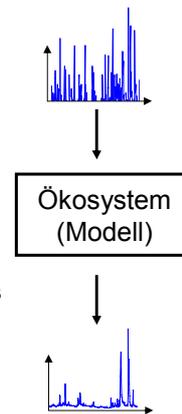


Informationsgehalt und Komplexität von Zeitreihen

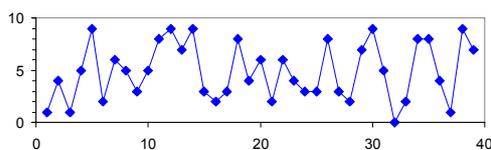
Motivation:

- Viele natürliche Systeme „verbrauchen“ Information: der Output ist weniger „komplex“ als der Input
- Thermodynamischer Ansatz: Belebte Systeme erhalten Strukturen (unter Energieverbrauch) und verhindern so die Zunahme von Entropie
- Verschiedene Zeitreihen sehen unterschiedlich „komplex“ aus => unterschiedlich schwer zu simulieren (?)
- Der Output von Modellen ist oft weniger „komplex“ als die beobachteten Daten

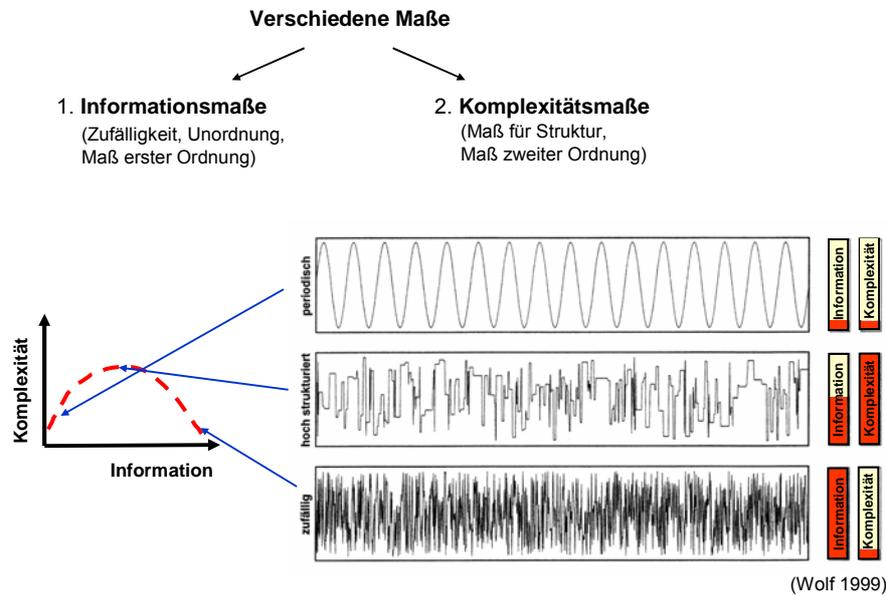


Informationsgehalt und Komplexität von Zeitreihen

- Keine einheitliche Definition, viele verschiedene Maße
- Zwei Kategorien:
 1. **Informationsmaß:** Maß für die „Überraschung“ oder den „Neuigkeitswert“ einer Beobachtung
 2. **Komplexitätsmaß:** Maß für die (Nicht-)Komprimierbarkeit der Daten
 - = Länge der kürzesten Beschreibung (**Kolmogorov-Komplexität**)
 - prinzipiell nicht exakt, sondern nur approximativ bestimmbar



Information und Komplexität von Zeitreihen



Chaos und Attraktoren

Chaos: kleine Abweichungen der Startbedingungen führen zu stark divergierendem Verhalten (*Lyapunov-Koeffizient*)

Attraktoren:

- *für eine Funktion:* Attraktoren sind Werte, denen sich die Funktionswerte annähernd, auch wenn die Startwerte sehr unterschiedlich sind
- *für eine univariate Zeitreihe:* Attraktoren sind Abfolgen von Werten, die exakt oder annähernd immer wieder auftreten
- „Gegenteil“: *Bifurkationen*

Logistische Abbildung

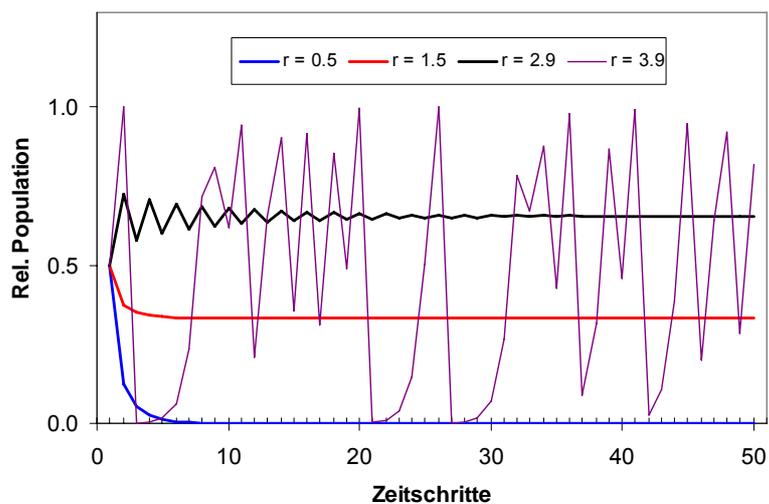
Beschreibt die Entwicklung der Population einer Art mit nicht überlappenden Generationen:

$$x_{n+1} = rx_n \cdot (1 - x_n)$$

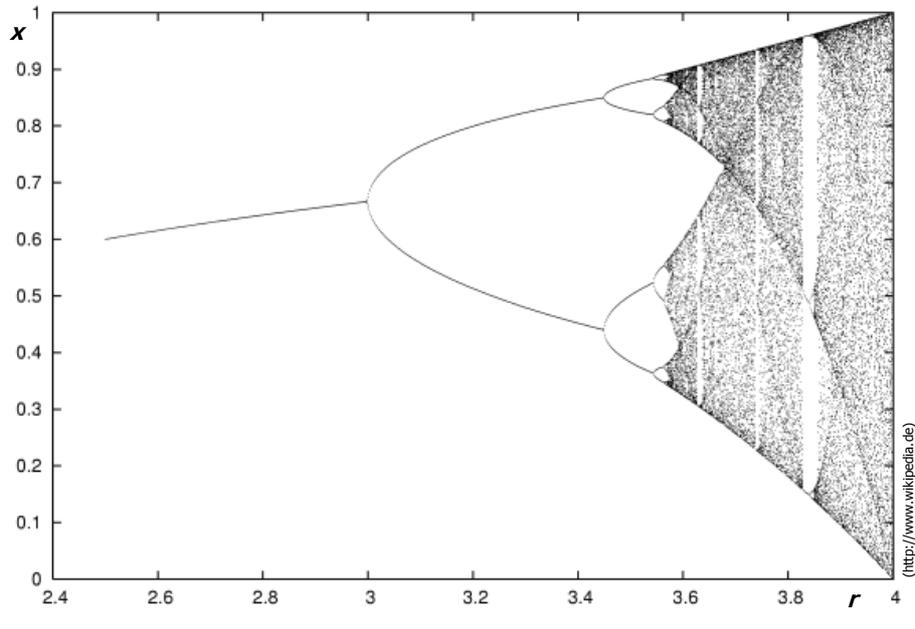
mit x = rel. Population (bezogen auf maximal mögliche Populationsgröße) und Wachstumsparameter r :

- $0 < r < 1$: Aussterben der Population
- $1 < r < 2$: monotone Annäherung an den Grenzwert
- $2 < r < 3$: oszillatorische Annäherung an den Grenzwert
- $r > 3$: deterministisches Chaos mit wiederholten Bifurkationen
- $r = 4$: voll entwickeltes Chaos
- Grafische Darstellung: Feigenbaum-Diagramm

Logistische Abbildung



Feigenbaum-Diagramm



Partitionierung

Informations- und Komplexitätsmaße werden nicht direkt auf die $x(t)$ -Werte angewendet, sondern auf *partitionierte* Daten

Definition: Eine *Partitionierung* P einer Menge X ist eine endliche Familie von Teilmengen $\{A_1, \dots, A_\lambda\}$, $A_i \subset X$ mit den Eigenschaften

- Disjunktheit: $A_i \cap A_j = \emptyset$ ($i \neq j$)

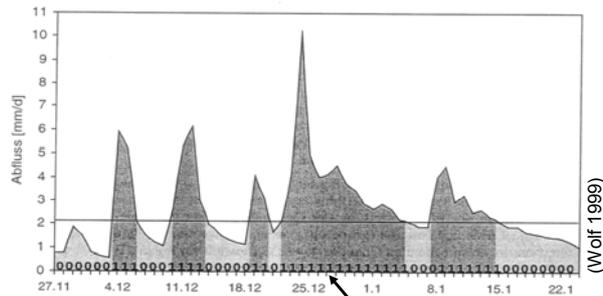
- Vollständigkeit: $\bigcup_{i=1}^{\lambda} A_i = X$

→ i.d. R. binäre Partitionierung $(0, 1)$

Statische Partitionierung

→ Vergleich mit Schwellwert (z.B. Median):

$$z = \begin{cases} 1 & \text{für } x(t) \geq x^* \\ 0 & \text{für } x(t) < x^* \end{cases}$$



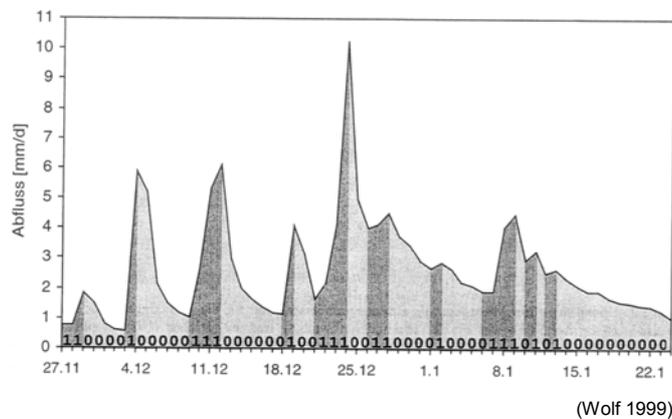
z.B. binäres Alphabet
Symbol $\in \{0,1\}$

Symbolfolge S
(00000011100...)

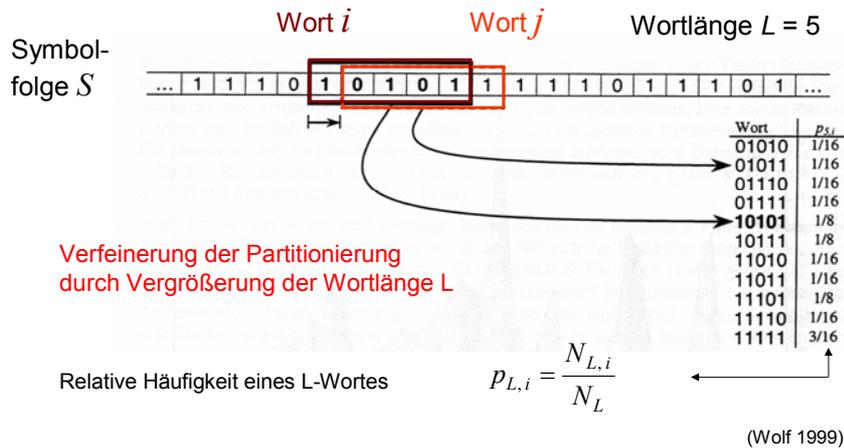
Dynamische Partitionierung

→ Transformation anhand der 1. Ableitung:

$$z = \begin{cases} 1 & \text{für } [x(t) - (x(t-1))] \geq 0 \\ 0 & \text{für } [x(t) - (x(t-1))] < 0 \end{cases}$$



L-Wörter und Verfeinerungen



Shannon-Entropie

Beispiel: Wurf einer (gezinkten) Münze (binärer *Bernoulli-Prozess*)

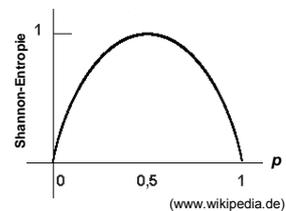
- Wahrscheinlichkeit, dass eine bestimmte Seite oben liegt: p mit $0 \leq p \leq 1$
- Wahrscheinlichkeit, dass die andere Seite oben liegt: $1-p$ mit $0 \leq p \leq 1$
- Maximale Unsicherheit für $p = 0.5$; minimale Unsicherheit für $p = 0$ bzw. $p = 1$

Definition:

$$H_S = -k \sum_{i=1}^{\lambda} p_i \log p_i$$

mit p_i = Wahrscheinlichkeit, mit der das i -te Symbol des Alphabets der Länge λ auftaucht

- k : Skalierungsparameter, für $k = 1/\log \lambda \Rightarrow 0 \leq I \leq 1 \Rightarrow$ Angabe von H_S in bits; dann für $H_S = 1$: zufällig

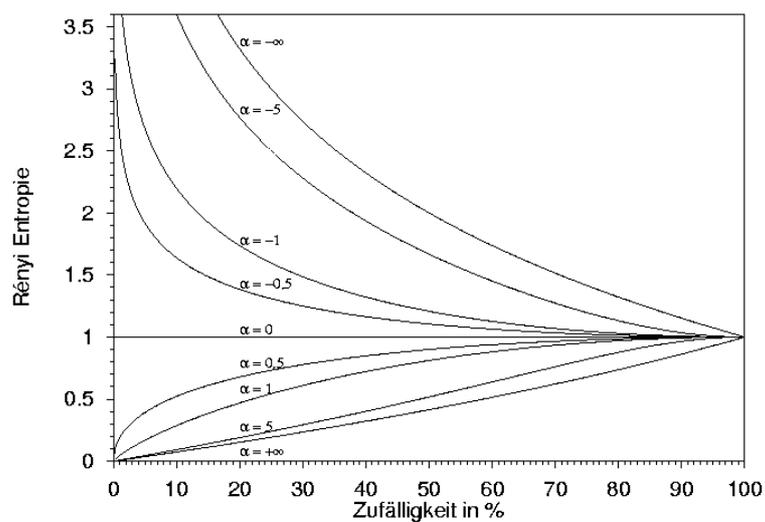


Weitere Informationsmaße

- **Shannon-Entropie:**
$$H_S = -\sum_{i=1}^{\lambda} p_i \log_2 p_i$$
- **Metrische Entropie:** (Normierung mit Wortlänge)
$$H_{\mu}(L) = \frac{H_S}{L} \quad (0 \leq H_{\mu} \leq 1 \quad \text{für } \lambda = 2)$$
- **Rényi-Entropie:**
$$H_R(\alpha) = \frac{1}{1-\alpha} \cdot \log_2 \sum_{i=1}^{\lambda} p_i^{\alpha} \quad (\text{für } \alpha \neq 1)$$
- **Mittlerer Informationsgewinn:**
$$H_G(L) = H_S(L+1) - H_S(L)$$

(MIG = Mean Information Gain)

Rényi-Entropie für den Bernoulli-Prozess



Weitere Informationsmaße

- **Shannon-Entropie:** $H_S = -\sum_{i=1}^{\lambda} p_i \log_2 p_i$
- **Metrische Entropie:** $H_{\mu}(L) = \frac{H_S}{L}$ ($0 \leq H_{\mu} \leq 1$ für $\lambda = 2$)
(Normierung mit Wortlänge)
- **Rényi-Entropie:** $H_R(\alpha) = \frac{1}{1-\alpha} \cdot \log_2 \sum_{i=1}^{\lambda} p_i^{\alpha}$ (für $\alpha \neq 1$)
- **Mittlerer Informationsgewinn:** $H_G(L) = H_S(L+1) - H_S(L)$
(MIG = Mean Information Gain)

Informationsgewinn G_{ij}

$\rightarrow G_{ij}$
 0000**10101**000 L = 5
 Wort *i* Wort *j*

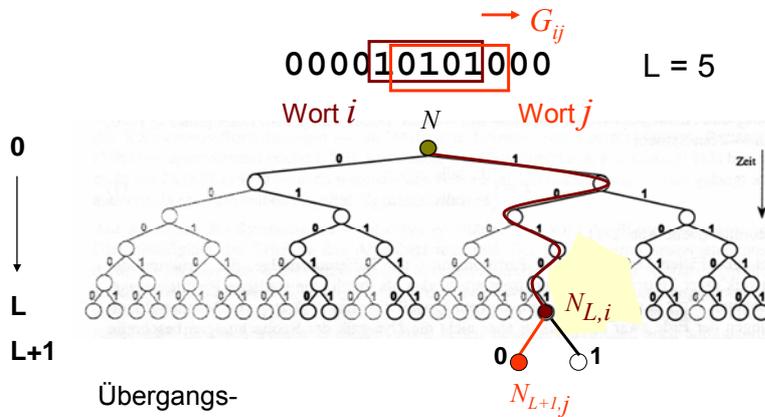
Veranschaulichung aus der
Nachrichtenübermittlung:

kleiner Informationsgewinn
(da „e“ sehr wahrscheinlich)

„... so gestatten wir uns zum
Abschluss dieser Erläuterung die
spekulative **Frage**, ob das Kontinuum
möglicherweise nur das Resultat
eines Abstraktionsprozesses durch
das menschliche Bewusstsein ist.“

großer Informationsgewinn
(da nach „das_“ jeder mögliche Buchstabe kommen kann)

Wortübergangs-Wahrscheinlichkeiten



Übergangswahrscheinlichkeit

$$P_{L,i \rightarrow j} = \frac{N_{L+1,j}}{N_{L,i}}$$

Informationsgewinn

$$G_{ij} = -\log_2 P_{L,i \rightarrow j}$$

Informationsmaß MIG

Wort *i* 000010101000 Wort *j*

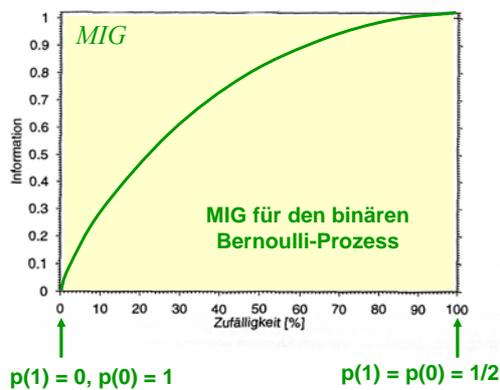
Informationsgewinn G_{ij}

Information = Maß für Zufälligkeit

mittlerer Informationsgewinn

$$MIG = \sum_{i,j=1}^n p_{L+1,j} \cdot G_{ij}$$

$$G_{ij} = -\log_2 p_{L,i \rightarrow j}$$



Quantifizierung von Komplexität

Komplexität = „Schwierigkeit einer kompakten (einfachen, kurzen) Beschreibung“

- kaum veränderliche Symbolfolgen sind *wenig komplex*
- ganz zufällige Datensätze lassen sich einfach beschreiben (Art des Rauschens, Mittelwert, Standardabweichung) und sind deshalb *wenig komplex*
- stark strukturierte Signale mit zufälligen und nicht-zufälligen Anteilen sind *komplex*

Komplexitätsmaße

Effektive Maßkomplexität: Geschwindigkeit des Informationszuwachses bei zunehmender Wortlänge

$$C_{EM} = \sum_{L=1}^{\infty} (H_G(L) - \lim_{L \rightarrow \infty} H_G)$$

Fluktuationskomplexität (FC): Schwankungen um den mittleren Netto-Informationsgewinn ($\lambda = \text{Alphabetgröße}$)

$$C_F = \sigma_{FC}^2 = \sum_{i,j=1}^{\lambda^L} p_{L,ij} \left(\log_2 \frac{p_{L,i}}{p_{L,j}} \right)^2$$

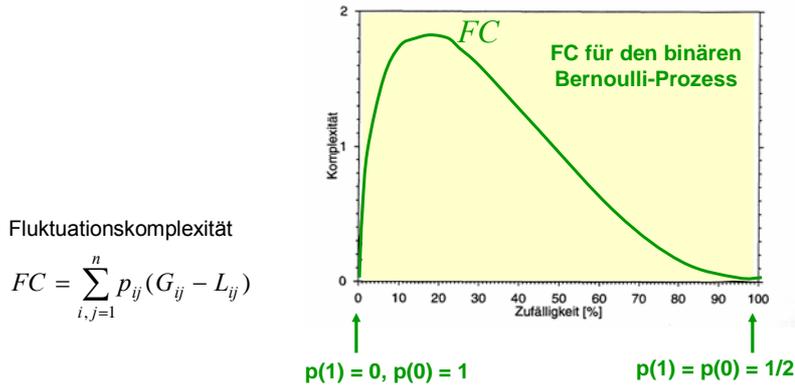
Rényi-Komplexität (RC): 2.-Ordnung-Verhalten durch Differenzenbildung (*entspricht FC für $\alpha=1$*)

$$C_R(\alpha, L) = \frac{2}{(\alpha - 1) \cdot L \cdot \ln 2} \cdot [H_R(\frac{1}{\alpha}) - H_R(\alpha)]$$

Komplexitätsmaß FC

Informationsverlust L_{ij} Informationsgewinn G_{ij}

Wort i 000010101000 Wort j



FC und Rényi-Komplexität

Informationsverlust L_{ij} Informationsgewinn G_{ij} Netto-Informationsgewinn = $G_{ij} - L_{ij}$

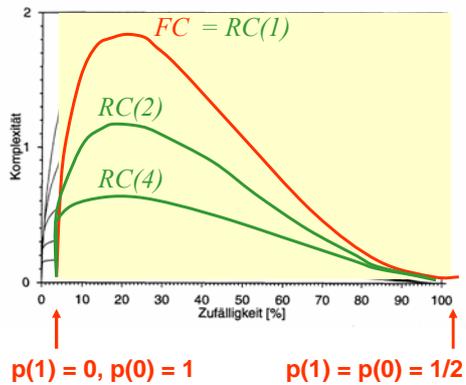
Wort i 000010101000 Wort j

Fluktuationskomplexität

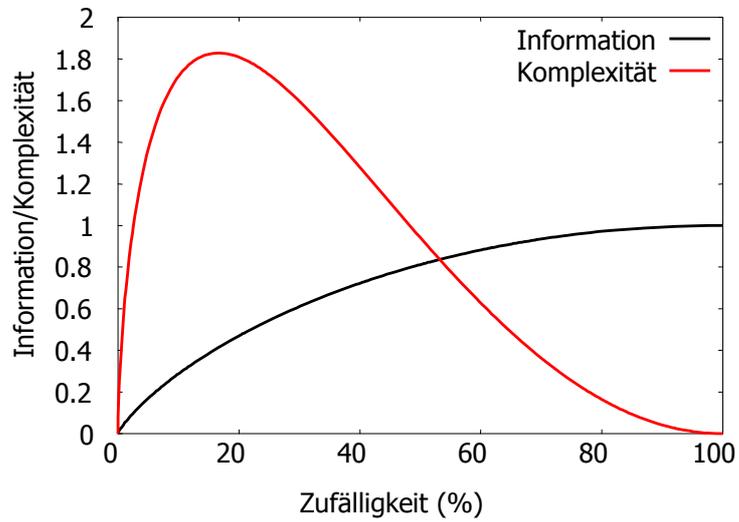
$$FC = \sum_{i,j=1}^n p_{ij}(G_{ij} - L_{ij})$$

Rényi-Komplexität

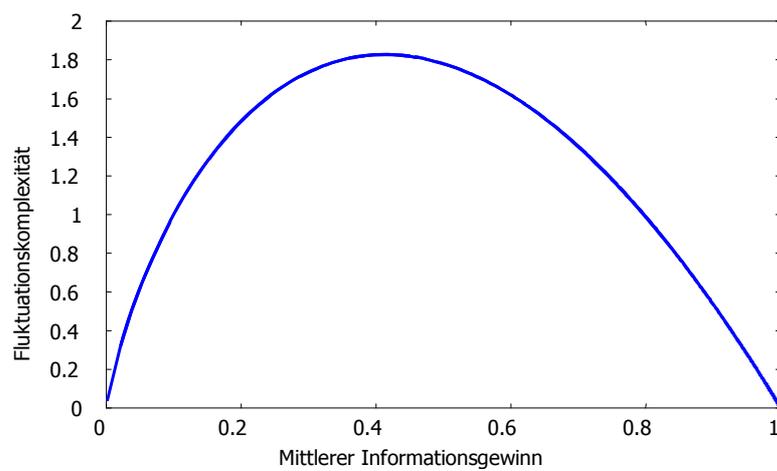
$$RC(\alpha) = \frac{1}{(\alpha - 1)L \ln 2} \left(H_R\left(\frac{1}{\alpha}\right) - H_R(\alpha) \right)$$



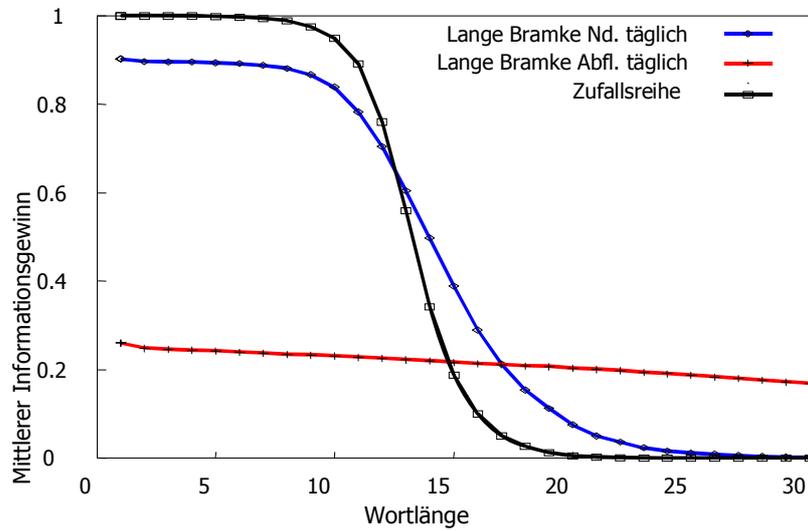
Ergebnisse für Bernoulli-Reihen



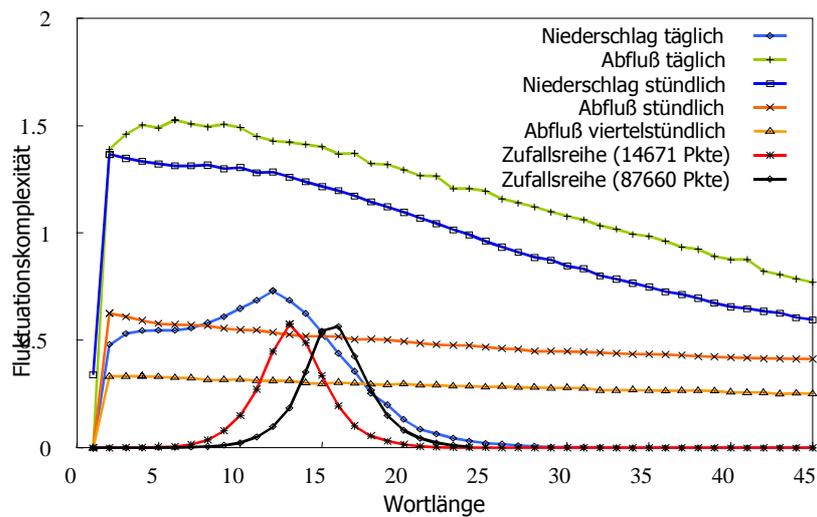
Informations-Komplexitäts-Diagramme: „MIG-FC-Plot“



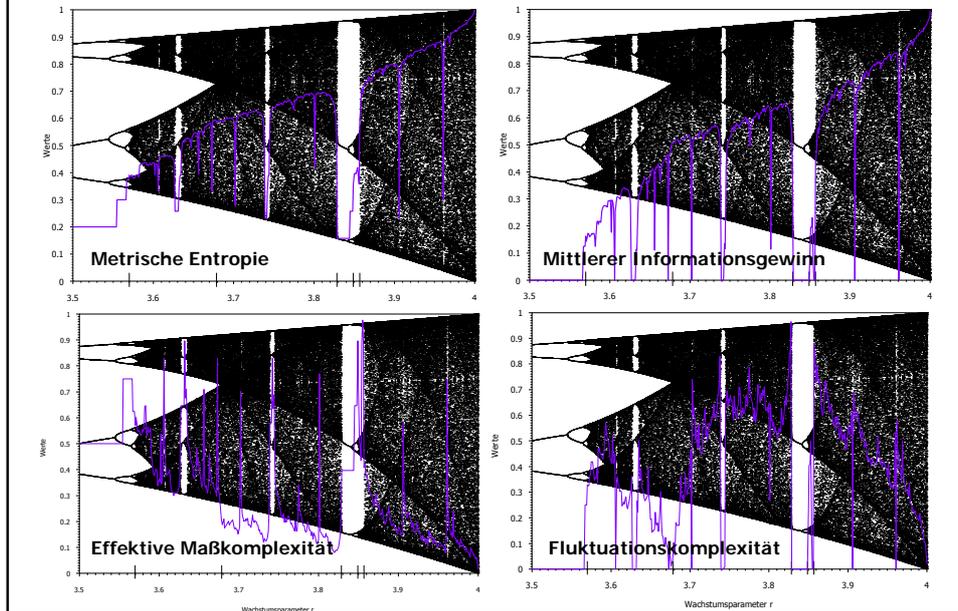
Der Einfluss der Wortlänge auf MIG



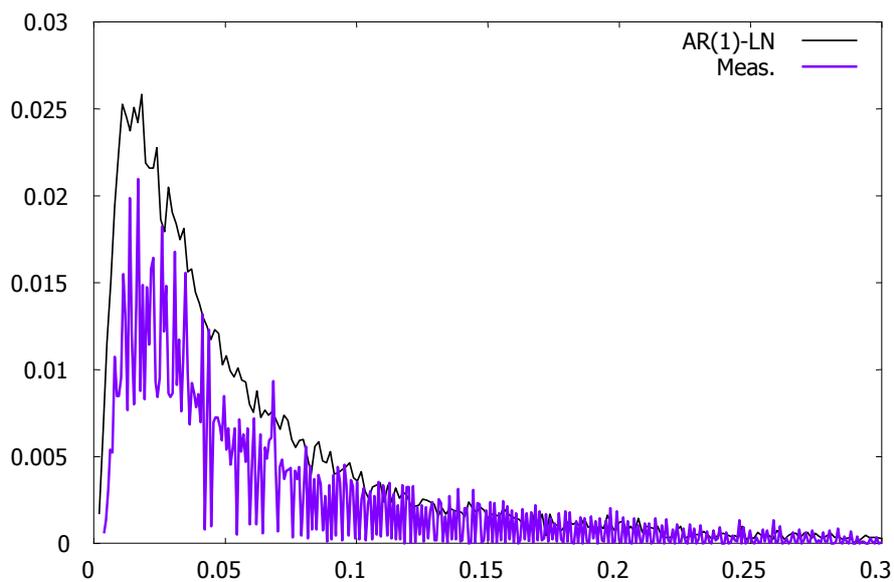
Der Einfluss der Wortlänge auf FC



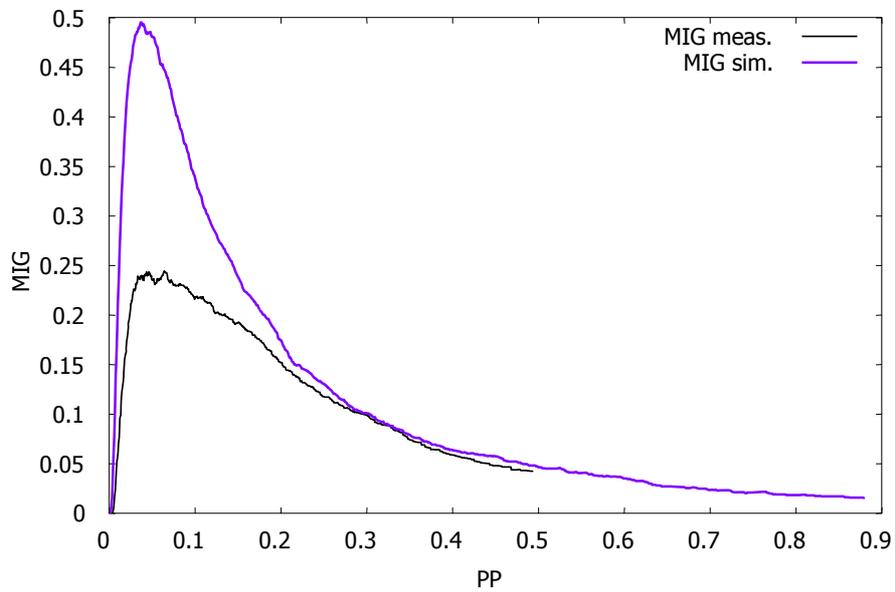
Bsp. Logistische Abbildung (Wolf 1999)



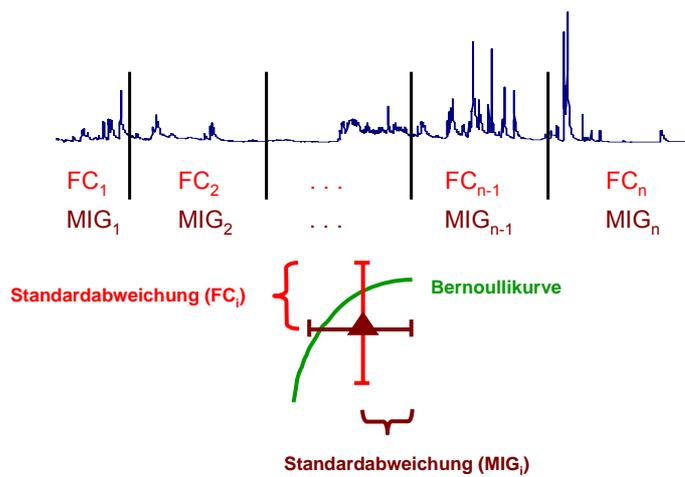
Histogramm LB Abfluss und AR(1)-LN Modell



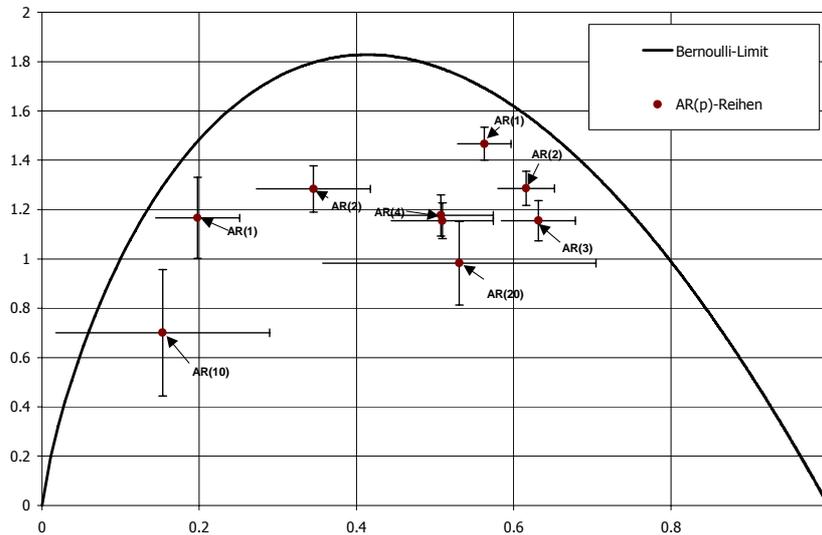
LB Abfluss: beobachtet und AR(1)-LN Model



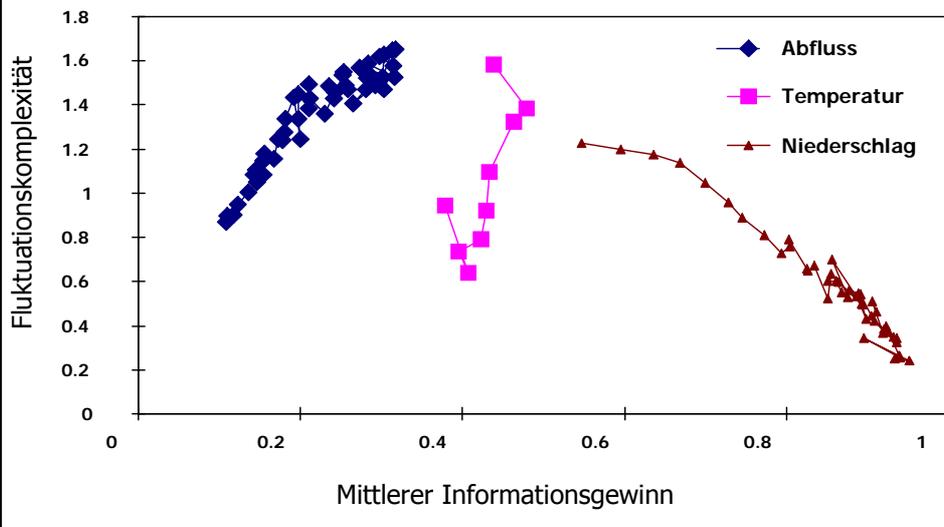
Fehlerschätzung der Maße



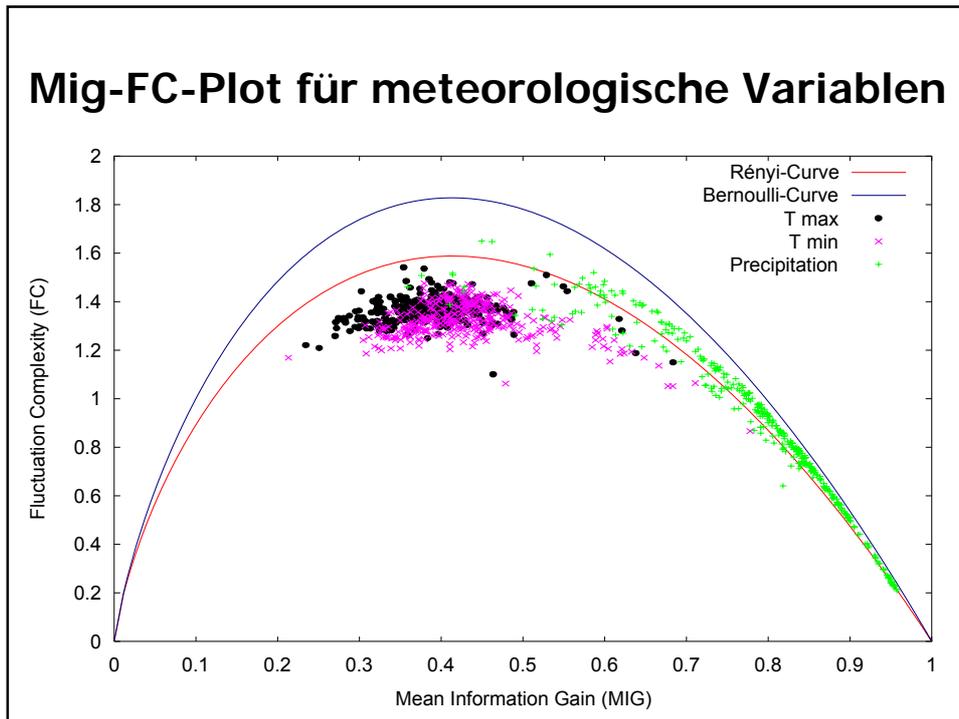
MIG-FC für AR-Rekonstruktionen



Mig-FC-Plot für unterschiedliche zeitliche Aggregation



Mig-FC-Plot für meteorologische Variablen



Abflüsse weltweit

