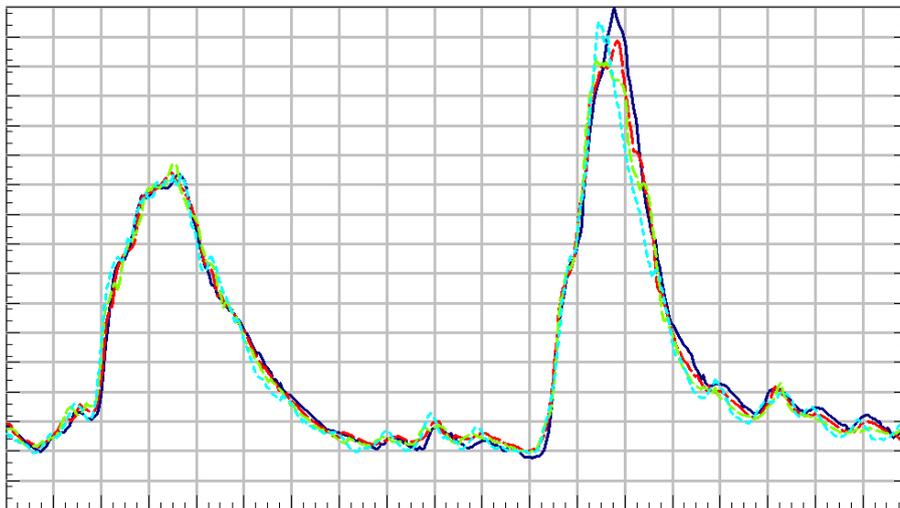


# ProFoUnD

## Programmbeschreibung und Anwendungshinweise



Stand: 15.12.2016 (Version 8.1)

Ingo Haag, Dirk Aigner

Herausgeber:

LARSIM-Entwicklergemeinschaft – Hochwasserzentralen LUBW, BIfU, LfU RP, HLNUG, BAFU

# ProFoUnD

## Programmbeschreibung und Anwendungshinweise

### Inhalt

Abbildungsverzeichnis	i
1 Aufgaben und prinzipielle Funktionsweise von ProFoUnD	1
1.1 Grundlegende Funktionsweise	1
1.2 Wesentliche Neuerungen ab Version 1	4
2 Grundlagen und Berechnungsalgorithmen des Programms	7
2.1 Ausgabe von Ganglinien	7
2.2 Klasseneinteilung und hydrologische Fälle	7
2.2.1 Hydrologische Fälle nach dem Verfahren des Bayerischen LfU	8
2.2.2 Hydrologische Fälle nach dem alten IKSMS-Verfahren (Stand 2009)	9
2.2.3 Hydrologische Fälle nach dem neuen IKSMS-Verfahren (Stand 2012)	11
2.3 Kontinuierliche Fehlermaße	12
2.3.1 Kontinuierliche Einzelfehler	12
2.3.2 Tolerierbares Zeitfenster bei kontinuierlichen Fehlern	13
2.3.3 Mittlere kontinuierliche Fehler	14
2.4 Häufigkeitsverteilungen der kontinuierlichen Fehler	17
2.4.1 Empirische Summenhäufigkeit und empirische Percentile	17
2.4.2 Momente der Häufigkeitsverteilungen	17
2.4.3 Polynome zur Beschreibung der Momente	19
2.4.4 Normalverteilung und theoretische Percentile	21
2.4.5 Anpassungstests	22
2.5 Kategorische Fehler	25
2.5.1 Standarddefinition der Kontingenztafel	25
2.5.2 Strenge Definition der Kontingenztafel	26
2.5.3 Kontingenztafel bei Berücksichtigung eines tolerierbaren Zeitfensters	27
2.5.4 Kategorische Gütemaße	29
2.6 Auswertung maximaler Werte für Vorhersagebereiche	32
2.7 Auswertung minimaler Werte für Vorhersagebereiche	33
2.8 Automatisierte Fortschreibung bestehender Auswertungen	33
2.9 Auswertung der Scheitelwerte	33

3	Hinweise zur Anwendung des Programms	34
3.1	Grundlegende Funktionsweise und Programmsteuerung	34
3.1.1	Hinweise zur Verwendung des LLA-Formats	38
3.1.2	Besonderheiten beim Fortschreiben bestehender Auswertungen	42
3.1.3	Übergabe des Ereigniszeitraums beim Programmaufruf	43
3.2	Eingangsdaten	43
3.3	Ausgaben	46
4	Literaturverzeichnis	54

## Abbildungsverzeichnis

Bild 1.1: Grundlegende Funktionsweise des Programms ProFoUnD.....	2
Bild 2.1: Flussdiagramm zur Einordnung der Wertepaare in die Kontingenztafel bei strenger HIT-Definition (ohne tolerierbares Zeitfenster; Ereignis-Definition = Überschreitung). ....	27
Bild 2.2: Flussdiagramm zur Einordnung der Wertepaare in die Kontingenztafel gemäß der Standarddefinition bei Vorgabe eines tolerierbaren Zeitfensters (Ereignis-Definition = Überschreitung).....	28
Bild 2.3: Flussdiagramm zur Einordnung der Wertepaare in die Kontingenztafel gemäß der strengen Definition bei Vorgabe eines tolerierbaren Zeitfensters (Ereignis-Definition = Überschreitung).....	29
Bild 3.1: Beispielhafter Ausschnitt aus der Datei <protokoll-vorhandene-vhs.csv>.....	46
Bild 3.2: Beispielhafter Ausschnitt aus der Datei <protokoll-fehlende-wertepaare.csv>.....	47
Bild 3.3: Beispielhafte Darstellung einer gemessenen Abflussganglinie mit zusammengesetzten Vorhersagen unterschiedlicher Vorhersagetiefe. ....	47
Bild 3.4: Beispielhafter Ausschnitt aus einer Datei mit Rangfolgen (ohne tolerierbares Zeitfenster). ....	48
Bild 3.5: Beispielhafter Ausschnitt aus einer Datei mit mittleren Fehlern. ....	49
Bild 3.6: Beispielhafter Ausschnitt aus einer Datei mit Momenten.....	49
Bild 3.7: Beispielhafter Ausschnitt aus einer Datei mit Polynomen. ....	50
Bild 3.8: Beispielhafter Ausschnitt aus einer Datei mit Percentilen. ....	51
Bild 3.9: Beispielhafter Ausschnitt aus einer Datei mit Ergebnissen des Anpassungstests. ....	52
Bild 3.10: Beispielhafter Ausschnitt aus einer Datei mit Kontingenztafeln und kategorischen Gütemaßen.....	52
Bild 3.11: Beispiel einer Datei mit Scheitelwerten (Wasserstände).....	53

# 1 Aufgaben und prinzipielle Funktionsweise von ProFoUnD

Das Programm ProFoUnD (program to assess the forecast uncertainty of discharge) dient der quantitativen Auswertung von Abfluss- und Wasserstandvorhersagen des Modells LARSIM. Die Ergebnisse können unter anderem zur quantitativen Beschreibung der Unsicherheiten von Vorhersagefehlern im operationellen Fall verwendet werden.

ProFoUnD ist eine grundlegende Weiterentwicklung des Programms GMD\_VHS\_GUETE, das im Auftrag und in Abstimmung mit der bayerischen Wasserwirtschaftsverwaltung (HND des Bayer. LfU) entwickelt worden war. Die grundlegenden Erweiterungen für ProFoUnD wurden im Auftrag und in Abstimmung mit dem LUWG Rheinland-Pfalz (Herr Demuth) durchgeführt, wobei die Anforderungen einer IKSMS-Arbeitsgruppe umgesetzt wurden (Versionen 2 bis 4). Die mit Version 5 von ProFoUnD implementierten Zusatzfunktionalitäten erfolgten wiederum im Auftrag und in Abstimmung mit dem HND des Bayer. LfU (Frau Moritz, Dr. Vogelbacher). Die Weiterentwicklungen in Version 6 zielten vor allem auf die automatisierte operationelle Nutzung des Programms und die damit einhergehende Fortschreibung der Ergebnisse auf der Grundlage früherer Auswertungen ab. Version 6 wurde in Abstimmung und im Auftrag des HUG (Dr. Brahmer) und des LUWG (Herr Demuth) durchgeführt. Mit Version 7 wurde neben der Möglichkeit zur Auswertung von Minimalwerten zur Niedrigwasser-Frühwarnung als weitere Funktionalität die Auswertung der Abweichungen der Scheitel der Vorhersagen vom gemessenen Scheitel eingeführt. Ab Version 8 können Daten im LILA-Format verarbeitet werden.

Zunächst wird in Abschnitt 1.1 die grundlegende Funktionsweise von ProFoUnD erläutert. In Abschnitt 1.2 werden die wesentlichen Neuerungen seit Version 1 kurz aufgeführt. Ausführliche Erläuterungen der Berechnungsverfahren folgen in den Kapitel 2. Kapitel 3 gewährt abschließend Hinweise zum Umgang mit dem Programm.

## 1.1 Grundlegende Funktionsweise

Die grundlegende Funktionsweise des Programms ist in Bild 1.1 schematisch dargestellt. Ausgehend von gemessenen und vorhergesagten Abfluss- oder Wasserstandganglinien vorzugebender Pegel liefert das Programm folgende Ergebnisse (vgl. Bild 1.1):

- Aus den Vorhersagen werden Ganglinien für ausgewählte Vorhersagetiefen extrahiert. Diese können getrennt für die angeforderten Pegel und Ereigniszeiträume zusammen mit der zugehörigen gemessenen Ganglinie in jeweils eine GMD- und/oder HMZ-Datei oder in eine LILA-Datei ausgegeben. Hierdurch ist ein direkter visueller Vergleich der gemessenen Ganglinie mit zusammengesetzten Ganglinien unterschiedlicher Vorhersagetiefen möglich.

Zudem können die zeitlichen Verläufe kontinuierlicher Einzelfehler der unterschiedlichen Vorhersagetiefen als zusätzliche Ganglinien ausgegeben werden. Hierdurch wird direkt deutlich, in welchen Zeiträumen besonders große oder kleine Fehler auftraten.

In den HMZ-Listen und LILA-Dateien kann zusätzlich der hydrologische Fall ausgegeben werden, der dem jeweiligen Vorhersagewert zugeordnet wurde. Hierdurch kann der Anwender einzelne Vorhersagen (z.B. extrem schlechte Vorhersagen) leichter identifizieren und zuordnen.

- Die Vorhersagewerte können mithilfe von drei unterschiedlichen Verfahren (gemäß den Vorgaben des Bayerischen LfU oder der IKSMS [alt und neu]) in unterschiedliche hydrologische Fälle gruppiert werden. Fehlerverteilungen und kontinuierliche Gütemaße der Vorhersagen werden sowohl gemeinsam als auch getrennt nach diesen unterschiedlichen Fällen berechnet.

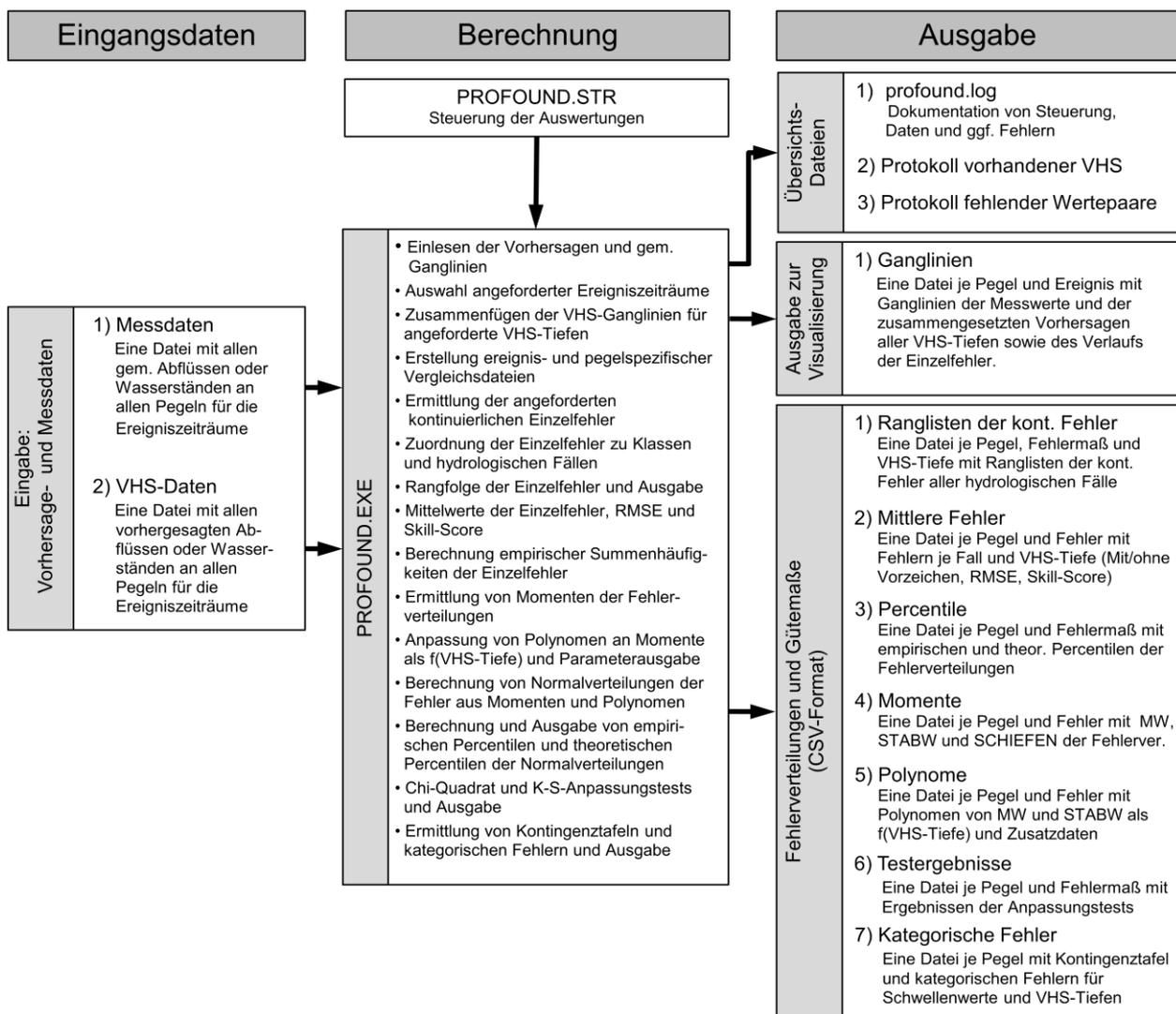


Bild 1.1: Grundlegende Funktionsweise des Programms ProFoUnD

- Im Kontext kontinuierlicher Fehler können die einzelnen Abweichungen zwischen gemessenen und vorhergesagten Abflüssen (getrennt nach hydrologischen Fällen und Vorhersagetiefen) mithilfe folgender Einzelfehler quantifiziert werden: Absolute Abweichung, prozentuale Abweichung, Quotienten aus Messung und Vorhersage, natürlicher Logarithmus des Quotienten, quadrierte Abweichung und Zeitfehler. Die jeweils zu verwendenden Fehlermaße können über die Steuerdatei ausgewählt werden.
- Die Ermittlung der angeführten Einzelfehler (mit Ausnahme des Zeitfehlers) kann dabei wahlweise mit oder ohne Verwendung eines tolerierbaren Zeitfensters erfolgen. Dabei kann für jede Vorhersagetiefe ein anderes tolerierbares Zeitfenster vorgegeben werden.
- Die genannten Einzelfehler werden (getrennt nach hydrologischen Fällen und Vorhersagetiefen) in aufsteigender Größe gemeinsam mit den zugehörigen Rängen als CSV-Dateien ausgegeben. Mit diesen Daten können empirische Summenhäufigkeiten grafisch dargestellt werden. Sofern ein tolerierbares Zeitfenster berücksichtigt wurde, wird auch der zur Ermittlung des jeweiligen Einzelfehlers verwendete Zeitverschub mit ausgegeben.
- Aus den Einzelfehlern werden (getrennt nach hydrologischen Fällen und Vorhersagetiefen) die jeweiligen mittleren Fehler berechnet und in CSV-Dateien ausgegeben. Die Mittelwerte werden dabei (soweit sinnvoll) sowohl unter Berücksichtigung der Vorzeichen als auch unter Verwendung der Beträge ermittelt.

- Bei Auswahl der quadrierten Abweichung werden zusätzlich zu den Mittelwerten auch der RMSE und ein Skill-Score in Bezug auf die Annahme von Persistenzvorhersagen berechnet und getrennt nach hydrologischen Fällen und Vorhersagetiefen) ausgegeben.
- Basierend auf der Annahme, dass die kontinuierlichen Fehler normalverteilt sind, werden die entsprechenden Normalverteilungen mit den für die jeweilige Vorhersagetiefe ermittelten Mittelwerten und Standardabweichungen berechnet.
- Falls über die Steuerdatei angefordert, werden Polynome 2. Grades an die Verläufe der Mittelwerte und Standardabweichungen über die Vorhersagetiefe angepasst. Die Parameterwerte der Polynome werden gemeinsam mit zusätzlichen Informationen in CSV-Dateien ausgegeben.
- Sofern die Polynome angefordert sind, werden die Normalverteilungen der Fehler zusätzlich mit den Mittelwerten und Standardabweichungen berechnet, die sich aus dem jeweiligen Polynom ergeben.
- Getrennt nach hydrologischen Fällen und Vorhersagetiefen werden die empirischen Percentile der Fehler, die Percentile der Normalverteilungen mit den für die jeweilige Vorhersagetiefe ermittelten Momenten sowie ggf. die Percentile der Normalverteilungen, die aus den Polynomen resultieren, berechnet und tabellarisch in CSV-Dateien ausgegeben. Anhand dieser Percentiltabellen können die Fehlerverteilungen näher analysiert oder (z.B. mittels Box-und-Whisker-Plots) visualisiert werden.
- Zudem werden Anpassungstests ( $\chi^2$ -Test und Kolmogoroff-Smirnoff-Test) durchgeführt, um zu quantifizieren, wie gut die empirischen Fehlerverteilungen durch die Normalverteilungen aus den Momenten sowie ggf. aus den Polynomen beschrieben werden. Die Ergebnisse der Tests werden ebenfalls als CSV-Dateien ausgegeben.
- Über die oben angeführten kontinuierlichen Gütemaße hinaus kann über die Steuerdatei auch die Berechnung und Ausgabe kategorischer Gütemaße angefordert werden. Für die kategorischen Maße wird untersucht, ob das Eintreten bzw. das Nichteintreten eines Ereignisses korrekt vorhergesagt wurde. Der Eintritt eines Ereignisses ist dabei durch das Überschreiten (Hochwasser) oder das Unterschreiten (Niedrigwasser) eines Schwellenwerts definiert. Dabei ist es möglich, zwischen zwei unterschiedlich strengen Definitionen für den Eintritt eines Ereignisses zu unterscheiden. Mit ProFoUnD werden neben den Kontingenztafeln fünf spezifische kategorische Gütemaße für die pegelspezifisch vorgegebenden Schwellenwerte in CSV-Dateien ausgegeben.
- Neben der bislang beschriebenen Standardauswertung kann ProFoUnD auch zur Bewertung der Hochwasser- oder Niedrigwasser-Frühwarnung verwendet werden. In diesem Fall werden anstelle der einzelnen Vorhersagewerte die maximal bzw. minimal vorhergesagten Werte für einen bestimmten Zeitbereich ausgewertet und mit den zugehörigen Maximal- bzw. Minimalwerten der Messungen verglichen. Hierfür können auch mehrere Zeitbereiche vorgegeben werden (z.B. 1h-24h VHS und 25h-48h VHS). Für diese Auswertungsmethode können weitgehend dieselben Gütemaße ermittelt werden wie für die Standardauswertung. Die Ermittlung eines Zeitfehlers oder die Berücksichtigung eines tolerierbaren Zeitfensters sind hierbei allerdings nicht möglich.
- Die Auswertungen können (sofern gewünscht) sukzessive auf der Basis neu verfügbarer Daten (z.B. neue Vorhersagen) fortgeschrieben werden. Hierzu werden die bisherigen Ergebnisse sowie die neuen Daten eingelesen, programmintern zusammengeführt und dann als neuer Gesamtdatensatz ausgewertet. Dies ermöglicht es, beispielsweise im Rahmen des operationellen Betriebs die Ergebnisse automatisiert kontinuierlich fortzuschreiben und die Statistiken somit sukzessive auf eine breitere Basis zu stellen.

Eine ausführlichere Beschreibung der verwendeten Methoden und Berechnungsansätze ist Kapitel 2 zu entnehmen. Kapitel 3 beinhaltet Hinweise zur Anwendung des Programms.

## 1.2 Wesentliche Neuerungen ab Version 1

Im Folgenden sind die wesentlichen Neuerungen von ProFoUnD nach Version 1 (Stand 12.01.2009, damaliger Programmname „GMD\_VHS\_GUETE“) bis zur aktuellen Version 8 aufgeführt:

- Anwendung auf Abflüsse und Wasserstände. Dabei ist zu beachten, dass Abflüsse und Wasserstände in getrennten Programmdurchläufen analysiert werden müssen. Hierfür wurde in der Steuerdatei im Block <PEGEL> ein entsprechender Schalter für die <Datenart> vorgesehen.
- Einlesen der Vorhersagewerte aus GMD-Dateien oder aus einem (CSV-basierten) Schnittstellenformat mit der Datenbank des HND (ab Version 5). Die Auswahl des Formats für die Vorhersagedaten erfolgt über den Schalter <Eingabeformat VHS:> im Block <PFADE> der Steuerdatei.
- Einlesen von Messwerten im Format der Wasser- und Schifffahrtsverwaltung (WSV) sowie der zugehörigen WAVOS-Vorhersagen für Wasserstände in einem spezifischen Listenformat („WSV-Format“) (ab Version 6). Die Auswahl des WSV-Formats für die Messwerte erfolgt über den Schalter <Eingabeformat GEM:> im Block <PFADE> der Steuerdatei mit dem Schlüsselwort <WSV>. Die Auswahl des WSV-Formats für die Vorhersagen erfolgt entsprechend über den Schalter <Eingabeformat VHS:>.
- Ausgabe der ereignisspezifischen vergleichenden Ganglinien von Messungen und zusammengesetzten Vorhersagen gleicher Vorhersagetiefe im GMD und/oder HMZ-Format sowie ab Version 8 auch im LILA-Format. Hierfür wurden im Block <GUETEMASSE> der Steuerdatei zwei (bzw. drei) zusätzliche Schalter vorgesehen, um die Ausgaben anzufordern.
- Mögliche Ausgabe des zeitlichen Verlaufs kontinuierlicher Einzelfehler („Fehlerganglinien“). Die Fehlerganglinien können ergänzend in die Dateien mit den ereignisspezifischen vergleichenden Ganglinien ausgegeben werden. Dabei werden alle angeforderten kontinuierlichen Einzelfehler getrennt nach Vorhersagetiefen ausgegeben. Um die Ausgabe der Fehlerganglinien auswählbar zu machen, wurde im Block <GUETEMASSE> der Steuerdatei ein zusätzlicher Schalter vorgesehen (<Verlauf Einzelfehler>).
- Mögliche Ausgabe des zugehörigen hydrologischen Falls in der HMZ-Liste (ab Version 5) oder im LILA-Format (ab Version 8) der vergleichenden Ganglinie. Für die Ausgabe des hydrologischen Falls wurde im Block <GUETEMASSE> der Steuerdatei ein zusätzlicher Schalter vorgesehen (<Verlauf hyd. Fall:>). Eine Ausgabe der hydrologischen Fälle in die GMD-Dateien ist nicht möglich.
- Auswahl des Verfahrens zur Ermittlung der hydrologischen Fälle für die kontinuierlichen Fehler. Neben dem bislang angewandten Verfahren können die hydrologischen Fälle auch nach den Kriterien der IKSMS ermittelt werden. Hierfür wurde im Block <HYDROLOGISCHE FAELLE> der Steuerdatei ein Schalter zur Auswahl der <Methode Schwellenwerte> eingeführt. Durch <blfu> wird das zuerst implementierte Verfahren des Bayer. LfU gewählt. Durch die Angabe <iksms> bzw. <iksms-alt> können zwei unterschiedliche Verfahren gemäß den Vorgaben der IKSMS ausgewählt werden. Durch die Wahl des Verfahrens ergeben sich auch unterschiedliche Anforderungen für die Vorgabe der Schwellenwerte in der Steuerdatei (Details siehe Kapitel 3).

- Die Möglichkeit tolerierbare Zeitfenster bei der Ermittlung der Fehlermaße zu berücksichtigen. Hierzu wurde ein Schalter <Zeitfenster tolerierbar> im Block <VORHERSAGEDAUERN> der Steuerdatei geschaffen. Sofern der Schalter auf <ja> gesetzt wird, kann für jede der zu untersuchenden Vorhersagetiefen ein tolerierbares Zeitfenster bei der Ermittlung der Fehlermaße angegeben werden.
- Als zusätzliche kontinuierliche Fehler können der Zeitfehler (optimaler Zeitverschub der Vorhersage) und die quadrierte Abweichung zwischen Vorhersage und Messung ermittelt werden. Dabei ist zu beachten, dass bei der Ermittlung des Zeitfehlers kein tolerierbares Zeitfenster verwendet wird. Für die Fehlermaße wurden zusätzliche Schalter im Block <GUETEMASSE> der Steuerdatei vorgesehen.
- Aus den kontinuierlichen Einzelfehlern werden getrennt nach hydrologischen Fällen und Vorhersagetiefen unterschiedliche mittlere Fehler berechnet und ausgegeben. Um sowohl Aussagen über zufällige als auch über systematische Fehler machen zu können, werden für die absolute Abweichung und den Zeitfehler jeweils vorzeichenbehaftete Mittelwerte und Mittelwerte der Beträge ermittelt und ausgegeben. Wird die quadrierte Abweichung angefordert, so werden neben den Mittelwerten auch der RMSE sowie ein Persistenz-Skill-Score (Skill-Score im Vergleich zur Persistenzannahme) ausgegeben.
- Neben den kontinuierlichen Fehlern können auch kategorische Fehler ermittelt werden. Hierzu wurde ein zusätzlicher Schalter im Block <GUETEMASSE> der Steuerdatei vorgesehen. Die kategorischen Gütemaße können nur in Verbindung mit den Kriterien der IKSMS zur Ermittlung der hydrologischen Fälle berechnet werden. Dabei können auch die oben genannten tolerierbaren Zeitfenster berücksichtigt werden. Ab Version 5 können über einen zusätzlichen Schalter im Block <GUETEMASSE> der Steuerdatei zwei unterschiedlich strenge Definitionen für den Eintritt eines Ereignisses verwendet werden. Für jeden der vorgegebenen Schwellenwerte und jede untersuchte Vorhersagetiefe wird eine Kontingenztafel ermittelt. Auf dieser Grundlage werden die kategorischen Gütemaße Probability of Detection (Hit rate), False Alarm Rate (Probability of False Detection), False Alarm Ratio, Threat Score (Critical Success Index) und Frequency Bias (Bias) ermittelt und ausgegeben.
- Ab Version 7 ist es möglich, die kategorischen Gütemaße nicht nur für das Überschreiten eines Schwellenwerts (Hochwasserauswertung) sondern auch für das Unterschreiten eines Schwellenwerts (Niedrigwasserauswertung) zu ermitteln. Hierfür wurde der Schalter <Ereignis-Definition> in der Steuerdatei eingeführt.
- ProFoUnD wurde programmtechnisch angepasst zur Auswertung umfangreicher Datensätze. Bei sehr großen Datenmengen wird die Verwendung der 64-bit-Version von ProFoUnD empfohlen (ab Version 6).
- Neben der Standardauswertung der einzelnen Vorhersagewerte können ab Version 4 auch die Maximalwerte der Vorhersagen für bestimmte Vorhersagebereiche (z.B. 1h-24h VHS, 25h-48h VHS etc.) mit den zugehörigen Maximalwerten der Messungen verglichen werden. Diese Methode ist insbesondere für die Bewertung der Hochwasser-Frühwarnung geeignet. Zur Auswahl der Auswertungsmethode wurde ein entsprechender Schalter im Block <PEGEL> der Steuerdatei eingefügt.
- Ab Version 7 können analog zu den Maximalwerten auch die Minimalwerte der Vorhersagen für bestimmte Vorhersagebereiche mit den zugehörigen Minimalwerten der Messung verglichen werden.

Dies ist insbesondere für die Bewertung der Niedrigwasser-Frühwarnung geeignet. Die Auswahl dieser Auswertungsmethode erfolgt über den gleichen Schalter in der Steuerdatei, wie für die Maximalwerte.

- Bei Auswertung der Minimalwerte zur Niedrigwasser-Frühwarnung sollte für die Ermittlung der kategorischen Gütemaße (siehe Kapitel 2.7) ein Ereignis als solches durch das Unterschreiten einer Klassengrenze bzw. eines Schwellenwerts definiert werden. Dazu wurde mit Version 7 im Block <GUETEMASSE> ein Schalter <Ereignis-Definition> geschaffen. Über diesen ist wahlweise <ueberschreitung> (Betrachtung Hochwasser; Standard) oder <unterschreitung> (Betrachtung Niedrigwasser) anzugeben.
- Ab Version 6 ist das Fortschreiben der Unsicherheits-Statistik auf Grundlage vorhandener Auswertungen und neuer Ereignisse möglich. Hierzu ist in der Steuerdatei im Block <PEGEL> ein Schalter <Fortschreiben> vorgesehen. Ist dieser auf <ja> gesetzt, so werden zu den Originaldaten des aktuellen Ereignisses und die Ergebnisse der bisherigen Auswertungen (CSV-Dateien) eingelesen. Auf dieser Grundlage werden die statistischen Auswertungen für den gesamten Zeitraum (zurückliegende Ereignisse + aktuelles Ereignis) fortgeführt (Details siehe Kapitel 3.1.1).
- Ab Version 6 ist es möglich, den auszuwertenden Ereigniszeitraum beim Programmaufruf zu übergeben. Dies ist insbesondere im Zusammenhang mit der kontinuierlichen Fortschreibung der Auswertungen im operationellen Betrieb sinnvoll. Hierfür ist in der Steuerdatei im Block <EREIGNISZEITRAEUME> ein Schalter <Ereignisse aus Aufruf:> vorgesehen (Details siehe Kapitel 3.1.2).
- Ab Version 7 ist eine Auswertung der Abweichung der Scheitelwerte der Vorhersagen vom Scheitelwert der Messung möglich. Dafür wurde in der Steuerdatei im Block <GUETEMASSE> ein Schalter <Abweichung Scheitel:> eingeführt. Wird dieser auf <ja> gesetzt, so werden für jeden Pegel die Scheitelwerte der Vorhersagen in einer CSV-Datei ausgegeben (Details siehe Kapitel 2.9).
- Mit Version 8 wurde das LILA-Format in ProFoUnD eingeführt (näheres in Kapitel 3.1.1).
- Neuerungen in Version 8.1:
  - Optional können nun spezifisch für jeden Pegel Schwellenwerte zur Erstellung der Kontingenztafeln bzw. zur Berechnung der kategorischen Gütemaße vorgegeben werden. Dazu sind für den entsprechenden Pegel in der Steuerdatei im Block <PEGEL> der Schalter <Schwellenw. kat. Guete:> einzufügen und die Schwellenwerte festzulegen. Somit können nun auch eigene Abfluss-Grenzwerte für die Erstellung der Kontingenztafeln festgelegt werden, die unabhängig von den pegelspezifischen Schwellenwerten für die Vereinbarung der hydrologischen Fälle sind.
  - In der Steuerdatei wurde ein Schalter <Sprache> bzw. <Langue> eingeführt. Mit diesem lässt sich die Sprache der Steuerdatei und wichtiger Log-Ausgaben des Programms auf Französisch stellen. Dazu ist als Sprache <FR> anzugeben. Ist die Zeile mit <Sprache> bzw. <Langue> nicht vorhanden, wird automatisch als Sprache Deutsch angenommen.

## 2 Grundlagen und Berechnungsalgorithmen des Programms

Die nachfolgenden Erläuterungen beziehen sich häufig auf Abflüsse (z.B. Abkürzung Q). Diese können aber analog auf Wasserstände angewandt werden.

### 2.1 Ausgabe von Ganglinien

Nach dem Einlesen der Daten werden aus den Vorhersagen zusammenhängende Ganglinien für ausgewählte Vorhersagetiefen extrahiert. Diese Ganglinien werden für die ausgewählten Pegel und Ereignisse jeweils zusammen mit der gemessenen Ganglinie wahlweise in eine Datei im GMD- und/oder HMZ-Format oder in eine Datei im LILA-Format geschrieben, sodass die Messung mit den Vorhersagen der unterschiedlichen Dauern verglichen werden kann (siehe Bild 3.3, S. 47). Das Ausgabeformat kann über die Steuerdatei ausgewählt werden.

Zusätzlich können optional die angeforderten kontinuierlichen Einzelfehler (siehe unten) als Ganglinien mit ausgegeben werden.

Darüber hinaus kann optional auch der hydrologische Fall ausgegeben werden, der dem jeweiligen Vorhersagewert zugeordnet ist. Dabei ist zu beachten, dass dieser hydrologische Fall nicht vom kontinuierlichen Einzelfehler abhängt. Somit wird je Vorhersagewert nur ein Mal der hydrologische Fall ausgegeben, unabhängig davon, wie viele kontinuierliche Fehlermaße angefordert wurden. Die zugehörigen hydrologischen Fälle werden nur in die HMZ- bzw. LILA-Dateien ausgegeben, da eine Ausgabe im GMD-Format nicht sinnvoll wäre. Anhand der Zusammenschau von Datumsangabe, Einzelfehler und hydrologischem Fall lassen sich einzelne auffällige Vorhersagen innerhalb der HMZ- bzw. LILA-Dateien leicht identifizieren.

### 2.2 Klasseneinteilung und hydrologische Fälle

Die Qualität der Vorhersagen hängt erfahrungsgemäß in erheblichem Maße von der aktuellen hydrologischen Situation ab. So sind Vorhersagen im stationären Mittel- und Niedrigwasserbereich naturgemäß einfacher und genauer als Vorhersagen im ansteigenden Ast eines Hochwassers. Aus diesem Grund können die einzelnen Vorhersagewerte mit ProFoUnD in unterschiedliche Klassen und daraus resultierende hydrologische Fälle unterteilt werden. Die Ermittlung kontinuierlicher Fehlermaße (siehe unten) erfolgt dabei zum einen für alle Vorhersagen (Fall 0) sowie nochmals getrennt für die unterschiedlichen hydrologischen Fälle.

Für die Einteilung in hydrologische Fälle stehen in ProFoUnD die drei folgenden Verfahren zur Verfügung:

- Nach Bayer. LfU (Stand 2008)
- Nach IKSMS, altes Verfahren (Stand 2009)
- Nach IKSMS, neues Verfahren (Stand 2012)

Die unterschiedlichen Vorgehensweisen werden nachfolgend ausführlich erläutert.

## 2.2.1 Hydrologische Fälle nach dem Verfahren des Bayerischen LfU

Sofern das Verfahren nach Bayerischen LfU gewählt wurde, erfolgt die Klasseneinteilung der Vorhersagewerte anhand von zwei Schwellenwerten je Pegel: dem Schwellenwert für den Übergang von NQ zu MQ (NQM) und dem Schwellenwert für den Übergang von MQ zu HQ (MQH). Hiermit erfolgt die Zuordnung der Vorhersagewerte zu einer von 9 möglichen Klassen. Die Zugehörigkeit des Vorhersagewerts zu einer der 9 Klassen ist durch den jeweiligen Vorhersagewert (aktuell vorhergesagter Abfluss) und den gemessenen Abfluss zum zugehörigen Vorhersagezeitpunkt definiert. Die nachfolgende Matrix definiert die 9 möglichen Klassen:

		zu klassifizierender Wert (vorhergesagter Q)		
		NQ	MQ	HQ
zugehöriger gem. Q zum Vorhersagezeitpunkt	NQ	1	2	3
	MQ	4	5	6
	HQ	7	8	9

Für die Bestimmung der Klassenzugehörigkeit muss grundsätzlich ein Messwert zum jeweiligen Vorhersagezeitpunkt der Vorhersagetiefe vorliegen, d.h. der erste Messwert sollte mindestens die Anzahl der Stunden der größten Vorhersagetiefe vor dem Ereignisbeginn liegen. Hierauf ist zu achten, wenn die Ganglinien der Messwerte bereitgestellt werden.

Die 9 Klassen können zu unterschiedlichen hydrologischen Fällen zusammengefasst werden. In Modellanwendungen, die im Auftrag des WWA Kempten durchgeführt wurden, haben sich folgende Definitionen als praktikabel und sinnvoll erwiesen:

- Mittel- und Niedrigwasser (MNW):  
Klassen 1, 2, 4 und 5.  
Sowohl der vorhergesagte Abflusswert als auch der gemessene Abfluss zum Vorhersagezeitpunkt liegen im Mittel- oder Niedrigwasserbereich.
- Hochwasser (HW):  
Klasse 9.  
Sowohl der vorhergesagte Abflusswert als auch der gemessene Abfluss zum Vorhersagezeitpunkt liegen im Hochwasserbereich.
- Vor dem Hochwasser (vHW):  
Klassen 3 und 6.  
Der gemessene Abfluss zum Vorhersagezeitpunkt liegt im Mittel- oder Niedrigwasserbereich, wogegen der vorhergesagte Wert im Hochwasserbereich liegt.
- Nach dem Hochwasser (nHW):  
Klassen 7 und 8.  
Der gemessene Abfluss zum Vorhersagezeitpunkt liegt im Hochwasserbereich, wogegen der vorhergesagte Wert im Mittel- oder Niedrigwasserbereich liegt.

Beim Fall MNW ist es möglich, dass sowohl der gemessene Abfluss zum VZP als auch der vorhergesagte Abfluss im MQ- oder NQ-Bereich liegen, zwischen beiden Zeitpunkten jedoch ein Hochwasser auftrat. Solche Sonderfälle werden dennoch im hydrologischen Fall MNW abgehandelt und nicht gesondert ausgewiesen, da sie nur in kleinen Einzugsgebieten bei großen Vorhersagetiefen auftreten. In solchen Fällen wird der Gesamtfehler von der Spreizung infolge der Unsicherheiten beim vorhergesagten Niederschlag dominiert. Der statische Fehler (Vorhersagetest mit gemessenen Niederschlägen) hat dann nur eine untergeordnete Bedeutung. Zudem liegt das Hauptaugenmerk auf dem Hochwasserbereich.

## 2.2.2 Hydrologische Fälle nach dem alten IKSMS-Verfahren (Stand 2009)

Bei den im Auftrag und in Abstimmung mit dem LUWG implementierten Verfahren zur Bestimmung der hydrologischen Fälle gemäß IKSMS werden aktuelle Vorhersagewerte sowie die Richtung der jeweiligen Vorhersage als Kriterien für die Klassenzugehörigkeit angewandt. Dabei kann zwischen einem älteren Ansatz aus dem Jahr 2009 und einem aktualisierten Ansatz aus dem Jahr 2012 (ab Version 4) unterschieden werden. Die beiden Ansätze unterschieden sich hinsichtlich der Ermittlung der Richtung der Vorhersage.

In beiden IKSMS-Ansätzen werden für die Klasseneinteilung an den einzelnen Pegel zunächst bis zu fünf Schwellenwerte vorgegeben. Alternativ kann auch ohne einen Schwellenwert gearbeitet werden. In diesem Fall liegen alle Vorhersagen in einem (Abfluss-) Bereich und es erfolgt lediglich eine Unterteilung nach der Richtung. Wenn kein Schwellenwert vorgegeben ist, können jedoch keine kategorischen Gütemaße bestimmt werden, da diese eine Klassengrenze erfordern (siehe unten).

In den nachfolgenden Erläuterungen wird davon ausgegangen, dass vier Schwellenwerte vorgegeben werden und diese Schwellenwerte folgenden Abfluss- (oder Wasserstands-) Kennwerten entsprechen: MQ, 2xMQ, HQ<sub>2</sub> und HQ<sub>5</sub>. Die Schwellenwerte sind aber grundsätzlich frei wählbar, so dass die Erläuterungen analog auch für andere Schwellenwerte bzw. eine andere Anzahl an Schwellenwerten gelten.

Die Vorhersagewerte werden gemäß der 4 Schwellenwerte zunächst in einen von 5 Bereichen eingeteilt:  $\leq MQ$ ,  $\leq 2xMQ$ ,  $\leq HQ_2$ ,  $\leq HQ_5$ ,  $> HQ_5$ .

Bei dem hier beschriebenen alten IKSMS-Verfahren wird die Richtung der jeweiligen Vorhersage anhand der Gesamtvorhersage bestimmt. Hierfür wird der erste Wert der Vorhersage, der Median aller Werte der Vorhersage und der Maximalwert aller Werte der Vorhersage ermittelt.

Hieraus werden folgende Zusammenhänge hergeleitet:

Richtung 1:  $Q_{\max,vhs} > Q_{1,vhs}$  UND  $Q_{\text{med},vhs} > Q_{1,vhs} \Rightarrow$  Überwiegend steigende Vorhersage

Richtung 2:  $Q_{\max,vhs} > Q_{1,vhs}$  UND  $Q_{\text{med},vhs} \leq Q_{1,vhs} \Rightarrow$  VHS mit steigendem und fallendem Anteil

Richtung 3:  $Q_{\max,vhs} = Q_{1,vhs} \Rightarrow$  Fallende Vorhersage

mit:

$Q_{\max,vhs}$  [m<sup>3</sup>/s] Maximaler Abfluss der untersuchten Vorhersage

$Q_{\text{med},vhs}$  [m<sup>3</sup>/s] Median aller Abflüsse der untersuchten Vorhersage

$Q_{1,vhs}$  [m<sup>3</sup>/s] Erster Vorhersagewert der untersuchten Vorhersage

Aus den beiden Kriterien „Bereich des Vorhersagewerts“ und „Richtung der Gesamtvorhersage“ ergibt sich folgende Matrix für die Klassifikation der Vorhersagewerte (Nummerierung der Klassen):

		Bereich des VHS-Werts				
		Ber1 $\leq MQ$	Ber2 $\leq 2xMQ$	Ber3 $\leq HQ_2$	Ber4 $\leq HQ_5$	Ber5 $> HQ_5$
Richtung der VHS	Ri1	1	2	3	4	5
	Ri2	6	7	8	9	10
	Ri3	11	12	13	14	15

Bei 4 Schwellenwerten (= 5 Bereiche) ergeben sich also 15 unterschiedliche Klassen. Im einfachsten Fall werden diese Klassen nicht weiter zu hydrologischen Fällen zusammengefasst. In diesem Fall entspricht die Klassennummerierung in obiger Matrix auch der Nummerierung der hydrologischen Fälle.

Die recht große Anzahl an Klassen lässt sich jedoch auch zu weniger hydrologischen Fällen zusammenfassen, indem man Bereiche auswählt, die nicht nochmal nach der Richtung differenziert werden. Fasst man also z.B. den Bereich 1 zusammen, so bilden die Klassen 1, 6 und 11 einen hydrologischen Fall. Alle anderen Bereiche werden weiterhin in die drei Richtungen differenziert, so dass in diesem Beispiel noch 13 hydrologische Fälle unterschieden werden. In dem aufgeführten Beispiel ergibt sich folgende Nummerierung für die hydrologischen Fälle:

		Bereich des VHS-Werts				
		Ber1 $\leq MQ$	Ber2 $\leq 2xMQ$	Ber3 $\leq HQ_2$	Ber4 $\leq HQ_5$	Ber5 $> HQ_5$
Richtung der VHS	Ri1	1	2	3	4	5
	Ri2	1	6	7	8	9
	Ri3	1	10	11	12	13

Generell ist zu beachten, dass alle kontinuierlichen Gütemaße einerseits getrennt für hydrologische Fälle ermittelt werden und andererseits auch immer für alle Werte (unabhängig von Bereich und Richtung). Bei der Betrachtung und Auswertung aller Werte spricht man vom Fall 0.

### 2.2.3 Hydrologische Fälle nach dem neuen IKSMS-Verfahren (Stand 2012)

Das im Jahr 2012 aktualisierte neue IKSMS-Verfahren unterscheidet sich vom alten Verfahren hinsichtlich der Ermittlung der Richtung der Vorhersage. In Abstimmung mit dem Auftraggeber wird im neuen Verfahren die Richtung der jeweiligen Vorhersage (aus der der einzelne Vorhersagewert entnommen wurde) anhand ihres ersten Wertes, eines vorzugebenden Perzentils und des Medians aller Werte bis zur aktuell betrachteten Vorhersagetiefe ermittelt.

Die Unterschiede zum alten Verfahren liegen also darin, dass die Vorhersage nur bis zum aktuellen Vorhersagewert betrachtet wird (und nicht die gesamte Vorhersage) und ein vorzugebender Perzentilwert anstelle des Maximalwerts verwendet wird.

Hieraus werden folgende Zusammenhänge hergeleitet:

Richtung 1:  $Q_{akt} \geq Q_1$  UND  $Q_{akt} \geq Q_{perz}$   $\Rightarrow$  Überwiegend steigende Vorhersage

Richtung 2:  $[Q_{akt} \geq Q_1$  UND  $Q_{akt} < Q_{perz}]$  ODER  $[Q_{akt} < Q_1$  UND  $Q_{akt} > Q_{med}]$   
 $\Rightarrow$  VHS mit steigendem und fallendem Anteil

Richtung 3:  $Q_{akt} < Q_1$  UND  $Q_{akt} \leq Q_{med}$   $\Rightarrow$  Fallende Vorhersage

mit:

$Q_{akt}$  [ $m^3/s$ ] Wert der Vorhersage zum aktuell untersuchten Zeitpunkt.

$Q_{perz}$  [ $m^3/s$ ] Perzentil aus den Werten der Vorhersage bis zum aktuell untersuchten Zeitpunkt.

$Q_{med}$  [ $m^3/s$ ] Median aus den Werten der Vorhersage bis zum aktuell untersuchten Zeitpunkt.

$Q_1$  [ $m^3/s$ ] Erster Vorhersagewert der untersuchten Vorhersage.

Zu beachten ist, dass für die Definition der Richtung nur die Werte in der aktuell untersuchten Vorhersage bis zur aktuell untersuchten Vorhersagetiefe verwendet werden. Bei einer Vorhersagetiefe von 1h muss daher die 2h-VHS als  $Q_{akt}$  verwendet werden. Das für  $Q_{perz}$  zu verwendende Perzentil wird über die Steuerdatei vorgegeben. Je größer das gewählte Perzentil ist, umso strikter ist die Auswahl für die Richtung 1 (überwiegend steigende Vorhersage). Empfohlen werden Perzentile größer/gleich 75. Ein Perzentil von 100 entspricht dem Maximalwert.

Die weitere Unterteilung in Klassen und hydrologische Fälle erfolgt im neuen IKSMS-Verfahren in gleicher Weise wie im alten IKSMS-Verfahren (siehe oben) und wird daher hier nicht nochmals beschrieben.

## 2.3 Kontinuierliche Fehlermaße

### 2.3.1 Kontinuierliche Einzelfehler

Für die spezifizierten Vorhersagetiefen werden die vorhergesagten Abflusswerte mit den jeweils zugehörigen Messwerten verglichen. Hierfür werden unterschiedliche Fehlermaße berechnet. Die Fehlermaße werden zum Einen immer für alle verfügbaren Wertepaare ermittelt (Fall 0). Zum Anderen erfolgt die Ermittlung getrennt für die benutzerdefinierten hydrologischen Fälle (siehe oben). Für die Wertepaare können im Einzelnen die nachfolgend aufgeführten Fehler ermittelt werden.

Die Abweichung (AbsAbw [ $\text{m}^3/\text{s}$ ]):

$$AbsAbw = Qgem_i - Qvhs_i \quad (Gl. 2.1)$$

mit:

$Qvhs_i$  [ $\text{m}^3/\text{s}$ ] vorhergesagter Abfluss für den Zeitpunkt  $i$

$Qgem_i$  [ $\text{m}^3/\text{s}$ ] gemessener Abfluss zum Zeitpunkt  $i$

Die prozentuale Abweichung (ProAbw [%]):

$$ProAbw = \frac{Qgem_i - Qvhs_i}{|Qvhs_i|} * 100 \quad (Gl. 2.2)$$

Der Quotient aus gemessenem und vorhergesagtem Wert (Quot [ ]):

$$Quot = \frac{Qgem_i}{Qvhs_i} \quad (Gl. 2.3)$$

Der natürliche Logarithmus des Quotienten (LnQuot [ ]):

$$LnQuot = \ln\left(\frac{Qgem_i}{Qvhs_i}\right) \quad (Gl. 2.4)$$

Die quadrierte Abweichung (QuadAbw [ $\text{m}^6/\text{s}^2$ ]):

$$QuadAbw = (Qgem_i - Qvhs_i)^2 \quad (Gl. 2.5)$$

Der Zeitfehler (Lag) [h]:

Zur Ermittlung des Zeitfehlers wird die gemessene Ganglinie innerhalb eines 12-stündigen Zeitfensters in 1-Stundenschritten so verschoben, dass die quadrierte Abweichung minimal wird. Der jeweilige Lag wird somit für jeden einzelnen Vorhersagewert in vollen Stundenwerten ermittelt. Positive Werte bedeuten dabei, dass die Messung auf einen früheren Zeitpunkt verschoben wurde (Voraneilen der Vorhersage) und negative Werte signalisieren, dass die Messung auf einen späteren Zeitpunkt verschoben wurde (Nachlaufen der Vorhersage).

Hinsichtlich der oben gewählten Fehlerdefinitionen sind die folgenden (ungewöhnlichen) Punkte zu beachten:

- Die Maße für die Abweichung zwischen vorhergesagtem und gemessenem Wert sind vorzeichenbehaftet. Eine Überschätzung durch die Vorhersage bedingt bei der Abweichung, der prozentualen Abweichung und beim Logarithmus des Quotienten negative Fehler.
- Die relativen Fehler werden auf den Vorhersagewert bezogen ( $Q_{vhs}$  im Nenner).

Diese ungewöhnliche, der Intuition widersprechende Fehlerdefinition wurde gewählt, da im operationellen Fall lediglich der vorhergesagte Abfluss (bzw. Wasserstand) bekannt ist. Der Messwert ist die gesuchte Größe, d.h. der Abflussbereich, in dem der Messwert liegt, muss mithilfe des Vorhersagewerts und des Fehlers abgeschätzt werden.

Bei der hier gewählten Definition von  $AbsAbw$  errechnet sich der (gesuchte) Messwert mit Unsicherheitszuschlag aus der Summe des Vorhersagewerts und  $AbsAbw$ . Beim Quotienten ergibt sich der (gesuchte) Messwert mit Unsicherheitszuschlag aus dem Produkt des Vorhersagewerts und  $Quot$ . Die gewählten Fehlerdefinitionen ermöglichen somit eine einfache Handhabung im operationellen Fall.

Die Einzelwerte werden getrennt nach hydrologischen Fällen (inkl. Fall 0 = alle Werte) in aufsteigender Reihenfolge sortiert in CSV-Dateien ausgegeben und stehen für weitergehende Auswertungen zur Verfügung.

### 2.3.2 Tolerierbares Zeitfenster bei kontinuierlichen Fehlern

Bei den oben angeführten Erläuterungen zu den kontinuierlichen Einzelfehlern wurde davon ausgegangen, dass Messwerte mit Vorhersagewerten für exakt denselben Zeitpunkt verglichen werden. Insbesondere bei größeren Vorhersagetiefen ist es jedoch nicht entscheidend, dass die Vorhersage zeitlich absolut exakt ist, vielmehr ist hier die korrekte Höhe der Vorhersage maßgebend. Daher gibt es für die kontinuierlichen Fehler (mit Ausnahme des Lag) die Möglichkeit, tolerierbare Zeitfenster vorzugeben.

Das jeweils tolerierbare Zeitfenster kann dabei als Funktion der Vorhersagetiefe angegeben werden, wobei dieses Zeitfenster maximal so groß sein darf wie die jeweilige Vorhersagetiefe.

Ist ein tolerierbares Zeitfenster vorgegeben, wird spezifisch für jedes Ereignis und jeden Fehlertyp der optimale Zeitverschub ermittelt. Da dieser optimale Zeitverschub ereignisspezifisch ermittelt wird, ist es auf jeden Fall empfehlenswert Ereigniszeiträume vorzugeben. Sofern nur ein Ereigniszeitraum (über den gesamten zu untersuchenden Zeitraum) vorgegeben ist, wird nur ein optimaler Verschub je Vorhersagetiefe für den gesamten Zeitraum ermittelt.

Um den ereignisspezifischen optimalen Zeitverschub (je Vorhersagetiefe) zu identifizieren, wird die gemessene Ganglinie (in Stundenschritten) innerhalb des tolerierbaren Zeitfensters gegen die (zusammengesetzte) Vorhersageganglinie eines Ereignisses verschoben. Spezifisch für den jeweils betrachteten Fehler wird der Zeitverschub festgehalten, bei dem der über das Ereignis gemittelte Wert des Fehlers sein Minimum erreicht. Konkret wird dieses Optimum des Zeitverschubs bei den unterschiedlichen Fehlermaßen dort festgelegt, wo folgende Bedingungen zutreffen:

AbsAbw: Minimum der über das Ereignis gemittelten Beträge der Einzelwerte

ProAbw: Minimum der über das Ereignis gemittelten Einzelwerte

Quot: Minimum der über das Ereignis gemittelten Beträge der natürlichen Logarithmen der Einzelwerte

LnQuot: Minimum der über das Ereignis gemittelten Beträge der Einzelwerte

QuadAbw: Minimum der über das Ereignis gemittelten Einzelwerte

Hieraus folgt, dass für unterschiedliche Fehlermaße jeweils unterschiedliche Zeitverschiebe für dasselbe Ereignis ermittelt werden können (mit Ausnahme von Quot und LnQuot).

Wenn mit tolerierbaren Zeitfenstern gearbeitet wird, werden die oben beschriebenen Einzelfehler (und alle daraus abgeleiteten Maße) für den (ereignisspezifisch ermittelten) optimalen Zeitverzug innerhalb des tolerierbaren Zeitfensters berechnet. Folglich sind die so ermittelten Einzelfehler im Allgemeinen kleiner, als bei der Auswertung ohne tolerierbares Zeitfenster. Bei der Ausgabe der Rangfolgen der Einzelfehler werden die jeweils verwendeten Zeitverschiebe mit angegeben.

Für den Zeitfehler (Lag) wird kein tolerierbares Zeitfenster zugelassen.

### 2.3.3 Mittlere kontinuierliche Fehler

#### Mittlere kontinuierliche Fehler:

Zusätzlich werden aus den oben genannten Einzelwerten der Fehler ereignisübergreifende mittlere Fehler je Pegel, Vorhersagetiefe und hydrologischem Fall (inkl. Fall 0) ermittelt und ausgegeben. Dies erfolgt nur, wenn die entsprechenden Einzelfehler angefordert wurden.

Die mittlere vorzeichenbehaftete Abweichung ( $MAbsAbw$  [ $m^3/s$ ]):

$$MAbsAbw = \frac{1}{n} * \sum_{i=1}^n Qgem_i - Qvhs_i \quad (Gl. 2.6)$$

mit:

n [ ] Gesamtzahl der Wertepaare

Die mittlere Abweichung (Mittelwert der Beträge der Abweichungen,  $MBetAbw$  [ $m^3/s$ ]):

$$MBetAbw = \frac{1}{n} * \sum_{i=1}^n |Qgem_i - Qvhs_i| \quad (Gl. 2.7)$$

Hierbei ist zu beachten, dass sich bei der mittleren vorzeichenbehafteten Abweichung ( $M_{AbsAbw}$ ) positive und negative Abweichungen gegeneinander aufrechnen. Ein positiver Mittelwert ist also als Anzeichen für eine systematische Unterschätzung der Messwerte durch die Vorhersagen anzusehen. Im Gegensatz hierzu werden bei der mittleren Abweichung ( $M_{BetAbw}$ ) die Beträge betrachtet. Sie ist somit ein Maß für die mittlere zufällige Abweichung. Wenn der Betrag von  $M_{AbsAbw}$  nahezu denselben Wert annimmt wie  $M_{BetAbw}$ , so weist das darauf hin, dass die beobachteten Abweichungen nahezu alle systematisch sind.

Die mittlere prozentuale Abweichung ( $M_{ProAbw}$  [%]):

$$M_{ProAbw} = \frac{100}{n} * \sum_{i=1}^n \frac{|Q_{gem_i} - Q_{vhs_i}|}{|Q_{vhs_i}|}$$

(Gl. 2.8)

Im Gegensatz zur absoluten Abweichung wird der Mittelwert der prozentualen Abweichung ( $M_{ProAbw}$ ) nur aus dem Betrag der prozentualen Abweichungen berechnet. Positive wie negative prozentuale Abweichungen gehen gleichermaßen in die Berechnung des mittleren Fehlers ein. Aussagen über eine mögliche systematische Abweichung sind hier folglich nicht möglich. Ist eine solche Aussage gewünscht, so kann sie jedoch leicht aus der Verteilung der einzelnen Fehlerwerte abgeleitet werden (siehe Percentile).

Der mittlere Quotient aus vorhergesagtem und gemessenem Abfluss ( $M_{Quot}$  [ ]):

$$M_{Quot} = \frac{1}{n} * \sum_{i=1}^n \frac{Q_{gem_i}}{Q_{vhs_i}}$$

(Gl. 2.9)

Der Mittelwert des natürlichen Logarithmus des Quotienten aus vorhergesagtem und gemessenem Abfluss ( $M_{LnQuot}$  [ ]):

$$M_{LnQuot} = \frac{1}{n} * \sum_{i=1}^n \ln\left(\frac{Q_{gem_i}}{Q_{vhs_i}}\right)$$

(Gl. 2.10)

Beim Quotienten und dem natürlichen Logarithmus des Quotienten gehen die einzelnen Werte mit ihrem jeweiligen Vorzeichen ein. Ein  $M_{Quot}$  größer 1 bzw. ein positiver  $M_{LnQuot}$  deuten somit auf eine systematische Unterschätzung der Messwerte durch die Vorhersagen hin. Umgekehrt weist ein  $M_{Quot}$  kleiner 1 (und somit ein negativer  $M_{LnQuot}$ ) auf eine systematische Überschätzung durch die Vorhersagen hin.

Die mittlere quadrierte Abweichung ( $M_{QadAbw}$ ) sowie der RMSE:

$$M_{QadAbw} = \frac{1}{n} * \sum_{i=1}^n (Q_{gem_i} - Q_{vhs_i})^2$$

(Gl. 2.11)

$$RMSE = \sqrt{\frac{1}{n} * \sum_{i=1}^n (Qgem_i - Qvhs_i)^2}$$

(Gl. 2.12)

Wenn die quadrierte Abweichung angefordert ist, wird neben der mittleren quadrierten Abweichung und dem RMSE ein Skill Score unter Einbeziehung der Persistenzvorhersage berechnet. Dieses Gütemaß basiert auf einem Vergleich zwischen den herkömmlichen quadrierten Abweichungen der Vorhersagen und quadrierten Abweichungen, die sich bei einer Persistenzannahme ergäben. Zur Berechnung der quadrierten Abweichungen unter Persistenzannahme wird davon ausgegangen, dass die Abflüsse (bzw. Wasserstände) nach dem Vorhersagezeitpunkt konstant bleiben und dem letzten Messwert entsprechen. Hieraus ergibt sich der Persistenz-Skill-Score (SkillScorePer):

$$SkillScore Per = \left( 1 - \frac{\sum_{i=1}^n (Qgem_i - Qvhs_i)^2}{\sum_{i=1}^n (Qgem_i - Qper_i)^2} \right)$$

(Gl. 2.13)

mit:

$Qper_i$  [m<sup>3</sup>/s] Abfluss zum zugehörigen Vorhersagezeitpunkt (Persistenzannahme)

Der SkillScorePer kann Werte zwischen  $-\infty$  und  $+1$  annehmen. Werte kleiner 0 weisen darauf hin, dass die Vorhersagen schlechtere Ergebnisse liefern als die Persistenzannahme. Ein Wert von  $+1$  steht für eine perfekte Übereinstimmung zwischen Vorhersage und Messung.

Für den Zeitfehler werden wie für die absolute Abweichung Mittelwerte aus den vorzeichenbehafteten Zeitfehlern (MAbsLag) und den Beträgen (MBetLag) der Zeitfehler berechnet:

$$MAbsLag = \frac{1}{n} * \sum_{i=1}^n LAGgem_i - LAGvhs_i$$

(Gl. 2.14)

$$MBetLag = \frac{1}{n} * \sum_{i=1}^n |LAGgem_i - LAGvhs_i|$$

(Gl. 2.15)

Somit sind auch hier differenzierte Aussagen hinsichtlich eines systematischen und zufälligen Zeitfehlers möglich.

## 2.4 Häufigkeitsverteilungen der kontinuierlichen Fehler

### 2.4.1 Empirische Summenhäufigkeit und empirische Percentile

Aus den Rangfolgen der Einzelfehler wird für alle hydrologischen Fälle (inkl. Fall 0 = alle Fehlerwerte) programmintern die empirische kumulative Summenhäufigkeit ermittelt. Die Berechnung erfolgt auf Grundlage des Rangs des Einzelfehlers und der Belegungszahl des hydrologischen Falls. Dabei wird davon ausgegangen, dass die Fehler näherungsweise normalverteilt sind, weshalb der Zähler um 0,375 verringert und der Nenner um 0,25 erhöht wird (Helsel, et al., 1992 S. 23):

$$S_N = \frac{m - 0,375}{n + 0,25}$$

(Gl. 2.16)

mit:

$S_N$	[ ]	Kumulierte Summenhäufigkeit des Einzelfehlers (0 ... 1)
$m$	[ ]	Rang des Einzelfehlers
$n$	[ ]	Anzahl aller Einzelfehler (Belegungszahl)

Mithilfe der kumulierten empirischen Summenhäufigkeit werden die empirischen Percentilwerte für den Bereich [0.05, 0.1, 0.2, ..., 0.9, 0.95] ermittelt und tabellarisch ausgegeben. Hierfür werden die dem jeweiligen Percentil benachbarten empirischen Summenhäufigkeiten gesucht und zwischen diesen linear interpoliert. Sofern nicht genügend Summenhäufigkeitswerte vorliegen, um alle gesuchten Percentile zu ermitteln, werden für die nicht bestimmbaren Percentilwerte Fehlwerte = -9999.0 ausgegeben.

Dabei ist zu beachten, dass die empirischen Percentile aus allen Fehlern im gesamten empirischen Summenhäufigkeitsbereich [0.0 ... 1.0] ermittelt werden, während alle weiteren Größen aus dem empirischen Summenhäufigkeitsbereich [0.05 ... 0.95] ermittelt werden (siehe unten).

### 2.4.2 Momente der Häufigkeitsverteilungen

Die Mittelwerte, Standardabweichungen und Schiefen der empirischen Fehlerverteilungen werden über die ersten drei Momente der Einzelfehler berechnet. Daher wird auch von Momenten gesprochen, wenngleich die Schiefe nicht dem 3. Moment entspricht, sondern lediglich aus diesem abgeleitet wird.

Die Momente werden dabei getrennt für die hydrologischen Fälle (inkl. Fall 0) und die zu untersuchenden Vorhersagetiefen ermittelt. Bei der Berechnung werden lediglich jene Fehler verwendet, die im Summenhäufigkeitsbereich von [0.05 ... 0.95] liegen. Kleinere und größere Fehler werden als Ausreißer interpretiert und daher bei der Momentenberechnung nicht berücksichtigt.

Die Berechnung und Ausgabe der Momente erfolgt sobald mindestens 2 (Mittelwert und Standardabweichung) bzw. 3 (Schiefe) Fehlerwerte vorliegen, wenngleich eine statistische Bewertung erst ab einer Belegungszahl von ca. 30 sinnvoll erscheint. Die jeweilige Belegungszahl wird zur besseren Beurteilung der Momente mit ausgegeben.

Die drei statistischen Maßzahlen berechnen sich wie folgt aus den Einzelfehlern (innerhalb des empirischen Summenhäufigkeitsbereichs [0.05 ... 0.95]):

$$MW = \frac{1}{n} * \sum_{i=1}^n fehler_i$$

(Gl. 2.17)

mit:

MW	Mittelwert der empirischen Fehlerverteilung (Einheit je nach Fehlermaß)
fehler <sub>i</sub>	i-ter Einzelfehler (Einheit je nach Fehlermaß)
n	Anzahl aller Einzelfehler in der jeweiligen Klasse (Belegungszahl)

$$STABW = \sqrt{\frac{1}{n-1} * \sum_{i=1}^n (fehler_i - MW)^2}$$

(Gl. 2.18)

mit:

STABW	Standardabweichung der empirischen Fehlerverteilung (Einheit je nach Fehlermaß)
-------	---

$$SCHIEFE = \frac{n}{(n-1) * (n-2)} * \sum_{i=1}^n \frac{(fehler_i - MW)^3}{STABW^3}$$

(Gl. 2.19)

mit:

SCHIEFE	Schiefe der empirischen Fehlerverteilung
---------	--

Hinsichtlich der Schiefe ist zu beachten, dass unterschiedliche Definitionen für deren Berechnung existieren (vgl. Helsel et al., 1992, Press et al., 1992). Im vorliegenden Programm wurde die Definition von Helsel et al., 1992 S. 10 übernommen. Vergleichsrechnungen ergaben, dass dieselbe Definition in EXCEL verwendet wird.

Geht man davon aus, dass die Einzelfehler der einzelnen Vorhersagetiefen normalverteilt sind, so lässt sich die theoretische Fehlerverteilung als Normalverteilung mit den Parametern MW und STABW ausdrücken.

### 2.4.3 Polynome zur Beschreibung der Momente

Sofern dies über die Steuerdatei angefordert ist, werden an den Verlauf der ersten beiden Momente über die Vorhersagetiefe Polynome 2. Grades (3 Parameter) angepasst. Somit erhält man für jeden Pegel und jeden hydrologischen Fall je zwei Polynome, die den Mittelwert und die Standardabweichung als Funktion der Vorhersagetiefe ausdrücken. Die Polynome sind wie folgt definiert:

$$MOMENT(x) = a0 + b1 * x + bw * x^2 \quad (Gl. 2.20)$$

mit:

MOMENT	Mit dem Polynom berechneter Mittelwert oder Standardabweichung der Fehlerverteilung (Einheit je nach Fehlermaß)
x	[h] Vorhersagetiefe
a0	Achsenabschnitt des Polynoms (Einheit je nach Fehlermaß)
b1, b2	Koeffizienten des Polynoms

Sind die Parameterwerte a0, b1 und b2 des Polynoms bekannt, so können die Mittelwerte und Standardabweichungen der Fehlerverteilungen als Funktion der Vorhersagetiefe (x in Gl. 2.20) berechnet werden. Unter der Annahme, dass die Einzelfehler normalverteilt sind, kann mithilfe der so ermittelten MW und STABW wiederum die theoretisch zu erwartende Fehlerverteilung für die jeweilige Vorhersagetiefe berechnet werden.

Die optimalen Werte für die drei Parameter (a0, b1, b2) des Polynoms werden unter Verwendung der vorliegenden Momente durch Minimierung der Fehlerquadrate ermittelt. Hierfür wird die folgende Zielfunktion minimiert:

$$X^2(\mathbf{a}) = \sum_{i=1}^n \left( \frac{y_i - y(x_i; \mathbf{a})}{\sigma_i} \right)^2 \quad (Gl. 2.21)$$

mit:

$\chi^2$	[ ] Zu minimierender Wert der Zielfunktion
$y_i$	[ ] Messwert am Punkt i (hier: für Vorhersagetiefe ermitteltes Moment)
$\sigma_i$	[ ] Standardabweichung des Messwertes am Punkt i (hier: Konstant = 1)
$x_i$	[ ] Unabhängige Variable am Punkt i (hier: Vorhersagetiefe)
$\mathbf{a}$	[ ] Vektor mit den Parametern der Anpassungsfunktion (hier: Parameter des Polynoms a0, b1 und b2)

Dabei ist zu beachten, dass über den Parameter  $\sigma_i$  eine Gewichtung der Vorhersagetiefen vorgenommen werden kann. Im vorliegenden Programm wurde in Absprache mit dem Auftraggeber auf eine solche Gewichtung jedoch verzichtet, so dass alle eingehenden MOMENTE unabhängig von ihrer Vorhersagetiefe in gleichem Maße in die Ermittlung der optimalen Parameter eingehen.

Für die Minimierung der Zielfunktion und die damit einhergehende Ermittlung der optimalen Parameterwerte des Polynoms wird im Programm die Singulärwertzerlegung angewandt. Dieses numerisch etwas aufwändigere Verfahren ist äußerst stabil. Vergleiche mit dem in EXCEL implementierten SOLVER haben gezeigt, dass mit der in ProFoUnD implementierten Singulärwertzerlegung das Minimum der Fehlerquadrate und die zugehörigen optimalen Parameterwerte zuverlässig ermittelt werden.

Für die Anpassung des Polynoms werden nur jene Momente verwendet, deren Wert aus mindestens 30 Einzelfehlern ermittelt wurde. Generell wird ein Polynom nur dann angepasst, wenn entsprechende Momente für mindestens fünf unterschiedliche Vorhersagetiefen vorliegen.

Die MW können dabei sowohl positive wie negative Werte annehmen, so dass hier keine weitere Überprüfung der Plausibilität der mit dem Polynom berechneten MW notwendig ist.

Hinsichtlich der STABW als Parameter der Normalverteilungen ist jedoch zu beachten, dass diese nicht kleiner als Null werden darf. Bei der Anpassung des Polynoms an die empirischen STABW wird programmintern daher wie folgt vorgegangen:

- Zunächst wird das Polynom mit allen drei Parametern angepasst.
- Anschließend wird überprüft, ob der Wert des Polynoms für die Vorhersagetiefe Null größer/gleich Null ist.
  - Ist dies der Fall, wird das Polynom so beibehalten.
  - Ist dies nicht der Fall, wird der Achsenabschnitt des Polynoms auf Null fixiert ( $a_0 = 0.0$ ). Die beiden verbleibenden Parameter werden erneut angepasst.
- Abschließend wird geprüft, ob die Werte des Polynoms für alle Vorhersagetiefen von eins bis zur maximal gültigen VHS-Tiefe positiv sind.
  - Falls dies nicht der Fall ist, wird eine Warnung in die Log-Datei ausgegeben. Die Parameterwerte des entsprechenden Polynoms und die Percentile der resultierenden Verteilungen werden als Fehlwerte ausgegeben.

Der Gültigkeitsbereich des Polynoms erstreckt sich auf Vorhersagetiefen zwischen 0 Stunden und der maximalen Vorhersagetiefe, die zur Ermittlung des Polynoms verwendet wurde. Eine Extrapolation auf größere Vorhersagetiefen ist nicht generell zulässig. Daher wird der Gültigkeitsbereich der Polynome mit ausgegeben.

Soll die Abschätzung der theoretischen Fehlerverteilung auch für größere Vorhersagetiefen außerhalb des Gültigkeitsbereichs erfolgen, so wird empfohlen, das Ergebnis des Polynoms für die maximal zulässige Vorhersagetiefe zu verwenden. Dieser Wert wird gemeinsam mit den Parameterwerten und dem Gültigkeitsbereich des Polynoms ausgegeben.

#### 2.4.4 Normalverteilung und theoretische Percentile

Wie oben beschrieben, werden aus der empirischen Fehlerverteilung (Summenhäufigkeiten) die empirischen Percentile im Bereich [0.05, 0.1, 0.2, ..., 0.9, 0.95] ermittelt. Dabei handelt es sich also um die Fehler, die in 5%, 10%, 20% usw. der Fälle unterschritten oder erreicht werden.

Auf der Grundlage früherer Untersuchungen wird trotz z.T. deutlicher Abweichungen im Programm davon ausgegangen, dass die empirischen Einzelfehler näherungsweise normalverteilt sind. Die Wahrscheinlichkeit der Fehler lässt sich mithilfe der Normalverteilung als Funktion des MW und der STABW ausdrücken.

Für die kumulierte Wahrscheinlichkeit, dass ein Fehler kleiner oder gleich seines Wertes ist, lässt sich für die Normalverteilung wie folgt ausdrücken:

$$F(x) = \frac{1}{STABW \sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{1}{2} * \left(\frac{y - MW}{STABW}\right)^2\right) dy \quad (\text{Gl. 2.22})$$

mit:

$F(x)$  [ ] Kumulierte Wahrscheinlichkeit des Wertes  $x$  (hier:  $x = \text{fehler}$ )

Für Gleichung 2.22 existiert keine analytische Lösung. Allerdings liegen für die sogenannte Standardnormalverteilung mit  $MW = 0$  und  $STABW = 1$  und  $x = z$  tabellierte Werte für  $F(z)$  vor, aus denen die kumulierten Wahrscheinlichkeiten für  $z$  (= normierte  $x$ -Werte) entnommen werden können. Genauso kann auch ausgehend von einer kumulierten Wahrscheinlichkeit  $F(x)$  der zugehörige  $z$ -Wert entnommen werden.

Die Umrechnung von  $x$ - in  $z$ -Werte erfolgt dabei wie folgt:

$$z = \frac{(x - MW)}{STABW} \quad (\text{Gl. 2.23})$$

mit:

$z$  [ ] Normierter Wert der Standardnormalverteilung

Im Programm ProFoUnD wird für kumulierte Wahrscheinlichkeiten der Normalverteilung ebenfalls auf intern vorgehaltene Tabellen für die Werte und Wahrscheinlichkeiten der Standardnormalverteilung zurückgegriffen. Die internen Tabellen umfassen insgesamt je 705 Werte für  $z$  und  $F(z)$ , so dass eine exakte Ermittlung der gesuchten Größe sichergestellt ist.

Sind der Fehler ( $x$ -Wert) sowie  $MW$  und  $STABW$  gegeben, wird  $x$  zunächst mittels Gleichung 2.16 normiert. Anschließend wird der zugehörige Wahrscheinlichkeitswert  $F(z) = F(x)$  aus den Tabellen ermittelt. Liegt der  $z$ -Wert zwischen zwei Tabelleneinträgen, so wird linear interpoliert. Liegen umgekehrt die Wahrscheinlichkeit  $F(x) = F(z)$  sowie  $MW$  und  $STABW$  vor, so wird der zugehörige  $z$ -Wert aus den Tabellen ermittelt, der anschließend in den  $x$ -Wert (= Fehler) transformiert wird.

Die Parameter der Normalverteilung MW und STABW liegen dabei zum Einen aus der direkten Ermittlung aus den empirischen Einzelfehlern vor. Zum Anderen werden (falls angefordert) die MW und STABW aus den Polynomen berechnet. Für beide Varianten werden die theoretischen Percentile der Normalverteilung im Bereich [0.05, 0.1, 0.2, ..., 0.9, 0.95] ermittelt.

Hierfür wird  $F(x) = F(z)$  dem gesuchten Percentil gleich gesetzt. Mit diesem  $F(z)$  wird dann wie oben erläutert der zugehörige z-Wert ermittelt. Dieser wird unter Verwendung von MW und STABW in den x-Wert überführt. Dieser x-Wert entspricht dem gesuchten Fehlerwert des Percentils.

Aus dem Vergleich der theoretischen Percentilwerte mit den empirischen Percentilwerten wird deutlich, wie gut die empirische Fehlerverteilung durch die Normalverteilung mit dem jeweiligen MW und der jeweiligen STABW beschrieben wird.

### 2.4.5 Anpassungstests

Wie oben beschrieben ist der Vergleich der Percentilwerte eine Möglichkeit zu überprüfen, wie gut die empirischen Fehlerverteilungen durch die Normalverteilungen beschrieben werden. Die Anpassung kann darüber hinaus auch mit geeigneten Hypothesentests überprüft werden. Die hierfür gebräuchlichsten und am besten geeigneten Testverfahren sind der  $\chi^2$ -Test und der Kolmogoroff-Smirnoff-Test (K-S-Test). Beide Testverfahren sind im Programm ProFoUnD implementiert.

Bei beiden Verfahren werden die folgenden Hypothesen geprüft:

- Nullhypothese (H<sub>0</sub>):

Die empirischen Fehler im Wahrscheinlichkeitsbereich [0.05 ... 0.95] stammen aus einer Grundgesamtheit, die durch eine Normalverteilung mit MW und STABW definiert ist.

- Alternativhypothese (A<sub>1</sub>):

Die empirischen Fehler im Wahrscheinlichkeitsbereich [0.05 ... 0.95] stammen nicht aus der Grundgesamtheit, die durch eine Normalverteilung mit MW und STABW definiert ist.

Zu beachten ist, dass in die Tests lediglich die empirischen Fehler im Wahrscheinlichkeitsbereich [0.05 ... 0.95] eingehen. Wie bei der Bestimmung der Momente wird davon ausgegangen, dass die kleineren und größeren Fehler Ausreißer sind.

Die meisten Hypothesentests werden mit dem Ziel durchgeführt, die Nullhypothese abzulehnen und die Alternativhypothese anzunehmen. Bei den vorliegenden Tests ist es umgekehrt: Das Ziel ist es, H<sub>0</sub> möglichst anzunehmen und A<sub>1</sub> abzulehnen (vgl. Schönwiese, 1992 S. 99ff.).

Bei beiden Testverfahren wird zunächst die jeweilige Teststatistik berechnet (siehe unten). Ausgehend von dieser Teststatistik und den Freiheitsgraden wird die zugehörige Irrtumswahrscheinlichkeit  $\alpha$  [%] berechnet. Die Irrtumswahrscheinlichkeit ist maßgeblich dafür, ob H<sub>0</sub> angenommen oder abgelehnt wird.

Bei klassischen Hypothesentests wird normalerweise zunächst eine maximal zulässige Irrtumswahrscheinlichkeit festgelegt. Diese liegt häufig bei 5% oder 1%. Ist die Irrtumswahrscheinlichkeit größer als 5% oder 1%, so wird die Nullhypothese beibehalten. Ist  $\alpha$  hingegen kleiner als 1%, kann mit einer Irrtumswahrscheinlichkeit unter 1% davon ausgegangen werden, dass die Alternativhypothese zutrifft.

In anderen Worten: Liegt  $\alpha$  über 5% kann davon ausgegangen werden, dass die empirischen Fehler aus der Grundgesamtheit der jeweiligen Normalverteilung stammen. Liegt  $\alpha$  unter 1% muss angenommen werden, dass die empirischen Fehler nicht aus der Grundgesamtheit der jeweiligen Normalverteilung stammen.

Ergänzend ist anzumerken, dass die Ergebnisse beider Testverfahren von der Belegungszahl  $n$  abhängen und je nach Verfahren unterschiedliche Schlüsse resultieren können.

Der  $\chi^2$ -Test wird nur dann berechnet, wenn mindestens 30 Fehlerwerte für die zu untersuchende Vorhersagetiefe vorliegen ( $n \geq 30$ ). Für eine geringere Belegungszahl wird die Anwendung des Tests nicht empfohlen (Haan, 1977 S. 174ff.).

Beim  $\chi^2$ -Test werden die empirischen Fehler und die theoretisch zu erwartenden Fehler Klassen zugeteilt. Für die Teststatistik muss ermittelt werden, wie viele der empirischen und der theoretischen Fehler innerhalb der einzelnen Klassen liegen (Belegungszahlen der Klassen).

Das Ergebnis des Tests kann insbesondere bei kleinen Belegungszahlen merklich von der Auswahl der Klassengrenzen abhängen. Den Empfehlungen von Haan (1977) folgend ist im Programm ProFoUnD daher die Anzahl der Klassen auf 10 festgelegt. Die Klassengrenzen werden dabei programmintern so gewählt, dass auf Grundlage der zu untersuchenden Normalverteilung jeweils 10% der Werte in den einzelnen Klassen liegen sollten.

Die  $\chi^2$ -Teststatistik berechnet sich aus den Differenzen zwischen den empirischen und theoretischen Klassenbelegungszahlen wie folgt:

$$\chi^2 = \sum_{i=1}^{10} \frac{(emp_i - theo_i)^2}{theo_i} \quad (\text{Gl. 2.24})$$

mit:

- $\chi^2$             [ ] Teststatistik
- $emp_i$         [ ] Anzahl der beobachteten empirischen Fehler in der  $i$ -ten Klasse
- $theo_i$         [ ] Anzahl der für die Normalverteilung erwarteten Fehler in der  $i$ -ten Klasse

Die zugehörige Irrtumswahrscheinlichkeit  $\alpha$  wird mithilfe der komplementären Gammafunktion berechnet (Press et al. 1992: S. 615 ff., 1996: S. 1089 ff.). Die Anzahl der Freiheitsgrade beträgt dabei immer 7 (10 Klassen – 2 Parameter der Normalverteilung – 1 = 7; Haan 1977).

Der Kolmogoroff-Smirnoff-Test (K-S-Test) wird nur dann berechnet, wenn mindestens 4 Fehlerwerte für die zu untersuchende Vorhersagetiefe vorliegen ( $n \geq 4$ ). Für eine geringere Belegungszahl wird die Anwendung des Tests nicht empfohlen (Haan, 1977 S. 176ff.).

Beim K-S-Test wird die empirische Summenhäufigkeit der Fehler mit der theoretischen kumulierten Wahrscheinlichkeit verglichen. Die Teststatistik  $d$  ergibt sich aus der maximalen Differenz dieser beiden Kurven:

$$d = \max_{-\infty < x < \infty} |S_N(x) - F(x)| \quad (\text{Gl. 2.25})$$

mit:

- $d$                 [ ] Teststatistik
- $S_N(x)$         [ ] Kumulierte empirische Summenhäufigkeit des Wertes  $x$   
(hier  $x = fehler_i$ ; Gl. 2.16)
- $F(x)$             [ ] Kumulierte Wahrscheinlichkeit des Wertes  $x$  der Normalverteilung  
(hier:  $x = fehler_i$ )

Die zugehörige Irrtumswahrscheinlichkeit  $\alpha$  wird mithilfe einer von Stephens (1970) vorgeschlagenen Berechnungsmethode ermittelt (siehe Press et al. 1992: S. 618 ff., Haan 1977: S. 338).

Beide Tests werden zum Einen für die direkt aus den empirischen Fehlern ermittelten MW und STABW als Parameterwerte der Normalverteilung durchgeführt. Sofern die Berechnung von Polynomen angefordert ist, werden die Tests zum Zweiten auch mit den aus den Polynomen abgeleiteten MW und STABW durchgeführt. Die Anpassung der aus den Polynomen abgeleiteten Normalverteilungen wird nur für den Gültigkeitsbereich des jeweiligen Polynoms getestet.

## 2.5 Kategorische Fehler

Grundlage der kategorischen Gütemaße ist die Frage, ob ein Ereignis eintritt oder nicht und ob dieses Eintreten bzw. Nicht-Eintreten korrekt vorhergesagt wurde. Das Eintreten eines Ereignisses ist dabei entweder durch das Überschreiten (Schalter <Ereignis-Definition> in Steuerdatei auf „Überschreitung“), oder durch das Unterschreiten einer Klassengrenze bzw. eines Schwellenwerts gekennzeichnet (Schalter <Ereignis-Definition> in Steuerdatei auf „Unterschreitung“).

Zur Ermittlung der kategorischen Fehler werden standardmäßig die pegelspezifischen Schwellenwerte verwendet, die für die Einordnung in Bereiche gemäß IKSMS in der Steuerdatei vorgegeben wurden. Eine Ermittlung kategorischer Fehler ist daher nur in Verbindung mit einer der beiden Schwellenwertmethoden der IKSMS möglich.

Alternativ können für jeden Pegel separate Schwellenwerte zur Berechnung der kategorischen Gütemaße definiert werden (Schalter <Schwellenw. kat. Güte:> in der Steuerdatei). Hier kann für jeden Pegel eine beliebige Anzahl an Schwellenwerten vorgegeben werden. Wenn für einen Pegel keine separaten Schwellenwerte angegeben sind, werden, wie oben beschrieben, die Schwellenwerte zur Einteilung in die hydrologischen Fälle für die Berechnung der kategorischen Gütemaße verwendet.

Anhand dieser Klassengrenzen (= Schwellenwerte) werden zunächst die Kontingenztafeln für die einzelnen Pegel, Vorhersagetiefen und Klassengrenzen ermittelt. Die Kontingenztafeln bilden die Grundlage für alle weiteren kategorischen Gütemaße.

Die folgenden Erläuterungen und Beispiele in den Kapiteln 2.5.1 bis 2.5.3 beziehen sich auf die Ereignis-Definition „Überschreitung“. Sofern die Option <Ereignis-Definition> auf „Unterschreitung“ gesetzt ist, sind die entsprechenden Fallunterscheidungen umgekehrt.

### 2.5.1 Standarddefinition der Kontingenztafel

Im einfachen, klassischen Fall werden die einzelnen Vorhersagen ausschließlich anhand der aktuellen Vorhersagewerte und der zugehörigen Messwerte einem der vier Elemente der Kontingenztafel zugeordnet. Für die Option „<Ereignis-Definition> = Überschreitung“ besteht damit jede Kontingenztafel aus den folgenden 4 Elementen:

- a. Hit:  
Eine Überschreitung der Klassengrenze wurde vorhergesagt und gemessen
- b. False alarm:  
Eine Überschreitung der Klassengrenze wurde vorhergesagt aber nicht gemessen
- c. Miss:  
Eine Überschreitung der Klassengrenze wurde nicht vorhergesagt aber gemessen
- d. Correct negative:  
Eine Überschreitung der Klassengrenze wurde nicht vorhergesagt und nicht gemessen

Diese Elemente lassen sich in folgende Matrix einteilen:

	Messung $\geq$ Klassengr.	Messung $<$ Klassengr.
VHS $\geq$ Klassengr.	a	b
VHS $<$ Klassengr.	c	d

Die entsprechende Zuordnung zu den Elementen der Kontingenztafel mit ProFoUnD kann über die HIT-Definition = Standard in der Steuerdatei vereinbart werden.

### 2.5.2 Strenge Definition der Kontingenztafel

Bei der Auswertung von Abfluss- bzw. Wasserstandsvorhersagen kann jedoch auch eine andere, strengere Definition für das Element „hit“ wünschenswert sein. Nach dieser strengeren Definition liegt ein „hit“ nur dann vor, wenn die beiden folgenden Bedingungen zutreffen (vgl. Bild 2.1):

- Vorhersage und Messwert liegen über dem definierten Schwellenwert
- Wert zum Vorhersageursprung bzw. Vorhersagezeitpunkt liegt unter dem Schwellenwert

Vorhersagen, die die erste Bedingung erfüllen, die zweite jedoch nicht, werden als „correct negative“ eingestuft.

In Analogie zu dieser strengeren Definition für einen „hit“ liegen auch ein „miss“ und ein „false alarm“ nur dann vor, wenn zusätzlich zur Standarddefinition gemäß der Matrix in Kapitel 2.5.1 auch der Messwert zum Vorhersageursprung unter dem Schwellenwert liegt (vgl. Bild 2.1).

Gemäß dieser strengeren Definition treten „hit“, „false alarm“ und „miss“ also nur dann auf, wenn der Messwert zum Vorhersagezeitpunkt unter dem jeweiligen Schwellenwert lag. Dementsprechend ist die Häufigkeit von „correct negatives“ bei dieser strengeren Definition wesentlich höher als bei der einfachen klassischen Definition. Dies hat entsprechende Auswirkungen auf die unten aufgeführten kategorischen Gütemaße.

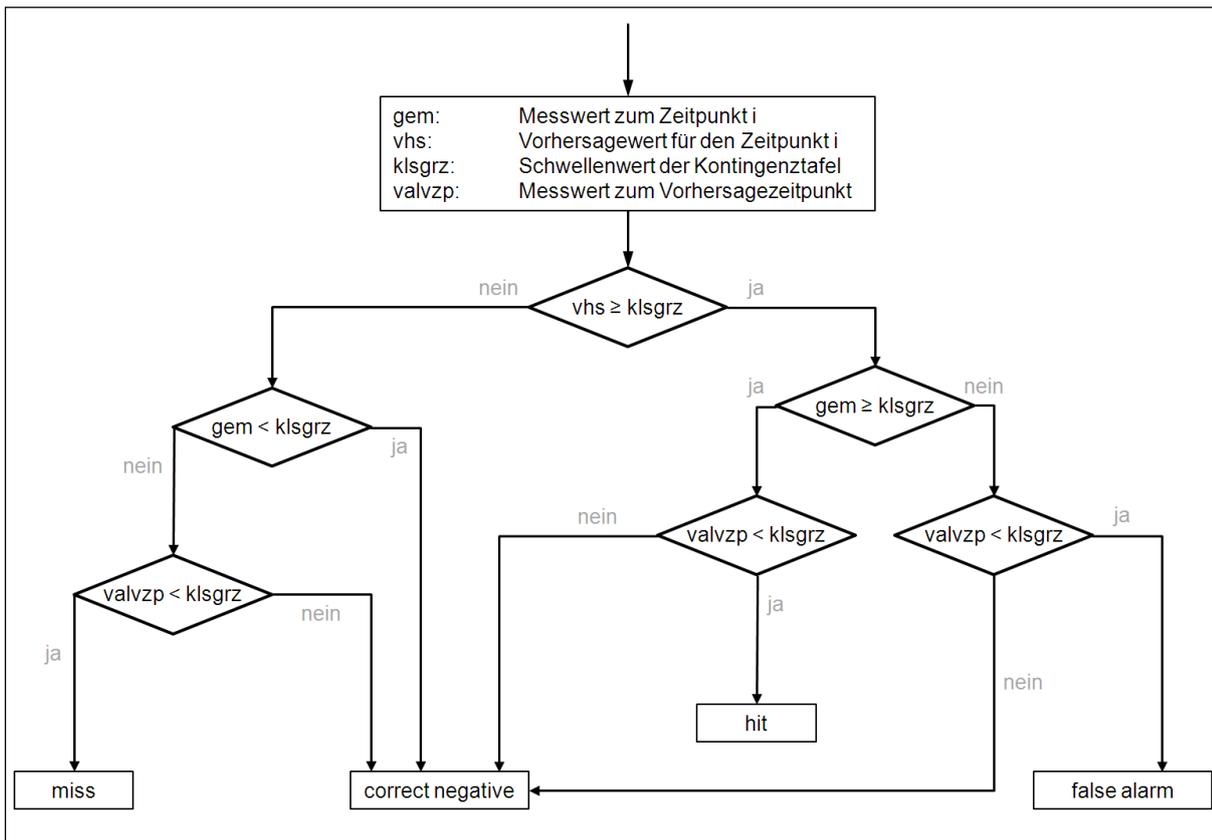


Bild 2.1: Flussdiagramm zur Einordnung der Wertepaare in die Kontingenztafel bei strenger HIT-Definition (ohne tolerierbares Zeitfenster; Ereignis-Definition = Überschreitung).

### 2.5.3 Kontingenztafel bei Berücksichtigung eines tolerierbaren Zeitfensters

Auch bei der Bestimmung der Kontingenztafeln (und den resultierenden kategorischen Gütemaßen) wird das tolerierbare Zeitfenster berücksichtigt, sofern ein solches in der Steuerdatei vereinbart ist. Die Einteilung in das zugehörige Element der Kontingenztafel wird dadurch jedoch sowohl für die Standarddefinition der Kontingenztafel als auch für die strengere Definition etwas komplizierter.

Bei der Standarddefinition der Kontingenztafel müssen neben dem Messwert und dem Vorhersagewert für den jeweiligen Zeitpunkt auch die minimalen und maximalen Werte innerhalb des tolerierbaren Zeitfensters berücksichtigt werden. Bild 2.2 verdeutlicht die resultierende Elemententeilung, die in Absprache mit dem LUWG definiert wurde, anhand eines Flussdiagramms.

Falls der Messwert für den zu untersuchenden Zeitpunkt über der Klassengrenze liegt, erfolgt im zweiten Schritt bereits die Einteilung in „hit“ und „miss“: Sofern der Maximalwert der Vorhersage innerhalb des tolerierbaren Zeitfensters ebenfalls über der Klassengrenze liegt, so wird das Wertepaar als „hit“ gewertet. Andernfalls handelt es sich um einen „miss“.

Liegt der Messwert des zu untersuchenden Zeitpunkts unterhalb der Klassengrenze, so wird das zugehörige Wertepaar entweder als „correct negative“ oder als „false alarm“ gewertet. Diese Zuordnung erfordert jedoch einige zusätzliche Entscheidungen: Sofern der Vorhersagewert für den zu untersuchenden Zeitpunkt ebenfalls kleiner als die Klassengrenze ist, handelt es sich um ein „correct negative“. Andernfalls wird im nächsten Schritt geprüft, ob der Minimalwert der Vorhersage im tolerierbaren Zeitfenster größer als die Klassengrenze ist. Ist dies der Fall, so handelt es sich um einen „false alarm“.

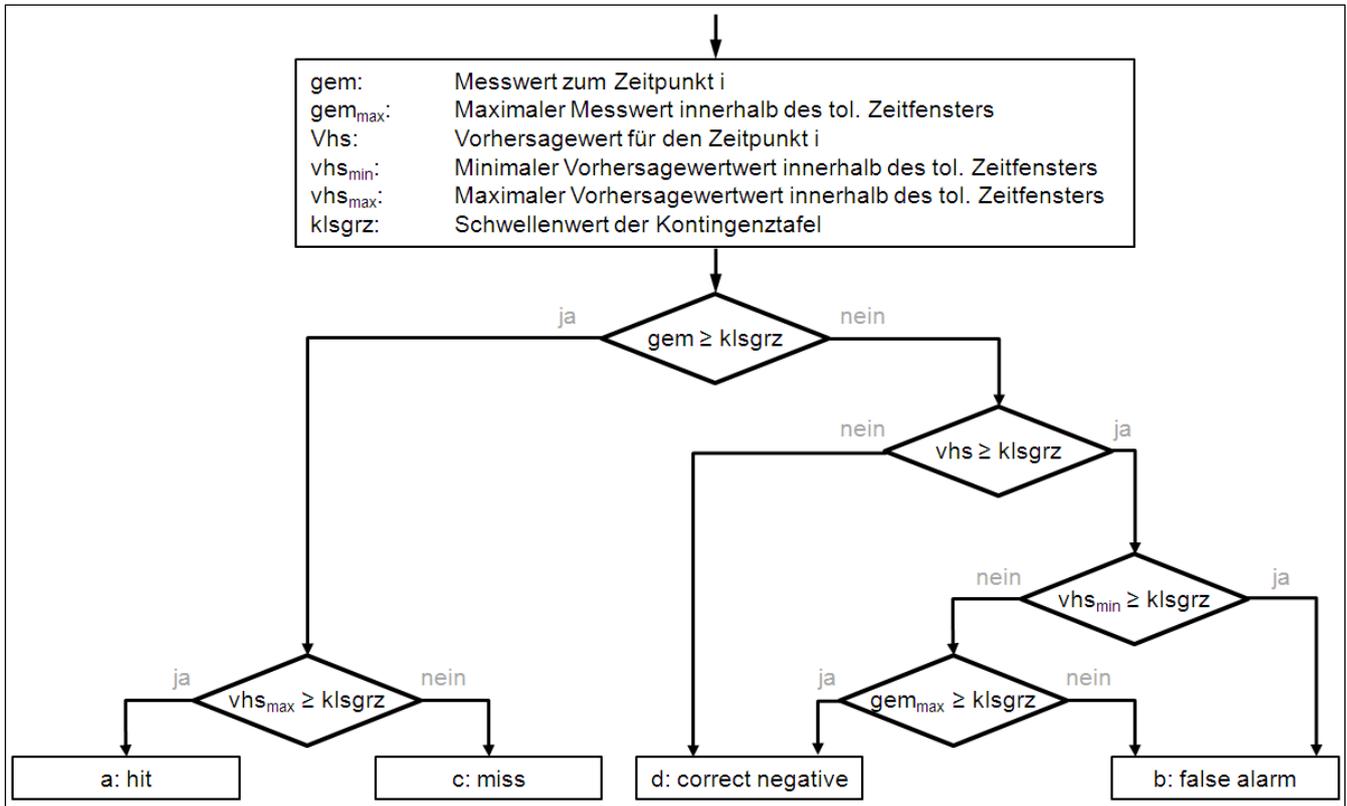


Bild 2.2: Flussdiagramm zur Einordnung der Wertepaare in die Kontingenztabelle gemäß der Standarddefinition bei Vorgabe eines tolerierbaren Zeitfensters (Ereignis-Definition = Überschreitung).

Ist der Minimalwert der Vorhersage im tolerierbaren Zeitfenster jedoch kleiner als die Klassengrenze, wird zusätzlich geprüft, ob der Maximalwert der Messungen im tolerierbaren Zeitfenster über der Klassengrenze liegt. Trifft dies nicht zu, so liegt ebenfalls ein „false alarm“ vor (der Vorhersagewert liegt über der Klassengrenze, obwohl innerhalb des tolerierbaren Zeitfensters der Messwert die Klassengrenze nie überschreitet). Liegt der maximale Messwert im tolerierbaren Zeitfenster über der Klassengrenze, liegt hingegen ein „correct negative“ vor.

Wenn ein tolerierbares Zeitfenster bei der Ermittlung kategorischer Gütemaße zugelassen wird, dann liegt also eine korrekte Ereignisvorhersage (hit) vor, sobald einer der Vorhersagewerte innerhalb des Zeitfensters über dem Schwellenwert liegt. Dies entspricht der grundlegenden Philosophie des tolerierbaren Zeitfensters, wonach nicht der exakte Zeitpunkt des Ereigniseintritts relevant ist, sondern die grundsätzlich korrekte Vorhersage. Bei Vorgabe eines tolerierbaren Zeitfensters ist folglich die Anzahl der hits in der Regel höher als ohne tolerierbares Zeitfenster. Dies gilt in ähnlicher Weise auch für die correct negatives. Folglich erhält man (im Einklang mit der zugrunde gelegten Philosophie) bei Anwendung des tolerierbaren Zeitfensters in der Regel bessere kategorische Gütemaße als ohne tolerierbares Zeitfenster.

Bei der strengen Definition der Kontingenztabelle erfolgt die Zuordnung zu den Elementen der Kontingenztabelle bei Berücksichtigung des tolerierbaren Zeitfensters analog wie für die Standarddefinition beschrieben. Hierbei sind jedoch noch zusätzlich die Messwerte zum Vorhersagezeitpunkt zu berücksichtigen. Hieraus folgt die Zuordnung gemäß dem Flussdiagramm in Bild 2.3.

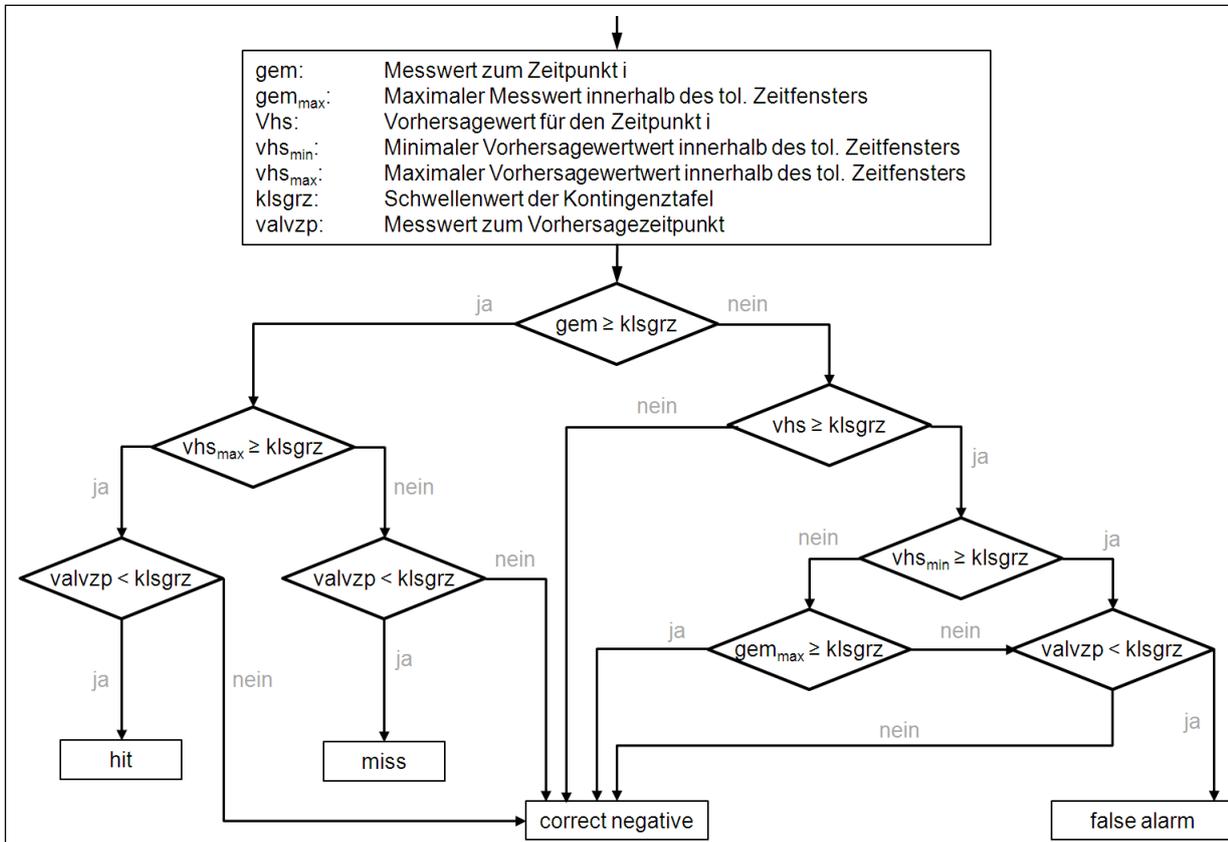


Bild 2.3: Flussdiagramm zur Einordnung der Wertepaare in die Kontingenztabelle gemäß der strengen Definition bei Vorgabe eines tolerierbaren Zeitfensters (Ereignis-Definition = Überschreitung).

### 2.5.4 Kategorische Gütemaße

Die Kontingenztabelle gewährt einen ersten Überblick darüber, ob das Eintreten oder Nicht-Eintreten eines Ereignisses (= Überschreiten der Klassengrenze) zumeist korrekt vorhergesagt wird. Je häufiger die Fälle a und d im Vergleich zu b und c auftreten, desto besser sind die Vorhersagen.

Aus den Kontingenztabelle werden die fünf nachfolgend näher erläuterten kategorischen Gütemaße ermittelt.

Die Probability of detection (Hit Rate) gibt an, welcher Anteil der gemessenen Ereignisse (Überschreitungen der Klassengrenzen) korrekt vorhergesagt wurde:

$$Probability\ Of\ Detection = \frac{a}{a + c}$$

(Gl. 2.26)

Die Probability of detection kann Werte zwischen 0 und 1 annehmen, wobei die optimale Güte durch einen Wert von 1 angezeigt wird. Zu beachten ist, dass bei diesem Gütemaß False alarms (b) nicht berücksichtigt werden. Die Probability of detection erreicht also z.B. bei einer systematischen Überschätzung hohe Gütewerte.

Die False alarm rate (Probability of False Detection) gibt an, welcher Anteil der gemessenen Klassengrenzenunterschreitungen fälschlicherweise als Ereignisse vorhergesagt wurden:

$$FalseAlarm Rate = \frac{b}{b + d}$$

(Gl. 2.27)

Die False alarm rate kann Werte zwischen 0 und 1 annehmen, wobei die optimale Güte durch einen Wert von 0 angezeigt wird. In der Berechnung des Gütemaßes werden Misses (c) nicht berücksichtigt, so dass z.B. bei einer systematischen Unterschätzung hohe Gütewerte (Werte nahe 0) für die False alarm rate erzielt werden können.

Die False alarm ratio gibt an, welcher Anteil der vorhergesagten Ereignisse nicht eintrat:

$$FalseAlarm Ratio = \frac{b}{a + b}$$

(Gl. 2.28)

Die False alarm ratio nimmt Werte zwischen 0 und 1 an, wobei die optimale Güte durch einen Wert von 0 angezeigt wird. Das Gütemaß bildet das Gegenstück zur Probability of detection und sollte gemeinsam mit dieser interpretiert werden. Misses (c) werden nicht berücksichtigt, so dass z.B. bei einer systematischen Unterschätzung hohe Gütewerte (Werte nahe 0) für die False alarm ratio erzielt werden können.

Der Threat score (Critical Success Index) gibt an, wie hoch der Anteil korrekt vorhergesagter Ereignisse an der Summe aller gemessenen und/oder vorhergesagten Ereignisse ist:

$$ThreatScore = \frac{a}{a + b + c}$$

(Gl. 2.29)

Auch der Threat Score nimmt Werte zwischen 0 und 1 an, wobei die optimale Güte durch einen Wert von 1 angezeigt wird. In die Berechnung des Threat Score gehen im Gegensatz zur hier nicht ausgewerteten Accuracy nur jene Elemente ein, bei denen ein Ereignis vorhergesagt und/oder gemessen wurde. Es kann somit als Accuracy unter alleiniger Einbeziehung der „relevanten“ Werte angesehen werden. Bei hohen Klassengrenzen ist mit einer Verschlechterung des Threat Score zu rechnen, da dann kaum noch zufällige Hits zu erwarten sind.

Der Frequency bias (Bias) zeigt, ob die Häufigkeit des Eintretens von Ereignissen korrekt vorhergesagt wird:

$$FrequencyBias = \frac{a + b}{a + c}$$

(Gl. 2.30)

Der Frequency bias kann Werte zwischen 0 und  $\infty$  annehmen, wobei die optimale Güte durch einen Wert von 1 angezeigt wird. Werte kleiner 1 stehen für eine Unterschätzung der Ereignishäufigkeit. Werte größer 1 deuten auf eine Überschätzung der Ereignishäufigkeit hin.

## 2.6 Auswertung maximaler Werte für Vorhersagebereiche

Alle im vorangegangenen Abschnitt beschriebenen Methoden beziehen sich zunächst auf die Auswertung konkreter Vorhersagewerte für bestimmte Vorhersagetiefen. Alternativ hierzu ist es (ab Version 4) möglich, auch die Maximalwerte vorzugebender Vorhersagebereiche mit denselben Methoden auszuwerten. Die Auswertung der <Maximalwerte> kann über die Steuerdatei ausgewählt werden (siehe unten).

Diese Form der Auswertung ist z.B. zur Beurteilung der Qualität der Hochwasser-Frühwarnung sehr nützlich. Bei der Frühwarnung ist der Zeitpunkt des Ereignisses von untergeordneter Bedeutung. Hier ist vielmehr entscheidend, dass der Maximalwert eines am nächsten oder übernächsten Tag zu erwartenden Hochwassers in etwa korrekt vorhergesagt wird.

Mit der Methode können (analog zu den Vorhersagetiefen) ein Vorhersagebereich oder mehrere Vorhersagebereiche ausgewertet werden. Die ausgewählten Vorhersagebereiche müssen alle gleich lang sein (äquidistant) sowie lückenlos und ohne Überlappung aneinander anschließen. Somit können also beispielsweise folgende Vorhersagebereiche mit einem ProFoUnD-Lauf ausgewertet werden: 1h-24h, 25h-48h, 49h-72h. Es ist aber z.B. nicht möglich, 1h-24h und 13h-24h oder 1h-12h und 25h-36h gleichzeitig auszuwerten.

Um einen eindeutigen Zeitbezug sicherzustellen, bezieht sich der Zeitstempel in dieser Methode immer auf den jeweils letzten Zeitpunkt des ausgewerteten Zeitbereichs. Um dies zu erläutern, wird hier ein Vorhersagebereich von 1h-24h angenommen. Der Maximalwert dieses Vorhersagebereichs (1. bis 24. Stunde) erhält also den Zeitstempel der 24. Stunde. Dasselbe gilt für die Messdaten: Hier wird für jeden Zeitstempel der Maximalwert der vorangegangenen 24 Stunden (inklusive des aktuellen Werts) ermittelt.

Zu beachten ist, dass für die Richtungsdefinition gemäß IKSMS der Vorhersage jeweils der letzte Vorhersagewert des ausgewerteten Zeitbereichs verwendet wird. Als  $Q_{akt}$  wird also der letzte Vorhersagewert des Bereichs verwendet. Ebenso werden  $Q_{med}$  und  $Q_{perz}$  aus den Vorhersagewerten bis zum Ende dieses Zeitbereichs ermittelt.

Hierdurch wird es möglich, die Maximalwerte voranschreitender Zeitbereiche mit denselben Methoden zu vergleichen, wie sie für die Einzelwerte angewandt werden. Diese Auswertemöglichkeit ist bislang jedoch nur für die Bestimmung der hydrologischen Fälle nach IKSMS implementiert. Eine Anpassung von Polynomen an die Momente der Fehlerverteilungen ist bislang nicht vorgesehen. Bei der Auswertung der Maximalwerte von Zeitbereichen können außerdem keine tolerierbaren Zeitfenster zugelassen werden und keine Zeitfehler bestimmt werden.

Alle anderen in den Kapiteln 2.1 bis 2.5 beschriebenen kontinuierlichen und kategorischen Fehlermaße können auch für die Maximalwerte von Vorhersagebereichen ermittelt und ausgegeben werden.

## 2.7 Auswertung minimaler Werte für Vorhersagebereiche

Ab Version 7 ist es auch möglich, die Minimalwerte vorzugebender Vorhersagebereiche auszuwerten. Dies ist insbesondere zur Beurteilung der Qualität der Niedrigwasser-Frühwarnung gedacht. Die Auswertung der Minimalwerte erfolgt analog zur Auswertung der Maximalwerte, d.h. es gelten hier die gleichen Besonderheiten und Beschränkungen, die im vorangegangenen Kapitel 2.6 beschrieben wurden.

Die Auswertung der <Minimalwerte> kann über die Steuerdatei ausgewählt werden (siehe unten).

## 2.8 Automatisierte Fortschreibung bestehender Auswertungen

Ursprünglich wurde ProFoUnD konzipiert, um die Gesamtheit (auswählbarer) Vorhersagen auszuwerten. Ab Version 6 ist es auch möglich, bestehende Auswertungen auf der Grundlage neu verfügbarer Vorhersagen fortzuschreiben. Hierzu werden die Originaldaten (Vorhersagen + Messungen) eines aktuellen Ereignisses und die Ergebnisse der bisherigen Auswertungen (CSV-Dateien) eingelesen. Auf dieser Grundlage werden die statistischen Auswertungen für den gesamten Zeitraum (zurückliegende Ereignisse + aktuelles Ereignis) fortgeführt (Steuerungsdetails siehe Kapitel 3.1.1).

Um die Fortschreibung im Rahmen des operationellen Betriebs weitestmöglich zu automatisieren, kann der auszuwertende Ereigniszeitraum (des aktuellen Ereignisses) alternativ zur Vorgabe in der Steuerdatei auch beim Programmaufruf übergeben werden. Hierfür ist in der Steuerdatei im Block <EREIGNISZEITRAEUME> ein Schalter <Ereignisse aus Aufruf:> vorgesehen (Steuerungsdetails siehe Kapitel 3.1.3).

## 2.9 Auswertung der Scheitelwerte

Ab Version 7 ist es möglich, die Scheitelwerte einzelner Vorhersagen im Vergleich zum Scheitelwert der Messung auszuwerten. Dabei werden nur diejenigen Vorhersagen berücksichtigt, deren Vorhersagezeitpunkt vor dem Auftreten des gemessenen Scheitels liegt.

Die Auswertung der Scheitelwerte ist nur in Verbindung mit der Option <Methode Auswertung> = <Standard> zulässig. Zudem dürfen nicht mehrere Ereignisse in einem Programmaufruf ausgewertet werden. Sofern bei der Eingabe das LILA, GMD-, oder CSV-Format verwendet wird, darf der untersuchte Ereigniszeitraum maximal zwei Wochen betragen. Im WSV-Format sind auch längere Auswertungszeiträume zugelassen. Weiter muss die Intervalllänge der auszuwertenden Vorhersage- und Messdaten eine Stunde betragen.

Ein Beispiel für eine Ausgabedatei der Scheitelwerte findet sich in Kapitel 3.3.

## 3 Hinweise zur Anwendung des Programms

### 3.1 Grundlegende Funktionsweise und Programmsteuerung

Die grundlegende Funktionsweise des Programms ProFoUnD ist in Bild 1.1 schematisch dargestellt.

Die Steuerung aller Berechnungsvorgänge erfolgt über die Steuerdatei <profound.str>. Diese muss im selben Verzeichnis vorgehalten werden, in dem die Berechnung aufgerufen wird. Für alle weiteren Eingabedateien können Pfade in der Steuerdatei definiert werden. Mit der Programmoption <Ereignisse aus Aufruf> ist es zudem möglich, die Zeiträume der auszuwertenden Ereignisse (alternativ zur Vorgabe in der Steuerdatei) beim Programmaufruf zu übergeben.

Nachfolgend ist eine exemplarische Steuerdatei für die Programmversion 8.1 abgebildet. Die darin enthaltenen Kommentare erläutern die Funktionsweise der Steuerdatei. In der Steuerdatei ist beispielhaft die Auswertung eines Pegels für ein Ereignis gemäß den Anforderungen der IKSMS (neu) konfiguriert. Als Eingabeformate werden im Beispiel das CSV- und das GMD-Format genutzt. Zusätzliche Zeilen, die im Falle einer anderen Konfiguration (oder die Auswertung von Maximalwerten) erforderlich sind, wurden in der Steuerdatei in auskommentierter Form beibehalten. Ein Beispiel für eine Steuerdatei für das LILA-Format befindet sich in Abschnitt 3.1.1.

Über die Option <Sprache:> bzw. <Langue :> kann vor Angabe des ersten Blocks die Sprache für die Steuerdatei und die Ausgabe wichtiger Log-Meldungen mit <FR> auf Französisch umgestellt werden.

#### Beispielhafte Steuerdatei für das GMD-Format:

```
*****
*
*                               Steuerdatei fuer das Programm                               *
*                               ProFoUnD                                                *
*                               Version 8.1                                              *
*
* HYDRON Ingenieurgesellschaft                                                         *
* Bearbeitung: Ingo Haag, Dirk Aigner                                                 *
* Stand: 09.09.2013                                                                    *
*
*****
*
* Die Steuerdatei ist im Arbeitsverzeichnis vorzuhalten.                             *
*
* Kommentarzeilen sind mit '*' oder einem Blank einzuleiten.                         *
*
* Vor Angabe des ersten Blocks kann die Sprache auf Franzoesisch umgestellt werden.   *
* Dazu ist die Option <Sprache: FR> oder <Langue: FR> zu setzen.                       *
*
* Die Bloecke in der Steuerdatei, die durch Schlüsselwoerter in Grossbuchstaben eingeleitet *
* werden, duerfen in beliebiger Reihenfolge angeordnet werden. Die Eintraege innerhalb eines *
* Blocks sind jedoch in einer festen Reihenfolge vorzugeben und duerfen nicht durch Kommentar- *
* zeilen oder Leerzeilen unterbrochen werden. Die Gross- und Kleinschreibung der Schluessel- *
* woerter ist fest definiert. Die Eintraege nach den Schluesselwoertern erfolgen ab Spalte 26. *
*
*****
* Pfade fuer die Ein- und Ausgabe
* Eingabeformat VHS:      Format der Datei mit den Vorhersagen GMD, CSV oder WSV
* Eingabeformat GEM:      Format der Datei mit den Vorhersagen GMD oder WSV
* Eingabe-Datei VHS:      Pfad und Name der Eingabedatei mit den vorhergesagten Ganglinien
*                          (Alternativ wird auch die alte Bezeichnung "GMD-Eingabe VHS:" akzeptiert)
* Eingabe-Datei GEM:      Pfad und Name der Eingabedatei mit den gemessenen Ganglinien
*                          (Alternativ wird auch die alte Bezeichnung "GMD-Eingabe GEM:" akzeptiert)
* Ausgabe Ganglinien:    Pfad fuer die Ausgabe der Ganglinien unterschiedlicher Vorhersagetiefe
*                          im GMD- und/oder HMZ-Format.
*                          (Alternativ kann auch die alte Kennung <GMD-Ausgabe:> verwendet werden.
* Ausgabe Statistik:      Pfad fuer die Ausgabe der Dateien mit den Ergebnissen CSV-Format
*                          (Guetemasse, Momente, Perzentile...)
*****
PFADE
Eingabeformat VHS:      csv
Eingabeformat GEM:      gmd
Eingabe-Datei VHS:      input\vhs-rgbg-test2.csv
```

```

Eingabe-Datei GEM:      input\qmess-rgbg.gmd
Ausgabe Ganglinien:    output\
Ausgabe Statistik:     output\
*****
* Definitionen fuer die auszuwertenden Pegel
* Datenart              Vorgabe der auszuwertenden Datenart: <Abfluss> oder <Wasserstand>.
*                       Andere Datenarten koennen bislang nicht ausgewertet werden.
* Methode Auswertung:   Auswahl zwischen <Standard> und <Maximalwerte>
*                       Standard:      Auswertung der allgemeinen Vorhersageguete
*                       Maximalwerte:  Auswertung von Maximalwerten fuer bestimmte
*                                       Vorhersagezeitraeume (fuer Fruehwarnung)
*                       Minimalwerte:  Auswertung von Minimalwerten fuer bestimmte
*                                       Vorhersagezeitraeume (fuer Fruehwarnung)
* Methode Schwellenwerte: Hier ist dieselbe Angabe wie in der entsprechenden Zeile unter
*                           HYDROLOGISCHE FAELLE zu machen.
*                           BLfU:      Bei <Anzahl Schwellenwerte> ist 2 anzugeben. Fuer jeden Pegel
*                                       ist dann eine Zeile mit <Schwelle NQM:> und eine Zeile mit
*                                       <Schwelle MQH:> vorzugeben.
*                           IKSMS:     Bei Anzahl Schwellenwerte ist ein Wert zwischen 0 und 5
*                                       anzugeben. Fuer jeden Pegel ist dann eine Zeile <Schwellenwerte>
*                                       mit der zugehoerigen Anzahl an Schwellenwerten anzugeben. Bei 0
*                                       Schwellenwerten wird die Angabe in der Zeile <Schwellenwerte>
*                                       ignoriert.
*                                       Die Zeilen <Schwelle NQM:> und <Schwelle MQH:> fuer jeden Pegel
*                                       entfallen dann.
*                           IKSMS-ALT: Hier gilt bzgl. der Eingaben dasselbe wie bei IKSMS.
*                                       Allerdings wird hier eine andere Methode zur Richtungsbest.
*                                       genutzt (siehe Programmdokumentation).
* Anzahl Schwellenwerte: Im Falle BLfU zwingend 2.
*                           Im Falle IKSMS oder IKSMS-ALT 0 bis 5 (Integer)
* VZP aus Kommentar:     Nein: VZP=1. Wert in VHS-Ganglinien
*                           Ja:        VZP muss in Kommentar angegeben sein
* VHS nur GMD-Kennung:   Nein: Die VHS-Ganglinie wird anhand des Stationsnamens erkannt. Die
*                           Zeile <Bezeichnung VHS:> ist erforderlich. Wenn die VHS-Ganglinie
*                           aus einer CSV-Datei stammt, muss hier NEIN gesetzt werden.
*                           Ja:        VHS-Gangline wird anhand der GMD-Kennung und einer
*                           Sonderkennzeichnung an der Stelle 21-23 des Stationsnamens erkannt
*                           (g+v, s+v oder kor). Die Zeile <Bezeichnung VHS:> darf nicht
*                           enthalten sein.
* GEM nur GMD-Kennung:   Nein: Die Messganglinie wird anhand des Stationsnamens erkannt.
*                           Die Zeile <Bezeichnung GEM:> ist erforderlich.
*                           Ja:        Messgangline wird nur anhand der GMD-Kennung erkannt.
*                           ACHTUNG mehrere Ganglinien mit derselben GMD-Kennung in der
*                           Messdatei sind unzuessaessig. Die Zeile <Bezeichnung GEM:> darf
*                           nicht enthalten sein.
* Fortschreiben:         Nein: Kein Fortschreiben der Guetemasse (Standardeinstellung)
*                           Ja:        Guetemasse werden fortgeschrieben. Naeheres dazu bitte der
*                                       Programmdokumentation entnehmen.
* Anzahl Pegel:          Anzahl der auszuwertenden Pegel (Integer)
*                           (entspricht Anzahl der Wiederholungen der folgenden Eintraege)
* Name Pegel:            Name des Pegels, wie er fuer die Ausgabe der Ganglinien
*                           und Guetemasse fuer
*                           den Dateinamen verwendet wird.
* Bezeichnung VHS:       Bezeichnung der Vorhersage in der GMD-Eingabedatei in Kennkarte 1
*                           (Spalte 5 bis 27) oder Pegelname in der CSV-Datei.
*                           Zeile nur erforderlich, sofern <VHS nur GMD-Kennung:> = <Nein>.
* Bezeichnung GEM:       Bezeichnung der Messung in der GMD-Eingabedatei in Kennkarte 1
*                           (Spalte 5 bis 27). Zeile nur erforderlich, sofern
*                           <GEM nur GMD-Kennung:> = <Nein>.
* GMD-Kennung:           GMD-Kennung fuer die Eingabe und Ausgabe im GMD-Format.
*Im Falle BLfU
* Schwelle NQM:          Schwellenwert fuer den uebergang von NQ zu MQ am jeweiligen
*                           Pegel (Real).
* Schwelle MQH:          Schwellenwert fuer den uebergang von MQ zu HQ am jeweiligen
*                           Pegel (Real).
*Im Falle IKSMS oder IKSMS-ALT
* Schwellenwerte:        Hier sind die Schwellenwerte fuer die Einteilung der Abflussklassen
*                           anzugeben (Real). Es muessen so viele Werte angegeben werden, wie unter
*                           <Anzahl Schwellenwerte> vorgegeben sind. Wenn <Anzahl Schwellenwerte>=0,
*                           wird der Eintrag in dieser Zeile ignoriert. Sofern im Block GUETEMASSE
*                           <Kategorische Guetemasse> = JA, dienen die Schwellenwerte auch als
*                           Klassengrenzen fuer die Kontingenztafeln.
* Schwellenw. kat. Guete: (optionale) Vorgabe von pegelspezifischen Abflussgrenzwerten zur
*                           Berechnung der kategorischen Guetemasse. Es duerfen maximal 10
*                           Schwellenwerte angegeben werden. Wenn die Angabe des Schluesselworts
*                           <Schwellenw. kat. Guete:> fehlt, werden die unter <Schwellenwerte>
*                           angegebenen Grenzwerte fuer die Berechnung der kategorischen
*                           Guetemasse beruecksichtigt.
*****

```

```

PEGEL
Datenart:                Abfluss
Methode Auswertung:     Standard
Methode Schwellenwerte: iksms
Anzahl Schwellenwerte:  1
VZP aus Kommentar:     ja
VHS nur GMD-Kennung:   nein
GEM nur GMD-Kennung:   ja
Fortschreiben:         nein
Anzahl Pegel:          1
Name Pegel:             Regensburg
Bezeichnung VHS:        Regensburg Eiserne B
GMD-Kennung:           RGBG
Schwellenwerte:         40.
Schwellenw. kat. Guete: 30. 60. 120.

```

```

*****
* Definition der auszuwertenden Ereigniszeitraeume
* Intervalllaenge in h:   Intervalllaenge der auszuwertenden Vorhersage- und Messdaten (Real)
* Anzahl Ereignisse:      Anzahl der auszuwertenden Ereignisse (Integer)
*                          (Entspricht Anzahl der Datumsangaben der folgenden Eintraege)
* Ereignisse aus Aufruf:  Ja: Ereignisbeginn und Ereignisende werden als Parameter beim
*                          Programmaufruf im Format <JJJJMMTHH JJJJMMTHH> uebergeben.
*                          Pro auszuwertendem Ereignis ist jeweils der Ereignisbeginn
*                          gefolgt von dem Ereignisende anzugeben.
*                          In diesem Falle werden die Zeilen <Ereignisbeginn> und
*                          <Ereignisende> ignoriert.
*                          Nein: Ereignisbeginn und Ereignisende werden ueber die folgenden
*                          beiden Zeilen <Ereignisbeginn> und <Ereignisende> angegeben.
* Ereignisbeginn:         Ereignisbeginn im Format TT MM JJJJ HH
* Ereignisende:           Ereignisende im Format TT MM JJJJ HH
*****

```

```

EREIGNISZEITRAEUME
Intervalllaenge in h:    1.
Anzahl Ereignisse:       1
Ereignisse aus Aufruf:   nein
Ereignisbeginn:         20 01 2012 00
Ereignisende:           24 01 2012 00
*****

```

```

*****
* Defintion der auszuwertenden Vorhersagedauern
* Fuer Methode Auswertung = Standard, werden die Zeilen <Anzahl VHS-Dauern:>, <VHS-Dauern in h:>,
* <Zeitfenster tolerierbar:> sowie ggf. <Zeitfenster in h:> erwartet.
* Anzahl VHS-Dauern:     Anzahl der auszuwertenden Vorhersagedauern (Integer)
* VHS-Dauern in h:       Angabe der Vorhersagedauern, wobei die Anzahl dem vorangegangenen
*                          Eintrag entsprechen muss (freies Format Real)
* Zeitfenster tolerierbar: ja: Bei der Ausw. wird eine von der VHS-Tiefe abhangige tolerierbare
*                          zeitliche Abweichung berucksichtigt. Der fuer die jeweilige
*                          VHS-Tiefe tolerierbare Zeitversatz ist in der Folgezeile (
*                          Zeitfenster in h:) anzugeben.
*                          nein: Bei der Auswertung wird kein zeitlicher Versatz zugelassen.
*                          Die Folgezeile entfallt.
* Zeitfenster in h:      Zeile ist nur zulassig, wenn bei "Zeifenster tolerierbar" = ja. Fur
*                          jede auszuwertende VHS-Tiefe ist ein zugehoriger Wert fuer das
*                          tolerierbare Zeitfenster anzugeben (freies Format Real)
* Fuer Methode Auswertung = Maximalwerte oder Minimalwerte werden die Zeilen <Anzahl VHS-Bereiche:>
* und <VHS-Bereiche in h:> erwartet
* Anzahl VHS-Bereiche:   Anzahl der auszuwertenden Vorhersagebereiche (Integer)
* VHS-Bereiche in h:     Angabe der Vorhersagebereiche von-bis (z.B. 1.-24.). Mehrere Bereiche
*                          sind durch Kommata zu trennen.
*                          Die VHS-Bereiche mussen luckenlos und aquidistant sein.
*****

```

```

VORHERSAGEDAUERN
Anzahl VHS-Dauern:       2
VHS-Dauern in h:         5. 10.
Zeitfenster tolerierbar: nein
*Zeitfenster in h:       2. 4.
*Anzahl VHS-Bereiche:    2
*VHS-Bereiche in h:      1.-1., 2.-2.
*****

```

```

*****
* Definition der zu berechnenden Guetemasse
* Ausgabe GMD-Format:    Zur Ausgabe vergleichender Ganglinien von Messung und Vorhersagen im
*                          GMD-Format auf JA setzen, ansonsten NEIN
* Ausgabe HMZ-Format:    Zur Ausgabe vergleichender Ganglinien von Messung und Vorhersagen im
*                          HMZ-Format auf JA setzen, ansonsten NEIN
* Verlauf Einzelfehler:  JA, sofern der zeitliche Verlauf der angeforderten Einzelfehler auch in
*                          der GMD- bzw. HMZ-Datei angezeigt werden soll, ansonsten NEIN. Die
*                          Ausgabe wird nur wirksam, wenn einer der beiden vorangestellten
*                          Schalter auch auf JA steht.
* Verlauf hyd. Fall:     JA, sofern der hyd. Fall zum jeweiligen VHS-Wert in der HMZ-Liste
*                          ebenfalls angezeigt werden soll. Die Ausgabe ist nur bei HMZ-Listen
*****

```

```

*
* Abweichung Scheitel:      möglich. Eine Ausgabe in die GMD-Datei wird nicht unterstützt.
*                           Zur Berechnung der Abweichung der VHS vom Scheitelwert der Messung,
*                           ansonsten NEIN.
*                           Die Option <Abweichung Scheitel> kann nur mit <Methode Auswertung>
*                           = STANDARD verwendet werden.
* Absolute Abweichung:      Zur Berechnung der absoluten Abweichung (Qgem-Qvhs)=JA, ansonsten NEIN
* Prozentuale Abweichung:   Zur Berechnung der prozentualen Abweichung (Qgem-Qvhs)/Qvhs*100=JA,
*                           ansonsten NEIN
* Qgem/Qvhs:                Zur Berechnung des Quotienten Qgem/Qvhs=JA, ansonsten NEIN
* ln(Qgem/Qvhs):           Zur Berechnung des natuerlichen Logarithmus von Qgem/Qvhs=JA,
*                           ansonsten NEIN
* Quadrierte Abweichung:    Zur Berechnung der quadrierten Abweichung (Qgem-Qvhs)^2=JA setzen,
*                           ansonsten NEIN
* Zeitfehler:              Zur Berechnung des Zeitfehlers JA setzen, ansonsten NEIN
* Funktion anpassen:        Zum Anpassen von Funktionen an Verlauf der empirischen Momente über
*                           VHS-Tiefe JA setzen, ansonsten NEIN
* Kategorische Guetemasse:  Kategorische Fehler können nur mit der Schwellenwertmethode nach IKSMS
*                           bzw. IKSMS-ALT berechnet werden.
*                           JA:   Es werden folgende kategorischen Gütemaße ber. und ausgegeben:
*                           Kontingenztafeln, False Alarm Rate, False Alarm Ratio,
*                           Threat Score, Frequency Bias. Als Klassengrenzen für die
*                           Kontingenztafeln werden die Schwellenwerte der IKSMS-
*                           Schwellenwertmethode verwendet (siehe Block PEGEL).
*                           NEIN: Keine Berechnung von kategorischen Guetemaßen.
* Ereignis-Definition:      Die Angabe der Zeile ist nur dann erforderlich, wenn
*                           <Kategorische Guetemasse> = JA. Bei NEIN kann die Zeile entfallen oder
*                           sie wird ueberlesen, falls vorhanden.
*                           Ueberschreitung: Ein Ereignis ist dann gegeben, wenn ein Schwellenwert
*                           ueberschritten wird (Hochwasser-Betrachtung) (Standard)
*                           Unterschreitung: Ein Ereignis ist dann gegeben, wenn ein Schwellenwert
*                           unterschritten wird (Niedrigwasserr-Betrachtung)
* HIT-Definition:          Die Angabe der Zeile ist nur obligatorisch, wenn
*                           <Kategorische Guetemasse> = JA. Bei NEIN kann die Zeile entfallen.
*                           Falls sie trotzdem da steht, wird sie ueberlesen.
*                           STANDARD: Standarddefinition für HIT, FALSE ALARM und MISS
*                           STRENG:   Strengere Def. für HIT, FALSE ALARM und MISS
*****
GUETEMASSE
Ausgabe GMD-Format:      ja
Ausgabe HMZ-Format:      ja
Verlauf Einzelfehler:    ja
Verlauf hyd. Fall:       ja
Abweichung Scheitel:     nein
Absolute Abweichung:     ja
Prozentuale Abweichung:  ja
Qgem/Qvhs:               nein
ln(Qgem/Qvhs):           nein
Quadrierte Abweichung:   nein
Zeitfehler:              nein
Funktion anpassen:        nein
Kategorische Guetemasse: ja
Ereignis-Definition:     ueberschreitung
HIT-Definition:          streng
*****
* Defintion der hydrologischen Faelle
* Methode Schwellenwerte: Hier ist dieselbe Angabe wie in entspr. Zeile unter PEGEL zu machen.
*                           BLfU:  Nachfolgend sind <Anzahl Faelle> sowie fuer die Faelle jeweils
*                           <Anzahl Klassen> und <Klassen> vorzugeben (siehe unten).
*                           IKSMS: Nachfolgend sind <Perzentil Richtungsdef.>,
*                           <Anzahl Zusammenfassung:> und <Klassen Zusammenfassung:>
*                           anzugeben (siehe unten)
*                           IKSMS-ALT: Nachfolgend sind <Anzahl Zusammenfassung:> und
*                           <Klassen Zusammenfassung:> anzugeben (siehe unten)
*Im Falle BLfU:
* Anzahl Faelle:          Anzahl der auszuwertenden Fälle (Integer) (entspricht Anzahl der 2
*                           darauffolgenden Einträge)
* Anzahl Klassen:         Anzahl der Klassen für hydrologischem Fall (Integer)
* Klassen:                Klassen des hydrologischen Falls, Anzahl entspr. vorangegangenen Eintrag
*                           (freies Format Integer, möglich sind 1, 2, 3, 4, 5, 6, 7, 8 und 9)
*Im Falle IKSMS:
* Perzentil Richtungsdef.: Angabe des Perzentils, der für die Richtungsdefinition gemäß IKSMS
*                           verwendet werden soll (Real). Bei IKSMS-ALT nicht anzugeben.
* Anzahl Zusammenfassung: Anzahl der Bereiche, bei denen die Auswertungen nicht nochmals nach der
*                           Richtung der VHS differenziert werden soll
*                           (maximal Anzahl der Schwellenwerte + 1) (Integer)
* Bereiche zusammengef.:   Bereiche die zusammengefasst werden sollen (Integer)
*                           (Bereich 1 alle Werte <= Schwellenwert1...letzter Bereich alle Werte >
*                           letzter Schwellenwert; => Anzahl Bereiche = Anzahl Schwellenwerte + 1).
*                           Bei <Anzahl Zusammenfassung>=0 wird der Eintrag in der Zeile ignoriert.
*****

```

```

HYDROLOGISCHE FAELLE
Methode Schwellenwerte:  iksms
Perzentil Richtungsdef.: 85.0
Anzahl Zusammenfassung:  1
Bereiche zusammengef.    1
*Anzahl Faelle:          4
*Anzahl Klassen:         4
*Klassen:                 1, 2, 4, 5
*Anzahl Klassen:         1
*Klassen:                 9
*Anzahl Klassen:         2
*Klassen:                 3, 6
*Anzahl Klassen:         2
*Klassen:                 7, 8

```

### 3.1.1 Hinweise zur Verwendung des LILA-Formats

#### Konfiguration der Steuerdatei

Für eine Auswertung von Eingabedaten im LILA-Format muss in der Steuerdatei <profound.str> im Block <PFADE> sowohl für <Eingabeformat VHS> als auch <Eingabeformat GEM> <LILA> angegeben werden. Eine Kombination des LILA-Formats mit anderen Eingabeformaten ist nicht möglich.

Um die Steuerdatei (Beispiel siehe unten) im Falle einer Auswertung im LILA-Format zu vereinfachen, dürfen die folgenden Angaben im Block <PEGEL> fehlen:

- VHS nur GMD-Kennung
- GEM nur GMD-Kennung
- Bezeichnung VHS
- Bezeichnung GEM
- GMD-Kennung

Diese Angaben werden nicht benötigt und werden ggf. überlesen, sofern sie bei einer Auswertung im LILA-Format dennoch in der Steuerdatei vorhanden sind.

Ebenso werden im Block <GUETEMASSE> folgende Angaben überlesen und dürfen auch weglassen werden:

- Ausgabe GMD-Format
- Ausgabe HMZ-Format

Anstelle dessen ist die Angabe <Ausgabe LILA-Format> mit ja/nein zu definieren. Eine Ausgabe im GMD- oder HMZ-Format ist bei einer Auswertung im LILA-Format nicht zulässig.

Beispiel einer Steuerdatei für die Auswertung von Daten im LILA-Format:

```

=====
*
*                               Steuerdatei fuer das Programm
*                               ProFoUnD
*                               Version 8.1
*
*                               Beispiel einer Steuerdatei zur Auswertung von Daten im LILA-Format
*
* HYDRON GmbH
* Bearbeitung: Ingo Haag, Dirk Aigner
* Stand: 29.07.2016
*
=====
*
* Die Steuerdatei ist im Arbeitsverzeichnis vorzuhalten.
*
* Kommentarzeilen sind mit '*' oder einem Leerzeichen ' ' einzuleiten.

```

```

* Vor Angabe des ersten Blocks kann die Sprache auf Franzoesisch umgestellt werden.
* Dazu ist die Option <Sprache: FR> oder <Langue: FR> zu setzen.
*
* Die Bloecke in der Steuerdatei, die durch Schluesselwoerter in Grossbuchstaben eingeleitet
* werden, duerfen in beliebiger Reihenfolge angeordnet werden. Die Eintraege innerhalb eines
* Blocks sind jedoch in einer festen Reihenfolge vorzugeben und duerfen nicht durch Kommentar-
* oder Leerzeilen unterbrochen werden. Die Gross- und Kleinschreibung der Schluesselwoerter
* muss beachtet werden. Die Eintraege nach den Schluesselwoertern erfolgen ab Spalte 26.
*
*=====
* PFADE -----
*=====
* Pfade fuer die Ein- und Ausgabe
* Eingabeformat VHS:      Format der Datei mit den Vorhersagen (LILA)
* Eingabeformat GEM:      Format der Datei mit den Vorhersagen (LILA)
* Eingabe-Datei VHS:      Pfad und Name der Eingabedatei mit den vorhergesagten Ganglinien
* Eingabe-Datei GEM:      Pfad und Name der Eingabedatei mit den gemessenen Ganglinien
* Ausgabe Ganglinien:     Pfad fuer die Ausgabe der Ganglinien unterschiedlicher Vorhersagedauern
* Ausgabe Statistik:      Pfad fuer die Ausgabe der Dateien mit den Ergebnissen CSV-Format
*                          (Guetemasse, Momente, Perzentile ...)
*=====
PFADE
Eingabeformat VHS:      LILA
Eingabeformat GEM:      LILA
Eingabe-Datei VHS:      .\input\pegel-qvhs.lila
Eingabe-Datei GEM:      .\input\pegel-qmes.lila
Ausgabe Ganglinien:     .\ergebnis\
Ausgabe Statistik:      .\ergebnis\
*=====
* PEGEL -----
*=====
* Defintionen fuer die auszuwertenden Pegel
* Datenart                Vorgabe der auszuwertenden Datenart: <Abfluss> oder <Wasserstand>.
*                          Andere Datenarten koennen bislang nicht ausgewertet werden.
* Methode Auswertung:     Auswahl zwischen <Standard> und <Maximalwerte>
*                          Standard:      Auswertung der allgemeinen Vorhersageguete
*                          Maximalwerte: Auswertung von Maximalwerten fuer bestimmte
*                          Vorhersagezeitraeume (fuer Fruehwarnung)
* Methode Schwellenwerte: Hier ist dieselbe Angabe wie in der entsprechenden Zeile unter
*                          HYDROLOGISCHE FAELLE zu machen.
*                          BLfU: Bei <Anzahl Schwellenwerte> ist 2 anzugeben. Fuer jeden Pegel
*                          ist dann eine Zeile mit <Schwelle NQM:> und eine Zeile mit
*                          <Schwelle MQH:> vorzugeben.
*                          IKSMS: Bei Anzahl Schwellenwerte ist ein Wert zwischen 0 und 5 anzugeben
*                          Fuer jeden Pegel ist dann eine Zeile <Schwellenwerte> mit der
*                          zugehoerigen Anzahl an Schwellenwerten anzugeben. Bei 0
*                          Schwellenwerten wird die Angabe in der Zeile <Schwellenwerte>
*                          ignoriert.
*                          Die Zeilen <Schwelle NQM:> und <Schwelle MQH:> fuer jeden Pegel
*                          entfallen dann.
*                          IKSMS-ALT: Hier gilt bzgl. der Eingaben dasselbe wie bei IKSMS.
*                          Allerdings wird hier eine andere Methode zur Richtungsbest.
*                          genutzt (siehe Programmdokumentation).
* Anzahl Schwellenwerte:  Im Falle BLfU zwingend 2. Im Falle IKSMS oder IKSMS-ALT 0 bis 5
*                          (Integer)
* VZP aus Kommentar:      NEIN: VZP = Erster Wert in VHS-Ganglinien
*                          JA: VZP muss in Metadaten angegeben sein
* Fortschreiben:          NEIN: Kein Fortschreiben der Guetemasse (Standardeinstellung)
*                          JA: Guetemasse werden fortgeschrieben. Naeheres dazu bitte der
*                          Programmdokumentation entnehmen.
* Anzahl Pegel:           Anzahl der auszuwertenden Pegel (Integer)
*                          (entspricht Anzahl der Wiederholungen der folgenden Eintraege)
* Name Pegel:             Name des Pegels, wie er fuer die Ausgabe der Ganglinien und Guetemasse
*                          fuer den Dateinamen verwendet wird.
*Im Falle BLfU:
* Schwelle NQM:           Schwellenwert fuer den Uebergang von NQ zu MQ am jeweiligen Pegel (Real)
* Schwelle MQH:           Schwellenwert fuer den Uebergang von MQ zu HQ am jeweiligen Pegel (Real)
*Im Falle IKSMS oder IKSMS-ALT:
* Schwellenwerte:        Hier sind die Schwellenwerte fuer die Einteilung der Abflussklassen
*                          anzugeben (Real). Es muessen so viele Werte angegeben werden, wie unter
*                          <Anzahl Schwellenwerte> vorgegeben sind. Wenn <Anzahl Schwellenwerte>=0,
*                          wird der Eintrag in dieser Zeile ignoriert. Sofern im Block GUETEMASSE
*                          <Kategorische Guetemasse> = JA, dienen die Schwellenwerte auch als
*                          Klassengrenzen fuer die Kontingenztafeln.
* Schwellenw. kat. Guete: (optionale) Vorgabe von pegelspezifischen Abflussgrenzwerten zur
*                          Berechnung der kategorischen Guetemasse. Es duerfen maximal 10
*                          Schwellenwerte angegeben werden. Wenn die Angabe des Schluesselworts
*                          <Schwellenw. kat. Guete:> fehlt, werden die unter <Schwellenwerte>
*                          angegebenen Grenzwerte fuer die Berechnung der kategorischen
*                          Guetemasse beruecksichtigt.

```

```

*=====
PEGEL
Datenart:                Abfluss
Methode Auswertung:     Standard
Methode Schwellenwerte: iksms
Anzahl Schwellenwerte:  2
VZP aus Kommentar:     nein
Fortschreiben:         nein
Anzahl Pegel:          2
Name Pegel:             Musterpegel
Schwellenwerte:         10. 50.
Schwellenw. kat. Guete: 10. 25. 50. 75.
Name Pegel:             Paradepegel
Schwellenwerte:         123. 1337.

```

```

*=====
* EREIGNISZEITRAEUME -----
*=====

```

```

* Definition der auszuwertenden Ereigniszeitraeume
* Intervalllaenge in h:  Intervalllaenge der auszuwertenden Vorhersage- und Messdaten (Real)
* Anzahl Ereignisse:      Anzahl der auszuwertenden Ereignisse (Integer)
*                         (entspricht Anzahl der Datumsangaben der folgenden Eintraege)
* Ereignisse aus Aufruf:  JA:  Ereignisbeginn und Ereignisende werden als Parameter beim
*                         Programmaufruf im Format <JJJJMMTHH JJJJMMTHH> uebergeben.
*                         Pro auszuwertendem Ereignis ist jeweils der Ereignisbeginn
*                         gefolgt von dem Ereignisende anzugeben.
*                         In diesem Falle werden die Zeilen <Ereignisbeginn> und
*                         <Ereignisende> ignoriert.
*                         NEIN: Ereignisbeginn und Ereignisende werden ueber die folgenden
*                         beiden Zeilen <Ereignisbeginn> und <Ereignisende> angegeben.
* Ereignisbeginn:        Ereignisbeginn im Format TT MM JJJJJ HH
* Ereignisende:          Ereignisende im Format TT MM JJJJJ HH

```

```

*=====
EREIGNISZEITRAEUME
Intervalllaenge in h:  1.
Anzahl Ereignisse:     1
Ereignisse aus Aufruf: nein
Ereignisbeginn:        01 06 2016 07
Ereignisende:          14 06 2016 07

```

```

*=====
* VORHERSAGEDAUERN -----
*=====

```

```

* Defintion der auszuwertenden Vorhersagedauern
* Fuer Methode Auswertung = Standard, werden die Zeilen <Anzahl VHS-Dauern:>, <VHS-Dauern in h:>,
* <Zeitfenster tolerierbar:> sowie ggf. <Zeitfenster in h:> erwartet.
* Anzahl VHS-Dauern:      Anzahl der auszuwertenden Vorhersagedauern (Integer)
* VHS-Dauern in h:        Angabe der Vorhersagedauern, wobei die Anzahl dem vorangegangenen
*                         Eintrag entsprechen muss (freies Format Real)
* Zeitfenster tolerierbar: JA: Bei der Auswertung wird eine von der VHS-Tiefe abhaengige
*                         tolerierbare zeitliche Abweichung beruecksichtigt. Der fuer die
*                         jeweilige VHS-Tiefe tolerierbare Zeitversatz ist in der Folgezeile
*                         (Zeitfenster in h:) anzugeben.
*                         NEIN: Bei der Auswertung wird kein zeitlicher Versatz zugelassen.
*                         Die Folgezeile entfaellt.
* Zeitfenster in h:       Zeile ist nur zulaessig, wenn bei "Zeifenster tolerierbar" = ja.
*                         Fuer jede auszuwertende VHS-Tiefe ist ein zugehoeriger Wert fuer das
*                         tolerierbare Zeitfenster anzugeben (freies Format Real)
* Fuer Methode Auswertung = Maximalwerte, werden die Zeilen <Anzahl VHS-Bereiche:> und
* <VHS-Bereiche in h:> erwartet
* Anzahl VHS-Bereiche:    Anzahl der auszuwertenden Vorhersagebereiche (Integer)
* VHS-Bereiche in h:      Angabe der Vorhersagebereiche von-bis (z.B. 1.-24.). Mehrere Bereiche
*                         sind durch Kommata zu trennen.
*                         Die VHS-Bereiche muessen lueckenlos und aequidistant sein.

```

```

*=====
VORHERSAGEDAUERN
Anzahl VHS-Dauern:      7
VHS-Dauern in h:        1.  6. 12. 24. 48. 72. 144.
Zeitfenster tolerierbar: nein

```

```

*=====
* GUETEMASSE -----
*=====

```

```

* Defintion der zu berechnenden Guetemasse
* Ausgabe LILA-Format:    Zur Ausgabe vergleichender Ganglinien von Messung und Vorhersagen im
*                         LILA-Format auf JA setzen, ansonsten NEIN
* Verlauf Einzelfehler:  JA, sofern der zeitliche Verlauf der angeforderten Einzelfehler auch in
*                         der LILA-Datei angezeigt werden soll, ansonsten NEIN. Die Ausgabe wird
*                         nur wirksam, wenn einer der beiden vorangestellten Schalter auch auf JA
*                         steht.

```

```

* Verlauf hyd. Fall:      JA, sofern der hyd. Fall zum jeweiligen VHS-Wert in der LILA-Datei
*                          ebenfalls angezeigt werden soll.
* Absolute Abweichung:   Zur Berechnung der absoluten Abweichung (Qgem-Qvhs)=JA, ansonsten NEIN
* Prozentuale Abweichung: Zur Berechnung der prozentualen Abweichung (Qgem-Qvhs)/Qvhs*100=JA,
*                          ansonsten NEIN
* Qgem/Qvhs:            Zur Berechnung des Quotienten Qgem/Qvhs=JA, ansonsten NEIN
* ln(Qgem/Qvhs):        Zur Berechnung des natuerlichen Logarithmus von Qgem/Qvhs=JA,
*                          ansonsten NEIN
* Quadrierte Abweichung: Zur Berechnung der quadrierten Abweichung (Qgem-Qvhs)^2=JA setzen,
*                          ansonsten NEIN
* Zeitfehler:           Zur Berechnung des Zeitfehlers JA setzen, ansonsten NEIN
* Funktion anpassen:    Zum Anpassen von Funktionen an Verlauf der empirischen Momente ueber
*                          VHS-Tiefe JA setzen, ansonsten NEIN
* Kategorische Guetemasse: Kategorische Fehler koennen nur mit der Schwellenwertmethode nach IKSMS
*                          bzw. IKSMS-ALT berechnet werden.
*                          JA: Es werden folgende kategorischen Guetemasse berechnet:
*                              Kontingenztafeln, False Alarm Rate, False Alarm Ratio,
*                              Threat Score, Frequency Bias. Als Klassengrenzen fuer die
*                              Kontingenztafeln werden die Schwellenwerte der IKSMS-
*                              Schwellenwertmethode verwendet (siehe Block PEGEL).
*                          NEIN: Keine Berechnung von kategorischen Guetemassen.
* HIT-Definition:       Die Angabe der Zeile ist nur obligatorisch, wenn
*                          <Kategorische Guetemasse> = JA. Bei NEIN kann die Zeile entfallen.
*                          Falls sie trotzdem da steht, wird sie ueberlesen.
*                          STANDARD: Standarddefinition fuer HIT, FALSE ALARM und MISS
*                          STRENG: Strengere Def. fuer HIT, FALSE ALARM und MISS
*=====

```

#### GUETEMASSE

```

Ausgabe LILA-Format:   ja
Verlauf Einzelfehler:  ja
Verlauf hyd. Fall:     ja
Abweichung Scheitel:   ja
Absolute Abweichung:   ja
Prozentuale Abweichung: ja
Qgem/Qvhs:             nein
ln(Qgem/Qvhs):         nein
Quadrierte Abweichung: nein
Zeitfehler:            nein
Funktion anpassen:     nein
Kategorische Guetemasse: ja
Ereignis-Definition:   ueberschreitung
HIT-Definition:        streng

```

#### HYDROLOGISCHE FAELLE -----

##### \* Defintion der hydrologischen Faelle

```

* Methode Schwellenwerte: Hier ist dieselbe Angabe wie in der entspr. Zeile unter PEGEL zu machen.
*                          BLfU: Nachfolgend sind <Anzahl Faelle> sowie fuer die Faelle jeweils
*                              <Anzahl Klassen> und <Klassen> vorzugeben (siehe unten).
*                          IKSMS: Nachfolgend sind <Perzentil Richtungsdef.>,
*                              <Anzahl Zusammenfassung:> und <Klassen Zusammenfassung:>
*                              anzugeben (siehe unten)
*                          IKSMS-ALT: Nachfolgend sind <Anzahl Zusammenfassung:> und
*                              <Klassen Zusammenfassung:> anzugeben (siehe unten)

```

##### \*Im Falle BLfU:

```

* Anzahl Faelle:         Anzahl der auszuwertenden Faelle (Integer) (entspricht Anzahl der 2
*                          darauffolgenden Eintraege)
* Anzahl Klassen:       Anzahl der Klassen fuer hydrologischem Fall (Integer)
* Klassen:              Klassen des hydrologischen Falls, Anzahl entspr. vorangegangenen Eintrag
*                          (freies Format Integer, moeglich sind 1, 2, 3, 4, 5, 6, 7, 8 und 9)

```

##### \*Im Falle IKSMS:

```

* Perzentil Richtungsdef.: Angabe des Perzentils, der fuer die Richtungsdefinition gemaess IKSMS
*                          verwendet werden soll (Real). Bei IKSMS-ALT nicht anzugeben.
* Anzahl Zusammenfassung: Anzahl der Bereiche, bei denen die Auswertungen nicht nochmals nach der
*                          Richtung der VHS differenziert werden soll
*                          (maximal Anzahl der Schwellenwerte + 1) (Integer)
* Bereiche zusammengef.: Bereiche die zusammengefasst werden sollen (Integer)
*                          (Bereich 1 alle Werte <= Schwellenwert1 ... letzter Bereich alle Werte >
*                          letzter Schwellenwert; => Anzahl Bereiche = Anzahl Schwellenwerte + 1).
* Bei <Anzahl Zusammenfassung>=0 wird der Eintrag in der Zeile ignoriert.
*=====

```

#### HYDROLOGISCHE FAELLE

```

Methode Schwellenwerte:  iksms
Perzentil Richtungsdef.: 90.0
Anzahl Zusammenfassung:  1
Bereiche zusammengef.:   1

```

## Formatspezifische Vorgaben

Es können sowohl Daten im LILA-Blockformat als auch im LILA-Spaltenformat ausgewertet werden. So ist es beispielsweise möglich, gemessene Abflüsse im Spaltenformat und Vorhersagen im Blockformat vorzugeben. Eine Auswertung von Einzeldateien (LILA-Einzeldateiformat) ist nicht möglich.

Folgende Metadaten müssen für eine Auswertung mit ProFoUnD vorhanden sein:

- Station
- Datenart (zugelassen sind Q oder W)
- Dimension
- Zeitintervall
- Datenursprung (\*)
- Ggf. Vorhersagezeitpunkt (\*\*)

\* Es werden nur Zeitreihen mit „vhs“ und „mes“ berücksichtigt (bzw. mit „mes+vhs“, wenn in der Steuerdatei <profound.str> im Block <PEGEL> die Option <VZP aus Kommentar> mit <ja> aktiviert ist).

\*\* Der Vorhersagezeitpunkt muss nur zwingend bei einer Auswertung von zusammenhängenden Zeitreihen von Messung und Vorhersage vorhanden sein.

Bei der Angabe der auszuwertenden Pegel in der Steuerdatei (Block <PEGEL>, Eintrag <Name Pegel>) ist darauf zu achten, dass die dort angegebenen Namen exakt den Stationsnamen in den LILA-Eingabedateien entsprechen (Metadatenart „Station“).

Es wird davon ausgegangen, dass der erste Zeitschritt einer Vorhersage im LILA-Format dem Vorhersagezeitpunkt entspricht (d.h. einem gemessenen bzw. simulierten Wert). Erst der zweite Zeitschritt der Vorhersage im LILA-Format entspricht dem ersten vorhergesagten Wert.

### **3.1.2 Besonderheiten beim Fortschreiben bestehender Auswertungen**

Ist im Block <PEGEL> der Steuerdatei der Schalter <Fortschreiben> auf <ja> gesetzt, werden die vorliegenden Auswertungen bisheriger Ereignisse auf der Grundlage eines aktuellen Ereignisses (neue Daten für Vorhersagen + Messungen) fortgeschrieben. Hierzu werden die Ergebnisse der bisherigen Auswertungen (CSV-Dateien) und die Originaldaten eines aktuellen Ereignisses eingelesen.

Die Fortschreibung der kontinuierlichen Gütemaße basiert auf den CSV-Dateien mit den Rangfolgen der Einzelfehler (eine Datei je Fehler, Pegel und Vorhersagetiefe), in denen die Einzelfehler je hydrologischem Fall in aufsteigender Reihenfolge sortiert sind (<\*\_Rang\_\*.csv>). Entsprechend müssen diese Dateien mit den Ergebnissen aller bisherigen Ereignisse vorgehalten werden. Die weiterführenden statistischen Auswertungen der bisherigen Ereignisse (z.B. mittlere Fehler, Häufigkeitsverteilungen, Momente, Anpassungstests) müssen zur Fortschreibung der kontinuierlichen Gütemaße nicht vorliegen.

Die Fortschreibung der kategorischen Gütemaße basiert auf den CSV-Dateien mit den kategorischen Fehlermaßen der bisherigen Ereignisse (eine Datei je Pegel), die die Kontingenztafeln und die daraus errechneten kategorischen Fehlermaße je Vorhersagetiefe und Schwellenwert enthalten (<\*KategorischeFehler.csv>). Die entsprechenden Dateien müssen somit zum Fortschreiben der kategorischen Fehlermaße vorgehalten werden.

Das Fortschreiben bestehender Auswertungen durch aktuelle Ereignisse ist nur mit gleichbleibenden Optionen in der Steuerdatei möglich. Dies bedeutet, dass z.B. die Methoden der Auswertung und Schwellenwerte sowie die Anzahl der Schwellenwerte unverändert bleiben müssen.

Weiter dürfen die Einstellungen zu den Vorhersagedauern, den Gütemaßen und den hydrologischen Fällen nicht verändert werden. Sofern neue Pegel zum Fortschreiben verwendet werden sollen, muss zuerst eine Auswertung ohne Fortschreiben für diese Pegel erfolgen, sodass entsprechende CSV-Dateien im Ausgabeordner zum zukünftigen Fortschreiben vorliegen.

Die fortzuschreibenden Berechnungsergebnisse (oben beschriebene CSV-Dateien) müssen im Ausgabeverzeichnis vorliegen, das in der Steuerdatei vereinbart ist. Diese Dateien (sowie Dateien mit weiterführenden Auswertungen) werden durch die fortgeschriebenen Auswertungen ersetzt. Daher wird dringend empfohlen, die Ergebnisse der bisherigen Auswertungen bei Verwendung der Option <Fortschreiben:> zu sichern (bspw. durch einen Kopierbefehl in einem Batch-Aufruf). Somit wird der Verlust zurückliegender Auswertungen, beispielsweise durch einen Programmabsturz oder fehlerhafter Einstellungen in der Steuerdatei, verhindert. Weiter wird die Möglichkeit geschaffen, ggf. Neuberechnungen für das aktuelle Ereignis durchzuführen. Dabei ist jedoch darauf zu achten, dass erneut die zuvor gesicherten Ergebnisse der bisherigen Auswertungen als Datengrundlage verwendet werden und nicht die Ergebnisse (CSV-Dateien) der zu wiederholenden Auswertung des aktuellen Ereignisses.

Die zur Visualisierung der Ergebnisse vorgesehenen Ganglinien-Dateien (GMD und/oder HMZ bzw. LILA) werden spezifisch für den gewählten Ereigniszeitraum ausgegeben. Da bei einer Fortschreibung im Normalfall ein spezifischer Ereigniszeitraum angegeben wird, werden die bisherigen Ganglinien-Dateien normalerweise nicht überschrieben. Für das aktuell auszuwertende Ereignis werden zusätzliche Ganglinien-Dateien mit dem Datumstempel des aktuellen Ereignisses erzeugt.

### 3.1.3 Übergabe des Ereigniszeitraums beim Programmaufruf

Alternativ zur Vorgabe in der Steuerdatei auch beim Programmaufruf übergeben werden. Ist im Block <EREIGNISZEITRAEUME> der Steuerdatei der Schalter <Ereignisse aus Aufruf:> auf <ja> gesetzt, so müssen Ereignisbeginn und Ereignisende als Parameter beim Programmaufruf im Format <JJJMMTTHH JJJMMTTHH> übergeben werden.

Sofern die Übergabe des Ereigniszeitraums per Programmaufruf erfolgt, werden die Angaben zu Ereignisbeginn und Ereignisende in der Steuerdatei nicht beachtet (es empfiehlt sich, diese dann auszukommentieren). Es können auch mehrere Ereignisse übergeben werden. In diesem Falle sind pro auszuwertenden Ereignis jeweils Ereignisbeginn gefolgt von Ereignisende anzugeben. Hier ein Beispiel für einen Batch-Aufruf von ProFoUnD mit gesetzter Option <Ereignisse aus Aufruf:> = <ja>:

```
profound.exe 2012010100 2012123123
```

## 3.2 Eingangsdaten

### Formate für die Dateneingabe

Für die in der Steuerdatei angegebenen Pegel und Ereigniszeiträume werden vorhergesagte Ganglinien aus einer Datei (GMD-, CSV- oder LILA-Format) bzw. mehreren Dateien (WSV-Format) und gemessene Ganglinien aus einer weiteren Datei (GMD-, WSV- oder LILA-Format) eingelesen. Pfade und Dateinamen der beiden erforderlichen Eingabedateien werden in der Steuerdatei angegeben.

Vorhersagewerte können neben dem GMD- oder LILA-Format in einem CSV-Format (einfaches Schnittstelleformat zur Datenbank des HND des Bayer. LfU) oder in einem speziellen Listenformat der WSV eingelesen werden. Die Messdaten können im GMD-, WSV- oder LILA-Format vorgegeben werden.

### Besonderheiten im WSV-Format

Bei der Verwendung des WSV-Formats ist bei der Angabe der Ereigniszeiträume zu beachten, dass die Stunden eines Tages in diesem Format von 1 Uhr bis 24 Uhr angegeben werden. ProFoUnD-intern werden die 24 Uhr Zeitpunkte auf 0 Uhr des Folgetags gesetzt. Bei der Angabe von Ereignisbeginn und Ereignisende in der Steuerdatei (oder ggf. beim Programmaufruf) muss unbedingt die von ProFoUnD verwendete Datumsformatierung der Stundenwerte (0-23 Uhr) verwendet werden. Eine Angabe des Ereigniszeitraums in der Formatierung 1-24 Uhr ist nicht zulässig und kann ggf. zu fehlerhaften Auswertungen führen.

Messdaten im WSV-Format dürfen neben den Werten zur vollen Stunde (z.B. 07:00 Uhr) auch Zwischenwerte enthalten, bspw. 07:15 Uhr oder 13:45 Uhr. Diese Zwischenwerte werden nicht berücksichtigt und es werden nur die Werte zur vollen Stunde eingelesen.

Weiter ist bei der Verwendung von Eingangsdaten im WSV-Format darauf zu achten, dass das Zeitintervall der Ganglinien dem Zeitintervall entspricht, das in der Steuerdatei unter „EREIGNISZEITRAEUME“ „Intervalllänge in h:“ angegeben ist.

### Allgemeine Hinweise

Zur gleichzeitigen Auswertung mehrerer Ereignisse muss der Anwender ggf. im Vorfeld die Ergebnisse aus mehreren Dateien in jeweils einer Datei für den vorhergesagten Abfluss und einer Datei für den gemessenen Abfluss zusammenfassen. Möglich ist auch eine Bereitstellung von vorhergesagten und gemessenen Abflüssen in einer einzigen GMD- oder LILA-Daten. In diesem Fall muss dieselbe Datei in der Steuerdatei zweimal angegeben werden.

Für die Bestimmung der Klassen nach Bayer. LfU (siehe Abschnitt 2) muss grundsätzlich ein Messwert zum zugehörigen Vorhersagezeitpunkt der Vorhersagetiefe vorliegen, d.h. der erste Messwert sollte mindestens die Anzahl der Stunden der größten Vorhersagetiefe vor dem Ereignisbeginn liegen. Hierauf ist zu achten, wenn die Ganglinien der Messwerte bereitgestellt werden.

Bei einer Auswertung gemäß IKSMS ist bei Vorgabe der Daten ggf. auf das tolerierbare Zeitfenster sowie den LAG zu achten. Die Messdaten müssen entsprechend bereits für Zeiträume vor dem Beginn der auszuwertenden Ereignisse vorliegen.

Das Programm ist nur zur Auswertung von Abflussganglinien oder Wasserstandganglinien vorgesehen. Sind als auszuwertende Datenart Abflüsse vorgegeben, so werden Wasserstände oder andere Datenarten (z.B. Niederschläge) in den Eingangsdateien überlesen. Dasselbe gilt umgekehrt, wenn Wasserstände als auszuwertende Datenart vorgegeben sind. Die entsprechenden Formatvorschriften für die Kennung der Datenart im GMD- bzw. LILA-Format sind daher in den Eingangsdaten zu beachten.

### Besonderheiten im GMD-Format

Das Programm geht standardmäßig davon aus, dass der erste Wert in den GMD-Vorhersageganglinien dem ersten Vorhersagewert entspricht. (Hingegen wird im LILA-Format der erste Zeitpunkt einer VHS-Ganglinie als Vorhersagezeitpunkt interpretiert, d.h. erst der zweite Zeitschritt der VHS-Ganglinie entspricht dem ersten vorhergesagten Wert.)

Im operationellen Betrieb werden die Vorhersagen jedoch häufig als zusammenhängende Ganglinien mit Messdaten bis zum VZP und Vorhersagedaten ab dem VZP + 1 abgespeichert. Im Block PEGEL der Steuerdatei ist daher der Ja/Nein-Schalter <VZP aus Kommentar> vorgesehen. Ist dieser auf <ja> gesetzt, so geht das Programm davon aus, dass die Ganglinien in der VHS-Datei aus Messungen und Vorhersagen zusammengesetzt sind.

Der Vorhersagezeitpunkt muss dann in einer Kommentarzeile im Kopf des GMD-Blocks enthalten sein. Diese Zeile ist durch den String <Vorhersagezeitpunkt> zu markieren. Der VZP muss in den Spalten 25 bis 40 dieser Kommentarzeile in folgendem Format angegeben sein:

TT.MM.JJJJ hh:mm

Der VZP entspricht dabei dem letzten Messwert. Der nächste Wert im GMD-Block wird als erster Vorhersagewert verwendet (siehe Erläuterung der Steuerdatei).

Wenn der Schalter <VZP aus Kommentar> auf <ja> gesetzt ist, aber in keiner der Kommentarzeilen der einzulesenden Ganglinie der String <Vorhersagezeitpunkt> vorhanden ist, so wird die entsprechende Vorhersageganglinie ignoriert.

Wenn die Schalter <VHS nur GMD-Kennung:> und <GEM nur GMD-Kennung:> auf <NEIN> gesetzt sind, erfolgt die Auswahl der Ganglinien aus den Vorhersage- und den Messwertdateien über die in der Steuerdatei vorzugebenden Stationsnamen (<Bezeichnung GEM:> bzw. <Bezeichnung VHS:>). Die ebenfalls anzugebende GMD-Kennung dient in diesem Fall lediglich als zusätzliche Kontrolle sowie für die Ausgabedateien.

Wird der Schalter <VHS nur GMD-Kennung:> auf <ja> gesetzt, so erfolgt die Auswahl der Vorhersageganglinien anhand der GMD-Kennung sowie einer Sonderkennzeichnung innerhalb des Stationsnamens. Diese Sonderkennzeichnung erlaubt die Differenzierung zwischen Vorhersagen und anderen Ganglinien (z.B. reine Simulationsganglinie). Als Sonderkennzeichnung muss bei der vorhergesagten Ganglinie an der Stelle 21 bis 23 des Stationsnamens die Zeichenfolge <g+v>, <s+v> oder <kor> stehen. Nur wenn eine dieser Zeichenfolgen an der angegebenen Stelle steht und zusätzlich die vorgegebene GMD-Kennung gefunden wird, wird die entsprechende Vorhersageganglinie eingelesen.

Ist der Schalter <VHS nur GMD-Kennung:> auf <ja> gesetzt (die Identifikation der Vorhersageganglinien erfolgt also anhand der GMD-Kennung und der fix vorgegebenen Sonderkennzeichnung), so müssen selbstverständlich keine Stationsnamen für die Vorhersageganglinien angegeben werden. Entsprechend geht das Programm davon aus, dass die Zeilen mit der <Bezeichnung VHS:> nicht in der Steuerdatei enthalten sind.

Wird der Schalter <GEM nur GMD-Kennung:> auf <ja> gesetzt, so erfolgt die Auswahl der gemessenen Ganglinien nur anhand der GMD-Kennung. In diesem Fall ist strengstens darauf zu achten, dass in der Messwertdatei nur Ganglinien für gemessene Werte stehen und keine anderen Ganglinien (z.B. Simulationen oder Vorhersagen) mit derselben GMD-Kennung.

Ist der Schalter <GEM nur GMD-Kennung:> auf <ja> gesetzt (die Identifikation der Vorhersageganglinien erfolgt also nur anhand der GMD-Kennung), so müssen selbstverständlich keine Stationsnamen für die Vorhersageganglinien angegeben werden. Entsprechend geht das Programm davon aus, dass die Zeilen mit der <Bezeichnung GEM:> nicht in der Steuerdatei enthalten sind.

Liegen in der Datei mit den vorhergesagten Ganglinien mehrere Vorhersagen mit demselben VZP vor, so wird generell nur die zuletzt eingelesene verwendet (d.h. die Vorhersageganglinie, die zuletzt in der Datei steht).

Nur wenn es sich bei einer der Vorhersageganglinien um eine korrigierte Vorhersage handelt (Sonderkennzeichnung <kor> an der Stelle 21-23 im Stationsnamen), dann wird diese Vorhersageganglinie prioritär behandelt. In diesem Fall wird immer die korrigierte Vorhersage verwendet, egal ob sie vor oder nach einer zeitgleichen unkorrigierten Vorhersage in der Vorhersagedatei steht.

Liegt keine zusammenhängende Ganglinie für die Messdaten eines Ereignisses vor (sondern mehrere Ganglinien jeweils bis zum Vorhersagezeitpunkt), werden die in mehreren Ganglinien vorliegenden Messdaten zu einer Ganglinie zusammengefügt. Für die Erstellung einer derartigen Ganglinie sind immer die für einen Pegel und Zeitschritt zuletzt eingelesenen Werte maßgeblich.

### 3.3 Ausgaben

Die im Programm durchgeführten Berechnungen und die verwendeten Methoden sind in Abschnitt 2 detailliert beschrieben. Ohne auf diese Berechnungsdetails einzugehen werden hier lediglich die resultierenden Ausgaben kurz erläutert.

In der Log-Datei <profound.log> werden wesentliche Informationen zum Programmablauf sowie ggf. Warnungen oder Fehlermeldungen ausgegeben. Die Log-Datei wird in das Verzeichnis ausgegeben, aus dem das Programm gestartet wird.

Die vorhandenen Vorhersagen und deren Verwendung für die statistische Auswertung werden in den beiden EXCEL-lesbaren Protokolldateien (<protokoll-vorhandene-vhs.csv> und <protokoll-fehlende-wertepaare.csv>) aufgelistet. Diese Protokolldateien werden ebenfalls in dem Verzeichnis abgelegt, aus dem das Programm aufgerufen wird.

Wie Bild 3.1 beispielhaft zeigt, werden in der Datei <protokoll-vorhandene-vhs.csv> alle prinzipiell verfügbaren Vorhersagen unter Angabe des Pegelnamens, der GMD-Kennung (nicht im LILA-Format) und des Vorhersagezeitpunkts aufgeführt. Eine Vorhersage ist dabei prinzipiell vorhanden, wenn mindestens ein Vorhersagewert in der eingelesenen Vorhersageganglinie vorliegt (d.h. mindestens einer der Vorhersagewerte ist kein Fehlwert).

In dieser Datei wird protokolliert, zu welchen VZP Vorhersagen vorliegen.			
Pegel	GMD-Kennung	Datum VZP	
Binzwangen	BINZ	02.01.2007 05:00	
Binzwangen	BINZ	03.01.2007 05:00	
Binzwangen	BINZ	04.01.2007 05:00	
Binzwangen	BINZ	05.01.2007 05:00	
Binzwangen	BINZ	08.01.2007 05:00	
Binzwangen	BINZ	11.01.2007 05:00	
Binzwangen	BINZ	16.01.2007 05:00	

Bild 3.1: Beispielhafter Ausschnitt aus der Datei <protokoll-vorhandene-vhs.csv>

In der Datei <protokoll-fehlende-wertepaare.csv> sind alle Wertepaare aufgelistet, für die die Berechnung klassenbezogener Fehler nicht möglich ist, wenngleich prinzipiell eine Vorhersage vorliegt. Wie Bild 3.2 beispielhaft zeigt, sind in der Auflistung der Pegelname, die GMD-Kennung (nicht im LILA-Format), der Vorhersagezeitpunkt, das Datum des Vorhersagewerts, die Vorhersagetiefe und die Ursache dafür, dass das Wertepaar nicht verwendet werden konnte, angegeben.

Wenn ein Wertepaar nicht für die statistische Auswertung genutzt werden kann, obwohl prinzipiell eine Vorhersage vorliegt, so kann dies folgende Ursachen haben: Zum Vorhersagezeitpunkt liegt kein Messwert vor (bzw. dieser ist Null) (nur bei Klasseneinteilung nach BLfU), zum Zeitpunkt des Vorhersagewerts liegt kein Messwert vor (bzw. dieser ist Null) oder es liegt kein Vorhersagewert vor (z.B. Vorhersage zu kurz oder vorhergesagter Wert beträgt Null).

Zu beachten ist, dass in keiner der beiden Protokolldateien jene Vorhersagezeitpunkte aufgelistet sind, für die in der Vorhersagedatei keine Vorhersageganglinie vorliegt.

Sofern in der Steuerdatei die Option <Fortschreiben> auf <ja> gesetzt ist, werden auch die beiden Protokoll-Dateien fortgeschrieben.

In dieser Datei wird protokolliert, welche Wertepaare nicht fuer die Ermittlung klassenbezogener Fehler verwendet werden konnten, obwohl prinzipiell eine VHS vorliegt. (Vgl. <protokoll-vorhandene-vhs.csv>)

Pegel	GMD-Kennung	Datum VZP	Datum VHS	VHS-Tiefe [h]	Ursache
Binzwangen	BINZ	19.01.2007 10:00	22.01.2007 10:00	72	Kein VHS-Wert
Binzwangen	BINZ	19.01.2007 15:00	22.01.2007 15:00	72	Kein Messwert
Binzwangen	BINZ	25.01.2007 05:00	26.01.2007 05:00	24	Kein Messwert zum VZP

Bild 3.2: Beispielhafter Ausschnitt aus der Datei <protokoll-fehlende-wertepaare.csv>.

Die zusammengesetzten Vorhersageganglinien unterschiedlicher Vorhersagetiefen werden für die ausgewählten Pegel und Ereignisse jeweils zusammen mit der gemessenen Ganglinie in eine Datei im GMD- und/oder HMZ-Format bzw. im LILA-Format geschrieben. Das Verzeichnis, in das diese Dateien ausgegeben werden, wird in der Steuerdatei vorgegeben. Die jeweiligen Dateinamen werden aus Namen des Pegels, Ereignisbeginn und Ereignisende zusammengesetzt. Die vergleichenden Ganglinien des Pegels Kempten für das Ereignis zwischen dem 10.05.1999 00:00 Uhr und dem 28.05.1999 23:00 Uhr werden somit in folgenden Dateien abgelegt:

<Kempten\_1999051000-1999052823.gmd>

<Kempten\_1999051000-1999052823.hmz>

<Kempten\_1999051000-1999052823.lila>

Mithilfe der Dateien wird ein unmittelbarer visueller Vergleich zwischen Messung und Vorhersagen unterschiedlicher Vorhersagetiefe möglich (Bild 3.3).

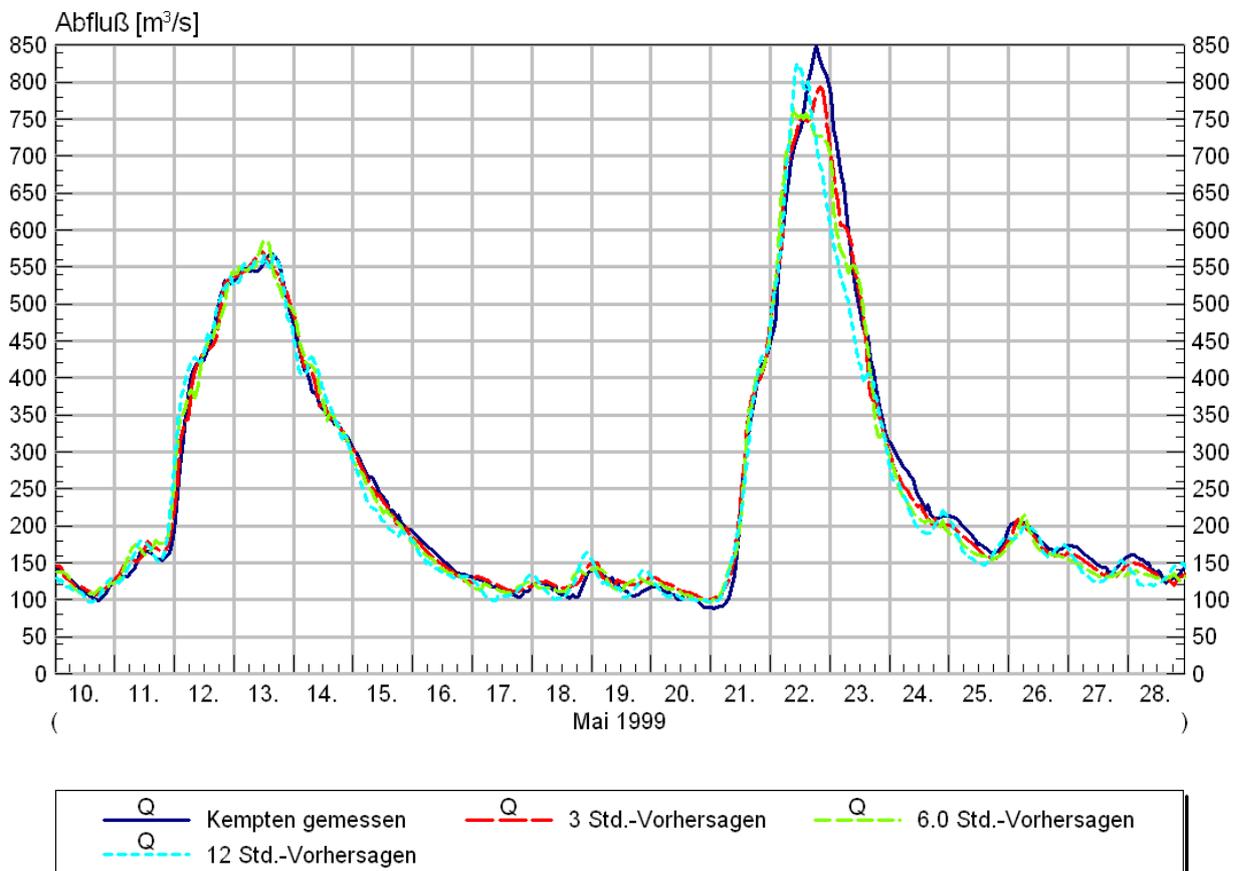


Bild 3.3: Beispielhafte Darstellung einer gemessenen Abflussganglinie mit zusammengesetzten Vorhersagen unterschiedlicher Vorhersagetiefe.

Sofern der Schalter <Verlauf Einzelfehler> auf <ja> gesetzt ist, werden zusätzlich die zeitlichen Verläufe der angeforderten kontinuierlichen Einzelfehler je Vorhersagetiefe in die angeführten Dateien ausgegeben. Somit können diese „Fehlerganglinien“ auf Wunsch mit visualisiert werden (nicht dargestellt). Dabei ist zu beachten, dass die meisten Einzelfehler negative Werte annehmen können. Negative Werte werden in den gängigen Programmen zur Visualisierung des GMD-Formats als Fehlwerte interpretiert. Daher wird für die Ausgabe der Fehlerganglinien das HMZ-Format empfohlen.

Ist der Schalter <Verlauf hyd. Fall> auf <ja> gesetzt, wird zusätzlich der hydrologische Fall des jeweiligen Vorhersagewerts in der HMZ-Liste mit ausgegeben. Hierdurch wird es möglich, die HMZ- bzw. LILA-Dateien unter Berücksichtigung der zusätzlichen Information hinsichtlich Besonderheiten auszuwerten.

Alle anderen Ausgaben erfolgen (wie bei den Protokolldateien) im Excel-lesbaren CSV-Format in ein zweites Ausgabeverzeichnis, das über die Steuerdatei definiert wird.

Die Rangfolgen werden für jedes angeforderte Fehlermaß, jeden Pegel und jede Vorhersagetiefe in eine separate Datei ausgegeben. Die Dateinamen werden aus folgenden Teilen zusammengesetzt: Name des Pegels, Bezeichnung Rang für die Rangfolge, Bezeichnung des Fehlermaßes und Vorhersagetiefe. Die Rangfolgen der prozentualen Abweichung der 1-Stundenvorhersagen am Pegel Kempten werden somit in folgender Datei abgelegt:

<Kempten\_Rang\_ProAbw\_001h.csv>

Wie Bild 3.4 beispielhaft zeigt, enthalten die ausgegebenen Dateien neben den Rangfolgen für die hydrologischen Fälle auch Informationen zum Pegel, der Vorhersagetiefe und die Definition der hydrologischen Fälle.

Sofern ein tolerierbares Zeitfenster vereinbart ist, wird neben dem Rang und dem Fehler in einer dritten Spalte noch der jeweils angesetzte Zeitverschub ausgegeben (nicht dargestellt).

```

Prozentuale Abweichung in Rangfolge fuer die Vorhersagetests fuer den Pegel Kempten mit einer Vorhersagedauer von 1.0 Stunden;
Klassen fuer Fall 0: Alle berechneten Guetewerte, auch wenn keine Klasse zugeordnet werden konnte
Klassen fuer Fall 1: 1 2 4 5;
Klassen fuer Fall 2: 9;
Klassen fuer Fall 3: 3 6;
Klassen fuer Fall 4: 7 8;
  Fall;      0;      Fall;      1;      Fall;      2;      Fall;      3;      Fall;      4;
  Rang;      Guete;   Rang;      Guete;   Rang;      Guete;   Rang;      Guete;   Rang;      Guete;
  1;    -40.538;    1;    -40.53846;    1;    -5.81064;    1;    -5.26031;    1;    -0.51799;
  2;    -33.890;    2;    -33.88951;    2;    -5.03790;    2;    -4.84271;    2;    0.03157;
  3;    -23.340;    3;    -23.33985;    3;    -4.99753;    3;    -2.19154;    3;    0.17942;
  4;    -23.206;    4;    -23.20639;    4;    -4.47317;    4;    -1.41210;    4;    0.63713;
  5;    -21.690;    5;    -21.69014;    5;    -3.66535;    5;    0.08461;    5;    0.69685;
  6;    -20.375;    6;    -20.37519;    6;    -3.65649;    6;    1.79893;    6;    4.49133;

```

Bild 3.4: Beispielhafter Ausschnitt aus einer Datei mit Rangfolgen (ohne tolerierbares Zeitfenster)

Die mittleren Fehler werden in eine Datei je Pegel und Fehlermaß ausgegeben. Die Dateinamen werden aus folgenden Teilen zusammengesetzt: Name des Pegels, Bezeichnung Mittel für die mittleren Fehler und Bezeichnung des Fehlermaßes. Die mittleren Fehler der prozentualen Abweichung am Pegel Kempten werden somit in folgender Datei abgelegt:

<Kempten\_Mittel\_ProAbw.csv>

Hierunter fallen auch die Ausgaben des RMSE und des SkillScorePer (wenn quadrierte Abweichungen angefordert sind) sowie die Mittelwerte der Beträge (bei Abw und Lag).

Wie Bild 3.5 zeigt, sind in den Dateien die mittleren Fehler je Vorhersagetiefe und hydrologischem Fall abgelegt.

```

Prozentuale Abweichung gemittelt fuer die Vorhersagetests fuer den Pegel Binzwangen;
Klassen fuer Fall 0: Alle berechneten Guetewerte, auch wenn keine Klasse zugeordnet werden konnte
Klassen fuer Fall 1: 1 2 4 5;
Klassen fuer Fall 2: 9;
Klassen fuer Fall 3: 3 6;
Klassen fuer Fall 4: 7 8;
Hyd. Fall; Vorhersagedauern in h;
; 1.0; 3.0; 6.0; 12.0; 24.0; 36.0; 48.0;
0; 37.837; 51.980; 68.519; 94.224; 86.716; 193.263; 117.871;
1; 37.837; 51.980; 68.519; 94.468; 86.716; 193.973; 118.205;
2; -9999.000; -9999.000; -9999.000; -9999.000; -9999.000; -9999.000; -9999.000;
3; -9999.000; -9999.000; -9999.000; 46.146; -9999.000; 56.148; 54.187;
4; -9999.000; -9999.000; -9999.000; -9999.000; -9999.000; -9999.000; -9999.000;

```

Bild 3.5: Beispielhafter Ausschnitt aus einer Datei mit mittleren Fehlern.

Die Momente (Mittelwert, Standardabweichung, Schiefe) der empirischen Fehlerverteilungen werden in einer Datei je Pegel und Fehlermaß ausgegeben. Die Dateinamen werden aus folgenden Teilen zusammengesetzt: Name des Pegels, Bezeichnung Momente und Bezeichnung des Fehlermaßes. Die Momente der prozentualen Abweichung am Pegel Kempten werden somit in folgender Datei abgelegt:

<Kempten\_Momente\_ProAbw.csv>

Bild 3.6 zeigt, dass die genannten Momente zusammen mit der jeweiligen Belegungszahl (n) in tabellarischer Form getrennt nach hydrologischen Fällen und Vorhersagetiefen ausgegeben werden. Wenn eines der Momente nicht bestimmt werden konnte, wird dies durch einen Wert von -9999.0 (= Fehlwertkennung) gekennzeichnet. Darüber hinaus enthalten die Dateien Informationen zum Pegel und zur Definition der hydrologischen Fälle.

```

Prozentuale Abweichung: Empirische Momente fuer den Pegel Kempten;
Klassen fuer Fall 0: Alle berechneten Guetewerte, auch wenn keine Klasse zugeordnet werden konnte
Klassen fuer Fall 1: 1 2 4 5;
Klassen fuer Fall 2: 9;
Klassen fuer Fall 3: 3 6;
Klassen fuer Fall 4: 7 8;
; Fall; 0; ; ; Fall; 1; ; ; Fall;
Vorhersagedauer; N; MW; STABW; SCHIEFE; N; MW; STABW; SCHIEFE; N;
1.0; 1346; -1.3954; 3.4790; -0.8245; 1068; -1.8571; 3.9252; -0.6175; 266;
2.0; 1346; -1.1555; 4.0409; -0.5852; 1061; -1.5327; 4.3900; -0.5527; 258;
3.0; 1346; -0.6099; 4.7634; -0.5674; 1055; -0.8175; 5.1290; -0.6718; 251;
4.0; 1346; -0.0581; 5.6375; -0.5352; 1050; -0.0550; 6.0194; -0.7061; 244;
5.0; 1346; 0.3685; 6.6046; -0.4272; 1042; 0.5632; 7.0086; -0.6094; 239;
6.0; 1346; 0.6399; 7.4457; -0.3186; 1038; 0.9359; 7.9277; -0.4949; 232;
8.0; 1346; 0.9811; 8.6419; -0.2632; 1027; 1.3544; 9.2457; -0.4627; 219;
10.0; 1346; 1.1064; 9.3881; -0.2265; 1016; 1.4558; 9.9308; -0.3932; 209;
12.0; 1346; 1.3622; 9.9851; -0.1970; 1002; 1.7915; 10.2551; -0.3258; 193;
18.0; 1346; 1.8885; 11.0462; -0.1040; 976; 2.4546; 10.9829; -0.2184; 168;
24.0; 1345; 1.6584; 11.3921; -0.2027; 953; 1.9727; 11.0903; -0.3596; 139;
36.0; 1346; 1.2343; 13.1384; -0.2663; 911; 0.8146; 12.7940; -0.4982; 80;
48.0; 1345; 1.1058; 13.8750; -0.2644; 876; 0.2161; 13.6856; -0.4752; 41;
60.0; 1346; 0.6476; 14.4520; -0.2247; 852; -0.6447; 13.7465; -0.3223; 12;
72.0; 1345; 0.5636; 14.5550; -0.1293; 839; -0.7553; 13.8669; -0.2287; 0;

```

Bild 3.6: Beispielhafter Ausschnitt aus einer Datei mit Momenten.

Die Polynome zur Beschreibung der Mittelwerte und Standardabweichungen als Funktion der Vorhersagetiefe werden in einer Datei je Pegel und Fehlermaß ausgegeben. Die Dateinamen werden aus folgenden Teilen zusammengesetzt: Name des Pegels, Bezeichnung Polynome und Bezeichnung des Fehlermaßes. Die Polynome zur Beschreibung des Mittelwerts und der Standardabweichung der prozentualen Abweichung am Pegel Kempten werden somit in folgender Datei abgelegt:

<Kempten\_Polynome\_ProAbw.csv>

Bild 3.7 zeigt einen beispielhaften Ausschnitt aus dieser Datei. Diese enthält Angaben zum Fehlermaß, zum Pegel, zur Definition der hydrologischen Fälle sowie weitere Angaben zu den Mindestanforderungen an die Anpassung eines Polynoms.

```

Prozentuale Abweichung: Parameterwerte der Polynome fuer unterschiedliche hydrologische Falle am Pegel Kempten
Klassen fuer Fall 0: Alle berechneten Guetewerte, auch wenn keine Klasse zugeordnet werden konnte
Klassen fuer Fall 1: 1 2 4 5;
Klassen fuer Fall 2: 9;
Klassen fuer Fall 3: 3 6;
Klassen fuer Fall 4: 7 8;

Mindestzahl der Werte je VHS-Dauer;          30;
Mindestzahl verwendbarer VHS-Dauern;         5;

          Fall;          0;
          ;          a0;          b1;          b2;          MAX VHS;Wert (MAX VHS);
Parameter Polynom MW;-0.54882          ;0.12878          ;-.16972E-02;          72.0;          -0.07513;
Parameter Polynom STABW; 4.4773          ;0.39408          ;-.36896E-02;          72.0;          13.72382;
Anzahl verwendbare VHS-Dauern;          15;
Verwendbare VHS-Dauern [h];          1.0          2.0          3.0          4.0          5.0          6.0          8.0          10.0          12.0          18.0          24.0          36.0          48.0          60.0          72.0
Nicht verwendbare VHS-Dauern [h];

          Fall;          1;
          ;          a0;          b1;          b2;          MAX VHS;Wert (MAX VHS);
Parameter Polynom MW;-0.56881          ;0.14860          ;-.23014E-02;          72.0;          -1.79971;
Parameter Polynom STABW; 5.0207          ;0.36605          ;-.35390E-02;          72.0;          13.03033;
Anzahl verwendbare VHS-Dauern;          15;
Verwendbare VHS-Dauern [h];          1.0          2.0          3.0          4.0          5.0          6.0          8.0          10.0          12.0          18.0          24.0          36.0          48.0          60.0          72.0
Nicht verwendbare VHS-Dauern [h];

          Fall;          2;

```

Bild 3.7: Beispielhafter Ausschnitt aus einer Datei mit Polynomen.

Die Parameterwerte  $a_0$ ,  $b_1$  und  $b_2$  der Polynome für die Mittelwerte und Standardabweichungen werden getrennt für hydrologische Fälle ausgegeben. Je nachdem womit eine höhere Genauigkeit erzielt wird, werden die Parameterwerte als Dezimalzahl oder als Exponentialzahl ausgegeben. Falls kein Polynom angepasst werden konnte, werden Fehlwerte von -9999.0 für die Parameterwerte ausgegeben.

In denselben Zeilen wie die Parameterwerte wird mit angegeben bis zu welcher maximalen Vorhersagetiefe (MAX VHS) das Polynom gültig ist und welchen Wert das aus dem Polynom berechnete Moment bei dieser maximalen Vorhersagetiefe hat (Wert(MAX VHS)).

Darüber hinaus wird für jeden hydrologischen Fall angegeben, wie viele und welche Vorhersagetiefen für die Anpassung der Polynome verwendet werden konnten.

Die Percentile der empirischen Fehlerverteilungen sowie der Normalverteilungen aus den direkt bestimmten Momente (MW und STABW) bzw. der aus den Polynomen abgeleiteten Momente, werden in eine Datei je Pegel und Fehlermaß ausgegeben. Die Dateinamen werden aus folgenden Teilen zusammengesetzt: Name des Pegels, Bezeichnung Percentile und Bezeichnung des Fehlermaßes. Die Percentile der prozentualen Abweichung am Pegel Kempten werden somit in folgender Datei abgelegt:

<Kempten\_Percentile\_ProAbw.csv>

Wie Bild 3.8 zeigt, enthalten diese Dateien Angaben zum Fehlermaß, dem Pegel und zur Definition der hydrologischen Fälle. Die Percentilwerte werden tabellarisch ausgegeben, wobei die angeforderten Vorhersagetiefen in Blöcken untereinander stehen. Die hydrologischen Fälle für die jeweilige Vorhersagetiefe werden nebeneinander ausgegeben. Wenn kein Percentil ermittelt werden konnte, wird ein Fehlwert = -9999.0 ausgegeben. Sofern in der Steuerdatei die Anpassung eines Polynoms nicht angefordert wurde, werden in der Spalte POLYNOM folglich nur Fehlwerte ausgegeben. Zusätzlich zu den Percentilwerten werden die jeweiligen Belegungszahlen (n) angegeben.

```

Prozentuale Abweichung: Percentile der Normalverteilung fuer unterschiedliche Vorhersagedauern am Pegel Kempton
Klassen fuer Fall 0: Alle berechneten Guetewerte, auch wenn keine Klasse zugeordnet werden konnte
Klassen fuer Fall 1: 1 2 4 5;
Klassen fuer Fall 2: 9;
Klassen fuer Fall 3: 3 6;
Klassen fuer Fall 4: 7 8;
VHS-Dauer [h]: 1.0;      Fall;      0;      ;      ;      Fall;      1;      ;      ;      Fall;
Percentil;      N;      EMPIRISCH;      NORMVERT;      POLYNOM;      N;      EMPIRISCH;      NORMVERT;      POLYNOM;      N;
0.050;      1346;      -10.7586;      -7.1180;      -8.4285;      1068;      -11.8628;      -8.3136;      -9.2772;      266;
0.100;      1346;      -8.4586;      -5.8540;      -6.6600;      1068;      -9.0810;      -6.8876;      -7.3214;      266;
0.200;      1346;      -5.0483;      -4.3235;      -4.5185;      1068;      -6.5703;      -5.1607;      -4.9532;      266;
0.300;      1346;      -3.0328;      -3.2198;      -2.9744;      1068;      -4.2265;      -3.9156;      -3.2455;      266;
0.400;      1346;      -1.3290;      -2.2768;      -1.6550;      1068;      -2.3270;      -2.8516;      -1.7863;      266;
0.500;      1346;      -0.3346;      -1.3954;      -0.4217;      1068;      -0.6520;      -1.8571;      -0.4225;      266;
0.600;      1346;      0.2982;      -0.5140;      0.8115;      1068;      0.2593;      -0.8627;      0.9413;      266;
0.700;      1346;      0.8680;      0.4290;      2.1309;      1068;      0.9044;      0.2013;      2.4005;      266;
0.800;      1346;      1.6720;      1.5326;      3.6751;      1068;      1.7894;      1.4464;      4.1081;      266;
0.900;      1346;      2.9115;      3.0632;      5.8165;      1068;      3.0588;      3.1733;      6.4764;      266;
0.950;      1346;      4.7893;      4.3272;      7.5850;      1068;      5.2532;      4.5994;      8.4322;      266;
;      ;      ;      ;      ;      ;      ;      ;      ;      ;
VHS-Dauer [h]: 2.0;      Fall;      0;      ;      ;      Fall;      1;      ;      ;      Fall;
Percentil;      N;      EMPIRISCH;      NORMVERT;      POLYNOM;      N;      EMPIRISCH;      NORMVERT;      POLYNOM;      N;
0.050;      1346;      -11.1495;      -7.8023;      -8.9348;      1061;      -11.9184;      -8.7536;      -9.7202;      258;
0.100;      1346;      -8.9655;      -6.3342;      -7.0272;      1061;      -9.8798;      -7.1587;      -7.6352;      258;
0.200;      1346;      -5.3025;      -4.5564;      -4.7172;      1061;      -6.4145;      -5.2274;      -5.1106;      258;
0.300;      1346;      -3.4541;      -3.2746;      -3.0516;      1061;      -4.2188;      -3.8348;      -3.2902;      258;
0.400;      1346;      -1.7363;      -2.1792;      -1.6283;      1061;      -2.4606;      -2.6449;      -1.7347;      258;
0.500;      1346;      -0.4014;      -1.1555;      -0.2980;      1061;      -0.5469;      -1.5327;      -0.2808;      258;
0.600;      1346;      0.7443;      -0.1317;      1.0322;      1061;      0.7865;      -0.4205;      1.1731;      258;
0.700;      1346;      1.7104;      0.9636;      2.4555;      1061;      1.8443;      0.7695;      2.7286;      258;
0.800;      1346;      2.8584;      2.2455;      4.1211;      1061;      2.9278;      2.1621;      4.5490;      258;
0.900;      1346;      4.1658;      4.0232;      6.4311;      1061;      4.0871;      4.0934;      7.0736;      258;

```

Bild 3.8: Beispielhafter Ausschnitt aus einer Datei mit Percentilen.

Die Ergebnisse der Anpassungstests werden in einer Datei je Pegel und Fehlermaß ausgegeben. Die Dateinamen werden aus folgenden Teilen zusammengesetzt: Name des Pegels, Bezeichnung Test für die Anpassungstests und Bezeichnung des Fehlermaßes. Die Ergebnisse der Anpassungstests für die prozentuale Abweichung am Pegel Kempton werden somit in folgender Datei abgelegt:

```
<Kempton_Tests_ProAbw.csv>
```

Wie Bild 3.9 zeigt, enthalten auch diese Dateien zunächst Informationen zum Fehlermaß und Pegel sowie zur Definition der hydrologischen Fälle. Die Ergebnisse der Anpassungstests werden in zwei getrennten Blöcken ausgegeben. Im oberen Block sind die Ergebnisse der Tests dargestellt, die für die Normalverteilungen durchgeführt wurden, bei denen MW und STABW direkt aus den empirischen Fehlern ermittelt wurden. Im unteren Block sind die Ergebnisse der Tests dargestellt, die für die Normalverteilungen durchgeführt wurden, bei denen MW und STABW mit den Polynomen berechnet wurden. Der untere Block wird nur ausgegeben, wenn in der Steuerdatei die Anpassung von Polynomen angefordert ist.

Getrennt für die Vorhersagetiefen und die hydrologischen Fälle werden jeweils die Belegungszahlen (n), die Teststatistik und die zugehörige Irrtumswahrscheinlichkeit  $\alpha$  des  $\chi^2$ -Tests sowie die Teststatistik (D) und die Irrtumswahrscheinlichkeit  $\alpha$  des Kolmogoroff-Smirnoff-Tests (K-S-Test) ausgegeben. Abschnitt 2 enthält weitergehende Hinweise zur Interpretation der Testergebnisse.

Die Kontingenztafeln und kategorischen Gütemaße werden in eine Datei je Pegel ausgegeben. Die Dateinamen ergeben sich aus dem Namen des Pegels und der Bezeichnung KategorischeFehler. Die Kontingenztafeln und kategorischen Fehler für den Pegel Kempton werden somit in folgender Datei abgelegt:

```
<KemptonKategorischeFehler.csv>
```

Wie Bild 3.10 zeigt, werden die Kontingenztafeln und die resultierenden kategorischen Fehler getrennt nach Vorhersagetiefen (Blöcke untereinander) und Schwellenwerten (Spalten) ausgegeben. Zunächst wird die Kontingenztafel ausgegeben, danach folgen die 5 kategorischen Fehlermaße. Neben dem jeweiligen Fehlermaß wird die Anzahl der für seine Ermittlung verwendeten Wertepaare (N) mit angegeben.



Die Scheitel für den Pegel Kempten werden somit in folgender Datei abgelegt:

<Kempten\_ScheitelAbw.csv>

Wie exemplarisch in Bild 3.11 zu erkennen, sind in der Kopfzeile der Scheitelwert der Messung und der zugehörige Zeitpunkt eingetragen. Darunter findet sich eine Kopfzeile, welche die Einträge der darauffolgenden Zeilen für die Scheitelwerte der Vorhersagen beschreibt. Für jede Vorhersage, deren Vorhersagezeitpunkt vor dem Auftreten des gemessenen Scheitels liegt, werden das Datum des Vorhersagezeitpunkts, das Datum des Scheitels der Vorhersage sowie der zugehörige Scheitelwert ausgegeben.

Die Ausgabedateien der Scheitelwerte dienen der Visualisierung der absoluten Abweichungen der Scheitel der Vorhersagen vom Scheitel der Messung mit dem Programm „PlotProfound“.

scheitel gem.;	11.10.2012 19;	753.000;
Datum VZP;	Datum Scheitel;	wert Scheitel;
08.10.2012 06;	11.10.2012 20;	705.000;
09.10.2012 06;	12.10.2012 04;	656.000;
10.10.2012 06;	11.10.2012 17;	737.000;
11.10.2012 06;	11.10.2012 22;	752.000;

Bild 3.11: Beispiel einer Datei mit Scheitelwerten (Wasserstände)

## 4 Literaturverzeichnis

- Bedient, P.B. und Huber, W.C. 1992. *Hydrology and Flood Plain Analysis*. Reading : Addison-Wesley, 2. Auflage, 1992.
- Evans, M., Hastings, N. und Peacock, B. 2000. *Statistical Distributions*. New York : John Wiley, 3. Auflage, 2000.
- Haan, C.T. 1977. *Statistical Methods in Hydrology*. Ames : Iowa State University Press, 1977.
- Helsel, D.R. und Hirsch, R.M. 1992. *Statistical Methods in Water Resources*. Amsterdam : Elsevier, 1992.
- Hirsch, R.M., Helsel, D.R., Cohn, T.A. und Gilroy, E.J. 1993. Statistical Analysis of Hydrologic Data. [Buchverf.] D.R. Maidment. *Handbook of Hydrology*. New York : McGraw-Hill, 1993, S. 18.1 - 18.66.
- Press, W.H., et al. 1992. *Numerical Recipes in Fortran 77: The Art of Scientific Computing*. Cambridge : Cambridge University Press, 1992.
- Press, W.H., et al. 1996. *Numerical Recipes in Fortran 90: The Art of Scientific Computing*. Cambridge : Cambridge University Press, 1996.
- Schönwiese, C.D. 1992. *Praktische Statistik für Meteorologen und Geowissenschaftler*. Berlin : Gebrüder Bornträger, 2. Auflage, 1992.
- Von Storch, H. und Zwiers, F.W. 1999. *Statistical analysis in climate research*. Cambridge : Cambridge University Press, 1999.