Einführung: Extremwertstatistik in der Hydrologie

- Aufgaben der Extremwertstatistik in der Hydrologie
- Grundbegriffe
- Ermittlung der Verteilungsfunktion für hydrologische Extremwerte
- Anwendung der bekannten Verteilungsfunktion

2002 S. Itzerott (GFZ) und M. Zebisch (TUB/PIK)

Aufgabe der Extremwertstatistik in der Hydrologie

- Viele wasserwirtschaftliche und verkehrstechnische Anlagen sind so bemessen, dass sie auch in Extremsituationen (hohe Niederschlags-intensitäten, Durchflüsse, Wasserstände) funktionstüchtig bleiben bzw. ihre Schutzwirkung erfüllen.
- Bei der Festlegung der Bemessungswerte ist der mögliche Schaden beim Versagen der Anlage (z.B. Deiche, Talsperren) entscheidend.
- Für eine unter dem Kosten-Nutzen Aspekt sinnvolle Bemessung (z.B. Deiche, Durchlässe, Entwässerungskanäle, Rückhaltebecken etc.) werden Aussagen zur Auftretenswahrscheinlichkeit von Ereignissen bestimmter Intensität benötigt:

- *Wie häufig tritt ein Niederschlag bestimmter Intensität und Dauer im Mittel auf?*
- *Mit welchem maximalen Durchfluss / Wasserstand muss innerhalb einer Zeitspanne statistisch gerechnet werden?*

- Die statistische Auswertung bereits beobachteter Extremwerte ist eine Möglichkeit, solche Aussagen zu liefern.

Begriffsdefinitionen Extremwertstatistik

Wahrscheinlichkeit: Anzahl der Fälle in denen ein Ereignis eintritt geteilt durch Anzahl der möglichen Fälle

Unterschreitungswahrscheinlichkeit P_U:	$P_U = P(x \leq x_i)$
Überschreitungswahrscheinlichkeit $P_Ü$:	$P_Ü = P(x > x_i)$
Beziehung zwischen P_U und $P_Ü$:	$P_Ü + P_U = 1$
Wiederkehrintervall oder Jährlichkeit:	$T = 1 / P_Ü$

T ist die durchschnittliche Zeitspanne innerhalb der ein Ereignis x auftritt, welches einen Schwellenwert x_i übertrifft.

Meist werden jährliche Extremwerte betrachtet und T wird in Jahren angegeben („Jährlichkeit“).

Beispiele zu den Begriffen P_U, $P_Ü$, T

$P_U = P(x \leq x_i)$ $P_Ü = P(x > x_i)$ $P_Ü + P_U = 1$ $T = 1 / P_Ü$

Frage 1

Ein Durchfluss von 100 m³/s wird an einem Pegel im Mittel einmal in 5 Jahren überschritten.

Wie groß sind Über- und Unterschreitungswahrscheinlichkeit für Q=100m³/s? Welche Einheiten besitzen diese Werte?

Frage 2

Die Unterschreitungswahrscheinlichkeit eines Durchflusses von 4250 m³/s wird mit P_U= 0.998 angegeben. Alle wieviel Jahre muss statistisch gesehen damit gerechnet werden, dass ein Wert > 4250 m³/s auftritt?

Ziel der extremwertstatistischen Auswertung & allg. Vorgehen

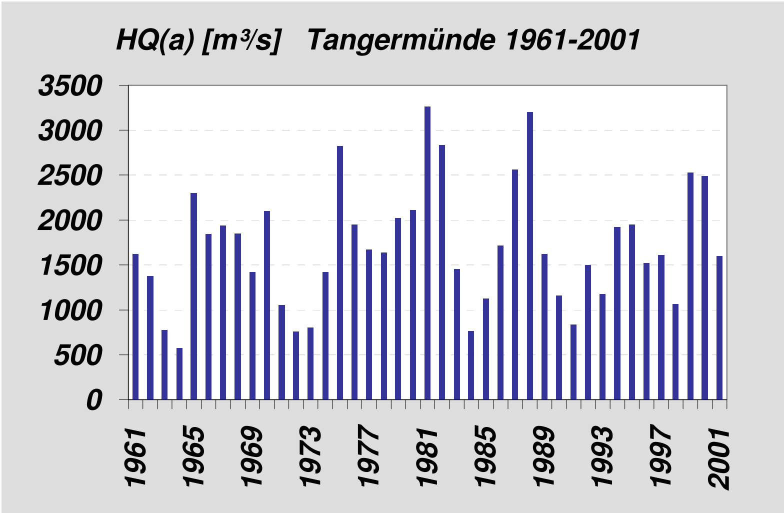
Untersuchte Variablen (meist)

- Durchfluss
- Niederschläge bestimmter Dauer und Intensität/Höhe

Ziel

- Es sollen Wahrscheinlichkeitsaussagen für beliebige Ereignisse getroffen werden (i.d.R. Extrapolation). Dazu muss die Verteilungsfunktion der Grundgesamtheit bekannt sein werden.

- Es steht aber nur eine begrenzte Zahl beobachteter Durchfluss- oder Niederschlagsereignisse zur Verfügung (=Stichprobe), denen empirische Werte für $P_Ü$, P_U bzw. T zugeordnet werden können.



→ Schätzung der Verteilungsfunktion der Grundgesamtheit auf Basis des Informationsgehalts der Stichprobe

Ziel der extremwertstatistischen Auswertung & allg. Vorgehen

Ablaufschema der Auswertung

1. Prüfung der Ausgangsdatenreihen (Konsistenz, Unabhängigkeit, Homogenität, Repräsentativität)
2. Aufstellung jährlicher oder partieller Serien (Stichprobe der beobachteten Extremwerte)
3. Berechnen empirischer Wahrscheinlichkeiten für die Werte der Stichprobe

ggf.

4. Wahl der theoretischen Verteilungsfunktion (Struktur der VF)
5. Bestimmung der Parameter der VF („Anpassung")
6. Anpassungstest (Ist die VF von ihrer Struktur her zur Beschreibung der empirischen Verteilung der Stichprobe geeignet?)
7. Berechnen von T für einen gegebenen Wert X (Frage nach der Jährlichkeit) oder Ermitteln von X für ein gegebenes T (Frage nach dem Bemessungswert).

Beispiel: Extremwerte des Durchflusses an einem Pegel

Auflaufschema der Auswertung

1. Prüfung der Ausgangsdatenreihen (Konsistenz, Unabhängigkeit, Homogenität, Repräsentativität)
2. Aufstellung jährlicher oder partieller Serien (Stichprobe der beobachteten Extremwerte)
3. Berechnen empirischer Wahrscheinlichkeiten für die Werte der Stichprobe
4. Wahl der theoretischen Verteilungsfunktion (Struktur der VF)
5. Bestimmung der Parameter der VF („Anpassung")
6. Anpassungstest (Ist die VF von ihrer Struktur her zur Beschreibung der empirischen Verteilung der Stichprobe geeignet?)
7. Berechnen von T für einen gegebenen Wert X (Frage nach der Jährlichkeit) oder Ermitteln von X für ein gegebenes T (Frage nach dem Bemessungswert).

ggf.

Beispiel: Extremwerte des Durchflusses an einem Pegel

Der Beispieldatensatz

Höchstwerte des Durchflusses (HQ(a)) am Elbpegel Tangermünde für die hydrologischen Jahre 1961-1980 (hier ersatzweise Tagesmittel statt Terminwerte)

	0	1	2	3	4	5	6	7	8	9
196x	–	1617	1374	775	572	2299	1841	1940	1850	1417
197x	2099	1051	757	800	1417	2820	1950	1670	1638	2020
198x	2111	–	–	–	–	–	–	–	–	–

- Unabhängigkeit der Werte: durch jährliche Serie gewährleistet
- Konsistenz wird unterstellt
- Prüfung von Homogenität und Repräsentativität: Zusatzinformationen oder längere Reihe notwendig

Exkurs: Gewässerkundliche Hauptwerte

Abk.	Bezeichnung	Berechnungsvorschrift
MQ	Mittlerer Durchfluss (Mittelwasserdurchfluss)	arithmetisches Mittel der Durchflüsse eines Zeitraumes
NQ	Niedrigwasserdurchfluss	niedrigster Wert des Durchflusses in einem Zeitabschnitt (Tagesmittel oder Terminwert)
HQ	Hochwasserdurchfluss	höchster Wert des Durchflusses in einem Zeitabschnitt (kein Tagesmittel, Terminwert!)
MNQ	mittlerer Niedrigwasser-durchfluss	arithmetisches Mittel der NQ-Werte mehrerer gleichartiger Zeitabschnitte (Monate, Jahre)
MHQ	mittlerer Hochwasser-durchfluss	arithmetisches Mittel der HQ-Werte mehrerer gleichartiger Zeitabschnitte (Monate, Jahre)
NNQ	niedrigster Niedrig-wasserdurchfluss	niedrigster bekannter Durchfluss (Minimum der NQ-Werte)
HHQ	höchster Hoch-wasserdurchfluss	höchster bekannter Durchfluss (Maximum der HQ-Werte)

- Welche Werte erfordern die Angabe einer Bezugsperiode bzw. eines Datums ?
- Warum darf man NQ aus Tagesmitteln berechnen, HQ aber nicht ?

Beispiel: Extremwerte des Durchflusses an einem Pegel

Auflaufschema der Auswertung

1. Test der Ausgangsdatenreihen (Konsistenz, Unabhängigkeit, Homogenität, Repräsentativität)
2. Aufstellung jährlicher oder partieller Serien (Stichprobe der beobachteten Extremwerte)
3. Berechnen empirischer Wahrscheinlichkeiten für die Werte der Stichprobe

ggf.

4. Wahl der theoretischen Verteilungsfunktion (Struktur der VF)
5. Bestimmung der Parameter der VF („Anpassung")
6. Anpassungstest (Ist die VF von ihrer Struktur her zur Beschreibung der empirischen Verteilung der Stichprobe geeignet?)
7. Berechnen von T für einen gegebenen Wert X (Frage nach der Jährlichkeit) oder Ermitteln von X für ein gegebenes T (Frage nach dem Bemessungswert).

Beispiel: Extremwerte des Durchflusses an einem Pegel

Berechnen empirischer Wahrscheinlichkeiten

Schritt 1: Ordnen der Stichprobe in aufsteigender Richtung

Schritt 2: Berechnen der empirischen P_U nach der Weibull-Formel

$$P_U(x_i) = \frac{m_i}{n+1}$$

„plotting positions“

Rang (m)	HQ-Wert	P_U	P_U [%]
1	572	0,0476	4,76
2	757	0,0952	9,52
3	775	0,1429	14,29
...	...	...	...
...	...	...	...
n	2820	0,9524	95,24

Beispiel: Extremwerte des Durchflusses an einem Pegel

Auflaufschema der Auswertung

ggf.

1. Test der Ausgangsdatenreihen (Konsistenz, Unabhängigkeit, Homogenität, Repräsentativität)
2. Aufstellung jährlicher oder partieller Serien (Stichprobe der beobachteten Extremwerte)
3. Berechnen empirischer Wahrscheinlichkeiten für die Werte der Stichprobe
4. Wahl der theoretischen Verteilungsfunktion (Struktur der VF)
5. Bestimmung der Parameter der VF („Anpassung")
6. Anpassungstest (Ist die VF von ihrer Struktur her zur Beschreibung der empirischen Verteilung der Stichprobe geeignet?)
7. Berechnen von T für einen gegebenen Wert X (Frage nach der Jährlichkeit) oder Ermitteln von X für ein gegebenes T (Frage nach dem Bemessungswert).

Beispiel: Extremwerte des Durchflusses an einem Pegel

Wahl der theoretischen Verteilungsfunktion

- Vielzahl von Verteilungsfunktionen verschiedener Gruppen
- Oft werden mehrere Verteilungsfunktionen getestet.
- Für einige Variablen haben sich bestimmte Typen von Verteilungs-funktionen besonders bewährt.

Hier vorgegeben: **Extremwertverteilung Typ 1** (Gumbel-V.)

- einfache Struktur von Verteilungs- und Dichtefunktion
- nur 2 Parameter
- analytisch lösbare Gleichung für die k-T-Beziehung

Verteilungsfunktion: $F(x) = \exp(-\exp(-a*(x-b)))$ mit $a>0$

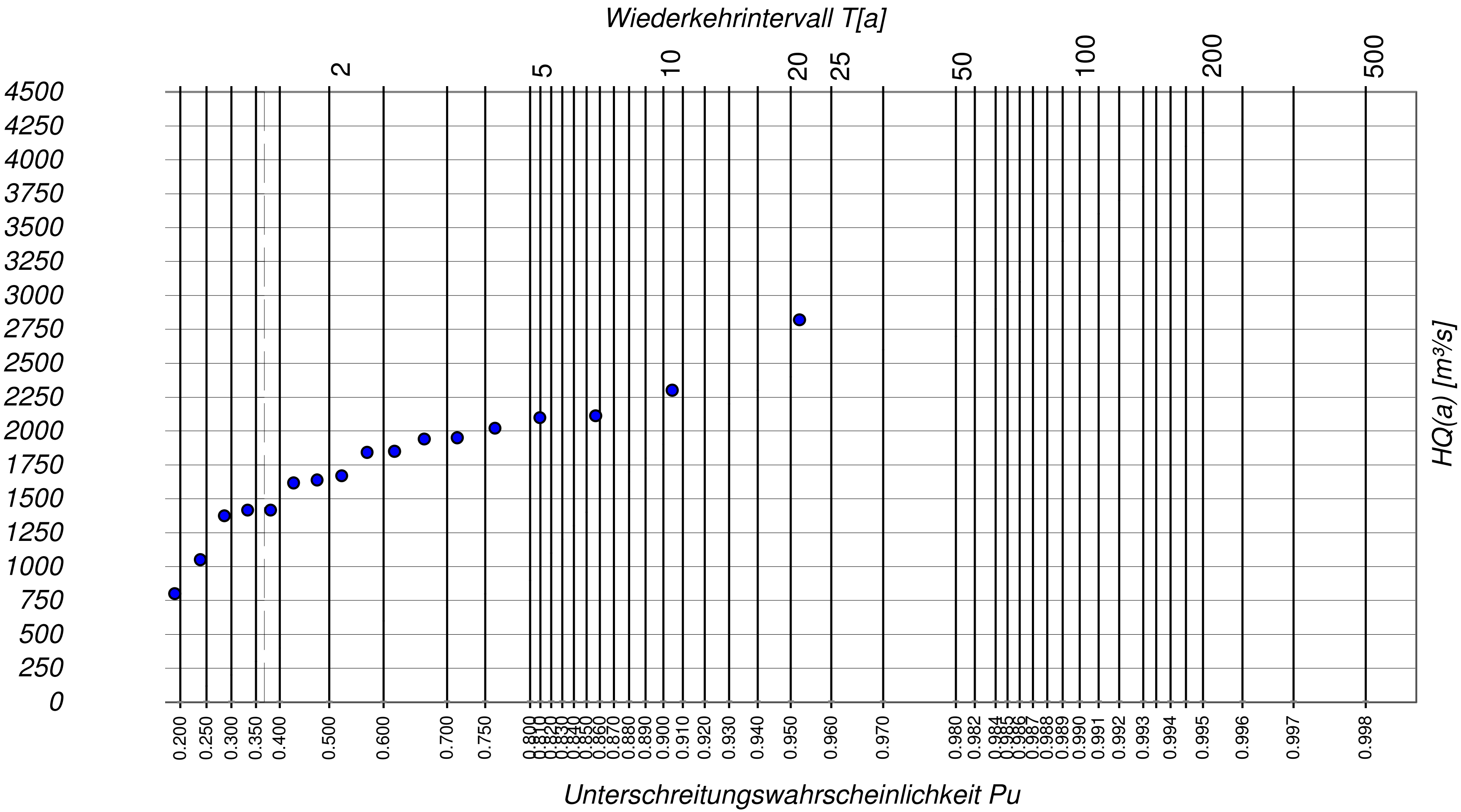
F(x): Funktionswert (Unterschreitungswahrscheinlichkeit)
x: Argument (Wert der Zufallsvariable)
a,b: Funktionsparameter

Wahrscheinlichkeitsnetze: Diagramme mit speziell skalierten Achsen, so dass die Verteilungsfunktion als Gerade dargestellt wird

Beispiel: Extremwerte des Durchflusses an einem Pegel

Wahl der theoretischen Verteilungsfunktion

Eintragen der „plotting positions“ in das Wahrscheinlichkeitsnetz



Beispiel: Extremwerte des Durchflusses an einem Pegel

Auflaufschema der Auswertung

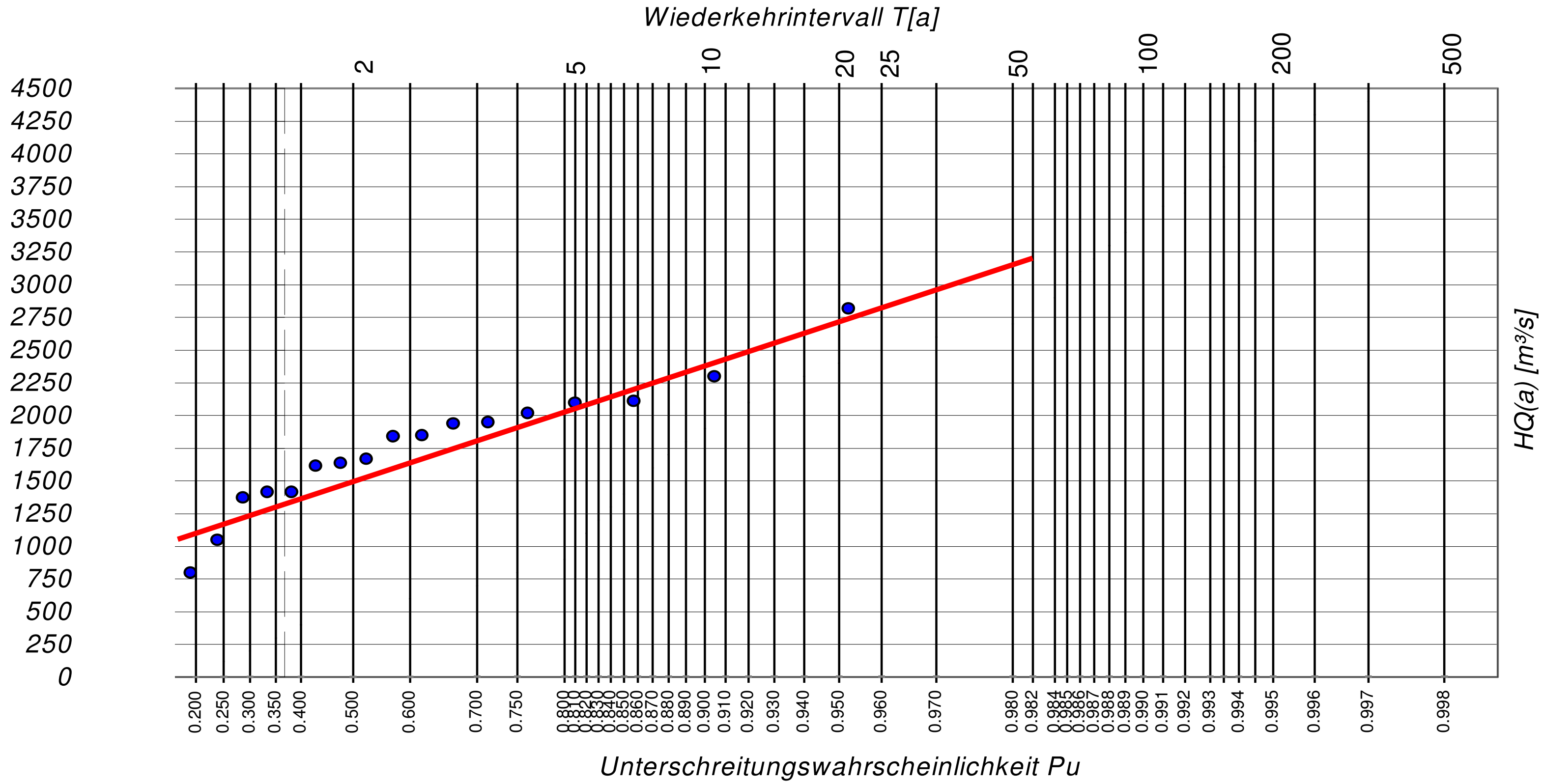
1. Test der Ausgangsdatenreihen (Konsistenz, Unabhängigkeit, Homogenität, Repräsentativität)
2. Aufstellung jährlicher oder partieller Serien (Stichprobe der beobachteten Extremwerte)
3. Berechnen empirischer Wahrscheinlichkeiten für die Werte der Stichprobe
4. Wahl der theoretischen Verteilungsfunktion (Struktur der VF)
5. Bestimmung der Parameter der VF („Anpassung")
6. Anpassungstest (Ist die VF von ihrer Struktur her zur Beschreibung der empirischen Verteilung der Stichprobe geeignet?)
7. Berechnen von T für einen gegebenen Wert X (Frage nach der Jährlichkeit) oder Ermitteln von X für ein gegebenes T (Frage nach dem Bemessungswert).

ggf.

Beispiel: Extremwerte des Durchflusses an einem Pegel

Bestimmung der Parameter der Verteilungsfunktion

Methode A: Freihand-Anpassung



- Vorteil: schnell, keine Rechnung erforderlich
- Nachteil: subjektiv, lediglich Abschätzung, geringe Extrapolation!

Beispiel: Extremwerte des Durchflusses an einem Pegel

Bestimmung der Parameter der Verteilungsfunktion

Methode B: Momenten-Methode

- Parameter der VF stehen mit statistischen Kenngrößen (Momenten) der Grundgesamtheit in Zusammenhang
- Idee: Verwendung der Momente der Stichprobe als Schätzung
  (d.h. es wird angenommen, dass die aus der Stichprobe berechneten statistischen Kenngrößen auch für die Grundgesamtheit gültig sind – „repräsentative Stichprobe")

Im Fall der Extremwertverteilung Typ 1:

$$F(x) = \exp(-\exp(-a*(x-b))) \qquad a = \frac{\pi}{\sqrt{6} * \sigma} \qquad b = \mu - \frac{0.5772}{a}$$

μ: Mittelwert (erstes Moment), σ: Standardabweichung (σ^2 : zweites Zentralmoment)

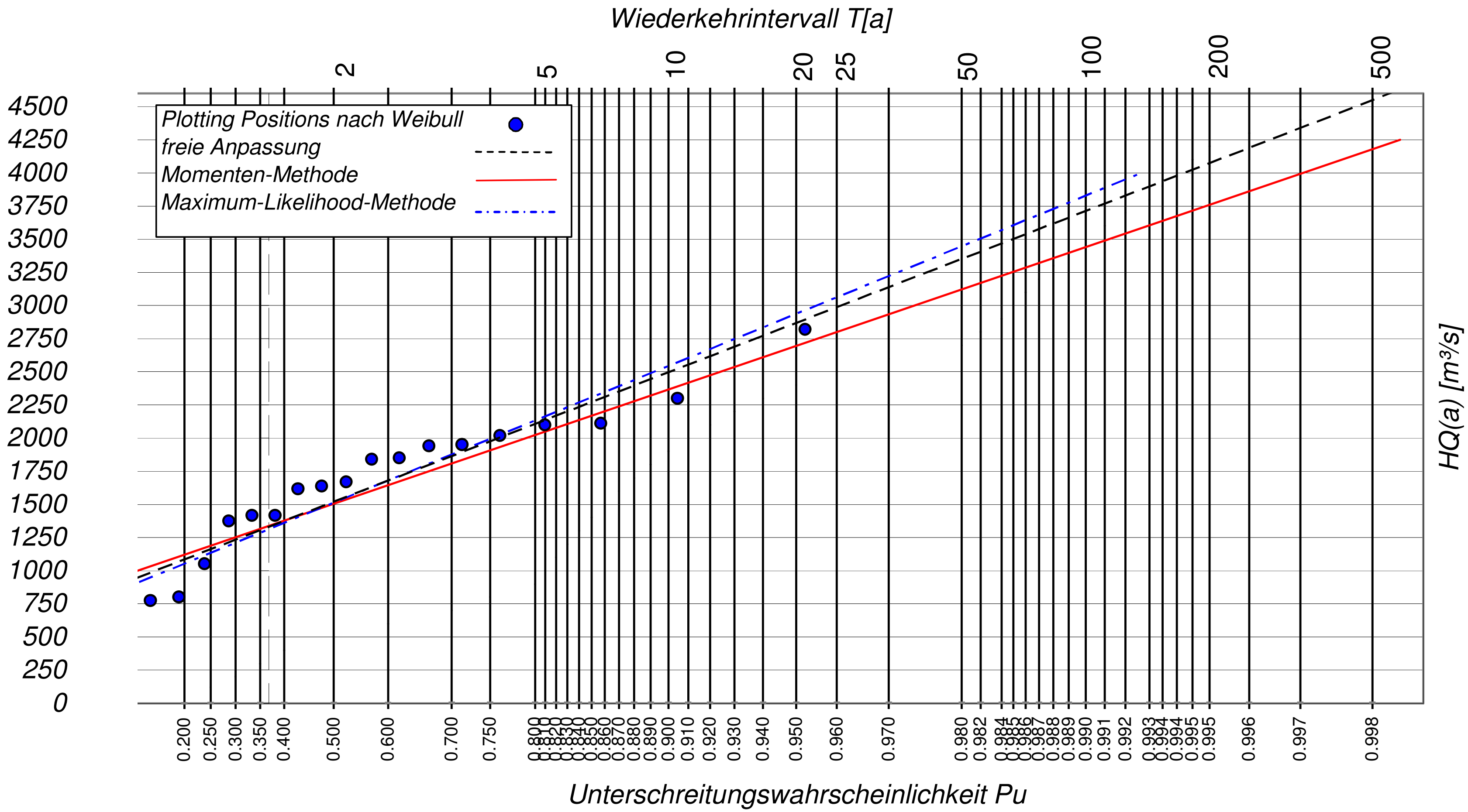
- Gerade im Netzdruck: ergibt sich aus 2 Wertepaaren (x,F(x))

weitere Methoden: Wahrscheinlichkeitsgewichtete Momenten-M.,
Maximum-Likelihood-Methode

Beispiel: Extremwerte des Durchflusses an einem Pegel

Bestimmung der Parameter der Verteilungsfunktion

Ergebnis der Anpassung mit 3 verschiedenen Methoden



Beispiel: Extremwerte des Durchflusses an einem Pegel

Auflaufschema der Auswertung

1. Test der Ausgangsdatenreihen (Konsistenz, Unabhängigkeit, Homogenität, Repräsentativität)
2. Aufstellung jährlicher oder partieller Serien (Stichprobe der beobachteten Extremwerte)
3. Berechnen empirischer Wahrscheinlichkeiten für die Werte der Stichprobe
4. Wahl der theoretischen Verteilungsfunktion (Struktur der VF)
5. Bestimmung der Parameter der VF („Anpassung")
6. Anpassungstest (Ist die VF von ihrer Struktur her zur Beschreibung der empirischen Verteilung der Stichprobe geeignet?)
7. Berechnen von T für einen gegebenen Wert X (Frage nach der Jährlichkeit) oder Ermitteln von X für ein gegebenes T (Frage nach dem Bemessungswert).

ggf.

Beispiel: Extremwerte des Durchflusses an einem Pegel

Anpassungstests

z.B. Kolmogorov-Smirnov-Test:

Es wird der Betrag der max. Abweichung zwischen der empirischen Wahrscheinlichkeit nach Weibull und dem zugehörigen Funktionswert der angepassten VF aus der Stichprobe bestimmt.

$$D = \max \mid emp.\ P_U(x) - F(x) \mid$$

Die Nullhypothese (angepasste VF repräsentiert die Stichprobe) wird abgelehnt, wenn die Abweichung D einen kritischen Betrag $\Delta(\alpha,n)$ überschreitet. Δ ist von der gewählten Irrtumswahrscheinlichkeit α und dem Stichprobenumfang n abhängig.

Aussagekraft des K-S-Tests im Bereich kleiner $P_Ü$ gering! Deshalb:

$\rightarrow$ weitere Testverfahren

$\rightarrow$ visueller Vergleich unterschiedlicher angepasster VF

Beispiel: Extremwerte des Durchflusses an einem Pegel

Auflaufschema der Auswertung

1. Test der Ausgangsdatenreihen (Konsistenz, Unabhängigkeit, Homogenität, Repräsentativität)
2. Aufstellung jährlicher oder partieller Serien (Stichprobe der beobachteten Extremwerte)
3. Berechnen empirischer Wahrscheinlichkeiten für die Werte der Stichprobe

ggf.

4. Wahl der theoretischen Verteilungsfunktion (Struktur der VF)
5. Bestimmung der Parameter der VF („Anpassung")
6. Anpassungstest (Ist die VF von ihrer Struktur her zur Beschreibung der empirischen Verteilung der Stichprobe geeignet?)
7. Berechnen von T für einen gegebenen Wert X (Frage nach der Jährlichkeit) oder Ermitteln von X für ein gegebenes T (Frage nach dem Bemessungswert).

Beispiel: Extremwerte des Durchflusses an einem Pegel

Berechnung von Jährlichkeiten und Bemessungswerten

Zulässiger Extrapolationsbereich:

- maximal bis zur 3-fachen Länge der Datenreihe, wenn keine Zusatzinformation einbezogen wird (z.B. historische HWs)
- für Ermittlung von HQ_{500} bräuchte man 170-jährige Reihe
- für einige Bemessungsaufgaben sind Werte >= HQ_{1000} nötig

Aus dem Wahrscheinlichkeitsnetz (grafisch) zu ermitteln:

- Jährlichkeit für einen Durchfluss Q> 3200 m³/s
- 50-jähriger Hochwasserdurchfluss (HQ_{50})
- Mit welcher Wahrscheinlichkeit liegt HQ(a) zw. 2000 & 3000 m³/s
- W.-keit, dass in einem Zeitraum von 5 Jahren Q> 2500 m³/s auftritt

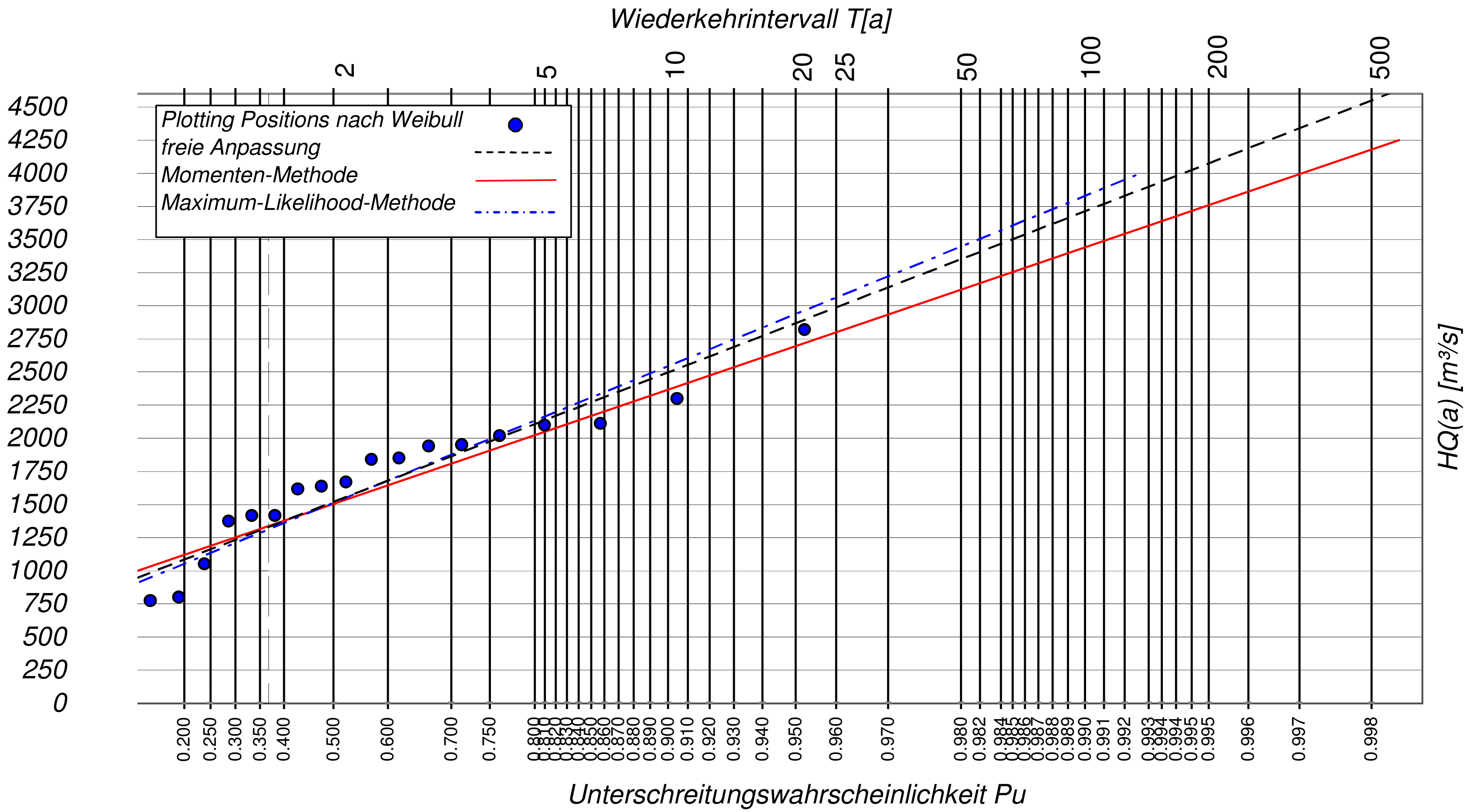
Mittels der VF zu berechnen (unerlaubte Extrapolationen):

- Jährlichkeit für Q> 3500 m³/s
- HQ_{100}

bei komplizierten VF bequemer: Anwendung der k-T-Beziehung

Beispiel: Extremwerte des Durchflusses an einem Pegel

Berechnung von Jährlichkeiten und Bemessungswerten



Vergleich: HQ_{50} in Abhängigkeit von gewählter Anpassungs-Methode

Extremwerte des Niederschlags (nur Übung 2004)

Quellen für Angaben zu Starkniederschlägen

- Niederschlags-(Höhen/Intensitäts)-Dauer-Häufigkeits-Beziehungen für konkrete Klima-/Niederschlagsstationen
- KOSTRA-Atlas (Übersichtswerte)
- früher oft verwendet: empirische Reinhold-Formel

Anwendungsbeispiel zur Reinhold-Formel

$$r(T,D) = r(1{,}15) * \frac{38 * (T^{0.25} - 0.369)}{D+9}$$

r(T,D): Regenspende eines Niederschlags der Dauer D [min] mit dem Wiederkehrintervall T [a] in Liter/Sekunde/Hektar

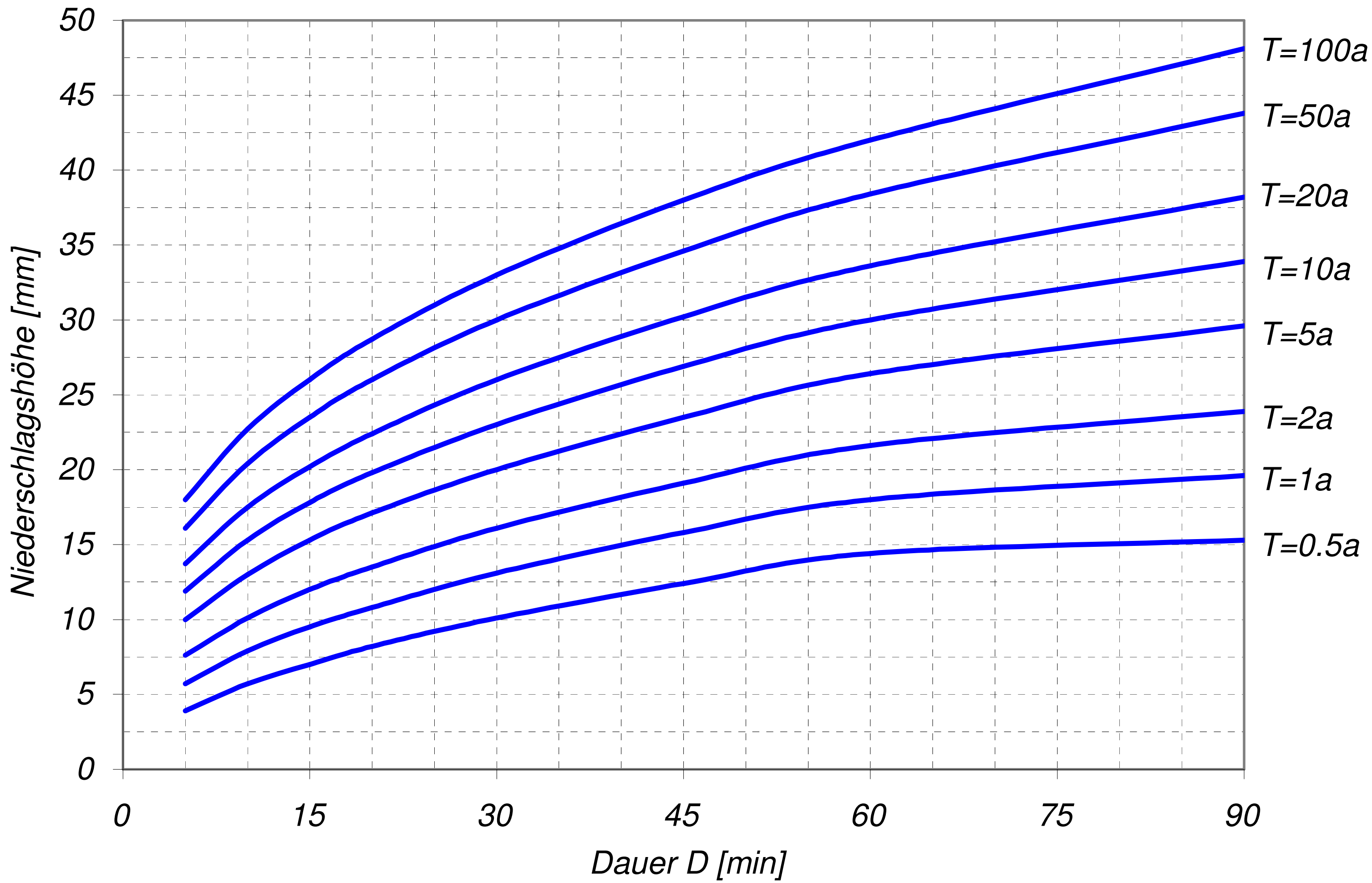
r(1,15): Basisregenspende (D=15 min, T=1 a)

Gesucht: Höhe des 10-jährigen 30-Minuten-Niederschlags in [mm]

Gegeben: Basisregenspende r(1,15)= 115 $l\ s^{-1}\ ha^{-1}$

Extremwerte des Niederschlags (nur Übung 2004)

Höhen-Dauer-Häufigkeits-Diagramm einer Klimastation



Gesucht: Höhe des 10-jährigen 30-Minuten-Niederschlags in [mm]


2002 S. Itzerott (GFZ) und M. Zebisch (TUB/PIK)