

Die Kombination von Prognosen mit SAS

Dietrich Eherler

Friedrich-Schiller-Universität Jena,
Lehrstuhl für Wirtschafts- und Sozialstatistik,
Carl-Zeiss-Straße 3
07743 Jena
e-mail: d.eherler@wiwi.uni-jena.de
Tel. 03641 94 33 04

Abstract

Diese Präsentation ist eine Realisierung der Idee der Kombination von Prognosen mittels SAS. Es läßt sich zeigen, daß durch eine Mischung von einzelnen Prognosen zu einer kombinierten eine Verbesserung der Prognosegüte zu erreichen ist, beispielsweise gemessen bezüglich der Prognosefehlervarianz bei erwartungstreuen Prognosefunktionen. Dem Benutzer eines solchen Kombinationsansatzes wird somit eine zuverlässigere und zusätzliche Informationen ausnützende Entscheidungsgrundlage geboten. Verschiedene Ansätze zur Kombination von Prognosen sind realisiert.

1. Theoretische Aspekte

Mit der Publikation des Artikels „*The Combination of Forecasts*“ lösten Bates und Granger (1969) eine rege Diskussion aus, deren Gegenstand es ist, durch Kombination von quantitativen Prognosemethoden zu besseren Vorhersagen zu kommen. Stehen mehrere Prognosen für die interessierende Größe zur Verfügung, so ist eine Auswahl für nur eine der verschiedenen Prognosen und die Ablehnung der anderen „... *not a wise procedure if the objective is to make as good a forecast as possible, since the discarded forecast always contains some useful independent information*“ (Bates/Granger, 1969, S. 451).

Die Motivation zur Kombination von Prognosen stützt sich also auf die grundsätzliche und einfache Idee, durch Aggregation von Informationen eine Reduzierung der Unsicherheit¹ sowie einer Erhöhung der Güte von Prognosen zu erreichen. Ein solches Vorgehen hat Parallelen z.B. in der Portefoliotheorie oder auch in der klassischen Statistik, in der eine Vergrößerung des Stichprobenumfangs stets als wünschenswert angesehen wird (Winkler, 1989, S. 606; Makridakis, 1989, S. 601).

Definition 1.1: Seien $f_{1,T+1}, \dots, f_{m,T+1}$ Prognosefunktionen für die zu prognostizierende Größe X_{T+1} . Dann ist eine kombinierte Prognosefunktion $f_{C,T+1}$ eine Verschmelzung der einzelnen Prognosefunktionen f_1, \dots, f_m zu

$$f_{C,T+1} = f_{C,T+1}(f_{1,T+1}, \dots, f_{m,T+1}). \rightsquigarrow$$

Unmittelbar stellt sich nun die Frage, wie eine solche Kombination vorzustellen ist, d.h., welche Struktur $f_{C,T+1}$ als Abbildung hat. Im allgemeinen wird die kombinierte Prognose als lineare Abbildung

$$\hat{f}_{C,T+1} = w_0 + w_1 \hat{f}_{1,T+1} + \dots + w_m \hat{f}_{m,T+1} \quad (Gl.1)$$

¹ Reduzierung der Unsicherheit: Im voraus ist nicht bekannt, welches Prognoseverfahren denn nun die beste Prognose liefert. Diese Unsicherheit soll durch eine Kombination reduziert werden (Gupta/Wilton, 1987, S. 368).

der Einzelprognosen $\hat{f}_{1,T+1}, \dots, \hat{f}_{m,T+1}$ angenommen mit den Gewichten w_i aus den reellen Zahlen. Meist werden für die Gewichte noch zusätzliche Bedingungen gestellt. Darauf wird beim Betrachten der einzelnen Ansätze näher eingegangen, ebenso wie auf die eigentliche Bestimmung der Gewichte. Die Einschränkung auf lineare Abbildungen ist keinesfalls zwingend. Weitere Gattungen von Abbildungen sowie die Bildung einer Mischung mit KI-Methoden sind durchaus denkbar.

Thiele (1993) hat die verschiedenen Ansätze zur Kombination systematisiert in

- ungewogene Kombination
- Kovarianzverfahren
- Regressionsverfahren
- Wahrscheinlichkeitsverfahren.

1.1 Ungewogene Kombination

Die Entdeckung verbesserter Prognosen durch eine ungewogene Kombination war Auslöser intensiver Studien zur Kombination von Prognosen. Doch auch nach empirischen Untersuchungen darauf entwickelter komplizierterer Kombinationsmethoden konnte Clemen (1989, S. 566) bestätigen, daß „*in many studies, the simple average of individual forecasts has performed best or almost best*“.

Die kombinierte ungewogene Prognose errechnet sich schlicht aus dem arithmetischen Mittel der Einzelprognosen

$$\hat{f}_{C,T+1} = \frac{\hat{f}_{1,T+1} + \dots + \hat{f}_{m,T+1}}{m} \quad (Gl.2)$$

Die Anwendung der ungewogenen Kombinationsmethode berechtigt sich nicht nur aus der Empirie, sondern auch aus einem entscheidungstheoretischen Ansatz. Thiele (1993, S. 110-7) zeigt die Überlegenheit der ungewogenen Durchschnittsprognose über sämtliche verfügbaren Einzelprognosen sowie auch der Durchschnittsbildung einer beliebigen Auswahl von weniger Einzelprognosen, falls ein betroffener Entscheidungsträger über keine Information über die Verteilungen der einzelnen Prognosefunktionen verfügt. Ist dies der Fall, so stehen dem Prognostiker die *Auswahlstrategien* s_k ($1 \leq k \leq m$) zur Verfügung der *zufälligen* Auswahl Verfahren von k der m Einzelverfahren. Es gelten die zwei Theoreme (Thiele; 1993, S. 116-7; Beweise ebd.):

Theorem 1.1: Alle Strategien s_1, \dots, s_m haben dieselben Prognosefehlererwartungen. \hookrightarrow

Theorem 1.2: Die Prognosefehlervarianz der Strategien s_m ist kleiner oder gleich der Prognosefehlervarianz der zu erwartenden Prognosefehlervarianzen der Strategien s_1, \dots, s_{m-1} . \hookrightarrow

Ferner hat die ungewogene Prognosefunktion die Eigenschaft mit allen weiteren Konvexkombinationen von Einzelprognosen, daß ihre prognostischen Qualitäten gemessen im Sinne des

MSE² nie unter die der schlechtesten Einzelprognosefunktion fällt ($\varepsilon_{i,T+1} := X_{T+1} - f_{i,T+1}$)

$$E[\varepsilon_{C,T+1}^2] \leq \max\{E[\varepsilon_{1,T+1}^2], \dots, E[\varepsilon_{m,T+1}^2]\} \quad (Gl.3)$$

und somit eine Absicherung nach unten besteht (Thiele, 1993, S. 120).

1.2 Kovarianzverfahren

Die Kovarianzverfahren gehen allesamt von einer linearen Kombination der Einzelprognosen aus. Das zugrundeliegende Modell ist die Menge

$$F = \{f_{C,T+1} = w_1 f_{1,T+1} + \dots + w_m f_{m,T+1} \mid w_1 + \dots + w_m = 1; w_i \in \mathbf{R}; i = 1, \dots, m\}$$

der linearen Abbildungen. Es wird gefordert, daß die einzelnen Prognosen erwartungstreu sind. Der Prognosefehler der gemischten Prognose formuliert sich zu

$$\varepsilon_{C,T+1} = \sum_{i=1}^m w_i \varepsilon_{i,T+1} \quad (Gl.4)$$

mit der Varianz

$$\sigma_C^2 = \sum_{i=1}^m w_i^2 \sigma_i^2 + \sum_{i,j=1(i \neq j)}^m w_i w_j \sigma_{ij} \quad (Gl.5)$$

Bates und Granger (1969, S. 452-3) wählten zur Gewichtsbestimmung die naheliegende Methode die Gewichte so zu terminieren, daß die Fehlervarianz der kombinierten Prognose minimal wird. Die Zielvorstellung³ läßt sich schreiben als

$$\begin{aligned} w^T \Sigma w &\rightarrow \min \\ w^T l &= 1 \\ w_i &\in \mathbf{R} \end{aligned}$$

Ist die Kovarianzmatrix invertierbar, so ergibt sich der optimale Gewichtsvektor w^* zu

$$w^* = \frac{\Sigma^{-1} l}{l^T \Sigma^{-1} l} \quad (Gl.6)$$

und die optimale Kombination der Einzelprognosefunktionen ist (Thiele, 1993, S. 20-2)

$$f_{C,T+1}^* = w_1^* f_{1,T+1} + \dots + w_m^* f_{m,T+1} \quad (Gl.7)$$

² Diese Aussage gilt auch für den RMSE und den MAPE (Thiele, 1993, S. 120).

³ Wegen der Definitheit von Σ gilt der Existenzsatz für quadratische Programme (Neumann, 1993, S. 553-4). Ferner ist $l = (1, \dots, 1)$ ein Vektor mit Einsen entsprechender Dimension.

Es gelten die beiden Aussagen (Thiele, 1993, S. 22-4):

Theorem 1.3: Die Prognosefehlervarianz einer mit obigen Gewichten w^* gewonnenen Kombination ist mindestens so klein wie die kleinste Fehlervarianz der einzelnen Ausgangsprognosefunktionen. Es gilt also:

$$\sigma_C^2 \leq \min(\sigma_1^2, \dots, \sigma_m^2). \quad (Gl.8)$$

Theorem 1.4: Eine zusätzliche Ausgangsprognosefunktion macht die Fehlervarianz ${}_{(m+1)}\sigma_C^{*2}$ der neu gewonnenen Kombination aus $m+1$ Einzelprognosefunktionen mindestens so klein wie die Fehlervarianz ${}_{(m)}\sigma_C^{*2}$ der vorherigen Kombination mit einer Ausgangsprognosefunktion weniger, d.h.

$${}_{(m+1)}\sigma_C^{*2} \leq {}_{(m)}\sigma_C^{*2}. \quad (Gl.9)$$

Vor allem bei kurzen Zeitreihen ist es wegen Schätzprobleme der Kovarianzmatrix angebracht, eine „Zweckmäßigkeitssannahme“ (Thiele, 1993, S. 35) einzuführen, nämlich die der Unkorreliertheit der Prognosefehler der jeweils betrachteten Ausgangsprognosefunktionen.

1.2.1 Spezialfall unkorrelierter Prognosefehler

Um den sich aus dem Schätzen von Σ ergebenden Unzulänglichkeiten zu begegnen werde nun angenommen, die kontemporären Prognosefehler $\varepsilon_{i,t}$ der i -ten Einzelprognosefunktion seien unkorreliert zu den Prognosefehlern $\varepsilon_{j,t}$ der j -ten Einzelprognosefunktion, kurz

$$\text{Cov}[\varepsilon_{i,t}, \varepsilon_{j,t}] = 0 \quad i, j = 1, \dots, m \quad i \neq j; t = 1, \dots, T$$

Die Kovarianzmatrix der Prognosefehler vereinfacht sich dann zu einer Diagonalmatrix mit den Prognosefehlervarianzen der einzelnen Ausgangsprognosefunktionen als Diagonalelemente. Das Invertieren dieser Matrix entspricht gerade der Kehrwertbildung, und die optimalen Gewichte genügen der Gleichung

$$w_i^* = \frac{\frac{1}{\sigma_i}}{\sum_{j=1}^m \frac{1}{\sigma_j}} \quad i = 1, \dots, m \quad (Gl.10)$$

Anschaulich läßt sich die Gleichung interpretieren, daß der Prognosefunktion $f_{i,T+1}$ ein größeres Gewicht zukommt, falls sie eine kleinere Fehlervarianz aufweist und umgekehrt. Dies entspricht auch einer intuitiven Vorgehensweise. Außerdem sind die Gewichte stets zwischen $[0,1]$, die gemischte Prognose ist also immer Konvexkombination der jeweiligen Einzelprognosefunktionen (Thiele, 1993, S. 36).

1.3 Regressionsverfahren

Beim Regressionsansatz wird angenommen, daß die Einzelfunktionen f_1, \dots, f_m die zu prognostizierende Größe X_{T+1} erklären. Die Regression, also der bedingte Erwartungswert,

liefert dann die beste Approximation im Sinne der kleinsten quadratischen Abweichung (Cramér, 1966, S. 270-3).

Der allgemeine lineare Ansatz für eine kombinierte Prognose genügt also der Gleichung

$$f_{C,T+1} = w_0 + w_1 f_{1,T+1} + \dots + w_m f_{m,T+1} \quad (Gl.11)$$

Die Gewichte w_0, \dots, w_m ($w_j \in \mathbf{R}$; $j=1, \dots, m$) lassen sich als Lösung des Minimierungsproblems

$$E [(X_{T+1} - (w_0 + w_1 f_{1,T+1} + \dots + w_m f_{m,T+1}))^2] \rightarrow \min$$

bestimmen.

Zwischen dem Regressionsansatz und dem allgemeinen Kovarianzverfahren besteht der folgende Zusammenhang (Thiele, 1993, S. 56):

Theorem 1.5: Sei $(X_{T+1}, f_{1,T+1}, \dots, f_{m,T+1})$ ein $(m+1)$ dimensionaler Zufallsvektor mit der linearen Regressionsfunktion

$$E[X_{T+1} | f_{1,T+1}, \dots, f_{m,T+1}] = w_1 f_{1,T+1} + \dots + w_m f_{m,T+1}$$

mit $w^T l = 1$ mit $l = (1, \dots, 1)$. Dann ist diese Regressionsfunktion identisch mit der optimal kombinierten Prognosefunktion

$$f_{CT+1}^* = w_1^* f_{1,T+1} + \dots + w_m^* f_{m,T+1}. \quad (Gl.12)$$

$$\text{mit } w^* = \frac{\Sigma^{-1} l}{l^T \Sigma^{-1} l} \text{ gemäß Gl.6 und Gl.7. } \heartsuit$$

Eine Folgerung aus diesem Theorem ist, daß der allgemeine Regressionsansatz mit Absolutglied und ohne Parameterrestriktion eine noch bessere gemischte Prognosefunktion bezüglich der mittleren quadratischen Abweichung bestimmt. Ferner generiert der allgemeine Regressionsansatz eine unverzerrte kombinierte Prognosefunktion selbst bei verzerrten Einzelprognosefunktionen (Granger/Ramanathan, 1984, S. 201).

1.4 Wahrscheinlichkeitsverfahren

Die unter dem Begriff Wahrscheinlichkeitsverfahren zusammengefaßten Ansätze erhalten ihren Namen von der Vorgehensweise bei der Bestimmung der Gewichte, mit denen die Einzelprognosen in die gemischte Prognose eingehen. Den Gewichten der Ausgangsprognosen kommt eine direkte Wahrscheinlichkeitsinterpretation zu als *Übertreffenswahrscheinlichkeiten*. Die grundlegende Idee ist die Bestimmung der Gewichte nach der Leistung der einzelnen Verfahren bezüglich ihrer Prognosegenauigkeit. Bunn (1975) hat eine plausible Operationalisierung vorgeschlagen unter Verwendung von absoluten Prognosefehlern als Leistungsindikator. Eine Prognosefunktion $f_{i,T+1}$ übertrifft eine Prognose $f_{j,T+1}$ ($i, j=1, \dots, m$; $i \neq j$), wenn sie einen kleineren absoluten Ein-Schritt Prognosefehler aufweist. Daraus resultiert dann unmittelbar

$$w_i = p(|\varepsilon_{i,T+1}| < |\varepsilon_{j,T+1}|; j=1,\dots,m; i \neq j) \quad (\text{Gl. 13})$$

als Gewicht für die Prognose $f_{i,T+1}$ $i=1,\dots,m$ in der kombinierten Prognose

$$f_{C,T+1} = \sum_{i=1}^m w_i f_{i,T+1}. \quad (\text{Gl. 14})$$

Die konkrete Bestimmung dieser Übertreffenswahrscheinlichkeiten erfolgt durch Schätzung aus den Daten. Die in der Literatur bekannten Ansätze (Bunn, 1978; Gupta/Wilton, 1987, 1988) erfordern teilweise aufwendige Matrizenoperationen, deren Realisierung das SAS IML Paket beanspruchen würde. Ein stark vereinfachter Ansatz vom Verfasser (Eherler, 1996, S. 102) ermöglicht dagegen die Schätzung von globalen Übertreffenswahrscheinlichkeiten in einem DATA-Step. Basierend auf einem Bayes'schen Schätzansatz wird für jede Einzelprognose \hat{f}_i die Häufigkeit gezählt, wie oft diese ex-post die beste Ein-Schritt-Prognose geliefert hat. Das Gewicht w_i mit dem die Prognose \hat{f}_i in die Mischung eingeht ist dann gerade die relative Häufigkeit bezüglich aller erstellten Ein-Schritt-Prognosen.

2. Implementierungsaspekte

2.1 Einzelprognosen

Zur Umsetzung des Kombinationsansatzes mit SAS sollen zunächst die Einzelprognosen in eine Datei generiert werden. Insgesamt sechs verschiedene Einzelprognosen bilden dafür die Grundlage. Teils werden diese Einzelprognosen in einem DATA-Step erzeugt und teils mit den Prozeduren ARIMA und FORECAST bereitgestellt. Die einzelnen Prognoseverfahren sind:

- Naive Prognose
- Gewogenes Mittel vom Grad 3
- *AdaptiveResponseRateSingleExponentialSmoothing* (Lewandowski, 1974, S. 143-65)
- Exponentielles Glätten nach Brown (z.B. Gardner, 1985, S. 4)
- Wintersverfahren (Winters, 1960, S. 324-42)
- Box-Jenkins Verfahren (Box/Jenkins, 1976).

Es sei hier allerdings gleich angemerkt, daß bei der Anwendung der einzelnen Verfahren kein Schwerpunkt auf die Erzielung optimaler Prognoseergebnisse bezüglich der zu lösenden Prognoseaufgabe gelegt wurde. Dies bietet Anlaß zu Kritik, weil so möglicherweise Prognosegüte verschonkt wird. Entlastend kann dagegen eingebracht werden, daß

1. durch die Kombination bessere Ergebnisse zu erwarten sind bei geringerem Aufwand,
2. die optimale Anpassung eines Modells, auf dem ein Verfahren beruht zwar ex-post möglich ist, dies jedoch keinesfalls für dessen ex-ante Validität und Ädaquanz garantiert und
3. die optimale Anpassung auch teilweise schwerlich (Schlittgen/Streitberg, 1991, S. 227-8) und mit deutlichem Mehraufwand verbunden ist, vor allem dann, wenn umfangreiche Prognoseaufgaben zu lösen sind.

So wurden bei der konkreten Implementierung für die Initialisierung der Verfahren des exponentiellen Glättens jeweils die erste Realisation x_1 für den Grundwert und für Trendwert bzw. Fehlerkorrektur die Differenz der ersten beiden Realisationen $x_2 - x_1$ verwendet. Für die Glättungsparameter sind feste Werte im Programm vorgegeben. Ebenso sind die Gewichte für das gewogene Mittel einstellbar. Für die Box-Jenkins Verfahren fiel die Entscheidung generell einen ARIMA(1,1,2)-Prozeß zu schätzen.

DATA-Statements:

```
naiv = xt-1; *Naive Prognose, verzögerter Wert;

retain exsm x1; *Einfaches Exponentielles Glätten;
if t > 1 then exsm = alpha * xt-1 + (1-alpha) * exsm;

gewog = w1 * xt-1 + w2 * xt-2 + w3 * xt-3; *Gewogenes Mittel;

retain err_ars x2-x1; *ARSES;
retain arses x1;
if t > 1 then arses = alpha1 * xt-1 + (1-alpha1) * arses;
retain mad 0; *Steuerung des von alpha1;
if t > 2 then mad = ceta * abs(err_ars) + (1-ceta) * mad;
retain md 0;
if t > 2 then md = ceta * err_ars + (1-ceta) * md;
if t > 3 then alpa1 = abs (md) / mad;

retain brown x1; *Exponentielles Glätten nach Brown;
retain steig x2-x1;
brown_1 = lag(brown);
if t > 1 then steig = beta * (brown-brown_1) + (1-beta) * steig;
if t > 1 then brown = exsm + brown;
```

PROC-Statements:

```
*Wintersverfahren;
proc forecast method=winters seasons=12 trend=2 weight=(.2 .1 .3) out=wintout
    out1step;
var x;
run;
*Ergebnisdatei wintout vereinigen mit Hauptdatei;

*Box-Jenkins Verfahren;
data boxjen;
set Hauptdatei;
if t > T then delete; *ab Zeitpunkt T ex-ante Ein-Schritt-
Prognosen;
proc arima;
identify var = x(1) noprint;
estimate p=1 q=2 method = ml noprint;
forecast out = bjout lead=1;
run;
*Datei bjout weiterverarbeiten, daß 1-Schritt-Prognose letztlich am richtigen Platz
in Hauptdatei steht und iterieren für t>T+j, j=1,...,11;
```

2.2 Kombinationsansätze

Für die programmtechnische Umsetzung der gemischten Prognosen wurde eine ähnliche Strategie wie bei der Erstellung der Einzelprognosen gewählt. Die ungewogene Kombination, die Kombination nach dem Wahrscheinlichkeitsverfahren sowie die Kombination unter der Annahme unkorrelierter Prognosefehler sind in einem DATA-Step umgesetzt, wobei die entsprechenden Fehlervarianzen bei der Erstellung der Einzelprognosen bereits iterativ ermittelt worden sind und somit zur Verfügung stehen.

```

*Ungewogene Kombination;
comb1 = 1/m * ( fc1 + fc2 + ... + fcm);

* Kombination unter Annahme unkorrelierter Prognosefehler;
comb2 = 1/ (1/var_fc1 + 1/var_fc2 + ... + 1/var_fcm) * (1/var_fc1*fc1 +
1/var_fc2 * fc2 + ... + 1/var_fcm * fcm);

```

Das Theorem 1.5 erlaubt die Bestimmung der kombinierten Prognose nach dem Kovarianzansatz über die SAS Prozedur REG zu bestimmen.

```

*Kovarianzansatz;
data kovar;
set Hauptdatei;
if n > T then delete;
proc reg outest=gewicht1 noprint;
model x = fc1 fc2 ... fcm / noint;
restrict fc1 + fc2 + ... + fcm = 1;
run;
* Mit den so bestimmten Gewichten w_fc1,...,w_fcm in Datei gewicht1 wird dann die
Prognose in der Periode T+1 erstellt, indem entsprechende Dateimanipulationen in
DATA-Steps durchgeführt werden. Der Vorgang wird dann iteriert für die Zeitpunkte
t=T+1,...,T+11;

```

Für die konkrete Bestimmung der gemischten Prognose müssen allerdings noch Dateioperationen durchgeführt werden, so daß mit ex-post ermittelten Gewichten echte ex-ante Prognosen geschätzt werden können.

```

*Kombination mit Kovarianzansatz;
comb3 = w_fc1 * fc1 + w_fc2 * fc2 + ... + w_fcm * fcm;

```

In analoger Weise werden die gemischten Prognosen mit dem Regressionsansatz bestimmt, wobei lediglich andere Restriktionen in der Prozedur REG zur Anwendung kommen.

```

*Kombination mit freier Regression ohne Absolutglied;
proc reg data = kovar outest = gewicht2 noprint;
model x = fc1 fc2 ... fcm / noint;

```

```

*Kombination mit freier Regression;
proc reg data = kovar outest = gewicht3 noprint;
model x = fc1 fc2 ... fcm;

```

Zur Implementierung des Wahrscheinlichkeitsansatzes bedarf es einer Verzweigung, die für jede Einzelprognose zählt, wie oft sie die beste Prognose gemessen am absoluten Prognosefehler geliefert hat. Dies kann, wie bereits oben erwähnt, in einem DATA-Step vollzogen werden. Aus den so ermittelten Gewichten berechnet sich dann die gemischte Prognose.

```

*Wahrscheinlichkeitsverfahren;
*1. Gewichtsbestimmung;
if (abs(x-fc1) <= abs(x-fc2) and ... (abs(x-fc1) <= abs (x-fcm)) then w_fc1 +1;
    • usw;
if (abs(x-fcm) <= abs(x-fc1) and ... (abs(x-fcm) <= abs (x-fc(m-1))) then w_fcm +1;
*2. Kombination;
comb6 = 1/(w_fc1 + ... + w_fcm) * (fc1 * w_fc1 + ... + fcm * w_fcm);

```


3. Evaluierung

3.1 Theil'scher Ungleichheitskoeffizient

Die Beurteilung der Leistungsfähigkeit der einzelnen als auch der gemischten Prognosen erfolgt unter Verwendung des Theil'schen Ungleichheitskoeffizienten U (Theil, 1971, S. 368)

$$U := \sqrt{\frac{\sum_{t=\tau+1}^{\tau+j} (x_t - \hat{x}_t)^2}{\sum_{t=\tau+1}^{\tau+j} (x_t - \hat{x}_t^N)^2}}. \quad (Gl.15)$$

Dieser setzt die summierte quadratische Abweichung der zu beurteilenden Prognosemethode ins Verhältnis zur summierten quadratischen Abweichung der naiven Prognose, welche lediglich eine Fortschreibung des zuletzt verfügbaren Wertes ist. Der so erhaltene Quotient wird noch mit der streng monotonen und konkaven Wurzelfunktion transformiert. Aufgrund dieser Transformation lassen Werte von U , die knapp kleiner als 1 sind, schon auf eine passable Performance der zu beurteilenden Prognosemethode schließen. Die Evaluierung erfolgt durch die Auswertung von 12 ex-anten Ein-Schritt-Prognosen.

3.2 Ergebnisse

Das Ausmaß der Leistungsfähigkeit der Kombinationsansätze soll mit einem Datensatz demonstriert werden, welcher 1982 von der Deutschen Gesellschaft für Operations Research (DGOR) speziell für einen Prognosewettbewerb zusammengestellt wurde (Schwarze/Weckerle, 1982). Der Datensatz soll einen Querschnitt über häufig in der Praxis auftretende Prognoseprobleme präsentieren. Tabelle 2.1 gibt Auskunft über den Datensatz.

Zeitreihe	Herkunft	Länge
ZR2	Arbeitslose in der BRD	103
ZR3	Index des Auftragseinganges der chemischen Industrie	103
ZR4	Verfremdete Absatzreihe (mit stochastischem „Schmutz“ beladen)	103
ZR05-07	Absatzzahlen verschiedener Produkte	72
ZR08-15	Umsatzzahlen eines Produktes	43-110
ZR16	Monatl. Stromverbrauch der Stadt Braunschweig	204

Tabelle 2.1 : Erläuterung der DGOR-Daten (Schwarze, 1982, A4)

Ein detaillierteres Studium der Ergebnistabelle 2.2 für die Einzelprognosen und der Ergebnistabelle 2.3 für die kombinierten Prognosen soll nicht Gegenstand dieses Aufsatzes sein. Es sei angemerkt, daß zwar die theoretisch zu erwartende Prognosegüteverbesserung eintritt, über das Ausmaß dieser möge sich der Leser jedoch selbst ein Bild machen.

Zeitreihe	Naive Prog.	Arses	gew. Mittel	Brown	Box-Jenkins	Winters
ZR2	1,000	1,529	1,292	1,523	0,710	1,764
ZR3	1,000	1,175	0,976	1,083	0,991	1,145
ZR4	1,000	0,978	0,946	1,018	0,974	0,944
ZR5	1,000	0,855	0,918	0,925	0,900	0,620
ZR6	1,000	1,022	0,996	1,083	1,033	0,700
ZR7	1,000	0,766	0,804	0,800	0,785	0,742
ZR8	1,000	0,743	0,831	0,798	0,793	0,865
ZR9	1,000	0,761	0,863	0,852	0,739	0,891
ZR10	1,000	1,101	1,078	1,190	0,850	0,775
ZR11	1,000	0,991	1,005	1,082	1,168	0,722
ZR12	1,000	0,811	0,905	0,894	0,790	0,780
ZR13	1,000	0,673	0,788	0,749	0,692	0,444
ZR14	1,000	0,717	0,789	0,747	0,689	0,414
ZR15	1,000	0,630	0,741	0,694	0,497	0,370
ZR16	1,000	1,325	1,086	1,295	0,868	0,737
Mittelwert	1,000	0,938	0,935	0,982	0,832	0,794
Varianz	0,000	0,066	0,021	0,054	0,028	0,114
Minimum	1,000	0,630	0,741	0,694	0,497	0,370
Maximum	1,000	1,529	1,292	1,523	1,168	1,764

Ergebnistabelle 2.2 : Auswertung der Teilwerte der Einzelprognosen

Zeitreihe	Ungewogene	Kovarianz	Regression	Freie Regres.	Unkorreliert	WK-Verfahren
ZR2	0,783	0,576	0,572	0,620	0,809	0,609
ZR3	0,783	0,570	0,568	0,593	0,805	0,610
ZR4	0,910	1,013	0,984	1,496	0,913	0,901
ZR5	0,784	0,699	0,668	0,850	0,764	0,728
ZR6	0,909	0,623	0,638	0,689	0,853	0,782
ZR7	0,711	0,768	0,751	1,169	0,742	0,705
ZR8	0,735	0,729	0,741	2,202	0,726	0,715
ZR9	0,750	0,748	0,742	2,063	0,743	0,734
ZR10	0,778	0,626	0,635	3,036	0,591	0,525
ZR11	0,907	0,736	0,737	0,768	0,815	0,775
ZR12	0,804	0,666	0,679	2,680	0,766	0,737
ZR13	0,626	0,628	0,619	0,471	0,604	0,600
ZR14	0,640	0,302	0,329	0,419	0,579	0,607
ZR15	0,562	0,296	0,294	0,282	0,524	0,480
ZR16	0,872	0,705	0,710	0,724	0,628	0,629
Mittelwert	0,770	0,646	0,644	1,204	0,724	0,676
Varianz	0,011	0,031	0,028	0,777	0,013	0,012
Minimum	0,562	0,296	0,294	0,282	0,524	0,480
Maximum	0,910	1,013	0,984	3,036	0,913	0,901

Ergebnistabelle 2.3 : Auswertung der Teilwerte der Kombinationsprognosen.

Zusammenfassung

Wie durch Kombination von verschiedenen Einzelprognosen zu einer gemischten verbesserte Prognosen erreicht werden können, ist Inhalt dieses Aufsatzes. Einerseits werden die theoretischen Grundlagen und Ergebnisse wiedergegeben und ferner diese mit SAS umgesetzt. Die wesentlichen Aufgaben zur Bestimmung der einzelnen sowie kombinierten Prognosen erfolgen in DATA-Steps sowie mit den SAS Prozeduren FORECAST, ARIMA und REG. Mit einer Demonstration der Leistungsfähigkeit der Kombinationsansätze schließt die Arbeit.

Literaturverzeichnis

- Bates, J. M.; Granger, C. W. J. (1969):**
The Combinations of Forecasts;
in: Operational Research Quarterly, Vol. 20, 1969, S. 451-468
- Box, G. E. P.; Jenkins, G. M. (1976):**
Time Series Analysis, forecasting and control;
Prentice Hall Inc., 1976
- Bunn, D. (1978):**
The Synthesis of Forecasting , Models in Decision Analysis;
Birkhaeuser, Basel, Stuttgart
- Cramér, H. (1996):**
Mathematical Methods of Statistics; 11th Printing;
Princeton University Press, Princeton
- Clemen, R. T. (1989):**
Combining Forecasts: A review and annotated bibliograpy;
in: International Journal of Forecasting, Vol. 5, 1989, S. 559-583
- Eherler, D. (1996):**
Prognose von Sonderausstattungen mit kombinierten Prognoseansätzen;
Universität Ulm, Diplomarbeit
- Granger, C. W. J.; Ramanathan, R. (1984):**
Improved Methods of Combining of Forecasts;
in: Journal of Forecasting, Vol.3, 1984, S. 197-204
- Gupta, S.; Wilton, P. C. (1987):**
Combination of Forecasts: An Extension
in: Management Science, Vol. 33, No. 3, 1987, S. 356-372
- Lewandowski, R. (1974):**
Prognose- und Informationssysteme und ihre Anwendungen;
Walter de Gruyter, Berlin, New York, 1974
- Makridakis, S. (1989):**
Why combining works?
in: International Journal of Forecasting, Vol. 5, 1989, S. 601-603
- Neumann, K. (1993):**
Operations Research;
Carl Hanser Verlag, München, Wien
- SAS Institute Inc. (1998):**
SAS/ETS User's Guide, Version 6, First Edition;
Cary, NC: SAS Institute Inc
- SAS Institute Inc. (1998):**
SAS/STAT User's Guide, Version 6, Fourth Edition, Volume 2;
Cary, NC: SAS Institute Inc
- Schlittgen, R.; Streitberg, B. H. J. (1991):**
Zeitreihenanalyse; 4. Unveränderte Auflage;
R. Oldenbourg Verlag, München Wien

Theil, H. (1971):

Applied Economic Forecasting, 2nd Edition;
North-Holand Publishing Co., Amsterdam

Thiele, J. (1993):

Kombination von Prognosen;
Physica Verlag, Heidelberg

Winkler, R. L. (1989):

Combining forecasts: A philosophical basis and some current issues;
in: International Journal of Forecasting; Vol. 5, 1989, S. 605-609

Winters, P. R. (1960):

Forecasting Sales by Exponentialy Weighted Moving Averages;
in: Management Science, Vol. 6, Apr. 1960, S. 324-342