

Die Absatzprognose einer Biermarke

Jul Martens

Institut für Statistik, Ökonometrie und Operations Research
Karl-Franzens-Universität Graz

Zusammenfassung: Die Absatzprognose einer Biermarke stellt die Anwendung eines saisonalen ARIMA-Modells dar, bei dem auch externe Effekte berücksichtigt werden. Dieses Modell wird aufgrund der vorliegenden Daten konstruiert und im weiteren Verlauf mit sparsameren Modellen verglichen. Die Prognosen haben sich in der Praxis als sehr gut erwiesen.

Abstract: The paper deals with the forecasting of sales data of a beer brand with seasonal ARIMA-Models, considering external effects. The construction of the model is applied to the data and afterwards compared to parsimonious models. The forecasts proved to be valid in the reality.

Schlüsselwörter: Box-Jenkins-Verfahren, ARIMA-Modelle, Absatzprognosen, Transfer-Funktions-Modelle, Intervention Analysis.

1 Einleitung

Das Ziel des vorliegenden Aufsatzes ist es, mit Hilfe von Transfer-Funktions-Modellen den Absatz einer Biermarke unter Berücksichtigung externer Effekte zu prognostizieren. Obgleich der Zeitraum dieser Prognose bereits mehr als 2 Jahre zurückliegt, sollen hier kurz die Ergebnisse und Erfahrungen wiedergegeben werden. Leider ist es aus Gründen der Diskretion und Sensibilität unternehmensinterner Daten aus der Praxis kaum möglich, aktuelle Ergebnisse in dieser Form zu präsentieren. Auf der anderen Seite liegt natürlich ein großer Reiz darin, theoretische Modelle der Zeitreihenanalyse auf reale Zeitreihen aus der Praxis anzuwenden, und die dabei auftretenden Probleme zu untersuchen.

Die Transfer-Funktions-Modelle bauen auf dem Verfahren von BOX und JENKINS auf, mit dessen Hilfe sich flexible Prognosemodelle für verschiedene Anwendungsgebiete individuell formulieren lassen (Pancratz, 1991). Für die ausführliche Darstellung des Verfahrens von BOX und JENKINS und der verschiedenen univariaten, multivariaten, saisonalen und nicht-saisonalen ARIMA-Modelle sei auf die entsprechende Literatur verwiesen (Box und Jenkins, 1976).

Das vorliegende Untersuchungsobjekt ist das Absatzprodukt Bier. Es gehört zu der Gruppe der Erfrischungsgetränke und ist durch eine begrenzte Haltbarkeit und

Lagerfähigkeit gekennzeichnet.¹ Die kurze Zeitspanne zwischen Produktion und Absatz machen nur eine kurze Prognose über wenige vorausliegende Monate für die Produktionsplanung und Absatzplanung notwendig und sinnvoll.

2 Die Untersuchung der Absatzzahlen

Bei der vorliegenden Zeitreihe handelt es sich um die monatlich erfaßten Absatzzahlen in Hektoliter einer Österreichischen Biermarke, die aus Gründen der Anonymität im folgenden mit X bezeichnet werden soll. Die Daten stammen aus einem Zeitraum vom 1. Jänner 1987 bis zum 30. Juni 1995. Die Absatzzahlen beziehen sich ausschließlich auf den Österreichischen Inlandsmarkt. Insgesamt handelt es sich um 102 äquidistante Zeitreihenwerte. Die sich daraus ergebende Problematik der relativ kleinen Stichprobe hat ebenfalls zur Folge, daß die Prognose nur über einen kurzen Prognosehorizont sinnvoll ist und sehr stark von der Stabilität der Zeitreihe abhängig ist.

Die Abbildung 1 zeigt die Zeitreihe der 102 monatlichen Absatzzahlen der Biermarke X auf dem Österreichischen Markt vom 1. Jänner 1987 bis zum 30. Juni 1995.

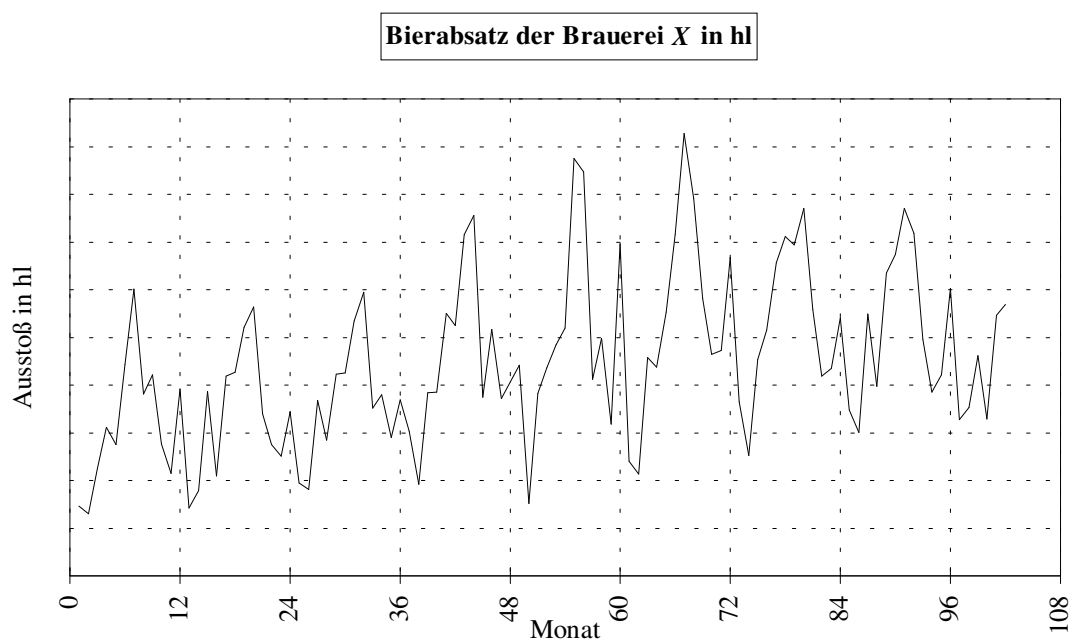


Abbildung 1: Monatliche Absatzzahlen der Marke X zwischen dem 1. Jänner 1987 und dem 30. Juni 1995

¹ In der Regel beträgt diese sechs Monate. Bier ist zwar noch über einen Zeitraum von sechs Monaten hinaus genießbar, jedoch ist dann mit einem Qualitätsverlust zu rechnen.

Der Absatz dieser Biermarke hat sich in den Jahren 1990, 1991 und 1992 stark erhöht. Die Amplituden in den Sommermonaten der Jahre 1990 und 1991 sind gegenüber denen der Vorperioden deutlich angestiegen. Die hohen Amplituden im Sommer legen die Vermutung nahe, das Wetter als grundlegende Ursache heranzuziehen. In diesen Jahren gab es überdurchschnittlich warme Sommermonate. Im Sommer 1991 gab es in Österreich eine große Hitzewelle, was sich in dem hohen Absatz im Monat August 1991 widerspiegelt. Bemerkenswert ist der sehr starke positive Ausschlag des Zeitreihenwertes Y_{60} (Dezember 1991). Hierbei ist als Erklärung anzumerken, daß im Jänner 1992 die Biersteuer in Österreich verdreifacht wurde, was eine Erhöhung des Bierpreises zur Folge hatte. Aufgrund der Vorankündigung im Jahre 1991 haben sich daraufhin insbesondere gewerbliche Großkunden Ende 1991 mit umfangreichen Vorräten eingedeckt, die dann in den folgenden Monaten abverkauft wurden. Der Grund für den verhältnismäßig hohen Absatz im Dezember 1992 dürfte in dem Bestellverhalten der Abnehmer liegen, die sich in der Höhe ihrer Bestellmengen offensichtlich im Sinne einer Naiven Prognose an denen der Vorperiode orientieren. Diese Monate stellen dennoch Ausreißerwerte dar, die aufgrund des Hintergrundes der Steuererhöhung und des (naiven) Bestellverhaltens des Handels eine besondere Berücksichtigung im Modell erfahren werden. Das erste Halbjahr 1995 führt den stagnierenden Absatz des Jahres 1994 fort, was der allgemeinen Entwicklung am gesamten Österreichischen Biermarkt entspricht.

3 Die Konstruktion des Modells

Ein bereits ex ante vermuteter Zusammenhang dieser Zeitreihe mit den Temperaturen wurde bereits mehrfach aus den Abbildungen offenkundig. Zur Überprüfung dieser Vermutung werden die durchschnittlichen Temperaturen für vier Regionen Österreichs zwischen Jänner 1987 und Juni 1995 herangezogen. Diese werden im weiteren mit einer regionalen Absatzkennzahl und der Bevölkerungsdichte in den entsprechenden Absatzgebieten gewichtet. Durch die Gewichtung mit der Bevölkerungsdichte wird der Tatsache Rechnung getragen, daß der überwiegende Teil der Bevölkerung Österreichs in den niedrigeren Lagen lebt, wodurch die Temperaturen dieser Lagen ein entsprechend höheres Gewicht bekommen. Für den Absatz der Biermarke X hat sich weiterhin gezeigt, daß die vier Regionen eine unterschiedliche Bedeutung für den Umsatz haben. Der Bedeutung dieser regionalen Umsätze entsprechend, werden die durchschnittlichen Temperaturen mit einer Umsatzmeßzahl gewichtet.

Lediglich die Dezembermonate der Jahre 1991, 1992, 1993 und 1994 stellen noch außergewöhnliche Ausreißer dar. Hier zeigt sich, daß es zwar aufgrund der Feiertage zu einem erhöhten Absatz kommt, die durchschnittlichen Temperaturen jedoch ihren niedrigsten Punkt erreichen, während in den anderen Monaten ein tendenzieller Absatzrückgang mit einem Temperaturrückgang verbunden ist. Der Grund dürfte die angekündigte Erhöhung der Biersteuer im Jahr 1991 sein, die zu Vorratskäufen besonders des umsatzstarken Großhandels und damit zu einem verstärkten Absatz im

vorausgegangenem Monat geführt hat. Dieser Effekt ist im Dezember 1992 in Folge des Bestellverhaltens der entsprechenden Großabnehmer ebenfalls aufgetreten, wie auch im Dezember 1993 und im Dezember 1994. Das Bestellsystem vieler umsatzstarker Großhändler orientiert sich offenkundig an der Vorperiode und reagiert auf Veränderungen in der Struktur entsprechend träge. Ob dieser Effekt eine im weiteren Verlauf nachlassende Wirkung hat, erscheint angesichts des Dezembermonats 1994 fraglich. Daher wollen wir hier nicht von einem Steuereffekt, sondern von einem *Dezembereffekt* sprechen. So soll dieser Effekt in Form einer Dummy-Variablen in der Modellgleichung Berücksichtigung finden.

Zur Identifikation der Ordnung der zugrunde liegenden ARMA-Prozesse der abhängigen Zeitreihe des Bierabsatzes zwischen den Jahren 1987 und 1994 werden die empirische Autokorrelationsfunktion $r(k)$ und die empirische partielle Autokorrelationsfunktion $r_p(k)$ der transformierten Untersuchungsschichtprobe der 96 Zeitreihendaten bis zum Lag $K = 20$ analysiert. Zur Transformation der Zeitreihe in eine schwach stationäre wurde, um Trend und Saison zu bereinigen, die 1. bzw. die 1. saisonale Differenz der Zeitreihe gebildet. Zur Identifikation und Messung des Einflusses der externen Zeitreihen Temperatur und Dezembereffekt wurden Scatterplots und Kreuzkorrelogramme verwendet.

Nach umfangreicher Prüfung wurde ein ARIMA(2,1,1)(1,1,1)₁₂-Modell mit den fünf ARMA-Prozeßparametern, f_1 , f_2 , q_1 , F_1 , und Q_1 spezifiziert:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \Phi_1 Y_{t-12} + u_t - \theta_1 u_{t-1} - \Theta_1 u_{t-12} + \omega_1 \xi_{t_1} + \omega_2 \xi_{t_2} . \quad (1)$$

Die beiden Prozeßparameter f_1 und f_2 ergeben sich aus der Ordnung $p = 2$ des autoregressiven Prozesses; der Parameter q_1 ergibt sich aus der Ordnung $q = 1$ des Moving-Average-Prozesses. Mit diesen Größen wird die Abhängigkeit der Zeitreihe von vorausliegenden Werten und von stochastischen Störgrößen berücksichtigt. Der starke saisonale Einfluß wird durch die Prozeßparameter F_1 und Q_1 berücksichtigt. Der Vergleich dieser Parameterkombination mit *sparsameren* Modellen hat die Festlegung dieser Parameterkombination ex post bestätigt. Die weiteren Zeitreihen ξ_{t_1} und ξ_{t_2} werden durch die Parameter w_1 und w_2 berücksichtigt. Der Ablauf der Prognose wird durch Abbildung 2 noch einmal verdeutlicht.

Als Untersuchungsschichtprobe (*out-of-sample*) dienen zunächst 96 Werte vom 31. Jänner 1987 bis um 31. Dezember 1994. Dazu wurden vorab die Prognosewerte der gewichteten Temperaturdaten als *Leading-Indicator* verwendet. Der Steuereffekt erhält als Dummy-Variable jeweils die Gewichtung 1. Die endgültige Prognose erfolgt mit Hilfe der gesamten Stichprobe der 102 Absatzdaten und der gewichteten Temperaturdaten.

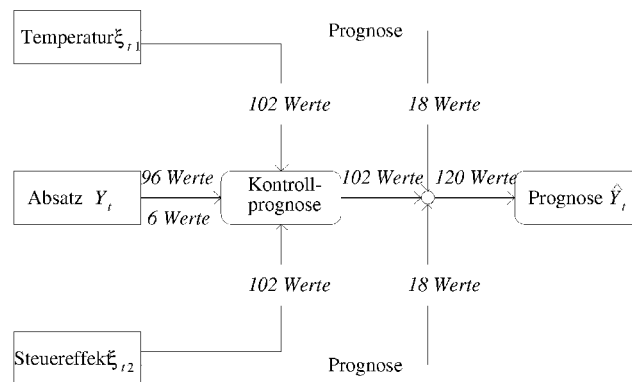


Abbildung 2: Schema zur Vorgangsweise bei der Prognose

4 Die Überprüfung des Modells

Wesentliches Kriterium für ein fehlerfrei geschätztes Modell ist die Annahme normalverteilter Residuen mit nur mehr zufälligen Schwankungen (*white noise*). Die Funktion der partiellen Autokorrelationskoeffizienten der Residuen, auf die an dieser Stelle nicht näher eingegangen wird, haben keinen Hinweis mehr auf eine mögliche systematische Abhängigkeit der Störgrößen gegeben. Die üblichen Tests der Residuen von BOX/PIERCE und BOX/LJUNG konnten ebenfalls keine signifikante Autokorrelation der Residuen mehr nachweisen. Wir betrachten die eingangs gezeigten Absatzzahlen und das geschätzte Modell (siehe Abbildung 3).

Das Sommerloch 1993 mit seinem außergewöhnlich kühlen und feuchten Sommer läßt sich trotz der Berücksichtigung der Temperaturdaten offensichtlich nicht exakt erfassen. Es können somit auch andere, nicht erklärbare Faktoren wie etwa Niedrigpreisaktionen von Konkurrenten für diesen einmaligen Absatzeinbruch verantwortlich sein.

Es wurden neben dem vorliegenden ARIMA(2,1,1)(1,1,1)₁₂-Modell mit zwei unabhängigen Variablen (Temperatur und Steuereffekt), das im folgenden mit ARIMA_2 bezeichnet werden soll, die folgenden, in der Tabelle 1 gezeigten Modelle mit der gleichen Stichprobe geschätzt und im Hinblick auf die Güte der Schätzung bzw. der Prognosefähigkeit überprüft. Die beiden alternativen Modelle enthalten neben den Absatzdaten nur die Temperatur als weitere unabhängige Zeitreihe (ARIMA_1) bzw. keine weitere Inputvariable (ARIMA). Diese beiden *sparsameren* Modelle wurden nach der gleichen oben skizzierten Vorgangsweise geschätzt.

Die Bezeichnung *Schätzung* bezieht sich auf den Schätzzeitraum mit den 96 Werten, und die Bezeichnung *Prognose* steht für die Abweichung der 6 mit dem Modell prognostizierten Werte und der 6 tatsächlichen Werte. Die Notation enthält in der Extension jeweils die Anzahl der unabhängigen Zeitreihen.

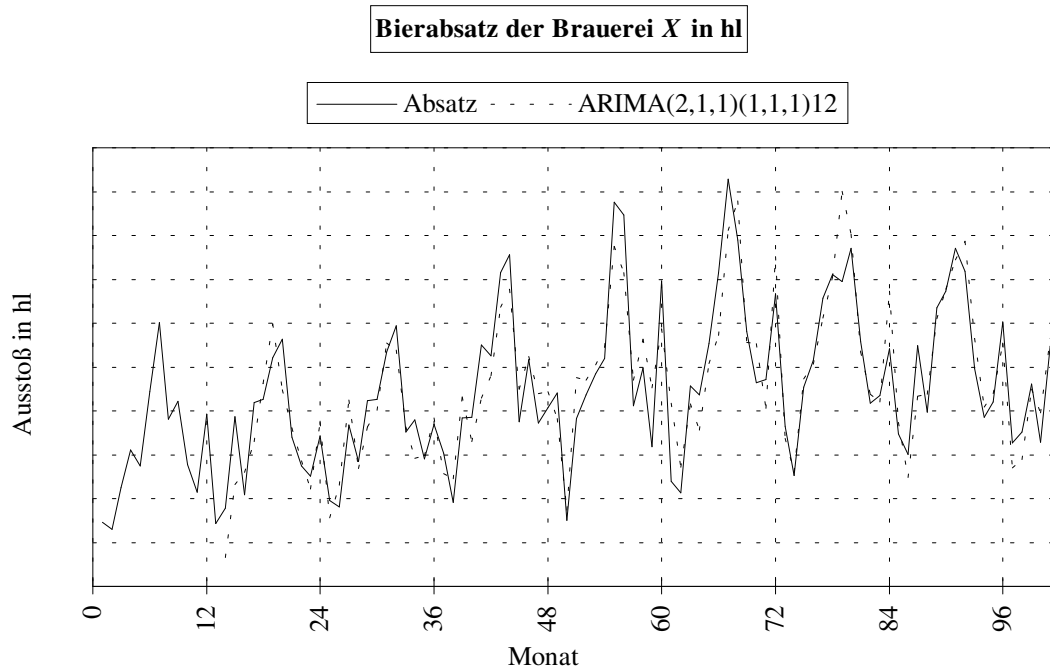


Abbildung 3: Absatz und ARIMA(2,1,1)(1,1,1)₁₂-Modell für den Zeitraum 1. Jänner 1987 bis 30. Juni 1995

Als Prognosefehlermaße wurden hier der Mittlere Absolute Prozentuale Fehler (*MAPE*), der Standardfehler der Residuen (*SE*) und der THEILSCHE Ungleichkoeffizient (*TU*) für die Schätzwerte und für die Prognosewerte ausgewählt. Der Mittlere Absolute Prozentuale Fehler (*MAPE*) setzt die geschätzten Werte in Beziehung zu den realen Werten in Form der Absolutbeträge, korrigiert um die Anzahl der Freiheitsgrade. Dieser Wert kann als Prozentzahl interpretiert werden. Der Standardfehler der Residuen (*SE*) entspricht der um die Freiheitsgrade bereinigten Standardabweichung zwischen Schätzwerten und den Beobachtungswerten. Der THEILSCHE Ungleichkoeffizient stellt die Prognosewerte des Modells in Relation zu den Werten der Naiven Prognose:

$$TU = \sqrt{\frac{\sum_{t=1}^T (Y_t - \hat{Y}_t)^2}{\sum_{t=1}^T (Y_t - Y_{t-12})^2}} \quad \text{mit } 0 \leq TU \leq 1.^2 \quad (2)$$

Für alle diese drei Fehlermaße weist das ARIMA(2,1,1)(1,1,1)₁₂-Modell mit zwei unabhängigen Variablen die jeweils günstigsten Werte auf. Das gilt sowohl für die 96 Schätzwerte, als auch für die 6 Prognosewerte (siehe Tabelle 1).

² Ein *TU* von > 1 ist ebenfalls denkbar, bedeutet aber eine Überlegenheit der Naiven Prognose gegenüber dem Prognosemodell.

Tabelle 1: Prognosefehlermaße für die Schätzwerte und für die Prognosewerte für alternative $ARIMA(2,1,1)(1,1,1)_{12}$ -Modelle

Fehlermaß		Modell	ARIMA	ARIMA_1	ARIMA_2
<i>MAPE</i> (in Prozent)	Schätzung		6,40	6,06	6,01
	Prognose		6,47	5,22	4,76
<i>SE</i> (in <i>hl</i>)	Schätzung		8.131,02	7.716,44	7.520,19
	Prognose		6.659,62	5.117,05	4.717,81
<i>TU</i>	Schätzung		0,84	0,79	0,76
	Prognose		0,88	0,68	0,62

Es wird insbesondere am THEILSCHEN Ungleichskoeffizient deutlich, wie sich die Berücksichtigung weiterer unabhängiger Variablen auf die Vorteilhaftigkeit der Modelle auswirkt (Abbildung 4).

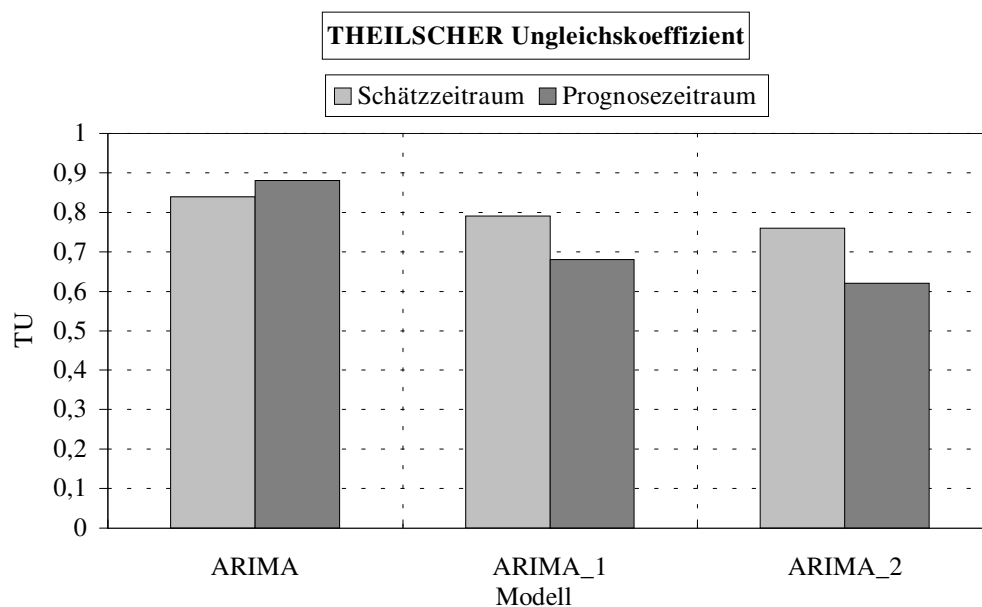


Abbildung 4: THEILSCHER Ungleichskoeffizient für die ARIMA-Modelle

5 Die Prognose für die Jahre 1995 und 1996

Die folgende Abbildung zeigt die monatliche Prognose für das Jahr 1995, für die alle 102 Werte der Ausgangsstichprobe verwendet wurden. Dazu sind wieder die gewichteten Temperaturdaten, jedoch keine Nachwirkung des Steuereffektes berücksichtigt.

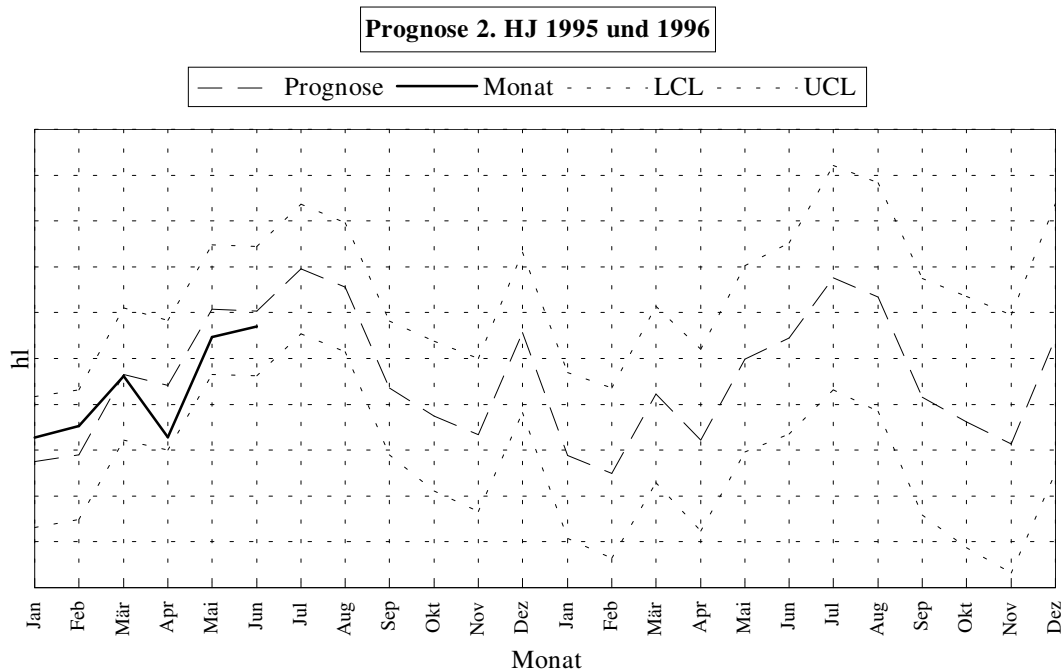


Abbildung 5: Absatz, Prognosewerte und Konfidenzintervall für die Prognose (LCL bzw. UCL) für das Jahr 1995 und 1996

Abbildung 5 zeigt die Prognose der Jahre 1995 und 1996 als strichlierte Linie. Zusätzlich ist ein 95%-Konfidenzintervall durch die gepunkteten Linien eingetragen. Die realen Werte des 1. Halbjahres 1995 sind ebenfalls als durchgezogene Linie in dem Diagramm eingetragen. Das verwendete ARIMA-Modell unterschätzt zunächst die Monate Jänner und Februar, trifft den Wert für März beinahe exakt und überschätzt die Monate April bis Juni 1995. Der Absatz wird sich nach dieser Darstellung in den folgenden Monaten stabilisieren, aber im weiteren Zeitablauf sinken.

Längerfristig prognostiziert das Modell einen Absatzrückgang dieser Biermarke, was sich analog zur gesamten Entwicklung des Bierabsatzes auf dem Österreichischen Markt verhält. Diese Entwicklung könnte auch durch die Anpassung eines quadratischen Trends an die gesamte Zeitreihe erkennbar gemacht werden.

6 Zusammenfassung und Ausblick

Der Aufsatz hat die Möglichkeit gezeigt, mit dem erweiterten Verfahren von BOX und JENKINS ein Modell zur kurzfristigen Absatzprognose zu erstellen und es sinnvoll einzusetzen. Obgleich die Prognose über einen Prognosehorizont von 6 Schritten hinaus an Aussagekraft verliert, antizipiert sie eindrucksvoll mögliche Entwicklungen.

Zu einer wachsenden Ungenauigkeit durch die Notwendigkeit der Prognose der externen Zeitreihen für die Erstellung der Prognose der abhängigen Reihe ist es nicht gekommen. Zur Messung des Einflusses dieser externen Größen haben sich Scatterplots und Kreuzkorrelationsanalysen als hilfreich erwiesen.

Für die aggregierten Absatzmengen eines Jahres ergaben sich bei dem verwendeten Prognosemodell jährliche Abweichungen von 0,1% Unterschied zur tatsächlich erzielten Absatzmenge. Auch dies spricht für das Modell.

Nicht berücksichtigen konnte das Modell unternehmenspolitische Entscheidungen und globale Veränderungen, die sich mittlerweile auf den Absatz der betrachteten Zeitreihe ausgewirkt haben.

Die Übernachtungszahlen als Indikator für den Tourismus wurden in dem Modell ebenfalls nicht berücksichtigt, weil hier weder ein direkter, noch ein zeitverschobener Zusammenhang nachgewiesen werden konnte.

Weitere externe Effekte wie etwa Niedrigpreisaktionen von Mitbewerbern oder auch interne Größen, wie "*Kanibalisierungseffekte*" mit eigenen Marken und Sonderpreisaktionen wären denkbare Erweiterungen dieses Modells. Lediglich die Wartung eines derartigen Prognosemodells erfordert Fingerspitzengefühl und Erfahrung im Umgang mit dieser Technik.

Literatur

- G.E.P. Box and G.M. Jenkins. *Time Series Analysis*. Holden Day, San Francisco, 2nd Edition, 1976.
- J. Hartung, B. Elpelt und K.-H. Klösener. *Statistik*. Oldenbourg, München, 1995.
- M. Hüttner. Absatzprognosen nach Box-Jenkins. *Marketing - Zeitschrift für Forschung und Praxis*, (1): 37-46, 1981.
- A. Pankratz. *Forecasting with Dynamic Regression Models*. New York, 1991.
- R. Schlittgen und B. Streitberg, *Zeitreihenanalyse*. Oldenbourg, München, 1991.

Adresse des Autors:

Mag. Dr. Jul Martens
Institut für Statistik, Ökonometrie und Operations Research
Karl-Franzens-Universität Graz
Universitätsstraße 15/E3
A-8010 Graz

Tel. +43 / 316 / 380 34 90

Fax +43 / 316 / 380 95 60