

Technischer Bericht 2/2006
November 2006

Leitfaden zur Ermittlung von Messunsicherheiten bei quantitativen Prüfergebnissen

Deutsche Ausgabe des EUROLAB Technical Report 1/2006
„Guide to the Evaluation of Measurement Uncertainty for
Quantitative Test Results“

Vorwort zur Deutschen Ausgabe

Dieses Dokument basiert auf dem *BAM-Leitfaden zur Ermittlung von Messunsicherheiten bei quantitativen Prüfergebnissen* (1. Auflage 2004), der von EUROLAB übernommen und in englischer Fassung als EUROLAB Technical Report 1/2006 – *Guide to the Evaluation of Measurement Uncertainty for Quantitative Test Results* veröffentlicht wurde. Dabei wurde der BAM-Leitfaden einer eingehenden Prüfung im EUROLAB-Ausschuss für Qualitätssicherung im Prüfwesen (Technical Committee for Quality Assurance in Testing – TCQA) unterzogen und im Ergebnis an einigen Stellen präzisiert und ergänzt. In der vorliegenden deutschen Ausgabe sind diese Änderungen berücksichtigt.

Die deutsche Fassung wurde am 9. November 2006 von den Ausschüssen von EUROLAB-D zur Veröffentlichung freigegeben.

Impressum

Deutsche Ausgabe des EUROLAB Technical Report 1/2006
„Guide to the Evaluation of Measurement Uncertainty for Quantitative Test Results“

November 2006

EUROLAB Deutschland
c/o BAM, Unter den Eichen 87
12205 Berlin, Germany
Tel.: +49-30-8104-3769
Fax: +49-30-8104-3717
e-mail: eurolab-d@bam.de

Inhaltsverzeichnis

Vorwort	5
1 Definitionen	7
1.1 Begriffe zur Messunsicherheit	7
1.2 Begriffe zur Prüfgenauigkeit	8
2 Grundlagen	9
2.1 Grundlegende messtechnische Begriffe und Konzepte	9
2.2 Genauigkeit, Richtigkeit und Präzision; das Zielscheibenmodell	12
2.3 Neue Aspekte im "Guide to the Expression of Uncertainty in Measurement"	13
2.3.1 Neue Definition der Messunsicherheit.....	13
2.3.2 Ermittlung von Unsicherheitskomponenten nach Typ A und Typ B	14
2.3.3 Gleichbehandlung aller Unsicherheitskomponenten	15
2.3.4 Erweiterte Unsicherheit.....	15
2.4 Abschätzung von Höchstwerten der Messunsicherheit.....	16
3 Analytisch-rechnerische Ermittlung von Messunsicherheiten	16
3.1 Übersicht.....	16
3.2 Klassifikation von Messunsicherheiten nach Art der Auswertung.....	17
3.3 Allgemeines Verfahren der Unsicherheitsermittlung	18
3.4 Hinweise zur Verwendung von Unsicherheitsbudgets	23
3.5 Abschätzung von Höchstwerten	23
4 Abschätzung von Messunsicherheiten mit laborinternen Validierungsdaten	24
4.1 Allgemeines	24
4.2 Ein-Punkt-Prozedur.....	25
4.2.1 Prüfung der Präzision	25
4.2.2 Prüfung auf systematische Abweichungen.....	26
4.2.3 Behandlung festgestellter systematischer Abweichungen	26
4.3 N-Punkt-Prozedur ($N \geq 2$)	28
4.3.1 Interpolation	28
4.3.2 Ausgleichsrechnung	30
5 Abschätzung von Messunsicherheiten mit Ringversuchsdaten	30
5.1 Validierungsringversuche	30
5.2 Ringversuche zur Eignungsprüfung (Proficiency testing).....	31
5.3 Ringversuche zur Referenzmaterialzertifizierung.....	33
6 Hybridstrategien zur Ermittlung von Messunsicherheiten	33
7 Angabe und Dokumentation von Messunsicherheiten	34
Literatur	34
Anhang	35
A.1 Häufig vorkommende Unsicherheitsquellen	35
A.2 Unsicherheit bei linearer Kalibrierung.....	36
A.2.1 Allgemeines	36
A.2.2 Bestimmung von Achsenabschnitt und Steigung	37
A.2.3 Ermittlung der Unsicherheit von Achsenabschnitt und Steigung.....	38
A.3 Modellierung von Verfahrensschritten durch Wirkungsgrade und Inkremente	41
A.4 Numerische Verfahren zur Unsicherheitsfortpflanzung.....	42
A.4.1 Differenzenrechnung	42
A.4.2 Monte-Carlo-Simulation	43
A.4.3 Software	44
A.5 Die Unsicherheit von Mittelwerten	44
A.5.1 Allgemeines	45
A.5.2 Korrelation innerhalb einer Messreihe.....	46
A.6 Ermittlung von Kovarianzen und Korrelationskoeffizienten	47
A.6.1 Allgemeines	47
A.6.2 Unsicherheitsfortpflanzung	48
A.6.3 Parallelmessungen	48
A.6.4 Korrelationskoeffizienten	48

Vorwort

Auch nach mehr als zehn Jahren seit dem ersten Auftritt der Messunsicherheit bei EUROLAB (wo der GUM 1994 auf dem Eurolab-Symposium in Florenz vorgestellt wurde), wird die Ermittlung der Messunsicherheit für Prüfergebnisse noch immer als ein problematisches Unterfangen betrachtet – nicht grundsätzlich, wohl aber in der täglichen Praxis.

Grundsätzlich wird der als der *GUM* bekannte *Guide to the Expression of Uncertainty in Measurement* (deutsche Fassung: *Leitfaden zur Angabe der Unsicherheit beim Messen*) im gesamten Prüfwesen als Bezugsdokument zum Thema Messunsicherheit anerkannt. Es besteht Einvernehmen darüber, dass der Begriff „Messunsicherheit“ auf alle Arten quantitativer Prüfergebnisse anwendbar ist, und die Prinzipien des GUM werden in vollem Umfang akzeptiert.

Wenn es jedoch konkret um die Ermittlung der der Unsicherheit für die Ergebnisse eines (quantitativen) Prüfverfahrens geht, wird der GUM häufig als nicht anwendbar kritisiert. Dieser Eindruck beruht darauf, dass im GUM fast ausschließlich eine einzige Methode zur Unsicherheitsermittlung behandelt wird: die analytisch-rechnerische Ermittlung auf Grundlage eines umfassenden mathematischen Modells für das Prüfverfahren. Dabei wird jedem Unsicherheitsbeitrag eine entsprechende Eingangsgröße zugeordnet; die Unsicherheitsbeiträge werden einzeln ausgewertet und quadratisch addiert. Diese Vorgehensweise wird daher oft als „die GUM-Methode schlechthin“ (miss)verstanden. Tatsächlich lassen die Prinzipien des GUM eine Vielzahl von Vorgehensweisen zu, aber diese Tatsache blieb lange Zeit unter einer Flut von Veröffentlichungen und Vorträgen verborgen, in denen die analytisch-rechnerische Methode als neues Paradigma der Qualitätssicherung von Messungen gefeiert wurde. Erst in letzter Zeit fanden alternative „empirische Methoden“ größere Beachtung. Sie beruhen auf Untersuchungen der Ergebnisqualität des *gesamten* Prüfverfahrens, die so gestaltet werden, dass die Einflüsse möglichst vieler Unsicherheitsquellen auf einmal erfasst werden. Bei dieser Vorgehensweise werden typischerweise Kennwerte für die Präzision und die Richtigkeit des Prüfverfahrens verwendet, die aus laborinternen Untersuchungen (Validierung, Qualitätskontrolle) oder aus Ringversuchen (Validierung von Standardverfahren, Eignungsprüfung) stammen. Sofern dabei die Prinzipien des GUM beachtet werden, sind solche Methoden in vollem Umfang mit dem GUM vereinbar.

Eurolab hat sich stets für die Nutzung empirischer Methoden zur Unsicherheitsermittlung als einer validen und häufig praktikableren Alternative zur analytisch-rechnerischen Methode eingesetzt, u.a. durch Publikation von Technical Reports zum Thema Messunsicherheit bei Prüfungen. Der erste in dieser Reihe (No. 1/2002) war eine Einführung in die Thematik für Neueinsteiger. Sie wird nun mit diesem Dokument durch einen umfassenden technischen Leitfaden fortgesetzt – gedacht für Anwender, die bereits Erfahrungen mit der Thematik gemacht haben. Dieser Leitfaden behandelt sowohl die analytisch-rechnerische („bottom-up“) Methode, die auf ein vollständiges mathematisches Modell des Prüfverfahrens aufsetzt, als auch empirische („top-down“) Methoden, die mit ganzheitlichen Verfahrenskenngrößen (Präzision, Richtigkeit) operieren. Ein weiterer Eurolab Technical Report, der von einer zu diesem Zweck eingesetzten Expertengruppe erarbeitet wird, hat den Ergebnisvergleich und die Kombination der gegenwärtig verfügbaren Methoden zur Unsicherheitsermittlung zum Thema. Dieser Bericht wird eine Beispielsammlung aus verschiedenen Prüfgebieten enthalten, wobei die mit verschiedenen Methoden erhaltenen Ergebnisse miteinander verglichen und die Schlussfolgerungen aus diesem Vergleich diskutiert werden.

Leitfaden zur Ermittlung von Messunsicherheiten bei quantitativen Prüfergebnissen

Dieses Dokument gibt technische Anleitungen zur Ermittlung von Messunsicherheiten bei quantitativen Prüfergebnissen. Die in diesem Leitfaden beschriebenen Methoden entsprechen in vollem Umfang den Grundsätzen des *Guide to the Expression of Uncertainty in Measurement* (GUM). Sie beschränken sich jedoch nicht auf die im GUM beschriebene analytisch-rechnerische („bottom-up“) Methode, die auf ein vollständiges mathematisches Modell des Messverfahrens aufsetzt. Als Alternativen werden empirische („top-down“) Methoden beschrieben, die mit ganzheitlichen Verfahrenskenngrößen aus Ringversuchen (Validierungsringversuche, Eignungsprüfungen) oder mit Daten aus der laborinternen Validierung und Qualitätskontrolle (Präzision, Richtigkeit) operieren. Ergänzende Informationen zu häufig vorkommenden Unsicherheitsquellen und häufig vorkommenden Datenauswertungsverfahren für die Unsicherheitsermittlung sind in einem Anhang zusammengestellt.

1 Definitionen

In diesem Leitfaden werden die Begriffe „quantitative Prüfung“ und „Messung“ synonym verwendet. Dabei wird, ebenso wie in den einschlägigen Normen, vorwiegend von Messungen und entsprechend z.B. von Messgröße, Messobjekt, Messergebnis und Messunsicherheit gesprochen. Ohne den gedanklichen Inhalt zu ändern, könnte man diese Begriffe durch Prüfung, Prüfgröße, Prüfobjekt, Prüfergebnis und Ergebnisunsicherheit ersetzen.

In diesem Dokument wird durchgängig der Begriff „Messverfahren“ anstelle des häufig verwendeten Begriffs „Messmethode“ in folgendem Sinn verwendet: Messverfahren steht für eine detailliert festgelegte Prozedur (Verfahrensanweisung), die für festgelegte Messobjekte und festgelegte Messbedingungen entwickelt und validiert wurde. Nur einem wohldefinierten Messverfahren lässt sich eine Messunsicherheit für die Ergebnisse bestimmungsgemäß durchgeführter Messungen zuordnen.

1.1 Begriffe zur Messunsicherheit

Das Ziel einer Messung (oder einer sonstigen quantitativen Untersuchung) besteht darin, einen Schätzwert für den wahren Wert der betreffenden Messgröße zu bestimmen. Dieser Schätzwert, das Messergebnis, kann ein einzelner Messwert sein. Im Allgemeinen wird das Messergebnis jedoch aus einer Reihe von Messwerten mittels statistischer Auswertungsverfahren gewonnen, z.B. als Mittelwert. Für jedes Messverfahren muss die Ergebnisgröße eindeutig definiert und das Auswertungsverfahren festgelegt sein.

Die Verwendung von Messergebnissen erfordert in der Regel die Kenntnis der Genauigkeit, d.h. das Ausmaß der möglichen Abweichung des Messergebnisses vom wahren Wert der Messgröße muss bekannt sein. In der Messtechnik wird als quantitatives Genauigkeitsmaß die "Messunsicherheit" verwendet. Dieser Begriff wird auch für die Unsicherheit quantitativer Prüfergebnisse verwendet.

Im Folgenden werden drei Definitionen aus grundlegenden Begriffsnormen (auszugsweise) wiedergegeben, die unterschiedliche Aspekte der Unsicherheit hervorheben, jedoch inhaltlich im Wesentlichen übereinstimmen.

Messunsicherheit

Dem Messergebnis zugeordneter Parameter, der die Streuung der Werte kennzeichnet, die vernünftigerweise der Messgröße zugeordnet werden könnten.

(Quelle: Internationales Wörterbuch der Metrologie)

Messunsicherheit

Aus Messungen gewonnener Kennwert, der zusammen mit dem Messergebnis zur Kennzeichnung eines Wertebereichs für den wahren Wert der Messgröße dient.

(Quelle: DIN 1319-1)

Ergebnisunsicherheit

Geschätzter Betrag zur Kennzeichnung eines Wertebereichs, innerhalb dessen der Bezugswert liegt, wobei dieser je nach Festlegung oder Vereinbarung der wahre Wert oder der Erwartungswert sein kann.

(Quelle: DIN 55350-13)

Die nachfolgenden Begriffe bilden das Begriffssystem des "Guide to the Expression of Uncertainty in Measurement" (GUM). Sie werden, ebenso wie die im GUM festgelegten Formelzeichen, auch für die nachfolgenden Abschnitte dieses Leitfadens zugrundegelegt.

Standardunsicherheit (u)

Als Standardabweichung ausgedrückte Unsicherheit des Ergebnisses einer Messung.

Kombinierte Standardunsicherheit (u)

Standardunsicherheit eines Messergebnisses, wenn dieses Ergebnis aus den Werten einer Anzahl anderer Größen gewonnen wird. Sie ist gleich der positiven Quadratwurzel einer Summe von Gliedern, wobei die Glieder Varianzen oder Kovarianzen dieser anderen Größen sind, gewichtet danach, wie das Messergebnis mit Änderungen dieser Größen variiert.

Anmerkung: Im GUM werden kombinierte Standardunsicherheiten durch einen Index mit u_c gekennzeichnet. Diese Kennzeichnung wird in dem Leitfaden nicht übernommen, da in der Prüftechnik die Unterscheidung zwischen kombinierten und nicht kombinierten Standardunsicherheiten nicht praxisrelevant ist.

Erweiterte Messunsicherheit (U)

Kennwert, der einen Bereich um das Messergebnis kennzeichnet, von dem erwartet werden kann, dass er einen großen Anteil der Verteilung der Werte umfasst, die der Messgröße vernünftigerweise zugeordnet werden könnten.

Erweiterungsfaktor (k)

Zahlenfaktor, mit dem die (ggf. kombinierte) Standardunsicherheit multipliziert wird, um eine erweiterte Messunsicherheit zu erhalten.

1.2 Begriffe zur Prüfgenauigkeit

Bei den Begriffen des vorigen Abschnitts handelt es sich überwiegend um Neuschöpfungen aus dem Bereich des Messwesens, denen im Bereich des Prüfwesens und der chemischen Analytik ein allgemein akzeptiertes Begriffssystem gegenübersteht. Da dieses Begriffssystem dort auch weiterhin Verwendung findet, werden in diesem Abschnitt die Hauptbegriffe aus der grundlegenden Begriffsnorm ISO 3534-1 (siehe auch DIN 55350-13), sinngemäß ins Deutsche übertragen, zusammengestellt. Auf das Verhältnis dieser beiden Begriffssysteme wird in Abschnitt 2 eingegangen.

Genauigkeit

Ausmaß der Übereinstimmung eines Messergebnisses mit einem anerkannten Bezugswert der Messgröße.

Richtigkeit

Ausmaß der Übereinstimmung des Mittelwertes der Ergebnisse einer großen Anzahl unabhängiger Messungen mit einem anerkannten Bezugswert der Messgröße.

Präzision

Ausmaß der Übereinstimmung zwischen den Ergebnissen unabhängiger Messungen bei vorgegebenen Bedingungen.

2 Grundlagen

2.1 Grundlegende messtechnische Begriffe und Konzepte

Die fett gedruckten Begriffe sind in den einschlägigen Normen definiert. Sofern nicht anders vermerkt, werden die Begriffe in Anlehnung an das Internationale Wörterbuch der Metrologie (VIM), 2. Auflage, 1994 erläutert. Ebenso wie in den einschlägigen Normen wird in diesem Abschnitt ausschließlich von Messungen und entsprechend von Messgröße, Messergebnis und Messunsicherheit gesprochen. Ohne den gedanklichen Inhalt zu ändern, könnte man diese Begriffe durch Prüfung, Prüfgröße, Prüfergebnis und Ergebnisunsicherheit ersetzen.

Im einfachsten Fall einer Messung hat man es mit einer einzigen **Messgröße** zu tun, d. h. mit nur einer speziellen Größe, die Gegenstand der Messung ist. Dabei könnte es sich z. B. um den Dampfdruck einer gegebenen Wasserprobe bei 20 °C handeln. Es ist von sehr großer Bedeutung, dass die Messaufgabe durch Angaben z. B. von Zeitpunkt, Temperatur oder Druck genau spezifiziert wird¹. Ist die Messgröße einer Messaufgabe auf diese Art und Weise exakt beschrieben, dann kommt ihr ein eindeutiger Wert zu, der sog. **wahre Wert**. Diesen wahren Wert würde man bei einer idealen Messung erhalten.

Da man es aber stets mit realen Messungen zu tun hat, besteht zwischen dem Messergebnis und dem wahren Wert eine (unbekannte) Differenz, die **Messabweichung** genannt wird. Wiederholt durchgeführte Messungen ergeben im Allgemeinen nicht jedes Mal den gleichen Wert, sondern mehr oder weniger dicht beieinander liegende Messergebnisse. Würde man diese Messung sehr häufig wiederholen und die Häufigkeit, mit der ein Messergebnis x auftritt, über x auftragen, so erhielte man in vielen Fällen eine glockenförmige Kurve, die durch die sog. Normalverteilung angenähert werden kann (siehe Abb. 2.1). Eine Normalverteilung wird durch zwei Kenngrößen charakterisiert: den Lageparameter μ , der die Position des Maximums angibt, und die Standardabweichung σ , die die Breite der Kurve beschreibt.

Wegen dieser Streuung der Messwerte führt man, sofern das möglich ist und der Aufwand vertretbar erscheint, eine Messung mehrmals (n mal) durch und bildet den arithmetischen Mittelwert \bar{x} der n Einzelwerte x_i gemäß Gleichung (2.1).

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.1)$$

(\bar{x} : arithmetischer Mittelwert; x_i : Wert der i -ten Messung; n : Zahl der Messungen, $n > 1$)

Die nach Gleichung (2.2) berechnete (empirische) Standardabweichung s ist ein Maß für die Streuung der Einzelwerte, d.h. für die Breite der Kurve in Abbildung 2.1.

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.2)$$

(s : empirische Standardabweichung; \bar{x} : arithmetischer Mittelwert; x_i : Wert der i -ten Messung; n : Zahl der Messungen, $n > 1$)

¹ Diese Erläuterung des Begriffs "Messgröße", die dem VIM folgt, weicht ab von der DIN 1319-1. In dieser deutschen Norm bedeutet Messgröße die physikalische Größe (z.B. Masse, Energie etc.), die Gegenstand der Messung ist. Sofern die speziellen Sachbezüge (Messobjekt, Messbedingungen, Einflussgrößen) einbezogen werden, wird von **spezieller Messgröße** gesprochen.

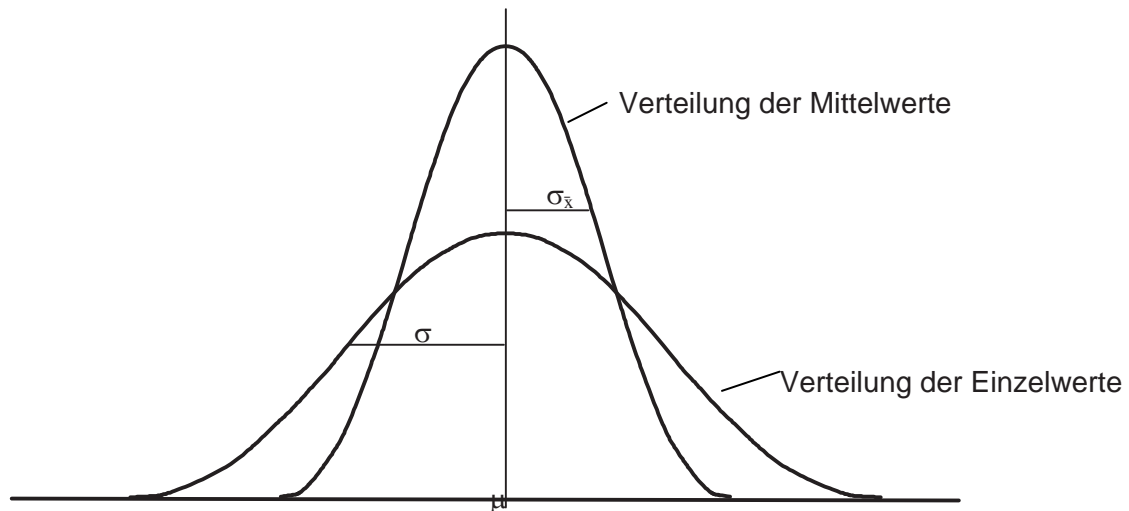


Abb. 2.1 Verteilung für die Einzelwerte der Messung x mit den Parametern μ und σ sowie für die Mittelwerte \bar{x} aus jeweils n Messungen mit den Parametern μ und $\sigma_{\bar{x}}$.

Auch wenn man solche Messreihen aus jeweils n Einzelmessungen sehr häufig wiederholt, jeweils den sich ergebenden Mittelwert berechnet und analog zu der Darstellung der Einzelwerte aufträgt, erhält man wieder eine Normalverteilung mit demselben Lageparameter μ , aber einer geringeren Breite (siehe Abb. 2.1). Für die Standardabweichung $\sigma_{\bar{x}}$ gilt:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad (2.3)$$

($\sigma_{\bar{x}}$: Standardabweichung (in der Grundgesamtheit) der Mittelwerte; σ : Standardabweichung (in der Grundgesamtheit) der Einzelwerte; n : Anzahl der Messwerte, die zur Mittelwertbildung herangezogen werden)

Die Streuung der Messwerte unter scheinbar identischen Bedingungen ist das Ergebnis einer Vielzahl von im Rahmen der Messbedingungen nicht kontrollierten Einflüssen, deren Wirkung sich bei wiederholter Durchführung der Messung ändert. Die dadurch bedingten Abweichungen der Messwerte vom Zentralwert μ , die einmal positiv, dann wieder negativ ausfallen, werden als **zufällige Messabweichungen** bezeichnet. Wenn ausschließlich zufällige Messabweichungen vorliegen, ist μ gleich dem wahren Wert der Messgröße. Man erhielte μ als Mittelwert \bar{x} , wenn man die Messung unbegrenzt wiederholen könnte, weil dann die Standardabweichung des Mittelwerts gegen Null ginge.

Da in der Praxis aber nur eine begrenzte Wiederholung der Messung möglich ist, verbleibt eine gewisse Streuung der Mittelwerte und damit eine gewisse Unkenntnis des Wertes der Messgröße, die man durch die **Messunsicherheit** abzuschätzen sucht. Sie ist nach DIN 1319-1 definiert als „Kennwert, der aus Messungen gewonnen wird und zusammen mit dem Messergebnis zur Kennzeichnung eines Wertebereiches für den wahren Wert der Messgröße dient“.

Zusätzlich zu diesen zufälligen Messabweichungen hat man es meist auch mit sog. **systematischen Messabweichungen** zu tun. Sie führen dazu, dass auch bei unendlicher Wiederholung das Zentrum der Verteilung gegenüber dem wahren Wert verschoben ist (siehe Abb. 2.2). Mögliche Ursachen für zufällige und systematische Abweichungen sind im Anhang A.1 aufgelistet. Festgestellte systematische Messabweichungen sollten, soweit das möglich ist, beseitigt oder durch geeignete Korrekturgrößen minimiert werden, wobei die Unsicherheit der Korrektur bei der Unsicherheitsbilanz zu berücksichtigen ist.

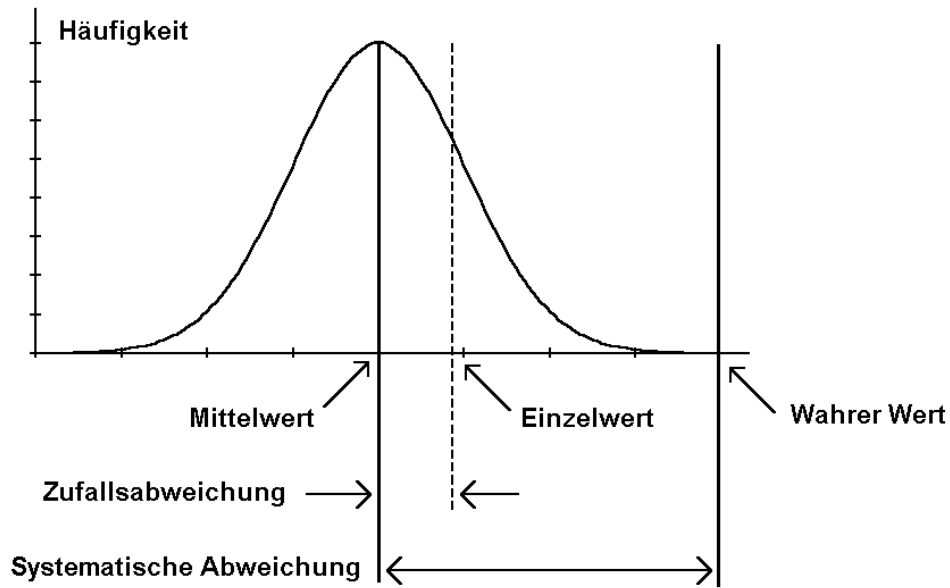


Abb. 2.2 Messwerte beim gleichzeitigen Auftreten von zufälligen und systematischen Abweichungen

In Abb. 2.3 werden die verschiedenen möglichen Komponenten der Messabweichung schematisch dargestellt. Es wird verdeutlicht, wie sie in das Messergebnis und die zugehörige Messunsicherheit eingehen.

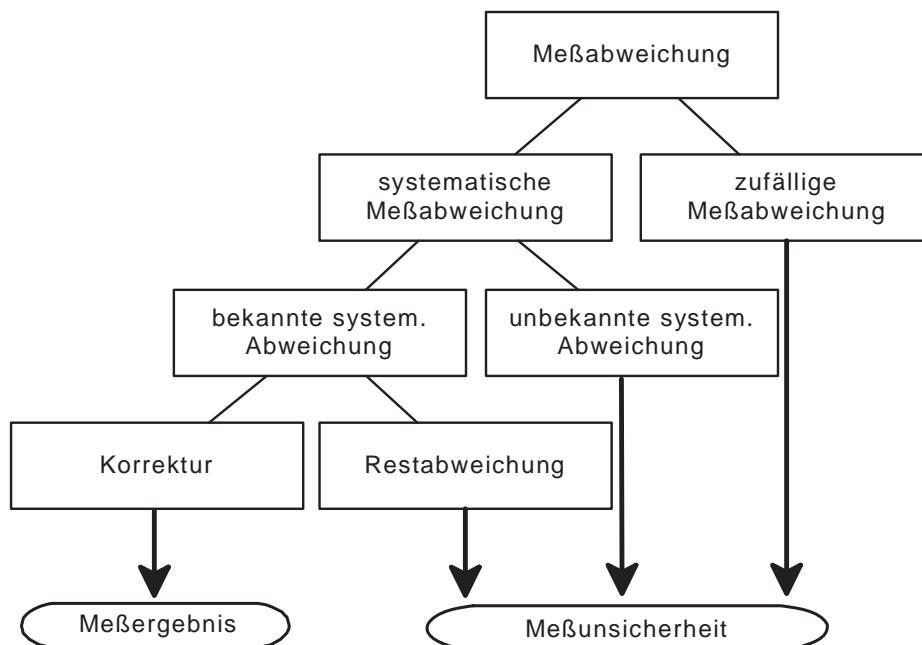


Abb. 2.3 Die verschiedenen möglichen Komponenten der Messabweichung und ihre Berücksichtigung bei der Ermittlung des Messergebnisses und der zugehörigen Messunsicherheit (Abbildung aus M. Hernla, QZ **41** (1996), 1156)

2.2 Genauigkeit, Richtigkeit und Präzision; das Zielscheibenmodell

Zur Charakterisierung eines Messverfahrens in Hinblick auf die damit verbundene Messunsicherheit dienen die Begriffe Genauigkeit, Richtigkeit und Präzision aus der ISO 3534-1, die in Abschnitt 1 dieses Leitfadens definiert sind.

Die Genauigkeit als Oberbegriff ist allgemein ein Maß für die Übereinstimmung eines Messergebnisses mit dem wahren Wert. Liegen aus einer Messreihe mehrere Messergebnisse für dieselbe Messgröße vor, so kann man einerseits die Übereinstimmung des Mittelwerts mit dem wahren Wert, die Richtigkeit, und andererseits die Übereinstimmung der Einzelwerte untereinander, die Präzision, betrachten (siehe Abb. 2.4).

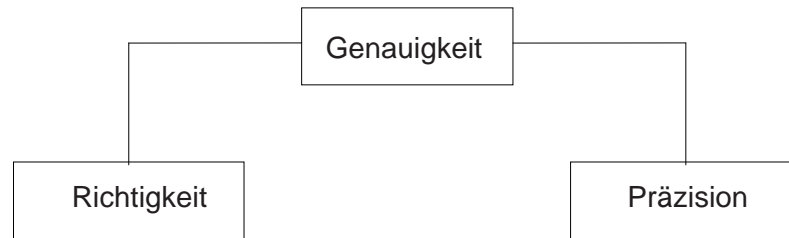


Abb. 2.4: Genauigkeit als Oberbegriff von Richtigkeit und Präzision

Die verschiedenen Kombinationsmöglichkeiten, die sich bei richtigen oder falschen bzw. präzisen oder unpräzisen Resultaten ergeben, lassen sich am Besten durch das Zielscheibenmodell beschreiben (Abb. 2.5).

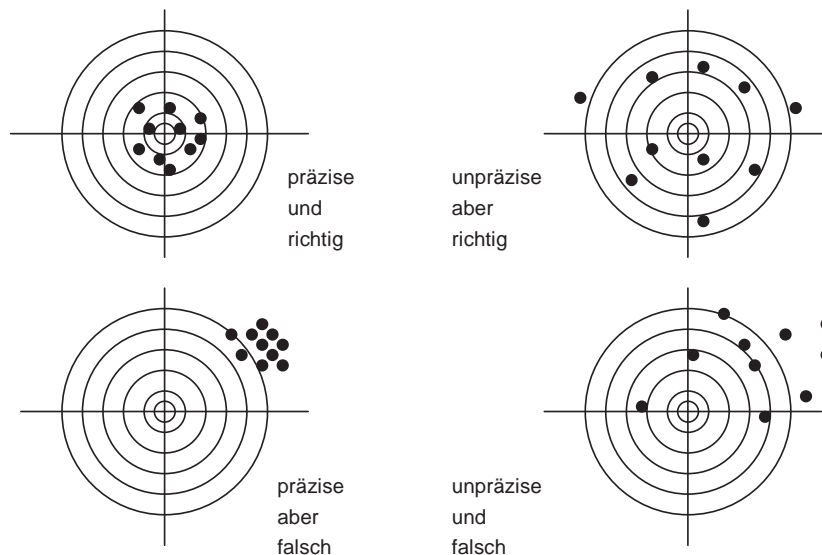


Abb. 2.5 Zielscheibenmodell zur Darstellung der Begriffe Richtigkeit und Präzision. Das Zentrum der Scheibe symbolisiert den (unbekannten) wahren Wert.

Bezüglich der Präzision unterscheidet man nach den Bedingungen, unter denen sie bestimmt wird, zwischen sog. **Wiederholbedingungen**, **Vergleichsbedingungen** und **Zwischenbedingungen**. Entsprechend spricht man von der **Wiederholpräzision**, der **Vergleichspräzision** bzw. der **Präzision unter Zwischenbedingungen**.

Wiederholbedingungen umfassen:

- dasselbe Messverfahren,
- dasselbe Labor
- denselben Bearbeiter,
- dasselbe Messgerät,
- Wiederholung innerhalb einer kurzen Zeitspanne.

Vergleichsbedingungen umfassen:

- dasselbe Messverfahren,
- verschiedene Labors,
- verschiedene Bearbeiter,
- verschiedene Messgeräte.

Wiederholbedingungen und Vergleichsbedingungen stellen die Fälle minimaler und maximaler Variabilität der Bedingungen für wiederholte Messungen dar. Bedingungen zwischen diesen Grenzfällen werden als Zwischenbedingungen bezeichnet. Bei der Verwendung von Zwischenbedingungen muss genau festgelegt sein, welche Faktoren variiert werden und welche identisch sind. Für die laborinterne Charakterisierung der Präzision von Prüfverfahren werden z.B. die folgenden Bedingungen verwendet:

- dasselbe Messverfahren,
- dasselbe Labor,
- verschiedene Bearbeiter,
- dasselbe Messgerät (alternativ: verschiedene Messgeräte),
- Wiederholung innerhalb längerer Zeiträume.

Bei diesem speziellen Fall von Zwischenbedingungen spricht man auch oft von „labor-internen Vergleichsbedingungen“.

Während die Vorgehensweise zur Ermittlung der Präzision eines Messverfahrens einsichtig ist, stellt sich die ungleich schwierigere Frage, wie man bei der Messung die Richtigkeit des Verfahrens bestimmen oder abschätzen soll, da der wahre Wert grundsätzlich unbekannt ist. Hier bietet es sich an, das Messverfahren auf geeignete Referenzobjekte (Normale, Maßverkörperungen, Referenzmaterialien) anzuwenden. Eine Alternative besteht darin, das Messverfahren parallel mit einem Referenzverfahren an geeigneten Messobjekten einzusetzen. Der den Referenzobjekten zugeordnete Wert bzw. das Ergebnis des Referenzverfahrens werden dabei jeweils als Referenzwert betrachtet, d.h. als ein Schätzwert für den unbekanntem wahren Wert, dessen Unsicherheit bekannt und für den vorgesehenen Zweck hinreichend klein ist. Die Richtigkeit wird dann auf diesen Referenzwert bezogen.

2.3 Neue Aspekte im "Guide to the Expression of Uncertainty in Measurement"

Der "Guide to the Expression of Uncertainty in Measurement" (GUM), der seit 1995 auch in deutscher Übersetzung vorliegt, verfolgt gegenüber der bisherigen Begriffsbildung eine etwas geänderte Betrachtungsweise und schlägt bei der Ermittlung der Messunsicherheit eine einheitliche und pragmatische Vorgehensweise vor, die im Folgenden erläutert werden soll.

2.3.1 Neue Definition der Messunsicherheit

Da es sich bei dem wahren Wert um eine Idealgröße handelt, die prinzipiell unbekannt ist, hat man im Rahmen der Erarbeitung des GUM auch eine neue Definition für den Begriff Messunsicherheit entwickelt, die sich nicht mehr auf den wahren Wert bezieht:

Messunsicherheit

Dem Messergebnis zugeordneter Parameter, der die Streuung der Werte kennzeichnet, die vernünftigerweise der Messgröße zugeordnet werden könnten.

Diese Definition, die im Anhang D des GUM näher erläutert wird, ist auch in die 2. Auflage des Internationalen Wörterbuchs der Metrologie (VIM) übernommen worden.

Zu den verschiedenen Werten gehören die, die man unter Wiederholbedingungen (siehe oben) erhält, und ggf. auch solche, die unter Vergleichsbedingungen gewonnen werden, d.h. die z.B. von einem anderen Beobachter, einem anderen Laboratorium oder aus einem anderen Messverfahren stammen und bei denen auch unbekannte systematische Abweichungen ins Spiel kommen können. Aber auch unterschiedliche Korrekturen und theoretische Ansätze können zu dieser Streuung beitragen (vgl. Abb. 2.3). Außerdem ist eventuell die Messgröße selbst nicht so genau definiert, dass ihr ein einzelner wahrer Wert zukommt.

Solange nicht einzelne dieser experimentell oder theoretisch erhaltenen Werte als falsch erkannt sind, muss man sie alle der Messgröße zuordnen. Die Messunsicherheit ist ein Maß für diese Bandbreite möglicher Werte und beschreibt zusammen mit dem Messergebnis den gesamten Kenntnisstand über die Messgröße. Wegen unseres begrenzten Wissens ist es durchaus möglich, dass aufgrund unberücksichtigter Unsicherheitskomponenten die Messunsicherheit unterschätzt wird.

Trotz dieser etwas anderen Betrachtungsweise besteht kein grundsätzlicher Widerspruch zwischen dem GUM und anderen Messunsicherheitsnormen.

2.3.2 Ermittlung von Unsicherheitskomponenten nach Typ A und Typ B

Der GUM klassifiziert die Unsicherheitskomponenten nach ihrer Ermittlungsmethode in Typ A und Typ B. Dabei bedeuten:

Typ A: Auswertung durch statistische Analyse von Messreihen,

Typ B: Auswertung mit anderen Mitteln als der statistischen Analyse von Messreihen.

Diese Klassifikation wird im Abschnitt 3.2 erläutert. Sie korrespondiert in gewisser Weise mit der Unterscheidung zwischen Unsicherheitskomponenten, die von zufälligen Effekten herrühren (kurz: „zufällige Unsicherheitskomponenten“) und Unsicherheitskomponenten, die von systematischen Effekten herrühren (kurz: „systematische Unsicherheitskomponenten“), ohne mit ihr identisch zu sein.

Im GUM wird hinsichtlich der vorgeschlagenen Vorgehensweise nicht zwischen zufälligen und systematischen Unsicherheitskomponenten unterschieden. Dabei ist allerdings vorausgesetzt, dass man erkannte systematische Messabweichungen, sofern das möglich ist, entweder durch technische Maßnahmen beseitigt oder aber rechnerisch korrigiert. Für die Unsicherheitsbilanz bleibt dann eine Komponente, die die Unsicherheit dieser messtechnischen Verbesserung oder Korrektur beinhaltet.

Für alle Komponenten ist eine einheitliche Behandlung vorgesehen (siehe Abschnitt 2.3.3). Ein Grund dafür ist, dass der systematische oder zufällige Charakter der mit einer Komponente verbundenen Messabweichungen nicht eindeutig ist, sondern vom jeweiligen Anwendungsfall abhängt. So wird aus einer durch zufällige Einflüsse bedingten Messabweichung eine systematische Messabweichung, wenn das Messergebnis als Eingangsgröße in eine weitere Messung einfließt.

Beispiel: Die Konzentration eines radioaktiven Isotops in einem Standardpräparat ist durch Aktivitätsmessungen bestimmt worden. Es wird vorausgesetzt, dass bei diesem Experiment ausschließlich zufällige Abweichungen auftreten.

Wenn nun in weiteren Versuchen der unbekannte Gehalt einer Probe durch Vergleich mit diesem Standardpräparat ermittelt wird, geht dessen Messabweichung jedes Mal im gleichen Sinne in das Messergebnis ein, wirkt sich also als systematische Abweichung aus.

Umgekehrt werden systematische Abweichungen, die ein Laboratorium bei einem bestimmten Messverfahren macht, zu zufälligen Abweichungen, wenn man z.B. bei der Auswertung eines Ringversuchs die Ergebnisse sehr vieler Laboratorien, die jeweils unterschiedliche systematische Abweichungen aufweisen, in einer Vergleichspräzision zusammenfasst.

2.3.3 Gleichbehandlung aller Unsicherheitskomponenten

Bei der Berechnung der kombinierten Standardunsicherheit werden alle Unsicherheitskomponenten gleich behandelt. Eine Diskussion dieser Vorgehensweise findet sich im Anhang E des GUM.

In der Vergangenheit hat man teilweise abweichend davon die zufälligen Komponenten quadratisch (entsprechend Gleichung (3.12) in Abschnitt 3.3) und die systematischen linear (entsprechend Gleichung (3.16) in Abschnitt 3.4) addiert und diese beiden Beiträge dann zusammengefasst. Das war darin motiviert, dass man eine „konservative“ Abschätzung der Unsicherheit erhalten, die Unsicherheit also keinesfalls unterschätzen wollte. Man nahm dafür eine eventuell zu große Unsicherheit in Kauf.

Beispiel: Die Unsicherheitsbilanz für eine Messung hat als zufällige Komponenten die Werte 3 und 2 und als systematische Komponenten 2 und 4 (in willkürlichen Einheiten) ergeben.

Nach dem GUM werden diese Komponenten alle quadratisch summiert (vgl. Abschnitt 3.3, Gleichung (3.12):

$$u = \sqrt{3^2 + 2^2 + 2^2 + 4^2} = \sqrt{9 + 4 + 4 + 16} = \sqrt{33} = 5,74$$

Wenn man dagegen die systematischen Komponenten linear addiert, ergibt sich mit

$$u' = \sqrt{3^2 + 2^2} + 2 + 4 = \sqrt{9 + 4} + 6 = 9,61$$

ein deutlich größerer Wert für die Unsicherheit.

Diesem Sicherheitsbedürfnis kann im Rahmen des GUM durch die Wahl eines geeigneten Erweiterungsfaktors k entsprochen werden (s. Abschnitt 3.3). Außerdem sind unter bestimmten Umständen, z. B. beim Vergleich mit Grenzwerten, auch Höchstwertabschätzungen der Ergebnisunsicherheit sinnvoll (s. dazu die Ausführungen in Abschnitt 2.4 und Abschnitt 3.5).

2.3.4 Erweiterte Unsicherheit

Eine Möglichkeit, die Messunsicherheit nach dem GUM anzugeben, ist die Angabe der erweiterten Unsicherheit

$$U(y) = k \times u(y)$$

also des Produkts der Standardunsicherheit $u(y)$ mit einem Erweiterungsfaktor k . Damit wird ein Intervall, der sog. Vertrauensbereich

$$y - U(y) \leq Y \leq y + U(y)$$

(y : Messergebnis; Y : wahrer Wert der Messgröße; U : erweiterte Unsicherheit)

festgelegt, von dem angenommen werden kann, dass es mit einer vorgegebenen Wahrscheinlichkeit p (z.B. $p = 95\%$) den wahren Wert Y der Messgröße enthält. In der Betrachtungsweise des GUM enthält dieses Intervall den Anteil p aller Werte, die der Messgröße vernünftigerweise zugeordnet werden können.

Die Berechnung des Vertrauensbereichs setzt die Kenntnis der Verteilungsfunktion der Messwerte voraus. Da diese Voraussetzung in der Regel nur sehr unvollkommen erfüllt ist, schlägt der GUM ganz pragmatisch die Wahl eines Erweiterungsfaktor k zwischen 2 und 3 vor. Empfohlen wird $k = 2$, was ganz grob einem Vertrauensniveau p von 95% entspricht. In jedem Fall muss der Faktor k explizit angegeben werden, damit auf die Standardunsicherheit u rückgeschlossen werden kann.

Statistisch strenger fundierte Vorgehensweisen zur Bestimmung der Erweiterungsfaktoren findet man im Anhang G des GUM.

Beispiel: In einem Prüfbericht wird die erweiterte Unsicherheit $U = 11,48$ angegeben bei einem Erweiterungsfaktor $k = 2$. Daraus ergibt sich die Standardunsicherheit u zu:

$$u = \frac{U}{k} = \frac{11,48}{2} = 5,74.$$

2.4 Abschätzung von Höchstwerten der Messunsicherheit

Die Abschätzung von Höchstwerten der Messunsicherheit kann z.B. dann von Interesse sein, wenn die Größe der Messunsicherheit für weitere Betrachtungen nur eine untergeordnete Rolle spielt oder wenn die Einhaltung bestimmter Grenzwerte oder Spezifikationen zu prüfen ist. In diesem Fall, der zu einer vereinfachten Ermittlung der Messunsicherheit führt, werden abweichend vom Prinzip der quadratischen Addition die einzelnen Beiträge zur Ergebnisunsicherheit linear addiert, und anstelle von Standardunsicherheiten können ggf. auch maximale Messabweichungen verwendet werden (vgl. dazu Abschnitt 3.5, Gleichungen (3.16) und (3.17)).

3 Analytisch-rechnerische Ermittlung von Messunsicherheiten

3.1 Übersicht

Die analytisch-rechnerische Ermittlung von Messunsicherheiten ist im Allgemeinen eine komplexe Prozedur, die sich aus mehreren Schritten zusammensetzt, und bei der viele - teilweise auch ungewohnte - Gesichtspunkte zu beachten sind. Zur Übersicht über die in den folgenden Abschnitten beschriebenen Verfahrensschritte und die dabei zu beachtenden Gesichtspunkte werden die wesentlichen Bausteine im Folgenden kurz skizziert.

Voraussetzung: Systematische Abweichungen - soweit bekannt - sind beseitigt bzw. korrigiert.

- Alle relevanten Unsicherheitsquellen werden identifiziert und aufgelistet.
- Die Beiträge der einzelnen Unsicherheitsquellen zur Ergebnisunsicherheit werden abgeschätzt und sortiert nach wesentlich/unwesentlich.
- Die wesentlichen Beiträge werden als Standardunsicherheiten (Standardabweichungen) quantifiziert. Dabei sind gleichberechtigt: Statistische Auswertung von Messreihen (Typ A Auswertung) und Schätzung nach Alternativverfahren (Typ B Auswertung).
- Es wird geprüft, ob Korrelationen zu berücksichtigen sind; ggf. werden die wesentlichen Korrelationen als Kovarianzen quantifiziert.
- Die Beiträge der wesentlichen Unsicherheitsquellen werden gemäß quadratischer Addition kombiniert; ggf. werden Kovarianzen einbezogen.
- Für die Ergebnisangabe wird die kombinierte Standardunsicherheit mit einem geeigneten Zahlenfaktor (in der Regel $k=2$) multipliziert.
- Bei der Ermittlung von Höchstwerten der Ergebnisunsicherheit (worst-case Abschätzungen) werden die Unsicherheitsbeiträge linear addiert; Kovarianzen entfallen.

In der Regel wird die Messunsicherheit nicht individuell für ein einzelnes Messergebnis bestimmt, sondern als Kenngröße für ein Messverfahren ermittelt. Sie gilt dann für alle Messobjekte und alle Messbedingungen, die bei der Ermittlung der Messunsicherheit berücksichtigt wurden. Vor der Verwendung muss daher in jedem Fall geprüft werden, ob das jeweilige Messobjekt und die jeweiligen Messbedingungen den Festlegungen für die Ermittlung der Messunsicherheit entsprechen. Sind in der Verfahrensunsicherheit wesentliche Unsicherheitskomponenten des Anwendungsfalls nicht berücksichtigt, so ist es häufig zweckmäßig, die Verfahrensunsicherheit als Bestandteil der Messunsicherheit zu übernehmen und die fehlenden Unsicherheitsbeiträge zu ergänzen.

Bei der Ermittlung von Messunsicherheiten ist das Verhältnis von Aufwand und Nutzen zu beachten. Es ist z. B. zweckmäßiger, alle wesentlichen Unsicherheitsbeiträge mit akzeptabler Genauigkeit zu erfassen statt einzelne Unsicherheitsbeiträge mit extremer Genauigkeit zu bestimmen.

3.2 Klassifikation von Messunsicherheiten nach Art der Auswertung

Entsprechend dem Vorschlag des GUM (siehe Abschnitt 2.3) werden alle Unsicherheiten durch Standardabweichungen ausgedrückt, unabhängig davon, ob sie auf zufälligen oder systematischen Einflüssen beruhen. Für die Ermittlung dieser Standardabweichung gibt es im Wesentlichen zwei verschiedene Verfahrensweisen. Die konventionelle Verfahrensweise (Typ A Auswertung) beruht auf der Annahme einer Häufigkeitsverteilung für die zufallsbedingte Streuung von Messergebnissen. Für die Standardabweichung dieser Häufigkeitsverteilung werden Schätzwerte ermittelt, indem Mehrfachmessungen durchgeführt und die Messwerte (Messreihen) statistisch ausgewertet werden. Die alternative Verfahrensweise (Typ B Auswertung) wird vorwiegend zur Schätzung von Unsicherheiten benutzt, die von unbekannt systematischen Abweichungen herrühren. Sie verwendet vernünftig angenommene Verteilungen möglicher Werte, die den Stand der unvollständigen Kenntnis über die betreffenden Größen wiedergeben und deren Standardabweichung. Die beiden Klassen von Unsicherheitsauswertungen werden im GUM wie folgt definiert.

Typ A: Auswertung durch statistische Analyse von Messreihen;

Typ B: Auswertung mit anderen Mitteln als der statistischen Analyse von Messreihen.

Typisches Beispiel für eine **Typ A Auswertung** ist die Ermittlung eines Schätzwertes für die Standardabweichung σ einer angenommenen Normalverteilung. Sind x_1, x_2, \dots, x_n die Ergebnisse wiederholter Messungen der betreffenden Größe X , so kann als Schätzwert für die Standardabweichung σ dieser Normalverteilung die empirische Standardabweichung s der Messreihe $\{x_1, x_2, \dots, x_n\}$ verwendet werden.

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (3.1)$$

wobei

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (3.2)$$

In Abwesenheit systematischer Messabweichungen ist dann der (arithmetische) Mittelwert \bar{x} der Messreihe $\{x_1, x_2, \dots, x_n\}$ ein geeigneter Schätzwert für den Wert der Messgröße X . Die Standardunsicherheit $u(\bar{x})$ dieser Ergebnisgröße ist gegeben durch

$$u(\bar{x}) = \frac{s}{\sqrt{n}} \quad (3.3)$$

Ist sichergestellt, dass das Messverfahren in dem betreffenden Messbereich frei von systematischen Abweichungen und mit konstanter Streuung arbeitet, so kann die empirische Standardabweichung der Messreihe $\{x_1, x_2, \dots, x_n\}$ auch zur Schätzung der Standardunsicherheit der Ergebnisse anderer Messungen in diesem Messbereich verwendet werden. Hierbei ist zu beachten, ob es sich bei der Ergebnisgröße um einen Einzelwert oder um den Mittelwert aus mehreren unabhängig gemessenen Einzelwerten handelt. Im Fall eines Einzelwertes ist die Standardunsicherheit gleich s , im Fall eines (arithmetischen) Mittelwertes aus m Einzelwerten gleich s/\sqrt{m} .

Anmerkung: Der Faktor $1/\sqrt{n}$ für die Standardabweichung eines Mittelwertes aus n Einzelwerten gilt nur für von einander unabhängige Einzelwerte. Bei von einander abhängigen Einzelwerten (bedingt durch korrelierte Messabweichungen) ist der Präzisionsgewinn geringer, siehe Anhang A.5.

Typisches Beispiel für eine **Typ B Auswertung** ist die Umwandlung einer Höchstwert/Mindestwert-Angabe in eine Standardunsicherheit. Angenommen, für den einem Referenzmaterial zugeschriebenen Merkmalswert (Referenzwert) sind nur ein Mindestwert x_{\min} und ein Höchstwert x_{\max} bekannt. Sind alle Werte in diesem Intervall als gleichwahrscheinlich anzunehmen, so können für den Referenzwert x und seine Standardunsicherheit $u(x)$ der Mittelwert und die Standardabweichung der Rechteckverteilung mit den Grenzen x_{\min} und x_{\max} verwendet werden.

$$x = \frac{(x_{\max} + x_{\min})}{2} \quad (3.4)$$

$$u(x) = \frac{(x_{\max} - x_{\min})}{\sqrt{12}} \quad (3.5)$$

Ist hingegen anzunehmen, dass Werte in der Mitte des Intervalls wahrscheinlicher sind als Werte am Rande, so kann z. B. anstelle der Rechteckverteilung (Gleichverteilung) eine symmetrische Dreiecksverteilung mit den Grenzen x_{\min} und x_{\max} gewählt werden. Dann folgt

$$x = \frac{(x_{\max} + x_{\min})}{2} \quad (3.6)$$

$$u(x) = \frac{(x_{\max} - x_{\min})}{\sqrt{24}} \quad (3.7)$$

Diese und andere Beispiele von Typ B Auswertungen sind z.B. im GUM, Abschnitt 4.3 und 4.4 enthalten.

Anmerkung: Für die quantitative Auswertung von Unsicherheiten wurden bislang vorwiegend nur Typ A Verfahren verwendet. Da diese Verfahren nicht universell anwendbar sind, wurden wesentliche Unsicherheitskomponenten häufig gar nicht oder nicht angemessen berücksichtigt. Diesem Missstand soll die Einführung der Typ B Verfahren abhelfen, die darauf abzielen, das bei den Fachleuten vorhandene Erfahrungswissen in Gestalt von regelbasierten Expertenschätzungen für Unsicherheitskomponenten nutzbar zu machen.

Im Allgemeinen setzen sich Ergebnisunsicherheiten aus mehreren Komponenten zusammen, von denen ein Teil nach Typ A, der andere Teil nach Typ B ausgewertet wurde. Deshalb ist die Klassifikation nach Typ A und Typ B in der Regel nur für die einzelnen Unsicherheitskomponenten sinnvoll.

3.3 Allgemeines Verfahren der Unsicherheitsermittlung

Die Unsicherheit eines quantitativen Ergebnisses setzt sich in der Regel aus mehreren Komponenten zusammen. Dementsprechend ist auch die Ermittlung der Messunsicherheit in der Regel eine komplexe Prozedur, die sich aus mehreren Schritten zusammensetzt.

In diesem Abschnitt wird eine allgemein anwendbare Abfolge von Verfahrensschritten beschrieben. Die hier gewählte Formulierung bezieht sich auf Messverfahren, lässt sich jedoch problemlos auf Prüfverfahren und Analysenverfahren übertragen.

Häufig vorkommende Unsicherheitsquellen und Datenauswertungstechniken werden im Anhang A.1 beschrieben.

Schritt 1: Festlegung der Ergebnisgröße und des Ermittlungsverfahrens

In diesem Schritt werden die Ergebnisgröße y und das Verfahren zur Ermittlung ihres Wertes festgelegt. Dies beinhaltet neben der eigentlichen Messung alle vorbereitenden Schritte, z.B. Probenahme und Probenpräparation, die bei Vorbereitung und Messung einzuhaltenden Bedingungen sowie die nachfolgende Auswertung.

Schritt 2: Definition der Eingangsgrößen, Identifikation der Unsicherheitsquellen

In diesem Schritt werden alle Eingangsgrößen x_i ($i = 1, 2, \dots, N$) definiert bzw. identifiziert, von denen das Ergebnis abhängt. Zugleich werden alle potentiellen Quellen für die Unsicherheit der Ergebnisgröße identifiziert (siehe Anhang A.1). Die Eingangsgrößen werden so definiert, dass alle relevanten Unsicherheitsquellen wirkungsmäßig erfasst werden.

Bei den Eingangsgrößen kann es sich handeln um:

- Konstituierende Messgrößen der Ergebnisgröße, z. B. Masse und Volumen, wenn die Dichte als Quotient dieser beiden Größen bestimmt wird;
- Einflussgrößen, d. h. Größen, die nicht Gegenstand der Messung sind, aber das Ergebnis beeinflussen, wie z. B. Druck und Temperatur der Probe bei einer Volumenmessung;
- Referenzgrößen, d. h. Größen, die der Kalibrierung oder der Korrektur systematischer Messabweichungen dienen, z. B. von Normalen oder Referenzmaterialien verkörperten Werte;
- Kenngrößen für das Eingangs/Ausgangsverhalten von Teilschritten des vollständigen Messverfahrens, z. B. Wirkungsgrade von Probenvorbereitungsverfahren, Korrekturfaktoren für festgestellte systematische Abweichungen, Parameter einer Kalibrierkurve etc.;
- Andere bei der Auswertung benutzte Größen, für die Literaturdaten herangezogen werden, z. B. Naturkonstanten oder Materialkenngrößen.

Unsicherheiten der Eingangsgrößen sind Quellen für die Unsicherheit der Ergebnisgröße. Umgekehrt lässt sich die Wirkung einer jeden Unsicherheitsquelle mittels geeigneter Eingangsgrößen (z. B. Wirkungsgraden oder Korrekturfaktoren) beschreiben. Eine solche Beschreibung wird im Folgenden vorausgesetzt. Zu diesem Zweck müssen die Eingangsgrößen so definiert werden, dass alle potentiellen Unsicherheitsquellen nach ihrer Wirkung erfasst werden. Hierfür wird die Verwendung von Flussdiagrammen empfohlen. Die Verwendung von Wirkungsgraden, Korrekturfaktoren o.ä. als Eingangsgrößen zur Modellierung von Verfahrensschritten wird im Anhang A.3 behandelt.

Zusammengefasst besteht die Aufgabe bei diesem Schritt in der Entwicklung eines mathematischen Modells für das vollständige Prüfverfahren, $y = F(x_1, x_2, \dots, x_N)$, d.h. einer Gleichung oder eines Algorithmus, die die Prüfgröße in Abhängigkeit von allen relevanten Einflussgrößen beschreiben.

Schritt 3: Ermittlung der maßgeblichen Unsicherheitsquellen

In diesem Schritt wird geprüft, welche von den identifizierten Unsicherheitsquellen maßgebliche Beiträge zur Ergebnisunsicherheit liefern. Hierzu wird für die betreffende Eingangsgröße der Beitrag zur Ergebnisunsicherheit als Produkt aus grob abgeschätzten Werten der Standardunsicherheit für die unter den gegebenen Bedingungen

zu erwartende Schwankungsbreite der Werte der Eingangsgröße sowie der Empfindlichkeit, mit der die Ergebnisgröße von der Eingangsgröße abhängt, überschlagsmäßig berechnet. Unterscheiden sich zwei Beiträge um den Faktor 1/5, so kann der kleinere in der Regel gegenüber dem größeren Beitrag vernachlässigt werden.

Anmerkung: Wegen der quadratischen Addition trägt eine um den Faktor $1/p$ kleinere Standardunsicherheit nur mit einem Anteil von ca. $1/(2p^2)$ der größeren Standardunsicherheit zur Ergebnisunsicherheit bei. Für $p = 5$ beträgt dieser Anteil ca. 2 %. Die Vernachlässigung kleiner Unsicherheitsbeiträge ist jedoch unzulässig, wenn sie in großer Zahl auftreten oder wenn Korrelationen vorliegen, die eine lineare Addition von Unsicherheitsbeiträgen anstelle der quadratischen Addition zur Folge haben.

Schritt 4: Quantifizierung der maßgeblichen Unsicherheitsbeiträge

In diesem Schritt werden die Beiträge der maßgeblichen Unsicherheitsquellen mittels der zugeordneten Eingangsgrößen quantifiziert. Hierzu wird für jede der betreffenden Eingangsgrößen x_i ($i = 1, 2, \dots, N$) die Standardunsicherheit $u(x_i)$ bestimmt: Je nach verfügbaren experimentellen Daten entweder als Standardabweichung der Werte einer Messreihe (Typ A Auswertung) oder als Standardabweichung einer vernünftig angenommenen Werteverteilung (Typ B Auswertung), z. B. einer Rechteckverteilung zwischen experimentell abgesicherten Extremwerten.

Anmerkung: Typ A Auswertungen haben vordergründig den Vorteil größerer Objektivität. Jedoch liefern empirische Standardabweichungen bei sehr kurzen Messreihen, wie sie in der Praxis häufig vorliegen, so ungenaue Schätzwerte für Standardunsicherheiten, dass eine erfahrungsbasierte Expertenschätzung (Typ B Auswertung) möglicherweise den Vorzug verdient. Beispielsweise beträgt die relative Standardabweichung einer empirischen Standardabweichung aus 5 Werten ca. 36 %, bei 10 Werten sind es immer noch 24 %. Hierbei wird vorausgesetzt, dass die Werte aus einer normalverteilten Grundgesamtheit stammen; bei Abweichungen von der Normalverteilung können die Ergebnisse noch ungünstiger ausfallen.

Weiterhin werden die Empfindlichkeitskoeffizienten c_i für die Abhängigkeit der Ergebnisgröße $y = F(x_1, x_2, \dots, x_N)$ von den Eingangsgrößen x_i bestimmt. Die Koeffizienten c_i sind gegeben durch die Differentialquotienten

$$c_i = \frac{\partial F}{\partial x_i} \quad (3.8)$$

Bei einfachen Modellfunktionen für die Ergebnisgröße y (Summen, Produkte o. ä.) können die Differentialquotienten direkt berechnet werden. Bei komplizierteren Modellfunktionen können numerisch berechnete Differenzenquotienten anstelle der Differentialquotienten verwendet werden (siehe Anhang A.4). Lässt sich der Einfluss einer Eingangsgröße x_i auf die Ergebnisgröße nicht modellmäßig erfassen, so verwendet man experimentell ermittelte Differenzenquotienten

$$c_i = \frac{\Delta y}{\Delta x_i} \quad (3.9)$$

Der Beitrag der Unsicherheit einer Eingangsgröße x_i zur kombinierten Standardunsicherheit der Ergebnisgröße y wird berechnet als Produkt $u_i = c_i \times u(x_i)$ aus der Standardunsicherheit $u(x_i)$ und dem Empfindlichkeitskoeffizienten c_i .

Schritt 5: Berücksichtigung von Korrelationen

In diesem Schritt wird zunächst geprüft, ob Korrelationen zwischen Unsicherheitsbeiträgen vorliegen. Solche Korrelationen entstehen, wenn Messabweichungen zweier Eingangsgrößen x_i und x_k nicht unabhängig voneinander, sondern stets gleichsinnig oder gegensinnig auftreten. Korrelationen sind zu erwarten,

wenn die betreffenden Eingangsgrößen voneinander abhängig sind oder beide von einer dritten Größe abhängen. Dies kann sich auf die Größen selbst, aber auch auf die Verfahren zur Ermittlung ihrer Werte beziehen.

Beispiel: Eine Korrelation liegt vor, wenn zur Kalibrierung für zwei verschiedene Messungen dasselbe Normal verwendet wird, oder wenn zwei Maßlösungen durch Verdünnung aus derselben Stammlösung hergestellt werden. Dann wirken sich Messabweichungen des Normals gleichsinnig auf die Ergebnisse der beiden Messungen aus. Ebenso wirkt sich eine Abweichung der Stammlösung vom angegebenen Gehalt gleichsinnig auf den Gehalt der beiden Maßlösungen aus.

Grundsätzlich sollten Korrelationen möglichst vermieden werden. Das heißt, es sind, soweit möglich, unabhängige Eingangsgrößen zugrunde zu legen und unabhängige Verfahren zur Ermittlung ihrer Werte zu verwenden.

Ist dies nicht möglich, so müssen die Korrelationen durch entsprechende Kovarianzen $u(x_i, x_k)$ quantifiziert und bei der Berechnung der kombinierten Standardunsicherheit der Ergebnisgröße y berücksichtigt werden.

Korrelationen tragen als Produkte $u_{ik} = c_i \times c_k \times u(x_i, x_k)$ der Kovarianz $u(x_i, x_k)$ mit den betreffenden Empfindlichkeitskoeffizienten c_i und c_k zur kombinierten Standardunsicherheit der Ergebnisgröße y bei.

Die Ermittlung von Kovarianzen wird einführend im Anhang A.6 und vertieft im GUM sowie in DIN 1319-4 behandelt. Einfach zu handhaben und für viele Zwecke ausreichend ist der folgende Fall: Zwei Eingangsgrößen x_i und x_k hängen von derselben Größe z ab. Dann ist die Kovarianz von x_i und x_k gegeben durch $u(x_i, x_k) = (\partial x_i / \partial z)(\partial x_k / \partial z)u(z)^2$. Hierbei ist $u(z)$ die Standardunsicherheit von z während $(\partial x_i / \partial z)$ und $(\partial x_k / \partial z)$ die Empfindlichkeitskoeffizienten für die Abhängigkeit der Größen x_i bzw. x_k von z sind. Hängen zwei Eingangsgrößen von mehreren gemeinsamen Größen ab, so ist die Kovarianz die Summe der entsprechenden Produkte.

Schritt 6: Berechnung der kombinierten Standardunsicherheit

In diesem Schritt werden die in den vorangegangenen Schritten ermittelten Beiträge zur Standardunsicherheit der Ergebnisgröße zusammengefasst. In der allgemeinsten Fassung, d.h. bei Berücksichtigung von Korrelationen zwischen sämtlichen Eingangsgrößen, lautet die Rechenvorschrift:

$$u(y)^2 = \sum_{i=1}^N u_i^2 + 2 \sum_{i=1}^{N-1} \sum_{k=i+1}^N u_{ik} \quad (3.10)$$

Ausführlicher dargestellt lautet diese Gleichung

$$u(y)^2 = \sum_{i=1}^N \left(\frac{\partial F}{\partial x_i} \right)^2 u(x_i)^2 + 2 \sum_{i=1}^{N-1} \sum_{k=i+1}^N \left(\frac{\partial F}{\partial x_i} \right) \left(\frac{\partial F}{\partial x_k} \right) \cdot u(x_i, x_k) \quad (3.11)$$

In den meisten Anwendungsfällen bestehen zwischen den Eingangsgrößen keine Korrelationen oder die Beiträge der Korrelationen können vernachlässigt werden. Dann reduziert sich die Gleichung (3.11) auf

$$u(y)^2 = \sum_{i=1}^N \left(\frac{\partial F}{\partial x_i} \right)^2 u(x_i)^2 \quad (3.12)$$

Die Standardunsicherheit der Ergebnisgröße $u(y)$ ergibt sich als positive Quadratwurzel aus der nach Gleichung (3.11) bzw. (3.12) berechneten Summe.

Anmerkung: Gleichung (3.12) ist die übliche Fassung des Gaußschen "Fehlerfortpflanzungsgesetzes" für unkorrelierte Größen, Gleichung (3.11) dessen Verallgemeinerung durch Berücksichtigung von Korrelationen.

Beide Gleichungen beruhen auf einer Reihenentwicklung der Ergebnisgröße nach Potenzen der Abweichungen der Eingangsgrößen vom betrachteten Wert, die nach dem linearen Glied abgebrochen wird. Bei Vorliegen ausgeprägter Nichtlinearitäten kann diese Näherung unzureichend sein. Dann müssen entweder weitere Terme der Reihenentwicklung (höhere Potenzen der Abweichungen) berücksichtigt werden, oder es müssen andere Auswertungsmethoden (numerische Simulation o.ä., siehe Anhang A.4.2) verwendet werden.

Besteht zwischen der Ergebnisgröße y und den Eingangsgrößen x_i ein einfacher Formelzusammenhang, so lassen sich die Empfindlichkeitsfaktoren mathematisch exakt durch Differentiation bestimmen.

Beispiel: Für Summen und Differenzen, $y = ax_1 + bx_2$ bzw. $y = ax_1 - bx_2$, gilt

$$u(y)^2 = a^2 u(x_1)^2 + b^2 u(x_2)^2$$

Für Produkte und Quotienten, $y = cx_1 x_2$ bzw. $y = cx_1/x_2$, gilt

$$\left(\frac{u(y)}{y}\right)^2 = \left(\frac{u(x_1)}{x_1}\right)^2 + \left(\frac{u(x_2)}{x_2}\right)^2$$

In allen anderen Fällen ist es zweckmäßiger, die Empfindlichkeitsfaktoren näherungsweise mittels Differenzenrechnung zu bestimmen. Diese Rechnungen wie auch die Zusammenführung durch quadratische Addition lassen sich sehr bequem mittels Tabellenkalkulation erledigen, siehe Abschnitt A.4.1.

Die in Gleichung (3.11) auftretenden Kovarianzen $u(x_i, x_k)$ hängen eng mit den Standardunsicherheiten der betreffenden Eingangsgrößen zusammen. Es gilt

$$u(x_i, x_k) = r(x_i, x_k) \cdot u(x_i) \cdot u(x_k) \quad (3.13)$$

Hierbei ist $r(x_i, x_k)$ der sog. Korrelationskoeffizient; sein Wert liegt zwischen -1 und 1. Die Werte 1 und -1 entsprechen dem Vorliegen totaler gleichsinniger bzw. totaler gegensinniger Korrelation, der Wert 0 der Abwesenheit von Korrelation. Wären alle Eingangsgrößen total gleichsinnig korreliert, so ergäbe sich die kombinierte Standardunsicherheit als lineare Summe $u(y) = \sum u_i$ der Unsicherheitsbeiträge. Hingegen werden bei vollständig unkorrelierten Eingangsgrößen die Unsicherheitsbeiträge quadratisch addiert gemäß $u(y)^2 = \sum u_i^2$. Die quadratische Addition ergibt in der Regel erheblich kleinere Werte für die kombinierte Standardunsicherheit $u(y)$ als die lineare Addition. Deshalb kann zur Abschätzung von Höchstwerten für die kombinierte Standardunsicherheiten die lineare Addition verwendet werden, ohne dass das Vorliegen von Korrelationen im Einzelnen geprüft wird. Als Verfahren zur Ermittlung von Unsicherheiten, die als Eingangsgrößen für die Ermittlung der Unsicherheit anderer Größen verwendet werden, ist die lineare Addition nicht geeignet, weil sie die kombinierte Standardunsicherheit in der Regel überschätzt.

Schritt 7: Festlegung von Erweiterungsfaktoren

Die Unsicherheit der Ergebnisgröße kann alternativ entweder als Standardunsicherheit $u(y)$ oder als erweiterte Standardunsicherheit $U(y) = k \times u(y)$, d. h. als Produkt der Standardunsicherheit mit einem geeignet gewählten Erweiterungsfaktor, angegeben werden.

Die erweiterte Standardunsicherheit wird gewählt, um einen Bereich abzugrenzen, von dem erwartet werden kann, dass er den wahren Wert der Ergebnisgröße mit hoher Sicherheit enthält.

Wenn keine nachvollziehbaren Gründe für eine andere Wahl vorliegen, soll der Wert von k zwischen 2 und 3 gewählt werden; empfohlen wird $k = 2$. Liegen hinreichende Kenntnisse über die Häufigkeitsverteilung der Ergebnisgröße vor,

so kann k als "Vertrauensfaktor" zu einem festgelegten Vertrauensniveau berechnet werden. Hierfür wird das Vertrauensniveau 0,95 (95 %) empfohlen.

Für derartige Rechnungen wird die Anzahl der effektiven Freiheitsgrade benötigt. Sie kann aus den Standardunsicherheiten und den Freiheitsgraden der Verteilungen für die Werte der Eingangsgrößen errechnet werden, siehe GUM, Abschnitt G.4.

3.4 Hinweise zur Verwendung von Unsicherheitsbudgets

Die analytisch-rechnerische Ermittlung der Messunsicherheit auf der Grundlage eines detaillierten Unsicherheitsbudgets eignet sich insbesondere für Messverfahren mit breitem Anwendungsbereich, d. h. bei erheblicher Variationsbreite der Messobjekte und der Messbedingungen. Dann lohnt der Aufwand für die Erstellung eines detaillierten Unsicherheitsbudgets, bei dem die Messunsicherheit in Abhängigkeit von den relevanten Einflussgrößen – insbesondere der Beschaffenheit der Messobjekte und der Messbedingungen – berechnet wird.

Für Messverfahren mit engem Anwendungsbereich – Messobjekte geringer Variationsbreite, standardisierte Messbedingungen – bieten die in Abschnitt 4 und 5 beschriebenen Verfahrensweisen zur Abschätzung der Messunsicherheit mit laborinternen Validierungsdaten und Ringversuchsdaten eine gute Alternative.

Unsicherheitsbudgets sind wertvolle diagnostische Werkzeuge bei der Entwicklung und Optimierung von Messverfahren. Für diesen Zweck ist die folgende Form der Gleichung (3.10) – zur Vereinfachung ohne Korrelationen – besonders gut geeignet:

$$1 = \sum_{i=1}^N \frac{u_i^2}{u(y)^2} \quad (3.14)$$

An den Varianzanteilen $u_i^2 / u(y)^2$ kann man ablesen, welche Einflussgrößen wesentlich zur kombinierten Unsicherheit des Messergebnisses beitragen, und für welche Einflussgrößen der Beitrag marginal ist. Nur für die wesentlichen Einflussgrößen lohnt der Aufwand für eine Verbesserung der Genauigkeit, während dies für die marginalen Einflussgrößen vergeudete Mühe wäre.

Eine weitere nützliche Form der Grundgleichung für die Unsicherheitsfortpflanzung – auch hier zur Vereinfachung ohne Korrelationen – lautet:

$$\left(\frac{u(y)}{y}\right)^2 = \sum_{i=1}^N d_i^2 \left(\frac{u(x_i)}{x_i}\right)^2 \quad \text{mit} \quad d_i = \frac{c_i \cdot x_i}{y} \quad (3.15)$$

Die Koeffizienten d_i geben an, wie stark die relative Unsicherheit einer Einflussgröße auf die relative Unsicherheit der Ergebnisgröße durchgreift.

3.5 Abschätzung von Höchstwerten

Die in Abschnitt 3.3 beschriebene Verfahrensweise zielt darauf ab, die Messunsicherheit mit angemessener Genauigkeit zu ermitteln. Im Einzelfall kann jedoch anstelle eines genauen Wertes der Messunsicherheit ein Höchstwert von Interesse sein, z. B. wenn die Größe der Messunsicherheit bei der weiteren Verwendung des Ergebnisses nur eine untergeordnete Rolle spielt, oder wenn die Einhaltung vorgegebener Spezifikationen bzw. Grenzwertforderungen sicherzustellen ist (siehe DIN 1319-3, Anhang D).

Bei der Abschätzung von Höchstwerten der Messunsicherheit ergeben sich gegenüber dem in Abschnitt 3.3 beschriebenen Auswertungsverfahren folgende Vereinfachungen:

- Die Unsicherheitsbeiträge $u_i = c_i \times u(x_i)$ der Eingangsgrößen werden linear addiert; die Korrelationsbeiträge $u_{ik} = c_i \times c_k \times u(x_i, x_k)$ entfallen.
- In den Unsicherheitsbeiträgen $u_i = c_i \times u(x_i)$ der Eingangsgrößen können anstelle der Standardunsicherheiten $u(x_i)$ Höchstwerte der möglichen Messabweichungen $|\Delta x_i|_{\max}$ verwendet werden.

Mit diesen Vereinfachungen ergeben sich die beiden folgenden Gleichungen, die wahlweise zur Berechnung von Höchstwerten der Messunsicherheit verwendet werden können.

$$|\Delta y|_{\max} = \sum_{i=1}^N \left| \frac{\partial F}{\partial x_i} \right| \cdot u(x_i) \quad (3.16)$$

$$|\Delta y|_{\max} = \sum_{i=1}^N \left| \frac{\partial F}{\partial x_i} \right| \cdot |\Delta x_i|_{\max} \quad (3.17)$$

4 Abschätzung von Messunsicherheiten mit laborinternen Validierungsdaten

4.1 Allgemeines

Eine direkte Methode der Messunsicherheitsermittlung besteht darin, das betreffende Messverfahren auf Referenzobjekte (Normale, Maßverkörperungen, Referenzmaterialien) anzuwenden und die unter laborinternen Vergleichsbedingungen (siehe Abschnitt 2.2) erhaltenen Ergebnisse mit den bekannten Referenzwerten zu vergleichen. Eine Variante, die weitgehend demselben Prinzip folgt, besteht darin, das Messverfahren parallel mit einem Referenzverfahren auf geeignete Messobjekte anzuwenden und die Ergebnisse des zu bewertenden Verfahrens mit denen des Referenzverfahrens zu vergleichen. Bei beiden Varianten wird die Messunsicherheit entsprechend dem Grundprinzip *Genauigkeit = Richtigkeit + Präzision* aus Kennwerten zur Richtigkeit und Kennwerten zur Präzision aufgebaut.

Die nachstehend beschriebene Prozedur besteht aus folgenden Schritten:

- Prüfung der Präzision;
- Prüfung auf systematische Abweichungen;
- ggf. Korrektur signifikanter systematischer Abweichungen;
- Ermittlung der Messunsicherheit für die (ggf. korrigierte) Messgröße.

In Abschnitt 4.2 wird als einfachster Fall die Verfahrensweise bei Verwendung eines einzigen Referenzobjekts beschrieben. Sind aus technischen Gründen mehrere Referenzobjekte erforderlich, z. B. bei der Ermittlung der Unsicherheit über einen weiten Messbereich, so ist die hier beschriebene Verfahrensweise entsprechend zu erweitern. Hierfür geeignete Verfahrensweisen werden in Abschnitt 4.3 beschrieben.

Üblicherweise werden Untersuchungen von Präzision und Richtigkeit eines Messverfahrens regelmäßig (bei besonderem Anlass ggf. auch zusätzlich) wiederholt. Hierbei ist es wichtig, sicherzustellen, dass die Daten einer aktuellen Untersuchung mit den Daten aus vorangegangenen Untersuchungen vergleichbar sind:

- Sind die Daten miteinander vereinbar, so können sie zusammengefasst werden, um die statistische Basis der betreffenden Schätzwerte (mittlere Abweichungen, mittlere Wiederfindungen und deren Standardabweichung) zu verbessern.
- Andernfalls liefert der Datenvergleich als diagnostisches Werkzeug Hinweise zur Aufklärung der festgestellten Diskrepanzen.

Aus diesem Grund sollten die Messungen an Referenzobjekten stets in gleicher Weise durchgeführt und ausgewertet werden – ohne Berücksichtigung von ggf. zuvor ermittelten Korrekturen.

4.2 Ein-Punkt-Prozedur

Die hier beschriebene Verfahrensweise ist nur dann anwendbar, wenn vorausgesetzt werden kann, dass der Befund am Referenzobjekt als repräsentativem Punkt für den gesamten Messbereich des Verfahrens (mit anderen Worten für die Gesamtheit der Messobjekte bzw. Messaufgaben) gilt. Andernfalls muss entweder der Messbereich entsprechend eingeschränkt oder eine Mehr-Punkt-Prozedur entsprechend Abschnitt 4.3 verwendet werden.

Das Referenzobjekt wird mehrfach (mindestens $n = 6$ mal) unter angemessenen laborinternen Vergleichsbedingungen (siehe Abschnitt 2.2), die denen beim späteren Einsatz des Messverfahrens entsprechen, gemessen. Für das verwendete Referenzobjekt sei

x_{ref}	Referenzwert der Messgröße;
$u(x_{\text{ref}})$	Standardunsicherheit des Referenzwertes;
x_{mess}	mit dem betreffenden Messverfahren erhaltener Messwert;
\bar{x}_{mess}	Mittelwert der n Messwerte x_{mess} ;
s_{mess}	Standardabweichung der n Messwerte x_{mess} ;
Δ	mittlere Abweichung ($\Delta = \bar{x}_{\text{mess}} - x_{\text{ref}}$) vom Referenzwert;
Q	mittlere Wiederfindung ($Q = \bar{x}_{\text{mess}} / x_{\text{ref}}$) des Referenzwertes.

Zunächst wird geprüft, ob die Standardabweichung der Messreihe mit der zuvor ermittelten und überwachten Verfahrensstandardabweichung vereinbar ist (Abschnitt 4.2.1). Anschließend wird der Mittelwert der Messreihe mit dem Referenzwert verglichen, um Aufschluss über systematische Abweichungen zu erhalten, und die festgestellte systematische Abweichung wird bewertet gemäß: „nicht akzeptabel“, „signifikant aber akzeptabel“ oder „nicht signifikant“ (Abschnitt 4.2.2). Je nach Ergebnis der Bewertung werden entsprechende Maßnahmen (Abschnitt 4.2.3) ergriffen:

Befund	nicht akzeptabel	signifikant aber akzeptabel	nicht signifikant
Maßnahme	Überprüfung und Verbesserung des Messverfahrens	Einführung einer Korrektur oder Einführung eines zusätzlichen Unsicherheitsbeitrags	Einführung eines zusätzlichen Unsicherheitsbeitrags

Tabelle 3.1 Befunde und Maßnahmen betr. systematische Abweichungen

Als Ergebnis der Untersuchung erhält man einen Schätzwert für die Messunsicherheit des (ggf. korrigierten) Messverfahrens (Abschnitt 4.2.3).

4.2.1 Prüfung der Präzision

Als Voruntersuchung wird die Präzision des Verfahrens unter laborinternen Vergleichsbedingungen (siehe Abschnitt 2.2), die den Bedingungen beim üblichen Laborbetrieb entsprechen, ermittelt. Hierfür eignet sich die Standardabweichung aus regelmäßigen Messungen an einem geeigneten Messobjekt (Präzisions-Regelkarte), ggf. auch die gepoolte Standardabweichung bei Verwendung mehrerer Messobjekte oder mehrerer Messgeräte. Diese Präzision wird im Folgenden als „Verfahrenspräzision“ bezeichnet. Für die zugehörige Standardabweichung werden die Bezeichnung „Verfahrensstandardabweichung“ und das Formelzeichen s_v verwendet.

Anmerkung: Für die Zusammenfassung (Pooling) von zwei Standardabweichungen gilt folgende Rechenvorschrift:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Hierbei sind n_1 und n_2 die jeweilige Anzahl von Messwerten, aus denen s_1 bzw. s_2 bestimmt wurden.

Die Standardabweichung s_{mess} der Messreihe am Referenzobjekt sollte in etwa mit der Verfahrensstandardabweichung s_v übereinstimmen, jedenfalls sollte s_{mess} nicht signifikant größer sein als s_v . Im Zweifelsfall kann das mit einem F-Test geprüft werden.

Anmerkung: Mit dem F-Test wird geprüft, ob zwei Standardabweichungen sich signifikant unterscheiden. Hierbei wird der quadrierte Quotient $(s_> / s_<)^2$, gebildet aus der größeren und der kleineren der beiden Standardabweichungen, mit dem Tabellenwert der F-Verteilung für die jeweiligen Freiheitsgrade und das gewünschte Signifikanzniveau verglichen. Der F-Test wird in allen einschlägigen Lehrbüchern der statistischen Datenauswertung beschrieben, siehe z.B. L. Sachs, *Angewandte Statistik*.

4.2.2 Prüfung auf systematische Abweichungen

Ist die Präzision am Referenzobjekt mit der zuvor bestimmten Verfahrenspräzision vereinbar, so wird als nächstes die Abweichung der am Referenzobjekt erhaltenen Messwerte vom Referenzwert untersucht und bewertet. Grundsätzlich kann das für jeden einzelnen Messwert erfolgen; hier wird jedoch der Einfachheit halber die mittlere Abweichung, d.h. die Abweichung des Mittelwertes untersucht. Dabei wird zunächst geprüft, ob die Abweichung des Mittelwerts akzeptabel oder nicht akzeptabel ist.

Eine nicht akzeptable Abweichung zeigt schwerwiegende Unzulänglichkeiten des Messverfahrens an, die eine eingehende Untersuchung aller Verfahrensschritte und Geräte auf Fehlerquellen und entsprechende Korrekturmaßnahmen erfordern.

Eine akzeptable Abweichung entspricht den Erwartungen hinsichtlich der Richtigkeit des Verfahrens und erfordert daher keine Revision des Verfahrens.

Ist die Abweichung des Mittelwerts akzeptabel, so wird sie des Weiteren auf (statistische) Signifikanz geprüft.

Eine signifikante systematische Abweichung liegt vor, wenn der Betrag der Abweichung des Mittelwertes \bar{x}_{mess} vom Referenzwert x_{ref} größer ist als die doppelte Standardunsicherheit dieser Abweichung,

$$|\bar{x}_{\text{mess}} - x_{\text{ref}}| > 2 \sqrt{\frac{s_{\text{mess}}^2}{n} + u(x_{\text{ref}})^2} \quad (4.1)$$

Andernfalls ist die Abweichung nicht signifikant.

Anmerkung: Anstatt zu prüfen, ob die mittlere Abweichung Δ signifikant von Null verschieden ist, kann auch geprüft werden, ob die mittlere Wiederfindung Q signifikant von Eins abweicht. Diese beiden Prüfungen sind i.W. äquivalent.

4.2.3 Behandlung festgestellter systematischer Abweichungen

Signifikante systematische Abweichungen werden je nach Datenlage entweder korrigiert oder die festgestellte Abweichung wird in die Messunsicherheit eingerechnet. Nicht signifikante systematische Abweichungen werden ausschließlich in die Messunsicherheit eingerechnet.

Korrektur

Bei Feststellung signifikanter systematischer Abweichungen ist eine Ein-Punkt-Korrektur nur dann sinnvoll, wenn vorausgesetzt werden kann, dass die (absoluten oder relativen) systematischen Abweichungen über den gesamten Messbereich konstant sind.

Ist mit konstanten absoluten Abweichungen zu rechnen, so wird die festgestellte mittlere Abweichung $\Delta = \bar{x}_{\text{mess}} - x_{\text{ref}}$ vom Messergebnis abgezogen.

$$y_{\text{korr}} = y_{\text{mess}} - \Delta \quad (4.2)$$

Hierbei ist y_{mess} das Messergebnis an einem Prüfobjekt, y_{korr} das korrigierte Messergebnis.

Ist mit konstanten relativen Abweichungen zu rechnen, so wird die Korrektur mit Hilfe der mittleren Wiederfindung $Q = \bar{x}_{\text{mess}} / x_{\text{ref}}$ wie folgt durchgeführt.

$$y_{\text{korr}} = \frac{y_{\text{mess}}}{Q} \quad (4.3)$$

Die Korrektur kann entweder durch Justierung der Messeinrichtung (Abgleich von Nullpunkt bzw. Empfindlichkeit) oder rechnerisch erfolgen.

Die Standardunsicherheit der korrigierten Messgröße wird nach den Regeln der Unsicherheitsfortpflanzung (siehe Abschnitt 3.3) bei Korrektur gemäß Gleichung (4.2) wie folgt berechnet.

$$u(y_{\text{korr}})^2 = s(y_{\text{mess}})^2 + u(\Delta)^2 = s_V^2 + \frac{s_{\text{mess}}^2}{n} + u(x_{\text{ref}})^2 \quad (4.4)$$

Hierbei wird für die Standardabweichung der unkorrigierten Messgröße die Verfahrensstandardabweichung angesetzt. Zur Vereinfachung kann auch s_{mess} durch s_V ersetzt werden.

Bei Korrektur gemäß Gleichung (4.3) gilt eine entsprechende Gleichung für die relative Standardunsicherheit, $u_{\text{rel}}(y) = u(y)/|y|$.

$$u_{\text{rel}}(y_{\text{korr}})^2 = s_{\text{rel}}(y_{\text{mess}})^2 + u_{\text{rel}}(Q)^2 = s_{\text{rel}_V}^2 + \frac{s_{\text{rel}_\text{mess}}^2}{n} + u_{\text{rel}}(x_{\text{ref}})^2 \quad (4.5)$$

Hierbei sind s_{rel_V} die relative Verfahrensstandardabweichung und $s_{\text{rel}_\text{mess}} = s_{\text{mess}} / \bar{x}_{\text{mess}}$ die relative Standardabweichung der Messwerte am Referenzobjekt. Auch hier kann zur Vereinfachung $s_{\text{rel}_\text{mess}}$ durch s_{rel_V} ersetzt werden.

Ist das Messergebnis y_{mess} der Mittelwert aus m gemessenen Einzelwerten, so ist in Gleichung (4.4) s_V^2 zu ersetzen durch s_V^2 / m . Entsprechendes gilt für Gleichung (4.5).

Einrechnung in die Messunsicherheit

Ist die direkte Übertragbarkeit der am Referenzobjekt festgestellten absoluten oder relativen systematischen Abweichungen auf die zu untersuchenden Messobjekte zweifelhaft, so wird von einer Korrektur abgesehen. Stattdessen wird die am Referenzobjekt festgestellte (mittlere) Abweichung als Unsicherheitsbeitrag in die Messunsicherheit eingerechnet. Hierbei wird die Vorgehensweise von Lira und Wöger [Literaturzitat am Ende des Abschnitts] empfohlen, mit dem Ergebnis

$$u(y_{\text{unkorr}})^2 = u(y_{\text{korr}})^2 + (y_{\text{korr}} - y_{\text{mess}})^2 \quad (4.6)$$

Genau so wird vorgegangen, wenn die am Referenzobjekt festgestellte (mittlere) Abweichung nicht signifikant ist.

Ausgehend von Gleichung (4.2) ergibt dieser Ansatz:

$$u(y_{\text{unkorr}})^2 = s(y_{\text{mess}})^2 + u(\Delta)^2 + \Delta^2 = s_V^2 + \frac{s_{\text{mess}}^2}{n} + u(x_{\text{ref}})^2 + (x_{\text{mess}} - x_{\text{ref}})^2 \quad (4.7)$$

Gleichung (4.7) setzt voraus, dass die Messunsicherheit im betrachteten Messbereich annähernd konstant ist. In der Regel nimmt die Messunsicherheit jedoch mit größeren Werten der Messgröße zu. Bei Proportionalität kann aus Gleichung (4.3) ein Ansatz analog Gleichung (4.7) für eine annähernd konstante relative Messunsicherheit abgeleitet werden. Häufig sind jedoch beide Ansätze zweifelhaft. In solchen Fällen ist die nachfolgende Abschätzung vorzuziehen, die darauf abzielt, die am Referenzobjekt ermittelte Messunsicherheit sowohl nach kleineren als auch nach größeren Werten der Messgröße sicher zu extrapolieren.

$$u(y_{\text{unkorr}})^2 = s_V^2 + \frac{s_{\text{mess}}^2}{n} + u(x_{\text{ref}})^2 + (\bar{x}_{\text{mess}} - x_{\text{ref}})^2 \quad \text{für } y_{\text{mess}} \leq \bar{x}_{\text{mess}} \quad (4.8)$$

$$u(y_{\text{unkorr}})^2 = \left(\frac{y_{\text{mess}}}{\bar{x}_{\text{mess}}} \right)^2 \left(s_V^2 + \frac{s_{\text{mess}}^2}{n} + u(x_{\text{ref}})^2 + (\bar{x}_{\text{mess}} - x_{\text{ref}})^2 \right) \quad \text{für } y_{\text{mess}} > \bar{x}_{\text{mess}}$$

Hierbei ist s_V die bei $y_{\text{mess}} \approx \bar{x}_{\text{mess}}$ gültige Verfahrensstandardabweichung.

Die Extrapolation der am Referenzobjekt ermittelten Messunsicherheit gemäß den beiden Gleichungen (4.8) beruht auf der folgenden empirischen Tatsache: Bei abnehmenden Werten der Messgröße nimmt in der Regel die Messunsicherheit ebenfalls ab oder bleibt im Höchstfall konstant. Bei zunehmenden Werten der Messgröße nimmt in der Regel die Messunsicherheit ebenfalls zu, wobei sie im Höchstfall proportional zum Wert der Messgröße wächst. Diese Regel gilt jedoch nicht ausnahmslos. Deshalb muss in Zweifelsfällen die Anwendbarkeit der Extrapolation nach den Gleichungen (4.8) überprüft werden.

Literatur: Lira, I.H. und Wöger, W., "Evaluation of the uncertainty associated with a measurement result not corrected for systematic effects", Meas Sci Technol 1998, **9**, S. 1010-1011.

4.3 N-Punkt-Prozedur ($N \geq 2$)

4.3.1 Interpolation

Sind die Voraussetzungen für die Anwendung der Ein-Punkt-Prozedur nicht gegeben, so müssen für den Ist/Soll-Vergleich und die ggf. erforderliche Korrektur mehrere Referenzobjekte verwendet werden. Kann von einem linearen Zusammenhang zwischen Messabweichung und Messgröße ausgegangen werden, so genügen für den Ist/Soll-Vergleich zwei Referenzobjekte.

Die Größen und Formelzeichen sind dieselben wie in Abschnitt 4.2, bis auf einen zusätzlichen Index A bzw. B zur Kennzeichnung von Größen für das jeweilige Referenzobjekt A und B. Der Einfachheit halber wird vorausgesetzt, dass für beide Referenzobjekte dieselbe Anzahl n von Einzelwerten vorliegt.

Prüfung der Präzision

Die Prüfung wird an beiden Referenzobjekten nach dem in 4.2.1 beschriebenen Verfahren durchgeführt. Liegen die Referenzwerte $x_{A\text{ref}}$ und $x_{B\text{ref}}$ dicht beieinander, so kann dieselbe Verfahrensstandardabweichung s_V zum Vergleich herangezogen werden. Andernfalls müssen zwei entsprechende Verfahrensstandardabweichungen s_{AV} und s_{BV} bestimmt und verwendet werden.

Wächst die Verfahrensstandardabweichung proportional zum Wert der Messgröße, so wird die Prüfung mit relativen Standardabweichungen durchgeführt. Dann genügt ein Wert $s_{rel,V}$ der relativen Verfahrensstandardabweichung.

Prüfung auf systematische Abweichungen

Die Prüfung wird an beiden Referenzobjekten nach dem in 4.2.2 beschriebenen Verfahren durchgeführt. Wird an einem Referenzobjekt eine nicht akzeptable systematische Abweichung festgestellt, so sind eine eingehende Untersuchung aller Verfahrensschritte und Geräte auf Fehlerquellen und entsprechende Korrekturmaßnahmen erforderlich.

Sind die an beiden Referenzobjekten gefundenen systematischen Abweichungen akzeptabel, so wird auf Signifikanz geprüft. Bei Signifikanz an einem oder beiden Referenzobjekten wird je nach Datenlage entweder eine Korrektur durchgeführt oder die festgestellte Abweichung in die Messunsicherheit eingerechnet. Nicht signifikante systematische Abweichungen an beiden Referenzobjekten werden ausschließlich in die Messunsicherheit eingerechnet.

Korrektur

Die Korrektur festgestellter systematischer Abweichungen wird mit Hilfe zweier Korrekturgrößen p und q entsprechend folgender Gleichung durchgeführt.

$$y_{\text{korr}} = p + q \cdot y_{\text{mess}} \quad (4.9)$$

Die Korrekturgrößen werden so bestimmt, dass y_{korr} für die beiden Referenzobjekte mit den entsprechenden Referenzwerten übereinstimmt. Die Rechnung ergibt

$$q = \frac{x_{\text{Bref}} - x_{\text{Aref}}}{\bar{x}_{\text{Bmess}} - \bar{x}_{\text{Amess}}} \quad (4.10)$$

$$p = \frac{1}{2} [(x_{\text{Aref}} + x_{\text{Bref}}) - q(\bar{x}_{\text{Amess}} + \bar{x}_{\text{Bmess}})] \quad (4.11)$$

Die Standardunsicherheit der korrigierten Messgröße wird nach den Regeln der Unsicherheitsfortpflanzung wie folgt aus den Standardunsicherheiten bzw. -abweichungen der beteiligten Größen, $s(y_{\text{mess}})$, $u(x_{\text{Aref}})$, $u(x_{\text{Bref}})$, $s(\bar{x}_{\text{Amess}}) = s_{\text{Amess}}/\sqrt{n}$, $s(\bar{x}_{\text{Bmess}}) = s_{\text{Bmess}}/\sqrt{n}$ berechnet.

$$u(y_{\text{korr}})^2 = q^2 s(y_{\text{mess}})^2 + \left(\frac{\bar{x}_{\text{Bmess}} - y_{\text{mess}}}{\bar{x}_{\text{Bmess}} - \bar{x}_{\text{Amess}}} \right)^2 \left(u(x_{\text{Aref}})^2 + p^2 \cdot \frac{s_{\text{Amess}}^2}{n} \right) + \left(\frac{y_{\text{mess}} - \bar{x}_{\text{Amess}}}{\bar{x}_{\text{Bmess}} - \bar{x}_{\text{Amess}}} \right)^2 \left(u(x_{\text{Bref}})^2 + p^2 \cdot \frac{s_{\text{Bmess}}^2}{n} \right) \quad (4.12)$$

Für die Standardabweichung $s(y_{\text{mess}})$ der unkorrigierten Messgröße wird die entsprechende Verfahrensstandardabweichung angesetzt. Hierbei ist zu beachten, ob die Ergebnisgröße als Einzelwert oder als Mittelwert aus einer festgelegten Anzahl von Einzelwerten definiert ist. Ist das Messergebnis ein gemessener Einzelwert, so gilt $s(y_{\text{mess}}) = s_V$; ist das Messergebnis der Mittelwert aus m gemessenen Einzelwerten, so gilt $s(y_{\text{mess}}) = s_V / \sqrt{m}$.

Einrechnung in die Messunsicherheit

Ist eine Korrektur gemäß den Gleichungen (4.9) - (4.11) zweifelhaft, oder sind die festgestellten Abweichungen nicht signifikant, so wird von einer Korrektur abgesehen. Stattdessen wird die an den Referenzobjekten festgestellte Abweichung als Unsicherheitsbeitrag in die Messunsicherheit eingerechnet. Wie in Abschnitt 4.2, wird hierfür die Vorgehensweise nach Lira und Wöger empfohlen:

$$u(y_{\text{unkorr}})^2 = u(y_{\text{korr}})^2 + (y_{\text{korr}} - y_{\text{mess}})^2 \quad (4.13)$$

Die weitere Rechnung erfolgt durch Einsetzen der Gleichungen (4.9) - (4.12). Aufgrund der Komplexität wird auf die Angabe der resultierenden Gleichung für die solchermaßen abgeschätzte Messunsicherheit $u(y_{\text{unkorr}})$ verzichtet.

4.3.2 Ausgleichsrechnung

Bei einem weiten Messbereich und erheblicher Variationsbreite der Messobjekte reichen häufig auch zwei Referenzobjekte nicht aus. Um systematische Abweichungen sicher einzuschätzen und ggf. auszugleichen, sind bereits im linearen Fall mindestens drei Referenzobjekte erforderlich. Die Auswertung eines derartigen Mehrfach-Vergleichs folgt einem anderen Muster als in den vorangegangenen Fällen, weil die Anzahl der Referenzobjekte größer ist als die Anzahl der ggf. zu bestimmenden Korrekturgrößen (Ausgleichsrechnung statt Interpolation).

Im Folgenden wird als häufigster Anwendungsfall die Ermittlung einer "Korrekturgeraden" behandelt. Insbesondere bei sehr ausgedehnten Messbereichen können jedoch nichtlineare "Korrekturkurven" erforderlich sein. Ihre rechnerische Behandlung erfordert Methoden der nichtlinearen Regression, z. B. unter Verwendung von Polynomansätzen.

Bei der N-Punkt-Prozedur werden die N Referenzobjekte gemessen und der resultierende Datensatz, bestehend aus den Referenzwerten und den entsprechenden Messwerten, mittels linearer Regression ausgewertet. Für diese Auswertung kann in den meisten Fällen die Standardform der Methode der kleinsten Quadrate verwendet werden. Die Voraussetzungen hierfür sind im Wesentlichen folgende:

- Die Unsicherheit der Referenzwerte ist erheblich kleiner als die Streuung der Messwerte;
- Die Streuung der Messwerte ist im betrachteten Messbereich annähernd konstant;
- Messwerte aus wiederholten Messungen an demselben Referenzobjekt sind annähernd normalverteilt.

Für die Rechnung werden die Referenzwerte als Werte der unabhängigen Variablen x_{ref} und die Messwerte als Werte der abhängigen Variablen x_{mess} , gewählt. Das Ergebnis ist eine Ausgleichsgerade, d. h. eine lineare Funktion

$$x_{\text{mess}} = \alpha + \beta \cdot x_{\text{ref}} \quad (4.14)$$

Bei Abwesenheit systematischer Abweichungen gilt $\alpha = 0$ und $\beta = 1$. Bei $\alpha \neq 0$ liegt eine additive systematische Abweichung vor, bei $\beta \neq 1$ eine multiplikative systematische Abweichung. Die Korrektur der festgestellten systematischen Abweichungen wird dann ggf. an vergleichbaren Messobjekten gemäß folgender Gleichung vorgenommen:

$$y_{\text{korr}} = \frac{y_{\text{mess}} - \alpha}{\beta} \quad (4.15)$$

Die Unsicherheit der korrigierten Messgröße wird nach den in Anhang A.2 behandelten Verfahren aus der Standardabweichung $s(y_{\text{mess}})$ und der Unsicherheit der Parameter der Ausgleichsgeraden, $u(\alpha)$, $u(\beta)$ sowie der Kovarianz $u(\alpha, \beta)$ berechnet.

5 Abschätzung von Messunsicherheiten mit Ringversuchsdaten

5.1 Validierungsringversuche

Bei Normverfahren werden Richtigkeit und Präzision in der Regel im Ringversuch ermittelt (siehe ISO 5725-2). Von den dabei ermittelten Verfahrenskennwerten eignet sich die sog. „Vergleichsstandardabweichung“ (s_R) als Schätzwert für die Messunsicherheit. Da sie bereits systematische Effekte durch unterschiedliche Arbeitsweisen der beteiligten Laboratorien beinhaltet, ist eine zusätzliche Einrechnung systematischer Unsicherheitsbeiträge i.A. nicht erforderlich.

Anmerkung: Eine „Wiederholstandardabweichung“, ermittelt im Ringversuch (s_r) oder laborintern unter möglichst identischen Bedingungen ist allein i.A. kein geeigneter Schätzwert für eine Messunsicherheit, da sie wesentliche Unsicherheitsbeiträge nicht erfasst.

In der internationalen technischen Spezifikation ISO/TS 21748 *Guide to the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation* vom Mai 2003 sind die genauen Bedingungen festgelegt, unter denen ein Laboratorium die Vergleichsstandardabweichung s_R als Schätzwert für die Messunsicherheit der mit dem festgelegten Verfahren erhaltenen Messergebnisse verwenden kann. Im Wesentlichen muss das Labor nachweisen,

- (a) dass es normkonform arbeitet, insbesondere
- (b) dass die Messbedingungen und Messobjekte mit denen im Ringversuch übereinstimmen, und
- (c) dass für seine Implementierung des Messverfahrens die Richtigkeit und die Präzision mit den Ringversuchsdaten verträglich sind.

Forderung (c) bedeutet, dass das Labor eine Überprüfung der Richtigkeit und Präzision (siehe Abschnitt 4) hinsichtlich Verträglichkeit mit den Ringversuchsdaten s_r und s_R vornehmen muss. Zu diesem Zweck kann das Labor z. B. Mehrfachmessungen an einem geeigneten Referenzobjekt durchführen. Sind n_{Lab} die Anzahl der Messungen, s_{Lab} die Standardabweichung der Messreihe und $\Delta = \bar{x}_{\text{Lab}} - x_{\text{ref}}$ die festgestellte Abweichung des Mittelwertes der Messreihe vom Referenzwert, so ist Verträglichkeit gegeben wenn:

$$s_{\text{Lab}} \approx s_r \quad \text{und}$$

$$|\Delta| \leq 2 \sqrt{\frac{s_r^2}{n_{\text{Lab}}} + (s_R^2 - s_r^2)}.$$

5.2 Ringversuche zur Eignungsprüfung (Proficiency testing)

Hat das Laboratorium sich mit einem Prüfverfahren im Rahmen von Eignungsprüfungen erfolgreich an Ringversuchen beteiligt, so kann es die Ergebnisse verwenden, um seine Messunsicherheit abzuschätzen. Hierfür werden im Folgenden einige einfache Ansätze beschrieben. Eine ähnliche Vorgehensweise wird in dem NORDTEST Technical Report 537 *Handbook for calculation of measurement uncertainty in environmental laboratories* beschrieben.

Alternativ bieten Ringversuche zur Eignungsprüfung die Möglichkeit, eine anderweitig ermittelte Messunsicherheit zu überprüfen.

a) Referenzwert gegeben

Vom Veranstalter des Ringversuchs ist ein Referenzwert (Sollwert) x_{ref} einschließlich Standardunsicherheit $u(x_{\text{ref}})$ vorgegeben. Das Labor hat n_{Lab} Ergebnisse mit dem Mittelwert \bar{x}_{Lab} und der Standardabweichung s_{Lab} vorgelegt. Aus diesen Daten werden analog Abschnitt 4.2 folgende Kenngrößen berechnet:

- die Differenz Δ_{RV} zwischen dem Mittelwert der Messwerte des Labors und dem Referenzwert

$$\Delta_{\text{RV}} = \bar{x}_{\text{Lab}} - x_{\text{ref}}$$

- die Standardunsicherheit $u(\Delta_{\text{RV}})$ dieser Differenz

$$u(\Delta_{\text{RV}}) = \sqrt{\frac{s_{\text{Lab}}^2}{n_{\text{Lab}}} + u(x_{\text{ref}})^2}$$

Diese Kenngrößen kommen grundsätzlich für eine Korrektur gemäß Abschnitt 4.2, Gleichung (4.2) bzw. mit der Wiederfindung $Q_{RV} = \bar{x}_{Lab} / x_{ref}$ gemäß Gleichung (4.3) in Betracht. Wegen der fraglichen Übertragbarkeit der in der „Momentaufnahme“ eines einzigen Ringversuchs festgestellten systematischen Abweichung auf künftige Messungen wird jedoch empfohlen, im Zweifelsfall von einer Korrektur abzusehen und stattdessen die festgestellte Abweichung wie in Abschnitt 4.2 nach Gleichung (4.7) bzw. (4.8) als Unsicherheitsbeitrag einzurechnen.

Anmerkung: Hat ein Labor die Messunsicherheit für sein Verfahren bereits anderweitig ermittelt, so können Ringversuchsergebnisse zur Überprüfung dieser Messunsicherheit verwendet werden. Dabei wird geprüft, ob die Differenz Δ_{RV} zwischen dem Mittelwert der Messwerte des Labors und dem Referenzwert signifikant ist. Zu diesem Zweck wird die Messunsicherheit $u(\bar{x}_{Lab})$ des Mittelwertes benötigt. Hierbei sind systematische Unsicherheitsbeiträge zu beachten, weshalb der bekannte Faktor $1/\sqrt{n}$ dafür in der Regel nicht anwendbar ist, siehe Anhang A.5.

Hat das Labor an mehreren Ringversuchen (für vergleichbare Messungen) teilgenommen, so sollten die Schätzwerte der Messunsicherheit aus den einzelnen Ringversuchen miteinander verglichen und, falls verträglich, kombiniert werden. Die Kombination erfolgt durch Pooling (quadratische Mittelung) der betreffenden Unsicherheitsdaten. Diese Vorgehensweise wird in dem NORDTEST Handbook verwendet; dabei werden Ergebnisse aus mindestens 6 vergleichbaren Ringversuchen vorausgesetzt.

b) Referenzwert nicht gegeben

In diesem Fall wird behelfsweise der vom Veranstalter des Ringversuchs ermittelte Bezugswert – in der Regel der Mittelwert $\langle x_{RV} \rangle$ der Ergebnisse aller Teilnehmer, ggf. ausreißerbereinigt – anstelle eines Referenzwertes verwendet, als dessen Unsicherheit die Standardabweichung $s_{RV} / \sqrt{n_{RV}}$ dieses Mittelwertes. Hierbei ist s_{RV} die Standardabweichung der Labormittelwerte, die zum Bezugswert beitragen, n_{RV} ihre Anzahl. Aus diesen Daten werden folgende Kenngrößen der (eingeschränkten) Richtigkeitskontrolle berechnet:

- die Differenz Δ_{RV} zwischen dem Mittelwert der Analyseergebnisse des Labors und dem Bezugswert

$$\Delta_{RV} = \bar{x}_{Lab} - \langle x_{RV} \rangle$$

- die Standardunsicherheit $u(\Delta_{RV})$ dieser Differenz

$$u(\Delta_{RV}) = \sqrt{\frac{s_{Lab}^2}{n_{Lab}} + \frac{s_{RV}^2}{n_{RV}}}$$

Auf Grundlage dieser Kenngrößen wird dann wie in Fall a) verfahren, wobei hier aufgrund der mangelnden Information über die Richtigkeit des Bezugswertes von einer Korrektur auf jeden Fall abgesehen werden sollte. Stattdessen wird die festgestellte Differenz zwischen dem Mittelwert der Analyseergebnisse des Labors und dem Bezugswert wie in Abschnitt 4.2 nach Gleichung (4.7) bzw. (4.8) als Unsicherheitsbeitrag eingerechnet.

c) Verfahrensspezifischer Ringversuch

Verfahrensspezifische Ringversuche werden häufig nach DIN ISO 5725-2 ausgewertet. Dann kann die Vergleichsstandardabweichung unter geeigneten Bedingungen (siehe 5.1) direkt als Schätzwert der Messunsicherheit verwendet werden. Sie beinhaltet sowohl zufällige als auch systematische Einflüsse, soweit diese auf unterschiedliche Arbeitsweisen der beteiligten Laboratorien zurückzuführen sind, nicht jedoch ggf. durch das Verfahren bedingte systematische Abweichungen.

Bei anderen Auswertungen ist anstelle der Vergleichsstandardabweichung die (ggf. ausreißerbereinigte) Standardabweichung der Einzelwerte aller Teilnehmer zu verwenden. Steht nur die Standardabweichung der Mittelwerte aller Teilnehmer zur Verfügung, so kann diese mit der im Labor ermittelten Standardabweichung unter laborinternen Vergleichsbedingungen kombiniert werden.

5.3 Ringversuche zur Referenzmaterialzertifizierung

Hat das Labor mit dem betreffenden Verfahren erfolgreich an einem Zertifizierungs-Ringversuch teilgenommen, so kann die Differenz zwischen dem Ergebnis des Labors – in der Regel der Mittelwert einer Reihe wiederholter Analysen – und dem zertifizierten Wert zur Bewertung der Richtigkeit verwendet werden.

Die Vorgehensweise ist grundsätzlich dieselbe wie in Abschnitt 5.2, wobei der zertifizierte Wert und dessen Unsicherheit die Rolle des Referenzwertes und seiner Unsicherheit übernehmen. Die Abweichung vom zertifizierten Wert wird auf jeden Fall als akzeptabel bewertet, wenn das Ergebnis des Labors in die Berechnung des zertifizierten Wertes einbezogen wird, aber auch größere Abweichungen können in Hinblick auf Routineanwendungen akzeptabel sein. Hinsichtlich der Berücksichtigung der festgestellten Abweichung kommen grundsätzlich sowohl Korrektur als auch Einrechnung in die Unsicherheit in Betracht, wobei im Zweifelsfall die Letztere vorzuziehen ist.

6 Hybridstrategien zur Ermittlung von Messunsicherheiten

Treten bei einer Messung Einflussfaktoren in Erscheinung, die bei der Validierung des Prüfverfahrens (laborintern oder im Ringversuch) nicht berücksichtigt wurden, so muss der in der Validierung ermittelte Schätzwert der Messunsicherheit entsprechend ergänzt werden. In diesem Sinne wird bei einer Hybridstrategie, soweit möglich, der kombinierte Effekt von Einflussfaktoren durch Daten aus Validierungsstudien erfasst, und notwendige Ergänzungen werden durch Modellierung der Einflüsse auf das Ergebnis und Unsicherheitsfortpflanzung (quadratische Addition) vorgenommen. Diese Vorgehensweise ist darauf angelegt, die Nutzung vorhandener Daten aus Validierungsuntersuchungen mit der Flexibilität der modellbasierten Ermittlung von Unsicherheitsbeiträgen zu verbinden.

Soll z.B. für ein Normverfahren die im Ringversuch ermittelte Vergleichsstandardabweichung s_R als Schätzwert für die Messunsicherheit verwendet werden (siehe Abschnitt 5.1), und weichen die Messbedingungen oder die Messobjekte erheblich von den im Ringversuch zugrundegelegten Vergleichsbedingungen oder Messobjekten ab, so muss der Effekt der Abweichungen abgeschätzt und mit der Vergleichsstandardabweichung kombiniert werden. Hierfür gilt die folgende summarische Gleichung:

$$u_{\text{komb}} = \sqrt{s_R^2 + \sum u_{\text{sonst}}^2}$$

Umgekehrt kann ein Unsicherheitsbudget durch Vergleich der errechneten Messunsicherheit mit der Streuung der Ergebnisse von Mehrfachprüfungen, bei Verfügbarkeit geeigneter Referenzobjekte auch unter Einbeziehung systematischer Abweichungen, auf Vollständigkeit geprüft und ggf. ergänzt werden.

Eine wesentliche Voraussetzung für alle derartigen Vorhaben ist eine detaillierte Betrachtung der Einflussgrößen (Unsicherheitsquellen) mit dem Ziel, diejenigen Einflussgrößen, deren Unsicherheitsbeiträge in einer gegebenen Verfahrenskenngröße, z. B. einer Vergleichsstandardabweichung s_R erfasst sind, von den damit nicht erfassten und daher anderweitig zu berücksichtigenden Einflussgrößen zu unterscheiden.

Anmerkung: Die in verschiedenen Beispielen des EURACHEM/CITAC-Guides Quantifying Uncertainty in Analytical Measurement praktizierte Vorgehensweise, auf ein detailliertes Unsicherheitsbudget eine Standardabweichung für die gesamte Verfahrensstreuung aufzuschlagen, ist kontraproduktiv in Hinblick auf einen Vergleich zwischen Unsicherheitsbudget und Präzisionsdaten und überschätzt zwangsläufig die Messunsicherheit.

7 Angabe und Dokumentation von Messunsicherheiten

Berechnete oder geschätzten Werte der Messunsicherheit werden zusammen mit dem Wert der Ergebnisgröße y entweder als Standardunsicherheit $u(y)$ oder als erweiterte Unsicherheit $U(y) = k \times u(y)$ angegeben. Wird die erweiterte Standardunsicherheit angegeben, so ist auch der benutzte Erweiterungsfaktor zu nennen und, soweit möglich, das zugehörige geschätzte Vertrauensniveau.

Sowohl die Standardunsicherheit als auch die erweiterte Standardunsicherheit können direkt oder, durch den Betrag der Ergebnisgröße dividiert, als Relativwert (z.B. in Prozent) angegeben werden.

Höchstwerte der Messunsicherheit werden zusammen mit dem Wert der Ergebnisgröße y als Zahlenwert $|\Delta y|_{\max}$ oder in Form einer Ungleichung für die anzunehmende Messabweichung $|\Delta y|$ angegeben. Die Angabe muss Verwechslungen mit Standardunsicherheiten oder erweiterten Standardunsicherheiten ausschließen.

Die Ermittlung der Messunsicherheit muss so detailliert und vollständig dokumentiert werden, dass sie in ihren wesentlichen Schritten nachvollzogen werden kann. Wurden als wesentlich eingeschätzte Unsicherheitsbeiträge nicht berücksichtigt, so muss dieser Umstand angegeben und begründet werden.

Literatur

In dieser Literaturliste sind die grundlegenden Dokumente (Normen und Leitfäden) aufgeführt, die im Hauptteil herangezogen werden. Angaben zu anderen Veröffentlichungen erfolgen direkt im Text.

- [1] *Guide to the Expression of Uncertainty in Measurement*
1st corr. Edition, ISO, Geneva 1995, ISBN 92-67-10188-9
deutsche Übersetzung: *Leitfaden zur Angabe der Unsicherheit beim Messen*
Beuth-Verlag, Berlin 1995, ISBN 3-410-13405-0
- [2] *International Vocabulary of Basic and General Terms in Metrology*
2nd Edition, ISO, Geneva 1993, ISBN 92-67-10188-9
deutsche Übersetzung: *Internationales Wörterbuch der Metrologie*
Beuth-Verlag, Berlin 1994, ISBN 3-410-11735-0
- [3] DIN 1319-1, *Grundlagen der Messtechnik – Teil 1: Grundbegriffe*
- [4] DIN 1319-4, *Grundlagen der Messtechnik – Teil 4: Auswertung von Messungen, Messunsicherheit*
- [5] DIN 55350-13, *Begriffe der Qualitätssicherung und Statistik – Teil 13: Begriffe zur Genauigkeit von Ermittlungsverfahren und Ermittlungsergebnissen*
- [6] ISO 3534-1, *Statistics – Vocabulary and symbols – Part 1: Probability and general statistical terms*
- [7] DIN ISO 5725-2, *Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen – Teil 2: Grundlegende Methode für die Ermittlung der Wiederhol- und Vergleichpräzision eines vereinheitlichten Messverfahrens*
- [8] ISO/TS 21748, *Guide to the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*
- [9] EURACHEM/CITAC GUIDE: *Quantifying Uncertainty in Analytical Measurement*
2nd Edition, EURACHEM / CITAC 2000
deutsche Übersetzung: *Ermittlung der Messunsicherheit bei analytischen Messungen*
EURACHEM-D 2002
- [10] NORDTEST Technical Report 537, *Handbook for calculation of measurement uncertainty in environmental laboratories*, 2nd Edition, NORDTEST 2004

Anhang

Dieser Anhang beschreibt häufig vorkommende Unsicherheitsquellen (A.1) und häufig vorkommende Datenauswertungsverfahren (A.2 – A.6).

In den Abschnitten A.2 bis A.6 wird bei der formelmäßigen Darstellung unterschieden zwischen Größen/Variablen X, Y, etc. und ihren Werten x, y, etc.

A.1 Häufig vorkommende Unsicherheitsquellen

In der Regel sind bei Messungen/Prüfungen eine Vielzahl von Unsicherheitsbeiträgen zu betrachten, die je nach Größe und Häufigkeit ihres Auftretens im Gesamtbudget (kombinierte Unsicherheit) zu berücksichtigen sind. Im Folgenden sind häufig vorhandene Unsicherheitsquellen aufgelistet. Sie lassen sich in vier Gruppen zusammenfassen:

1. Unsicherheiten, die mit der Probennahme/Probenaufbereitung zusammenhängen
 - a) Entnahme von Proben, die das Untersuchungsobjekt nur eingeschränkt repräsentieren
 - b) Kontamination/Veränderung von Proben bei ihrer Entnahme
 - c) Kontamination/Veränderung von Proben bei ihrer physikalischen Aufbereitung
 - d) Homogenisierung (Unvollständigkeit)
 - e) Kontamination/Veränderung von Proben bei ihrer Lagerung
 - f) Aufschluss (Unvollständigkeit, Kontamination, Interferenzen)
 - g) Chemische Probenvorbereitung/Trennungstechniken (Unvollständigkeit, Kontamination)

2. Unsicherheitsbeiträge, die mit den Eigenschaften des untersuchten Objekts zusammenhängen
 - a) Rauschen, Instabilität des untersuchten Objekts (zeitl. Änderung der Kennwerte)
 - b) Veränderung oder Alterung des untersuchten Objekts (zeitl. Änderung der Kennwerte)
 - c) Inhomogenität/Ungleichförmigkeit des untersuchten Objekts (örtl. Änderung der Kennwerte)
 - d) Matrixeffekte/Wechselwirkungen

3. Unsicherheitsbeiträge, die dem angewandten Mess-/Prüfverfahren anhaften
 - a) unzureichende Realisierung bzw. Definition der Mess-/Ergebnisgröße (Näherungen /Idealisierungen/Hypothesen)
 - b) Unsicherheit von Verfahrensparametern (z. B. auch Umgebungsbedingungen) und der berücksichtigten Einflussgrößen
 - c) Vernachlässigte Einflussgrößen (z. B. Temperatur, Luftdruck, magn. Feldstärke)
 - d) Begrenzte Instrumentenauflösung, Unschärfe bzw. Unsicherheit der Lage von Diskriminatorschwellen etc.
 - e) Nachweisgrenzen, begrenzte Empfindlichkeit
 - f) Rauschen und Drift der Apparatur
 - g) Zufällige Störeinflüsse (wie Störfelder etc.)
 - h) Unzureichende Impedanzanpassung und Übertragung/Wandlung der Messgröße

- i) Totzeiten der Apparatur (Fehler durch Koinzidenzen)
 - j) Dynamik der Apparatur (Frequenzgang/Überschwingen/Resonanzen)
 - k) Unterschiedliche Wahrnehmung/Visualisierung der Messgrößen
 - l) Auswertung, Rechengenauigkeit etc.
 - m) Unsicherheit ermittelt aus Ringversuchen
4. Unsicherheit der Referenzwerte, die der Messung/Prüfung zugrunde liegen
- a) Unsicherheit zertifizierter Werte/Kalibrierwerte
 - b) Drift/Alterung von Referenzwerten/Referenzmaterialien
 - c) importierte Werte aus der Literatur (Tabellenwerke, Konstanten, ggf. aktuelle Veröffentlichungen, etc.)
 - d) Unsicherheit ermittelt aus Ringversuchen

Die einzelnen Unsicherheitsbeiträge sind nicht notwendig unabhängig voneinander. Sie haben teilweise zufälligen und teilweise systematischen Charakter. Zufällige Einflüsse führen zur Variation der Einzelergebnisse bei wiederholten Messungen unter vergleichbaren Messbedingungen. Die entsprechenden Unsicherheitsbeiträge können mit empirisch statistischen Verfahren (z. B. als empirische Standardabweichung eines Mittelwertes) abgeschätzt werden (Typ A). Unsicherheitsbeiträge, die von systematischen Einflüssen herrühren, müssen geeignet geschätzt bzw. im Falle importierter Werte aus den Angaben ermittelt werden (Typ B).

A.2 Unsicherheit bei linearer Kalibrierung

A.2.1 Allgemeines

Bei der linearen Kalibrierung im Sinne dieses Abschnitts wird ein linearer Zusammenhang zwischen den Werten einer Größe X und den Werten einer zweiten Größe Y ermittelt, z. B. zwischen

- Zugfestigkeit und Härte von Stahlwerkstoffen;
- Thermospannung und Temperatur bei Thermoelementen;
- Response und Analytgehalt bei einem instrumentellen Analysenverfahren;
- Messergebnissen eines Prüfverfahrens und entsprechenden Referenzwerten bei der Validierung von Prüfverfahren.

Der ermittelte Zusammenhang wird dazu verwendet, für gegebene Werte der einen Größe (Eingangsgröße, unabhängige Variable) die entsprechenden Werte der anderen Größe (Ausgangsgröße, abhängige Variable) zu berechnen. Je nachdem, ob X oder Y als Eingangsgröße gewählt wird, lässt sich ein linearer Zusammenhang in zwei verschiedenen Formen ($Y = \alpha + \beta X$ bzw. $X = \gamma + \delta Y$) darstellen.

Im Folgenden wird davon ausgegangen, dass durch Kalibrierung Y als Funktion von X bestimmt wird,

$$Y = \alpha + \beta X \quad (\text{A.2.1})$$

d. h., dass die Werte der Parameter α (Achsenabschnitt) und β (Steigung) ermittelt werden.

Bei der Anwendung des durch Kalibrierung ermittelten linearen Zusammenhanges ist zwischen zwei Fällen zu unterscheiden:

Direkte Kalibrierung: Y ist die Ergebnisgröße, deren Unsicherheit zu ermitteln ist, X ist Eingangsgröße von Y.

Indirekte Kalibrierung: X ist die Ergebnisgröße, deren Unsicherheit zu ermitteln ist, Y ist Eingangsgröße von X.

Bei der direkten Kalibrierung wird der ermittelte Zusammenhang $Y = \alpha + \beta X$ direkt zur Berechnung der Ergebnisgröße Y verwendet. Nach den Regeln der Unsicherheitsfortpflanzung [siehe Abschnitt 3.3, Gleichung (3.11)] setzt sich die Unsicherheit $u(Y)$ der Ergebnisgröße wie folgt aus der Unsicherheit $u(X)$ der Eingangsgröße X, den Unsicherheiten $u(\alpha)$, $u(\beta)$ der Parameter α , β sowie ihrer Kovarianz $u(\alpha, \beta)$ zusammen:

$$u(Y)^2 = \beta^2 u(X)^2 + u(\alpha)^2 + X^2 u(\beta)^2 + 2Xu(\alpha, \beta) \quad (\text{A.2.2})$$

Bei der indirekten Kalibrierung wird der ermittelte Zusammenhang zwecks Berechnung der Ergebnisgröße X invertiert.

$$X = \frac{Y - \alpha}{\beta} \quad (\text{A.2.3})$$

Die Unsicherheit $u(X)$ der Ergebnisgröße ergibt sich wie folgt:

$$u(X)^2 = \frac{u(Y)^2 + u(\alpha)^2 + X^2 u(\beta)^2 + 2Xu(\alpha, \beta)}{\beta^2} \quad (\text{A.2.4})$$

Anmerkung 1: Werden die Parameter α und β wie üblich gemeinsam bestimmt, so sind diese Größen korreliert, denn sie hängen von denselben (unsicherheitsbehafteten) Größen, den Kalibrierdaten (x_1, y_1) , (x_2, y_2) , ..., (x_K, y_K) ab. Die Kovarianz $u(\alpha, \beta)$ liefert i.A. einen wesentlichen Beitrag zur Ergebnisunsicherheit und darf deshalb nicht außer Betracht bleiben.

Anmerkung 2: Die nach Gleichung (A.2.2) bzw. (A.2.4) ermittelte Unsicherheit ist nur dann vertrauenswürdig, wenn die Gültigkeit des Regressionsmodells geprüft und bestätigt ist. Diese Prüfung kann durch statistische Analyse der Reststreuung der Kalibrierpunkte um die Ausgleichsgerade (z.B. F-Test) durchgeführt werden.

In den nachfolgenden Abschnitten wird die Bestimmung der Parameter α , β und die Ermittlung ihrer Unsicherheiten, d. h. der Größen $u(\alpha)$, $u(\beta)$ und $u(\alpha, \beta)$ behandelt.

A.2.2 Bestimmung von Achsenabschnitt und Steigung

Die Parameter α (Achsenabschnitt) und β (Steigung) werden mittels statistischer Auswertungsverfahren aus entsprechenden Kalibrierdatensätzen $\{(x_1, y_1), (x_2, y_2), \dots, (x_K, y_K)\}$ gewonnen. Theoretisch genügen zwei Punkte (x_1, y_1) , (x_2, y_2) zur Festlegung einer Geraden. Um eventuelle Messabweichungen festzustellen und soweit möglich auszugleichen, wird jedoch in der Regel mehr als diese Minimalzahl von Kalibrierpunkten (x_i, y_i) ($i = 1, 2, \dots, K$) zugrundegelegt und mittels linearer Regression Achsenabschnitt und Steigung einer Ausgleichsgeraden bestimmt. Zu diesem Zweck wird in den meisten Fällen die Standard-version der Methode der kleinsten Quadrate verwendet. Nach dieser Methode werden die Parameter α und β so bestimmt, dass die "Fehlerquadratsumme" $S = \sum [y_i - (\alpha + \beta x_i)]^2$ minimal ist. Die Lösung dieses Optimierungsproblems ist durch die folgenden, allgemein bekannten Formeln für die Schätzwerte des Achsenabschnitts α und der Steigung β gegeben.

$$\beta = \frac{Q_{xy}}{Q_{xx}} \quad (\text{A.2.5})$$

$$\alpha = \bar{y} - \beta \bar{x} \quad (\text{A.2.6})$$

Hierbei bedeuten

$\bar{x} = [\sum x_i]/K$ Mittelwert der x_1, x_2, \dots, x_K ;

$\bar{y} = [\sum y_i]/K$ Mittelwert der y_1, y_2, \dots, y_K ;

$Q_{xx} = \sum (x_i - \bar{x})^2$ Summe der Abweichungsquadrate in x;

$Q_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$ Summe der Abweichungsprodukte in x und y.

Die Summen laufen über $i = 1, 2, \dots, K$.

Die Standardversion der Methode der kleinsten Quadrate ist im Wesentlichen an folgende Voraussetzungen geknüpft:

- Die Unsicherheit der Werte der unabhängigen Variablen X ist gegenüber der Streuung der Werte der abhängigen Variablen Y vernachlässigbar;
- Die Streuung der Werte von Y ist im betrachteten Wertebereich konstant;
- Die Werte von Y zu festem X sind normalverteilt.

Unter diesen Voraussetzungen liefert die Methode optimale Schätzwerte für die Parameter α und β . Sie ist jedoch auch bei mäßigen Abweichungen von diesen Voraussetzungen anwendbar (siehe einschlägige statistische Lehrbücher).

Sind die o.g. Voraussetzungen auch nicht annähernd gegeben, so müssen andere Regressionsmethoden verwendet werden. Für viele Anwendungsfälle sind Abwandlungen der Standardmethode der kleinsten Quadrate geeignet, z.B.

- Die gewichtete Methode der kleinsten Quadrate, wenn die Streuung der Werte von Y im betrachteten Wertebereich stark variiert. Hierbei werden die Abweichungsquadrate mit den inversen Varianzen der Kalibrierwerte gewichtet.
- Die verallgemeinerte Methode der kleinsten Quadrate, wenn die Unsicherheit der Werte von X und die Streuung der Werte von Y vergleichbar sind. Hierbei wird die Summe der Abweichungsquadrate in X und Y, ggf. mit den inversen Varianzen der Kalibrierwerte gewichtet, minimiert.
- Robuste Regressionsmethoden, wenn erhebliche Abweichungen von der Normalverteilung oder Kontaminationen durch Ausreißer vorliegen. Hierbei wird z. B. mit Medianen und Quantilen anstelle von Mittelwerten und Standardabweichungen gerechnet.

Für nähere Informationen zu den vorstehend genannten Regressionsmethoden muss auf die Spezialliteratur zum Thema Regression verwiesen werden, z. B. den Abschnitt "Modelling of data" in: Press, W. H. et al.: Numerical Recipes in Fortran, 2. Auflage, Cambridge University Press (1992)

A.2.3 Ermittlung der Unsicherheit von Achsenabschnitt und Steigung

Für die Ermittlung der Unsicherheit der Parameter α und β gibt es zwei grundsätzlich verschiedene Methoden:

- Statistische Analyse der Streuung der Kalibrierpunkte;
- Fortpflanzung der Unsicherheit der Kalibrierdaten.

Diese beiden Methoden werden im Folgenden beschrieben.

Statistische Analyse

Die statistische Analyse macht nur implizit Gebrauch von der Unsicherheit der Kalibrierdaten. Bei dieser Methode wird im ersten Schritt geprüft, ob die Streuung der Kalibrierpunkte um die errechnete Ausgleichsgerade annähernd konstant ist. Ist dies der Fall, so wird als Grundgröße der weiteren Rechnung die sog. Reststandardabweichung s_R der Messwerte y_i gemäß folgender Gleichung bestimmt:

$$s_R^2 = \sum \frac{[y_i - (\alpha + \beta x_i)]^2}{K - 2} \quad (\text{A.2.7})$$

Mit dieser Grundgröße lassen sich die Standardunsicherheiten $u(\alpha)$, $u(\beta)$ und die Kovarianz $u(\alpha, \beta)$ wie folgt ausdrücken.

$$u(\beta)^2 = \frac{s_R^2}{Q_{xx}} \quad (\text{A.2.8})$$

$$u(\alpha)^2 = s_R^2 \left[\frac{1}{K} + \frac{\bar{x}^2}{Q_{xx}} \right] \quad (\text{A.2.9})$$

$$u(\alpha, \beta) = -s_R^2 \frac{\bar{x}}{Q_{xx}} \quad (\text{A.2.10})$$

Zur Berechnung der Ergebnisunsicherheit bei direkter Kalibrierung gemäß (A.2.2)

$$u(Y)^2 = \beta^2 u(X)^2 + u(\alpha)^2 + X^2 u(\beta)^2 + 2Xu(\alpha, \beta)$$

oder bei indirekter Kalibrierung gemäß (A.2.4)

$$u(X)^2 = \frac{u(Y)^2 + u(\alpha)^2 + X^2 u(\beta)^2 + 2Xu(\alpha, \beta)}{\beta^2}$$

wird nun lediglich noch die Unsicherheit der jeweiligen Eingangsgröße, $u(X)$ bzw. $u(Y)$ benötigt.

Bei der direkten Kalibrierung wird in der Regel die Unsicherheit $u(X)$ der Eingangsgröße als vernachlässigbar klein angenommen. Damit erhält man für die Standardunsicherheit der Ergebnisgröße Y

$$u(Y)^2 = s_R^2 \left[\frac{1}{K} + \frac{(X - \bar{x})^2}{Q_{xx}} \right] \quad (\text{A.2.11})$$

Bei der indirekten Kalibrierung ist hingegen Y die Eingangsgröße. Die Standardunsicherheit $u(Y)$ kann durch die Reststreuung s_R der Kalibrierpunkte um die Ausgleichsgerade geschätzt werden. Hierbei ist zu beachten, ob die Eingangsgröße Y als Einzelwert oder als Mittelwert einer festgelegten Anzahl von Einzelwerten definiert ist. Es gilt

$$u(Y) = s_R \quad \text{wenn } Y \text{ ein Einzelwert ist;}$$

$$u(Y) = s_R / \sqrt{m} \quad \text{wenn } Y \text{ ein Mittelwert aus } m \text{ unabhängigen Einzelwerten ist.}$$

Damit erhält man für die Standardunsicherheit der Ergebnisgröße X im ersten Fall (Einzelwert)

$$u(X)^2 = \frac{s_R^2}{\beta^2} \left[1 + \frac{1}{K} + \frac{(X - \bar{x})^2}{Q_{xx}} \right] \quad (\text{A.2.12})$$

Im zweiten Fall (Mittelwert) erhält man hingegen

$$u(X)^2 = \frac{s_R^2}{\beta^2} \left[\frac{1}{m} + \frac{1}{K} + \frac{(X - \bar{x})^2}{Q_{xx}} \right] \quad (\text{A.2.13})$$

Die Ergebnisunsicherheit $u(Y)$ nach Gleichung (A.2.11) bzw. $u(X)$ nach Gleichung (A.2.12) oder (A.2.13) erreicht am Rand des Kalibrierintervalls ihren Maximalwert. Bei annähernd äquidistanten Kalibrierwerten x_1, x_2, \dots, x_K kann dieser Maximalwert wie folgt abgeschätzt werden:

$$u(Y)^2 < s_R^2 \cdot \frac{4}{K} \quad (\text{A.2.11a})$$

$$u(X)^2 < \frac{s_R^2}{\beta^2} \left(1 + \frac{4}{K} \right) \quad (\text{A.2.12a})$$

$$u(X)^2 < \frac{s_R^2}{\beta^2} \left(\frac{1}{m} + \frac{4}{K} \right) \quad (\text{A.2.13a})$$

Diese Maximalwerte können verwendet werden, um den Beitrag der Kalibrierung zur Gesamtunsicherheit eines Verfahrens abzuschätzen, oder um für diese Gesamtunsicherheit eine obere Schranke zu ermitteln.

Für die in Abschnitt A.2.2 erwähnten Varianten der Standardmethode der kleinsten Quadrate können unter bestimmten Voraussetzungen entsprechende Rechenformeln verwendet werden.

Bei der Literaturrecherche ist die in der Regel abweichende Bezeichnungsweise zu berücksichtigen: Anstelle von $u(x)^2$ wird zumeist $\text{var}(x)$, gelegentlich auch $s(x)^2$ verwendet, anstelle von $u(x,y)$ zumeist $\text{cov}(x,y)$, gelegentlich auch $s(x,y)$.

Unsicherheitsfortpflanzung

Bei der im vorigen Abschnitt behandelten Verfahrensweise wird die Ergebnisunsicherheit aus der Streuung der Kalibrierpunkte um die Ausgleichsgerade abgeleitet. Die Unsicherheit der Kalibrierdaten geht in diese Rechnung nicht ein. Im Gegensatz dazu wird bei der in diesem Abschnitt beschriebenen Verfahrensweise die Ergebnisunsicherheit auf die Unsicherheit der Kalibrierdaten zurückgeführt.

Bei der Unsicherheitsfortpflanzung wird die Unsicherheit von Achsenabschnitt und Steigung der Kalibriergeraden sowie die Kovarianz dieser beiden Größen aus den Unsicherheiten $u(x_i)$, $u(y_i)$ der Koordinaten der Kalibrierpunkte sowie den ggf. zu berücksichtigenden Kovarianzen $u(x_i, x_j)$ berechnet. Hierfür gelten folgende Gleichungen.

$$u(\alpha)^2 = \sum_i \left(\frac{\partial \alpha}{\partial x_i} \right)^2 u(x_i)^2 + \sum_i \left(\frac{\partial \alpha}{\partial y_i} \right)^2 u(y_i)^2 + \sum_{i \neq j} \left(\frac{\partial \alpha}{\partial x_i} \right) \left(\frac{\partial \alpha}{\partial x_j} \right) u(x_i, x_j) \quad (\text{A.2.14})$$

mit einer entsprechenden Gleichung für $u(\beta)$ sowie

$$u(\alpha, \beta) = \sum_i \left(\frac{\partial \alpha}{\partial x_i} \right) \left(\frac{\partial \beta}{\partial x_i} \right) u(x_i)^2 + \sum_i \left(\frac{\partial \alpha}{\partial y_i} \right) \left(\frac{\partial \beta}{\partial y_i} \right) u(y_i)^2 + \sum_{i \neq j} \left(\frac{\partial \alpha}{\partial x_i} \right) \left(\frac{\partial \beta}{\partial x_j} \right) u(x_i, x_j) \quad (\text{A.2.15})$$

Der dritte Term in diesen Gleichungen kann in der Regel gleich Null gesetzt werden, denn die Kalibrierwerte x_i der unabhängigen Variablen werden in der Regel unabhängig voneinander bestimmt.

In der Praxis treten jedoch auch Fälle auf, in denen diese Unabhängigkeit nicht gegeben ist, z. B. bei Verdünnungsreihen von Kalibrierlösungen. Stammen zwei "Tochterlösungen" von derselben "Stammlösung" ab, so pflanzt sich die Unsicherheit der Zusammensetzung der Stammlösung gleichsinnig in den Tochterlösungen fort. Diese Korrelation muss durch entsprechende Kovarianzen berücksichtigt werden.

Die in den Gleichungen (A.2.14) und (A.2.15) auftretenden Empfindlichkeitskoeffizienten $(\partial\alpha/\partial x_i)$, $(\partial\alpha/\partial y_i)$, $(\partial\beta/\partial x_i)$, $(\partial\beta/\partial y_i)$ können nur in Ausnahmefällen durch Differentialrechnung bestimmt werden. Deshalb müssen diese Gleichungen in der Regel unter Verwendung numerischer Differentiationstechniken ausgewertet werden. Hierfür eignet sich das in Abschnitt A.4 beschriebene Verfahren.

A.3 Modellierung von Verfahrensschritten durch Wirkungsgrade und Inkremente

Für die Anwendung des Unsicherheitsfortpflanzungsgesetzes

$$u(Y)^2 = \sum_{i=1}^N c_i^2 u(X_i)^2 + 2 \sum_{i=1}^{N-1} \sum_{k=i+1}^N c_i \cdot c_k \cdot u(X_i, X_k) \quad (\text{A.3.1})$$

$$\text{mit } c_i = \left(\frac{\partial Y}{\partial X_i} \right) \quad (\text{A.3.2})$$

ist es erforderlich, die Ergebnisgröße Y als Funktion der relevanten Eingangsgrößen X_1, X_2, \dots, X_N zu beschreiben. Diese Funktionsdarstellung wird benötigt, um die Empfindlichkeitskoeffizienten $c_i = (\partial Y/\partial X_i)$ zu bestimmen.

Sind die Eingangsgrößen X_i Bausteine (konstitutive Größen) der Ergebnisgröße – z. B. Probenmasse und Probenvolumen für die Ergebnisgröße Dichte – und ist Y als mathematische Funktion der Eingangsgrößen gegeben, so ist die Bildung der Ableitungen $(\partial Y/\partial X_i)$ grundsätzlich kein Problem. Ist der Zusammenhang zwischen Ergebnisgröße und Eingangsgrößen durch einen Algorithmus gegeben, so können die Ableitungen mittels numerischer Verfahren berechnet werden (siehe Abschnitt A.4). Häufig jedoch sind die Unsicherheitsquellen Verfahrensschritte – z. B. Probenahme, Probenvorbereitung aber auch Korrektur festgestellter systematischer Abweichungen –, bei denen zunächst unklar ist, wie eine Beschreibung durch Eingangsgrößen angesetzt werden kann. Im Folgenden wird die Beschreibung von Verfahrensschritten durch Wirkungsgrade und Inkremente vorgestellt. Diese Beschreibung eignet sich für solche Verfahrensschritte, bei denen Eingangsgröße und Ausgangsgröße identisch sind und sich lediglich in ihren Werten unterscheiden, z.B. der Analytgehalt einer Probe vor und nach der Probenaufbereitung.

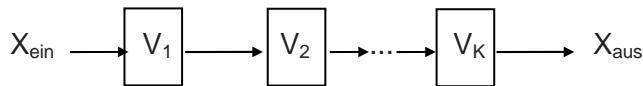
Bei der Beschreibung durch Wirkungsgrade wird die Wirkung des Verfahrensschrittes V durch Multiplikation der Eingangsgröße mit einem dimensionslosen Zahlenfaktor f_V dargestellt.

$$X_{\text{ein}} \longrightarrow \boxed{V} \longrightarrow X_{\text{aus}} = f_V \cdot X_{\text{ein}}$$

Bei der Beschreibung durch Inkremente wird die Wirkung des Verfahrensschrittes V durch Addition eines Zusatzterms a_V (derselben Dimension wie der Eingangsgröße) dargestellt.

$$X_{\text{ein}} \longrightarrow \boxed{V} \longrightarrow X_{\text{aus}} = X_{\text{ein}} + a_V$$

In einfachen Fällen lässt sich das gesamte Verfahren durch eine Kette von Verfahrensschritten V_1, V_2, \dots, V_K beschreiben:



Dann gilt bei Verwendung von Wirkungsgraden

$$X_{\text{aus}} = f_1 \cdot f_2 \cdot \dots \cdot f_K \cdot X_{\text{ein}} \quad (\text{A.3.3})$$

bzw. bei Verwendung von Inkrementen

$$X_{\text{aus}} = X_{\text{ein}} + a_1 + a_2 + \dots + a_K \quad (\text{A.3.4})$$

Hierbei ist X_{aus} die gemessene Größe, X_{ein} die tatsächlich vorliegende Größe, deren Wert zu bestimmen ist. Damit ergibt sich die Ergebnisgröße, $Y = X_{\text{ein}}$, als Funktion der am Ende der Kette gemessenen Größe, $X_{\text{mess}} = X_{\text{aus}}$, sowie der Kenngrößen für die einzelnen Verfahrensschritte wie folgt:

$$Y = \frac{X_{\text{mess}}}{f_1 \cdot f_2 \cdot \dots \cdot f_K} \quad (\text{A.3.5})$$

bzw.

$$Y = X_{\text{mess}} - (a_1 + a_2 + \dots + a_K) \quad (\text{A.3.6})$$

Für die Unsicherheit der Ergebnisgröße gilt dementsprechend

$$\left(\frac{u(Y)}{Y} \right)^2 = \left(\frac{u(X_{\text{mess}})}{X_{\text{mess}}} \right)^2 + \left(\frac{u(f_1)}{f_1} \right)^2 + \dots + \left(\frac{u(f_K)}{f_K} \right)^2 \quad (\text{A.3.7})$$

bzw.

$$u(Y)^2 = u(X_{\text{mess}})^2 + u(a_1)^2 + \dots + u(a_K)^2 \quad (\text{A.3.8})$$

Diese Unsicherheitsbilanzen gelten grundsätzlich auch dann, wenn die Wirkungsgrade $f_i = 1$ bzw. wenn die Inkremente $a_i = 0$ sind, denn auch die Werte Eins bzw. Null sind in der Regel mit Unsicherheiten behaftet. Häufig liegt nicht genug Information vor, um einen Wert $a_i \neq 0$ bzw. $f_i \neq 1$ festzulegen, aber es bestehen Vorstellungen über das Intervall der in Frage kommenden Werte. Dann ist der Ansatz $a_i = 0 \pm u(0)$ bzw. $f_i = 1 \pm u(1)$ sinnvoll.

In komplizierteren Fällen kann es zweckmäßig sein, sowohl mit Wirkungsgraden als auch mit Inkrementen zu arbeiten. Dann ist die Unsicherheitsbilanz entsprechend anzusetzen.

A.4 Numerische Verfahren zur Unsicherheitsfortpflanzung

A.4.1 Differenzenrechnung

Die Standardunsicherheit $u(Y)$ einer Ergebnisgröße Y , die von mehreren Eingangsgrößen X_1, X_2, \dots, X_N abhängt, setzt sich wie folgt aus den Standardunsicherheiten $u(X_1), u(X_2), \dots, u(X_N)$ der Eingangsgrößen sowie ggf. den Kovarianzen $u(X_i, X_k)$ zwischen korrelierten Eingangsgrößen zusammen.

$$u(Y)^2 = \sum_{i=1}^N u_i^2 + 2 \sum_{i=1}^{N-1} \sum_{k=i+1}^N u_{ik} \quad (\text{A.4.1})$$

Hierbei sind die Größen u_i und u_{ik} gegeben durch

$$u_i = \left(\frac{\partial Y}{\partial X_i} \right) \cdot u(X_i) \quad (\text{A.4.2})$$

und

$$u_{ik} = \left(\frac{\partial Y}{\partial X_i} \right) \cdot \left(\frac{\partial Y}{\partial X_k} \right) \cdot u(X_i, X_k) \quad (\text{A.4.3})$$

Bei der praktischen Anwendung dieser Gleichungen können bei den beiden folgenden Verfahrensschritten Probleme auftreten.

- (1) Ermittlung der Empfindlichkeitskoeffizienten $(\partial Y/\partial X_i)$ mittels Differentialrechnung;
- (2) "Manuelle" Einzelermittlung und Zusammenführung der Unsicherheitsbeiträge bei hoher Anzahl von Einflussgrößen.

Die Empfindlichkeitskoeffizienten $(\partial Y/\partial X_i)$ lassen sich nur dann explizit mittels Differentialrechnung bestimmen, wenn die Ergebnisgröße Y eine einfache mathematische Funktion der Eingangsgrößen X_i ist, z. B. eine Summe oder ein Produkt. Bei komplizierten Funktionen ist die Bildung der Ableitungen mühsam und fehleranfällig. Ist der Zusammenhang zwischen der Ergebnisgröße Y und den Eingangsgrößen nicht als mathematische Funktion, sondern durch ein Rechenprogramm gegeben, so ist die Bildung der Ableitungen nicht möglich.

In derartigen Fällen können die Empfindlichkeitskoeffizienten $(\partial Y/\partial X_i)$ näherungsweise als Differenzenquotienten berechnet werden. Hierfür bietet sich der folgende Rechenweg an, bei dem in einem Schritt ein Näherungswert Δ_i für das Produkt $u_i = (\partial Y/\partial X_i)u(X_i)$ bestimmt wird.

$$\Delta_i = Y\left(X_1, \dots, X_i + \frac{u(X_i)}{2}, \dots, X_N\right) - Y\left(X_1, \dots, X_i - \frac{u(X_i)}{2}, \dots, X_N\right) \quad (\text{A.4.4})$$

Mit Hilfe dieser Näherungswerte erhält man einen Näherungswert für die Standardunsicherheit der Ergebnisgröße wie folgt:

$$u(Y)^2 = \sum_{i=1}^N \Delta_i^2 + 2 \sum_{i=1}^{N-1} \sum_{k=i+1}^N \Delta_i \cdot \Delta_k \cdot r(X_i, X_k) \quad (\text{A.4.5})$$

Hierbei ist $r(X_i, X_k)$ der Korrelationskoeffizient der Größen X_i und X_k , der wie folgt mit den Standardunsicherheiten und der Kovarianz dieser Größen zusammenhängt.

$$u(X_i, X_k) = r(X_i, X_k) \cdot u(X_i) \cdot u(X_k) \quad (\text{A.4.6})$$

Die oben beschriebene Rechnung lässt sich bequem und zuverlässig mit Hilfe eines Tabellenkalkulationsprogramms durchführen. Die Vorgehensweise wird in dem EURACHEM/CITAC-Guide *Quantifying Uncertainty in Analytical Measurement*, Appendix E, Abschnitt E.2 ausführlich beschrieben.

Die Ermittlung der ggf. benötigten Korrelationskoeffizienten $r(X_i, X_k)$ zwischen korrelierten Eingangsgrößen wird in Abschnitt A.6 behandelt.

A.4.2 Monte-Carlo-Simulation

Mit dem im vorigen Abschnitt beschriebenen Verfahren lässt sich die kombinierte Standardunsicherheit in der linearen Approximation gemäß Gleichung (A.4.1) für alle praktischen Anwendungsfälle berechnen. Wie jedoch schon im Haupttext erwähnt, kann die lineare Approximation bei nicht-linearen Zusammenhängen zwischen der Ergebnisgröße Y und einer Eingangsgröße X_i zu merklichen Fehlern führen. Des Weiteren kann die Wahrscheinlichkeitsverteilung der Ergebnisgröße Y erheblich von einer Normalverteilung abweichen, mit der Folge, dass $k = 2$ den Erweiterungsfaktor für einen Überdeckungsgrad von 95 % erheblich unterschätzt.

Diese Probleme werden vermieden, wenn anstelle der Standardunsicherheiten die den Eingangsgrößen zugeordneten Wahrscheinlichkeitsverteilungen kombiniert (fortgepflanzt) werden. Bei der Monte-Carlo-Technik wird für jede Eingangsgröße eine passende Verteilung (in der Regel eine Normalverteilung, Rechteckverteilung oder Dreiecksverteilung) angesetzt. Aus diesen Verteilungen wird je ein „Zufallswert“ simuliert und aus dem so erzeugten Satz von Eingangsdaten ein Wert der Ergebnisgröße berechnet. Dieser Vorgang wird vielfach wiederholt, so dass als Ergebnis ein Datensatz resultiert, der eine Stichprobe der „möglichen“ Werte der Ergebnisgröße in Abhängigkeit von Variationen der Eingangsgrößen entsprechend ihrer Verteilung darstellt. Der Mittelwert und die Standardabweichung dieser Stichprobe sind Schätzwerte für den Wert der Ergebnisgröße und dessen Standardunsicherheit. Um verlässliche Werte zu erhalten, ist eine hohe Zahl von Wiederholungen nötig (ab 10^3); die erforderliche Größenordnung muss in der Regel ausprobiert werden.

Die Monte-Carlo-Technik liefert jedoch weitaus mehr als einen Schätzwert für die Ergebnisgröße und deren Standardunsicherheit: eine geschätzte Verteilung von Werten, die der Ergebnisgröße aufgrund der vorhandenen Informationen über die Eingangsgrößen zugeordnet werden. Aus dieser Verteilung kann u.a. bei größeren Abweichungen von der Normalverteilung ein realistischeres Vertrauensintervall bestimmt werden als nach dem Ansatz $x \pm 2u(x)$, z.B. als kleinstes Intervall, das 95 % der Verteilung umfasst.

Die Anwendung der Monte-Carlo-Technik für den o.a. Zweck wird in zahlreichen Publikationen beschrieben, z.B. in: Siebert, B.: „Berechnung der Messunsicherheit mit der Monte-Carlo-Methode“, PTB-Mitteilungen 3/4, 51 – 65 (2001).

A.4.3 Software

Für Berechnungen der Messunsicherheit durch Unsicherheitsfortpflanzung werden verschiedene Rechenprogramme angeboten.

In dem Nordtest Technical Report 430 Tools for the test laboratory to implement measurement uncertainty budgets (1999) wurde eine Übersicht über allgemein verfügbare Software gegeben.

Bezüglich neuerer Entwicklungen empfiehlt sich eine Recherche im Internet.

A.5 Die Unsicherheit von Mittelwerten

Die Mittelwertbildung ist die bei weitem häufigste Operation bei der Auswertung experimenteller Daten. Demzufolge verdient die Unsicherheitsfortpflanzung bei der Mittelwertbildung besondere Aufmerksamkeit. Üblicherweise wird jedoch bedenkenlos das bekannte $(1/\sqrt{n})$ -Gesetz benutzt, wonach die Standardabweichung eines Mittelwertes aus n Einzelwerten das $(1/\sqrt{n})$ -fache der (gemeinsamen) Standardabweichung der Einzelwerte beträgt. Das ist jedoch nur richtig bei unkorrelierten (d.h. statistisch unabhängigen) Einzelwerten. Bei korrelierten Einzelwerten müssen auch die Kovarianzen zwischen den Einzelwerten berücksichtigt werden.

Korrelationen zwischen den Einzelwerten einer Datenreihe (genauer: zwischen den Messabweichungen der Einzelwerte) treten immer dann auf, wenn Komponenten der Messabweichung zwischen den Einzelwerten der Datenreihe nicht zufällig schwanken sondern konstant sind oder sich systematisch ändern. Aufschluss gibt eine sorgfältige Analyse der Messabweichungskomponenten oder eine statistische Untersuchung geeigneter Messreihen (siehe Abschnitt A.6).

A.5.1 Allgemeines

Die Standardunsicherheit (Standardabweichung) des Mittelwertes \bar{y} aus n unabhängigen Einzelwerten y_1, y_2, \dots, y_n ist gegeben durch

$$u(\bar{y})^2 = \frac{1}{n^2} \sum_{i=1}^n u(y_i)^2 \quad (\text{A.5.1})$$

Hierbei sind die $u(y_i)$ die Standardunsicherheiten (Standardabweichungen) der Einzelwerte y_i . Stammen die Einzelwerte alle aus derselben Verteilung mit Standardabweichung σ , und ist diese Zufallsstreuung die einzige Unsicherheitsquelle, so gilt $u(y_i) = \sigma$, und es folgt

$$u(\bar{y})^2 = \frac{\sigma^2}{n} \quad (\text{A.5.2})$$

Damit ergibt sich das bekannte $(1/\sqrt{n})$ -Gesetz

$$u(\bar{y}) = \frac{\sigma}{\sqrt{n}} \quad (\text{A.5.3})$$

demzufolge mit wachsendem Stichprobenumfang n die Standardabweichung eines Mittelwertes aus n unabhängigen Einzelwerten gegen Null strebt.

Sind die Einzelwerte jedoch korreliert – z. B. wenn sie von einer oder mehreren gemeinsamen fehlerbehafteten Größen abhängen – dann gilt

$$u(\bar{y})^2 = \frac{1}{n^2} \left[\sum_{i=1}^n u(y_i)^2 + 2 \sum_{i=1}^{n-1} \sum_{k=i+1}^n u(y_i, y_k) \right] \quad (\text{A.5.4})$$

Hierbei sind die $u(y_i, y_k)$ die Kovarianzen zwischen den Einzelwerten.

Sind die Werte alle gleich stark miteinander korreliert, so ergibt sich mit $u(y_i) = \sigma$ und einem Korrelationskoeffizienten r für die Unsicherheit des Mittelwertes

$$u(\bar{y})^2 = \left(r + \frac{1-r}{n} \right) \cdot \sigma^2 \quad (\text{A.5.5})$$

Bereits bei mäßiger Korrelation wird $u(\bar{y})$ durch den Ansatz σ / \sqrt{n} erheblich unterschätzt: mit $r = 0,5$ gilt $u(\bar{y}) \geq \sigma / \sqrt{2}$.

Korrelationen treten u.a. dann auf, wenn die Einzelwerte nicht nur zufällig streuen, sondern zusätzlich mit einer systematischen Abweichung behaftet sind, z. B. durch Verwendung derselben Kalibrierung. Dann mitteln sich zwar die zufälligen Abweichungen heraus (je größer der Stichprobenumfang n , desto perfekter). Die systematische Abweichung wird jedoch durch Mittelung nicht vermindert und wird mit wachsendem Stichprobenumfang n zum beherrschenden Teil der Ergebnisunsicherheit.

Beispiel: Ein Labor will für ein genormtes Messverfahren die im Ringversuch ermittelte Vergleichsstandardabweichung s_R verwenden. Für einen besonderen Anwendungsfall werden jedoch zwecks Steigerung der Genauigkeit abweichend von der Norm Mehrfachmessungen (n) durchgeführt und gemittelt. Dann darf nicht etwa $u(\bar{y}) = s_R / \sqrt{n}$ angesetzt werden, denn die Vergleichsstandardabweichung s_R besteht aus zwei Komponenten gemäß $s_R^2 = s_r^2 + s_L^2$. Bei der Mittelung mehrerer Messergebnisse im Labor verhält sich nur die Wiederholstreuung zufällig, während die Laborabweichungskomponente konstant bleibt. Deshalb ist anzusetzen: $u(\bar{y})^2 = s_r^2 / n + s_L^2$.

Zusammenfassend ist festzustellen, dass die Gleichungen (A.5.1) bzw. (A.5.2) und (A.5.3) nur dann verwendet werden dürfen, wenn sichergestellt ist, dass die Einzelwerte unkorreliert sind, oder wenn der Beitrag der Korrelationen nachweislich geringfügig ist.

Andernfalls müssen die Korrelationen untersucht und ggf. quantitativ berücksichtigt werden. Hierfür können alternativ die beiden folgenden Verfahrensweisen verwendet werden:

- (1) Rückführung auf möglichst unkorrelierte Eingangsgrößen, Unsicherheitsfortpflanzung in Bezug auf diese Eingangsgrößen;
- (2) Unsicherheitsfortpflanzung nach Gleichung (A.5.4), wobei die Kovarianzen zwischen den Einzelwerten nach den in Abschnitt A.6 beschriebenen Verfahren ermittelt werden.

A.5.2 Korrelation innerhalb einer Messreihe

In diesem Abschnitt wird als wichtigster Anwendungsfall die Unsicherheit eines Mittelwertes \bar{y} aus den Einzelwerten einer Messreihe y_1, y_2, \dots, y_n , die mit demselben Messverfahren am selben Messobjekt (oder an verschiedenen i.W. identischen Messobjekten) erhalten wurden.

Unter diesen Umständen ist die Standardunsicherheit für alle Einzelwerte identisch, d.h. es gilt $u(y_i) = u(y)$ für alle y_i . Neben den rein zufälligen Schwankungen von Messung zu Messung gibt es jedoch in der Regel Einflüsse auf die Messung, die während einer Messreihe unverändert bleiben: z. B. Kalibrierung, Messbedingungen, Charakteristika der Messobjekte. Die Messunsicherheit $u(y)$ besteht somit aus zwei Komponenten gemäß

$$u(y)^2 = u_{\text{var}}(y)^2 + u_{\text{inv}}(y)^2 \quad (\text{A.5.6})$$

Hierbei erfasst $u_{\text{var}}(y)$ die Beiträge der von Messung zu Messung variierenden Einflüsse, während $u_{\text{inv}}(y)$ die Beiträge der von Messung zu Messung gleichbleibenden (invarianten) Einflüsse erfasst.

Für $u_{\text{var}}(y)$ ist die „Verfahrensstandardabweichung“ s_V , d.h. die Standardabweichung unter laborinternen Vergleichsbedingungen, ein angemessener Schätzwert; behelfsweise kann auch die Standardabweichung der Einzelwerte der Messreihe verwendet werden. Falls für $u_{\text{inv}}(y)$ kein separater Wert verfügbar ist, aber ein Schätzwert für die Gesamtunsicherheit $u(y)$ vorliegt (z.B. als Vergleichsstandardabweichung s_R), kann für $u_{\text{inv}}(y)$ die Wurzel aus der Differenz $u(y)^2 - s_V^2$ verwendet werden:

$$u_{\text{inv}}(y) = \sqrt{u(y)^2 - s_V^2} \quad (\text{A.5.7})$$

Die Kovarianz $u(y, y')$ zwischen zwei Einzelwerten erhält man gemäß

$$u(y, y') = u_{\text{inv}}(y)^2 = u(y)^2 - s_V^2 \quad (\text{A.5.8})$$

Damit ergibt sich für die Unsicherheit des Mittelwertes

$$u(\bar{y})^2 = \frac{u_{\text{var}}(y)^2}{n} + u_{\text{inv}}(y)^2 = \frac{s_V^2}{n} + (u(y)^2 - s_V^2) \quad (\text{A.5.9})$$

Durch die Mittelwertbildung wird also nur die Zufallskomponente der Messunsicherheit um den Faktor $1 / \sqrt{n}$ verringert, während die systematische Unsicherheitskomponente unverändert bleibt.

Die Korrelation wirkt sich nicht nur bei Mittelwerten, sondern auch bei anderen Kombinationen von Messwerten aus, die mit demselben Messverfahren am selben Messobjekt oder an vergleichbaren Messobjekten erhalten werden. So gilt z.B. für eine Differenz

$$u(y - y')^2 = u(y)^2 + u(y')^2 - 2u(y, y') \quad (\text{A.5.10})$$

Ist die Messunsicherheit für y und y' gleich, so erhält man mit Gleichung (A.5.8)

$$u(y - y')^2 = 2s_V^2 \quad (\text{A.5.11})$$

anstelle von $2u(y)^2$ ohne Berücksichtigung von Korrelationen. Der anschauliche Grund für diesen Genauigkeitsgewinn besteht darin, dass sich bei Differenzen (und ebenso bei Quotienten) systematische Abweichungen kompensieren, wie z.B. bei Differenzwägungen.

A.6 Ermittlung von Kovarianzen und Korrelationskoeffizienten

A.6.1 Allgemeines

Bei der Berechnung der Standardunsicherheit $u(Y)$ einer Ergebnisgröße Y , die von mehreren Eingangsgrößen X_1, X_2, \dots, X_N abhängt, werden neben den Standardunsicherheiten $u(X_1), u(X_2), \dots, u(X_N)$ der Eingangsgrößen auch die Kovarianzen $u(X_i, X_j)$ zwischen korrelierten Eingangsgrößen X_i, X_j benötigt.

Mit Korrelationen ist immer dann zu rechnen, wenn zwei Größen voneinander oder von einer gemeinsamen dritten (ggf. versteckten) Größe bzw. von mehreren solchen Größen abhängen. Diese Abhängigkeit kann sich direkt auf die physikalischen Größen selbst beziehen. So sind z.B. die Massenanteile der Bestandteile eines Stoffgemisches voneinander abhängig, denn ihre Summe ist gleich Eins. Häufiger jedoch sind zwar die physikalischen Größen voneinander unabhängig, aber ihre Werte wurden nicht unabhängig voneinander ermittelt. Das ist der Fall, wenn zwei Größen in demselben Experiment bestimmt werden – z.B. Achsenabschnitt und Steigung einer Kalibriergeraden – oder wenn bei verschiedenen Messungen dasselbe Normal verwendet wird. Dann hängen die ermittelten Größen von gemeinsamen Größen ab: Dem Kalibrierdatensatz bzw. dem Wert des Normals o.ä.

Die zu zwei Eingangsgrößen X_i und X_j gehörende Kovarianz $u(X_i, X_j)$ kann gleich Null gesetzt werden, wenn

- X_i und X_j unabhängige physikalische Größen sind und ihre Werte unabhängig voneinander ermittelt wurden;
- mindestens eine der beiden Unsicherheiten $u(X_i), u(X_j)$ vernachlässigbar klein ist.

Ist die zur Ermittlung einer Kovarianz erforderliche Information nicht verfügbar und nicht mit vertretbarem Aufwand beschaffbar, so ist man auf Abschätzungen angewiesen. Grundlage hierfür ist die Tatsache, dass $u(X_i, X_j)$ zwischen $u(X_i)u(X_j)$ und $-u(X_i)u(X_j)$ liegt. Gibt es darüber hinaus keinerlei Anhaltspunkte für die Größe der Kovarianz, nicht einmal für das Vorzeichen, so kann $u(X_i, X_j) = 0$ angesetzt werden, wenn Überschätzungen und Unterschätzungen gleichermaßen vermieden werden sollen. Ist es hingegen vor allem wichtig, Unterschätzungen zu vermeiden, so kann $u(X_i, X_j) = u(X_i)u(X_j)$ angesetzt werden. Weniger grobe Abschätzungen sind möglich, wenn Vorstellungen über den Mechanismus einer Korrelation bestehen.

Häufig sind Eingangsgrößen korreliert, weil man von unkorrelierten Urvariablen zu abgeleiteten Variablen übergegangen ist, die bequemer handhabbar sind (geringere Anzahl, treffendere Systemcharakterisierung etc.). Dieser Vorteil wird jedoch mit dem Auftreten von Korrelationen bezahlt. Falls dies bei der Ermittlung der Ergebnisunsicherheit zu Schwierigkeiten führt, ist es zweckmäßig, von den korrelierten Eingangsgrößen zu den unkorrelierten Urvariablen zurückzugehen.

Wenn es sich denn nicht vermeiden lässt, Kovarianzen zu ermitteln, stehen hierfür im Wesentlichen zwei Verfahrensweisen zur Verfügung:

- Unsicherheitsfortpflanzung bezüglich der gemeinsamen Variablen;
- experimentelle Ermittlung aus Parallelmessungen.

A.6.2 Unsicherheitsfortpflanzung

Hängen zwei Größen X_i und X_j von denselben unkorrelierten Größen Z_1, Z_2, \dots, Z_K ab, so gilt

$$u(X_i, X_j) = \sum_{k=1}^K \left(\frac{\partial X_i}{\partial Z_k} \right) \cdot \left(\frac{\partial X_j}{\partial Z_k} \right) \cdot u(Z_k)^2 \quad (\text{A.6.1})$$

Sind die Größen Z_1, Z_2, \dots, Z_K ihrerseits korreliert, so sind auch deren Kovarianzen zu berücksichtigen. In diesem Falle gilt

$$u(X_i, X_j) = \sum_{k=1}^K \left(\frac{\partial X_i}{\partial Z_k} \right) \cdot \left(\frac{\partial X_j}{\partial Z_k} \right) \cdot u(Z_k)^2 + 2 \sum_{k=1}^{K-1} \sum_{l=k+1}^K \left(\frac{\partial X_i}{\partial Z_k} \right) \cdot \left(\frac{\partial X_j}{\partial Z_l} \right) \cdot u(Z_k, Z_l) \quad (\text{A.6.2})$$

A.6.3 Parallelmessungen

Werden zwei Größen X_i und X_j unter Wiederholbedingungen mehrmals gemeinsam in n unabhängigen Versuchen gemessen, und ergeben sich dabei die Einzelwertpaare $(x_{i1}, x_{j1}), (x_{i2}, x_{j2}), \dots, (x_{in}, x_{jn})$, so können die Werte von X_i und X_j durch die Mittelwerte \bar{x}_i und \bar{x}_j von $x_{i1}, x_{i2}, \dots, x_{in}$ bzw. von $x_{j1}, x_{j2}, \dots, x_{jn}$ geschätzt werden. Dementsprechend wird die Kovarianz $u(X_i, X_j)$ durch die empirische Kovarianz der Mittelwerte, $s(\bar{x}_i, \bar{x}_j)$ geschätzt. Sie ist gegeben durch

$$s(\bar{x}_i, \bar{x}_j) = \frac{1}{n(n-1)} \cdot \sum_{p=1}^n (x_{ip} - \bar{x}_i)(x_{jp} - \bar{x}_j) \quad (\text{A.6.3})$$

Die Standardunsicherheiten $u(X_i)$ und $u(X_j)$ werden dann durch die empirischen Standardabweichungen $s(\bar{x}_i)$ bzw. $s(\bar{x}_j)$ der Mittelwerte geschätzt, die wie folgt gegeben sind:

$$s(\bar{x}_i)^2 = \frac{1}{n(n-1)} \cdot \sum_{p=1}^n (x_{ip} - \bar{x}_i)^2 \quad (\text{A.6.4})$$

mit einer entsprechenden Gleichung für $s(\bar{x}_j)$.

Sollen die Werte für $u(X_i)$, $u(X_j)$ und $u(X_i, X_j)$ als Verfahrensunsicherheiten bei künftigen Messungen der Größen X_i und X_j verwendet werden, so ist zu beachten, ob Einzelwerte oder Mittelwerte bestimmt werden. Bei Einzelwerten sind die nach (A.6.3) und (A.6.4) bestimmten Schätzwerte für $u(X_i, X_j)$ und $u(X_i)^2$ bzw. $u(X_j)^2$ mit dem Faktor n zu multiplizieren, bei Mittelwerten aus m Einzelwerten mit dem Faktor n/m .

Anmerkung: Bei geringer Anzahl n von Messdaten sind die Schätzungen gemäß Gleichung (A.6.3) und (A.6.4) sehr ungenau. Dann ist es u.U. zweckmäßiger, auf Erfahrungen oder Modellen beruhende Schätzwerte zu verwenden.

A.6.4 Korrelationskoeffizienten

Korrelationskoeffizienten $r(X_i, X_j)$ ergeben sich aus den entsprechenden Kovarianzen $u(X_i, X_j)$ durch Normierung mit den Standardunsicherheiten:

$$r(X_i, X_j) = \frac{u(X_i, X_j)}{u(X_i) \cdot u(X_j)} \quad (\text{A.6.5})$$

Falls benötigt, können Schätzwerte für Korrelationskoeffizienten aus Schätzwerten für Kovarianzen und Standardunsicherheiten bestimmt werden, z. B. durch Kombination der Gleichungen (A.6.3) und (A.6.4). Sind Informationen über Richtung und Stärke der Korrelation vorhanden, so können auf dieser Grundlage Korrelationskoeffizienten direkt abgeschätzt werden.

Beispiel: Korrelationen zwischen den Einzelwerten einer Messreihe (siehe Abschnitt A.5) können nur positiv sein, d.h. $0 < r \leq 1$. Dann kann bei Fehlen näherer Information $r = 0,5$ angesetzt werden.

