

Originalarbeiten

Meteorol. Zeitschrift, N.F. 4, 139–149 (August 1995)
© by Gebrüder Borntraeger 1995

Praktische Vorhersagbarkeit: Persistenz in rotem Rauschen

KLAUS FRAEDRICH und CHRISTINE ZIEHMANN, Hamburg

Zusammenfassung. Persistenz-Vorhersagen von rotem Rauschen sind ein einfacher konzeptioneller Ansatz für die Untersuchung der Vorhersagbarkeit imperfekter Prognose-Modelle. Die Methoden der Analyse der praktischen Vorhersagbarkeit numerischer Wetterprognosen (NWP) und ihre Ergebnisse lassen sich damit exemplarisch nachvollziehen.

External predictability: persistence in red noise

Summary. Persistence forecasts in a red-noise atmosphere provide a conceptual model for the predictability analysis of imperfect forecasts. The methods applied in analysing numerical weather prediction (NWP) and the results obtained are qualitatively comparable with those of this conceptual model.

1. Einführung: Vorhersagbarkeit

Vorhersagbarkeit beschreibt die sensitive (empfindliche) Abhängigkeit des Wetter- und Klimasystems von Anfangs- und Randbedingungen. Man unterscheidet die Vorhersagen erster und zweiter Art (LORENZ 1975):

Vorhersage 1. Art beschreibt die sensitive Abhängigkeit von Anfangsbedingungen, deren Ursache interne Instabilitäten des Systems sind. In der Praxis tritt das in der Wetterprognose auf. Ein entsprechendes „Weltbild“ wird durch das deterministische Chaos eines Systems mit vielen Freiheitsgraden geliefert.

Vorhersage 2. Art beschreibt die Abhängigkeit von Randbedingungen. Ursachen sind die strukturelle Instabilität bei sich nur langsam ändernden Parametern (Tab. 1). In der Praxis tritt das bei Langfrist- und Klimaprognosen auf, wobei das vorherzusagende System deterministisch ist und die Systeme mit dem längeren Gedächtnis die Randbedingungen stellen.

Tab. 1. Interne und externe Klimasysteme beim Modellbau (nach HASSELMANN 1979).

System	Zeitskalen (Tage)			
	0–10	100	1 000	10 000
Atmosphäre	Vorh. 1. Art →			Vorhersage 2. Art
Atmosphären+ Ozean	Vorhersage 1. Art →			Vorhersage 2. Art
Atmosphäre+ Ozean+Eis	Vorhersage	1. Art	—————→	

Analyse der Vorhersagbarkeit: Vorhersagbarkeit wird traditionell durch den Prognose-Fehler analysiert, zum Beispiel den Abstand im Euklidischen Raum. Damit kann das Fehlerbudget in Analogie zur Diffusion (im Phasenraum) untersucht werden, wobei der Abstand diffundierender Partikel dem Fehler entspricht: Ein-Partikel Diffusion beschreibt nur die Verifikations- oder Beobachtungstrajektorie, $X(t)$, d. h. die zeitliche Entwicklung der Atmosphäre im Modell oder in der Natur. Der Abstand, $|X(t)-X(t_0)|$, entspricht dem Fehler der Persistenz-Prognose, die von t_0 startet. In der Zwei-Partikel Diffusion werden Verifikations- und Vorhersage-Trajektorien assoziiert, und viele Partikel (Punkt-Wolke) stehen für Ensemble-Prognosen.

Vorhersagbarkeitsexperimente: Die Daten zur Diagnose des Fehlerbudgets werden aus Experimenten gewonnen, von denen man zwei Arten unterscheidet: *Externe* oder praktische Vorhersagbarkeit charakterisiert die Wetterprognose (Modell-Trajektorie) und ihre Verifikation (die Wetter-Trajektorie). Hier tritt der systematische Fehler (Bias) wie bei jedem imperfekten Modell besonders nach längeren Prognose-Zeiten auf. *Interne* oder theoretische Vorhersagbarkeitsexperimente werden nur mit dem Modell gemacht:

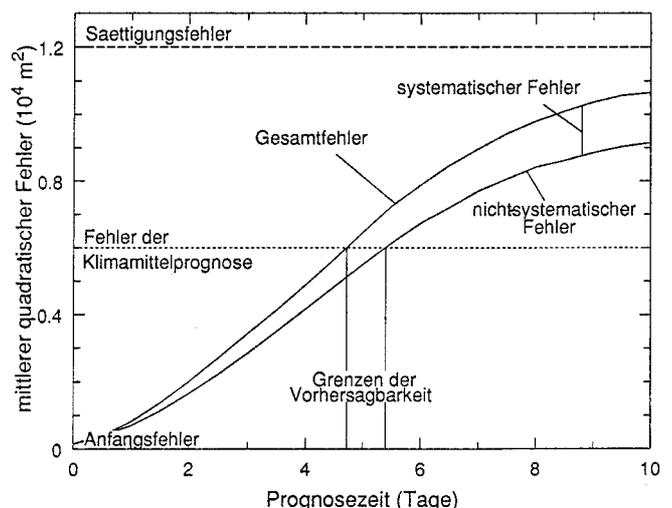


Abb. 1. Systematischer und nicht-systematischer Vorhersage-Fehler in Abhängigkeit von der Prognose-Zeit eines numerischen Wetter-Prognose-Modells (nach DALCHER und KALNAY 1987).

Fig. 1. Systematic and random forecast error as a function of forecast time; analysis and forecast variance are also shown for comparison.

Prognosen werden mit künstlichen Anfangsstörungen initialisiert, die Verifikation ist die ungestörte Modell-Trajektorie („eineiige Zwillinge“, identical twin Experiment) in perfekter Modell-Umgebung.

Ein typisches Beispiel für die zeitliche Entwicklung des Prognose-Fehlers ist in Abb. 1 dargestellt (nach DALCHER und KALNAY 1987, Fig. 4); drei Bereiche werden in solchen Fehlerkurven diskutiert:

- der Anfangsfehler und sein Wachstum,
- die Vorhersagbarkeit und ihre Grenzen,
- die Sättigung (für große Vorhersagezeiten).

Anfangsfehler und -wachstum hängen eng mit dem Analyse-System zusammen. Die Vorhersagbarkeit und ihre Grenze kann entweder (mehr oder weniger) willkürlich definiert oder durch den mittleren Fehler einer Referenz-Prognose, wie zum Beispiel das Klima, festgelegt werden. Sättigung ist erreicht, wenn Prognose und Verifikation unabhängig voneinander sind. Hier zeigt sich auch die Größe des systematischen Modell-Fehlers, der zusammen mit dem nichtsystematischen Fehler den Gesamt-Fehler ausmacht. Im folgenden werden Vorhersagbarkeitsexperimente mit der Persistenz-Prognose (als imperfektes Prognose-Modell) für eine sich zeitlich entwickelnde „Surrogat“-Atmosphäre durchgeführt, die durch rotes Rauschen bzw. den autoregressiven Prozeß repräsentiert wird. Dieser Ansatz ermöglicht analytische Ergebnisse für ein weites Spektrum von Vorhersagbarkeitsexperimenten. Der Aufbau ist wie folgt: Die Ersatz-Atmosphäre wird in Abschnitt 2 eingeführt; dann folgen individuelle (Abschnitt 3) und Ensemble-Prognosen (Abschnitt 4) und zum Abschluß (Abschnitt 5) wird eine Parameterisierung für den Zusammenhang zwischen Anfangsfehler und Vorhersagbarkeitsgrenze diskutiert. Vergleich mit den Ergebnissen der numerischen Wetterprognose wird ebenfalls versucht.

2. Ersatz-Atmosphäre: Autoregressiver Prozeß 1. Ordnung

Stochastische dynamische Systeme sind in der Meteorologie und Klimatologie in zweifacher Hinsicht von Bedeutung: Als stochastische Modelle werden sie den Daten angepaßt und dienen der Prognose; als Prozesse simulieren sie natürliche Fluktuationen. Insbesondere wird der autoregressive Prozeß, AR(1) oder rotes Rauschen, oft als Referenz (Null-Hypothese) zum Test atmosphärischer Zeitreihen eingeführt. So kann man mit nur einem geeignet zu wählenden Parameter das zeitliche Verhalten unterschiedlicher atmosphärischer Vorgänge von synoptischen Störungen bis hin zu persistenten Anomalien, qualitativ beschreiben (zum Beispiel DOLE und GORDON 1983, FRAEDRICH und DÜMMEL 1983, GUTZLER und MO 1983).

Der AR(1) als stationärer Gauß-Prozeß erzeugt hier eine univariate Zeitreihe, die eine „Ersatz“-Atmosphäre charakterisieren soll. Damit erhält man Sequenzen einer Variablen, $X(t) = \langle X \rangle + X'(t)$, deren Mittel und Varianz konstant sind. Die normal-verteilten Anomalien $X'(t)$, treten in diskreten Zeitschritten, $t, t-1, t-2, \dots$ auf und fluktuieren um einen

verschwindenden Mittelwert $\langle X \rangle = 0$. Im folgenden wird der Strich weggelassen, da nur noch Anomalien X' betrachtet werden. Der AR(1)-Prozeß besteht aus einem deterministischen, $aX(t-1)$, und einem additiven stochastischen Teil z_1

$$X(t) = aX(t-1) + z_1. \quad (2.1)$$

Der stochastische Anteil wird von einem Gauß'schen Zufallsprozeß bestimmt, dessen Mittel und Auto-Korrelation verschwinden, $\langle z \rangle = 0$ und $\langle z_i z_j \rangle = 0$, wenn $i \neq j$ („weißes Rauschen“); die Varianz ist $\langle z^2 \rangle = s_z^2$. So ist die Größe X Gauß-verteilt mit der Wahrscheinlichkeitsdichte $(2\pi/s_x^2)^{1/2} \exp\{-(X - \langle X \rangle)^2/2s_x^2\}$ und der Varianz s_x^2 ; sie wird durch die Standard-Abweichung s_x normiert, X/s_x . Der deterministische Anteil wird durch die Auto-Kovarianz Funktion (bzw. Auto-Korrelation, da X normiert ist) charakterisiert, $a = a(\text{lag})$, die von der Zeitverschiebung, lag, abhängt. Beim AR-Prozeß erster Ordnung gilt:

$$a(\text{lag}) = \langle X(t)X(t-\text{lag}) \rangle / s_x^2, \quad (2.2)$$

Nun kann man Zeitreihen der Ersatz-Atmosphäre durch AR(1)-„Vorhersagen“ definieren: Sie gehen aus von einem Initial-Zeitpunkt ($t-r$) und reichen bis zur Verifikationszeit (t), wobei r identisch mit dem Prognose-Intervall des im Abschnitt (3.2) eingeführten Persistenz-Vorhersage-Modells ist: Aus $X(t) = aX(t-1) + z_t$ und $X(t-1) = aX(t-2) + z_{t-1}$ ergibt sich $X(t) = a^2X(t-2) + az_{t-1} + z_t = a^3X(t-3) + a^2z_{t-2} + az_{t-1} + z_t$, etc., so daß

$$X(t) = a^r X(t-r) + \sum_{i=1}^r a^{i-1} z_{t-i}. \quad (2.3)$$

Integraler Zeit-Scale: Ein Zeit-Scale, der das Gedächtnis des Prozesses charakterisiert, kann mit den Fluktuationen, $X(t)$, des AR(1) assoziiert werden. Die Auto-Korrelation ist, $a(r) = \langle X(t)X(t-r) \rangle / s_x^2 = a^r$, da $\langle \{a^r X(t-r) + \sum_{i=1}^r a^{i-1} z_{t-i}\} \cdot X(t-r) \rangle = a^r \langle X^2(t-r) \rangle = a^r s_x^2$, und der stochastische Term, $\langle \{ \sum_{i=1}^r a^{i-1} z_{t-i} \} \cdot X(t-r) \rangle = 0$ entfällt. Integration von a^r über beliebig lange Zeitschritte r liefert den integralen Zeit-Scale τ :

$$\tau = 1/(1-a), \quad (2.4)$$

da $\sum_{i=1}^N a^i = (1-a^{N+1})/(1-a) \sim 1/(1-a) = \tau$ für große N ; τ ist ein Maß für das Gedächtnis des Prozesses, bzw. der Zeitreihe. Es ist groß bei großer Autokorrelation, und umgekehrt. Damit kann qualitativ der AR(1)-Prozeß auch meteorologisch interpretiert werden wie schematisch in Tab. 2 dargestellt. Für diskrete lineare Systeme verläuft der Abkling-Prozeß nach einer geometrischen Progression, $X_0 a^r$; für kontinuierliche lineare Systeme gilt ein exponentielles Gesetz, $X_0 \exp(-\lambda(t-t_0))$, mit der Zerfallsrate λ .

Folgende Spezial-Fälle ($a=0, 0 < a < 1, a=1$) sollen erwähnt werden:

- (i) Weißes Rauschen ($a=0$): Der Zufallsprozeß „Weißes Rauschen“ hat ein verschwindendes Mittel $\langle z \rangle = 0$, ist unkorreliert mit verschwindender Auto-Kovarianz $\langle z_i z_j \rangle = 0$ wenn $i \neq j$, und die Varianz ist $\langle z^2 \rangle = s_z^2$. Im Spektrum wird keine Frequenz bevorzugt (weiß).
- (ii) Rotes Rauschen ($0 < a < 1$): Der oben beschriebene AR(1)-Prozeß wird auch als „rotes Rauschen“ bezeichnet. In diesem Spektrum treten lange Perioden bevorzugt auf (rot).

(iii) Irrweg ($a=1$, random walk): Nach dem Start bei $X(t) = 0$, erhält man nach r Schritten den Zufallsprozeß, $X(t+r) = \sum_{i=1}^r z_i$ für $i=1, \dots, r$. Das Mittel ist $\langle X(t+r) \rangle = r \langle z_i \rangle = 0$ wenn $\langle z \rangle = 0$, und die Varianz $\langle X^2(t+r) \rangle = \langle \sum_{i=1}^r z_i^2 \rangle$ mit $i=1, \dots, r$ wächst linear mit der Zeit, $\langle X^2(t+r) \rangle = r s_z^2$.

Mittel, Median bzw. andere Quantile charakterisiert (wie in Abb. 2). Es gibt andere Fehlermaße: der Quadrat-Wurzel-Fehler (rms-error), \sqrt{e} , als Euklidischer Abstand zwischen Prognose und Verifikation, die Anomalie-Korrelation, $\langle F'X' \rangle$, und andere.

Tab. 2. Qualitative Interpretation des AR(1).

$a \rightarrow 1$	a groß	a klein	$a \rightarrow 0$
IRRWEG $s_x^2 = r s_z^2$	ROTES τ groß	RAUSCHEN τ klein	WEISSES RAUSCHEN $s_x^2 = s_z^2 / (1 - a^2)$
	Lärm Signal	Lärm Signal	$s_x^2 = s_z^2 (1 - a^2)$ $a^2 s_x^2$
	„Blok- kierung“	„Zyklonen- familie“	

3. Individuelle Vorhersagen

Gegeben sei eine Vorhersage $F(r)$ für das Zeitintervall r . Die Vorhersage-Fehler, $e = X(t) - F(r)$, zwischen den Prognosen $F(r)$ und ihren Verifikationen $X(t)$ lassen sich als Realisierungen einer Zufallsvariablen interpretieren. Oft werden die Fehler durch den mittleren quadratischen Fehler, $E(r) = \langle e^2 \rangle = \langle (X(t) - F(r))^2 \rangle$ normiert, $y = e / \sqrt{E}$. Sind die Fehler y Gauß-verteilt, so ist das normierte Fehler-Quadrat y^2 Chi-Quadrat-verteilt; solche Verteilungen werden meist durch

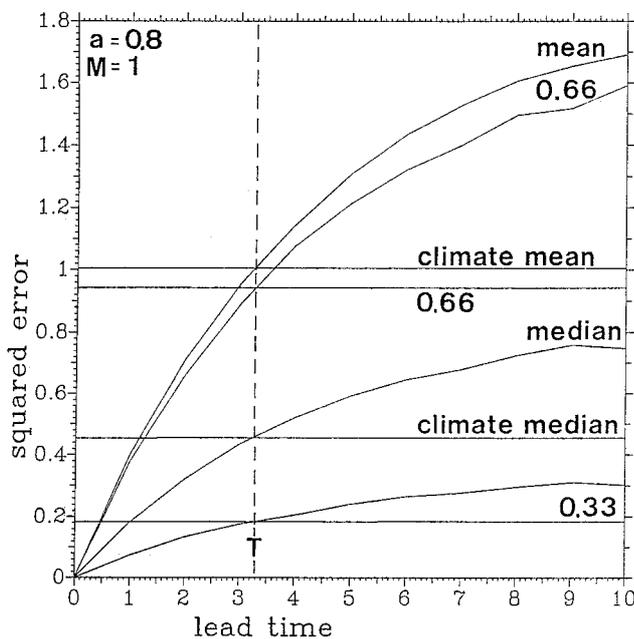


Abb. 2. Quadratischer Fehler von Klima- und Persistenz-Vorhersagen von rotem Rauschen: Mittlerer quadratischer Fehler, Median, oberes und unteres Terzil (FRAEDRICH und ZIEHMANN-SCHLUMBOHM 1994, Fig. 1).

Fig. 2. Error budget of individual persistence forecasts ($M=1$) in a red-noise atmosphere (autocorrelation $a=0.8$). The time evolution of the error variance, the median and the upper and lower terciles (0.66; 0.33) of the squared-error distribution. The same is shown for climate-mean forecasts (horizontal lines). The predictability limit, T , is also indicated.

3.1. Referenz-Prognosen: Klima und Zufall

Das Klimamittel, $\langle X \rangle$, ist eine wichtige Referenz-Vorhersage. Hier wird die Vorhersagbarkeitsgrenze definiert als die Zeit, nach der der Vorhersage-Fehler eines zu untersuchenden Modells den Fehler der Klima-Mittel-Prognose überschreitet. In der AR(1)-Atmosphäre hat die Klimaprognose, $F(r) = \langle X \rangle = 0$, den mittleren quadratischen Fehler,

$$E_K(r) = E_K = \langle X(t)^2 \rangle = s_x^2,$$

der unabhängig von der Prognose-Zeit ist. Das heißt, der mittlere quadratische Abstand von einem Wetterzustand zum Klima entspricht der Varianz s_x^2 der Zeitreihe.

Wird dagegen ein beliebiger Wetterzustand zufällig aus der klimatologischen Verteilung der Zustände (entsprechend ihrem natürlichen Maß) für eine Prognose gezogen, $F(r) = X_i$, die unabhängig von der Verifikation ist, so ergibt der mittlere quadratische Abstand oder Fehler dieser Vorhersagen

$$E_Z(r) = E_Z = \langle (X_i - X_j)^2 \rangle = 2s_x^2$$

die zweifache Varianz der Zeitreihe (gleich dem Abstand zweier beliebiger Wetterzustände) und ist ebenfalls unabhängig vom Prognose-Intervall. Diesem Sättigungswert, $2s_x^2$, nähern sich die Fehler von numerischen Wetterprognosen wie auch von Persistenz-Vorhersagen nach längerer Prognose-Zeit.

3.2. Individuelle Persistenz-Prognose

Auch die Persistenz-Vorhersage wird vielfach als Referenz-Prognose herangezogen; nur bessere Vorhersagen als Persistenz haben Güte in der Prognose der Zeitableitung, dem Ziel aller Prognosen. Hier wird Persistenz jedoch nicht als Referenz- sondern als imperfektes Vorhersage-Modell für die rote Ersatz-Atmosphäre benutzt. Dabei ist $F(r)$ wieder die Prognose für das Zeitintervall r , $X(t)$ die dazu gehörige Verifikation, $e(r)$ der Fehler, $\langle \rangle$ das Ensemble-Mittel über die individuellen Vorhersage-Realisierungen (Abb. 2):

$$F(r) = X(t-r); X(t) = a^r X(t-r) + \sum_{i=1}^{r-1} a^{i-1} z_{t-i};$$

$$e(r) = F(r) - X(t)$$

$$E_1(r) = \langle e^2(r) \rangle = 2s_x^2(1 - a^r) = E_1(\infty)[1 - a^r] \quad (3.1)$$

da

$$\langle (F(r) - X(t))^2 \rangle = \langle (X(t-r) - X(t))^2 \rangle = 2s_x^2 - 2s_x^2 a^r.$$

Persistenz zeigt am Anfang ($r=0$) keine Fehler, $E_1(r=0) = 0$. Diese besondere Eigenschaft der Persistenz-Prognose kann in Kombinationsvorhersagen genutzt werden (FRAEDRICH 1992). Bei Sättigung ($r \rightarrow \infty$) entspricht der Persistenz-Fehler

der mittleren quadratischen Differenz von zwei zufällig gewählten Wetterzuständen, $E_1(\infty) = E_z = 2s_x^2$, wobei die Varianz $s_x^2 = \langle X^2 \rangle$ ist. Die Verteilungsdichten der normierten Fehler y bzw. deren Quadrate, y^2 sind Gauß bzw. χ^2 :

$$g(y) = (2\pi E_1(r))^{-1/2} \exp(-y^2/2),$$

$$f(y^2) = (2\pi)^{-1/2} y^{-1} \exp(-y^2/2)$$

Vorhersagbarkeitsgrenze (limit of predictability): Der wachsende Fehler erreicht bei $r = T$ die Größenordnung der Varianz. Dann ist der Fehler der Persistenz-Prognose gleich dem Fehler der Prognose mit dem Klima-Mittel:

$$E_1(r=T) = s_x^2 \rightarrow T = \ln(1/2) / \ln(a) \sim \tau \ln 2 \rightarrow T \sim \tau \quad (3.2)$$

Dabei ist der integrale Zeit-Scale, $\tau = 1/1-a$, zur Umformung benutzt worden: $\ln(1/a) \sim 1/\tau$. Das heißt, der effektive Vorhersagebereich komplexer Systeme ist endlich und begrenzt durch die Lebensspanne des energiereichsten Phänomens (TENNEKES 1991); die Lebensspanne von Anomalien entspricht hier dem integralen Zeit-Scale der Ersatz-Atmosphäre.

Systematischer und nicht-systematischer Fehler: Für ein imperfektes Modell kann der mittlere quadratische Fehler, E , in einen systematischen Fehler oder Bias, $SE = \langle e \rangle^2 = \langle \langle F \rangle - \langle X \rangle \rangle^2$, und einen nicht-systematischen, $RE = \langle (e - \langle e \rangle)^2 \rangle = \langle (F - X) - (\langle F \rangle - \langle X \rangle) \rangle^2$, aufgeteilt werden:

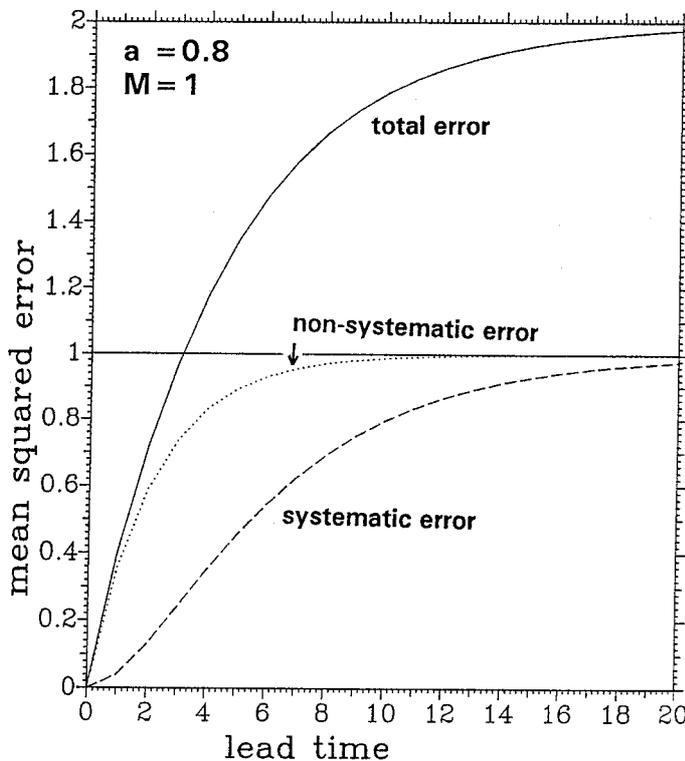


Abb. 3. Systematischer und nicht-systematischer Fehler individueller Persistenz-Prognosen (FRAEDRICH und ZIEHMANN-SCHLUMBOHM 1994, Fig. 6a).

Fig. 3. Ensemble-mean predictions: Systematic, non-systematic and total error variance changing with lead time r for individual persistence forecasts, $M=1$.

$$E = SE + RE. \quad (3.3)$$

Es wird ausgegangen vom bedingten Fehler bei vorgeschriebenem Anfangswert X_0 :

$$E_1(r|X_0) = X_0^2(1-a^r)^2 + s_x^2(1-a^{2r}). \quad (3.4)$$

Die Herleitung erfolgt durch bedingte Mittel-Bildung, $E = \langle E_1(r|X_0) \rangle$, über feste X_0 's: $SE(r|X_0) = (\langle F \rangle - \langle X \rangle)^2 = X_0^2(1-a^r)^2$. Nach Mittelung ist $SE = \langle SE(r|X_0) \rangle = \langle X_0^2 \rangle (1-a^r)^2 = s_x^2(1-a^r)^2$, da $\langle X(t) \rangle = a^r \langle X(t-r) \rangle + \langle \sum_i^{t-1} a^i z_{t-i} \rangle = a^r \langle X(t-r) \rangle = a^r X_0$, wobei $\langle z_i \rangle = 0$; ferner ist $\langle F \rangle = \langle X(t-r) \rangle$. Somit folgt $RE = \langle RE(r|X_0) \rangle = E - SE = s_x^2(1-a^{2r})$, da der Gesamt-Fehler $E = 2s_x^2(1-a^r)$ und $\langle X_0^2 \rangle = s_x^2$. Dann zeigt sich, daß der systematische Fehler der Persistenz-Prognose den Sättigungswert vom nicht-systematischen Niveau, $RE(\infty) = s_x^2$, um $SE(\infty) = s_x^2$ auf $E(\infty) = 2s_x^2$ erhöht; das ist der Abstand unabhängiger Wetterzustände (Abb. 3). Ohne ihren systematischen Fehler würde die Persistenz-Prognose bei Sättigung den Fehler der Klima-Mittel-Prognose, E_K , erreichen. Bei $r=0$ zeigt der systematische Fehler Null-Wachstum, während der nicht-systematische wie der Gesamt-Fehler umgekehrt proportional zur Lebensspanne, $RE_r(r=0) \sim 1/\tau$ wachsen, da $RE_r = 2s_x^2 a^r \ln(1/a)$.

3.3. Gesetz für das Fehlerwachstum

Fehlerwachstum beschreibt die Änderung des Fehlers pro zunehmende Prognose-Zeit, $E_{1r} = dE_1/dr = 2s_x^2 a^r \ln(1/a)$. Dieser Ansatz „Persistenz in rotem Rauschen“ liefert beschränktes Wachstum mit einem *linearen* Gesetz für den Fehler $E_1(r)$

$$dE_1/dr = \ln(a^{-1}) E_1(\infty) (1 - E_1/E_1(\infty)) \sim E_1(\infty) \{1 - E_1/E_1(\infty)\} / \tau \quad (3.5)$$

während das von LORENZ (1969, siehe auch review von STROE und ROYER 1993) analog zur Verhulst-Gleichung eingeführte Fehler-Modell beschränktes Wachstum durch einen *quadratischen* Term erreicht. Das Anfangswachstum in (3.5) ist $E_{1r}(r=0) = 2s_x^2 \ln(1/a)$, denn $d[2s_x^2(1-a^2)]/dr = 2s_x^2 a^r \ln(a)$. Je länger das Gedächtnis τ des atmosphärischen Prozesses, desto geringer das Fehlerwachstum: $E_{1r}(r=0) \sim 2s_x^2/\tau$, da $\ln(1/a) \sim 1/\tau$. Das heißt, die Lebensspanne τ des zu prognostizierenden Prozesses und damit die Grenze der Vorhersagbarkeit $T \sim \tau$, sind umgekehrt proportional zum Anfangswachstum des Prognose-Fehlers $\sim 1/\tau$.

Die Zeit $D(r)$ zwischen r und $r+D$, in der Fehler-Verdopplung stattfindet, hängt von der Prognose-Zeit r ab. Aus dieser Definition, $E_1(r+D) = 2E_1(r)$, folgt für die Fehler-Verdopplungszeit $D(r) = \ln(2 - a^{-r}) / \ln a$, da $2s_x^2(1-a^{r+D}) = 4s_x^2(1-a^r)$. Sie ist klein zu Beginn der Prognose und wird für große Vorhersagezeiten r unendlich groß, wenn der Fehler sich der Sättigung nähert. Deshalb wird dieses Fehlermaß meist nur für das Wachstum kleiner Fehler zu Beginn der Prognose verwendet.

4. Ensemble-Vorhersagen

Individuelle Vorhersagen werden durch ein Ensemble von Einzel-Prognosen ergänzt, um durch die folgende Statistik drei Ziele zu erreichen:

- (1) Ensemble-Mittel: Verbesserung der Vorhersage-Güte,
- (2) Ensemble-Varianz: Prognose des Vorhersage-Fehlers,
- (3) Ensemble-Verteilung: Vorhersage der Wahrscheinlichkeit atmosphärischer Zustände.

Die Hoffnung, diese Ziele zu erreichen, wird mit folgenden statistischen Überlegungen begründet, wobei die Bedingungen eines perfekten Modells und eines perfekten Ensembles erfüllt sein müssen (siehe z. B. BRANKOVIC et al. 1990).

Allgemein gilt zunächst folgende Statistik: Gegeben ein Ensemble von M individuellen Vorhersagen, $F_i, i=1, \dots, M$, mit dem Ensemble-Mittel, $[\] = M^{-1} \sum_{i=1}^M$. Zusammen mit einer weiteren Variablen X (z. B. der Verifikation, des Klima-Mittels, etc.), erhält man die mittlere quadratische Differenz:

$$[(F_i - X)^2] = [(F_i - [\])^2] + [(F_i - [\])^2] \quad (4.1)$$

Denn $[F_i - [\] + [\] - X]^2 = [(F_i - X)^2] + [(F_i - [\])^2] + 2(F_i - [\])([\] - X)$; hier verschwindet der letzte Term und der erste ist unabhängig von der Ensemble-Mittelung über M Einzel-Prognosen $[F_i] = \sum_{i=1}^M F_i / M$. Folgende Interpretationen von (4.1) sind möglich:

- $X = 0$ liefert die Varianz (Dispersion oder „spread“) der Ensemble-Prognosen um das Ensemble-Mittel $[F_i]$: $s_M^2 = [F_i^2] - [F_i]^2$.
- Für $X = X(t)$ als Verifikation stellt (4.1) den Zusammenhang zwischen dem mittleren Fehler der individuellen Vorhersagen (linke Seite) und dem Fehler der Ensemble-Mittel-Prognose (erster Term rechts) her. Es zeigt sich, daß der Fehler der Ensemble-Mittel-Prognosen, $e_M^2 = [(F_i - X)^2]$, gegenüber dem mittleren Fehler der M Einzel-Prognosen, $[e_i^2] = [(F_i - X)^2]$, um den spread, $s_M^2 = [F_i^2] - [F_i]^2$, verbessert wird: $[e_i^2] = e_M^2 + s_M^2$.
- Mit $X = F_i$ in (4.1), Summation über das Ensemble und Division durch $M-1$ erhält man den mittleren quadratischen Abstand zwischen allen nicht-identischen Paaren von individuellen Vorhersagen: $[d_i^2] = \sum_i \sum_j (F_i - F_j)^2 / M(M-1) = 2Ms_M^2 / (M-1)$; die Mittel-Bildung $[\]$ erfolgt nicht über M Abstände zur Verifikation sondern über $M(M-1)$ Abstände zwischen Einzel-Vorhersagen.

Nun werden die Hypothesen vom perfekten Modell ($r > 0$) und perfekten Ensemble ($r = 0$) eingeführt:

- Perfektes Ensemble ($r = 0$): Mitglieder eines perfekten Ensembles sind dadurch charakterisiert, daß sie so um das Ensemble-Mittel verteilt sind wie durch Anfangsfehler verfälschte Analysen um den wahren unbekanntem Zustand. Dann sollte auch gelten $[e_i^2] = [d_i^2] = 2Ms_M^2 / (M-1)$ bei $r = 0$.
- Perfektes Modell ($r > 0$): Für Prognosen eines perfekten Modells gilt, daß der mittlere Abstand zwischen den Ensemble-Mitgliedern, $[d_i^2]$, wächst wie internes deterministisches Fehlerwachstum. Dann gilt $[e_i^2] = [d_i^2] = 2Ms_M^2 / (M-1)$ auch für $r > 0$. Kombination mit obigen Formeln zeigt dann, wie sich der Fehler der Prognose

des Ensemble-Mittels, e_M^2 , gegenüber den Einzel-Fehlern, $[e_i^2]$, verhält (1. Ziel der Ensemble-Prognose), und wie dieser Vorhersage-Fehler, e_M^2 , mit der Ensemble-Varianz, s_M^2 , zusammenhängt (2. Ziel):

$$\begin{aligned} (1) \quad e_M^2 &= 1/2[e_i^2](M+1)/M \sim 1/2[e_i^2] \\ (2) \quad e_M^2 &= s_M^2(M+1)/(M-1) \sim s_M^2. \end{aligned} \quad (4.2)$$

Das heißt, für eine große Zahl M von Ensemble-Elementen führt die Ensemble-Mittelung im besten Fall zu einer Halbierung des Fehlers von individuellen Prognosen. Weiter ist der Fehler der Ensemble-Mittel-Vorhersage nahezu identisch der Ensemble-Dispersion. In der Praxis sind die Hypothesen vom perfekten Modell und Ensemble nicht erfüllt, so daß bei imperfekten Bedingungen die Ziele (1) und (2) nur bedingt erreicht werden.

Ensemble-Prognosen in der Praxis: Ensemble-Prognosen werden mit unterschiedlichen Techniken durchgeführt: Monte Carlo-Ensemble bestehen aus zufälligen Anfangsfehlern (LEITH 1974). LAF-Prognosen (oder lagged average forecast; HOFFMAN and KALNAY 1983) bestehen aus aufeinanderfolgenden Analysen und Vorhersagen. Sie sind in der Praxis leicht durchzuführen, da sie bei den großen Wetterdiensten in der Routine vorliegen (Tab. 3). Analysiert werden hierbei Ensembles, deren Elemente Prognosen sind, die von zeitverschobenen Anfangszuständen ausgehen. Beide Verfahren erfüllen nicht die o. a. Voraussetzungen von perfekten Ensembles.

Tab. 3. Ensemble-Prognosen mit zeitverschobener Initialisierung.

t=1	t=2	t=3	t=4	t=4+r
1	----->			
	2	----->		
		3	----->	
			4	----->

4.1. Verbesserung der Vorhersage-Güte

In unserem Experiment wird die LAF-Methode auf Persistenz-Prognosen angewandt. Eine Ensemble-Mittel-Prognose, $[F(r+i)] = \sum_{i=0}^{M-1} F(r+i) / M$, besteht somit aus M individuellen Persistenz-Vorhersagen für X(t) ausgehend von der jüngsten $X(t-r)$ plus $M-1$ vorangegangenen $X(t-r-i)$ mit $i=1, \dots, M-1$. Der mittlere Fehler aller Ensemble-Mittel-Prognosen, $E_M(r) = \langle e_M^2 \rangle$, und sein Wachstum, $dE_M(r) / dr$ lassen sich im Fall der hier betrachteten Persistenz-Prognose in einer roten Atmosphäre analytisch berechnen (Abb. 4):

$$\begin{aligned} E_M(r) &= E_M(\infty) \{ 1 - a^r (1 - a^M) M^{-1} / [1 - a] E_M(\infty) \} \\ dE_M / dr &= E_M(\infty) \{ 1 - E_M / E_M(\infty) \} / \tau \end{aligned} \quad (4.3)$$

wobei

$$E_M(\infty) = 2s_x^2 [1/2 + 1/2(1+a)M^{-1} / (1-a) + a(1-a^M)M^{-2} (1-a)^{-2}].$$

Es zeigt sich eine Fehler-Reduktion durch Ensemble-Mittelung bei längeren Prognose-Zeiten nach dem Überschrei-

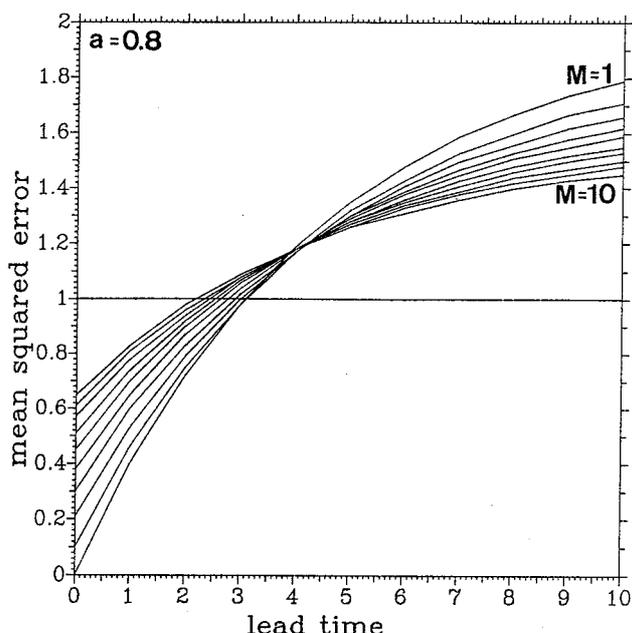


Abb. 4. Quadratischer Fehler von Ensemble-Mittel-Prognosen verschiedener Ensemble-Größen $M=1, \dots, 10$ für $a=0,8$ (FRAEDRICH UND ZIEHMANN-SCHLUMBOHM 1994, Fig. 3a).

Fig. 4. Ensemble-mean error budget, $E_M(r)$. The error variance of ensemble averaged lagged persistence forecasts for varying ensemble size $M=1$ to 10 in red-noise atmospheres with the autocorrelation $a=0.8$.

ten der Vorhersagbarkeitsgrenze. Prognosen vom Ensemble-Mittel und, zum Beispiel, vom „jüngsten“ Ensemble-Element können auch mit Hilfe der Anomalie-Korrelation verglichen werden. Dieses Fehlermaß wird oft zur Beschreibung der Unterschiede von prognostizierten und beobachteten Wetterkarten benutzt; $A_1 \sim 0,6$ wird dabei als Grenze der Vorhersagbarkeit empirisch definiert. Die Anomalie-Korrelation zwischen univariaten Persistenz-Prognosen und rotem Rauschen ist $A_1(r) = \langle X(t)X(t-r) \rangle / s_x^2 = a^r$; für Ensemble-Mittel gilt mit $s_F^2 = \langle [F(r+i)]^2 \rangle - \langle [F(r+i)] \rangle^2 = s_x^2 [(1+a)/M(1-a) - 2a(1-a^M)/(-M^2(1-a)^2)]$:

$$A_M(r) = \frac{\langle x(t) [F(r+i)] \rangle}{s_x \sqrt{\langle [F(r+i)]^2 \rangle}} = \frac{a^r (1-a^M)}{\sqrt{M(1-a^2) - 2a(1-a^M)}}$$

Den Vergleich zwischen letzter Einzelprognose und Ensemble-Mittel liefert das Verhältnis:

$$\frac{A_1}{A_M} = \frac{\sqrt{M(1-a^2) - 2a(1-a^M)}}{1-a^M} > 1 \text{ für } M > 1 \quad (4.4)$$

Das heißt, hier wie auch im Fall von praktischen Vorhersagen mit zeitverschobenen Ensemble-Elementen (lagged average forecasts, LAF) von NWP-Modellen zeigt sich, daß im Mittel und für kurze Zeiten die letzte Prognose besser ist als die LAF-Ensemble-Mittel-Vorhersage (PALMER et al. 1993): „unweighted lagged averaged forecasts give no significant advantage over single forecasts from the latest initialization“. Umgekehrt, für große Vorhersagezeiten tendiert das Ensemble-Mittel zu kleineren Fehlern als die letzte Einzel-Prognose. Dies geht aber zum Teil auf Kosten der

Varianz-Reduktion der Prognosen durch Ensemble-Mittelung (MURPHY 1988). Die Gleichungen (4.3 und 4.4, und Abb. 4) machen deutlich, daß der Fehler, $E(r \rightarrow \infty)$, nahe der Sättigung reduziert wird, so daß die mittleren Prognose-Fehler des Ensemble-Mittels nach langen Vorhersage-Intervallen geringer sind als die der Einzel-Prognosen. Im Vergleich dazu wird im Quotienten (4.4) der Anomalie-Korrelationen die Änderung des Signal/Rausch-Verhältnisses berücksichtigt, so daß im Mittel $\langle \rangle$ die Ensemble-Mittel-Prognose (unabhängig von den Anfangswerten X_0) zu keiner Vorhersagezeit besser als die letzte Einzel-Prognose ist.

4.2. Vorhersage der Vorhersagbarkeit (Prognose des Fehlers)

Ensemble-Dispersion (Varianz, spread): Die Varianz zu einem Ensemble-Mittel, $s_M^2 = \langle [F(r+i) - \langle F(r+i) \rangle]^2 \rangle = \langle [F(r+i)]^2 \rangle - \langle [F(r+i)] \rangle^2$ führt nach Mittelung über alle Anfangszustände zu $S_M = \langle s_M^2 \rangle = \langle [F(r+i)]^2 \rangle - \langle [F(r+i)] \rangle^2$, bzw.

$$S_M = s_x^2 \left\{ 1 - \frac{1+a}{(1-a)M} + \frac{2a(1-a^M)}{(1-a^2)M^2} \right\} \quad (4.5)$$

“Error-spread relation”: In der perfekten Modell und Ensemble Umgebung hängt jeder Prognose-Fehler des Ensemble-Mittels und jede Dispersion (Varianz) der Einzel-Prognosen um dieses Mittel zusammen, $e_M^2 = s_M^2(M+1)/(M-1) \sim s_M^2$. In der Praxis ist das nicht notwendig der Fall. Dies wird an zwei Fällen verdeutlicht: Abb. 5a zeigt ein „scatter“-Diagramm zwischen den Vorhersage-Fehlern von Ensemble-Mitteln, e_M^2 , und deren Dispersion, s_M^2 für $a=0,8$, $r=1$ und $M=8$ von LAF-Persistenz-Prognosen von rotem Rauschen. Die Korrelation

$$\langle (e_M^2 - E_M) (s_M^2 - S_M) \rangle / \{ \langle (e_M^2 - E_M)^2 \rangle \langle (s_M^2 - S_M)^2 \rangle \}^{1/2}$$

zeigt die Stärke des linearen Zusammenhangs an und kann als Regressionsbeziehung zur Prognose des Vorhersage-Fehlers aus der Dispersion der Prognosen verwendet werden. Hier und für LAF-NWP Modelle erreicht diese Korrelation selten Werte über 0,5. In Abb. 5b ist im Vergleich dazu ein scatter-Diagramm zwischen Fehler und spread eines NWP-Modells dargestellt (MOLTENI et al. 1995, Fig. 5a). Hier ist der rms-Fehler der Kontroll-Vorhersage basierend auf der operationellen Analyse gegen den rms-spread eines aus $M=32$ Elementen bestehenden Ensembles aufgetragen. Die Ensemble-Elemente werden durch Addition von dynamisch definierten Störungen zur operationellen Analyse gebildet. Hierzu verwendet man die sogenannten singulären Vektoren als die am schnellsten wachsenden Moden (siehe z. B. LACARRA und TALAGRAND 1988, MUREAU et al. 1993). Der dargestellte Fall bezieht sich auf Sieben-Tage-Prognosen im Winter 1993 für die Nordhemisphäre. Die error-spread Korrelation beträgt in diesem Fall 0,56, einem relativ hohen Wert, wenn man bedenkt, daß auch in einer perfekten Modell Ensemble Umgebung der spread mit dem Fehler einer individuellen Vorhersage nicht perfekt korreliert sein kann (MURPHY 1988, BARKER 1991).

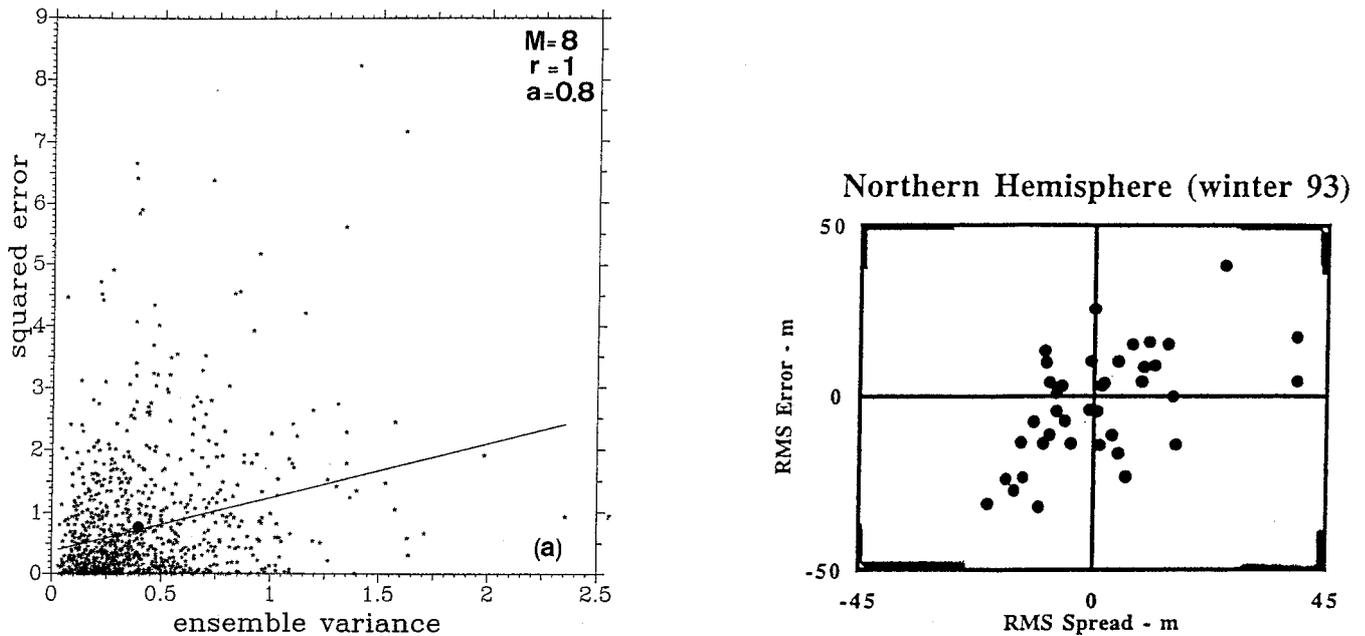


Abb. 5. (a) Scatter-Diagramm der Fehler von Ensemble-Mittel-Vorhersagen ($M=8, a=0,8, r=1$) und der entsprechenden Ensemble-Varianz (spread) (FRAEDRICH und ZIEHMANN-SCHLUMBOHM 1994, Fig. 10a); (b) Scatter-Diagramm der rms-Fehler von Ensemble-Mittel-Vorhersagen des EZMWF ($M=32, r=7$ Tage, Nordhemisphäre Winter 92/93) und dem „spread“ zwischen Ensemble-Mitgliedern und Kontroll-Prognosen (MOLTENI et al. 1995).

Fig. 5. (a) Scatter diagrams of squared errors versus ensemble variance (spread) of a sample of 1000 individual ensemble mean forecasts of lead time $r=1$ in a red-noise atmosphere with autocorrelation $a=0.8$ and ensemble size $M=8$. (b) Scatter-diagram between day-7 rms-error of control forecast (ordinate) versus day-7 rms-spread between ensemble members and control forecasts (abscissa) for winter 1992/93 in the Northern Hemisphere (MOLTENI et al. 1995).

Es zeigt sich, daß die Verwendung von LAF-Ensembles zur Verbesserung der Prognose (1. Ziel) kein optimaler Weg ist. Ebenso wenig ist die Fehler-Prognose aus dem spread zeitverschobener oder LAF-Vorhersagen (2. Ziel) zufriedenstellend, da Dispersion und Fehler nicht stark genug korreliert sind. Der Grund ist, daß die Voraussetzungen vom perfekten Ensemble und Modell nicht erfüllt sind. Ist das Modell schlecht, so können Fälle auftreten, in denen das Ensemble nur eine geringe Dispersion zeigt und die Verifikationstrajektorie einen ganz anderen Weg geht. So bleiben noch (a) alle Ensemble-Mitglieder zu einer Wahrscheinlichkeitsverteilung für zukünftige Wetterzustände zusammenzufassen (das dritte Ziel der Ensemble-Prognose) und (b) dazu geeigneter Ensemble-Mitglieder zu finden. Diese Wege werden zur Zeit am ECMWF (PALMER et al. 1993, MOLTENI et al. 1995) und NMC (TRACTON and KALNAY 1993) erfolgreich besprochen (wie in Abb. 5b gezeigt).

5. Grenzen der Vorhersagbarkeit

Die einfachste Ensemble-Prognose besteht aus $M=2$ aufeinander folgenden Persistenz-Vorhersagen, $F_1(r) = X(t-r) = X(t_0)$ und $F_2(r+R) = X(t-(r+1)) = X(t_0-1)$ und liefert das Modell „Persistenz plus halber Trend“:

$$F = [F_i] = (F_1 + F_2)/2 = F_1 + (F_2 - F_1)/2. \quad (5.1)$$

Das Fehler-Budget $E_2(r) = \langle \{X(t) - F(t)\}^2 \rangle$ sowie seine

Änderung $dE_2/dr = E_{2r}$ sind abhängig von der Prognosezeit

$$E_2(r|X_0) = \langle X_0^2 \rangle \left\{ a^r - \frac{1+a}{2} \right\}^2 + s_x^2 \left\{ 1 - a^{2r} + \frac{1-a^2}{4} \right\}$$

$$E_{2r}(r|X_0) = 2 \langle X_0^2 \rangle \left\{ \frac{1+a}{2} - a^r \right\} a^r \ln \frac{1}{a} + 2s_x^2 a^{2r} \ln \frac{1}{a} \quad (5.2)$$

Im Vergleich zur individuellen Persistenz-Prognose verschwindet hier nicht der Anfangsfehler, sondern er ändert sich, wie auch das Sättigungsniveau, in Abhängigkeit vom Wetterregime. Damit wird dieses Modell in seinem qualitativen Fehler-Verhalten ähnlich den numerischen Wettervorhersagen in der Praxis. Zudem haben DALCHER et al. (1988) gefunden, daß ein Ensemble der Größe $M=2$ für regionale Vorhersagen bereits eine gute Vorhersage des Fehlers (skill) individueller Prognosen liefern kann (siehe auch HOUTEKAMER 1994). Folgende Ergebnisse der Persistenz plus Halb-Trend Vorhersagen erlauben weitergehende Interpretationen.

5.1. Vom Anfangsfehler zur Sättigung

Gleichung (5.2) gemittelt über die Anfangsfehler, d. h. $\langle X_0^2 \rangle = s_x^2$, ergibt den gesamten quadratischen Fehler $E_2(r) = s_x^2 \{ 1 + (1/2 - a^r)(1-a) \}$ (Abb. 6). Anfangs- und Sättigungswert, $E_2(r=0)$ und $E_2(r \rightarrow \infty)$ sind beschränkt: $0 < E_2(r=0) =$

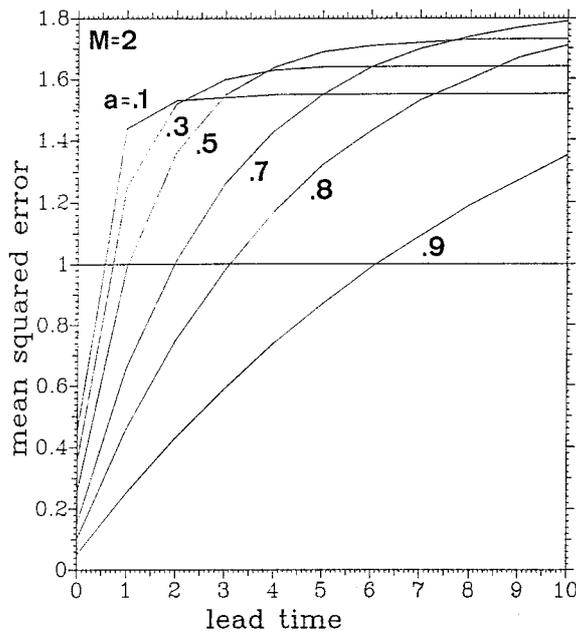


Abb. 6. Fehlerwachstum in Abhängigkeit vom Wetterregime bzw. Anfangsfehler (FRAEDRICH und ZIEHMANN-SCHLUMBOHM 1994, Fig. 7b).

Fig. 6. The lead-time-dependent error variance for various red-noise atmospheres: $a=0.1$ to 0.9 ; note the associated initial errors.

$\frac{1}{2}s_x^2(1-a) < \frac{1}{2}s_x^2$ und $1,5s_x^2 < E(r \rightarrow \infty) = s_x^2\{1,5 + \frac{1}{2}a\} < 2s_x^2$ und hängen vom Wetterregime ab: Mit wachsendem Gedächtnis des Regimes (Zeit-Scale τ oder Auto-Korrelation a) verringert sich der Anfangsfehler *und* das Niveau der Sättigung wird erhöht. Je größer der Anfangsfehler $E_2(r=0) = \frac{1}{2}s_x^2(1-a)$ desto größer ist das Fehlerwachstum $E_{2r}(r=0) = s_x^2(1+a)\ln 1/a$. Das setzt sich fort bei individuellen Prognosen, die bei gleichem Regime geringeres Fehlerwachstum zeigen (da ihr Anfangsfehler verschwindet). Die Ursachen liegen in den großen (kleinen) Fluktuationen des additiven weißen Rauschens, $s_z^2 = s_x^2(1-a^2)$, in Regimen mit kurzem (langem) Gedächtnis: Diese Fluktuationen erzeugen große (kleine) Anfangsfehler $E_2(r=0)$ und auch Wachstumsraten $E_{2r}(r=0)$. Weiter sieht man, daß Regimes mit längerem Gedächtnis oder größerem „scale“ zu kleineren Anfangsfehlern führen. Die Vorhersagbarkeitsgrenze T ist dieselbe wie bei individueller Persistenz-Prognose, $T = \ln \frac{1}{2} / \ln a$. An dieser Grenze ist der systematische Fehler $SE(r=T) = \frac{1}{4}a^2s_x^2$, der nicht-systematische $RE = 1 - SE$. Der Sättigungswert nach langer Vorhersagezeit, $E_2(r \rightarrow \infty)$, ist wegen der Ensemble-Mittelung kleiner als der individueller Prognosen. Großskalige Regimes (mit langem Gedächtnis τ) haben höhere Sättigung $E(r \rightarrow \infty) = s_x^2\{1,5 + \frac{1}{2}a\}$ als die mit kürzerer Zeit-Skala. Der nicht-systematische Fehler ist unbeeinflusst vom systematischen Modell-Fehler und charakterisiert deshalb die Dynamik des Systems. In (5.2) liefert der erste Term mit $\langle X_0^2 \rangle = s_x^2$ den systematischen, der zweite den nicht-systematischen Fehler. Für nicht-systematische Fehler zeigt sich das folgende Bild: Je größer der (integrale Zeit-), „scale“ des Regimes, desto geringer Anfangsfehler und Anfangswachstum. Das gilt auch für das

Gesamt-Fehler-Budget und damit für die praktischen Vorhersage-Experimente, denn der systematische Fehler hat für kurze Prognose-Zeiträume einen vernachlässigbar kleinen Einfluß: Für $r=0$ ist hier der systematische Fehler (und sein Wachstum) kleiner als der nicht-systematische Anteil: $SE_2(0)/s_x^2 = \frac{1}{4}(1-a)^2 < \frac{1}{4}(1-a^2)$ und $d/dr[SE_2(r=0)]/s_x^2 = (a-1)\ln 1/a < 2\ln(1/a)$. Ähnliches gilt auch für praktische NWP-Prognosen, bei denen bekanntlich der systematische NWP-Fehler erst nach längerer Prognose-Zeit an Bedeutung gewinnt (siehe Abb. 1).

5.2. Praktische numerische Wetterprognosen (NWP)

„Upper and lower bound of predictability“: Eine Grenze der Vorhersagbarkeit wird erreicht, wenn, wie hier definiert, die Fehler-Varianz bis zur Klima-Varianz angewachsen ist (Sektion 3). Liefße sich der immer auftretende Anfangsfehler bis auf Null reduzieren, so erreicht die Vorhersagbarkeit ihre hypothetische Untergrenze (lower bound). Die maximal mögliche hypothetische Obergrenze (upper bound) würde erreicht, wenn sich neben dem verschwindenden Anfangsfehler noch der systematische Modell-Fehler auf Null reduzieren liefße. Wie für NWPs lassen auch „Persistenz-plus-Halb-Trend“-Prognosen solche Abschätzungen zu (Tab. 4, Abb. 7):

Gesamt- und nicht-systematischer Anfangsfehler, $E(r=0)$ und $RE(r=0)$, sowie die entsprechenden Prognose-Grenzen, T_E und T_{RE} , bei Erreichen der Klima-Varianz, hängen vom Wetterregime (τ bzw. a) ab (Tab. 4). Wetterregimes zwischen weißem Rauschen und Irrweg-Prozess ($0,8 < a < 0,2$) liefern die Punkte-Paare $\{E(r=0); T_E\}$ und $\{RE(r=0); T_{RE}\}$ in dem Anfangsfehler-Prognoselimit-Diagramm (Abb. 7a,b). Werden $E(r=0)$ bzw. $RE(r=0)$ auf einen verschwindenden Anfangsfehler linear extrapoliert, erhält man Abschätzungen für die Untergrenze $T_E \rightarrow T_U \sim 4$ bzw. Obergrenze $T_{RE} \rightarrow T_0 \sim 8$ der Vorhersagbarkeit. CHEN (1989) hat diesen linearen Zusammenhang zwischen T_E und $E(r=0)$ postuliert aus Analysen NWP-Fehler-Budgets; diese Ergebnisse hat TOTH (1991, Fig. 2) bei Analog-Prognosen bestätigen können: Die Untergrenze der praktischen Wetterprognosen liegt danach bei etwa 25 Tagen.

Im Verlauf einer Prognose reduziert sich das Fehlerwachstum wie sich auch die Fehler-Verdopplung verringert; das bedeutet nicht notwendig, daß am Beginn der Vorhersage die kleineren Fehler immer schneller wachsen als größere; die Fehler-Analyse bei unterschiedlichen Wetterregimes zeigt folgendes Bild:

Regime-Abhängigkeit: PALMER und TIBALDI (1988) assoziieren meteorologische Regimes, deren Stabilitätseigenschaften und unterschiedliche Vorhersagbarkeit. Die größten Anfangsfehler und Wachstumsraten beobachtet man in den kleineren räumlichen Skalen (Wellenzahl > 18). Insbesondere CHEN (1989, Fig. 5,7,9) hat dies demonstrieren können für die Experimente mit praktischen NWP-Vorhersagen (Abb. 8c-d). Untersuchungen von HOREL und ROADS (1988, Abb. 8a-b) mit theoretischen und von TOTH (1991,

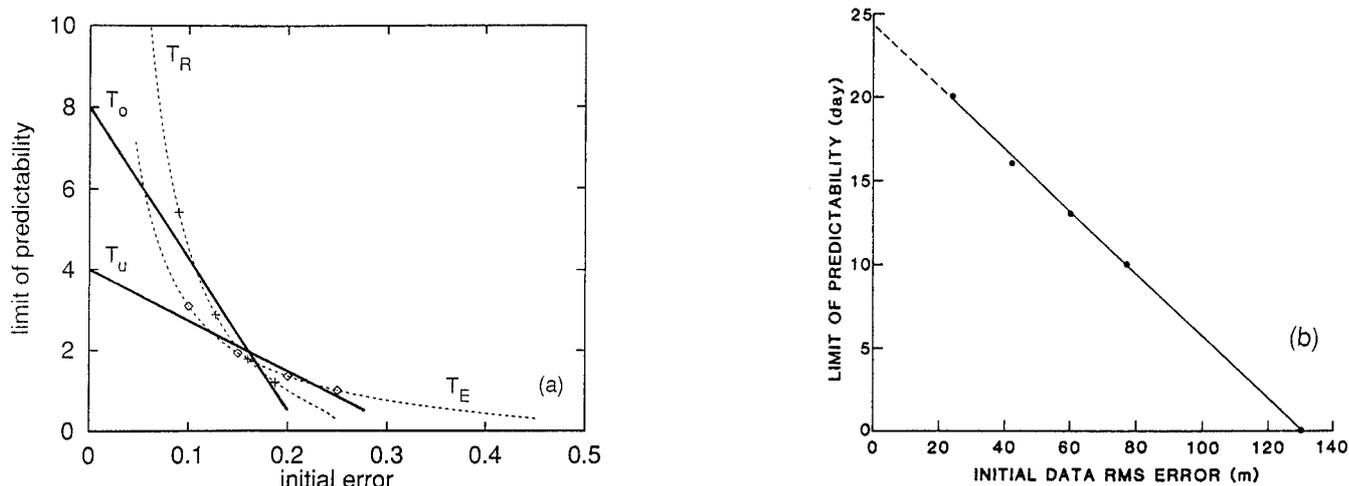


Abb. 7. Grenzen der Vorhersagbarkeit und Anfangsfehler: (a) Grenzen der Vorhersagbarkeit des nicht-systematischen, T_R , und des Gesamt-Fehlers, T_E . Die Abhängigkeit vom Anfangsfehler, $RE(r=0)$ bzw. $E(r=0)$, wird im Persistenz-plus-Halb-Trend-Modell durch die unterschiedlichen Wetterregimes parametrisiert. Lineare Extrapolationen auf verschwindenden Anfangsfehler liefern die Ober- und Untergrenze der Vorhersagbarkeit. (b) Schätzung der Untergrenze der Vorhersagbarkeit (CHEN 1989, Fig. 8).

Fig. 7. Limit of predictability versus initial error: a) Limit of predictability of the random, T_R , and total error, T_E . In the persistence-half-trend-model, the initial error, $RE(r=0)$ bzw. $E(r=0)$, is parameterized through different weather regimes or time scales. Linear extrapolation leads to the upper and lower bounds of the limit of predictability. (b) Estimate of the lower bound of predictability (CHEN 1989, Fig. 8).

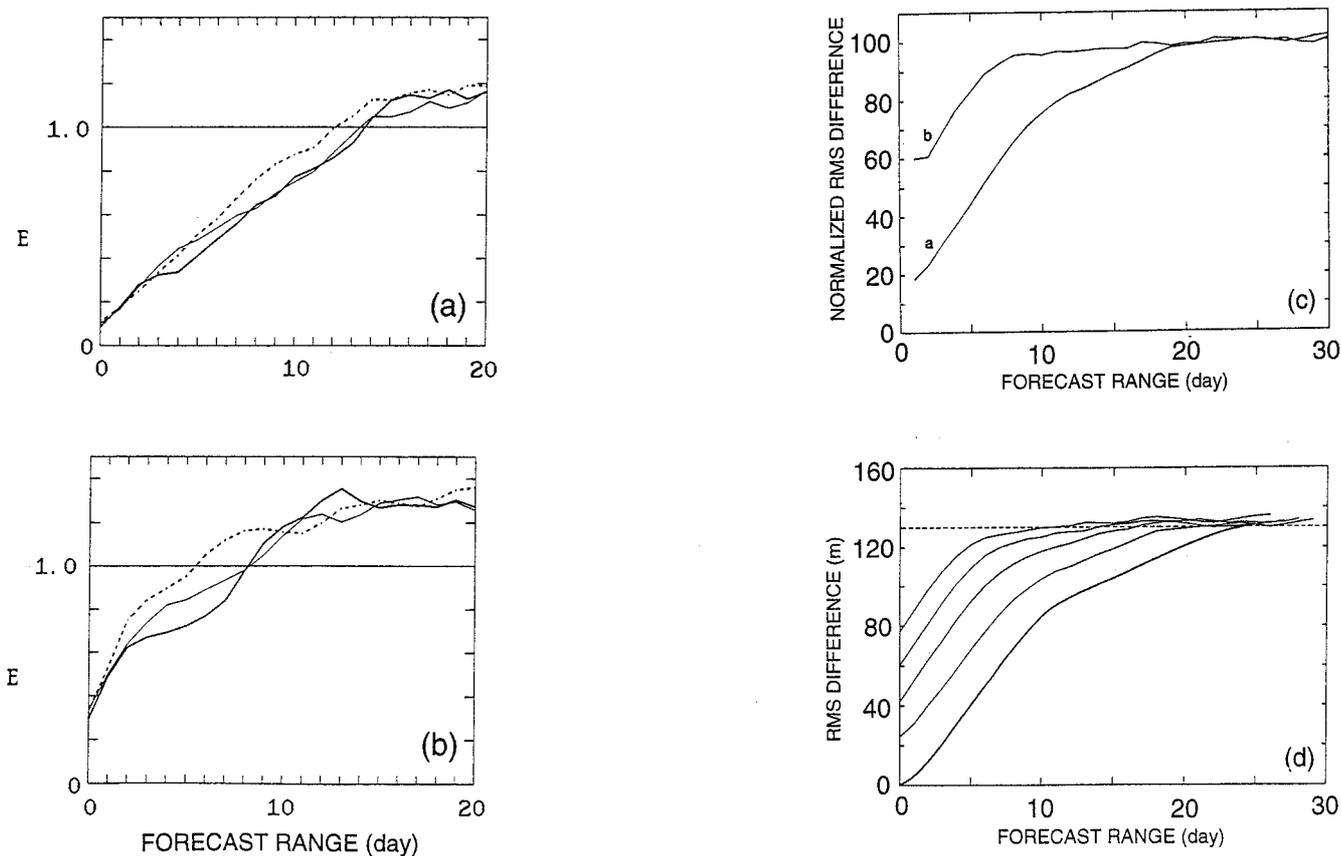


Abb. 8. Skalen-Abhängigkeit des Fehlerwachstums in NWP-Modellen: Fehler-Wachstumskurven von kleinen (a) und großen (b) Anfangsfehlern (HOREL und ROADS 1988, Fig. 9); (c) Wachstumskurven von Fehlern in kleineren „scales“ (Wellenzahl >18 , Kurve b) im Vergleich zum Gesamt-Fehler (CHEN 1989, Fig. 4); (d) Zusammenfassung von Wachstumskurven eines Spektrums von vier verschiedenen Anfangsfehler-Klassen einschließlich einer qualitativen Extrapolation auf verschwindenden Anfangsfehlern dargestellt als dicke Linie (CHEN 1989, Fig. 9).

Fig. 8. Scale dependence of error-growth in NWP-models: error growth of small (a) and large (b) initial errors (HOREL and ROADS 1988, Fig. 9); (c) error growth of the small scale regime (wavenumber $k>18$) compared to the total (CHEN 1988, Fig. 4); (d) spectrum of error growth curves commencing at different initial errors including a qualitative extrapolation to vanishing initial error (bold line, CHEN 1989, Fig. 9).

Tab. 4. Anfangsfehler ($r=0$) und Vorhersagbarkeitsgrenze ($r=T$).

Fehler	Anfangsfehler	Vorhersagbarkeitsgrenze	RE, E(0)→0
Gesamt E_2	$E_2(0) = \frac{1}{2}(1-a)$	$E_2(T) = s_x^2 : T_E = \ln \frac{1}{2} / \ln a$	$T_U \sim 4$
Nicht-Sys. RE ₂	$RE_2(0) = \frac{1}{4}(1-a^2)$	$RE_2(T) = s_x^2 : T_R = \ln \frac{1}{4}(1-a^2) / \ln a^2$	$T_O \sim 8$

Fig. 5) mit historischen Wetter-Analog-Prognosen zeigen ähnliches: „One can only guess about the cause of the phenomenon. The structure of the same magnitude of error, its wavenumber or geographical distribution, might be different with various initial errors“ (TOTH 1991). Die hier dargestellten Experimente von Persistenz-Prognosen für rotes Rauschen könnten die Antwort liefern: Je größer der (integrale Zeit-) „scale“ τ des dynamischen Regimes, innerhalb dessen prognostiziert wird, desto kleiner ist der Anteil interner Fluktuationen und desto geringer der Analyse- oder Anfangsfehler für das Prognose-Modell und damit das Wachstum der Prognose-Fehler (Abb. 6).

6. Diskussion und Zusammenfassung

Die praktische (oder externe) Vorhersagbarkeit von imperfekten numerischen Wetterprognosen (NWP) wird durch konzeptionelle Prognose-Experimente interpretiert: Dabei dient Persistenz als Prognose-Modell und rotes Rauschen als Ersatz-Atmosphäre, deren Auto-Korrelationszeiten unterschiedliche Wetterregimes charakterisieren. Individuelle und Ensemble-Vorhersagen werden analysiert: (i) Aus dem Fehler-Budget individueller Prognosen werden unterschiedliche Maße der Vorhersagbarkeit (analytisch) bestimmt: Anfangsfehler, Fehlerwachstum, Grenzen der Vorhersagbarkeit, Fehler und Fehler-Quadrat-Verteilungen in Abhängigkeit von den Anfangsbedingungen, systematische und nicht-systematische Fehler, etc. (ii) Ensemble-Prognosen verbessern die Vorhersage, prognostizieren Vorhersage-Fehler und Wahrscheinlichkeiten atmosphärischer Zustände unter perfekten Bedingungen. Unter imperfekten Bedingungen (mit den zeitlich verschobenen Ensemble-Vorhersagen von imperfekten Modellen) werden diese Ziele nur bedingt erfüllt. Trotz der Einfachheit dieser konzeptionellen Experimente zur praktischen Vorhersagbarkeit zeigen die Fehler-Budgets Merkmale, die qualitativ mit denen numerischer Wetterprognosen übereinstimmen.

Danksagung

Diese Arbeit entstand auf Anregung der Herausgeber der Meteorologischen Zeitschrift; sie ist eine Art deutschsprachige Zusammenfassung einer Publikation von FRAEDRICH und ZIEHMANN-SCHLUMBOHM (1994) und Teil der Vorlesung „Theoretische Meteorologie“ von FRAEDRICH, auf das wesentliche reduziert und um Anwendungsbeispiele erweitert. Ein zweiter Teil „internal or theoretical predictability, ensemble forecasting etc.“ ist von uns geplant.

Literatur

- Barker, T. W., 1991: The relationship between spread and forecast error in extended-range forecasts. — *J. Climate* **4**, 733–742.
- Brankovic, C., T. N. Palmer, F. Molteni, S. Tibaldi, U. Cubasch, 1990: Extended range prediction with ECMWF models: Time lagged ensemble forecasting. — *Quart. J. Roy. Meteor. Soc.* **116**, 867–912.
- Chen, W. Y., 1989: Estimate of dynamical predictability from NMC DERF experiments. — *Mon. Wea. Rev.* **117**, 1227–1236.
- Dalcher, A., E. Kalnay, 1987: Error growth and predictability in operational ECMWF forecasts. — *Tellus* **39A**, 474–491.
- Dalcher, A., E. Kalnay, R. Hoffman, 1988: Medium range lagged average forecasts. — *Mon. Wea. Rev.* **116**, 402–416.
- Dole, R., N. Gordon, 1983: Persistent anomalies of the extratropical northern hemisphere wintertime circulation: Geographical distribution and regional persistence characteristics. — *Mon. Wea. Rev.* **111**, 1567–1586.
- Fraedrich, K., 1992: Combination of weather forecasts: Some applications. — 12th Conference on Probability and Statistics, 22–26 June 1992, Toronto, J45–J49.
- Fraedrich, K., Th. Dümmel, 1983: On single station forecasting: the geopotential height, its vertical and time structure and 500mbar ARMA prediction. — *Contrib. Atmos. Phys.* **56**, 221–239.
- Fraedrich, K., Chr. Ziehmann-Schlumbohm, 1994: Predictability experiments of persistence forecasts in a red noise atmosphere. — *Quart. J. Roy. Meteor. Soc.* **120**, 387–428.
- Gutzler, D. S., K. C. Mo, 1983: Autocorrelation of northern hemisphere geopotential heights. — *Mon. Wea. Rev.* **111**, 155–164.
- Hasselmann, K. 1979: On the problem of multiple time scales in climate modelling. — In: Bach, J., J. Pankrath, W. Kellog (eds.): *Man's Impact on Climate*, 43–56. — Elsevier.
- Hoffman, R., E. Kalnay, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. — *Tellus* **35A**, 100–118.
- Horel, J. D., J. O. Roads, 1988: Sensitivity of regional predictability to flow characteristics. — *J. Geophys. Res.* **93**, 11005–11014.
- Houtekamer, P. L., J. Derome, 1994: Prediction experiments with two-member ensembles. — *Mon. Wea. Rev.* **122**, 2179–2191.
- Lacarra, J.-F., O. Talagrand, 1988: Short-range evolution of small perturbations in a barotropic model. — *Tellus* **40A**, 81–95.
- Leith, C. C., 1974: Theoretical skill of Monte Carlo forecasts. — *Mon. Wea. Rev.* **102**, 409–418.
- Lorenz, E. N., 1969: Atmospheric predictability as revealed by naturally occurring analogues. — *J. Atmos. Sci.* **26**, 636–646.
- 1975: Climatic predictability. *The Physical Basis of Climate Modelling*. — WMO, GARP Publ. Ser. **16**, 132–136.
- Molteni, F., R. Buizza, T. N. Palmer, T. Petroliaigis, 1995: The ECMWF prediction system: Methodology and validation. — *Quart. J. Roy. Meteor. Soc.* **120**, in press.
- Mureau, R., F. Molteni, T. N. Palmer, 1993: Ensemble prediction using dynamically-conditioned perturbations. — *Quart. J. Roy. Meteor. Soc.* **119**, 299–324.
- Murphy, J. M., 1988: The impact of ensemble forecasts on predictability. — *Quart. J. Roy. Meteorol. Soc.* **114**, 463–493.

- Palmer, T. N., S. Tibaldi, 1988: On the prediction of forecast skill. — *Mon. Wea. Rev.* 116, 2453–2480.
- Palmer, T. N., F. Molteni, R. Mureau, R. Buizza, P. Chapelet, J. Tribbia, 1993: Ensemble prediction. — ECMWF Seminar Proceedings 'Validation of models over Europe: Vol. 1', ECMWF, Shinfield Park, UK, 285 pp.
- Stroe, R., J. F. Royer, 1993: Comparison of different error growth formulas and predictability estimation numerical extended range forecasts. — *Ann. Geophys.* 11, 296–316.
- Tennekes, H., 1991: Karl Popper and the accountability of numerical forecasting. — *New Developments in Predictability, ECMWF Workshop*, 21–28.
- Toth, Z., 1991: Estimation of atmospheric predictability by circulation analogs. — *Mon. Wea. Rev.* 119, 65–72.
- Tracton, M. S., E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: practical aspects. — *Weather and Forecasting* 8, 379–398.

K. FRAEDRICH
CHR. ZIEHMANN
Meteorologisches Institut
Universität Hamburg
Bundesstr. 55
D-20146 Hamburg

Eingereicht 5. Januar 1995, in revidierter Form: 21. April 1995