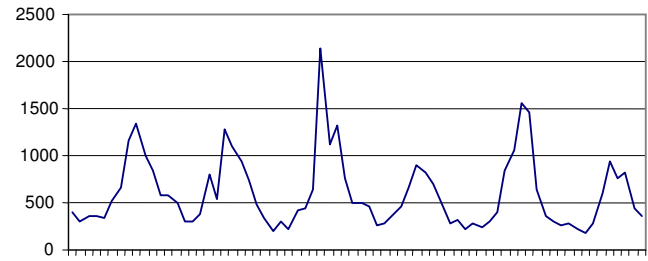
<i>Modell</i>	<i>ACF</i>	<i>PACF</i>
MA(q)	Slutar efter lagg (q)	Exponentiellt avtagande och eller en dämpadesinuskurva
AR(p)	Exponentiellt avtagande och eller en dämpad sinuskurva	Slutar efter lagg (p)
ARMA (p,q)	Exponentiellt avtagande och eller en dämpad sinuskurva	Exponentiellt avtagande och eller en dämpad sinuskurva

Appendix 3: Grafisk representation av våra tidsserier.

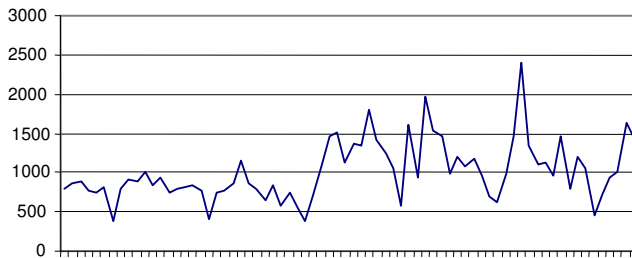
2224 Fristående spisar



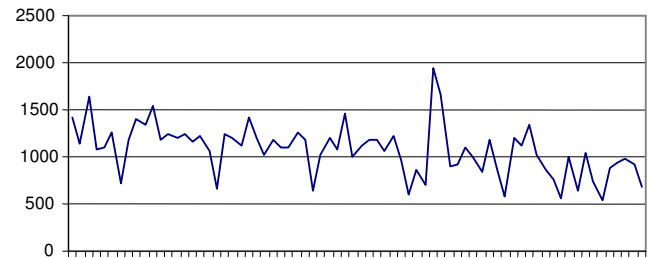
3503 Frysack



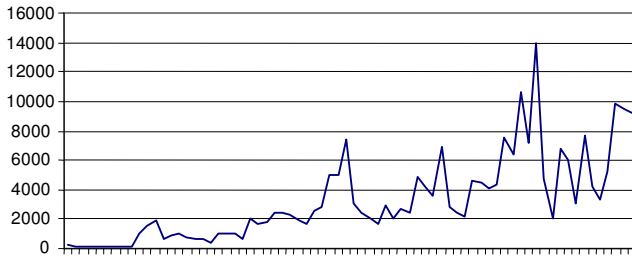
4022 Tillbehör till spisar



4033 Tillbehör till köksfläktar



920 Handhållna golvdammsugare



Appendix 4: Korrelations- och residualplotar

