

DIN EN ISO 4259-5



ICS 75.080

Einsprüche bis 2022-12-21

Entwurf

**Mineralölerzeugnisse –
Präzision von Messverfahren und Ergebnissen –
Teil 5: Statistische Bewertung der Übereinstimmung zweier
verschiedener Messverfahren die vorgeben, dieselbe Eigenschaft zu
messen (ISO/DIS 4259-5:2022);
Deutsche und Englische Fassung prEN ISO 4259-5:2022**

Petroleum and related products –

Precision of measurement methods and results –

Part 5: Statistical assessment of agreement between two different measurement methods that claim to measure the same property (ISO/DIS 4259-5:2022);

German and English version prEN ISO 4259-5:2022

Produits pétroliers et connexes –

Fidélité des méthodes de mesure et de leurs résultats –

Partie 5: Évaluation statistique de l'accord entre deux méthodes de mesure différentes qui prétendent mesurer la même propriété (ISO/DIS 4259-5:2022);

Version allemande et anglaise prEN ISO 4259-5:2022

Anwendungswarnvermerk

Dieser Norm-Entwurf mit Erscheinungsdatum 2022-10-21 wird der Öffentlichkeit zur Prüfung und Stellungnahme vorgelegt.

Weil die beabsichtigte Norm von der vorliegenden Fassung abweichen kann, ist die Anwendung dieses Entwurfs besonders zu vereinbaren.

Stellungnahmen werden erbeten

- vorzugsweise online im Norm-Entwurfs-Portal von DIN unter www.din.de/go/entwuerfe bzw. für Norm-Entwürfe der DKE auch im Norm-Entwurfs-Portal der DKE unter www.entwuerfe.normenbibliothek.de, sofern dort wiedergegeben;
- oder als Datei per E-Mail an stellungnahme@fam-hamburg.de möglichst in Form einer Tabelle. Die Vorlage dieser Tabelle kann im Internet unter www.din.de/go/stellungnahmen-norm-entwuerfe oder für Stellungnahmen zu Norm-Entwürfen der DKE unter www.dke.de/stellungnahme abgerufen werden;
- oder in Papierform an den Fachausschuss Mineralöl- und Brennstoffnormung (FAM) des NMP, 22205 Hamburg, Postfach 60 05 49 oder Große Elbstr. 131, 22767 Hamburg.

Die Empfänger dieses Norm-Entwurfs werden gebeten, mit ihren Kommentaren jegliche relevanten Patentrechte, die sie kennen, mitzuteilen und unterstützende Dokumentationen zur Verfügung zu stellen.

Gesamtumfang 134 Seiten

DIN-Normenausschuss Materialprüfung (NMP)
Fachausschuss Mineralöl- und Brennstoffnormung (FAM) des NMP

Nationales Vorwort

Dieses Dokument (prEN ISO 4259-5:2022) wurde vom Technischen Komitee ISO/TC 28 „Petroleum and related products, fuels and lubricants from natural or synthetic sources“ in Zusammenarbeit mit dem Technischen Komitee CEN/TC 19 „Gasförmige und flüssige Kraft- und Brennstoffe, Schmierstoffe und verwandte Produkte aus Erdöl und mit biologischem oder synthetischem Ursprung“ erarbeitet, dessen Sekretariat von NEN (Niederlande) gehalten wird.

Das zuständige nationale Normungsgremium ist der Arbeitsausschuss NA 062-06-16 AA „Präzision von Prüfverfahren“ im DIN-Normenausschuss Materialprüfung (NMP).

Um Zweifelsfälle in der Übersetzung auszuschließen, ist die englische Originalfassung beigelegt. Die Nutzungsbedingungen für den deutschen Text des Norm-Entwurfes gelten gleichermaßen auch für den englischen Text.

Für die in diesem Dokument zitierten Dokumente wird im Folgenden auf die entsprechenden deutschen Dokumente hingewiesen:

ISO 4259-1	siehe	DIN EN ISO 4259-1
ISO 4259-3	siehe	DIN EN ISO 4259-3
ISO 4259-4	siehe	DIN EN ISO 4259-4
ISO 5165:2017	siehe	DIN EN ISO 5165:2018-04

Aktuelle Informationen zu diesem Dokument können über die Internetseiten von DIN (www.din.de) durch eine Suche nach der Dokumentennummer aufgerufen werden.

Nationaler Anhang NA (informativ)

Literaturhinweise

DIN EN ISO 4259-1, *Mineralölerzeugnisse — Präzision von Messverfahren und Ergebnissen — Teil 1: Bestimmung der Präzisionsdaten von Prüfverfahren*

DIN EN ISO 4259-3, *Mineralölerzeugnisse — Präzision von Messverfahren und Ergebnissen — Teil 3: Monitoring und Management veröffentlichter Präzisionsdaten in Bezug auf Prüfverfahren*

DIN EN ISO 4259-4, *Mineralölerzeugnisse — Präzision von Messverfahren und Ergebnissen — Teil 4: Verwendung von Kontrollkarten zur Validierung des Status der statistischen Kontrolle bei der Durchführung von genormten Prüfverfahren in einem einzelnen Labor*

DIN EN ISO 5165:2018-04, *Mineralölerzeugnisse — Bestimmung der Zündwilligkeit von Dieselmotoren — Cetan-Verfahren mit dem CFR-Motor (ISO 5165:2017); Deutsche Fassung EN ISO 5165:2018*

– *Entwurf* –

E DIN EN ISO 4259-5:2022-11

– Leerseite –

**Mineralölerzeugnisse – Präzision von Messverfahren und Ergebnissen – Teil 5:
Statistische Bewertung der Übereinstimmung zweier verschiedener Messverfahren
die vorgeben, dieselbe Eigenschaft zu messen (ISO/DIS 4259-5:2022)**

Petroleum and related products – Precision of measurement methods and results – Part 5: Statistical assessment of agreement between two different measurement methods that claim to measure the same property (ISO/DIS 4259-5:2022)

Produits pétroliers et connexes – Fidélité des méthodes de mesure et de leurs résultats – Partie 5: Évaluation statistique de l'accord entre deux méthodes de mesure différentes qui prétendent mesurer la même propriété (ISO/DIS 4259-5:2022)

Inhalt

	Seite
Europäisches Vorwort	5
Vorwort	6
Einleitung	7
1 Anwendungsbereich	8
2 Normative Verweisungen	8
3 Begriffe	8
4 Symbole und Einheiten	10
5 Überblick zum Verfahren	11
5.1 Allgemeine Anforderungen	11
5.2 Zusätzliche Anforderungen für PTP-Daten	12
5.2.1 Prüfung auf das Vorliegen extremer Proben	13
5.2.2 Prüfung auf Verteilung der Laborergebnisse	14
5.2.3 Vergleich der Präzision	14
5.3 Kurze Abfolge der Verfahrensschritte	15
5.4 Flussdiagramm des Verfahrens	17
6 Durchführung	19
6.1 Probenmittelwert und Standardfehler	19
6.1.1 Berechnung der Mittelwerte	19
6.1.2 Berechnung der Standardfehler	19
6.2 Eignung der Daten	20
6.2.1 Prüfung auf Eigenschaftsstreuung	20
6.2.2 Korrelation der Prüfverfahren	21
6.3 Bias-Korrektur-Statistik	21
6.3.1 Allgemeines	21
6.3.2 Klasse 0 — Keine Bias-Korrektur.	21
6.3.3 Klasse 1a — Konstante Bias-Korrektur.	22
6.3.4 Klasse 1b — Proportionale Bias-Korrektur.	22
6.3.5 Klasse 2 — Proportionale und konstante Bias-Korrektur.	23
6.4 Auswahl der geeigneten Bias-Korrektur-Klasse	24
6.5 Bestätigung der Normalverteilung der Residuen	25
6.6 Probenspezifische systematische Abweichungen	25
7 Prüfbericht	27
8 Bestätigung der Korrelation	27
Anhang A (informativ) Berechnungsbeispiel unter Verwendung von Ringversuchsdaten	29
A.1 Auswahl relevanter Eigenschaften	29
A.1.1 Prüfverfahren X	29
A.1.2 Prüfverfahren Y	30
A.2 Labordaten	30
A.3 Berechnung der Mittelwerte über die Wiederholungen und Neuordnung	32
A.4 Grundlegende Überprüfungen des Datensatzes durchführen.	35
A.5 Prüfung auf Einfluss der Proben	35
A.6 Erstellung von Übersichtstabellen mit zusammenfassenden Statistiken je Probe	36
A.7 Prüfung auf ausreichende Streuung	39
A.8 Korrelationsprüfung	40
A.9 Berechnung der Bias-Korrektur-Klassen	42
A.10 Auswahl der am besten geeigneten Bias-Korrektur-Klasse	48
A.11 Vergleichbarkeit zwischen den Verfahren	51
A.12 Bestätigung	51
Anhang B (informativ) Berechnungsbeispiel unter Verwendung von PTP-Daten	52
B.1 Auswahl relevanter Eigenschaften	52
B.1.1 Prüfverfahren X	52

B.1.2	Prüfverfahren Y	52
B.2	Labordaten	53
B.3	Neuanordnung des Datensatzes	54
B.4	Grundlegende Überprüfungen des Datensatzes durchführen:	57
B.5	Prüfung auf Einfluss der Proben	57
B.6	Erstellung von Übersichtstabellen mit zusammenfassenden Statistiken je Probe	58
B.7	Prüfung auf ausreichende Streuung	61
B.8	Korrelationsprüfung	62
B.9	Berechnung der Bias-Korrektur-Klassen	64
B.10	Auswahl der am besten geeigneten Bias-Korrektur-Klasse	70
B.11	Vergleichbarkeit zwischen den Verfahren	71
B.12	Bestätigung	72
	Literaturhinweise	74

Bilder

Bild A.1	Diagramm der Mittelwerte Verfahren X (ISO 5165) versus Verfahren Y (EN 16906)	41
Bild A.2	Residuendiagramm von X_i und Y_i	41
Bild A.3	Gewichtete Residuen für eine Korrektur der Klasse 0	43
Bild A.4	Quantil-Diagramm für eine Korrektur der Klasse 0	43
Bild A.5	Gewichtete Residuen für eine Korrektur der Klasse 1a	44
Bild A.6	Quantil-Diagramm für eine Korrektur der Klasse 1a	45
Bild A.7	Gewichtete Residuen für eine Korrektur der Klasse 1b	46
Bild A.8	Quantil-Diagramm für eine Korrektur der Klasse 1b	46
Bild A.9	Gewichtete Residuen für eine Korrektur der Klasse 2	47
Bild A.10	Quantil-Diagramm für eine Korrektur der Klasse 2	48
Bild B.1	Diagramm der Mittelwerte Verfahren X versus Verfahren Y	63
Bild B.2	Residuendiagramm von X_i und Y_i	63
Bild B.3	Gewichtete Residuen für eine Korrektur der Klasse 0	65
Bild B.4	Quantil-Diagramm für eine Korrektur der Klasse 0	65
Bild B.5	Gewichtete Residuen für eine Korrektur der Klasse 1a	66
Bild B.6	Quantil-Diagramm für eine Korrektur der Klasse 1a	67
Bild B.7	Gewichtete Residuen für eine Korrektur der Klasse 1b	68
Bild B.8	Quantil-Diagramm für eine Korrektur der Klasse 1b	68
Bild B.9	Gewichtete Residuen für eine Korrektur der Klasse 2	69
Bild B.10	Quantil-Diagramm für eine Korrektur der Klasse 2	70

Tabellen

Tabelle 1	Bias-Korrektur-Klassen	24
Tabelle A.1	Anwendungsbereich der ISO 5165:2017	29
Tabelle A.2	Veröffentlichte Präzision für ISO 5165:2017	29
Tabelle A.3	Anwendungsbereich der EN 16906:2020	30
Tabelle A.4	Cetanzahldaten für Verfahren X (ISO 5165:2017)	31
Tabelle A.5	Cetanzahl für Verfahren Y (EN 16906:2020)	32
Tabelle A.6	Cetanzahldaten für Verfahren X (ISO 5165:2017)	33
Tabelle A.7	Cetanzahldaten für Verfahren Y (EN 16906:2020)	34
Tabelle A.8	Anwendungs- und Präzisionsbereiche der Prüfverfahren im Vergleich mit dem Datensatz	35
Tabelle A.9	Einfluss der einzelnen Proben, gemittelt über beide Verfahren	35
Tabelle A.10	Übersichtstabelle für Verfahren X (ISO 5165:2017)	37
Tabelle A.11	Übersichtstabelle für Verfahren Y (EN 16906:2020)	38
Tabelle A.12	Zusammenfassende Statistiken für die Prüfung auf ausreichende Streuung	39

Tabelle A.13 — Zusammenfassende Statistiken für die Korrelationsprüfung	40
Tabelle A.14 — Keine Bias-Korrektur, Klasse 0	42
Tabelle A.15 — Konstante Korrektur, Klasse 1a	43
Tabelle A.16 — Proportionale Korrektur, Klasse 1b	45
Tabelle A.17 — Lineare Korrektur, Klasse 2	46
Tabelle A.18 — Zusammenfassung der Ergebnisse für die Korrekturklassen	49
Tabelle A.19 — Zusammengefasste Statistiken für die Auswahl der optimalen Bias-Korrektur . .	50
Tabelle B.1 — Anwendungsbereich für Benzol ASTM D6839:2017	52
Tabelle B.2 — Anwendungsbereich der ASTM D5580:2020	52
Tabelle B.3 — Benzolanteil für Verfahren X (ASTM D6839:2017)	53
Tabelle B.4 — Benzolanteil nach Verfahren Y (ASTM D5580:2020)	54
Tabelle B.5 — Neu angeordnete Daten für Benzol für Verfahren X (ASTM D6839:2017)	55
Tabelle B.6 — Neu angeordnete Daten für Benzol für Verfahren Y (ASTM D5580:2020)	56
Tabelle B.7 — Anwendungs- und Präzisionsbereiche für beide Prüfverfahren	57
Tabelle B.8 — Einfluss der einzelnen Proben, gemittelt über beide Verfahren	57
Tabelle B.9 — Überblickstabelle für Verfahren X (ASTM D6839:2017)	59
Tabelle B.10 — Überblickstabelle für Verfahren Y (ASTM D5580:2020)	60
Tabelle B.11 — Zusammengefasste Statistiken für die Prüfung auf ausreichende Streuung	61
Tabelle B.12 — Zusammenfassende Statistik für die Korrelationsprüfung	62
Tabelle B.13 — Keine Bias-Korrektur, Klasse 0	64
Tabelle B.14 — Konstante Korrektur, Klasse 1a	65
Tabelle B.15 — Proportionale Korrektur, Klasse 1b	67
Tabelle B.16 — Lineare Korrektur, Klasse 2	68
Tabelle B.18 — Zusammenfassung der Ergebnisse der Korrekturklassen	70
Tabelle B.19 — Zusammengefasste Statistiken für die Auswahl der optimalen Bias-Korrektur . .	70
Tabelle B.20 — Benzolanteil für Verfahren Y (D5580:2020)	72
Tabelle B.21 — Benzolanteil für Verfahren X (D6839:2017)	72
Tabelle B.22 — Ergebnisse der Differenz-Statistik D für die einzelnen Proben S_i	73

Europäisches Vorwort

Dieses Dokument (prEN ISO 4259-5:2022) wurde vom Technischen Komitee ISO/TC 28 „Mineralölerzeugnisse und verwandte Produkte, Kraft-, Brenn- und Schmierstoffe mit natürlichem oder synthetischem Ursprung“ in Zusammenarbeit mit dem Technischen Komitee CEN/TC 19 „Gasförmige und flüssige Kraft- und Brennstoffe, Schmierstoffe und verwandte Produkte aus Erdöl und mit biologischem oder synthetischem Ursprung“ erarbeitet, dessen Sekretariat von NEN gehalten wird.

Dieses Dokument ist derzeit zur ersten parallelen Umfrage vorgelegt.

Rückmeldungen oder Fragen zu diesem Dokument sollten an das jeweilige nationale Normungsinstitut des Anwenders gerichtet werden. Eine vollständige Liste dieser Institute ist auf den Internetseiten von CEN abrufbar.

Anerkennungsnotiz

Der Text von ISO/DIS 4259-5:2022 wurde von CEN als prEN ISO 4259-5:2022 ohne irgendeine Abänderung genehmigt.

Vorwort

ISO (die Internationale Organisation für Normung) ist eine weltweite Vereinigung nationaler Normungsinstitute (ISO-Mitgliedsorganisationen). Die Erstellung von Internationalen Normen wird üblicherweise von Technischen Komitees von ISO durchgeführt. Jede Mitgliedsorganisation, die Interesse an einem Thema hat, für welches ein Technisches Komitee gegründet wurde, hat das Recht, in diesem Komitee vertreten zu sein. Internationale staatliche und nichtstaatliche Organisationen, die in engem Kontakt mit ISO stehen, nehmen ebenfalls an der Arbeit teil. ISO arbeitet bei allen elektrotechnischen Normungsthemen eng mit der Internationalen Elektrotechnischen Kommission (IEC) zusammen.

Die Verfahren, die bei der Entwicklung dieses Dokuments angewendet wurden und die für die weitere Pflege vorgesehen sind, werden in den ISO/IEC-Direktiven, Teil 1 beschrieben. Es sollten insbesondere die unterschiedlichen Annahmekriterien für die verschiedenen ISO-Dokumentenarten beachtet werden. Dieses Dokument wurde in Übereinstimmung mit den Gestaltungsregeln der ISO/IEC-Direktiven, Teil 2 erarbeitet (siehe www.iso.org/directives).

Es wird auf die Möglichkeit hingewiesen, dass einige Elemente dieses Dokuments Patentrechte berühren können. ISO ist nicht dafür verantwortlich, einige oder alle diesbezüglichen Patentrechte zu identifizieren. Details zu allen während der Entwicklung des Dokuments identifizierten Patentrechten finden sich in der Einleitung und/oder in der ISO-Liste der erhaltenen Patenterklärungen (siehe www.iso.org/patents).

Jeder in diesem Dokument verwendete Handelsname dient nur zur Unterrichtung der Anwender und bedeutet keine Anerkennung.

Für eine Erläuterung des freiwilligen Charakters von Normen, der Bedeutung ISO-spezifischer Begriffe und Ausdrücke in Bezug auf Konformitätsbewertungen sowie Informationen darüber, wie ISO die Grundsätze der Welthandelsorganisation (WTO, en: World Trade Organization) hinsichtlich technischer Handelshemmnisse (TBT, en: Technical Barriers to Trade) berücksichtigt, siehe www.iso.org/iso/foreword.html.

Dieses Dokument wurde vom Technischen Komitee ISO/TC 28, *Petroleum and related products, fuels and lubricants from natural or synthetic sources* erarbeitet.

Eine Auflistung aller Teile der Normenreihe ISO 4259 ist auf der ISO-Internetseite abrufbar.

Rückmeldungen oder Fragen zu diesem Dokument sollten an das jeweilige nationale Normungsinstitut des Anwenders gerichtet werden. Eine vollständige Auflistung dieser Institute ist unter www.iso.org/members.html zu finden.

Einleitung

In diesem Dokument wird die statistische Methodik zur Beurteilung der erwarteten Übereinstimmung zweier genormten Prüfverfahren, die vorgeben, dieselbe Eigenschaft eines Materials zu messen, erläutert. Anschließend wird untersucht, ob eine lineare Bias-Korrektur die erwartete Übereinstimmung signifikant verbessern kann. Der Grad der Übereinstimmung wird als Wiederholpräzision zwischen den Verfahren angegeben, nachdem eine Bias-Korrektur (falls erforderlich) vorgenommen wurde.

Das Verfahren verwendet numerische Ergebnisse aus einer Reihe von Proben, die unabhängig voneinander mit beiden Prüfverfahren durch unterschiedliche Labore ausgewertet wurden. Die mit den Ergebnissen der einzelnen Prüfverfahren verbundene Abweichung wird zur Beurteilung der erforderlichen Bias-Korrektur herangezogen.

Anhang A und Anhang B enthalten Berechnungsbeispiele für die Anwendung der Methodik.

1 Anwendungsbereich

In diesem Dokument wird die statistische Methodik zur Beurteilung der erwarteten Übereinstimmung zweier verschiedener Prüfverfahren festgelegt, die vorgeben, dieselbe Eigenschaft eines Materials zu messen, und es wird entschieden, ob eine einfache lineare Bias-Korrektur (Korrektur der systematischen Abweichung) die erwartete Übereinstimmung weiter verbessern kann.

Dieses Dokument wird für Analyseverfahren verwendet, bei denen die quantitativen Eigenschaften von Mineralölerzeugnissen gemessen werden, die sich aus einer Multi-Sample-Multi-Lab-Studie (MSMSL, en: multi-sample-multi-lab study) ergeben. Zu dieser Art von Studien gehören Ringversuche (ILS, en: interlaboratory study), die den Anforderungen der ISO 4259-1 oder einer gleichwertigen Norm entsprechen, und Eignungsprüfungsprogramme (PTP, en: proficiency testing program), die den Anforderungen der ISO 4259-3 oder einer gleichwertigen Norm entsprechen, sind aber nicht auf diese Datenlieferanten beschränkt, solange die Anforderungen an die Datenqualität erfüllt werden.

Es wird eine Methodik dargestellt für die Festlegung des Grenzwerts, der als Vergleichbarkeit zwischen den Verfahren bezeichnet wird. Für diesen Wert wird erwartet, dass er bei üblicher und korrekter Durchführung beider Prüfverfahren aufgrund zufälliger Schwankung für die Differenz zwischen zwei Ergebnissen mit einer Wahrscheinlichkeit von 5 % überschritten wird, wenn die einzelnen Ergebnisse durch einen anderen Bearbeiter unter Verwendung einer anderen Ausrüstung erhalten werden und jeweils eines der beiden Verfahren X und Y an identischem Material angewendet wird, wobei das Ergebnis des Verfahrens X in geeigneter Weise dementsprechend Bias-korrigiert wurde.

ANMERKUNG Weitere Bedingungen für die Anwendung dieser Methodik sind in 5.1 und 5.2 angegeben.

2 Normative Verweisungen

Die folgenden Dokumente werden im Text in solcher Weise in Bezug genommen, dass einige Teile davon oder ihr gesamter Inhalt Anforderungen des vorliegenden Dokuments darstellen. Bei datierten Verweisungen gilt nur die in Bezug genommene Ausgabe. Bei undatierten Verweisungen gilt die letzte Ausgabe des in Bezug genommenen Dokuments (einschließlich aller Änderungen).

ISO 4259-1, *Petroleum and related products — Precision of measurement methods and results — Part 1: Determination of precision data in relation to methods of test*

ISO 4259-3, *Petroleum and related products — Precision of measurement methods and results — Part 3: Monitoring and verification of published precision data in relation to methods of test*

ISO 4259-4, *Petroleum and related products — Precision of measurement methods and results — Part 4: Use of statistical control charts to validate 'in-statistical-control' status for the execution of a standard test method in a single laboratory*

3 Begriffe

Für die Anwendung dieses Dokuments gelten die Begriffe nach ISO 4259-1 und die folgenden Begriffe.

ISO und IEC stellen terminologische Datenbanken für die Verwendung in der Normung unter den folgenden Adressen bereit:

- ISO Online Browsing Platform: verfügbar unter <https://www.iso.org/obp>
- IEC Electropedia: verfügbar unter <https://www.electropedia.org/>

3.1

Multi-Sample-Multi-Lab-Studie

Vergleichsstudie zwischen Laboren

MSMLS, en: multi-sample-multi-lab study

Untersuchung, in der ein oder mehrere Leistungsmerkmale auf Grundlage von Analyseergebnissen mehrerer Proben und mehrerer Labore bestimmt werden

Anmerkung 1 zum Begriff: Unter bestimmten Bedingungen erfüllen Ringversuche und Eignungsprüfungsprogramme diese Definition einer Multi-Sample-Multi-Lab-Studie.

3.2

Ringversuch

ILS, en: inter laboratory study

Untersuchung, die so angelegt ist, dass durch die statistische Auswertung der Ergebnisse, die für Teilproben verschiedener Materialien an einem festgelegten Zeitpunkt in verschiedenen Laboren mit einem Standard-Prüfverfahren ermittelt wurden, die Abschätzung der Wiederholbarkeit und der Vergleichbarkeit eines Prüfverfahrens ermöglicht wird

3.3

Eignungsprüfungsprogramm

PTP, en: proficiency testing program

Programm, das für die regelmäßige Bewertung der Eignung teilnehmender Prüflabore zur Durchführung eines Standard-Prüfverfahrens durch statistische Analyse ihrer Prüfergebnisse, die an Teilproben einer Charge homogenen Materials ermittelt wurden, konzipiert wurde. Es wird gelegentlich auch als PT-Studie oder Ringversuchsprogramm (ILCP, en: Inter Laboratory Cross Check Program) bezeichnet

3.4

systematische Abweichung (Bias) zwischen den Verfahren

quantitativer Ausdruck für die Differenz in der Reaktion zweier Prüfverfahren, die durch Anwendung einer Bias-Korrektur, die den Grad der Übereinstimmung zwischen den erwarteten Werten zweier Prüfverfahren, die vorgeben, dieselbe Eigenschaft zu messen, erheblich verbessert, einander angenähert werden können

3.5

Korrektur der systematischen Abweichung

Bias-Korrektur

quantitativer Ausdruck, beruhend auf einem festen und/oder proportionalen Term, der den Grad der Übereinstimmung zwischen den erwarteten Werten zweier Prüfverfahren, die vorgeben, dieselbe Eigenschaft zu messen, signifikant verbessert

3.6

Korrelationskoeffizient

r

statistisches Maß für die Stärke und Richtung der Beziehung zweier Variablen

Anmerkung 1 zum Begriff: Die Werte liegen immer im Bereich -1 (starke negative Beziehung) bis +1 (starke positive Beziehung). Werte bei oder nahe Null deuten auf eine schwache oder nichtlineare Beziehung hin.

Anmerkung 2 zum Begriff: In diesem Dokument wird das Symbol r für die Bezeichnung des Korrelationskoeffizienten entsprechend den statistischen Konventionen verwendet. Dieses Symbol r sollte nicht mit der Wiederholgrenze verwechselt werden. Nur, wenn das Symbol r tiefergestellt ist, bezieht es sich auf die Wiederholgrenze.

3.7

Standardfehler (des Mittelwerts)

SE, en: standard error

Inferenzstatistik zur Abschätzung der Standardabweichung einer Probenverteilung des Mittelwerts

3.8

Probenstandardabweichung

s_i
Abschätzung der Standardabweichung der Grundgesamtheit unter Verwendung des Probenmittelwerts und des Probenumfangs

Anmerkung 1 zum Begriff: Wird auch als Standardabweichung der Probe bezeichnet.

3.9

Vergleichbarkeit zwischen den Verfahren

R_{XY}
quantitativer Ausdruck des Zufallsfehlers, verbunden mit der Differenz zwischen zwei Ergebnissen, welche durch verschiedene Bearbeiter unter Verwendung unterschiedlicher Ausrüstungen und unter Anwendung der beiden Verfahren X bzw. Y erhalten werden; bei jedem dieser Verfahren wird ein Einzelergebnis an einer identischen Prüfprobe erhalten, wenn die Verfahren beurteilt wurden und eine geeignete Bias-Korrektur auf das Ergebnis von Verfahren X dementsprechend angewandt wurde; er ist definiert als der Grenzwert für die Differenz zwischen zwei solchen einzelnen und unabhängigen Ergebnissen, für die erwartet wird, dass aufgrund der zufälligen Schwankung bei üblicher und korrekter Durchführung beider Prüfverfahren eine Wahrscheinlichkeit von 5 % für ihre Überschreitung besteht

3.10

Residuensumme der Quadrate

Summe der quadrierten Residuen

SSR, en: sum of squared residuals

Statistik zur Quantifizierung des Grades der Übereinstimmung zwischen den Ergebnissen zweier Prüfverfahren nach einer Bias-Korrektur unter Verwendung der dementsprechenden Methodik

Anmerkung 1 zum Begriff: Die SSR wird als Optimalitätskriterium bei der Auswahl von Parametern und Modellen für die Bias-Korrektur verwendet.

3.11

Gesamtsumme der Quadrate

SST, en: total sum of squares

Statistik zur Quantifizierung des Informationsgehalts der Ringversuche in Form der Gesamtvariation der Probenmittelwerte im Verhältnis zum Standardfehler der einzelnen Probenmittelwerte

4 Symbole und Einheiten

Symbol	Erläuterung
X, Y	Bezug auf das Verfahren X bzw. Y
X_{ijk}, Y_{ijk}	Einzelergebnis des Verfahrens X und des Verfahrens Y
X_i, Y_i	Mittelwert der Ergebnisse für die i -te Probe
$\bar{X}, \bar{Y}$	gewichteter Mittelwert über Proben, verwendet für die Berechnung der Gesamtsumme der Quadrate SST_X und SST_Y
$\ddot{X}, \ddot{Y}$	gewichteter Mittelwert über Proben, verwendet für die für die Berechnung des gewichteten Korrelationskoeffizienten r
x_i, y_i	Abweichung der Mittelwerte der Ergebnisse der i -ten Probe von $\bar{X}$ bzw. $\bar{Y}$
$\hat{Y}$	vorhergesagter Wert des Verfahrens Y für eine Probe durch Anwendung der dementsprechend festgelegten Bias-Korrektur auf ein tatsächliches Ergebnis des Verfahrens X für dieselbe Probe
$\hat{Y}_i$	vorhergesagter Mittelwert der i -ten Probe des Verfahrens Y durch Anwendung der dementsprechend festgelegten Bias-Korrektur auf seinen zugehörigen Mittelwert des Verfahrens X
S	Anzahl der Proben im Multi-Lab-Multi-Sample-Datensatz

L_{Xi}, L_{Yi}	Anzahl der Labore, die Ergebnisse für die i -te Probe geliefert haben
n_{Xij}, n_{Yij}	Anzahl der wiederholten Ergebnisse für die i -te Probe von la j der Verfahren X bzw. Y
R_X, R_Y	Vergleichbarkeit der Verfahren X bzw. Y
R_{Xi}, R_{Yi}	Vergleichbarkeit der Verfahren X und Y, bewertet anhand der Mittelwerte der Verfahren X und Y für die i -te Probe
R_{XY}	Abschätzung der Vergleichbarkeit zwischen den Verfahren
s_{RXi}, s_{RYi}	Standardabweichung der Vergleichbarkeit, bewertet anhand der Mittelwerte der Verfahren X und Y für die i -te Probe
s_{rXi}, s_{rYi}	Standardabweichung der Wiederholbarkeit, bewertet anhand der Mittelwerte der Verfahren X und Y für die i -te Probe
SE_{Xi}, SE_{Yi}	Standardfehler der Mittelwerte der i -ten Probe
SSR	Summe der quadrierten Residuen, gewichtete Summe der quadrierten (korrigierten) Differenzen zwischen Mittelweltergebnissen
SST_X, SST_Y	Gesamtsumme der Quadrate, um die gewichteten Mittelwerte X und Y
F	Teststatistik für den Vergleich von Varianzen, definiert durch den Quotienten aus zwei Varianzen
t	ist ein Student-t-Wert bei einem bestimmten Vertrauensniveau und bestimmten Freiheitsgraden
k	Nummer der für die Bias-Korrektur ausgewählten Klasse
ν_X, ν_Y	Freiheitsgrade für die Vergleichbarkeit von Varianzen
w_i	Gewichtung, verbunden mit der Differenz zwischen (korrigierten) Mittelwerten der i -ten Probe
a, b	Parameter einer linearen Korrektur: $\hat{Y} = a + bX$
t_1, t_2	Verhältnis zur Bewertung der Reduzierungen der Quadratsummen
ε_i	genormte Differenz zwischen Y_i und $\hat{Y}_i$, gelegentlich auch als Fehler bezeichnet
A, B, C	Koeffizient einer quadratischen Funktion, der für die iterative Berechnung des Proportionalitätskoeffizienten b für Klasse 1b- und Klasse 2-Korrekturklasse verwendet wird
D	Differenzstatistik zur Bestätigung der Korrelation
A_i^2, A_i^{2*}	Anderson-Darling-Test-Statistik bzw. Anderson-Darling-Anpassungstest-Statistik

5 Überblick zum Verfahren

5.1 Allgemeine Anforderungen

Die statistische Methodik beruht auf der Annahme, dass eine Bias-Korrektur nicht erforderlich ist. Wenn es keine statistischen Hinweise darauf gibt, dass eine Bias-Korrektur die erwartete Übereinstimmung zwischen den beiden Verfahren verbessern würde, wird keine Bias-Korrektur vorgenommen.

Ist eine Bias-Korrektur erforderlich, wird das Parsimonie-Prinzip¹ befolgt, wonach eine einfache Korrektur einer komplexeren vorzuziehen ist. Wird dies nicht beachtet, führt dies im Allgemeinen zu einem Modell, das überangepasst ist und in der Praxis nicht gut funktioniert.

Die dementsprechenden Bias-Korrekturen sind auf eine konstante Korrektur, eine proportionale Korrektur oder eine lineare (proportionale + konstante) Korrektur beschränkt.

1 Der Grundsatz, dass die annehmbarste Erklärung für ein Ereignis, ein Phänomen oder eine Begebenheit die einfachste ist, die die wenigsten Einheiten, Annahmen oder Änderungen umfasst.

Die dementsprechenden Bias-Korrektur-Verfahren sind methodensymmetrisch in dem Sinne, dass äquivalente Korrekturen erhalten werden, unabhängig davon, welches Verfahren zur Anpassung an das andere Verfahren korrigiert wird.

Die in diesem Dokument beschriebene Methodik ist nur anwendbar, wenn der mit den einzelnen Mittelwerten verbundene Standardfehler bekannt ist oder berechnet werden kann und die mit allen Standardfehlern verbundenen Freiheitsgrade mindestens 30 betragen.

Diese Methodik wird auf eine Datenquelle angewandt, die aus einer MSMLS abgeleitet wurde. Die Untersuchung muss an mindestens zehn (10) unabhängigen Materialien durchgeführt werden, welche die sich überschneidenden Anwendungsbereiche der Prüfverfahren abdecken. Die Ergebnisse müssen von mindestens sechs (6) Laboren stammen, die die einzelnen Verfahren anwenden.

Die Ergebnisse werden mit demselben Vergleichsprobensatz erhalten, und es wird empfohlen, dass nicht beide Prüfverfahren von demselben Labor durchgeführt werden. Ist dies der Fall, muss darauf geachtet werden, dass die Unabhängigkeit der Prüfergebnisse sichergestellt ist, z. B. durch Doppelblindprüfungen von Proben in zufälliger Reihenfolge.

Dieses Verfahren darf nicht auf Grundlage einer vorläufig oder vorübergehend veröffentlichten Präzisionserklärung angewendet werden. Einer vorläufigen oder vorübergehenden Präzisionserklärung fehlt in der Regel die Größenordnung der verwendeten Datenmenge, so dass nicht genügend Freiheitsgrade zur Verfügung stehen.

Die Kombination mehrerer Datenquellen ist zulässig, solange die für den Datensatz festgelegten Qualitätsanforderungen erfüllt werden.

Die von den einzelnen Laboren angewandten Prüfverfahren müssen einer statistischen Kontrolle unterliegen, die den Anforderungen der ISO 4259-4 entspricht.

Diese Methodik erfordert Daten mit einer ausreichenden Auflösung, um eine statistisch aussagekräftige Beobachtung der Streuung zu ermöglichen. Eine statistisch aussagekräftige Streuung setzt voraus, dass die Gesamtanzahl der eindeutigen Werte in einem Datensatz, d. h. die Laborergebnisse der einzelnen Proben für die einzelnen Prüfverfahren, ausreichend groß sein sollte. Ist die Anzahl der Einzelwerte im Datensatz nach Ansicht des Analytikers unzureichend, müssen die Daten erneut bei den betreffenden Laboren mit ausreichender Auflösung angefordert werden. Stehen die Daten nur mit unzureichender Auflösung zur Verfügung, sollte diese Auswertung nicht fortgesetzt werden.

Falls die Daten für das Verfahren aus einem Ringversuch stammen, müssen die Anforderungen der ISO 4259-1 oder einer gleichwertigen Norm erfüllt sein, einschließlich Prüfung auf das Vorliegen extremer Proben auf Grundlage von Einfluss² und Cook-Abstand³ sowie Probenverteilung. Unter diesen Bedingungen gelten die zusätzlichen Anforderungen an PTP-Daten nicht.

Die Verfahren sollen von einem Analytiker durchgeführt werden, der über ausreichende Kenntnisse der in diesem Dokument beschriebenen statistischen Werkzeuge und Theorien verfügt.

5.2 Zusätzliche Anforderungen für PTP-Daten

Die statistischen Berechnungen sind auch für diese Beurteilung anwendbar, vorausgesetzt, die Ergebnisse und die zugehörigen Statistiken der einzelnen Prüfverfahren stammen aus einem Eignungsprüfungsprogramm (PTP), das die Anforderungen der ISO 4259-3 oder eines gleichwertigen Verfahrens erfüllt. Ein Merkmal der

2 Der Einfluss ist ein Maß dafür, wie weit die unabhängigen Variablen einer Beobachtung von denen der anderen Beobachtungen entfernt sind.

3 Der Cook-Abstand ist eine Abschätzung des Einflusses eines Datenpunktes und wird im Rahmen der Referenz verwendet, um einflussreiche Datenpunkte zu identifizieren, deren Gültigkeit besonders prüfenswert ist.

aus einem derartigen PTP gewonnenen Daten ist, dass für die einzelnen Proben ein Einzelergebnis durch die einzelnen Labore für die einzelnen Prüfverfahren geliefert wird.

Für die Verwendung von Daten aus Eignungsprüfungsprogrammen (PTP) gelten die folgenden zusätzlichen und übergeordneten Bedingungen:

- die Ergebnisse müssen von mindestens zehn (10) Laboren stammen, die die einzelnen Verfahren anwenden;
- der Einfluss der einzelnen Proben im Datensatz überschreitet nicht den Grenzwert von 0,5, siehe Prüfung auf das Vorliegen extremer Proben in 5.2.1;
- die Anderson-Darling-Statistiken für die Prüfungen auf Normalverteilung der Laborergebnisse je Probe $\leq 1,12$, siehe Prüfung auf Verteilung der Laborergebnisse in 5.2.2;
- die Probenstandardabweichungen liegen bei mindestens 80 % der Proben auf dem Signifikanzniveau 0,05 nicht signifikant über den veröffentlichten Standardabweichungen für die Vergleichbarkeit, siehe Vergleich der Präzision in 5.2.3.

5.2.1 Prüfung auf das Vorliegen extremer Proben

Der Einflusswert h_i für die einzelnen Proben i im Datensatz wird untersucht und darf den Grenzwert von 0,5 nicht überschreiten. Für jedes der beiden Verfahren wird der Mittelwert der Laborergebnisse für die einzelnen Proben berechnet. Anschließend werden die einzelnen Labormittelwerte je Probe über beide Prüfverfahren gemittelt.

Der Einflusswert h_i ergibt sich aus:

$$h_i = \frac{1}{S} + \frac{(Z_i - \bar{Z})^2}{\sum_{k=1}^S (Z_k - \bar{Z})^2} \quad (1)$$

Dabei ist

h_i der Einfluss der Probe i , $i=1 \dots S$;

S die Gesamtanzahl der Proben;

Z_i der natürliche Logarithmus (Ln) des Probenmittelwerts, gemittelt über beide Verfahren;

$\bar{Z}$ der Gesamtmittelwert des natürlichen Logarithmus von Z_i .

Werden eine oder mehrere Proben als extrem eingestuft, so müssen sie eliminiert werden, und das Verfahren sollte wiederholt werden. Die Mindestanzahl der verbleibenden Proben muss berücksichtigt werden. Kann die Mindestanzahl der Proben nicht mehr eingehalten werden, muss das Verfahren abgebrochen werden.

5.2.2 Prüfung auf Verteilung der Laborergebnisse

Die Verteilung der Laborergebnisse für die einzelnen Proben wird auf Normalität geprüft, indem die Anpassungsgüte der Normalverteilung unter Verwendung der Anderson-Darling-Statistik⁴ für die einzelnen Proben bestätigt wird.

Es wird angenommen, dass die Verteilung der Ergebnisse einer Normalverteilung folgt, wenn der entsprechende Wert $\leq 1,12^5$ ist.

Die Teststatistik A_i^{2*} wird wie folgt berechnet:

$$A_i^{2*} = A_i^2 \left(1 + \frac{0,75}{N_i} + \frac{2,25}{N_i^2} \right) \quad (2)$$

Dabei ist

N_i die Gesamtanzahl der Laborergebnisse im Datensatz;

$$A_i^2 = -N - \frac{1}{N} \sum_{i=1}^N (2i - 1) [\ln(F(x_i)) + \ln(1 - F(x_{N-i+1}))];$$

$F(x_i)$ die kumulative Normalverteilungsfunktion auf Grundlage von Probenmittelwert und Standardabweichung;

x_i die in aufsteigender Reihenfolge sortierten Daten, $x_1 \leq x_2 \leq x_3 \dots \leq x_N$.

Ergibt diese Prüfung, dass die Verteilung einer oder mehrerer Proben das zuvor genannte Kriterium nicht erfüllt, muss diese Probe eliminiert werden. Die Mindestanzahl der Proben für dieses Verfahren sollte berücksichtigt werden. Kann die Mindestanzahl der Proben nicht mehr eingehalten werden, muss das Verfahren abgebrochen werden.

Daten mit unzureichender Auflösung aufgrund von Rundungen können die Normalitätsbewertungsstatistiken überbewerten. Zu Auflösungsbestimmungen siehe 5.1.

5.2.3 Vergleich der Präzision

Die Probenstandardabweichungen s_i sollten die veröffentlichten Vergleichbarkeitsstandardabweichungen s_{Ri} für mindestens 80 % der Proben bei einem Signifikanzniveau von 0,05 unter Verwendung eines F-Tests nicht signifikant überschreiten.

Da für die einzelnen Proben, für die $s_i \leq s_{Ri}$ ist, die Standardabweichung s_i niemals signifikant größer als s_{Ri} sein wird, braucht das folgende F -Verhältnis nur für die Proben berechnet werden, für die $s_i > s_{Ri}$ ist:

$$F = \frac{s_i}{s_{Ri}} \quad (3)$$

Dabei ist

s_i die Standardabweichung der Probe i , berechnet aus den Laborergebnissen;

⁴ Der Anderson-Darling-Test ist ein statistischer Test, mit dem geprüft wird, ob eine bestimmte Datenprobe einer bestimmten Wahrscheinlichkeitsverteilung entnommen ist. Im Rahmen dieses Dokuments wird dieser Test als Prüfung auf Normalität verwendet, wobei die Parameter der Wahrscheinlichkeitsverteilung (Mittelwert und Standardabweichung) anhand der Probe geschätzt werden. Siehe [7].

⁵ Der kritische Wert von 1,12 basiert auf einem Signifikanzniveau von etwa 1 %, wobei die Auswirkungen der Rundung der Eingabedaten auf die Auflösung berücksichtigt werden.

s_{Ri} die veröffentlichte Standardabweichung der Vergleichbarkeit, bewertet auf Konzentrationsniveau der gemittelten Ergebnisse für Probe i .

Die Anzahl der Freiheitsgrade in Verbindung mit s_i ist gleich $N-1$, wobei N die Anzahl der Ergebnisse für die Probe i ist.

Die Anzahl der zu s_{Ri} gehörenden Freiheitsgrade wird vorzugsweise aus der veröffentlichten Präzisionserklärung des Prüfverfahrens oder dem zugrunde liegenden Forschungsbericht entnommen. Wird s_{Ri} nicht als solches angegeben, kann s_{Ri} auf Grundlage der veröffentlichten Vergleichbarkeit R_i geschätzt werden, und zwar nach der Gleichung $s_{Ri} = R_i / (t\sqrt{2})$, wobei t den Student-t-Wert bei einem Vertrauensniveau von 0,05 und die mit R_i verbundenen Freiheitsgrade darstellt.

Sind im letztgenannten Fall die Freiheitsgrade für R_i nicht bekannt, darf der Mindestwert von 30 angesetzt werden, und die veröffentlichte Standardabweichung der Vergleichbarkeit wird durch $s_{Ri} = (R_i/2,888)$ geschätzt.

Wird das vorgenannte Kriterium bei einer oder mehreren Proben nicht erfüllt, so sind die fehlerhaften Proben zu eliminieren. Die Mindestanzahl der Proben für dieses Verfahren sollte berücksichtigt werden. Kann die Mindestanzahl der Proben nicht mehr eingehalten werden, muss das Verfahren abgebrochen werden.

5.3 Kurze Abfolge der Verfahrensschritte

In der folgenden komprimierten Übersicht sind die Verfahrensschritte zusammengefasst. 5.4 enthält ein Flussdiagramm dieser Verfahrensschritte.

1) Prüfung der Eignung der verfügbaren Daten

Die verfügbaren Daten werden in Bezug auf die allgemeinen Anforderungen (siehe 5.1) und gegebenenfalls auf die zusätzlichen Anforderungen bei Verwendung von PTP-Daten (siehe 5.2, 5.2.1, 5.2.2 und 5.2.3) überprüft.

2) Berechnung der Mittelwerte und des Standardfehlers der Proben

Die arithmetischen Mittelwerte der Ergebnisse für die einzelnen gemeinsamen Proben, die mit der einzelnen Verfahren erhalten wurden, werden berechnet (siehe 6.1.1) und Schätzwerte für die Standardfehler dieser Mittelwerte werden berechnet (siehe 6.1.2).

3) Prüfung der Eignung der Daten

Es wird geprüft, ob die Eigenschaften beider Verfahren ausreichend streuen, indem für die einzelnen Verfahren die gewichteten Summen der quadrierten Residuen für die Gesamtabweichung der mittleren Ergebnisse über alle gemeinsamen Proben berechnet werden. Diese Quadratsummen werden mit den Standardfehlern der Mittelwerte für die einzelnen Verfahren verglichen, um sicherzustellen, dass die Proben ausreichend streuen, bevor mit dem Verfahren fortgefahren wird (siehe 6.2.1).

Es wird geprüft, ob eine ausreichende Korrelation zwischen den beiden Verfahren besteht, indem die gewichteten Summen der quadrierten Residuen für die lineare Korrektur mit der Gesamtabweichung der Mittelwerte für beide Verfahren verglichen werden, um sicherzustellen, dass eine ausreichende Korrelation zwischen den beiden Verfahren besteht (siehe 6.2.2).

4) Berechnung der Bias-Korrektur-Statistiken für die einzelnen Bias-Korrektur-Klassen

Der Grad der Übereinstimmung der Mittelwerte der einzelnen Verfahren wird unter Verwendung geeigneter gewichteter Summen der quadrierten Residuen beurteilt. Diese Quadratsummen werden aus den Daten zunächst ohne Bias-Korrektur, dann mit einer konstanten Bias-Korrektur, dann ggf. mit einer proportionalen Korrektur und schließlich mit einer linearen (proportionalen + konstanten) Korrektur berechnet (siehe 6.3).

5) Auswahl der geeigneten Bias-Korrektur-Klasse

Die am besten geeignete Bias-Korrektur wird auf Grundlage der gewichteten Summe der quadrierten Residuen aus den einzelnen Bias-Korrekturen und den entsprechenden t - und F -Tests ausgewählt (siehe 6.4).

6) Prüfung auf Normalverteilung der Residuen

Die (gewichteten) Residuen je Probe werden auf Zufälligkeit geprüft. Die Residuen sind definiert als die Differenz zwischen jedem einzelnen Y_i und dem Bias-korrigierten X_i . Die Prüfung auf Zufälligkeit wird mit dem Anderson-Darling-Test auf Normalverteilung der Residuen durchgeführt. Wenn sich herausstellt, dass die gewichteten Residuen nicht normalverteilt sind, wird dieses Verfahren als beendet angesehen (siehe 6.5).

7) Prüfung auf probenspezifische systematische Abweichungen

Die gewichtete Summe der quadrierten Residuen wird bewertet, um festzustellen, ob zusätzliche ungeklärte Abweichungsquellen in den Residualdaten verbleiben (siehe 6.6).

Jede verbleibende, ungeklärte Abweichung wird probenspezifischen systematischen Abweichungen zugeschrieben, die auch als Verfahren/Material-Wechselwirkungen oder Matrixeffekte bezeichnet werden. Wenn sich probenspezifische systematische Abweichungen als konsistent mit einem Modell mit zufälligen Auswirkungen erweisen, wird ihr Beitrag zur Vergleichbarkeit zwischen den Verfahren geschätzt und zu einer Gesamtschätzung der Vergleichbarkeit zwischen den Verfahren kumuliert.

8) Bestimmung der Vergleichbarkeit zwischen den Verfahren

Es wird die Vergleichbarkeit zwischen den Verfahren unter Berücksichtigung möglicher probenspezifischer Abweichungen berechnet.

Falls sich herausstellt, dass die Residuen normalverteilt sind und keine probenspezifischen systematischen Abweichungen vorliegen, wird die Vergleichbarkeit zwischen den Verfahren nach Gleichung (40) berechnet.

Falls die Residuen normalverteilt sind und probenspezifische systematische Abweichungen vorliegen, wird die Vergleichbarkeit zwischen den Verfahren nach Gleichung (41) berechnet.

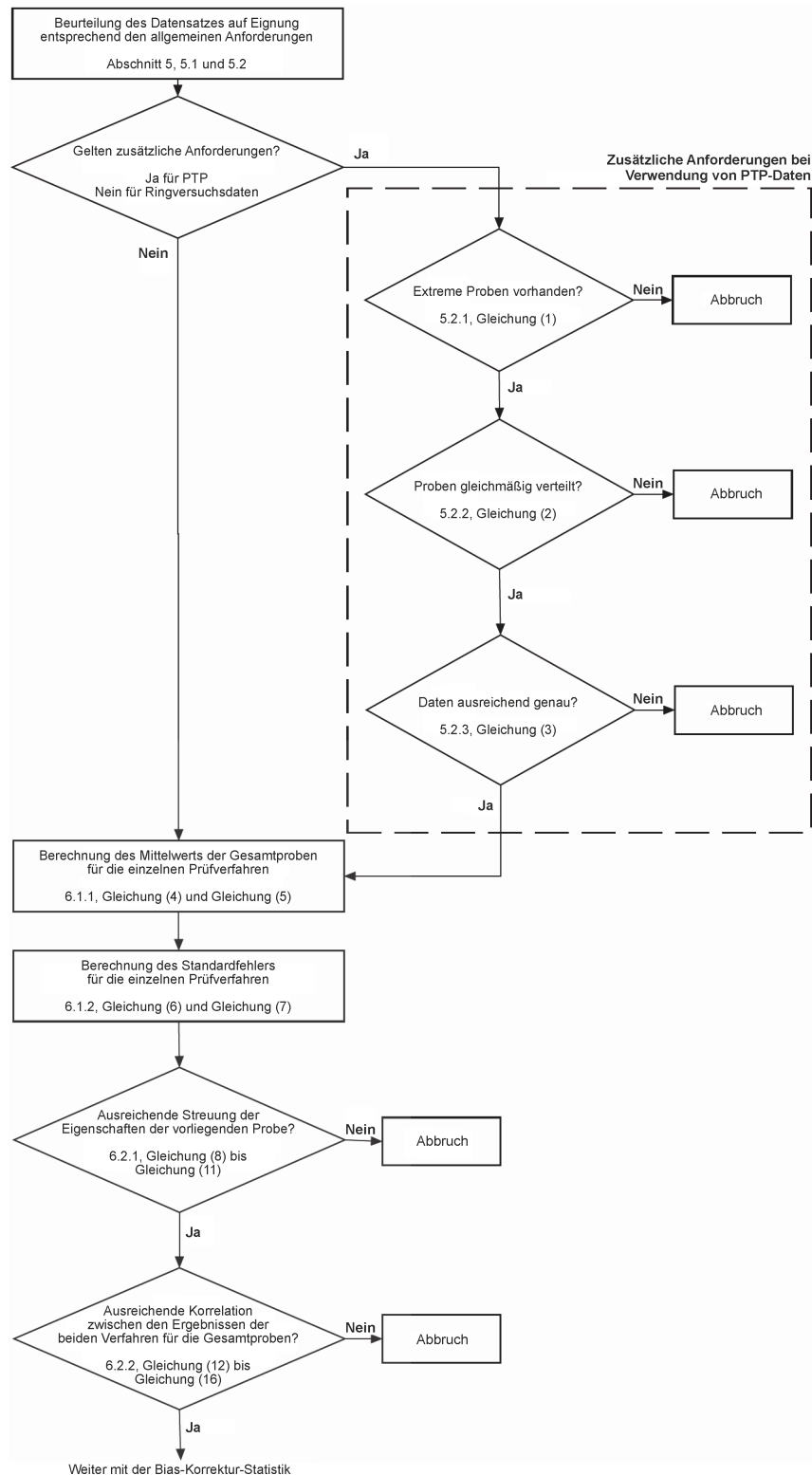
9) Angaben im Prüfbericht

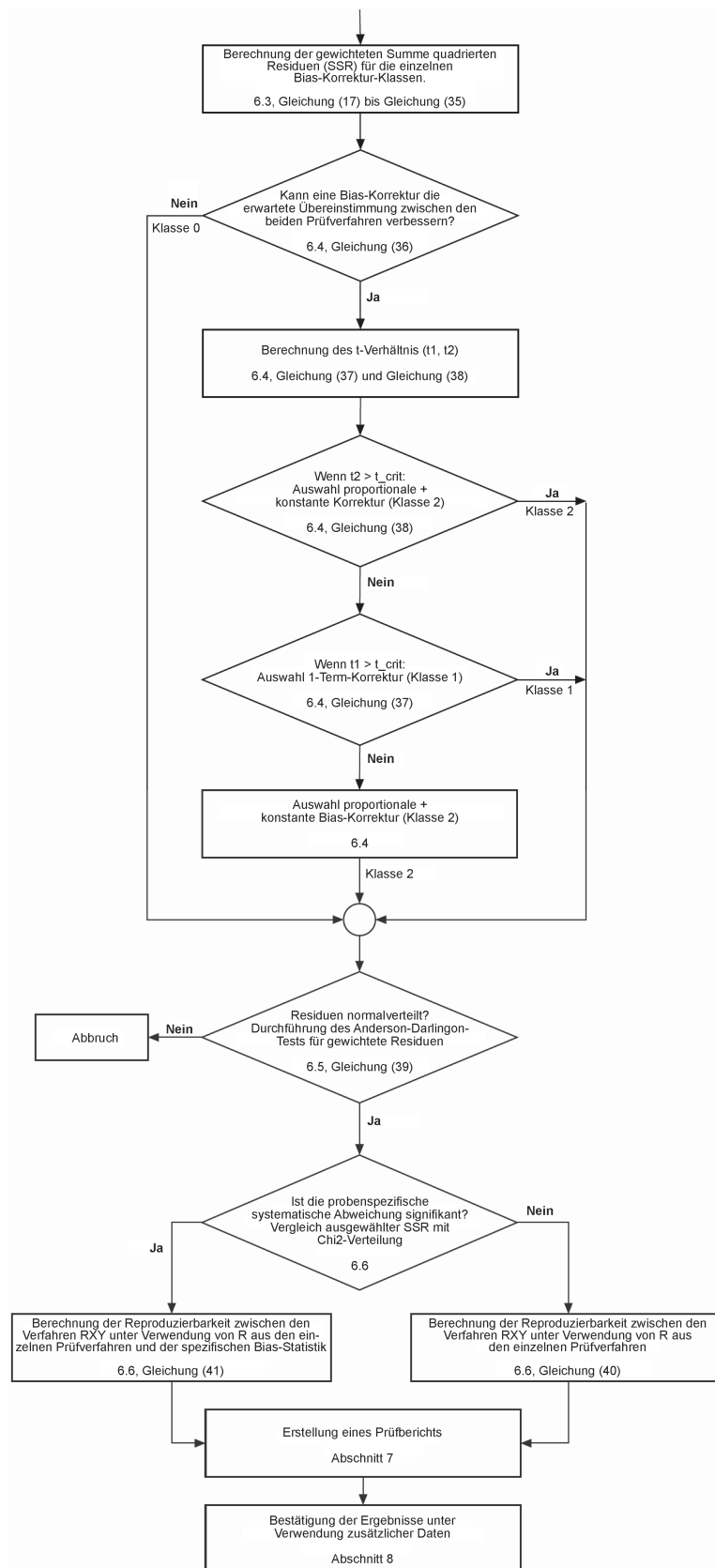
Die Ergebnisse dieses Verfahrens werden im Abschnitt über Präzision und Abweichung der entsprechenden Norm(en) angegeben (siehe Abschnitt 7).

10) Bestätigung der Korrelation

Die Ergebnisse der Beurteilung werden in regelmäßigen Abständen von den Anwendern der Korrelation bestätigt, indem die Differenzstatistiken mittels Regelkarten überwacht werden (siehe Abschnitt 8).

5.4 Flussdiagramm des Verfahrens





6 Durchführung

6.1 Probenmittelwert und Standardfehler

Die Probenmittelwerte und der Standardfehler werden aus den Ergebnissen der MSMLS berechnet. Zur Abschätzung der Standardfehler dieser Mittelwerte werden veröffentlichte Schätzwerte für die Präzision verwendet.

Die Probenmittelwerte X_i und Y_i werden berechnet und der Standardfehler SE_{X_i} und SE_{Y_i} werden bestimmt. Dieser Prozess umfasst möglicherweise die Eliminierung von Ergebnissen als Ausreißer.

X_{ijk} sei das k -te Ergebnis für das i -te übliche Material des j -ten Labors für die Verfahren X. Gleiches gilt für Y_{ijk} .

n_{Xij} sei die Anzahl der Ergebnisse (Wiederholungen) für das i -te Material aus dem j -ten Labor für das Verfahren X. Gleiches gilt für n_{Yij} .

L_{Xi} sei die Anzahl der Labore in dem Verfahren X, die mindestens ein Ergebnis für das i -te Material haben. Gleiches gilt für L_{Yi} .

S sei die Gesamtanzahl der Proben, die beiden Datensätzen gemeinsam sind.

ANMERKUNG Das i -te Material ist für beide Datensätze gleich, aber das j -te Labor in einem Datensatz ist nicht notwendigerweise das gleiche Labor wie das j -te Labor in einem anderen Datensatz.

6.1.1 Berechnung der Mittelwerte

Der arithmetische Mittelwert der Ergebnisse des Verfahrens X für die i -te Probe ist:

$$X_i = \frac{1}{L_{Xi}} \sum_j \frac{\sum_k X_{ijk}}{n_{Xij}} \quad (4)$$

wobei X_i der Mittelwert der Feldmittelwerte für die i -te Probe nach Verfahren X ist.

Analog dazu ist der Mittelwert der Ergebnisse des Verfahrens Y für die i -te Probe:

$$Y_i = \frac{1}{L_{Yi}} \sum_j \frac{\sum_k Y_{ijk}}{n_{Yij}} \quad (5)$$

6.1.2 Berechnung der Standardfehler

Die Standardfehler sind den Standardabweichungen der Mittelwerte zugeordnet und werden wie folgt berechnet.

Wenn s_{RXi} die Standardabweichung der Vergleichbarkeit des Verfahrens X und s_{rXi} die Standardabweichung der Wiederholbarkeit ist, dann berechnet sich der Standardfehler für X_i wie folgt:

$$SE_{X_i} = \sqrt{\frac{1}{L_{Xi}} \left[s_{RXi}^2 - s_{rXi}^2 \left(1 - \frac{1}{L_{Xi}} \sum_j \frac{1}{n_{Yij}} \right) \right]} \quad (6)$$

Der geschätzte Standardfehler für Y_i berechnet sich nach der analogen Gleichung:

$$SE_{Y_i} = \sqrt{\frac{1}{L_{Yi}} \left[s_{RYi}^2 - s_{rYi}^2 \left(1 - \frac{1}{L_{Yi}} \sum_j \frac{1}{n_{Yij}} \right) \right]} \quad (7)$$

Die Standardabweichungen der Wiederholbarkeit und der Vergleichbarkeit werden aus der veröffentlichten Wiederholbarkeit und der veröffentlichten Vergleichbarkeit berechnet, indem diese durch $t\sqrt{2}$ dividiert werden. Dabei bezieht sich t auf den Student- t -Wert bei einem Vertrauensniveau von 0,05 und die Anzahl der Freiheitsgrade, die mit den Präzisionswerten verbunden sind.

Für den Fall, dass die veröffentlichte Wiederholbarkeit und Vergleichbarkeit bekannt sind, aber die Anzahl der Freiheitsgrade, die mit diesen Präzisionswerten verbunden sind, unbekannt sind, ist ein Wert von 30 für die Anzahl der Freiheitsgrade zulässig.

Da die veröffentlichte Wiederholbarkeit und Vergleichbarkeit mit den Mittelwerten der mit dem Verfahren X erhaltenen Ergebnisse X_i streuen können, könnten die SE_{Xi} selbst dann, wenn die L_{Xi} für alle Materialien und die n_{Xij} für alle Labore und alle Materialien gleich wären, von einem Material zum anderen verschieden sein. Das Gleiche gilt auch für das Verfahren Y.

6.2 Eignung der Daten

6.2.1 Prüfung auf Eigenschaftsstreuung

Für die einzelnen Verfahren wird die gewichtete Gesamtquadratsumme (SST) berechnet und bestimmt, ob die Proben durch beide Verfahren voneinander unterschieden werden können. Die Gesamtquadratsummen werden wie folgt berechnet:

$$SST_X = \sum_i \left(\frac{X_i - \bar{X}}{SE_{Xi}} \right)^2 \quad (8)$$

$$SST_Y = \sum_i \left(\frac{Y_i - \bar{Y}}{SE_{Yi}} \right)^2 \quad (9)$$

Die gewichteten Mittelwerte $\bar{X}$ und $\bar{Y}$ berücksichtigen ihre zugehörigen Standardfehler SE^2_{Xi} und SE^2_{Yi} und werden wie folgt berechnet:

$$\bar{X} = \frac{\sum_i \left(\frac{X_i}{SE^2_{Xi}} \right)}{\sum_i \left(\frac{1}{SE^2_{Xi}} \right)} \quad (10)$$

$$\bar{Y} = \frac{\sum_i \left(\frac{Y_i}{SE^2_{Yi}} \right)}{\sum_i \left(\frac{1}{SE^2_{Yi}} \right)} \quad (11)$$

$F = SS_{TX}/(S-1)$ wird mit dem 95. Perzentil der F -Verteilung nach Fisher mit $(S-1)$ und v_X Freiheitsgraden für den Zähler bzw. den Nenner verglichen, wobei v_X die Freiheitsgrade für die Vergleichbarkeitsvarianz für das Verfahren X ist.

- Liegt F nicht über dem 95. Perzentil, dann ist das Verfahren X nicht ausreichend genau, um zwischen den S Proben zu unterscheiden. Mit diesem Verfahren nicht fortfahren, da keine aussagekräftigen Ergebnisse erzielt werden können.
- Liegt F über dem 95. Perzentil, dann ist das Verfahren X ausreichend genau, um zwischen den S Proben zu unterscheiden. Mit der Prüfung auf Korrelation in 6.2.2 fortfahren.

In ähnlicher Weise wird $F = SST_Y / (S-1)$ mit dem 95. Perzentil der F -Verteilung nach Fisher verglichen, wobei die Freiheitsgrade der Vergleichbarkeitsvarianz des Verfahrens Y v_Y anstelle von v_X verwendet werden. Analog dazu mit dem Verfahren nicht fortfahren, wenn F das 95. Perzentil nicht überschreitet.

6.2.2 Korrelation der Prüfverfahren

Um zu prüfen, ob die beiden Verfahren ausreichend korreliert sind, wird der gewichtete Korrelationskoeffizient r wie folgt berechnet:

$$r = \frac{\sum_i (w_i (X_i - \bar{X}) (Y_i - \bar{Y}))}{\sqrt{\sum_i (w_i (X_i - \bar{X})^2) \sum_i (w_i (Y_i - \bar{Y})^2)}} \quad (12)$$

Die gewichteten Mittelwerte $\bar{X}$ und $\bar{Y}$ berücksichtigen jeweils die beiden Standardfehler $SE^2_{X_i}$ und $SE^2_{Y_i}$ und werden wie folgt berechnet:

$$\bar{X} = \frac{\sum w_i X_i}{\sum w_i} \quad (13)$$

$$\bar{Y} = \frac{\sum w_i Y_i}{\sum w_i} \quad (14)$$

$$w_i = \frac{1}{SE^2_{Y_i} + SE^2_{X_i}} \quad (15)$$

Der gewichtete Korrelationskoeffizient r wird zur Berechnung der F -Statistik verwendet:

$$F = \frac{(S-2) r^2}{(1-r^2)} \quad (16)$$

F wird mit dem 99. Perzentil der F -Verteilung nach Fisher mit 1 und $S-2$ Freiheitsgraden im Zähler bzw. im Nenner verglichen.

- Liegt F unter dem 99. Perzentilwert, dann lässt dieses Verfahren den Schluss zu, dass die Prüfverfahren zu uneinheitlich sind, um die Ergebnisse des einen Verfahrens zur Vorhersage der Ergebnisse des anderen Verfahrens zu verwenden. Mit diesem Verfahren nicht fortfahren.
- Liegt F über dem 99. Perzentilwert, kann angenommen werden, dass die beiden Prüfverfahren ausreichend korreliert sind, um das Verfahren mit den Bias-Korrektur-Statistiken in 6.3 fortzusetzen.

An diesem Punkt des Verfahrens kann es aufschlussreich sein, die Daten graphisch darzustellen: Mittelwerte des Verfahrens X (X_i) versus Mittelwerte des Y-Verfahrens (Y_i) der Proben.

6.3 Bias-Korrektur-Statistik

6.3.1 Allgemeines

Die gewichtete Summe der quadrierten Residuen wird für jede der folgenden Klassen der Bias-Korrektur-Methodik berechnet.

6.3.2 Klasse 0 — Keine Bias-Korrektur.

Für die einzelnen Proben i werden die Gewichtungen w_i berechnet:

$$w_i = \frac{1}{SE^2_{Y_i} + SE^2_{X_i}} \quad (17)$$

SSR_0 wird berechnet:

$$SSR_0 = \sum_i w_i (X_i - Y_i)^2 \quad (18)$$

6.3.3 Klasse 1a —Konstante Bias-Korrektur.

Unter Verwendung der Gewichtungen w_i aus der Klasse-0-Korrektur (17) wird die konstante Bias-Korrektur a errechnet:

$$a = \frac{\sum_i w_i (Y_i - X_i)}{\sum_i w_i} = \frac{\sum_i w_i Y_i}{\sum_i w_i} - \frac{\sum_i w_i X_i}{\sum_i w_i} \quad (19)$$

SSR_{1a} wird berechnet:

$$SSR_{1a} = \sum_i w_i (Y_i - (X_i + a))^2 \quad (20)$$

6.3.4 Klasse 1b —Proportionale Bias-Korrektur.

Die Berechnungen dieses Unterabschnitts sind nur dann sinnvoll, wenn beide der folgenden Bedingungen zutreffen:

1. Die gemessene Eigenschaft nimmt nur positive Werte an, und
2. Ein Eigenschaftswert von Null hat eine physikalische Bedeutung (z. B. Konzentrationen bestimmter Bestandteile)

Die Berechnungen umfassen die iterative Berechnung der Gewichtungen w_i und der proportionalen Korrektur b .

$b = 1$ setzen.

Die Gewichtung w_i wird für die einzelnen Proben i berechnet:

$$w_i = \frac{1}{SE_{Yi}^2 + b^2 SE_{Xi}^2} \quad (21)$$

Die drei folgenden Summen werden berechnet:

$$A = \sum_i w_i^2 X_i Y_i SE_{Xi}^2 \quad (22)$$

$$B = \sum_i w_i^2 (X_i^2 SE_{Yi}^2 - Y_i^2 SE_{Xi}^2) \quad (23)$$

$$C = - \sum_i w_i^2 X_i Y_i SE_{Yi}^2 \quad (24)$$

b_0 wird berechnet:

$$b_0 = \frac{-B + \sqrt{B^2 - 4AC}}{2A} \quad (25)$$

Wenn $|b - b_0| > 0,001 \cdot b$, wird b durch b_0 ersetzt und zu (21) zurückgekehrt. Andernfalls kann die Iteration abgebrochen werden, da eine weitere Iteration keine sinnvolle Verbesserung bringt. b wird durch b_0 ersetzt und die Iteration wird abgebrochen.

Die endgültigen Gewichtungen w_i werden wie in (21) berechnet.

SSR_{1b} wird berechnet:

$$SSR_{1b} = \sum_i w_i (Y_i - bX_i)^2 \quad (26)$$

6.3.5 Klasse 2 — Proportionale und konstante Bias-Korrektur.

Dies umfasst die iterative Berechnung der Gewichtungen w_i , der gewichteten Mittelwerte von X_i und Y_i und des proportionalen Terms b .

$b = 1$ setzen.

Die Gewichtung w_i wird für die einzelnen Proben i berechnet:

$$w_i = \frac{1}{SE_{yi}^2 + b^2 SE_{xi}^2} \quad (27)$$

Die gewichteten Mittelwerte von X_i und Y_i , $\bar{X}$ bzw. $\bar{Y}$, werden berechnet:

$$\bar{X} = \frac{\sum_i w_i X_i}{\sum_i w_i}, \bar{Y} = \frac{\sum_i w_i Y_i}{\sum_i w_i} \quad (28)$$

Die Abweichungen von den gewichteten Mittelwerten werden berechnet:

$$\Delta x_i = X_i - \bar{X}, \Delta y_i = Y_i - \bar{Y} \quad (29)$$

Die drei Summen A , B und C werden berechnet:

$$A = \sum_i w_i^2 \Delta x_i \Delta y_i SE_{xi}^2 \quad (30)$$

$$B = \sum_i w_i^2 (\Delta x_i^2 SE_{yi}^2 - \Delta y_i^2 SE_{xi}^2) \quad (31)$$

$$C = - \sum_i w_i^2 \Delta x_i \Delta y_i SE_{yi}^2 \quad (32)$$

b_0 wird berechnet:

$$b_0 = \frac{-B + \sqrt{B^2 - 4AC}}{2A} \quad (33)$$

Wenn $|b - b_0| > 0,001 \cdot b$, wird b durch b_0 ersetzt und zu (27) zurückgekehrt, um neue Werte für die Gewichtungen w_i , $\bar{X}$, $a = \bar{Y} - b\bar{X}$, x_i , y_i , und b_0 zu berechnen. Andernfalls kann die Iteration abgebrochen werden, da eine weitere Iteration keine sinnvolle Verbesserung bringt. b wird durch b_0 ersetzt, die Iteration wird abgebrochen und die endgültigen Gewichtungen w_i werden wie in (27) berechnet.

SSR_2 und a werden berechnet:

$$SSR_2 = \sum_i w_i (\Delta y_i - b \cdot \Delta x_i)^2 \quad (34)$$

$$a = \bar{Y} - b\bar{X} \quad (35)$$

6.4 Auswahl der geeigneten Bias-Korrektur-Klasse

Die Summe der quadrierten Residuen aus den einzelnen Bias-Korrektur-Klassen wird verwendet, um die am besten geeignete Bias-Korrektur-Klasse auszuwählen, die den erwarteten Grad der Übereinstimmung zwischen $\hat{Y}$ (dem vorhergesagten Ergebnis des Verfahrens Y unter Verwendung des Ergebnisses des Verfahrens X) und dem tatsächlichen Ergebnis des Verfahrens Y für dasselbe Material verbessern kann. Die Bias-Korrektur-Klassen und die zugehörigen SSR, wie sie zuvor berechnet wurden, sind in Tabelle 1 wiedergegeben.

Tabelle 1 — Bias-Korrektur-Klassen

Bias-Korrektur-Klasse	Geeignete Korrektur	Summe der quadrierten Residuen
<i>Klasse 0</i>	keine Korrektur	SSR_0
<i>Klasse 1a</i>	konstante Bias-Korrektur	SSR_{1a}
<i>Klasse 1b</i>	proportionale Korrektur der systematischen Abweichung	SSR_{1b}
<i>Klasse 2</i>	proportionale Bias-Korrektur	SSR_2

Um zu bestimmen, ob *eine* Bias-Korrektur (Klasse 1a, 1b oder 2) die erwartete Übereinstimmung zwischen den beiden Verfahren signifikant verbessern kann, wird das folgende Verhältnis berechnet:

$$F = \frac{(SSR_0 - SSR_2) / 2}{\frac{SSR_2}{(S-2)}} \quad (36)$$

F wird mit dem oberen 95. Perzentil der F -Verteilung mit 2 und $S-2$ Freiheitsgraden für den Zähler bzw. Nenner verglichen.

- Ist das berechnete F kleiner, wird davon ausgegangen, dass eine Bias-Korrektur der Klasse 1a, 1b oder 2 die erwartete Übereinstimmung zwischen den beiden Verfahren im Vergleich mit *Klasse 0* (keine Bias-Korrektur) nicht ausreichend verbessert. Mit der Prüfung auf Normalverteilung der Residuen in 6.5 fortfahren.
- Ist das berechnete F größer, wird davon ausgegangen, dass eine Korrektur die erwartete Übereinstimmung zwischen den beiden Verfahren verbessern kann, und es werden die folgenden t -Verhältnisse berechnet:

$$t_1 = \sqrt{\frac{SSR_0 - SSR_1}{\frac{SSR_2}{(S-2)}}} \quad (37)$$

$$t_2 = \sqrt{\frac{SSR_1 - SSR_2}{\frac{SSR_2}{(S-2)}}} \quad (38)$$

wobei SSR_1 der kleinere Wert von SSR_{1a} oder SSR_{1b} ist, sofern letzterer geeignet ist und berechnet wurde.

t_2 wird mit dem oberen 97,5. Perzentil der t -Verteilung mit $S-2$ Freiheitsgraden verglichen.

- Ist t_2 größer, wird davon ausgegangen, dass eine Bias-Korrektur der Klasse 2 (proportionale + konstante Korrektur) die erwartete Übereinstimmung im Vergleich mit einer Korrektur mit nur einem Term (konstant oder proportional) verbessern kann (Klasse 1). Mit der Prüfung auf Normalverteilung der Residuen in 6.5 fortfahren.
- Ist t_2 kleiner als das t -Perzentil, wird t_1 mit demselben oberen 97,5. Perzentil der t -Verteilung mit $(S-2)$ Freiheitsgraden verglichen.

- Ist t_1 größer, wird davon ausgegangen, dass eine 1-Term-Bias-Korrektur der Klasse 1 einer Bias-Korrektur der Klasse 2 vorzuziehen ist. Die konstante Korrektur wird verwendet, es sei denn, SSR_{1b} ist geeignet und ist kleiner als SSR_{1a} . Mit der Prüfung auf Normalverteilung der Residuen in 6.5 fortfahren.
- Ist t_1 kleiner, dann ist weder t_1 noch t_2 statistisch signifikant. Eine Bias-Korrektur der Klasse 2 ist einer 1-Term-(konstanten oder proportionalen)Korrektur der Klasse 1 vorzuziehen.

6.5 Bestätigung der Normalverteilung der Residuen

Um gültige Rückschlüsse aus der gewählten Bias-Korrektur zu ziehen, sollten die gewichteten Residuen einer Normalverteilung folgen. $\hat{Y}_i$ seien die Mittelwerte des Verfahrens Y, die aus den entsprechenden Mittelwerten des Verfahrens X X_i unter Verwendung der in 6.4 gewählten Bias-Korrektur vorhergesagt werden.

Die gewichteten Residuen ε_i sind gegeben durch:

$$\varepsilon_i = \sqrt{w_i} (Y_i - \hat{Y}_i) \quad (39)$$

wobei w_i die entsprechenden Gewichtungen sind, die mit der entsprechenden Bias-Korrektur-Klasse verbunden sind.

Der Anderson-Darling-Test auf Normalität wird für die Residuen ε_i durchgeführt. Siehe 5.2.2 für Hinweise zur Berechnung und Auswertung dieser Statistik. Es wird angenommen, dass die Verteilung der Residuen auf dem 5 %-Signifikanzniveau einer Normalverteilung folgt, wenn $A_i^{2*} \leq 0,752$ ist.

- Erfüllt diese Prüfung das zuvor genannte Kriterium, gelten die Residuen als normalverteilt. Mit 6.6 fortfahren, um auf das Vorliegen probenspezifischer Abweichungen zu prüfen.
- Erfüllt diese Prüfung das zuvor genannte Kriterium nicht, gelten die Residuen als nicht normalverteilt. Und dieses Verfahren wird an dieser Stelle als beendet angesehen, da die statistischen Beweise darauf hindeuten, dass eine einzige Vergleichbarkeit zwischen den Verfahren (R_{XY}) nicht gefunden werden kann, die auf alle Materialien, die durch den sich überschneidenden Anwendungsbereich beider Prüfverfahren abgedeckt werden, anwendbar ist. Dies lässt den Schluss zu, dass die Prüfverfahren zumindest bei einigen Materialien nicht dieselbe Eigenschaft messen. Nicht mit 6.6 fortfahren.

6.6 Probenspezifische systematische Abweichungen

Probenspezifische systematische Abweichungen sind durch eine funktionale Abhängigkeit der (gewichteten) Residuen vom Probenwert festgelegt. Für die Prüfung auf das Vorliegen probenspezifischer Abweichungen wird ein χ^2 -Unabhängigkeitstest durchgeführt, bei dem die SSR der in 6.4 ausgewählten Bias-Korrektur-Klasse mit dem 95. Perzentilwert einer χ^2 -Verteilung mit ν Freiheitsgraden verglichen wird.

Dabei ist

ν S für Klasse 0 (keine Bias-Korrektur);

ν $S - 1$ für Klasse 1a oder Klasse 1b (konstante oder proportionale Korrektur);

ν $S - 2$ für Klasse 2 (lineare Korrektur).

- Ist $SSR \leq \chi^2$ -Perzentil, wird die Nullhypothese angenommen. Dies lässt den Schluss zu, dass es keine probenspezifischen systematischen Abweichungen gibt, d. h. dass es keine anderen Abweichungsquellen gibt, die über den Messfehler hinaus statistisch beobachtbar sind.
- Die Vergleichbarkeit zwischen den Verfahren ist nun festgelegt durch:

$$R_{XY} = \sqrt{\frac{R_Y^2 + b^2 R_X^2}{2}} \quad (40)$$

— Dabei ist

R_Y die Vergleichbarkeit der Verfahren Y, bewertet mit dem Wert der Einzelergebnisse Y;

R_X die Vergleichbarkeit der Verfahren X, bewertet mit dem Wert der Einzelergebnisse X;

b der Koeffizient der geeigneten Bias-Korrektur. (Für Bias-Korrekturen der Klasse 0 und der Klasse 1a ist $b=1$).

— Ist $SSR > \chi^2$ -Perzentil, wird die Nullhypothese angenommen. Es gibt deutliche Hinweise darauf, dass die systematischen Abweichungen zwischen den Verfahren durch die Bias-Korrekturen von 6.3 nicht angemessen korrigiert wurden. Mit anderen Worten, die systematischen Abweichungen sind über die gemeinsamen S Proben nicht konsistent. Der Anwender möchte möglicherweise untersuchen, ob die systematischen Abweichungen auf andere beobachtbare Eigenschaften der Proben zurückgeführt werden können. Es könnte sinnvoll sein, sich auf eine kleinere Klasse von Materialien zu beschränken, um eine SSR zu erreichen, die unter dem zuvor erwähnten Chi-Quadrat-Perzentil liegt, um eine Vergleichbarkeit zwischen den Verfahren sicherzustellen. Diese Untersuchungen liegen außerhalb des Anwendungsbereichs dieses Verfahrens, da die Probleme in der Regel nicht statistischer Natur sind.

Überschreitet die SSR den 95. Perzentilwert der entsprechenden χ^2 -Verteilung, spricht vieles dafür, dass andere Quellen als Messfehler zur Abweichung von der erwarteten Übereinstimmung zwischen den beiden Verfahren beitragen. Diese Quellen werden in der Regel probenspezifischen Effekten zugeschrieben (auch bekannt als Matrixeffekte oder Verfahren/Material-Wechselwirkungen). In einigen Fällen können diese probenspezifischen Effekte als *zufällige* Auswirkungen behandelt und somit als zusätzliche Abweichungsquelle in die in diesem Abschnitt beschriebene Vergleichbarkeit zwischen den Verfahren aufgenommen werden. Selbst wenn es angemessen ist, diese probenspezifischen Effekte als zufällig zu behandeln, kann die Zusatzabweichung dazu führen, dass die Vergleichbarkeit zwischen den Verfahren weitaus größer ist als der quadratische Mittelwert der Vergleichbarkeit der Verfahren in Gleichung (40).

Entscheidet sich der Anwender dafür, alle Proben in der Studie beizubehalten, wird die probenspezifische systematische Abweichung als zufällige Varianzkomponente behandelt, und die Vergleichbarkeit zwischen den Verfahren (R_{XY}) wird wie folgt berechnet:

$$R_{XY} = \sqrt{\left(\frac{b^2 R_X^2}{2} + \frac{R_Y^2}{2} \right) \left(1 + \frac{2t^2 (SSR - S + k) S}{(S - k) \sum_i \frac{b^2 R_{Xi}^2 + R_{Yi}^2}{b^2 SE_{Xi}^2 + SE_{Yi}^2}} \right)} \quad (41)$$

Dabei ist

b der Koeffizient der geeigneten Bias-Korrektur. Für Klasse 0 und Klasse 1a gilt $b = 1$;

SSR die SSR der geeigneten Bias-Korrektur-Klasse;

S ist die Anzahl der Proben;

k ist die Nummer der ausgewählten Klasse: Für Klasse 0 $k=0$, für Klasse 1a oder Klasse 1b $k=1$ und für Klasse 2 $k=2$;

SE_X^2, SE_Y^2 die Standardfehler für Verfahren X und Verfahren Y;

R_X, R_Y die Vergleichbarkeit der Verfahren X und Y, bewertet beim Wert der Einzelergebnisse X und Y;
 t der Student- t -Wert bei einem Vertrauensniveau von 0,05 und unendlichen Freiheitsgraden.

Gleichung (40) und Gleichung (41) liefern beide eine Schätzung des Grenzwerts, den etwa 95 % der Differenzen bei üblicher und korrekter Durchführung beider Verfahren überschreiten dürften, wenn eine Partei das Bias-korrigierte Verfahren X und eine andere Partei das Verfahren Y anwendet, auf Materialien, die den für diese Studie verwendeten Proben ähnlich sind. Die Anwendung der Verfahren auf Materialien, die sich wesentlich von den für diese Studie verwendeten Materialien unterscheiden, kann sowohl die mittlere Abweichung als auch die Varianz der Zufallskomponente beeinflussen. Labore, die routinemäßig ein Verfahren durch eine anderes ersetzen, wird empfohlen, die Abweichungen zwischen den Verfahren als regelmäßigen Bestandteil ihres Qualitätssicherungsprogramms zu überwachen.

7 Prüfbericht

Nach Abschluss der Berechnungen wird empfohlen, die Ergebnisse der Bewertung in den Abschnitt „Präzision und Abweichung“ der entsprechenden Norm(en) und/oder Prüfverfahren aufzunehmen.

Für den Fall, dass eines der bewerteten Prüfverfahren als Referenzprüfverfahren angeführt wird und das andere Prüfverfahren eine Alternative darstellt, wird folgende Bezeichnung empfohlen, wobei das Veröffentlichungsjahr für die Norm YYYY durch den Zusatz „-YY“ und das Veröffentlichungsjahr für die Norm XXXX durch den Zusatz „-XX“ angegeben wird.

Berechnung eines 95 % Konfidenzintervalls, das ein einzelnes Ergebnis des Verfahrens Y (bei Durchführung an derselben Probe) unter Verwendung eines einzelnen Bias-korrigierten Ergebnisses des Verfahrens X (vorhergesagtes Y , $\hat{Y}$) und der Vergleichbarkeit zwischen den Verfahren R_{XY} enthalten würde.

$\hat{Y}$ sei ein einzelnes Bias-korrigiertes Ergebnis des Verfahrens X. Ein Intervall, das durch $Y \pm R_{XY}$ begrenzt wird, enthält mit einer Wahrscheinlichkeit von etwa 95 % ein einziges entsprechendes Ergebnis des Verfahrens Y, das mit demselben Material erhalten wurde. Dazu wird $R_{X\hat{Y}}$ nach Gleichung (40) bzw. Gleichung (41) berechnet, wenn R_Y das Bias-korrigierte Ergebnis des Verfahrens X bewertet.

8 Bestätigung der Korrelation

Anwendern, die beabsichtigen, die Bias-korrigierten Ergebnisse des Prüfverfahrens X als Prädiktor für die Ergebnisse des Prüfverfahrens Y (oder umgekehrt) unter Verwendung der durch dieses Prüfverfahren hergestellten Korrelation zu verwenden, wird empfohlen, regelmäßig die Übereinstimmung zwischen dem Prädiktionsergebnis Y ($\hat{Y}$ oder Bias-korrigiertes X) und dem tatsächlichen Y mit den Anforderungen ihrer Anwendung unter Verwendung von Regelkartenverfahren zu überprüfen, wie in ISO 4259-4 für den Prädiktionsfehler, definiert als $(\hat{Y} - Y)$, beschrieben.

Die Bestätigung sollte mit der nachstehenden Differenzstatistik D oder anderen statistisch gleichwertigen Techniken durchgeführt werden. Bei einem Einzelwert D gelten die Beurteilungsergebnisse als validiert, wenn der absolute Wert kleiner oder gleich 3 ist. Bei einer Regelkarte wird erwartet, dass die D -Werte auf beiden Seiten von Null zufällig streuen.

Nachhaltige Werte von D auf der positiven oder negativen Seite von Null sollten eine Neubeurteilung auslösen.

$$D = \frac{(\bar{Y} - \hat{Y})}{\sqrt{SE_{\bar{Y}}^2 + SE_{\hat{Y}}^2}} \quad (42)$$

wobei $\hat{Y}$ die Prädiktion des Mittelwerts Y ($=Y$) ist, berechnet aus dem gemittelten Ergebnis X ($=X$):

$$\hat{Y} = b\bar{X} + a \quad (43)$$

Dabei ist

- Y der Mittelwert des Verfahrens Y für dasselbe Material;
- X der Mittelwert des Verfahrens X für dasselbe Material;
- a, b die Bias-Korrektur-Koeffizienten aus der Bias-Beurteilung, falls zutreffend;
- SE_Y der Standardfehler von Y , gegeben durch $SE_Y = s_{RY} / \sqrt{L_Y}$ mit $s_{RY} = R_Y / (t\sqrt{2})$ Veröffentlichte Standardabweichung der Wiederholbarkeit des Verfahrens Y;
- $SE_{\hat{Y}}$ der Standardfehler von $\hat{Y}$, $SE_{\hat{Y}}$, berechnet aus dem Bias-korrigierten X , gegeben durch $SE_{\hat{Y}} = b \cdot (s_{RX} / \sqrt{L_X})$ mit $s_{RX} = R_X / (t\sqrt{2})$ Veröffentlichte Standardabweichung der Wiederholbarkeit des Verfahrens X;
- R_X, R_Y die veröffentlichten Vergleichbarkeiten für die Verfahren X und Y;
- L_X, L_Y die Anzahl der nicht eliminierten Ergebnisse, die zur Berechnung des Mittelwerts für die Verfahren X und Y verwendet wurden, wobei das Protokoll vorsieht, dass die einzelnen Teilnehmer ein einziges Prüfergebnis mitteilen;
- t der Student-t-Wert mit einem Vertrauensniveau von 0,05 und den Freiheitsgraden aus der veröffentlichten Präzisionserklärung des Verfahrens Y.

Falls der Freiheitsgrad nicht bekannt ist, darf der Mindestwert von 30 verwendet werden.

Anhang A (informativ)

Berechnungsbeispiel unter Verwendung von Ringversuchsdaten

In diesem Anhang soll das Verfahren unter Verwendung eines Ringversuchsdatensatzes verdeutlicht werden, der Ergebnisse für die Bestimmung der Cetanzahl (CN) von Dieselmotoren nach den Prüfverfahren ISO 5165:2017[3] und EN 16906:2020[4] enthält. Dieser Datensatz ist eine Teilmenge eines größeren Datensatzes, der zur Erstellung einer Präzisionserklärung für beide Prüfverfahren verwendet wurde.

Innerhalb dieser Auswertung wird das Prüfverfahren ISO 5165 mit X und das Prüfverfahren EN 16906 mit Y bezeichnet.

A.1 Auswahl relevanter Eigenschaften

Die relevanten Eigenschaften der beiden Prüfverfahren werden berücksichtigt und zusammengefasst.

A.1.1 Prüfverfahren X

Die Norm ISO 5165:2017 legt die Bewertung von Dieselmotoren mittels einer willkürlichen Skala von Cetanzahlen (CN) unter Verwendung eines genormten Viertakt-Einzylinder-Dieselmotors mit variablem Verdichtungsverhältnis und indirekter Einspritzung fest. Die Anwendungs- und Präzisionsbereiche sind in Tabelle A.1 angegeben.

Tabelle A.1 — Anwendungsbereich der ISO 5165:2017

Üblicher Anwendungsbereich, CN	30 bis 65
Präzisionsbereich	52,4 bis 73,8
Veröffentlichte Vergleichbarkeit	$0,125 \cdot CN - 2,2$
Veröffentlichte Wiederholbarkeit	$0,01 \cdot CN + 0,42$

Beachten, dass der Ausdruck der Wiederholbarkeits- und der Vergleichbarkeitsfunktionen durch Regression fester Präzisionswerte bei bestimmten CN-Niveaus wie in Tabelle A.2 konstruiert wird.

Tabelle A.2 — Veröffentlichte Präzision für ISO 5165:2017

CN	r	R
40	0,8	2,8
44	0,9	3,3
48	0,9	3,8
52	0,9	4,3
56	1,0	4,8

Zusätzlich ist in der Präzisionserklärung folgendes vermerkt: „Die mittlere Standardabweichung für jedes CN-Niveau wurde mit 2,772 multipliziert, um die entsprechenden Werte für Wiederholbarkeit und Vergleichbarkeit zu erhalten.“ Das bedeutet, dass für die Berechnung der Standardabweichung der Vergleichbarkeit und Wiederholbarkeit nicht der Faktor 2,888, sondern 2,772 verwendet werden sollte.

A.1.2 Prüfverfahren Y

Die Norm EN 16906:2020 legt ein Prüfverfahren für die Bestimmung der Cetanzahl („CN“) von Dieselmotorkraftstoff im Bereich von CN 0 bis CN 100 (üblicherweise CN 40 bis CN 75) unter Verwendung eines genormten Viertakt-Einzylinder-Motors mit indirekter Einspritzung fest. Die Cetanzahl ist ein Maß für die Zündwilligkeit von Dieselmotorkraftstoffen in Selbstzündungsmotoren. Die relevanten Einzelheiten zum Anwendungsbereich sind in Tabelle A.3 aufgeführt.

Tabelle A.3 — Anwendungsbereich der EN 16906:2020

Üblicher Anwendungsbereich	0 bis 100 (üblicherweise 40 bis 75)
Präzisionsbereich	44 bis 66
Veröffentlichte Vergleichbarkeit	1,5
Veröffentlichte Wiederholbarkeit	0,64

A.2 Labordaten

Die Labordaten aller Proben werden gesammelt und alle Ausreißer eliminiert. Da der gesamte Datensatz bereits mit einem Verfahren nach ISO 4259-1 oder einem gleichwertigen Verfahren beurteilt wurde, um eine Präzisionserklärung zu erstellen, kann davon ausgegangen werden, dass keine Ausreißer im Datensatz vorliegen.

Für die einzelnen Proben wird eine Zusammenfassung mit statistischen Angaben, einschließlich Mittelwert, Standardabweichung und Anzahl der Labore, angegeben. Entsprechend der Berichtsform werden die Daten mit 1 Nachkommastelle angegeben.

- Entwurf -

Tabelle A.4 — Cetanzahlendaten für Verfahren X (ISO 5165:2017)

		Probe														
		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
Labor	L1	52,8	52,1	67,6	61,3	54,4	49,1	43,8	53,9	51,9	52,1	53,8	53,5	56,5	51,4	52,5
		52,6	52,8	68,6	62,2	54,9	49,4	43,5	53,3	52,1	51,7	53,2	54,0	56,5	51,2	52,1
	L2	52,0	52,3	65,9	60,8	53,7	47,9	43,4	56,5	53,3	52,2	54,8	51,6	55,4	50,7	52,9
		52,3	52,1	66,5	60,1	53,1	48,0	43,6	56,8	52,9	52,3	53,8	51,2	54,9	51,2	53,4
	L3	51,8	53,3	66,4	60,0	53,8	49,4	44,1	56,3	52,4	51,9	52,7	51,7	55,0	51,0	53,6
		52,7	52,4	65,7	60,6	54,2	49,7	44,2	55,9	51,9	51,3	52,7	52,0	54,9	50,9	53,0
	L4	52,3	50,5	67,0	59,7	55,0	50,6	42,7	57,9	53,0	52,7	52,2	51,3	56,9	51,3	54,4
		52,4	51,0	66,4	60,0	55,1	50,5	41,7	56,7	53,6	53,2	52,3	51,5	57,3	51,3	54,6
	L5	51,9	51,7	65,7	60,5	53,7	49,7	44,4	53,8	53,2	53,3	54,0	52,0	55,4	50,2	52,4
		51,6	51,9	64,7	61,1	54,1	49,6	44,9	53,9	52,8	53,4	53,7	52,3	55,0	49,4	52,3
	L6	53,0	52,6	66,9	59,5	54,1	48,6	42,5	54,2	52,8	53,5	52,3	52,4	55,5	50,7	52,8
		53,7	52,6	66,4	59,9	54,3	49,2	43,5	54,5	52,0	53,6	53,0	53,3	55,7	51,6	53,0
	L7	52,5	50,1	63,3	60,8	53,9	46,4	42,6	56,5	51,8	53,7	53,7	51,8	54,2	51,7	54,0
		52,3	51,3	63,4	61,3	54,6	48,3	42,5	56,3	51,8	53,5	54,1	52,1	53,8	51,2	53,5
	L8	51,7	51,6	66,6	60,9	53,8	48,2	43,9	55,2	52,5	52,9	53,8	51,8	55,1	52,6	52,4
		51,1	51,7	65,8	60,2	53,5	47,4	44,4	55,1	52,9	52,1	54,0	52,5	54,6	52,1	52,4
	L9	52,0	52,6	66,9	58,9	52,4	48,9	42,4	53,8	51,2	53,4	55,0	52,5	55,0	51,0	53,3
		51,9	52,1	67,7	58,0	51,2	47,8	42,9	54,9	51,5	53,8	54,5	53,2	55,9	51,6	53,8

Tabelle A.5 — Cetanzahl für Verfahren Y (EN 16906:2020)

		Probe														
		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
Labor	L1	51,6	50,8	65,3	60,0	53,7	48,0	44,6	53,8	52,7	52,6	53,8	51,6	55,2	51,0	53,0
		51,7	51,6	66,1	59,9	53,8	48,2	44,1	54,1	52,5	52,5	53,7	51,6	54,9	50,9	52,8
	L2	52,3	52,2	66,1	59,2	52,8	47,9	43,4	55,0	52,1	52,6	53,4	52,1	54,8	50,7	53,0
		52,2	51,9	66,3	59,0	53,3	47,4	43,8	55,2	52,2	52,9	53,3	52,4	54,8	50,8	52,8
	L3	52,1	51,5	66,9	59,9	52,8	47,7	43,3	54,5	52,1	52,8	53,7	51,2	55,4	51,8	53,0
		52,2	52,0	67,4	60,1	52,8	47,8	43,3	54,7	52,3	52,6	53,0	51,2	55,8	52,0	53,9
	L4	52,2	52,1	65,8	59,3	54,4	48,7	43,9	53,9	51,8	52,3	52,8	52,8	54,8	50,0	52,4
		52,0	52,0	65,3	59,6	54,8	48,8	44,0	54,2	51,3	53,0	52,8	52,6	55,0	49,7	52,3
	L5	51,8	52,4	65,5	60,0	53,2	48,0	43,6	55,0	51,1	52,7	52,9	52,6	55,1	50,8	52,2
		51,9	52,1	65,4	60,1	53,0	48,4	44,1	54,4	51,5	52,7	52,8	52,2	55,0	50,9	52,2
	L6	52,6	51,0	66,1	61,3	53,4	49,2	43,5	55,6	52,8	53,3	53,2	52,3	56,9	52,0	52,2
		52,4	50,6	65,8	61,1	53,2	48,3	43,4	55,5	52,9	53,4	53,0	52,3	56,6	52,2	52,3
	L7	51,9	52,9	63,6	59,6	53,8	47,3	44,1	56,2	52,2	52,4	51,9	51,9	56,1	51,4	52,3
		52,0	52,5	64,6	59,5	53,5	48,0	43,3	55,9	52,2	52,7	52,2	52,1	56,4	51,1	52,4
	L8	51,7	51,4	65,9	59,7	54,2	48,1	42,9	55,1	51,5	51,6	53,4	51,6	55,2	51,7	53,4
		52,0	52,0	66,4	60,2	54,7	48,0	42,2	54,9	51,9	51,9	53,4	52,2	55,4	51,8	53,6
	L9	51,6	50,5	65,6	59,3	53,6	49,1	42,8	54,2	52,5	52,1	54,3	51,6	54,6	50,5	52,9
		52,3	51,1	65,4	59,8	53,7	49,5	42,3	54,9	53,2	51,6	53,7	52,0	55,3	50,3	52,5

Zur Prüfung der Variationsbreite oder Auflösung der Daten wird die Anzahl der eindeutigen Werte innerhalb des Datensatzes gezählt und ausgewertet. Mit 117 (43 %) eindeutigen Werten von insgesamt 270 Werten für das Prüfverfahren ISO 5165 und 114 (42 %) eindeutigen Werten von insgesamt 270 Werten wird die Variation im Datensatz als ausreichend angesehen.

A.3 Berechnung der Mittelwerte über die Wiederholungen und Neuordnung

Es werden neue Tabellen mit Zellen, die den Mittelwert über die Wiederholungen aller Labor/Proben-Kombinationen mit einer zusätzlichen Nachkommastellen darstellen, erstellt. Um die Datenverarbeitung zu erleichtern, wird die Tabelle so gedreht, dass die Proben horizontal und die Labore vertikal angeordnet sind.

Tabelle A.6 — Cetanzahl Daten für Verfahren X (ISO 5165:2017)

		Labor									Mittelwert	Standardab- weichung	Anzahl Labore	Relative Standardabwei- chung
		L1	L2	L3	L4	L5	L6	L7	L8	L9	X_i	s_{Xi}	L_{Yi}	$s_{Xi} \%$
Probe	S1	52,70	52,15	52,25	52,35	51,75	53,35	52,40	51,40	51,95	52,26	0,561	9	1,1 %
	S2	52,45	52,20	52,85	50,75	51,80	52,60	50,70	51,65	52,35	51,93	0,776	9	1,5 %
	S3	68,10	66,20	66,05	66,70	65,20	66,65	63,35	66,20	67,30	66,19	1,342	9	2,0 %
	S4	61,75	60,45	60,30	59,85	60,80	59,70	61,05	60,55	58,45	60,32	0,934	9	1,5 %
	S5	54,65	53,40	54,00	55,05	53,90	54,20	54,25	53,65	51,80	53,88	0,924	9	1,7 %
	S6	49,25	47,95	49,55	50,55	49,65	48,90	47,35	47,80	48,35	48,82	1,036	9	2,1 %
	S7	43,65	43,50	44,15	42,20	44,65	43,00	42,55	44,15	42,65	43,39	0,840	9	1,9 %
	S8	53,60	56,65	56,10	57,30	53,85	54,35	56,40	55,15	54,35	55,31	1,346	9	2,4 %
	S9	52,00	53,10	52,15	53,30	53,00	52,40	51,80	52,70	51,35	52,42	0,655	9	1,2 %
	S10	51,90	52,25	51,60	52,95	53,35	53,55	53,60	52,50	53,60	52,81	0,775	9	1,5 %
	S11	53,50	54,30	52,70	52,25	53,85	52,65	53,90	53,90	54,75	53,53	0,834	9	1,6 %
	S12	53,75	51,40	51,85	51,40	52,15	52,85	51,95	52,15	52,85	52,26	0,765	9	1,5 %
	S13	56,50	55,15	54,95	57,10	55,20	55,60	54,00	54,85	55,45	55,42	0,916	9	1,7 %
	S14	51,30	50,95	50,95	51,30	49,80	51,15	51,45	52,35	51,30	51,17	0,661	9	1,3 %
	S15	52,30	53,15	53,30	54,50	52,35	52,90	53,75	52,40	53,55	53,13	0,737	9	1,4 %

- Entwurf -

Tabelle A.7 — Cetanzahl Daten für Verfahren Y (EN 16906:2020)

		Labor									Mittelwert	Standardabweichung	Anzahl Labore	Relative Standardabweichung
		L1	L2	L3	L4	L5	L6	L7	L8	L9				
Probe	S1	51,65	52,25	52,15	52,10	51,85	52,50	51,95	51,85	51,95	52,028	0,253	9	0,5 %
	S2	51,20	52,05	51,75	52,05	52,25	50,80	52,70	51,70	50,80	51,700	0,654	9	1,3 %
	S3	65,70	66,20	67,15	65,55	65,45	65,95	64,10	66,15	65,50	65,750	0,812	9	1,2 %
	S4	59,95	59,10	60,00	59,45	60,05	61,20	59,55	59,95	59,55	59,867	0,592	9	1,0 %
	S5	53,75	53,05	52,80	54,60	53,10	53,30	53,65	54,45	53,65	53,594	0,616	9	1,1 %
	S6	48,10	47,65	47,75	48,75	48,20	48,75	47,65	48,05	49,30	48,244	0,574	9	1,2 %
	S7	44,35	43,60	43,30	43,95	43,85	43,45	43,70	42,55	42,55	43,478	0,606	9	1,4 %
	S8	53,95	55,10	54,60	54,05	54,70	55,55	56,05	55,00	54,55	54,839	0,674	9	1,2 %
	S9	52,60	52,15	52,20	51,55	51,30	52,85	52,20	51,70	52,85	52,156	0,556	9	1,1 %
	S10	52,55	52,75	52,70	52,65	52,70	53,35	52,55	51,75	51,85	52,539	0,483	9	0,9 %
	S11	53,75	53,35	53,35	52,80	52,85	53,10	52,05	53,40	54,00	53,183	0,574	9	1,1 %
	S12	51,60	52,25	51,20	52,70	52,40	52,30	52,00	51,90	51,80	52,017	0,454	9	0,9 %
	S13	55,05	54,80	55,60	54,90	55,05	56,75	56,25	55,30	54,95	55,406	0,676	9	1,2 %
	S14	50,95	50,75	51,90	49,85	50,85	52,10	51,25	51,75	50,40	51,089	0,737	9	1,4 %
	S15	52,90	52,90	53,45	52,35	52,20	52,25	52,35	53,50	52,70	52,733	0,496	9	0,9 %

Aus qualitativer Sicht lässt sich die Datenintegrität einigermaßen beurteilen, wenn die relativen Standardabweichungen an dieser Stelle für beide Prüfverfahren betrachtet werden. Ausreißer sollten näher betrachtet werden.

A.4 Grundlegende Überprüfungen des Datensatzes durchführen.

A.4.1 Der Datensatz stammt von 15 Proben aus 9 Laboren und enthält Mittelwerte über alle wiederholten Ergebnisse. Damit sind die Anforderungen an Ringversuchsdaten, dass die Ergebnisse von mindestens 10 Proben aus 6 Laboren stammen, erfüllt.

A.4.2 Die Proben im Datensatz umfassen den CN-Bereich von 43,6 bis 66,2 für ISO 5165 und von 43,5 bis 65,8 für EN 16906. Diese Bereiche werden mit den veröffentlichten Anwendungs- und Präzisionsbereichen der Prüfverfahren verglichen.

Tabelle A.8 — Anwendungs- und Präzisionsbereiche der Prüfverfahren im Vergleich mit dem Datensatz

Cetanzahl	ISO 5165	EN 16906
Üblicher veröffentlichter Anwendungsbereich	30 bis 65	45 bis 63
Veröffentlichter Präzisionsbereich	52,4 bis 73,8	47 bis 61
Datensatzbereich	43,6 bis 66,2	43,5 bis 65,8

Der Vergleich dieser Bereiche zeigt, dass die Ergebnisse der Bewertung auf einen CN-Bereich von 52,4 bis 61 begrenzt sind.

A.5 Prüfung auf Einfluss der Proben

Die Einflüsse der Proben werden auf Grundlage der Mittelwerte der Verfahren berechnet. In diesem Fall wird der Einfluss nicht für die einzelnen Proben für beide Prüfverfahren bestimmt, sondern es wird der Mittelwert über die Prüfverfahren verwendet, um den Einfluss je Probe zu bestimmen.

Tabelle A.9 — Einfluss der einzelnen Proben, gemittelt über beide Verfahren

Probe	Mittelwert je Probe	$Z_i = \text{Ln}(\text{Mittelwert})$	Mittelwert über Z_i	SQ-Differenz	Summe der Quadrate	Einfluss
	$(X_i + Y_i)/2$	Z_i	Z_0	$(Z_i - Z_0)^2$	$\sum (Z_i - Z_0)^2$	h_i
S1	52,14	3,954	3,974	0,000	0,117	0,07
S2	51,81	3,948	3,974	0,001	0,117	0,07
S3	65,97	4,189	3,974	0,047	0,117	0,46
S4	60,09	4,096	3,974	0,015	0,117	0,19
S5	53,74	3,984	3,974	0,000	0,117	0,07
S6	48,53	3,882	3,974	0,008	0,117	0,14
S7	43,43	3,771	3,974	0,041	0,117	0,42
S8	55,07	4,009	3,974	0,001	0,117	0,08
S9	52,29	3,957	3,974	0,000	0,117	0,07
S10	52,68	3,964	3,974	0,000	0,117	0,07
S11	53,36	3,977	3,974	0,000	0,117	0,07
S12	52,14	3,954	3,974	0,000	0,117	0,07

Tabelle A.9 (fortgesetzt)

Probe	Mittelwert je Probe	$Z_i = \ln(\text{Mittelwert})$	Mittelwert über Z_i	SQ-Differenz	Summe der Quadrate	Einfluss
	$(X_i + Y_i)/2$	Z_i	Z_0	$(Z_i - Z_0)^2$	$\sum (Z_i - Z_0)^2$	h_i
S13	55,41	4,015	3,974	0,002	0,117	0,08
S14	51,13	3,934	3,974	0,002	0,117	0,08
S15	52,93	3,969	3,974	0,000	0,117	0,07

bestanden

Für die einzelnen Proben bleibt der berechnete Einflusswert h_i unter dem kritischen Wert von 0,5, so dass die Prüfung bestanden ist.

Die Probenverteilung wird unter Verwendung des AD-Tests und eines Teils der Übersichtstabelle auf Normalität geprüft.

Beachten, dass diese Prüfung für Daten, die aus einem Ringversuch stammen und bereits ein vollständiges Verfahren nach ISO 4259-1 oder einem gleichwertigen Verfahren durchlaufen haben, nicht erforderlich ist; aus Gründen der Einheitlichkeit mit der PTP-Datensatzauswertung in Anhang B ist es jedoch ratsam, diese Prüfung trotzdem durchzuführen.

A.6 Erstellung von Übersichtstabellen mit zusammenfassenden Statistiken je Probe

Um die Daten auf Einhaltung der Anforderungen zu prüfen, wird die Verteilung der Laborergebnisse je Probe untersucht, wobei für die einzelnen Proben mindestens 6 Labore ein Ergebnis geliefert haben müssen, und die Präzision der Daten wird mit den veröffentlichten Präzisionswerten verglichen. Beachten, dass für Daten aus einem Ringversuch nur geprüft werden muss, ob die Mindestanzahl der Labore ein Ergebnis geliefert hat.

Tabelle A.10 — Übersichtstabelle für Verfahren X (ISO 5165:2017)

Probe	Mittelwert	Standardabweichung	Labor Nr.	Veröffentlichte Präzision				Standardfehler	Einfluss	Verteilung	Prüfung	F-Test
S_i	X_i	s_{Xi}	L_{Xi}	R_{Xi}	r_{Xi}	s_{RXi}	s_{rXi}	SE_{Xi}	h_i	AD-Wert	$L \geq 6$	$s_{Xi} \leq s_{RXi}$
S1	52,256	0,561	9	= 0,125 · X-2,2	= 0,01 · X+0,42	1,563	0,340	0,515	0,070	0,229	ja	ja
S2	51,928	0,776	9			1,548	0,339	0,510	0,072	0,466	ja	ja
S3	66,194	1,342	9			2,191	0,390	0,726	0,464	0,439	ja	ja
S4	60,322	0,934	9			1,927	0,369	0,637	0,194	0,268	ja	ja
S5	53,878	0,924	9			1,636	0,346	0,539	0,068	0,512	ja	ja
S6	48,817	1,036	9			1,408	0,328	0,462	0,138	0,199	ja	ja
S7	43,389	0,840	9			1,163	0,308	0,379	0,416	0,253	ja	ja
S8	55,306	1,346	9			1,700	0,351	0,561	0,077	0,370	ja	ja
S9	52,422	0,655	9			1,570	0,341	0,517	0,069	0,194	ja	ja
S10	52,811	0,775	9			1,588	0,342	0,523	0,067	0,458	ja	ja
S11	53,533	0,834	9			1,620	0,345	0,534	0,067	0,378	ja	ja
S12	52,261	0,765	9			1,563	0,340	0,515	0,070	0,396	ja	ja
S13	55,422	0,916	9			1,706	0,351	0,563	0,081	0,380	ja	ja
S14	51,172	0,661	9			1,514	0,336	0,498	0,080	0,753	ja	ja
S15	53,133	0,737	9			1,602	0,343	0,528	0,067	0,290	ja	ja

bestanden

bestanden

bestanden

100 %
bestanden