



GEFÖRDERT VOM



Flexible und zuverlässige Konzepte für eine nachhaltige Wasserwiederverwendung in der Landwirtschaft (*FlexTreat*)

Validierungsleitfaden für die uneingeschränkte Bewässerung

Wolfgang Seis¹, Nicole Zacharias², Benedikt Aumeier³, Lia Freier², Michael Stapf¹, Ulf Miehe¹, Thomas Wintgens⁴

¹ Kompetenzzentrum Wasser Berlin gGmbH, Grunewaldstr. 61-62, 10825 Berlin

² Institut für Hygiene und öffentliche Gesundheit, Universitätsklinikum Bonn, 53127 Bonn, Deutschland

³ Lehrstuhl für Siedlungswasserwirtschaft, Technische Universität München, Am Coulombwall 3, 85748 Garching

⁴ Lehrstuhl für Siedlungswasserwirtschaft und Wassergütewirtschaft und Institut für Siedlungswasserwirtschaft

Inhalt

1	Einleitung und Zielstellung	1
2	Zusammenfassung der gesetzlichen Anforderungen	1
2.1	Routineüberwachung	2
2.2	Validierungsmonitoring	2
2.3	Leitlinien zur Anwendung der Verordnung 2020/741 über Mindestanforderungen an die Wasserwiederverwendung	4
3	Auswahl und Beschreibung der Validierungsparameter	7
3.1	<i>Escherichia coli</i>	7
3.1.1	Messverfahren:.....	8
3.1.2	Hinweise:	8
3.2	Bakteriophagen.....	8
3.2.1	Messverfahren:.....	8
3.2.2	Hinweise:	8
3.3	<i>Clostridium perfringens</i>	9
3.3.1	Messverfahren:.....	9
3.3.2	Hinweise:	9
3.4	Vorkommen mikrobieller Indikatoren im Rohabwasser	9
4	Versuchsplanung und Durchführung.....	12
4.1	Bestimmung des Validierungszeitraums und zu validierender Betriebsfenster	12
4.2	Betriebsstörungen	12
4.3	Ort der Probennahme	13
4.4	Art der Probennahme.....	14
4.5	Auswahl des Analysenverfahrens.....	17
4.6	Anreicherung großer Probenvolumina im Klärwerksablauf bei zu niedrigen Zulaufkonzentrationen	17
4.6.1	Allgemeine Beschreibung des Auswertungsansatzes	17
4.6.2	Bestimmung des zu filtrierenden Volumens	19
4.6.3	Anreicherungsverfahren.....	21
4.6.4	Anreicherungsversuche im Projekt FlexTreat	21
4.6.5	Weitere Anwendungsfälle	24
5	Datenauswertung	25
5.1	Log ₁₀ -Reduktion	25

5.1.1	Messwerte kleiner Bestimmungsgrenze	25
5.2	Zufallsvariablen, statistische Parameter, und Unsicherheitsintervalle.....	26
5.3	Überblick über Auswertungsalternativen	28
5.3.1	10. Perzentil vs. Erfolgsrate	28
5.3.2	Gepaarte vs. ungepaarte Auswertung.....	29
5.3.3	Parametrische vs. nicht-parametrische Ansätze	29
5.4	Berechnung von Punktschätzern und Unsicherheitsintervallen für die Erfolgsrate r.....	30
5.4.1	Konfidenzintervall nach Wilson.....	31
5.4.2	Bayes'sches Konfidenzintervall	31
5.4.3	Vergleich der Konfidenzintervalle nach Wilson und Bayes	32
5.5	Berechnung von Punktschätzern und Unsicherheitsintervallen für das 10. Perzentil	35
5.5.1	Punktschätzer des parametrischen 10. Perzentil für normalverteilte Log_{10} -Reduktionswerte 35	
5.5.2	Berechnung der einseitigen unteren Toleranzgrenze für normalerteilte Log_{10} - Reduktionswerte	36
5.5.3	Vergleich von Punktschätzern und Unsicherheitsintervallen für normalerteilte Log_{10} - Reduktionswerte	37
5.5.4	Ansätze nicht-normalverteilte Log_{10} -Reduktionswerte.....	38
5.6	Anwendungsbeispiel	42
5.6.1	Beschreibung des Datensatzes und zugrundeliegende Annahmen (Versuchsphase 1).....	43
5.6.2	Ergebnisse (Versuchsphase 1)	46
5.6.3	Anreicherungsversuche (Versuchsphase 2)	47
5.6.4	Ergebnisse (Versuchsphase 2)	48
5.7	Zusammenfassung, Diskussion und Handlungsempfehlungen	50
5.7.1	Zielparameter	50
5.7.2	Quantifizierung der statistischen (Un-)Sicherheit.....	50
5.7.3	Erforderliche Anzahl an Stichproben.....	51
5.7.4	Ungepaarte Auswertung	53
6	Schlussfolgerung und Handlungsempfehlungen.....	55
7	Literatur.....	58
Anhang	59
I.	Vorversuch zur Bestimmung des Korrekturfaktors „repräsentatives Ablaufvolumen V_r “	59
II.	Korrektur k Werte Toleranzintervall	61

Abbildungen

Abbildung 1: Zweiphasiger Validierungsprozess eines angenommenen Genehmigungsverfahrens einer Anlage zur landwirtschaftlichen Wasserwiederverwendung.(DWA, 2024)	6
Abbildung 2: Präzision verschiedener Nachweisverfahren und Einordnung relevanter Validierungsparameter nach (WHO (2016), Anhang C, Seite 139)	7
Abbildung 3: Vergleich zwischen somatischen Coliphagen, F-spezifischen Coliphagen und Coliphagen (gesamt) im Zulauf der vier FlexTreat Standorte und des Standorts Schweinfurt aus dem Schwesterprojekt Nutzwasser	10
Abbildung 4: Zulaufkonzentration verschiedener Indikatoren in zehn Klärwerken in Deutschland. Blaue Punkte: Messwerte, Unsicherheitsintervalle (schwarz) entsprechen dem unteren und oberen einseitigen Toleranzintervall aus den verfügbaren Daten. Rote Linien entsprechen dem jeweiligen Validierungsziel. (Quelle: IHPH, KWB, TZW)	11
Abbildung 5: Bild eines gekühlten automatischen Probennehmers am Zulauf der Kläranlage Steinhof in Braunschweig.	14
Abbildung 6: Stabilitätsversuche für E. coli (IHPH).....	16
Abbildung 7: Illustrative Darstellungen eines Anreicherungsverfahrens über Hohlfasermembranen.....	21
Abbildung 8: Beispielhafte Darstellung einer Dialysemembran zur Aufkonzentrierung von mikrobiologischen Parametern.	22
Abbildung 9: Darstellung der Korrektur der Zulaufkonzentration über eine Erhöhung des Bezugsvolumens bzw. repräsentativen Ablaufvolumens.	24
Abbildung 10: Anzahl benötigter Erfolgsmessungen in Abhängigkeit der Anzahl an Misserfolgen (Bayes'scher Beta-Binomialansatz)	33
Abbildung 11: Vergleich von Punktschätzern und Unsicherheitsintervallen von normalverteilten Werten.	38
Abbildung 12: Illustration des MCMC Monte Carlo Verfahrens zur Bestimmung Bayes'scher Toleranzintervalle.....	40
Abbildung 13 Vergleich verschiedener Auswertungsansätze für Daten aus dem Projekt FlexTreat. Die Unsicherheitsintervalle beschreiben das 10. und 90. Perzentil (schwarzes Intervall), bzw. die einseitigen unteren und oberen Toleranzgrenzen mit $\alpha = 0,05$. Zahlenwerte X/Y illustrieren den Binomialansatz mit X = Anzahl der Erfolge und Y = Gesamtanzahl. Werte die aufgrund von zu niedrigen Zulaufkonzentrationen mit „> WERT“ unterhalb des Zielwerts liegend wurden als Misserfolg gewertet (relevant für somatische Coliphagen – som. CP).....	45
Abbildung 14: Simulierte Ablaufverteilung.auf Basis realer Messwerte, sowie unter der Annahme von ausschließlichen Negativbefunden im Ablauf der Anlage.....	49
Abbildung 15: Darstellung des Punktschätzers des 10. Perzentils auf Basis der empirischen Methode nach Badegewässerrichtlinie auf Basis nicht-angereicherter, gepaarter Werte für den Betriebszustand Δ SAK = 47 %, und den berechneten Konfidenzintervallen für das 10. Perzentil auf Basis angereicherter Proben, wobei die untere Grenze des Konfidenzintervalls der unteren Toleranzgrenze entspricht.	49
Abbildung 16: Zusammenhang zwischen mittlerer \log_{10} -Reduktion, Streuung der \log_{10} -Reduktion und der Anzahl benötigter Stichproben.	52

Tabellen

Tabelle 1: Überwachung zur Validierung bei aufbereitetem Wasser für die landwirtschaftliche Bewässerung nach 2020/741	3
Tabelle 2: Empfohlene und akzeptable maximale Lagerungsdauern für verschiedene Indikatororganismen nach DIN ISO 19458:2006.....	15
Tabelle 3: Abschätzung des benötigten, zu filtrierenden Volumens für Coliphagen auf Basis der Messwerte aus Abbildung 4	20
Tabelle 4: Eingangsgrößen für die Abschätzung des zu filtrierenden Volumens.....	22
Tabelle 5: Zusammenfassung der ermittelten Größen, Messwerte und Korrekturfaktoren.....	23
Tabelle 6: Vergleich \log_{10} -Reduktion gegenüber Reduktion in %.....	25
Tabelle 7: Überblick über potenzielle Zielparameter und Validierungsansätze	29
Tabelle 8: Vergleich zwischen Intervallgrenzen der einseitigen Wilson und Bayes'schen Konfidenzintervalle. Grüne Schattierung: untere Konfidenzgrenze > Zielwert, Orange Schattierung: untere Konfidenzgrenze < Zielwert.....	32
Tabelle 9: Mögliche Kombinationen aus Zu- und Ablaufwerten, sowie deren Effekt auf die resultierende Verteilung der \log_{10} -Reduktionswerte	39
Tabelle 10: Zusammenfassung der wesentlichen Schritte des MCMC Monte-Carlo Verfahren.....	41
Tabelle 11: Anzahl der benötigten Stichproben, bei vorgegebener statistischer Sicherheit, damit kleinster Beobachtungswert einer Schätzung des 10. Perzentils entspricht.....	42
Tabelle 12. Vergleich verschiedener Auswertungsalternativen.....	54

1 Einleitung und Zielstellung

Mit der Europäischen Verordnung 2020/741 existiert die erste rechtlich bindende Festlegung von Mindestanforderungen im Bereich der Wasserwiederverwendung auf europäischer Ebene. Ein wesentlicher Aspekt ist die Sicherstellung der hygienischen Wasserqualität. Ein Baustein, um diese sicherzustellen, ist es, die Reinigungsleistung von Wasseraufbereitungsanlagen hinsichtlich bakterieller, parasitärer und viraler Krankheitserreger und Indikatororganismen zu validieren. Die Verordnung 2020/741 gibt hierzu den Rahmen in Form von zu erreichenden Mindestanforderungen und Validierungszielen vor. Detailliertere und verallgemeinerbare Vorgaben über die Art und Weise wie eine solche Validierung durchgeführt und ausgewertet werden soll, sind jedoch bisher nicht Bestandteil der Verordnung. Letztere sind jedoch maßgeblich für eine erfolgreiche Umsetzung und somit für die Implementierung der Verordnung 2020/741 in der Praxis.

Als weiterer wichtiger Aspekt ist die Umsetzung der Verordnung 2020/741 auf nationaler Ebene zu nennen. Als europäische Verordnung entwickelt 2020/741 zwar direkte Gesetzeswirkung, jedoch haben die Mitgliedstaaten die Möglichkeit, auf nationaler Ebene weitergehende Vorgaben zu machen. Dies gilt auch für die geforderte Validierung. Im deutschen Kontext steht hierzu vor allem die Zusatzanforderung zur Diskussion, dass selbst bei niedrigen Konzentrationen von mikrobiellen Indikatororganismen im Zulauf einer Kläranlage, die Reinigungsleistung in jedem Fall quantitativ validiert werden muss.

Vor diesem Hintergrund sind es die Ziele dieses Leitfadens:

- Existierende gesetzliche Regelungen zusammenfassen und zu analysieren
- Auf Basis der durchgeführten Analyse Regelungslücken und Fragestellungen zu identifizieren, die für die praktische Validierung maßgeblich sind.
- Lösungsansätze für die identifizierten Fragestellungen aufzuzeigen und zu vergleichen
- Praxisnahe Handlungsempfehlungen für die Durchführung der Validierung auf Basis der durchgeführten Analyse abzuleiten

Die erarbeiteten Empfehlungen berücksichtigen dabei explizit auch die im deutschen Kontext diskutierten Zusatzanforderungen. Bei den Handlungsempfehlungen handelt es sich um Einschätzungen des Projekts FlexTreat, in Abstimmung mit seinem Schwesterprojekt Nutzwasser. Sie sollen sowohl Betreibern als auch Entscheidungsträgern dabei unterstützen, die Anforderungen der europäischen Verordnung zur landwirtschaftlichen Wasserwiederverwendung bzgl. der Prozessvalidierung umzusetzen.

2 Zusammenfassung der gesetzlichen Anforderungen

Am 25. Mai 2020 verabschiedete das Europäische Parlament und der Europäische Rat die Verordnung 2020/741 über die Mindestanforderungen an die Wasserwiederverwendung. Ziel der Verordnung ist es, durch Mindestanforderung an die Wasserqualität, die Überwachung sowie Bestimmungen über das Risikomanagement, ein hohes Schutzniveau für die Gesundheit von Mensch, Tier und Umwelt zu gewährleisten. Dadurch soll garantiert werden, dass das aufbereitete Wasser, d.h. kommunales Abwasser, das gemäß den Anforderungen der Richtlinie 91/271/EWG behandelt und in einer Aufbereitungseinrichtung gemäß Anhang I Abschnitt 2 der EU-Verordnung 2020/741 weiterbehandelt wurde, für die landwirtschaftliche Bewässerung sicher ist und ein einheitlicher Mindeststandard in der EU erreicht wird. Die Regelungen aus Verordnung 2020/741 traten zum 26. Juni 2023 in Kraft und erlangten

somit aufgrund der Form des Rechtsaktes den Status des unmittelbar geltenden Rechts in den Mitgliedsstaaten, daher wird im vorliegenden Dokument von „gesetzlichen Anforderungen“ gesprochen. Um die mikrobiologische Sicherheit von landwirtschaftlichen Produkten zu gewährleisten, legt Verordnung 2020/741 Anforderungen auf verschiedenen Ebenen fest:

1. Eingrenzung der anbaubaren landwirtschaftlichen Produkte in Abhängigkeit der Wasserqualität (Güteklassen A-D).
2. Definition von Mindestanforderungen an die Aufbereitungstechnik, die in Abhängigkeit der Wasserqualitätsklasse implementiert sein muss.
3. Sicherstellung der Bewässerungswasserqualität und durch eine regelmäßige Routineüberwachung (siehe Kapitel 2.1).
4. Validierung der Reinigungsleistung durch ein zeitlich befristetes, intensives Validierungsmonitoring (siehe Kapitel 2.2)

2.1 Routineüberwachung

Die Anforderungen an die Routineüberwachung sind in Abschnitt 2a und 2b des Anhangs der Verordnung EC 2020/740 geregelt. Als Überwachungsparameter sind darin die Parameter *Escherichia coli* (*E. coli*), der biologische Sauerstoffbedarf (BSB₅), abfiltrierbare Stoffe (AFS) sowie die Trübung aufgeführt. Die ersten drei Überwachungsparameter sind wöchentlich zu bestimmen, während die Trübung kontinuierlich zu messen ist. Die Qualitätsanforderungen sind in Abhängigkeit der Güteklasse in Tabelle 2, Abschnitt 2 festgelegt. Die Anforderungen an die tolerierbaren Ablaufkonzentrationen für den Parameter *E. coli* von <10 KBE/100mL sind in mindestens 90 % der Proben die Überwachungswerte einzuhalten. Für die übrigen 10 % darf die Überschreitungshöhe für *E. coli* und *Legionella* spp. maximal eine logarithmische Einheit betragen. Grenzwerte für intestinale Nematoden sind in 100 % der Proben einzuhalten. Die Parameter *Legionella* spp. und intestinale Nematoden sind mit dem Vermerk „falls zutreffend“ versehen, da diese Parameter nur bei einem Risiko der Aerosolbildung (*Legionella* spp.), beziehungsweise der Bewässerung von Weideflächen oder Futterpflanzen (Nematoden) zu bestimmen sind. Für Deutschland werden weitere Parameter für die Routineüberwachung festgelegt, die u. a. im technischen Regelwerk DWA-M 1200 zusammengefasst werden. Eine Umsetzung in die nationale Wassergesetzgebung ist ebenfalls geplant, jedoch liegt diese zum Zeitpunkt der Erstellung dieses Leitfadens noch nicht vor.

2.2 Validierungsmonitoring

Das Validierungsmonitoring dient dem Zweck die Effizienz der Entfernungsleistung einer Aufbereitungseinrichtung hinsichtlich mikrobiologischer Parameter sicherzustellen und zu validieren. Unter einer Aufbereitungseinrichtung wird die Kombination aus konventioneller und weitergehender Abwasserbehandlung verstanden. Die Überwachung zur Validierung ist nicht notwendig, falls eine Aufbereitungseinrichtung bereits vor dem 25. Juni 2020 in Betrieb war und die Anforderungen der Routineüberwachung einhält. Hier existiert ein Bestandsschutz. Die Überwachung zur Validierung ist jedoch notwendig, falls es durch Modernisierungen, Erweiterungen, oder anderen Gründen, wie beispielsweise einer Vergrößerung des Einzugsgebietes zu wesentlichen Änderungen des existierenden Aufbereitungsprozesses kommt. Laut Verordnung 2020/741 ist das Validierungsmonitoring nur für Bewässerungswasser der Güteklasse A notwendig. Im deutschen Kontext wird jedoch diskutiert, auch für weitere Güteklassen ein Validierungsmonitoring durchzuführen, wenn damit beispielsweise Pflanzen für

die Nahrungsmittelproduktion bewässert werden sollen. Die Zielvorgaben der Überwachung zur Validierung sind in Tabelle 1 dargestellt.

Tabelle 1: Überwachung zur Validierung bei aufbereitetem Wasser für die landwirtschaftliche Bewässerung nach 2020/741

Indikator-Mikroorganismen*	Leistungsziele für die Behandlungskette (log ₁₀ -Reduktion)
<i>E. coli</i>	> 5,0
Coliphagen insgesamt/f-spezifische Coliphagen/somatische Coliphagen/Coliphagen**	> 6,0
<i>Clostridium perfringens</i> -Sporen / sporenbildende sulfatreduzierende Bakterien ***	> 4,0 (<i>Clostridium perfringens</i> -Sporen) > 5,0 (sporenbildende sulfatreduzierende Bakterien)
<p>(*) Anstelle der vorgeschlagenen Indikator-Mikroorganismen können für die Überwachung zur Validierung auch die Referenzpathogene <i>Campylobacter</i>, Rotavirus und <i>Cryptosporidium</i> herangezogen werden. In diesem Fall gelten die folgenden log₁₀-Reduktionsziele: <i>Campylobacter</i> (≥ 5,0), Rotavirus (≥ 6,0) und <i>Cryptosporidium</i> (≥ 5,0).</p> <p>(**) „Coliphagen insgesamt“ wurde als der am besten geeignete Virenindikator ausgewählt. Wenn jedoch die Analyse der Coliphagen insgesamt nicht möglich ist, wird mindestens ein Coliphagentyp (f-spezifische Coliphagen oder somatische Coliphagen) analysiert.</p> <p>(***) <i>Clostridium perfringens</i>-Sporen werden als der am besten geeignete Indikator für Protozoen ausgewählt. Sporenbildende sulfatreduzierende Bakterien sind jedoch eine Alternative, wenn die Konzentration von <i>Clostridium perfringens</i>-Sporen nicht ausreicht, um die erforderliche log₁₀-Reduktion zu validieren.</p>	

Tabelle 1 zeigt, dass Verordnung 2020/741 für die Überwachung zur Validierung Aufbereitungsziele für Indikatororganismen aus den Kategorien Bakterien, Viren und sporenbildende Bakterien, bzw. deren Sporen definiert. Letztere dienen als Indikatororganismen für Entfernung von Dauerstadien (Zysten, Oozysten) parasitärer Eukaryoten, v.a. *Giardia* spp. und *Cryptosporidium* spp. Darüber hinaus sind die Anforderungen mit verschiedenen Hinweisen versehen (gekennzeichnet mit *), die entweder auf Präferenzen hinsichtlich der Indikatorauswahl oder auf die Nutzung alternativer Parameter, v. a. die direkte Bestimmung pathogener Krankheitserreger, hinweisen.

Bezüglich der Art und Weise wie eine Validierung durchzuführen ist, unterscheidet die Verordnung zwischen drei Alternativen:

1. Analytische Kontrolle, also eine Validierung auf Basis lokaler Messwerte.
2. Addition von log₁₀-Reduktionswerten, die einzelnen Behandlungsschritten auf Grundlage wissenschaftlicher Erkenntnisse für etablierte Standardprozesse zuerkannt werden.
3. Bestimmung der log₁₀-Reduktion unter kontrollierten Testbedingungen im Labor.

Des Weiteren weist die Verordnung 2020/741 darauf hin, dass mindestens 90 % der Validierungsproben die in Tabelle 1 definierten Aufbereitungsziele an der Stelle der Einhaltung, erreichen oder übersteigen müssen. Als obere Systemgrenze wird im folgenden Bericht der Rohzulauf des Klärwerks, und als untere Systemgrenze der Ablauf der weitergehenden Abwasserbehandlung verstanden. Liegt ein Indikator nicht

in ausreichend hoher Konzentration im Zulauf der Anlage vor, kann laut EC 2020/741 ein Negativbefund am Ende der Aufbereitungskette als ein Einhalten der Anforderungen gewertet werden. Im Rahmen der Diskussion über notwendige Anpassungen der Verordnung in Deutschland wird dieser Ansatz kritisch gesehen, weswegen im vorliegenden Leitfaden stets das Szenario mitbetrachtet wird, dass die Regelungen der Verordnung verschärft angewendet werden und eine quantitative Validierung der Einhaltung der Zielwerte auch bei zu niedrigen Zulaufkonzentrationen, zwingend erforderlich ist.

Anforderungen an die Art der Probenahme oder Ansätze für die statistische Auswertung werden in der Verordnung selbst nicht weiter definiert. Dies stellt eine wesentliche Regelungslücke der existierenden Verordnung dar, da die Auswahl des Auswertungsverfahrens das erzielte Ergebnis stark beeinflussen kann. Auswertungsalternativen sowie Empfehlungen werden in Kapitel 5 detailliert dargestellt.

2.3 Leitlinien zur Anwendung der Verordnung 2020/741 über Mindestanforderungen an die Wasserwiederverwendung

Als Ergänzung zur Verordnung 2020/741 veröffentlichte die Europäischen Kommission am 05.08.2022 Leitlinien zur Anwendung der Verordnung 2020/741 über Mindestanforderungen an die Wasserwiederverwendung (2022/C 298/01). Innerhalb der Leitlinien finden sich ergänzende Hinweise zur Überwachung zur Validierung. Diese sind in Kapitel 3.3 der Leitlinien zusammengefasst, und umfassen die Bereiche:

1. Allgemeine Grundsätze
2. Protokollierung
3. Beispiele

Als zentrale Zusatzinformation wird hervorgehoben, dass unabhängig von vorhandener Vorinformation „die in Tabelle 4 in Anhang I der Verordnung vorgeschriebene Analyse für die Überwachung zur Validierung während der Anlaufphase am Ein- und Auslass durchzuführen [ist], um nachzuweisen, dass die mikrobielle \log_{10} -Reduktion erreicht wird“ (2022/C 298/01, S:28, Abs.:3).

Dies bedeutet, dass eine Validierung der \log_{10} -Reduktion nicht ausschließlich auf Basis von Vorversuchen und / oder existierender Information etablierter Reinigungsprozesse stattfinden kann, sondern **in jedem Fall** durch Vor-Ort-Untersuchung während der Inbetriebnahme und Anlaufphase analytisch validiert werden muss. Eine alleinige Validierung, die nur auf Basis von Vorversuchen und Informationen zu etablierten Reinigungsprozessen basiert, ist somit ausgeschlossen.

Vorinformationen zur Reinigungsleistung etablierter Reinigungsprozesse und Vorversuchen sind folglich vor allem für Vor- und Auslegungsstudien, sowie behördliche Genehmigungsprozesse für den Bau einer Anlage als relevant zu betrachten, was als erste Phase der Validierung betrachtet werden kann. Die analytische Prozessvalidierung, die bei Inbetriebnahme der Anlage erforderlich ist, nimmt somit eine hervorgehobene Stellung ein, da sie in jedem Fall durchzuführen ist. Sie stellt die zweite Validierungsphase dar. Abbildung 1 stellt die verschiedenen Validierungsvarianten im Rahmen eines verallgemeinerten Genehmigungsablaufs dar. Der vorliegende Leitfaden befasst sich ausschließlich mit der Planung, Durchführung und Auswertung des analytischen Validierungsmonitorings (Phase 2).

Neben der Definition der herausgehobenen Stellung des Validierungsmonitorings, beinhalten die Leitlinien zur Validierung auch Kommentare zur Datenqualität. Demzufolge ist es „für die mikrobielle Überwachung

[...] wichtig, eine Reihe statistisch gültiger Proben zu analysieren, sodass mindestens drei Proben an jedem Stichprobenpunkt die Berechnung von Mittelwerten und Standardabweichungen ermöglichen“ (2022/C 298/01, S:28, Abs.:6, Anstrich 2). Darüber hinaus sollte die Standardabweichung kleiner sein als eine logarithmische Einheit.

Eine genauere Erläuterung darüber, wie mit Hilfe dreier Proben und den daraus berechneten Mittelwerten und Standardabweichungen nachgewiesen werden soll, dass 90 % der Validierungsproben die Zielwerte einhalten, fehlt. Aufgrund fehlender Auswertungsansätze und der sehr geringen minimalen Stichprobenanzahl von $N \geq 3$, sorgen die Leitlinien für wenig Klarheit hinsichtlich der Versuchsdurchführung und Datenauswertung. Aufgrund dieser fehlenden Klarheit wird in diesem Bericht nicht gemutmaßt, wie die genannten Kommentare inhaltlich interpretieren werden könnten. Stattdessen soll durch eine klare und verständliche Darlegung von Auswertungsansätzen eine eindeutige, konforme und inhaltlich nachvollziehbare Alternative erarbeitet werden.

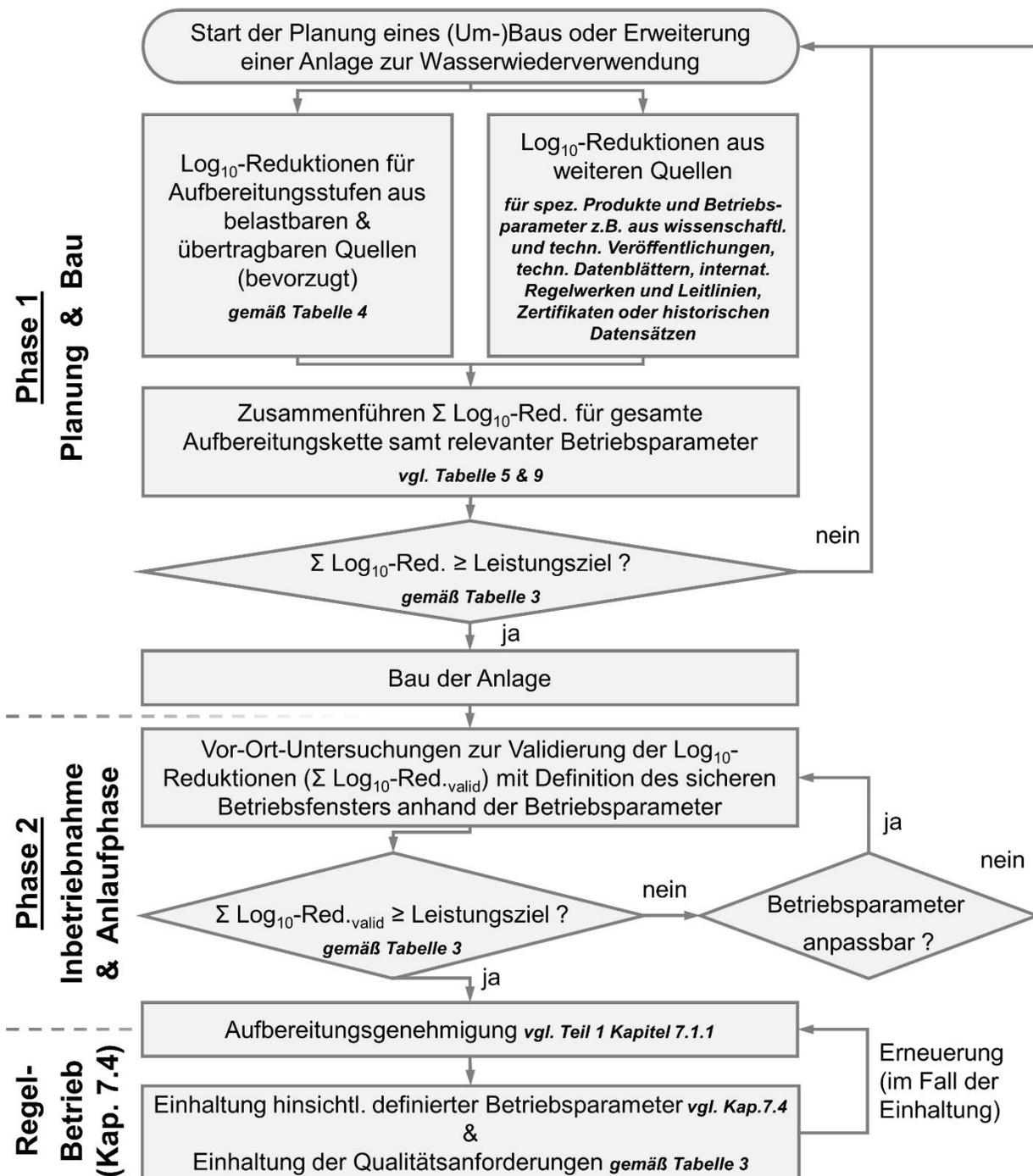


Abbildung 1: Zweiphasiger Validierungsprozess eines angenommenen Genehmigungsverfahrens einer Anlage zur landwirtschaftlichen Wasserrückverwendung. (DWA, 2024)

3 Auswahl und Beschreibung der Validierungsparameter

Die Europäische Verordnung zur Wasserwiederverwendung 2020/741 lässt eine Validierung grundsätzlich über Messungen von Indikatororganismen oder geeigneten Referenzpathogenen zu. Für die Durchführung der Validierung wird die Nutzung von Indikatororganismen empfohlen. Neben den geringeren Kosten, und der leichteren Handhabung, spricht vor allem das Vorhandensein standardisierter Messverfahren für die Nutzung von Indikatororganismen. WHO (2016) weist beispielsweise in Anhang C, Seite 139 auch auf die höhere relative Präzision von standardisierten Methoden im Vergleich zu weniger oder nicht-standardisierten Methoden zum Nachweis pathogener Mikroorganismen hin (Abbildung 2). Aus den genannten Gründen erscheint die Nutzung etablierter Indikatororganismen und Nachweismethoden sinnvoll.

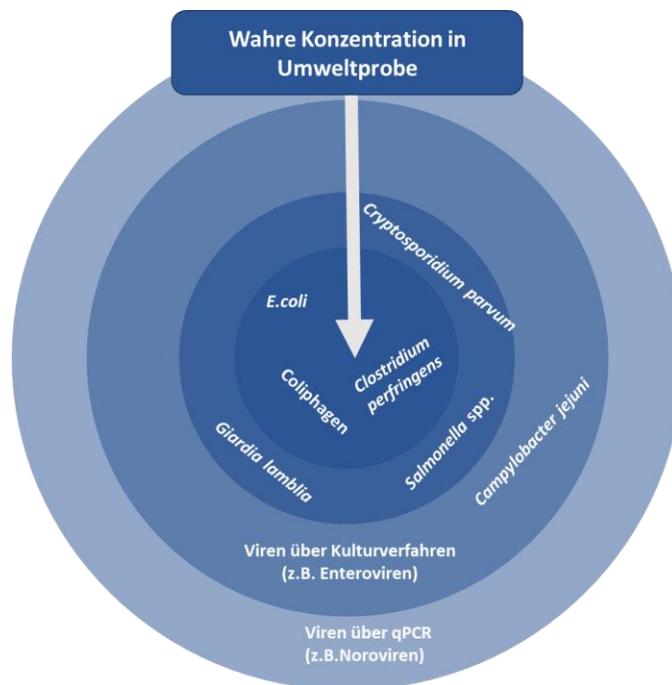


Abbildung 2: Präzision verschiedener Nachweisverfahren und Einordnung relevanter Validierungsparameter nach (WHO (2016), Anhang C, Seite 139)

3.1 *Escherichia coli*

Escherichia coli (*E. coli*) ist ein gramnegatives, stäbchenförmiges Bakterium. Die typische Größe liegt bei etwa 1-2 μm in der Länge und 0,25-1 μm im Durchmesser. *E. coli* ist ein fakultativ anaerobes Bakterium, was bedeutet, dass es sowohl in Gegenwart als auch in Abwesenheit von Sauerstoff wachsen kann. Als Indikatororganismus weist *E. coli* auf eine kürzlich erfolgte fäkale Kontamination hin, da *E. coli* zwar in großer Zahl im menschlichen und tierischen Darm vorkommt, jedoch außerhalb des Körpers in der Umwelt nicht dauerhaft überlebt. Im Rahmen des Validierungsmonitorings dient *E. coli* als Indikator für die Entfernung bakterieller Pathogene.

3.1.1 Messverfahren:

Für den Nachweis von *E. coli* existieren drei genormte Messverfahren

- ISO 9308-1 (Membranfiltrationsverfahren, für Gewässer mit wenig Begleitflora <100 KBE/Untersuchungsvolumen)
- ISO 9308-2 (Colilert Quanti-Tray, für alle Wässer geeignet, unabhängig der Begleitflora)
- ISO 9308-3: das Miniaturverfahren ist grundsätzlich für Abwasseranalysen geeignet, jedoch weist es eine untere Nachweisgrenze von < 15 MPN/100mL auf.

Je nach Messverfahren werden Messwerte entweder als Kolonie-bildende Einheiten (KBE/100mL) oder most-probable Number (MPN/100mL) angegeben.

3.1.2 Hinweise:

Für die Validierung der Log₁₀-Reduktion sind im Ablauf einer Aufbereitungsanlage Verfahren mit einer unteren Bestimmungsgrenze von < 1/100mL zu bevorzugen. Diese kann sowohl über Verfahren nach ISO 9308-1 oder 9308-2 erreicht werden. Eine Messung basierend auf ISO 9308-3 erscheint weniger geeignet, da sie bei einem Unterschreiten der Nachweisgrenze durch Anwendung eines zusätzlichen Verfahrens zur Absicherung der <1/100mL ergänzt werden müsste. Für den Zulauf einer Aufbereitungsanlage besteht diese Einschränkungen nicht, da von ausreichend hohen Konzentrationen ausgegangen werden kann.

3.2 Bakteriophagen

Bei Bakteriophagen handelt es sich um Viren, die Bakterien infizieren. Bakteriophagen existieren für alle bekannten Bakterienstämme. Für die Nutzung als Fäkalindikator dienen zumeist Coliphagen, die coliforme Bakterien (inkl. *E. coli*) infizieren. Dabei wird in den Regelwerken und gültigen Normen zwischen somatischen und F-spezifischen Coliphagen unterschieden. Somatische Coliphagen sind DNA-Viren während es sich bei F-spezifischen Coliphagen hauptsächlich um RNA-Viren handelt. Im Rahmen des Validierungsmonitorings dienen Coliphagen als Indikator für die Entfernung viraler Pathogene.

3.2.1 Messverfahren:

Für den Nachweis von Bakteriophagen existieren zwei genormte Messverfahren:

- DIN ISO 10705-1 2001 (F-spezifische RNA-Bakteriophagen)
- DIN ISO 10705-2 (somatische Coliphagen)

Messwerte von Coliphagen werden als Plaque-bildende Einheiten (PBE/100 mL) angegeben.

3.2.2 Hinweise:

Für die Reduktion viraler Indikatoren wird Verordnung 2020/741 folgend, die Analyse von Gesamtcoliphagen (Summe aus somatischen Coliphagen und F-spezifischen Coliphagen) am geeignetsten angesehen, da die Reduktionsleistung je nach technischem Verfahren spezifisch zwischen somatischen (DNA Viren) und F-spezifischen Coliphagen (RNA Viren) variieren kann. Neben der Empfehlung beide Parameter zu bestimmen, gibt Verordnung 2020/741 auch die Möglichkeit die Validierung mit nur einem Parameter durchzuführen.

3.3 *Clostridium perfringens*

Clostridium perfringens (*C. perfringens*) ist ein grampositives, anaerobes Bakterium aus der Gattung *Clostridium* und kommt im Darm sowohl in Form von vegetativen Zellen als auch in Form von Sporen, also Dauerstadien, vor. Im Gegensatz zu vegetativen Zellen sind die Sporen resistent gegenüber diversen Umwelteinflüssen wie UV-Strahlung, Hitze, oder Chemikalien und können monatelang im Wasser überleben. Die Sporen von *C. perfringens* werden im Rahmen vieler Regularien und Normen, sowie des Validierungsmonitoring der Verordnung 2020/741 als Indikator für die Entfernung von Zysten und Oozysten parasitärer Protozoen (v.a. *Giardia*, *Cryptosporidium*) verwendet.

3.3.1 Messverfahren:

Der Nachweis der Sporen von *C. perfringens* erfolgt mittels Kultivierung nach DIN ISO 141189:2016 über ein Membranfiltrationsverfahren. Die Methode hat eine unter Bestimmungsgrenze von $<1/\text{Untersuchungsvolumen}$ und wird als KBE/100mL angegeben.

3.3.2 Hinweise:

Mit dem oben beschriebenen Membranfiltrationsverfahren, können sowohl *C. perfringens* gesamt (vegetativen Bakterien und Sporen) als auch nur die *C. perfringens* Sporen analysiert werden. Für die Analyse von Sporen wird die entsprechende Wasserprobe vorab bei $(60 \pm 2)^\circ\text{C}$ pasteurisiert, um die vegetativen Bakterien zu inaktivieren. In der Praxis wird häufig nur von *C. perfringens* gesprochen, sodass es leicht zu Verwechslungen kommen kann. Bei der Beauftragung der Analytik sollten daher explizit auf die Analyse von *C. perfringens* Sporen hingewiesen und diese eingefordert werden.

3.4 Vorkommen mikrobieller Indikatoren im Rohabwasser

Laut Verordnung 2020/741 sollen mind. 90 % der berechneten Log_{10} -Reduktionswerte die indikatorspezifischen Zielwerte einhalten. Damit dieses Ziel mit quantifizierbaren Messungen hinterlegt werden kann, ist es unter Annahme einer unteren Bestimmungsgrenze von $<1/\text{Volumeneinheit}$ notwendig, dass 90 % der logarithmierten Zulaufwerte größer sind als der entsprechende Zielwert. Zulaufwerte von Coliphagen der vier FlexTreat-Standorte und des Nutzwasser-Standorts Schweinfurt sind in Abbildung 3 dargestellt. Die Abbildung zeigt, dass die Messwerte F-spezifischer Coliphagen auf einem ähnlichen Niveau wie somatische Coliphagen anzusiedeln sind. Messwerte aller Validierungsindikatoren aus zehn verschiedenen Klärwerken in Deutschland sind in Abbildung 4 dargestellt. Für die meisten Klärwerke liegen ausschließlich Werte für somatische Coliphagen vor. Da, wie gezeigt, F-spezifische Coliphagen auf einem ähnlichen Niveau liegen wie somatische Coliphagen, wurde für eine grobe Schätzung des Parameters „Coliphagen gesamt“, der Messwert für somatische Coliphagen verdoppelt und ebenfalls dargestellt.

Die untere Grenze des dargestellten Unsicherheitsintervalls (hier: Toleranzintervall) beschreibt den Schätzwert, für den mit einer statistischen Konfidenz von 95% gesagt werden kann, dass 90% der Werte darüber liegen.

Für KA4 wird darauf hingewiesen, dass es sich bei den dargestellten Werten um den Parameter *C. perfringens* gesamt (Sporen + vegetative Zellen) handelt. Die unten beschriebenen Aussagen über die generell ausreichend hohen Konzentration im Zulauf von Klärwerken bleiben gültig. Die Daten wurden jedoch hinzugefügt, um zu zeigen, dass durch die Analyse von *C. perfringens* gesamt, die gemessenen

Konzentration um mehr als eine Größenordnung höher liegen können, als bei der ausschließlichen Analyse von Sporen.

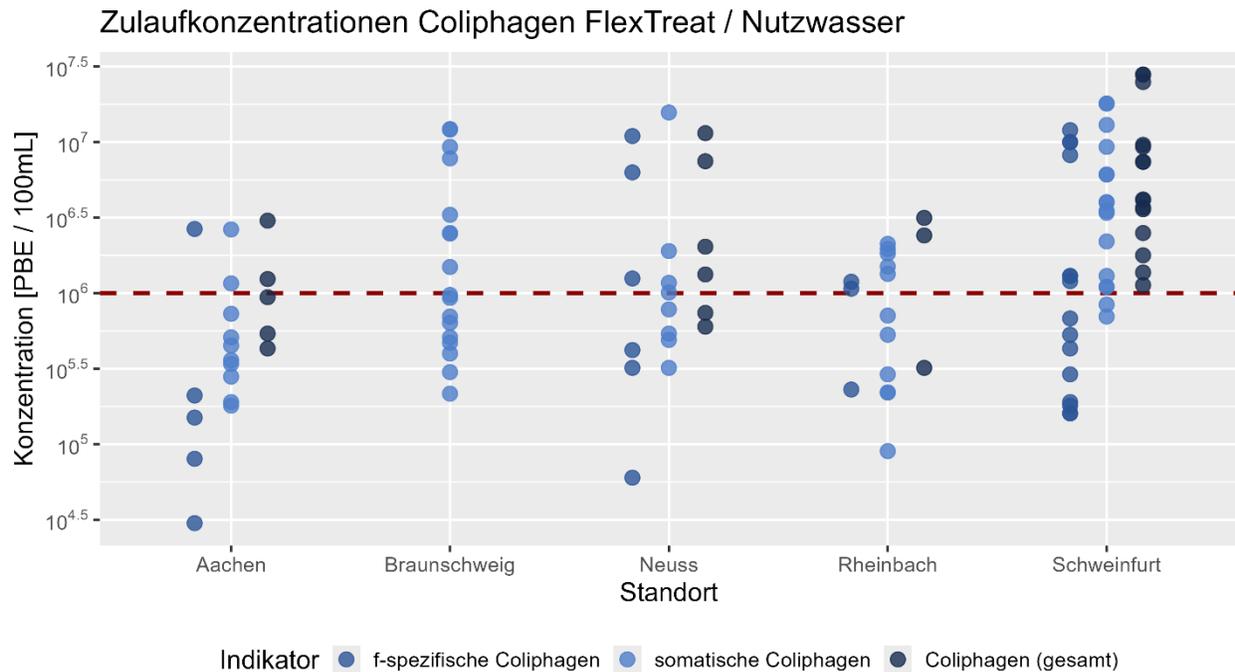


Abbildung 3: Vergleich zwischen somatischen Coliphagen, F-spezifischen Coliphagen und Coliphagen (gesamt) im Zulauf der vier FlexTreat Standorte und des Standorts Schweinfurt aus dem Schwesterprojekt Nutzwasser

Aus Abbildung 4 wird deutlich, dass für *E. coli* und *C. perfringens* in der Regel von hinreichend hohen Zulaufkonzentrationen ausgegangen werden kann. Für somatische Coliphagen und geschätzte „Coliphagen gesamt“ ist festzustellen, dass nur bei drei bis vier der zehn Kläranlagen die Zulaufkonzentration hinreichend hoch ist, um mit Standardmethoden und einem Analysevolumen von 100 mL eine \log_{10} -Reduktion in mindestens 90 % der Proben quantitativ validieren zu können, wenn am Ende der Aufbereitungskette ein Negativbefund für den bzw. die Parameter gemessen werden. Es ist folglich davon auszugehen, dass die mit Hilfe von Standardmethoden ermittelten Zulaufkonzentrationen an Coliphagen in den meisten Fällen nicht ausreichend hoch sind, um das Validierungsziel quantitativ validieren zu können.

Der Umgang mit berechneten \log_{10} -Reduktionswerten, die aufgrund zu niedriger Zulaufkonzentrationen als „> WERT“, angegeben werden müssen, da im Ablauf < 1 PFU/Volumeneinheit gemessen wurde, hat erheblichen Einfluss auf die verfügbaren Optionen der Versuchsplanung und Durchführung. Laut Verordnung 2020/741 können Werte dieser Art, als solche betrachtet werden, die die Validierungsanforderungen einhalten. Im Wortlaut:

„Wenn ein biologischer Indikator nicht in ausreichender Menge im Rohabwasser vorhanden ist, um die \log_{10} -Reduktion zu erreichen, bedeutet das Fehlen eines solchen biologischen Indikators im aufbereiteten Wasser, dass die Validierungsanforderungen eingehalten werden.“ (EU 2020/741 Anhang 2)

Im Gegensatz dazu wird in deutschen Kontext diskutiert, dass eine Validierung der Reinigungsleistung in jedem Fall, also auch für Fall, dass die Anzahl an Indikatororganismen pro 100 mL im Zulauf der Kläranlage

nicht ausreichend hoch ist, quantitativ validiert werden muss. Falls dies so umgesetzt werden sollte, ist die Versuchsplanung und Durchführung entsprechend anzupassen. Für eine solche Anpassung ist vor allem eine Anreicherung größerer Probenvolumina im Ablauf des Klärwerks zu nennen (vgl. Kapitel 4.6). Zwar ist für die Validierung einzelner Behandlungsstufen (z.B. UV-Desinfektion) theoretisch auch eine Erhöhung der Zulaufkonzentration durch Spiking von Indikatororganismen denkbar, für eine Validierung eines Gesamtsystems im Großmaßstab sind Spikingverfahren jedoch nicht praktikabel.

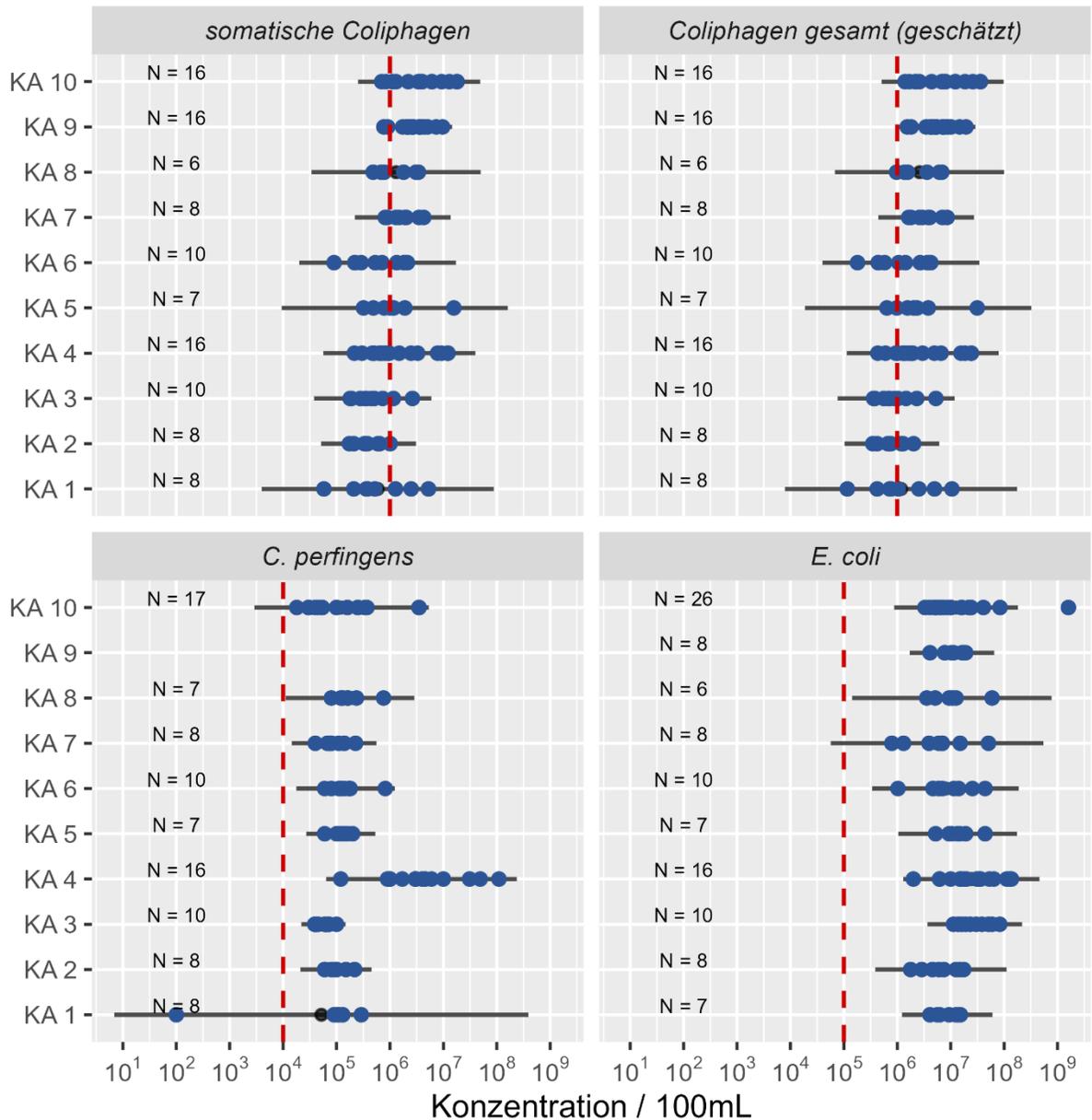


Abbildung 4: Zulaufkonzentration verschiedener Indikatoren in zehn Klärwerken in Deutschland. Blaue Punkte: Messwerte, Unsicherheitsintervalle (schwarz) entsprechen dem unteren und oberen einseitigen Toleranzintervall aus den verfügbaren Daten. Rote Linien entsprechen dem jeweiligen Validierungsziel. (Quelle: IHPH, KWB, TZW)

4 Versuchsplanung und Durchführung

Im Folgenden werden die wesentlichen Aspekte und die Vorgehensweise eines Validierungsmonitorings zusammengefasst und Empfehlungen zu dessen Durchführung gegeben. Es wird davon ausgegangen, dass bereits die erforderlichen Vorversuche stattgefunden haben, um die notwendigen Betriebsparameter festzulegen (z.B. UV-Dosis), sodass es sich um die Phase 2 der Validierung des Systems handelt (vgl. Abbildung 1). Alle wesentlichen Schritte des Validierungsmonitorings sowie die angewendete Methodik (Probenahme, Analytik, Auswertung) sind im Vorfeld in einem detaillierten Protokoll festzuhalten.

4.1 Bestimmung des Validierungszeitraums und zu validierender Betriebsfenster

Das Validierungsmonitoring sollte idealerweise die wesentlichen Zeiträume und Betriebsbedingungen abdecken, in denen mit der Aufbereitung für die Wasserwiederverwendung zu rechnen ist. Dabei sollten mittelfristige periodische Schwankungen, wie beispielsweise saisonale Schwankungen ebenso berücksichtigt werden, wie kurzfristige, jedoch regelmäßig auftretende Schwankungen, wie beispielsweise aufgrund von Regenwetter.

Findet eine Aufbereitung zur Wasserwiederverwendung ausschließlich während der Vegetationsperiode im Trockenwetter statt, kann sich auf diese Zeiträume beschränkt werden. Wird jedoch angestrebt Wasser ganzjährig oder auch während des Regenwetterbetriebs aufzubereiten, und für die Bewässerung während der Vegetationsperiode oder Trockenphasen zu speichern, sollte das Validierungsmonitoring die genannten Schwankungen mit einbeziehen. Dies bedeutet, dass die genannten Betriebszustände entweder separat zu validieren sind oder glaubhaft dargelegt werden kann, warum in einem spezifischen Fall die genannten Schwankungsursachen unwesentlich für das Validierungsergebnis sind. Da zu erwarten ist, dass der Bedarf an Bewässerungswasser vor allem während der Vegetationsperiode bei Trockenheit am höchsten ist, kann durch eine Validierung der Trockenwettersituation im Frühjahr und Sommer bereits das Gros der Anwendungsfälle effizient erfasst und validiert werden, und eine entsprechende Betriebserlaubnis erteilt werden.

Aus der Dokumentation sollte hervorgehen auf welche „Betriebsfenster“ und „Betriebszustände“ sich die Validierung bezieht. Die Validierung ist ausschließlich für dieses Betriebsfenster gültig. Die Verordnung EC 2020/741 gibt mit den Parametern zur Routineüberwachung bereits einen Rahmen für die Definition von Betriebsfenstern vor. Für Rahmenbedingungen im deutschen Kontext sei an dieser Stelle auf das DWA Merkblatt M1200 verwiesen (DWA, 2024).

4.2 Betriebsstörungen

Kurzzeitige Betriebsstörungen der Kläranlage (z.B. Havarien oder Schlammabtrieb) oder der weitergehenden Aufbereitungsstufen, können die Reinigungsleistung des Gesamtprozesses beeinträchtigen. Da solche Beeinträchtigungen im Vorfeld weder ausgeschlossen noch vollständig verhindert werden können, sind im Rahmen eines zu erarbeitenden Risikomanagementplans, geeignete Maßnahmen zur Erkennung von Betriebsstörungen festzulegen. Darauf aufbauend sind geeignete Handlungsmaßnahmen zu definieren, um die Qualität des aufbereiteten Wassers sicherzustellen. Ist dies nicht der Fall, kann das Wasser nicht für Bewässerungszwecke genutzt werden und ist entsprechend in das Gewässer abzuleiten.

Betriebsstörungen, die während des Validierungsmonitorings stattfinden, können das Validierungsergebnis unter Umständen stark negativ beeinflussen. Daher ist im Vorfeld klar zu definieren, wie mit den unter diesen Umständen erhobenen Messwerten zu verfahren ist.

Die Europäische Badegewässerrichtlinie kann hier als Vorlage für einen pragmatischen Ansatz dienen. Gemäß dieser Richtlinie könnten Überwachungswerte in bestimmten Fällen aus der Bewertung der Badegewässerqualität „herausgenommen“ werden. Voraussetzung hierfür ist, dass die zuständige Behörde rechtzeitig über die Verschmutzung informiert, und so eine Exposition von Badenden zu kontaminiertem Wasser verhindert wird. Darüber hinaus muss die Ursache der Verschmutzung bekannt sein.

Überträgt man dieses Konzept auf das Validierungsmonitoring, kann wie folgt vorgegangen werden.

1. Während der Planungsphase der Validierung werden die Probennahmetermine für die Durchführung des Validierungsprogramms im Validierungsprotokoll festgehalten.
2. Der Betreiber der Anlage hat die Möglichkeit kurzfristig, rechtzeitig im Vorfeld der Analyse einer Probe, auf eine Betriebsstörung hinzuweisen. Die genommene Probe dann wird nicht für die Validierung verwendet.
3. Falls eine Betriebsstörung angezeigt wird, darf eine potenziell trotzdem analysierte Probe nicht in die Auswertung zu Validierung mit einfließen.
4. Die Art der Betriebsstörung ist zu dokumentieren.

4.3 Ort der Probennahme

Zur Validierung der Log_{10} -Reduktion sollten Proben im Zulauf des Klärwerks sowie im Ablauf der weitergehenden Abwasserbehandlung erfolgen, also am Anfang und am Ende der Ausbereitungskette. Für die Probennahme mikrobiologischer Untersuchungen existiert mit ISO 19458:2006 eine internationale Norm, die eine Anleitung zur Planung eines Probennahmeprogramms, dessen Durchführung und den Anforderungen zu Transport und Lagerung bereitstellt. Der Norm folgend, sollten grundsätzlich Probennahmeorte gewählt werden, die repräsentativ sind und alle vertikalen, horizontalen und zeitlichen Veränderungen berücksichtigen. Probennahmestellen, an denen die Bedingungen instabil sind, sollten vermieden werden. Täglich auftretenden Schwankungen können auch durch die Art der Probennahme verringert werden. Beispielsweise können 24 h-Mischproben genutzt werden, um auftretende Schwankungen zu verringern (vgl. Kapitel 4.4). Im Projekt Flextrat durchgeführt Stabilitätsversuche zeigen, dass die Konzentrationen von 24h-Mischproben auch bei längerem Transport und Lagerungszeiten stabil blieben (siehe Abbildung 6).



Abbildung 5: Bild eines gekühlten automatischen Probennehmers am Zulauf der Kläranlage Steinhof in Braunschweig.

4.4 Art der Probennahme

Wie in Kapitel 4.3 erwähnt, sollten bei der Probennahme zeitliche Schwankungen an der Probennahmestelle berücksichtigt werden, um einen repräsentativen Datensatz zu generieren. Hinsichtlich der Art der Probennahme kann grundsätzlich zwischen Mischproben und Stichproben unterschieden werden. Werte aus Mischproben stellen das arithmetische Mittel der zugrunde liegenden Einzelproben dar, aus denen die Mischprobe besteht. Das arithmetische Mittel minimiert als Schätzwert den quadratischen Fehler der zugrunde liegenden Daten. Folglich stellt ein aus 24 h-Mischproben gewonnener Messwert eine weniger fehlerbehaftete und somit robustere Schätzung für die Zulaufkonzentration des beprobten Tages da, als dies bei der Verwendung von einzelnen Stichproben der Fall wäre. Eine möglichst robuste Schätzung der Zu- und Ablaufkonzentration sollte Ziel bei der Bestimmung der Log_{10} -Reduktion sein. Vor diesem Hintergrund sind Mischproben grundsätzlich Stichproben vorzuziehen. Mischproben können sowohl kontinuierlich als auch diskontinuierlich genommen werden. Beide Ansätze werden als geeignet angesehen.

Als weitere Argumente für die Nutzung von Mischproben sind zu nennen:

1. Durch die Schätzung der mittleren Zulaufkonzentration im Zulauf und Ablauf, sind die gewonnenen Werte klarer miteinander in Verbindung zu bringen (siehe gepaarte Auswertung in Kapitel 5.3.1).
2. Die Verteilung der mittleren Zulaufkonzentration hat eine kleinere Varianz als die Populationsverteilung von einzelnen Stichproben. Dies reduziert die Streuung der berechneten Log_{10} -Reduktionsverteilung aufgrund von zufälligen Fehlern.

Als Kritikpunkt hinsichtlich der Nutzung von 24 h-Mischproben bei der Bestimmung von mikrobiologischen Parametern, wird teilweise die Verweildauer der ersten Aliquote der Probe im Probennehmer angeführt. Tabelle 2 zeigt die empfohlenen und akzeptablen Lagerungsdauern (einschließlich Transport) für die relevanten Validierungsparameter.

Tabelle 2: Empfohlene und akzeptable maximale Lagerungsdauern für verschiedene Indikatororganismen nach DIN ISO 19458:2006

Indikator	Empfohlene maximale Lagerungsdauer (T = 2-8 °C) [h]	Akzeptable maximale Lagerungsdauer (T = 2-8 °C) [h]
Fäkalindikatoren (<i>E. coli</i> , intestinale Enterokokken)	12	18
Sporen (<i>Clostridium</i> spp.)	24	72
Viren (Bakteriophagen)	48	72

Wie aus der Tabelle hervorgeht, bewegen sich die durch 24 h-Mischproben entstehenden Lagerungszeiten für Sporen und Viren grundsätzlich noch im akzeptablen Bereich, für Viren kann sogar der empfohlene Bereich leicht eingehalten werden. Kritischer sind in diesem Zusammenhang die Lagerungsdauern für den Parameter *E. coli* zu bewerten. Hier werden für die ersten Aliquote der 24h-Mischprobe die maximalen Lagerungszeiten nach DIN ISO 19458:2006 überschritten. Zu bemerken ist, dass innerhalb der DIN ISO 19458:2006 auf die unterschiedliche Informationslage hinsichtlich der maximalen Transportzeit hingewiesen wird.

Vor diesem Hintergrund wurden im Projekt FlexTreat zusätzliche Stabilitätsversuche durchgeführt, um die Stabilität des Parameters *E. coli* besser zu untersuchen. Die Ergebnisse aus fünf Experimenten zeigen, dass es nach 48-72 h in keiner der analysierten Proben zu einer statistisch signifikanten Abnahme der Ausgangskonzentration kam (Abbildung 6). Die Unsicherheit, die durch eine längere Verweilzeit der ersten Aliquote im Probennehmer entsteht, wird demzufolge als tolerierbar angesehen. Die Vorteile durch die Nutzung von Mischproben die Probenahmeunsicherheit im Vergleich zu Stichproben zu reduzieren, und so eine robustere Schätzung der Zulaufkonzentration zu erzielen, überwiegen hier deutlich.

Erfahrung aus FlexTreat

Im Projekt FlexTreat wurde der Effekt der Transportzeit auf die Stabilität der Messwerte von *E. coli* im Labor anhand von fünf Stichproben untersucht. Hierfür wurde dieselbe Probe mehrere Male nach definierten Zeitschritten (0, 8h, 24h, 48h, 72h) analysiert. Zwischen den Messungen wurde die Probe gekühlt, um einen realen Transportvorgang zu simulieren. Die Startkonzentration schwankten zwischen 800 (Stichprobe 3) und 120.000 KBE/100mL (Stichprobe 1), sodass eine große Bandbreite abgedeckt wurde. Die Ergebnisse sind als Abweichung bezüglich der Ausgangskonzentration in der unteren Abbildung dargestellt. Die Ergebnisse zeigen, dass die Messwerte auch nach 72 Transportzeit als weitestgehend stabil anzusehen sind. Die Varianz innerhalb der gewonnenen Trendlinien ist auf die Empfindlichkeit zurückzuführen, die durch die niedrige Anzahl an Datenpunkten ($n=4$) entsteht. Die Fehlerbereiche schließen bis auf Stichprobe 2 stets die 0 mit ein, für Stichprobe 2 ist ein leichter Anstieg zu beobachten, also keine Erniedrigung. Die Ergebnisse zeigen, dass auch für den Parameter *E. coli* 24-Mischproben eine potenzielle Möglichkeit darstellt, die Validierungsmessung durchzuführen.

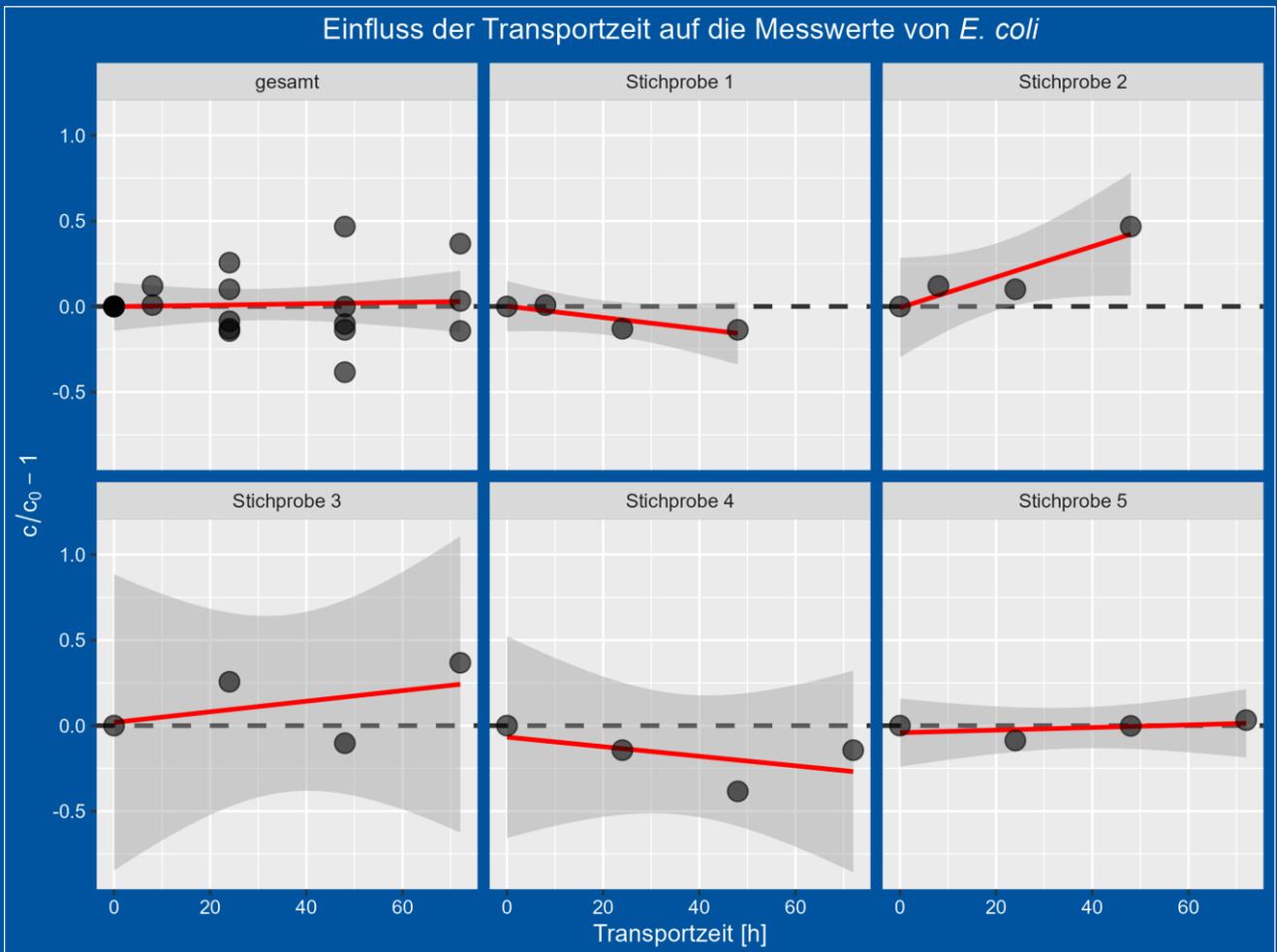


Abbildung 6: Stabilitätsversuche für *E. coli* (IHPH).

4.5 Auswahl des Analysenverfahrens

Für die genannten Parameter stehen genormte Analysenverfahren zur Verfügung. Für das Validierungsmonitoring hinsichtlich Güteklasse A, sollte für die Analyse des *Klärwerkablaufs* auf Methoden zurückgegriffen werden, die als untere Bestimmungsgrenze einen Wert von < 1 aufweisen. Dies erleichtert die statistische Auswertung, da Werte kleiner 1 für diskrete Größen wie mikrobiologische Indikatororganismen gleich 0/Volumeneinheit gesetzt werden können.

Messparameter	Zulauf	Ablauf (mit BG < 1)
<i>Clostridium perfringens</i> Sporen		DIN ISO 141189:2016
Somatische Coliphagen,		DIN ISO 10705-2:2001
F-spezifische Coliphagen		DIN ISO 10705-1 2001
<i>E. coli</i>	DIN ISO 9308-2 DIN ISO 9308-3	DIN ISO 9308-1 DIN ISO 9308-2

Für den Zulauf des Klärwerks sind die Verdünnungen ausreichend hoch zu wählen, um die zu erwartenden Zulaufkonzentrationen (siehe Abbildung 4) abzudecken. Anders als im Ablauf, werden alle gängigen Methoden, die im Abwasserbereich eingesetzt werden, als geeignet angesehen.

4.6 Anreicherung großer Probenvolumina im Klärwerkablauf bei zu niedrigen Zulaufkonzentrationen

Das folgende Kapitel befasst sich explizit mit der Fragestellung, wie die Log_{10} -Reduktion einer Anlage quantitativ validiert werden kann, falls die vorliegenden Zulaufkonzentrationen zu niedrig sind, um die notwendige Log_{10} -Reduktion in 90 % der Fälle nachweisen zu können. Wie in Abbildung 4 dargestellt, ist diese Problemstellung vor allem für die Validierung der Virenreduktion mit Hilfe von Coliphagen zu erwarten. Die folgenden Beispielanwendungen beziehen sich von daher stark auf die Anreicherung von Coliphagen, sind aber grundsätzlich auch für andere Parameter anwendbar. Da, wie im Folgenden gezeigt wird, nur ein Teil des angereicherten Eluats für die Analyse von Bakteriophagen benötigt wird, bietet sich an, das übrige Volumen für die Bestimmung der anderen mikrobiologischen Parameter *E. coli* und *C. perfringens* Sporen zu nutzen, um höhere Log_{10} -Reduktionswerte validieren zu können.

4.6.1 Allgemeine Beschreibung des Auswertungsansatzes

Die hier vorgestellte Methode basiert auf einer Anreicherung großer Probenvolumina im Ablauf der Aufbereitungsanlage, und daraus abgeleiteten Konzentrations- und Korrekturfaktoren. Sie basiert auf folgenden Schritten:

Schritt 1: Bestimmung des Konzentrationsfaktors durch Filtration und Elution großer Probenvolumina

Ziel der Anreicherung der Viren ist es, die Anzahl der vorliegenden Coliphagen von einem großen auf ein kleineres, analysierbares Volumen zu überführen und somit einer mikrobiologischen Analyse zugänglich zu machen.

Werden beispielsweise 15 L Probenvolumen über eine Ultrafiltrations-Hohlfasermembran (vgl. Kapitel 4.6.3) gefiltert und letztere anschließend mit 400 mL ausgewaschen, wurde unter Annahme einer fiktiven Wiederfindungsrate von 100 % die Anzahl von Viruspartikeln von einem Volumen von 15 L auf ein Volumen von 400 mL überführt. Die Probe wurde angereichert bzw. aufkonzentriert. Durch das Verhältnis der Volumina lässt sich ein Konzentrationsfaktor (K) ermitteln, der im vorliegenden Beispiel einen Wert von 37,5 annimmt (Gleichung (4-1)).

$$K = \frac{V_{Probe}}{V_{Eluat}} = \frac{15000 \text{ mL}}{400 \text{ mL}} = 37,5 \quad (4-1)$$

Schritt 2: Berücksichtigung der Wiederfindungsrate und des analysierten Eluatvolumens zur Bestimmung des „repräsentativen Ablaufvolumens“ (V_R)

Würde das gesamte Eluat (im Beispiel 400 mL) analysiert und wäre die Annahme einer Wiederfindungsrate von 100 % korrekt, wäre der ermittelte Messwert repräsentativ für das gesamte filtrierte Volumen (hier 15 L). Das Ergebnis könnte als PBE/15L angegeben werden.

In der Regel ist jedoch beides nicht der Fall, sodass sowohl das analysierte Eluatvolumen ($V_{Eluat,analysiert}$), als auch die Wiederfindungsrate (R_{WF}) berücksichtigt werden müssen. Im folgenden Rechenbeispiel wird angenommen, dass 100 mL des Eluats analysiert werden und die Wiederfindungsrate 0,5 beträgt.

$$V_R = K \times V_{Eluat,analysiert} \times R_{WF} = \quad (4-2)$$

$$37,5 \times 100 \text{ mL} \times 0,5 = 1875 \text{ mL} \approx 1,9 \text{ L}$$

Da im Beispiel nur 100 mL des Eluats analysiert werden, und die Wiederfindung nur 50% beträgt, ist der gewonnene Messwerte nicht repräsentativ für 15 L, sondern nur für ca. 1,9 L.

Eine solche Anpassung über das Probenvolumen ist vor allem deswegen notwendig, da bei negativen Befunden < 1 , also 0 / Volumeneinheit, die so ermittelte 0 nicht auf größere Volumina hochgerechnet werden darf. Falls im Rechenbeispiel in 100 ml des analysierten Eluats ein Negativbefund von 0 PBE/100 mL ermittelt würde, wäre es falsch, diesen Wert mit 4 zu multiplizieren und so für das gesamte Eluat von 400 mL von einem Wert von 0 / Volumeneinheit auszugehen¹.

Von daher findet innerhalb der vorgestellten Methodik findet die Hochrechnung im *Zulauf* der Anlage über das repräsentative Ablaufvolumen statt (Schritt 3), wo von Negativbefunden nicht auszugehen ist.

Schritt 3: Anpassung der Zulaufkonzentration auf Basis des „repräsentativen Ablaufvolumens“

Wie im Beispiel von Schritt 2 erläutert, sind die durch die Anreicherung gewonnenen Messwerte repräsentativ für 1,9 L Probenvolumen, oder präzise 1875 mL. Für die Berechnung der Log₁₀-Reduktion sind im nächsten Schritt die Zulauf- und Ablaufwerte auf dasselbe Probenvolumen zu beziehen. Zulaufwerte werden nach Standardverfahren in der Regel auf 100 mL bezogen. Um die

¹ anders verhält es sich falls in 100 mL ein Wert > 0 gemessen wird. Hier wäre eine Hochrechnung möglich.

Zulaufkonzentration ebenfalls auf das ermittelte „repräsentative Ablaufvolumen V_r “ von 1875 mL zu beziehen, kann die Zulaufkonzentration $c_{zu,100ml}$ entsprechend korrigiert werden.

$$c_{zu,1875mL} = c_{zu,100mL} \frac{V_r}{100 mL} =$$

$$c_{zu,100mL} \frac{1875ml}{100ml} = c_{zu,100mL} \times 18,75 \quad (4-3)$$

Im dargestellten Beispiel liegt die korrigierte Zulaufkonzentration um den Faktor 18,75 oberhalb des Konzentrationswerts einer Standardprobe von 100 mL. Auf diese Weise wird der beobachtbare Unterschied zwischen Zulauf- und Ablaufwerten also um den Faktor 18,75 erhöht, und die Daten für eine Validierung zugänglich. Werden beispielsweise in der angereicherten Ablaufprobe keine Phagen nachgewiesen, so kann durch die Anreicherung mehr als eine Größenordnung gewonnen werden.

4.6.2 Bestimmung des zu filtrierenden Volumens

In Kapitel 4.6.1 wurde das allgemeine Vorgehen beschrieben, wie durch eine Anreicherung größerer Probenvolumen im Klärwerksablauf, der quantifizierbare Unterschied zwischen Zu- und Ablauf vergrößert werden kann, um den Validierungsprozess zu ermöglichen (bei zu niedrigen Zulaufwerten) bzw. zu vereinfachen. Für eine praktische Anwendung dieser Methode muss das minimale Volumen abgeschätzt werden, das für eine erfolgreiche Validierung gefiltert werden muss. Es ist hervorzuheben, dass es sich bei dieser Abschätzung um das minimale Volumen handelt. Größere Volumina können stets genommen werden.

Um das minimale, zu filtrierende Volumen ($V_{min, \text{filtriert}}$) abzuschätzen, werden die folgenden Eingangsdaten benötigt:

1. Eine Abschätzung des „repräsentativen Ablaufvolumens“ V_r
2. Eine Abschätzung der Wiederfindungsrate R_{WF}
3. Das Volumen des Eluats V_{eluat} , sowie das Volumen des tatsächlich analysierten Eluats

Das zu filtrierende Ablaufvolumen kann wie folgt berechnet werden:

$$V_{min, \text{filtriert}} = \frac{V_r}{R_{WF} R_A} \quad (4-4)$$

mit:

$$R_A = \frac{V_{Eluat, \text{analysiert}}}{V_{Eluat, \text{gesamt}}} \quad (4-5)$$

V_r : repräsentatives Ablaufvolumen [mL]

R_{WF}: Wiederfindungsrate der Aufkonzentrierungsmethode [-]

R_A: tatsächlich analysierter Anteil der aufkonzentrierten Probe [-]

4.6.2.1 Repräsentatives Ablaufvolumen V_r

Verallgemeinerbare Werte für das repräsentative Ablaufvolumen sind in Tabelle 3 dargestellt. Die Werte wurden mit Hilfe eines Bayes'schen hierarchischen, statistischen Modells anhand von Daten von neun Klärwerken abgeleitet die in Abbildung 4 dargestellt sind. Neben den Werten für das benötigte repräsentative Ablaufvolumen ist auch der Grad der statistischen Sicherheit angegeben mit dem 90 % der Zulaufwerte größer sind als der zu validierende Wert. Um 95 % sicher sein zu können, dass 90 % der Zulaufwerte für Coliphagen größer als die zu validierende \log_{10} -Reduktion von 6,0 sind, wäre für das repräsentative Ablaufvolumen ein Wert von 1,3 L anzusetzen, für eine Sicherheit von 99 % ein Wert von 2,4 L.

Die hier dargestellten Werte können als konservative Schätzung verstanden werden, da sie ausschließlich auf Zulaufwerten für somatische Coliphagen beruhen. Die ermittelten Volumina für die Filtration sind folglich auch für den Fall ausreichend, wenn ausschließlich auf den Parameter somatische Coliphagen zurückgegriffen würde. Da für die Validierung empfohlen wird, die Summe aus somatischen und F-spezifischen Coliphagen zu verwenden, sind die hier dargestellten Werte auch als ausreichend hoch für eine Validierung des Parameter Coliphagen (gesamt) anzusehen. Als Alternative zu den hier dargestellten Referenzwerten, kann das benötigte repräsentative Ablaufvolumen auch in einem Vorversuch ermittelt werden (siehe Anhang I). Dies ist jedoch mit einem erhöhten Aufwand verbunden.

Tabelle 3: Abschätzung des benötigten, zu filtrierenden Volumens für Coliphagen auf Basis der Messwerte aus Abbildung 4

Zielwert	Anteil	Konfidenzniveau (1- α)	Einseitige untere Toleranzgrenze*	Korrekturfaktor repräsentatives Volumen [L]**
6	90%	80%	5,13	0,8
6	90%	90%	5,01	1
6	90%	95%	4,9	1,3
6	90%	99%	4,6	2,4

* Schätzverfahren über Bayes'sches hierarchisches Modell unter Berücksichtigung der statistischen Parameterunsicherheit
** aufgerundete Werte

4.6.2.2 Wiederfindungsrate der Anreicherungs-methode

Während des Anreicherungsprozesses von Mikroorganismen kommt es in der Regel zu Verlusten, sodass sich nicht alle Organismen, die sich in der ursprünglichen Probe befanden, wiederfinden lassen. Bei einer Anreicherung unter Verwendung von Hohlfasermembranen (vgl. Kapitel 4.6.3 und 4.6.4) wäre es denkbar, dass ein gewisser Anteil der Phagen die Membran passiert. Des Weiteren ist es denkbar, dass ein gewisser Anteil an der Membran haften bleibt und sich nicht eluieren lässt. Die Wiederfindungsrate R_{WF} beschreibt den Anteil des Zielparameters, der sich in der aufbereiteten Probe wiederfinden lässt.

Für eine Anreicherung mittels Hohlfasermembranen wurde eine weite Spanne an Wiederfindungsraten berichtet. Hill et al. (2009) berichten sehr hohe Wiederfindungsraten von im Mittel 77% (+/- 15%) und 100% (+/- 20%) für die Phagen vom Typ phi X174 (somatisch) und MS2 (F-spezifisch). Die niedrigsten

Wiederfindungen wurden für Echovirus ermittelt mit im Mittel 58% (+/- 8%), woraus sich eine konservative Spanne von 40-100% abschätzen lässt.

Im Projekt FlexTreat wurde für somatische Coliphagen ein Wert von 41% +/- 11% bestimmt. Innerhalb persönlicher Kommunikation mit dem Technologiezentrum Wasser (TZW), dass die Mikrosens Methode zur schnellen Anreicherung und besseren Bestimmung hygienischer Parameter entwickelt hat (die auch in FlexTreat verwendet wurde), wurde die Wiederfindung auf 50% für somatischen Coliphagen geschätzt.

Vor diesem Hintergrund wird bei der Verwendung der Mikrosens-Methode ein Wert von 0,4 als plausibler und hinreichend konservativer Schätzwert angesehen, der für die Wiederfindungsrate angesetzt werden kann.

4.6.2.3 Analysierter Anteil der aufkonzentrierten Probe

Der analysierte Anteil der aufkonzentrierten Probe R_A ist die am einfachsten zu ermittelnde Eingangsgröße, da sowohl das Volumen des Eluats als auch das analysierte Volumen leicht erfasst werden kann, beziehungsweise Teil der zugrunde liegenden Methode ist.

4.6.3 Anreicherungsverfahren

Die Anreicherung von Viren passiert in den meisten Fällen entweder über die Adsorption und anschließende Elution von Viren (VIRADEL), oder über die Filtration des Probenvolumens über Ultrafiltrationsmembranen (Hill et al., 2009; Shi et al., 2017). Der Anreicherung über Membranen werden in der Literatur höhere Wiederfindungsraten zugeschrieben als der VIRADEL Methode (Hill et al., 2009).

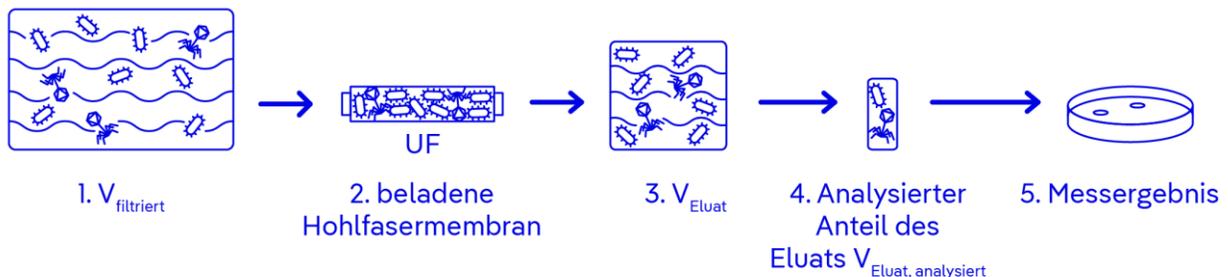


Abbildung 7: Illustrative Darstellungen eines Anreicherungsverfahrens über Hohlfasermembranen

4.6.4 Anreicherungsversuche im Projekt FlexTreat

Im Projekt FlexTreat wurde eine Anreicherung von somatischen und F-spezifischen Coliphagen über Hohlfasermembranen vom Typ REXEED 25 A, mit einem Molecular Weight Cut Off (MWCO) von 30kDa, und einer Filteroberfläche von 2,5 m² durchgeführt. Die Methode folgt der am Technologiezentrum Wasser (TZW) entwickelten Mikrosens Methodik (Kort et al., 2017).



Abbildung 8: Beispielhafte Darstellung einer Dialysemembran zur Aufkonzentrierung von mikrobiologischen Parametern.

Versuchsplanung

Tabelle 4 fasst die wesentlichen Eingangsgrößen zusammen, die für die Bestimmung des minimal zu filtrierenden Probenvolumens genutzt wurden. Die Berechnung ergab einen Wert von 13 L. Der berechnete Wert wurde auf einen Wert von 15 L aufgerundet. Größere Probenvolumina verbessern stets die Sensitivität des Verfahrens, sodass ein Aufrunden der berechneten Volumina möglich ist.

Tabelle 4: Eingangsgrößen für die Abschätzung des zu filtrierenden Volumens

Eingangsgrößen	Wert	Herkunft
V_{Eluat}	400 mL	Mikrosens
$V_{\text{Eluat, analysiert}}$	100 mL	ISO Methode
R_A	0,25	Gleichung (4-5)
R_{WF}	0,4	Siehe Kapitel 4.6.2.2
V_r	1,9 L	Tabelle 3
$V_{\text{min, filtriert}}$	13 L	Gleichung (4-4)
$V_{\text{real, filtriert}}$	15 L	aufgerundet von 13 Litern

Verifizierung von Annahmen

Um die Annahme von 40% für die Wiederfindungsrate zu verifizieren, wurden am Institut für Hygiene und Public Health (IHPH), des Bonner Universitätsklinikums, Versuche zur Wiederfindung von somatischen und F-spezifischen Bakteriophagen durchgeführt. Für somatische Coliphagen wurde ein Wert von 41% +/- 11% ermittelt, was innerhalb des Erwartungsbereichs liegt. Für F-spezifische Bakteriophagen wurden teilweise sehr hohe Wiederfindungsraten von > 85% ermittelt, was sich ebenfalls mit den Werten von Hill et al. (2009) deckt. Während hohe Wiederfindungen grundsätzlich zu begrüßen sind, erscheinen die ermittelten Werte ohne weitere Plausibilisierung, als sehr hoch, sodass weiterhin konservativ mit der niedrigeren Wiederfindung von 40% gerechnet wurde. Die Bestimmung der Wiederfindung für F-spezifische Phagen für die Anreicherung mittels Mikrosens ist somit ein noch offener Punkt.

Durchführung

Für die Anreicherung wurden insgesamt 7 Proben des Klärwerkablaufs im Klärwerk Braunschweig analysiert. Hierfür wurden 6 x 15 L und 1 x 16 L über eine Hohlfasermembran des Typs Typ REXEED 25 A, mit einem Molecular Weight Cut Off (MWCO) von 30 kDa, und einer Filteroberfläche von 2,5 m² gefiltert. Die beladenen Membranen wurden für die weitere Analyse per Übernachtkurier an das IHPH in Bonn geschickt. Die beladenen Membranen wurden mit einer angesetzten Pufferlösung eluiert und zur

Ermittlung des Konzentrationsfaktors das Gesamtvolumen des Eluats bestimmt. Daraufhin wurden 100 mL des Eluats nach Standardverfahren (vgl. Kapitel 4.5) auf somatische und F-spezifische Coliphagen analysiert.

Ergebnisse

Tabelle 5 fasst die Messergebnisse, das Eluatvolumen, sowie die darauf basierende Berechnung des repräsentativen Ablaufvolumen, und des Korrekturfaktors für die Zulaufkonzentration zusammen. Die Ergebnisse können so interpretiert werden, dass die in 100 mL Eluat gemessenen Werte (Spalte 1, Tabelle 5) repräsentativ für ein Volumen von 2020 mL angesehen werden können. Auf dieser Basis können die gemessenen Zulaufkonzentrationen pro 100 mL entsprechend angepasst werden. Der berechnete Korrekturfaktor beträgt 20,2. Die Korrektur wurde auf die Messergebnisse der Zulaufkonzentration angewendet und in Abbildung 9 dargestellt. Aus Abbildung 9 wird deutlich, dass durch die Anreicherung und die Erhöhung des Bezugsvolumens von 100 mL auf 2020 mL das 10. Perzentil der Verteilung von somatischen Coliphagen im Zulauf von einem Wert von $10^{5,4}$ PBE auf $10^{6,7}$ PBE pro Volumeneinheit erhöht werden konnte. Dieses liegt nun deutlich oberhalb des Zielwerts, sodass eine quantitative Validierung möglich ist.

Tabelle 5: Zusammenfassung der ermittelten Größen, Messwerte und Korrekturfaktoren.

Messwert Eluat	V _{Eluat, analysiert}	V _{Eluat}	V _{filtriert}	K	R _{WF}	V _r	Mittleres V _r	Korrektur Zulauf
PBE / 100mL	mL	mL	mL	-	-	mL	mL	-
0	100	325	15000	46,2	0,41	1892	2020 +/- 178	20,2 +/- 1,7
27	100	322	15000	46,6	0,41	1910		
12	100	329	15000	45,6	0,41	1869		
0	100	275	16000	58,2	0,41	2385		
0	100	305	15000	49,2	0,41	2016		
0	100	310	15000	48,4	0,41	1983		
0	100	295	15000	50,8	0,41	2084		

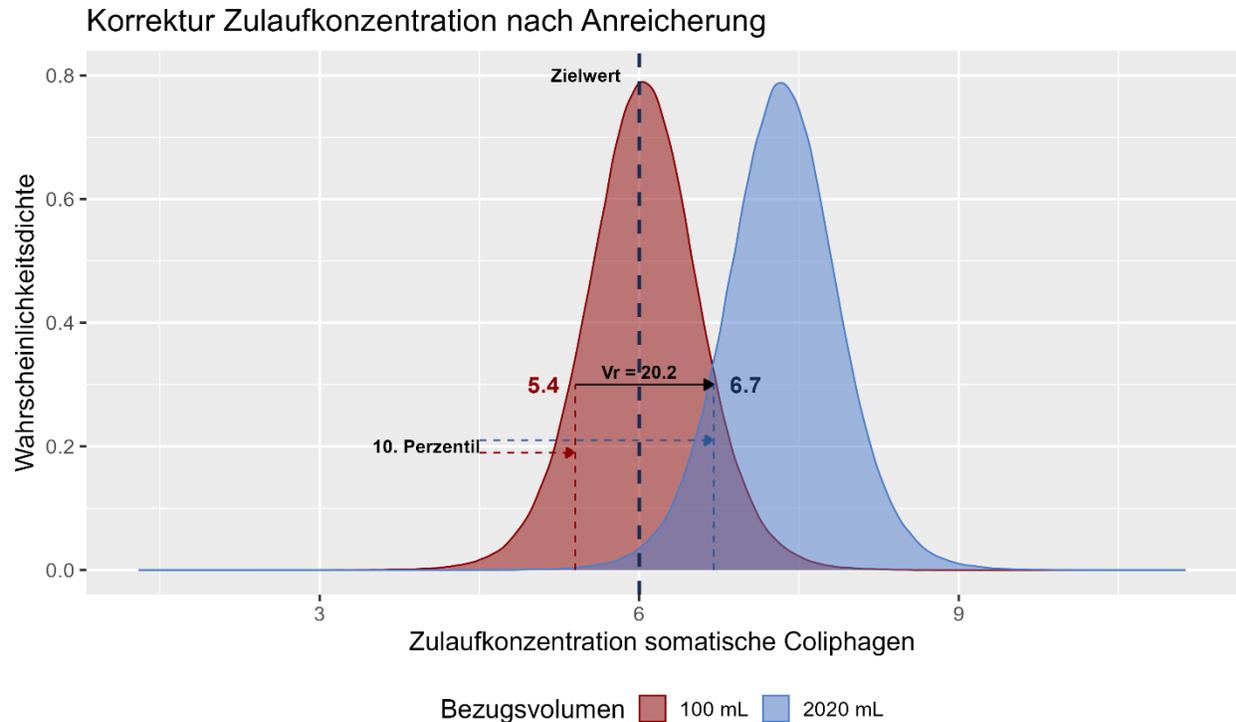


Abbildung 9: Darstellung der Korrektur der Zulaufkonzentration über eine Erhöhung des Bezugsvolumens bzw. repräsentativen Ablaufvolumens.

4.6.5 Weitere Anwendungsfälle

Die Anreicherung größerer Probenvolumina wurde vor allem als mögliche Methode eingeführt, um dem Problem zu niedriger Zulaufkonzentrationen zu begegnen. Wie in Kapitel 5.6 für das Beispiel *E. coli* gezeigt und diskutiert wird, sind für die Validierung umso weniger Validierungsproben notwendig, desto größer die Differenz zwischen dem zu validierenden Zielwert und den beobachteten, bzw. ermittelten Log_{10} -Reduktionswerten ist. Dies hängt damit zusammen, dass bei einer größeren Differenz zwischen dem Zielwert und Beobachtungswerten, auch mehr statistische Unsicherheit toleriert werden kann (vgl. Kapitel 5.1 und 5.5).

Vor diesem Hintergrund kann es für die Validierung günstig sein, auch bei ausreichend hohen Zulaufkonzentrationen größere Probenvolumina anzureichern, da hierdurch die Anzahl an Validierungsproben verringert werden kann. Dies ist insbesondere dann zu erwarten, wenn aufgrund der installierten Reinigungstechnik eine Log_{10} -Reduktion erwartet werden kann, die deutlich über dem zu validierenden Zielwert liegt. Da in der Regel nicht das gesamte Eluat für die Analyse der Bakteriophagen verwendet wird, kann das nicht-verwendete Eluatvolumen beispielsweise für die Analyse von *E. coli* und *C. perfringens* Sporen verwendet werden, ohne dass dadurch ein zusätzlicher analytischer Aufwand entsteht.

5 Datenauswertung

Im Vorfeld des Validierungsmonitorings muss festgehalten werden, wie die erhobenen Daten ausgewertet werden. Die Methodik der Datenauswertung ist Bestandteil des Validierungsprotokolls. Wie in Kapitel 2 bereits angedeutet, stellt das Fehlen klarer Vorgaben zur statistischen Auswertung eine wesentliche Regelungslücke dar. Im Folgenden werden unterschiedliche Interpretationsmöglichkeiten dargestellt und deren Vor- und Nachteile diskutiert. Abschließend werden Empfehlungen gegeben, welches Verfahren vor dem Hintergrund der existierenden Vor- und Nachteile empfohlen wird. Zunächst werden grundsätzliche statistische Begrifflichkeiten und Zielgrößen beschrieben.

5.1 Log₁₀-Reduktion

Die wesentliche Messgröße für das Validierungsmonitoring ist die Log₁₀-Reduktion von mikrobiologischen Indikatororganismen durch die Abwasserbehandlung. Eine Reduktion der Konzentration um eine logarithmische Einheit zur Basis 10, also um eine Größenordnung, beschreibt eine Reduzierung der Indikatorkonzentration um 90 % (Tabelle 6).

Die Log₁₀-Reduktion kann selbst nicht direkt gemessen werden, sondern wird auf Basis von Zulauf- und Ablaufwerten berechnet. Sie ist somit eine Funktion der Zulauf und Ablaufkonzentration.

$$\text{Log}_{10}\text{Reduktion} = f(c(\text{Zulauf}), c(\text{Ablauf})) = \log_{10} \frac{c(\text{Zulauf})}{c(\text{Ablauf})} \quad (5-1)$$

Tabelle 6: Vergleich Log₁₀-Reduktion gegenüber Reduktion in %

Log ₁₀ -Reduktion	Reduktion in %
1	90
2	99
3	99,9
4	99,99
5	99,999
6	99,9999

5.1.1 Messwerte kleiner Bestimmungsgrenze

Messwerte mikrobiologischer Indikatoren werden in der Regel als ganzzahlige, diskrete Größen pro Volumeneinheit beschrieben, deren Einheiten sich je nach Messverfahren unterscheiden (siehe Kapitel 3). Konzentrationsangaben können, falls es sich um berechnete Größen handelt, auch als Dezimalgrößen angegeben sein.

Bei negativen Befunden am Ablauf des Klärwerks, also bei Werten <1 (= 0 für diskrete Größen), kann die tatsächliche Log₁₀-Reduktion nicht abschließend bestimmt werden. Für die Validierung kann nur eine Log₁₀-Reduktion validiert werden, die dem Wert der Zulaufkonzentration entspricht, sodass, für $c_{\text{Ablauf}} < 1$ gilt:

$$LRV_{validiert} = \log_{10}(c(\text{Zulauf}))^2 \quad (5-2)$$

Derselbe Wert kann gewonnen werden, indem die untere Nachweisgrenze in die Gleichung (5-2) für die Berechnung der \log_{10} -Reduktion eingesetzt wird. Das Verwenden der halben Bestimmungsgrenze wird nicht empfohlen, da dies zu einer Erhöhung der berechneten \log_{10} -Reduktion um 0,3 führen würde, was sich aus den Daten alleine nicht begründen lässt.

5.2 Zufallsvariablen, statistische Parameter, und Unsicherheitsintervalle

Zufallsvariable: Messwerte unterliegen stets einer gewissen Streuung oder Variabilität. Die Streuung, oder Verteilung von Messwerten wird häufig in Form von statistischen Verteilungen abgebildet, von denen die Gauß'sche Normalverteilung das bekannteste Beispiel ist. Wird eine Messgröße durch eine statistische Verteilung abgebildet, ist dieser Messgröße kein spezifischer Wert zugeordnet, sondern ein Zufallswert aus einer definierten Verteilung. Eine so spezifizierte Größe wird demzufolge Zufallsvariable genannt. Eine Verteilung von Messwerten ist somit eine Zufallsvariable. Beispielsweise können die zur Validierung berechneten \log_{10} -Reduktionswerte als Zufallsvariable dargestellt werden.

Parameter: Das Wort *Parameter* wird in verschiedenen Kontexten unterschiedlich genutzt. Im Kontext der Wasseraufbereitung werden Messgrößen wie beispielsweise die Sauerstoffkonzentration oder der pH-Wert oft als *Parameter* oder *Messparameter* bezeichnet. Davon zu unterscheiden sind *statistische Parameter*. Statistische Parameter beziehen sich auf nicht messbare Größen, die Eigenschaften einer Zufallsvariable oder statistischen Modells beschreiben. Für das Beispiel der Gauß'schen Normalverteilung sind die beiden Parameter, die diese Verteilung eindeutig beschreiben, der Mittelwert μ und die Standardabweichung σ . Auch das 10. Perzentil, also der Wert oberhalb dessen 90% der Werte der Grundgesamtheit erwartet werden, ist ein statistischer Parameter. In diesem Kapitel beschreibt der Begriff Parameter einen statistischen Parameter.

Parameterunsicherheit: Statistische Parameter werden auf Basis der beobachteten Messwerte geschätzt; sie sind selbst nicht direkt messbar. Das Wort *Parameterunsicherheit* bezieht sich in diesem Kontext auf die Unsicherheit hinsichtlich eines zu schätzenden Parameters, zum Beispiel der Unsicherheit über den „wahren Wert“ des Mittelwerts vor dem Hintergrund endlich vieler Messwerte. Die Frage ist, wie genau kann der wahre Mittelwert bestimmt werden, wenn nur eine limitierte Anzahl N an Messwerten vorliegt?

Grundsätzlich gilt, dass je größer die Anzahl der Messwerte bzw. Stichproben, desto kleiner die Parameterunsicherheit. Die Parameterunsicherheit wird in der klassischen Statistik meist über Konfidenzintervalle ausgedrückt, in der Bayes'schen Statistik über Wahrscheinlichkeitsverteilungen. Die Parameterunsicherheit ist entscheidend für Analysen, wie zum Beispiel Signifikanztests, durch die untersucht werden soll, ob beobachtete Unterschiede mehr sind als purer Zufall. In den meisten Fällen beziehen sich diese Analysen auf Unterschiede hinsichtlich des Mittelwerts.

Liegt das Konfidenzintervall des Mittelwerts beispielsweise in seiner Gesamtheit oberhalb eines bestimmten Wertes, ist der entsprechende Mittelwert statistisch signifikant unterschiedlich von diesem

² Der Nenner wird hier nicht berücksichtigt, da nicht durch 0 geteilt werden kann.

Wert. In der Regel wird eine Nullhypothese definiert, die annimmt, dass es keinen Unterschied zwischen beiden Größen gäbe. Letztere wird dann verworfen.

Für die Berechnung eines Konfidenzintervall ist es notwendig, das Konfidenzniveau oder Vertrauensmaß der statistischen Sicherheit anzugeben, das durch das Intervall abgedeckt werden soll. In der Wissenschaft wird hierfür meist 95% gewählt.

Wichtig ist, dass der aus den Daten berechnete Mittelwert als Punktschätzer „nur“ den wahrscheinlichsten Wert auf Basis der erhobenen Daten darstellt. Dies zählt auch für alle anderen Parameter wie Standardabweichungen oder Perzentile.

Punktschätzer alleine geben keine Auskunft über die statistische Unsicherheit und demzufolge auch keinerlei Auskunft über die Vertrauenswürdigkeit des ermittelten Wertes. Beispielsweise kann eine Stichprobe mit $N = 3$ und eine Stichprobe $N = 300$ zufällig identische Punktschätzer liefern, jedoch ist dem Wert mit $N = 300$ mehr Vertrauen entgegenzubringen, als demselben Wert aus drei Stichproben.

Um ein Aufbereitungsverfahren statistisch zu validieren, gilt es vor diesem Hintergrund mit einer zu definierenden statistischen Sicherheit bzw. Vertrauensmaß (z.B. 80%, 90% oder 95%) zu zeigen, dass ein zu definierender Zielparameter größer ist als der zu validierende Wert.

Unsicherheitsintervalle für zukünftige/nicht beobachtete Messwerte: Während sich Konfidenzintervalle auf die Unsicherheit statistischer Parameter beziehen, existieren auch Unsicherheitsintervalle, die sich auf die Verteilung unbeobachteter oder zukünftiger Messwerte beziehen. In diesem Zusammenhang sind vor allem Vorhersage- und Toleranzintervalle zu nennen. Vorhersageintervalle beziehen sich auf die Vorhersagen einzelner oder weniger Messwerte, während sich Toleranzintervalle sich auf einen definierten Anteil einer Grundgesamtheit beziehen, zum Beispiel 90% für den Anteil der Log_{10} -Reduktionswerte, die nach Verordnung 2020/741 die Zielwerte überschreiten sollen. Wie für Konfidenzintervalle ist es notwendig das Vertrauensmaß $1-\alpha$ festzulegen, mit dem ein Toleranzintervall einen bestimmten Anteil an Messwerten abdecken soll.

Einseitige und zweiseitige Unsicherheitsintervalle: Unsicherheitsintervalle können einseitig oder zweiseitig beschrieben werden.

Zweiseitige Unsicherheitsintervalle haben eine obere und eine untere Begrenzung, sodass zum Beispiel ein zweiseitiges 95%-Unsicherheitsintervall den Bereich zwischen 2,5% und 97,5% abdeckt. Zwischen diesen beiden Grenzen liegen 95% der Wahrscheinlichkeit.

Einseitige Unsicherheitsintervalle haben entweder nur eine untere oder eine obere Begrenzung, sodass ein einseitiges unteres Unsicherheitsintervall den Wahrscheinlichkeitsbereich zwischen 5% bis 100% abdeckt, während ein einseitiges oberes Unsicherheitsintervall den Wahrscheinlichkeitsbereich zwischen 0% und 95% abdeckt.

Um mit einem Vertrauensmaß von $1-\alpha$ zu berechnen, dass 90% der Messwerte größer sind als ein bestimmter Zielwert kann folglich die einseitige untere Toleranzgrenze berechnet werden. An dieser Stelle sei auf die Ähnlichkeit einer solchen einseitigen unteren Toleranzgrenze zur Definition eines 10. Perzentils hingewiesen. Letzteres beschreibt den Punktschätzer oberhalb dessen 90% der Messwerte erwartet werden. Wichtig ist jedoch, dass das 10. Perzentil alleine, im Gegensatz zur unteren Toleranzgrenze, keine

Aussage über das Vertrauensmaß $1-\alpha$ trifft. Die Parameterunsicherheit bleibt hier unberücksichtigt. Um diese zu berücksichtigen, müsste das Konfidenzintervall des 10. Perzentils berechnet werden.

Eine einseitige, untere Toleranzgrenze kann mit der unteren Grenze eines einseitigen, unteren Konfidenzintervalls des 10. Perzentils gleichgesetzt werden (vgl. Hahn et al. (1991), Kapitel 2.4.2). Das 10. Perzentil als Punktschätzer kann als untere Toleranzgrenze mit $\alpha = 0,5$ angesehen werden.

5.3 Überblick über Auswertungsalternativen

Im Folgenden werden verschiedene Validierungsansätze detaillierter betrachtet. Jeder dieser Ansätze setzt sich aus einem **Zielparameter θ** und dem dazugehörigen **Konfidenzniveau ($1-\alpha$) mit Unsicherheitsintervall** zusammen. Für alle Ansätze gilt, dass stets die einseitig untere Grenze des jeweiligen Unsicherheitsintervalls größer als der jeweilige Vergleichswert sein muss, damit die Validierung erfolgreich ist.

5.3.1 10. Perzentil vs. Erfolgsrate

Als Zielparameter kommt zum einen das 10. Perzentil der Verteilung der Log_{10} -Reduktionswerte in Frage. Das 10. Perzentil einer Verteilung beschreibt den Wert oberhalb dessen 90% der Werte zu erwarten sind. Für eine analytische Validierung sollte das 10. Perzentil mit ausreichender statistischer Sicherheit größer sein, als der zu validierende Zielwert, also der indikatorspezifischen Log_{10} -Reduktion nach Tabelle 1. Das 10. Perzentil wird dabei aus der Verteilung der ermittelten Log_{10} -Reduktionswerte ermittelt.

Alternativ lassen sich die ermittelten Log_{10} -Reduktionswerte auch in die Klassen „Zielwert eingehalten“ oder „Zielwert nicht eingehalten“ untergliedern. Der Anteil der Werte die den Zielwert einhalten, also die Rate „erfolgreicher“ Validierungsmessungen, kann über eine Binomialverteilung mit der Erfolgsrate r ausgedrückt werden. Zielparameter wäre folglich die Erfolgsrate r , die mit ausreichender statistischer Sicherheit größer sein sollte als 90% bzw. 0,9.

Tabelle 7 fasst die beiden Zielparameter und Validierungsansätze überblicksartig zusammen.

Neben den Unterscheidungen hinsichtlich des Zielparameters (10. Perzentil vs. Erfolgsrate r), gibt es zudem mehrere Möglichkeiten diese Parameter zu bestimmen bzw. deren Unsicherheit zu berücksichtigen. Dabei werden verschiedene Unterscheidungsmerkmale berücksichtigt.

Tabelle 7: Überblick über potenzielle Zielparameter und Validierungsansätze

Zielparameter	Vergleichswert	Bewertungsansatz	Geeignet für Log ₁₀ -Reduktion
10. Perzentil	Indikatorspezifische Log ₁₀ -Reduktion, wie in Verordnung 2020/741 definiert (6 Viren, 5 Bakterien, 4 Sporen von <i>C. Perfringens</i> oder 5 für sporenbildende Sulfatreduzierer)	Einseitige untere Toleranzgrenze > Vergleichswert (einseitiges unteres Toleranzintervall entspricht einseitigem Konfidenzintervall des 10. Perzentils)	empfohlen
Erfolgsrate r	90% oder 0,9	Einseitiges unteres Konfidenzintervall der Erfolgsrate r einer Binomialverteilung > Vergleichswert	Kann unter bestimmten regulatorischen und technischen Randbedingungen empfohlen werden*

*in Deutschland in der Regel nicht empfohlen

5.3.2 Gepaarte vs. ungepaarte Auswertung

Bei der gepaarten Auswertung werden Log₁₀-Reduktionswerte aus Wertepaaren von Zu- und Ablaufkonzentrationen berechnet. Auf Basis der so berechneten Log₁₀-Reduktionswerte werden dann entweder die Erfolgsrate r oder das 10. Perzentil (untere Toleranzgrenze) geschätzt. Der häufigste Fall ist eine Paarung nach Datum. Dies kann vor allem bei der Nutzung von 24h-Mischproben gut gerechtfertigt werden.

Alternativen beinhalten zufällige Paarungen oder Paarungen nach Rang, wobei letztere mit sehr starken statistischen Annahmen einhergehen. Charakteristisch für die gepaarte Auswertung ist, dass dieselbe Anzahl an Messwerten für die Zu- und Ablaufkonzentration benötigt werden. Falls im Zu- oder Ablauf eine ungleiche Anzahl an Werten vorliegt, geht durch die gepaarte Auswertung potenziell Information verloren, da nicht alle Messwerte genutzt werden können. Vorteilhaft ist eine vergleichsweise einfachere statistische Auswertung.

Neben der gepaarten Auswertung kann das 10. Perzentil (untere Toleranzgrenze) der Verteilung der Log₁₀-Reduktionswerte auch unter Annahme ungepaarter Messwerte erfolgen. Hierbei werden aus den Messwerten der Zu- und Ablaufkonzentrationen die statistischen Parameter der Verteilungen der Zu- und Ablaufwerte ermittelt. Über eine Monte-Carlo Simulation wird im Anschluss das Unsicherheitsintervall über das 10. Perzentil der Verteilung der Log₁₀-Reduktionswerte ermittelt (vgl. Kapitel 5.5.4.1).

5.3.3 Parametrische vs. nicht-parametrische Ansätze

Die Auswertung von N gepaarten Zu- und Ablaufwerte führt zu einer bestimmten Anzahl N an Log₁₀-Reduktionswerten. Aus diesen Werten wird als Zielparameter entweder die Erfolgsrate oder das 10. Perzentil bestimmt.

Für die Bestimmung des 10. Perzentils existieren sowohl parametrische Ansätze, als auch nicht-parametrische Ansätze. Bei parametrischen Ansätzen wird angenommen, dass sich die ermittelten Log_{10} -Reduktionswerte mit Hilfe einer bestimmten statistische Verteilung hinreichend gut abbilden lassen.

Aus dieser Annahme heraus lassen sich sowohl Punktschätzer für das 10. Perzentil, als auch die statistische Unsicherheit über das 10. Perzentil ableiten, die für die Konstruktion von Konfidenzintervallen benötigt wird.

Als Alternative zur parametrischen Bestimmung des 10. Perzentils, kann dieses auch über nicht-parametrische Verfahren geschätzt werden. Nicht-parametrische Verfahren ermitteln Perzentile über den statistischen Rang. Dies bedeutet, dass die gewonnen Messwert zunächst der Größe aufsteigend geordnet werden. Nicht-parametrische Schätzverfahren des 10. Perzentils geben als Ergebnis den Rang, also die Ordnungszahl, desjenigen Wertes, der am ehesten dem gesuchten Perzentil entspricht. Die Ordnungszahl kann dabei auch zwischen zwei Werten liegen. In diesem Fall wird zwischen dem darunter und darüber liegenden Wert linear interpoliert, wie dies beispielsweise bei der Ermittlung des Medians (50. Perzentil) einer Verteilung der Fall ist.

Nicht-parametrische Verfahren haben den Vorteil, dass keine spezifischen Annahmen über die Verteilung der Messwerte getroffen werden müssen. Nachteilig ist, dass für Aussagen mit gleichwertiger statistischer Sicherheit in der Regel deutlich mehr Messwerte benötigt werden als für Ansätze auf Basis parametrischer Perzentile. Darüber hinaus ist für Aussagen mit einem bestimmten Grad an statistischer Sicherheit ein Mindestmaß an Werten benötigt wird. Beispielsweise werden für die Aussage, dass mit einer statischen Sicherheit von 95%, 90% einer Grundgesamtheit oberhalb eines beobachteten Minimalwerts liegen mindestens 29 Stichproben benötigt.

5.4 Berechnung von Punktschätzern und Unsicherheitsintervallen für die Erfolgsrate r

Als Punktschätzer für die Erfolgsrate r kann in den meisten Fällen die folgende Gleichung herangezogen werden:

$$r(\text{Erfolg}) = \frac{N(\text{Erfolg})}{N(\text{Gesamt})} \quad (5-3)$$

mit:

$N(\text{Erfolg})$: Anzahl der berechneten Log_{10} -Reduktionswerte, die den Zielwert einhalten

$N(\text{Gesamt})$: Anzahl der Versuche

Für $N(\text{Misserfolg}) < 5$, wird für die Schätzung der Erfolgsrate, teilweise eine Korrektur angesetzt (vgl. z.B. Gelman et al. (2020), Kapitel 4.2):

$$r(\text{Erfolg}) = \frac{N(\text{Erfolg}) + 2}{N(\text{Gesamt}) + 4} \quad (5-4)$$

Wichtiger als der Punktschätzer ist jedoch die Berechnung von Unsicherheitsintervallen, deren untere einseitige Grenze für eine Validierung größer sein sollte als 0,9. Für die Berechnung von Unsicherheitsintervallen existieren verschiedenen Ansätze aus der klassischen und Bayes'schen Statistik, die im Folgenden kurz dargestellt werden.

5.4.1 Konfidenzintervall nach Wilson

Aus dem Bereich der klassischen Statistik ist das Konfidenzintervall nach Wilson zu empfehlen, dessen untere und obere Grenzen wie folgt berechnet werden können:

$$CI = \frac{1}{1 + \frac{z^2}{n}} \left(r + \frac{z^2}{2n} \pm z \sqrt{\frac{r(1-r)}{n} + \frac{z^2}{4n^2}} \right) \quad (5-5)$$

r = Anteil der Erfolge (Erfolgsrate) in der Stichprobe

n = Gesamtzahl der Beobachtungen in der Stichprobe

z = z-Wert der Normalverteilung, der dem gewünschten Konfidenzniveau entspricht (hier 1,645 für ein einseitiges unteres 95% Konfidenzintervall)

Softwareimplementierungen: In Excel existiert keine Funktion, die das Konfidenzintervall nach Wilson direkt berechnet. Formel (5-5) muss also manuell implementiert werden. Eine einfache Implementierung ist unter folgendem Link zu finden:

<https://github.com/KWB-R/Logremoval>

In anderen gängigen Skriptsprachen wie R oder Python existieren direkte Implementierungen, sodass Werte mit Hilfe eines Befehls erhalten werden können.

5.4.2 Bayes'sches Konfidenzintervall

In der Bayes'schen Statistik wird die Parameterunsicherheit über Wahrscheinlichkeitsverteilungen ausgedrückt. Für die Binomialverteilung wird die Erfolgsrate r über eine Beta-Verteilung beschrieben.

$$r \sim \text{Beta}(\alpha + N(\text{Erfolge}), \beta + N(\text{Misserfolge})) \quad (5-6)$$

α, β : Verteilungsparameter der a-priori Verteilung (hier: $\alpha = \beta = 1$)

N (Erfolge): Anzahl Validierungsmessung \geq Zielwert

N (Misserfolge): Anzahl Validierung $<$ Zielwert

Ein Vorteil von Bayes'schen Verfahren liegt darin, dass sich nicht nur bestimmen lässt, ob die untere Intervallgrenze größer ist als der Vergleichswert, sondern sich direkt auch die Wahrscheinlichkeit abschätzen lässt, mit der die wahre Erfolgsrate größer ist, als der Vergleichswert. Gegeben der aktuellen Datenlage also:

$$P(r > 0,9 \mid \text{Daten}) \quad (5-7)$$

Softwareimplementierung: Eine Excel basierte Softwareimplementierung ist unter folgendem Link zu finden <https://github.com/KWB-R/Logremoval>

5.4.3 Vergleich der Konfidenzintervalle nach Wilson und Bayes

Tabelle 8 illustriert anhand verschiedener Kombinationen für die Anzahl an Erfolgen und Misserfolgen das Verhalten der einseitigen unteren Konfidenzgrenze nach Wilson und der eines Bayes'schen Beta-Binomialansatzes. Es ist festzustellen, dass die Grenzen des Bayes'schen Intervalls etwas weiter als die des Wilson-Intervalls sind. Dies führt dazu, dass im Bayes'schen Ansatz 28 Erfolge aus 28 Versuchen notwendig sind, damit die untere Grenze des einseitigen 95%-Konfidenzintervalls oberhalb von 0,9 liegt, während für das Wilson-Intervall 25 Erfolge bei 25 Versuchen ausreichend sind. Im Allgemeinen sind die Ergebnisse jedoch sehr ähnlich.

Des Weiteren ist zu betonen, dass für die Kombination 9 Erfolge aus 10 Versuchen zwar der Punktschätzer bei 0,9 liegt, die untere Intervallgrenze beider Ansätze jedoch mit 0,65 und 0,63 weit darunter liegen. Mit Hilfe des Bayes'schen Ansatzes kann hierfür auch die Wahrscheinlichkeit quantifiziert werden, dass $r > 0,9$ liegt. Diese beträgt nur 33%, sodass die wahre Erfolgsrate mit einer Wahrscheinlichkeit von 67% darunter liegt. Im Allgemeinen ist festzustellen, dass für eine Validierung über die Erfolgsrate mindestens 25 Proben benötigt werden. Diese Anzahl ist jedoch nur für den Fall gültig, dass alle Werte die Anforderungen erfüllen. Kommt es zu zusätzlichen Misserfolgen, müssen diese Misserfolge durch weitere „Erfolge“ ausgeglichen werden. Der Zusammenhang zwischen benötigten Erfolgsmessungen in Abhängigkeit der Anzahl an Misserfolgen ist in Abbildung 10 dargestellt. Die Abbildung unterstreicht, dass jeder Misserfolg mit mehr als dem 10-fachen an Erfolgen ausgeglichen werden muss. Die potenziell hohe Anzahl an Stichproben stellt einen wesentlichen Nachteil dieser Methode dar.

Tabelle 8: Vergleich zwischen Intervallgrenzen der einseitigen Wilson und Bayes'schen Konfidenzintervalle. Grüne Schattierung: untere Konfidenzgrenze > Zielwert, Orange Schattierung: untere Konfidenzgrenze < Zielwert.

Erfolge / Gesamt (Daten)	Punktschätzer Erfolgsrate r	Einseitiges 95% : Konfidenzintervall (Wilson)	Einseitiges 95% Konfidenzintervall (Bayes) Prior: Beta(1, 1) (uniform)
3 / 3	1	[0,53; 1]	[0,47; 1] $P(r > 0.9) = 34 \%$
9 / 10	0.9	[0,65; 1]	[0,63; 1] $P(r > 0.9) = 33 \%$
18 / 18	1	[0,87; 1]	[0,85; 1] $P(r > 0.9) = 87 \%$
25 / 25	1	[0,90; 1]	[0,89; 1] $P(r > 0.9) = 93 \%$
28 / 28	1	[0,91; 1]	[0,90; 1] $P(r > 0.9) = 95 \%$

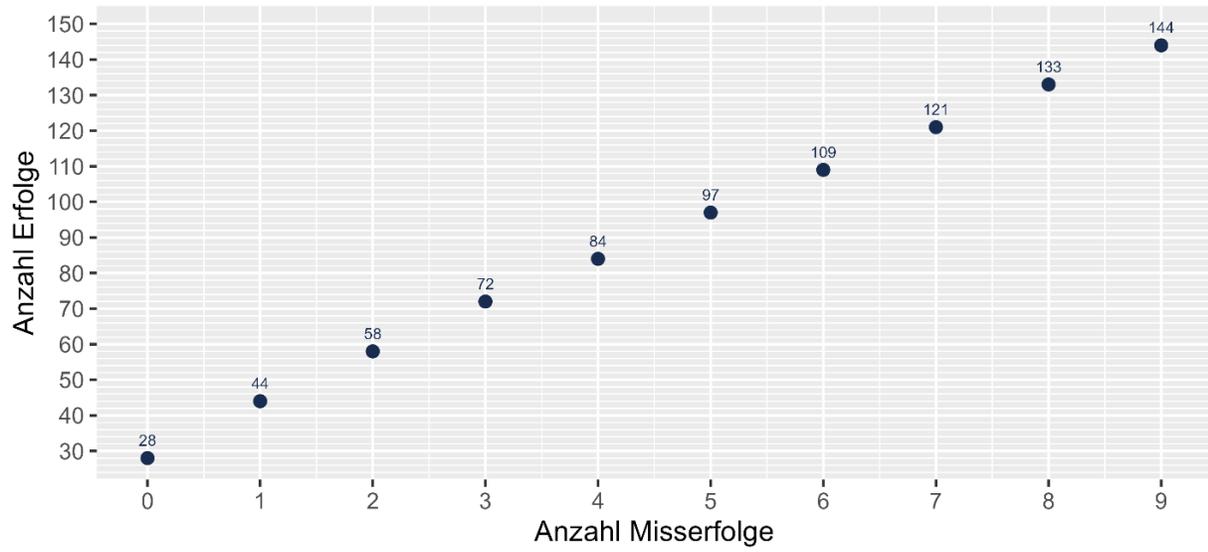


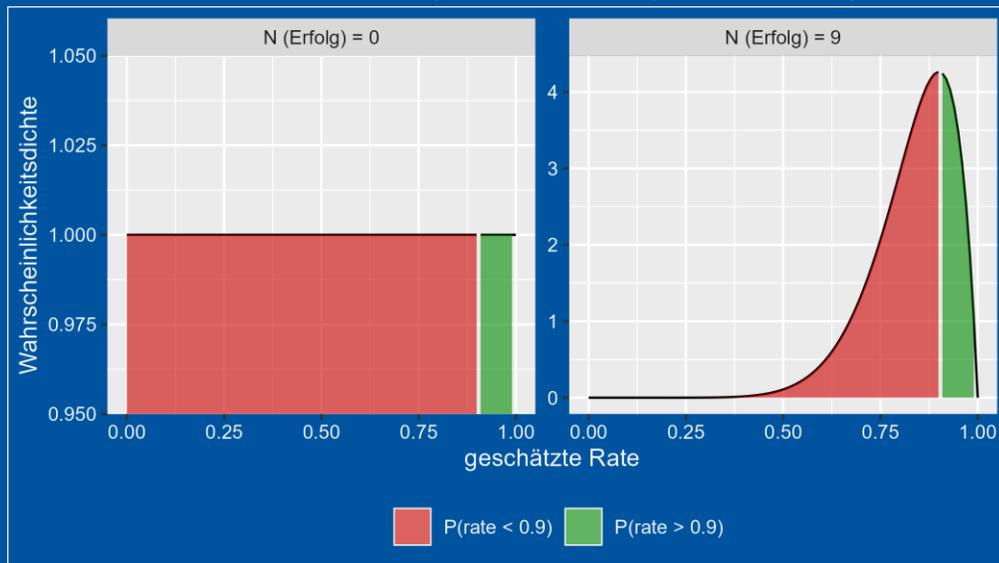
Abbildung 10: Anzahl benötigter Erfolgsmessungen in Abhängigkeit der Anzahl an Misserfolgen (Bayes'scher Beta-Binomialansatz)

Beispiel aus dem Projekt FlexTreat

Die folgende Tabelle zeigt die Zu- und Ablaufwerte für den Parameter *Clostridium perfringens*.

Zulauf [KBE/100mL]	Ablauf [KBE/100mL]	Log ₁₀ -Reduktion	Zielwert eingehalten
4.90E+07	26	6.28	Ja (Erfolg)
1.00E+06	64	4.19	Ja (Erfolg)
3.00E+06	11	5.44	Ja (Erfolg)
3.00E+06	0	6.48	Ja (Erfolg)
3.00E+06	69	4.64	Ja (Erfolg)
6.00E+06	39	5.19	Ja (Erfolg)
4600000	74	4.79	Ja (Erfolg)
1.00E+07	47	5.33	Ja (Erfolg)
1.00E+05	47	3.33	Nein (Misserfolg)
1.00E+06	31	4.51	Ja (Erfolg)

Aus den Daten wird deutlich, dass von 10 berechneten Log₁₀-Reduktionswerten, 9 die gesetzlichen Anforderungen einhalten. Zusammen entspricht dies 9 beobachteten Erfolgen bei 1 beobachteten Misserfolg. Gesucht wird die wahre zugrundeliegende Erfolgsrate r . Die Abbildung links zeigt die Apriori-Verteilung für r , während die rechte Abbildung die resultierende Unsicherheitsverteilung der wahren Erfolgsrate bei 9 Erfolgen und 1 Misserfolg zeigt.



Mit Hilfe des zugrundeliegenden Beta-Binomial Modells kann errechnet werden, dass das untere 95% Konfidenzlimit einen Wert von 0.64 annimmt, also deutlich unterhalb des geforderten Werts von 0,9 liegt. Die Wahrscheinlichkeit, dass die wahre Erfolgsrate größer ist als 0.9 beträgt nur etwa 33% (grüne Fläche). Bei 9 Erfolgen und 1 Misserfolg ist es somit wahrscheinlicher, dass der Zielwert von 90% nicht eingehalten wird, als dass er dies tut.

5.5 Berechnung von Punktschätzern und Unsicherheitsintervallen für das 10. Perzentil

Das 10. Perzentil gibt den Wert an, oberhalb dessen 90% der Werte einer Verteilung erwartet werden. Wie jeder andere Parameter einer Verteilung (z.B. der Mittelwert) ist auch das 10. Perzentil der statistischen Unsicherheit unterworfen. Für eine statistische Validierung ist es demzufolge notwendig, den Grad der statistischen Sicherheit mit anzugeben, der für die Einhaltung des Validierungsziels maßgeblich sein soll. Das Validierungsziel für die Log_{10} -Reduktion von Viren kann beispielsweise sein, dass mit einer statistischen Sicherheit von 95% das 10. Perzentil größer sein soll der Vergleichswert von 6. Für die Validierung wäre demzufolge sicherzustellen, dass das einseitige untere Konfidenzintervall des 10. Perzentils größer ist als 6.

Um das einseitige untere Konfidenzintervall des 10. Perzentils zu beschreiben, wird in der Statistik gleichgestellt das Konzept der einseitigen unteren Toleranzgrenze $TL_u(P, 1 - \alpha)$ verwendet, wobei sich P auf den Anteil der Verteilungswerte bezieht, der oberhalb der Toleranzgrenze liegen soll (für das 10. Perzentil also 90% oder 0,9) und α das Vertrauensmaß definiert.

Die untere einseitige Toleranzgrenze ist somit gleichzusetzen mit der unteren Grenze des einseitigen Konfidenzintervalls des 10. Perzentils (siehe Hahn et al. (1991), Kapitel 2.4.2).

5.5.1 Punktschätzer des parametrischen 10. Perzentil für normalverteilte Log_{10} -Reduktionswerte

Als Punktschätzer des 10. Perzentils einer Gauß'schen Normalverteilung kann folgende Gleichung herangezogen werden.

$$10. \text{ Perzentil} = MW - k SD \quad (5-8)$$

mit:

$$k = 1.282$$

MW = berechneter Mittelwert

SD = berechnete Standardabweichung

Gleichung (5-8) wird beispielsweise innerhalb der Europäischen Badegewässerrichtlinie (BG-RL) oder auch den US-amerikanischen Richtlinie für die Bewässerungswasserqualität genutzt, um die Qualität von Badegewässern und Bewässerungswasser zu bewerten.

Da Punktschätzer die statistische Parameterunsicherheit zunächst unberücksichtigt lassen (vgl. Kapitel 5.2), wird in den genannten Richtlinien eine Mindestanzahl an Proben vorgegeben, um ein Mindestmaß an statistischer Sicherheit zu gewährleisten.

In der Europäischen Badegewässerrichtlinie wird beispielsweise eine Mindestanzahl von 16 Werten herangezogen, in den USA sind 20 Proben für die Berechnung erforderlich.

Das grundsätzliche Vorgehen zur Berechnung des Perzentils besteht aus folgenden Schritten (mit 20 Stichproben als Beispiel):

1. Berechnung von 20 Log₁₀-Reduktionswerten aus 20 gepaarten Zu- und Ablaufproben
2. Berechnung des Mittelwerts (MW) und der Standardabweichung (SD) aus den Log₁₀-Reduktionswerten.
3. Berechnung des 10. Perzentils nach Formel (5-8)

5.5.2 Berechnung der einseitigen unteren Toleranzgrenze für normalerteilte Log₁₀-Reduktionswerte

Für die Validierung wird empfohlen, die untere Toleranzgrenze zu verwenden. Diese ist identisch mit der unteren Grenze eines einseitigen unteren Konfidenzintervalls des 10. Perzentils. Bei vorgegebener statistischer Sicherheit (1-α) erfolgt die Berechnung der unteren Toleranzgrenze derselben Struktur wie die Berechnung des Punktschätzers (vgl. Kapitel (5-8), mit dem Unterschied, dass der Intervallfaktor k (hier k₁) nicht konstant ist, sondern sich in Abhängigkeit von der Anzahl der Stichproben ändert. Die untere Toleranzgrenze kann wie folgt analytisch angenähert werden (Gleichungen (5-9) bis (5-12)).³

$$a = 1 - \frac{z_{\alpha}^2}{2(N-1)} \quad (5-9)$$

$$b = z_p^2 - \frac{z_{\alpha}^2}{N} \quad (5-10)$$

$$k_1 = \frac{z_p + \sqrt{z_p^2 - ab}}{a} \quad (5-11)$$

$$TI_U = MW - k_1 SD \quad (5-12)$$

Mit:

z_{α} : z-Wert des Konfidenzniveaus α (für α = 0.05 → $z_{\alpha} = 1,65$)

z_p : z-Wert des zu schätzenden Perzentils (hier P = 0,9 → $z_p = 1,282$)

N: Anzahl der Stichproben

MW: Mittelwert der log₁₀-Reduktionswerte

SD: Standardabweichung der log₁₀-Reduktionswerte

TI_u: einseitige untere Toleranzgrenze

Werte für k₁, für α = 0.95 und P = 0,9, sind für verschiedene Werte für N im Anhang II gelistet.

Softwareimplementierung: eine einfache, Excel-basierte Umsetzung ist unter folgendem Link zu finden:

<https://github.com/KWB-R/Logremoval>

³ <https://www.itl.nist.gov/div898/handbook/prc/section2/prc263.htm>

5.5.3 Vergleich von Punktschätzern und Unsicherheitsintervallen für normalerteilte Log_{10} -Reduktionswerte

Zur Illustration des Effekts der Stichprobenanzahl auf Punktschätzer, Konfidenz- und Toleranzintervalle (TI) wurden 5, 10, 20 und 50 Zahlenwerte so generiert, dass für alle Datensätze der berechnete Mittelwert (MW) exakt bei 6,5 und die berechnete Standardabweichung (SD) exakt bei 0,3 liegt. Aus jedem Datensatz wurden die folgenden Schätzer und Intervalle berechnet:

1. Punktschätzer analog zur europäischen Badegewässerrichtlinie (blau gestrichelte, vertikale Linie)
2. Konfidenzintervall des 10. Perzentils (violettes Intervall)
3. Untere und obere Toleranzgrenze mit $P = 0,9$ und $\alpha = 0,05$ (graues Toleranzintervall)
4. Untere und obere Toleranzgrenze mit $P = 0,9$ und $\alpha = 0,5$ (schwarzes Toleranzintervall)

Als Zielwert für die Validierung wurde die Log_{10} -Reduktion für Viren von 6 eingezeichnet. Aus der Abbildung werden folgende Aspekte deutlich. Die untere einseitige Toleranzgrenze für $\alpha = 0,05$ (Statistische Sicherheit = 95%) entspricht dem einseitigen unteren Konfidenzintervall für das 10. Perzentil. Die untere einseitige Toleranzgrenze mit $\alpha = 0,5$ (Statistische Sicherheit = 50%) entspricht dem Punktschätzer. Die berechneten Punktschätzer sind für alle Datensätze identisch, da die statistische Unsicherheit unberücksichtigt bleibt. Die Breite der Toleranzintervalle mit $\alpha \neq 0,5$, sowie die der Konfidenzintervalle nehmen mit Anzahl der Stichproben ab. So wäre im gegebenen Beispiel eine Stichprobenanzahl von $N > 10$ notwendig, um das System zu validieren, da zuvor die untere Toleranzgrenze noch unterhalb des Zielwerts liegt.

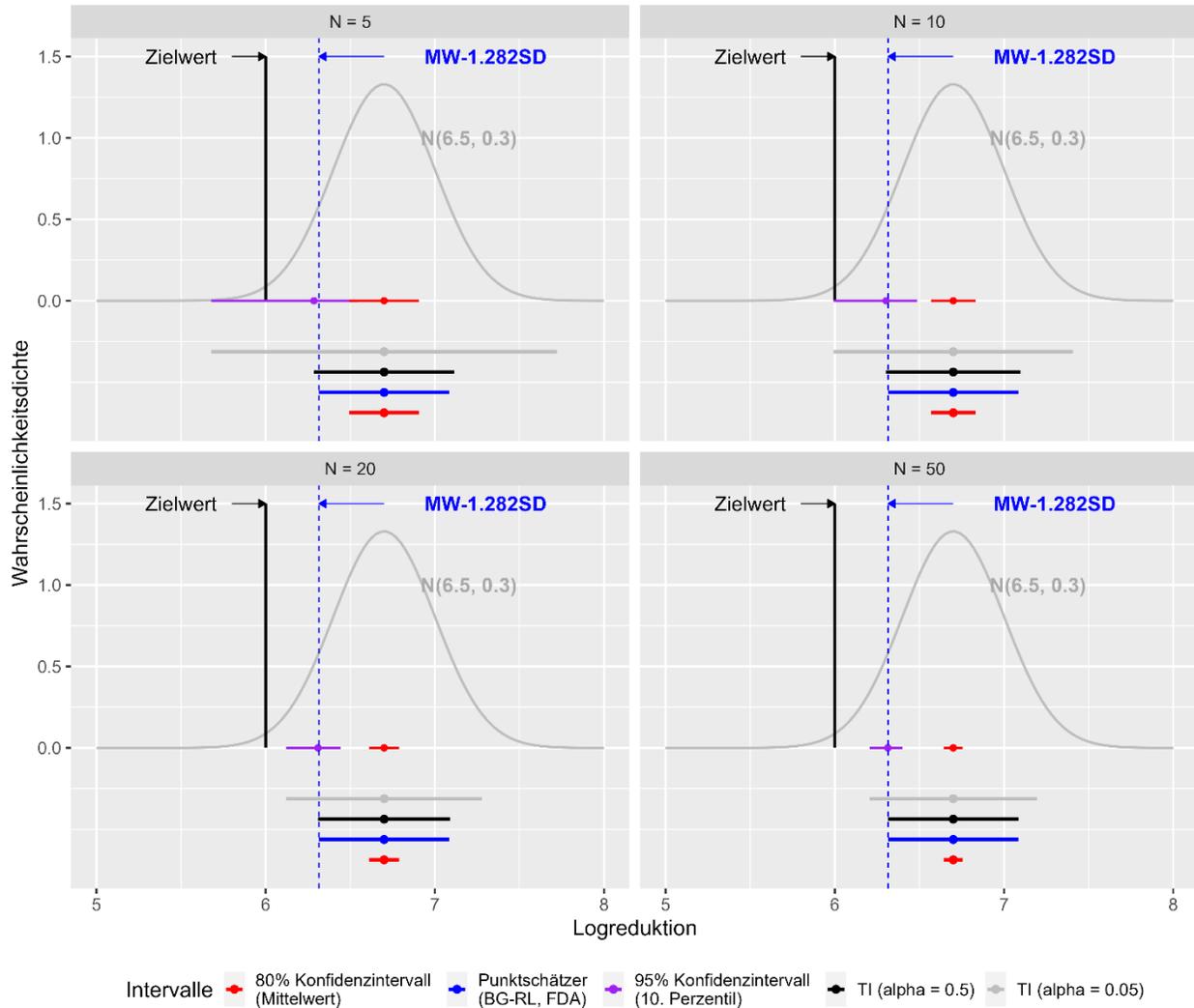


Abbildung 11: Vergleich von Punktschätzern und Unsicherheitsintervallen von normalverteilten Werten.

5.5.4 Ansätze nicht-normalverteilte Log_{10} -Reduktionswerte

Eine Validierung auf Basis von Toleranzgrenzen normalverteilter Log_{10} -Reduktionswerte kann in Zweifel gezogen werden, falls die berechneten Log_{10} -Reduktionswerte der Annahme einer Normalverteilung nicht standhalten. Letzteres kann durch statistische Testverfahren wie beispielsweise dem Shapiro-Wilk Test geprüft werden.

Nicht-normalverteilte Werte sind insbesondere dann zu erwarten, wenn ein bestimmter Prozentsatz der Ablaufwerte kleiner als die Bestimmungsgrenze von 1, also 0, sein sollte. Dies ist damit zu begründen, dass sich mikrobiologische Daten meist gut über eine Lognormalverteilung abbilden lassen. Lineare Transformationen von Normalverteilungen sowie Verrechnungen mit Konstanten sind ebenfalls normalverteilt. Liegen im Ablauf Messwerte jedoch unterhalb der Bestimmungsgrenze, ist die resultierende Verteilung weder symmetrisch noch konstant. Daraus berechnete Log_{10} -Reduktionswerte

sind demzufolge ebenfalls nicht mehr zwangsläufig normalverteilt. Tabelle 9 fasst die verschiedenen Kombinationen aus zu- und Ablaufverteilungen zusammen.

Tabelle 9: Mögliche Kombinationen aus Zu- und Ablaufwerten, sowie deren Effekt auf die resultierende Verteilung der Log_{10} -Reduktionswerte

Zulaufwerte	Ablaufwerte	Resultierende Verteilung der Log_{10} -Reduktionswerte
Lognormal	Lognormal	normal
Lognormal	Alle Werte < BG* = konst.	normal
Lognormal	1-99% der Werte < BG*	voraus. nicht-normal

*BG: Bestimmungsgrenze

Bei nicht-normalverteilten Log_{10} -Reduktionswerten können für die Bestimmung der unteren Toleranzgrenze alternative Auswertungsvarianten in Betracht kommen (vgl. Kapitel 5.5.4.1 und 5.5.4.2). Darüber hinaus kann eine Auswertung über die Nutzung der Erfolgsrate r in Betracht kommen.

5.5.4.1 Bayes'sches Toleranzintervall auf Basis ungepaarter Auswertung und Markov-Ketten-Monte-Carlo (MCMC) Schätzverfahren

Dieses Schätzverfahren ist das flexibelste der hier dargestellten Verfahren, da es sowohl bei einer beliebigen Zahl von Ablaufwerten <BG angewendet werden kann, als auch bei einer ungleichen Zahl von Werten im Zu- und Ablauf des Klärwerks. Falls beispielsweise die Betriebsparameter einer Aufbereitungsanlage nochmals angepasst werden müssen, da mit einer gegebenen Konfiguration die Validierungsziele noch nicht eingehalten werden konnten, können im weiteren Verlauf die bisher gewonnenen Zulaufwerte weiter verwendet werden.

Wesentlicher Nachteil des Verfahrens ist die vergleichsweise komplizierte Implementierung, da diese gewisse statistische und programmiertechnische Kompetenzen erfordert. Eine Implementierung in der Programmiersprache R wird öffentlich bereitgestellt (<https://github.com/KWB-R/Logremoval>), sodass der Nutzer prinzipiell nur die erhobenen Messdaten in eine Excel-Datei eingeben muss. Eine einfache, Excel-basierte Lösung steht für diese Methode nicht zur Verfügung. Im Folgenden werden die wesentlichen Schritte, und Zwischenschritte erläutert. Der Simulationsansatz folgt im Wesentlichen dem Ansatz, der bei Stoudt et al 2021⁴ für normalverteilte Werte beschrieben ist, jedoch für Ablaufwerte angepasst wurde die einer negativen Binomialverteilung folgen. Durch die negative Binomialverteilung können diskrete, überdispertierte Werte, einschließlich der 0 als Messwert beschreiben werden. Sie ist somit geeignet für mikrobiologische Werte im Ablauf, die diese Charakteristika aufweisen.

⁴ <https://nvlpubs.nist.gov/nistpubs/jres/126/jres.126.004.pdf>

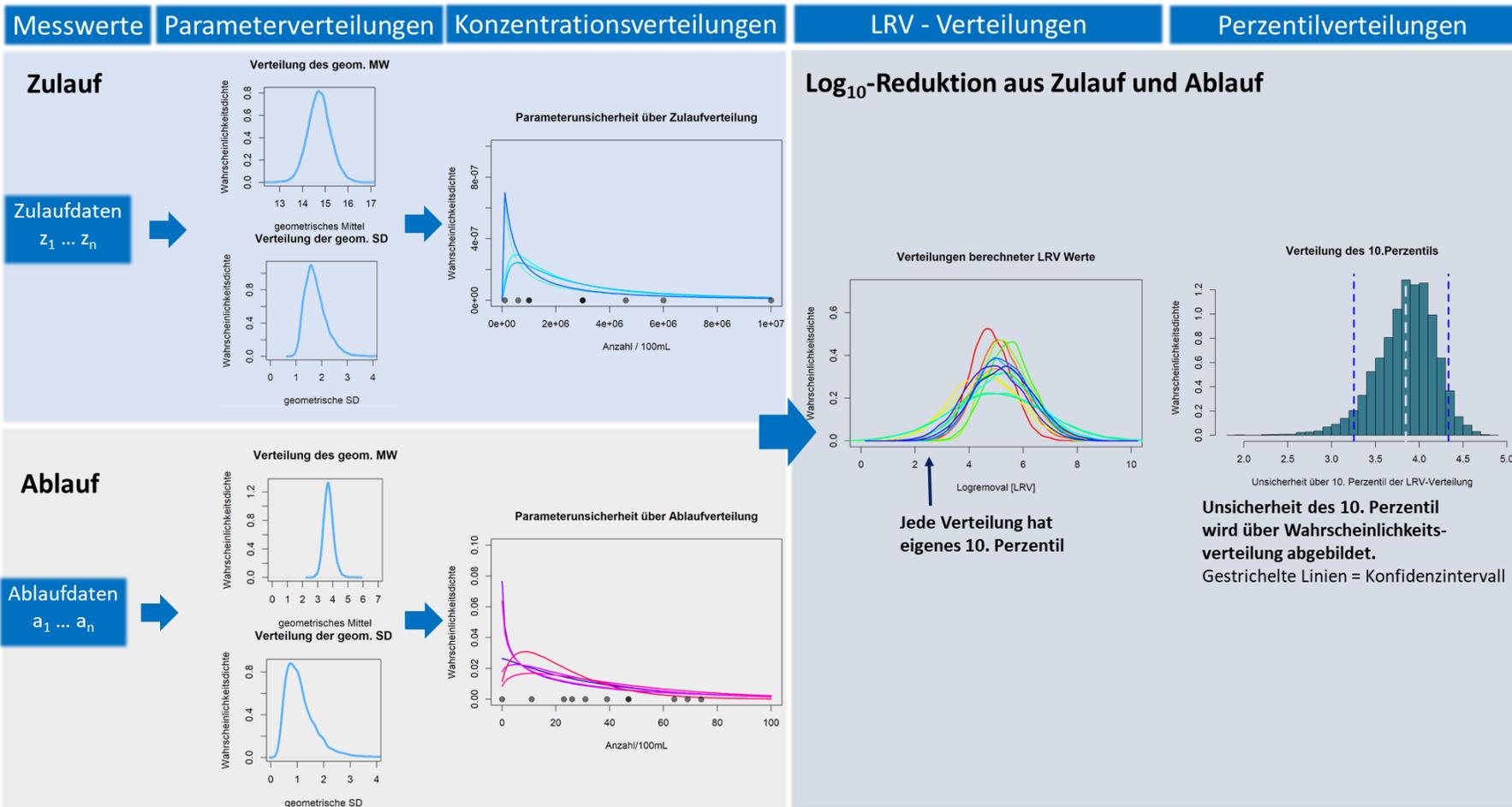


Abbildung 12: Illustration des MCMC Monte Carlo Verfahrens zur Bestimmung Bayes'scher Toleranzintervalle

Tabelle 10: Zusammenfassung der wesentlichen Schritte des MCMC Monte-Carlo Verfahrene

Schritt	Beschreibung	Input	Output	R Code
1a.	Fitten einer Lognormalverteilung auf die gewonnenen Zulaufwerte (mind. 10000 Iterationen) über Markov-Chain Monte Carlo (MCMC)	Zulaufdaten des Validierungsindikators	Verteilung des Mittelwerts und Verteilung der Standardabweichung, MCMC Samples	https://github.com/KWB-R/Logremoval
1b.	Fitten einer negativen Binomialverteilung auf die gewonnenen Ablaufwerte (mind. 10000 Iterationen) über Markov-Chain Monte Carlo (MCMC)	Ablaufdaten des Validierungsindikators	Verteilungen des Lage- und Streuungsparameters, MCMC Samples	
2	Konstruktion von 10000 Zulauf und Ablaufverteilungen aus korrelierten Parameterverteilungen	Parameterverteilungen (MCMC Samples) der Parameter Zulauf und Ablaufverteilungen	10000 Zulaufverteilungen + 10000 Ablaufverteilungen	
3	Simulation von 10000 Log ₁₀ -Reduktionsverteilungen	10000 Zulaufverteilungen 10000 Ablaufverteilungen	10000 Simulierte Verteilungen für die Log ₁₀ -Reduktion	
4	Extraktion des 10. Perzentils	10000 Simulierte Verteilungen für die Log ₁₀ -Reduktion	Verteilung des 10. Perzentils (10000 iterationen)	

5.5.4.2 Berechnung von nicht-parametrischen Toleranzintervallen für das 10. Perzentil

Nicht-parametrische Ansätze zur Bestimmung von Toleranzintervallen unterscheiden sich deutlich von parametrischen Ansätzen. Sie sind teilweise verwandt mit Ansätzen die als Zielparameter die Erfolgsrate nutzen vor allem Hinsichtlich der Bestimmung des Konfidenzniveaus ($1-\alpha$), die ebenfalls auf einer Binomialverteilung beruhen.

Nicht-parametrische Ansätze machen keine Annahmen über die Verteilung der erhobenen Messgrößen, sondern sortieren berechnete Werte aufsteigend nach Größe und bestimmen den statistischen Rang, also die Ordnungszahl, desjenigen Wertes, der am ehesten dem gesuchten Perzentil entspricht. Auf diese Art sind sie unabhängig von Annahmen über die Verteilung von Messdaten. Wie bei parametrischen Verfahren existiert das grundsätzliche Problem, dass auf Basis von endlich vielen Beobachtungen der Messgröße,

verallgemeinerbare Aussagen über die Grundgesamtheit getroffen werden sollen. Auch hier ist es intuitiv leicht nachvollziehbar, dass der Grad der statistischen Sicherheit zunimmt, je mehr Messwerte vorliegen.

Im Gegensatz zu parametrischen Verfahren bei denen bei kleinen Stichproben, die Spannweite eines Intervalls zunimmt, um Aussagen mit einem vorgegebenen Konfidenzniveau $(1-\alpha)$ zu treffen, ist die Spannweite eines nicht-parametrischen Intervalls durch den kleinsten und größten Beobachtungswert limitiert. Die Frage ist, wie sicher man sich sein kann, dass der kleinste gemessene Wert einem 10. Perzentil entspricht?

Die Anzahl an Messwerten die benötigt wird, damit der kleinste Beobachtungswert mit einer vorgegebenen statistischen Sicherheit $(1-\alpha)$, einem 10. Perzentil entspricht, also ein bestimmter Anteil der Grundgesamtheit P (hier 0,9) größer ist als dieser Wert, und somit eine untere Grenze eines einseitigen unteren Toleranzintervalls darstellt, kann nach Krishnamoorthy and Mathew (2009) über folgende Formel abgeschätzt werden.

$$n = \frac{\ln(\alpha)}{\ln(P)} \quad (5-13)$$

Für $\alpha = 0.05$ und $P = 0,9$ ergibt sich ein Wert von $n = 28,4$. Es sind somit mindestens 29 Werte notwendig damit der kleinste beobachtete Wert, als eine einseitige, untere Toleranzgrenze angesehen werden kann, oberhalb derer 90% der Werte, mit einer statistischen Sicherheit von 95%, zu erwarten sind.

Für die Praxis würde dies bedeuten, dass bei 29 Messwerten alle Werte oberhalb des Zielwerts liegen müssen, damit das so ermittelte 10. Perzentil ebenfalls oberhalb des Zielwerts liegt. Ebenso wie bei der Schätzung der Erfolgsrate liegt ein wesentlicher Nachteil von nicht-parametrischen Verfahren in der vergleichsweise hohen Anzahl an benötigten Stichproben.

Da die Ergebnisse im Wesentlichen mit denen auf Basis der Erfolgsrate vergleichbar sind, wird auf eine detailliertere Darstellung von nicht-parametrischen Toleranzintervallen verzichtet.

Tabelle 11: Anzahl der benötigten Stichproben, bei vorgegebener statistischer Sicherheit, damit kleinster Beobachtungswert einer Schätzung des 10. Perzentils entspricht.

Grad der statistischen Sicherheit $(1-\alpha)$	Stichproben
80%	16
90%	22
95%	29
99%	44

5.6 Anwendungsbeispiel

Im folgenden Kapitel werden die verschiedenen Auswertungsansätze anhand realer Daten vom Standort Braunschweig verglichen und gegenübergestellt. Ziel dieses Vergleichs ist nicht die abschließende Bewertung oder Validierung der Anlage, sondern die Illustration, wie sich verschiedene Auswertungsansätze hinsichtlich der Anlagenbewertung verhalten. Für die beispielhafte Darstellung werden zwei unterschiedliche Versuchsphasen genutzt, um verschiedene Aspekte zu illustrieren. Die erste

Versuchsphase nutzt dabei die Daten, die während des Pilotbetriebs gewonnen wurden. In dieser Phase fand noch keine Anreicherung von Coliphagen (vgl. Kapitel 4.6) statt, jedoch wurden zeitlich gepaarten Proben im Zulauf und Ablauf der Aufbereitungsanlage genommen, sodass Auswertungsalternativen die auf gepaarten Werten beruhen verglichen werden können. Die hier beschriebene Versuchsphase beschreibt den Effekt der Anreicherungsversuche (vgl. Kapitel 4.6.4) auf die Validierung der Anlage hinsichtlich der Entfernungsleistung viraler Indikatoren am Beispiel somatischer Coliphagen.

5.6.1 Beschreibung des Datensatzes und zugrundeliegende Annahmen (Versuchsphase 1)

Während des Projekts FlexTreat wurden Zulauf- und Ablaufdaten auf Basis von 24 h - Mischproben für die Parameter *E. coli*, *Clostridium perfringens* und somatische Coliphagen gesammelt. Eine Anreicherung fand zunächst nicht statt. Während der Versuchsphase wurden verschiedene Betriebszustände untersucht, die sich durch die Steuerungsgröße der Ozonanlage unterschieden ($\Delta\text{SAK} = 34\%$, $\Delta\text{SAK} = 47\%$). Für jeden Indikator wurden verschiedene Validierungsansätze durchgeführt.

1. Gepaarte Auswertung

Für die paarweise Auswertung werden diejenigen Proben paarweise ausgewertet, die am selben Tag im Zulauf und Ablauf genommen wurden. Unvollständige Datensätze werden entfernt. Die Auswertung beschränkt sich auf den als für die Desinfektion günstiger angesehenen Betriebszustand von $\Delta\text{SAK} = 47\%$. Obwohl die Zulaufwerte des Klärwerks unabhängig vom Betriebszustand sind, werden in der paarweisen Auswertung die Zulaufwerte des Betriebszustands $\Delta\text{SAK} = 34\%$ verworfen, da kein vergleichbarer Ablaufwert existiert.

Für jedes Wertepaar wurde die Log_{10} -Reduktion nach Kapitel 5.1 berechnet. Die so ermittelten Log_{10} -Reduktionswerte werden wie folgt ausgewertet:

1. Berechnung der Erfolgsrate r und der dazugehörigen Parameterunsicherheit mit Hilfe des Bayes'schen Verfahrens (Beta-Binomial nach Kapitel 5.4.2)
2. Berechnung der einseitigen unteren Toleranzgrenze mit $P = 0,9$ und $\alpha = 0,05$ (vgl. Kapitel 5.5.2)
3. Berechnung des Punktschätzers für das 10. Perzentil nach Kapitel 5.5.1

Für den Binomialansatz wurde zudem die Wahrscheinlichkeit berechnet mit dem die wahre Erfolgsrate größer ist als 0,9. Für die untere Toleranzgrenze als auch das 10. Perzentil wurde für eine verbesserte Visualisierung zusätzlich das 90. Perzentil, als auch die obere Toleranzgrenze berechnet.

Für den Betriebszustand $\Delta\text{SAK} = 47\%$ standen insgesamt 14 Wertepaare zur Verfügung. Um die Abhängigkeit der berechneten Unsicherheitsintervall von der Stichprobenanzahl N zu illustrieren, wurden alle Auswertungen chronologisch aufsteigend für $N = 3$ bis $N = 14$ durchgeführt.

2. Ungepaarte Auswertung

Zusätzlich zur gepaarten Auswertung wurden die unteren und oberen Toleranzgrenzen mittels einer Bayes'schen Auswertung nach Kapitel 5.6.1 ermittelt. In einem ersten Ansatz wurde sich, wie für die gepaarte Auswertung, ebenfalls auf den Betriebszustand $\Delta\text{SAK} = 47\%$ beschränkt. In einem zweiten Ansatz (alle Werte) wurde für die Bestimmung der Verteilung der Zulaufwerte die Messdaten der Versuchsreihen mit $\Delta\text{SAK} = 34\%$ mit hinzugezogen. Auf diese Weise stehen im Zulauf 10 zusätzliche Messwerte zur Verfügung. Wie für die gepaarte Auswertung wurde die Auswertung für $N = 3$ bis $N = 14$ durchgeführt. Für

die Auswertung in der alle Werte genutzt werden startet die Auswertung mit 13 Zulauf und 3 Ablaufwerten und steigert sich zu 24 Zulauf und 14 Ablaufwerten.

Vergleich Auswertungsansätze

Untere Intervallgrenze: 10. Perzentil, obere Intervallgrenze: 90. Perzentil

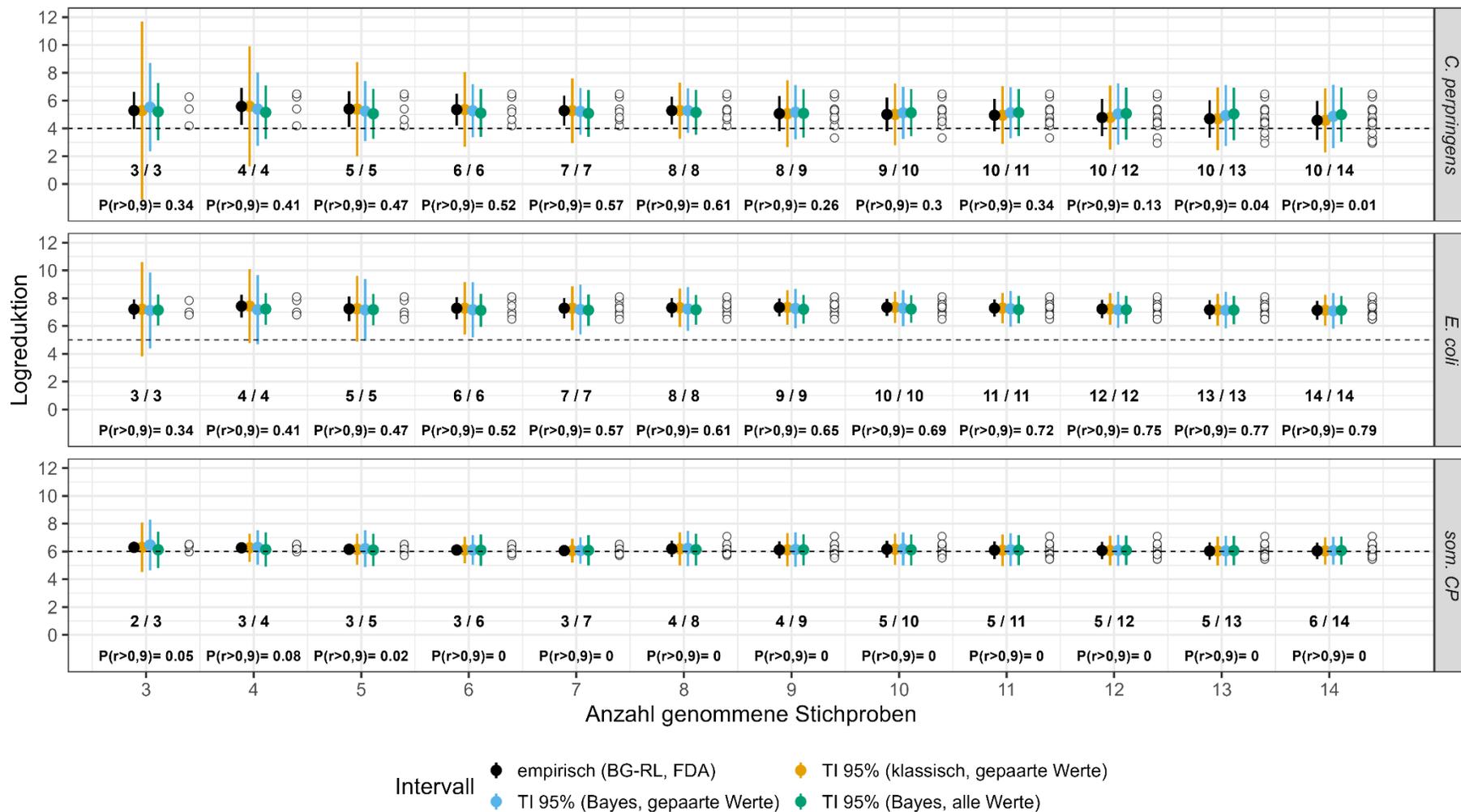


Abbildung 13 Vergleich verschiedener Auswertungsansätze für Daten aus dem Projekt FlexTreat. Die Unsicherheitsintervalle beschreiben das 10. und 90. Perzentil (schwarzes Intervall), bzw. die einseitigen unteren und oberen Toleranzgrenzen mit $\alpha = 0,05$. Zahlenwerte X/Y illustrieren den Binomialansatz mit X = Anzahl der Erfolge und Y = Gesamtanzahl. Werte die aufgrund von zu niedrigen Zulaufkonzentrationen mit „> WERT“ unterhalb des Zielwerts liegend wurden als Misserfolg gewertet (relevant für somatische Coliphagen – som. CP).

5.6.2 Ergebnisse (Versuchsphase 1)

Die Ergebnisse der verschiedenen Auswertungsvarianten sind in Abbildung 13 dargestellt. Auffällig ist zunächst, dass die empirischen Unsicherheitsintervalle (analog zum Ansatz der Badegewässerrichtlinie) bereits bei einer geringen Anzahl an Stichproben von $N = 3$ sehr schmal sind. Dies ist darauf zurückzuführen, dass die Parameterunsicherheit bei dieser Auswertungsvariante nicht berücksichtigt wird. Bei $N = 3$ liegt die untere Grenze des Unsicherheitsintervalls für *E. coli* und *C. perfringens* oberhalb der Validierungsgrenze. Für somatische Coliphagen liegt die untere Grenze bei 5,92, was der Validierungsgrenze bereits sehr nahekommt.

Für *C. perfringens* wird die Validierungsgrenze bis zu einem Wert von $N = 8$ eingehalten, danach nicht mehr. Dies bedeutet, dass sich die Validierung von „positiv“ nach „negativ“ ändert. Dies zeigt, dass für die empirische Auswertungsvariante ein Minimum an Stichproben festzulegen ist, da eine Bewertung aufgrund von zu enger Intervallgrenzen zu optimistisch ausfallen kann.

Als weiteres Indiz dafür, dass die Intervallgrenzen der empirischen Variante zu eng bemessen sind, stellt der Anteil an tatsächlichen Messwerten dar, die durch das Intervall abgedeckt werden. Nimmt man die berechneten \log_{10} Reduktionswerte für $N = 14$ und vergleicht, wie viele dieser Daten oberhalb der unteren Intervallgrenze bei $N = 3$ liegen, ist festzustellen, dass im Beispiel von somatischen Coliphagen und *E. coli* nur ca. 50% bzw. 80% größer sind als die untere Intervallgrenze. Da es sich bei der unteren Intervallgrenze um das 10. Perzentil handeln sollte, müssten eigentlich 90% oberhalb dieser Grenze liegen. Das Intervall ist folglich zu eng, um die tatsächlich gemessenen Werte abzubilden.

Im Gegensatz dazu sind die berechneten Toleranzgrenzen von Anfang an ($N = 3$) weit genug, um die tatsächlich gemessenen Daten abzudecken. Für somatischen Coliphagen und *E. coli* liegen stets 100% der gemessenen Werte ($N = 14$) oberhalb der unteren Grenze, unabhängig der Anzahl an gepaarten Werten. Für *C. perfringens* liegen stets mindestens 90% oberhalb der unteren Toleranzgrenze für das Bayes'sche Verfahren, und 100% oberhalb der unteren Toleranzgrenze für das klassische Verfahren. Hinsichtlich der Gesamtbewertung bleiben die berechneten Toleranzintervalle von $N = 3$ bis $N = 14$ konsistent in ihrer negativen Bewertung für *C. perfringens*. Für *C. perfringens* basiert diese Bewertung auf tatsächlich gemessenen Zu- und Ablaufkonzentrationen und ist somit aussagekräftig für die betriebene Pilotanlage im beobachteten Betriebszustand.

Für somatische Coliphagen liegen die berechneten unteren Toleranzgrenzen ebenfalls unterhalb des Zielwerts einer \log_{10} -Reduktion von 6. Im Fall der somatischen Coliphagen ist dies jedoch auf zu niedrige Zulaufkonzentrationen zurückzuführen. Für den Betriebszustand $\Delta\text{SAK} = 47\%$ lagen die Ablaufkonzentrationen stets unterhalb der Bestimmungsgrenze von <1 PBE / 100mL. Die ermittelten unteren Toleranzgrenzen sind somit nicht aussagekräftig, und müssen zur besseren Quantifizierung durch Anreicherungsversuche ermittelt werden (vgl. Kapitel 4.6 und Kapitel 5.6.3). Zum Zeitpunkt der Planungsphase und Durchführung dieses Hauptmessprogramms, war das Problem zu niedriger Zulaufkonzentrationen für Coliphagen im Zulauf der meisten Klärwerke noch nicht bekannt, sodass innerhalb dieses Messprogramms keine Anreicherung größerer Probenvolumina stattfand. Aus diesem Grund weisen ca. 40-50% der berechneten \log_{10} -Reduktionswerte Werte von $> \text{WERT}$ auf, wobei der numerische $\text{WERT} < 6$ ist (Beispiel $> 5,6$).

Für den Parameter *E. coli* zeigt sich, dass auch unter Berücksichtigung der statistischen Unsicherheit eine Validierung mit deutlich weniger als 10 Stichproben möglich ist. Dies ist möglich, da die beobachtete Log_{10} -Reduktion im Mittel um zwei Log_{10} -Stufen höher liegt, als der zu validierende Zielwert. Dabei ist zu unterstreichen, dass es sich bei zwei log_{10} -Stufen, um einen Faktor von 100 handelt, mit dem der zu validierende Zielwert übererfüllt wird. Zwar ist eine genaue Aussage über den „wahren“ Wert immer noch mit großen Unsicherheiten behaftet, die Aussage, dass 90% der Messwerte oberhalb des deutlich niedriger liegende Validierungswerts liegen, kann jedoch dennoch getroffen werden. Die Auswertung von *E. coli* zeigt, dass eine Validierung umso einfacher zu erreichen ist, je größer die Differenz zwischen dem zu validierendem Wert und den erhobenen Messwerten ist. Aus diesem Grund kann es beispielsweise sinnvoll sein, durch eine Anreicherung großer Probenvolumina (vgl. Kapitel 4.6) die beobachtbare Differenz zwischen Zulauf und Ablauf zu vergrößern. Dies gilt insbesondere dann, wenn aufgrund der eingesetzten Technologien davon ausgegangen werden kann, dass die reale Log_{10} -Reduktion deutlich oberhalb des zu validierenden Zielwerts liegt.

Wie gezeigt, ist die Differenz zwischen den analytisch ermittelten Log_{10} -Reduktionswerten maßgeblich für die Anzahl an notwendigen Stichproben, da je größer die Differenz ist, desto mehr statistische Unsicherheit toleriert werden kann, und desto weniger Stichproben benötigt werden. Vor diesem Hintergrund ist es folgerichtig, dass Methoden, die diese Information unberücksichtigt lassen, eine wichtige Zusatzinformation ignorieren, welche wiederum durch eine höhere Anzahl an Stichproben kompensiert werden muss.

So zeigt die Auswertung auf Basis der Erfolgsrate r (Binomialansatz), dass im Beispiel von *E. coli* selbst bei 14 Erfolgen aus 14 Versuchen, immer noch mit einer Wahrscheinlichkeit von 21% das Validierungsziel nicht erreicht wird, und die Validierung mit Hilfe eines Binomialansatzes noch nicht erfolgreich ist. Zur Erinnerung, bei der Ermittlung der Erfolgsrate (vgl. Kapitel 5.4) wird kein Unterschied gemacht, ob das Validierungsziel um 0,1 Log_{10} -Stufen oder 2 Log_{10} -Stufen übererfüllt wurde. In beiden Fällen geht der ermittelte Wert als „Zielwert eingehalten“ bzw. „Erfolg“ in die statistische Auswertung ein.

5.6.3 Anreicherungsversuche (Versuchsphase 2)

Versuchsphase 2 illustriert den Effekt der Probenanreicherung auf die Validierung der Anlage hinsichtlich der Reduktion von viralen Indikatoren. Die Probenanreicherung sowie die gewonnen Messwerte sind im Detail in Kapitel 4.6.4 beschrieben, insbesondere in Tabelle 5. Tabelle 5 zeigt, dass es sich bei 5 von insgesamt 7 genommenen und angereicherten Proben um Negativbefunde handelte. In den beiden übrigen Proben wurden in 100 mL Eluat 12 bzw. 27 PBE an somatischen Coliphagen nachgewiesen. Das mittlere repräsentative Ablaufvolumen betrug 2020 mL, sodass die gemessenen Zulaufkonzentration pro 100 mL mit einem Faktor von 20,2 multipliziert wurde, um die Zulaufwerte und Ablaufwert auf das gleiche Volumen zu beziehen (vgl. Abbildung 9).

Da für die Anreicherungsversuche keine zeitlich gepaarten Zulaufwerte zu Verfügung stehen, und zudem aufgrund der vergleichsweise niedrigen Stichprobenanzahl von ($N = 7$), selbst für den Fall dass solche Werte zur Verfügung stünden, nur wenige gepaarte Werte vorlägen, wird für die Auswerten der ungepaarte Bayes'sche Auswertungsansatz zur Berechnung von Toleranzgrenzen verwendet. Da während des Pilotbetriebs im Zulauf zudem ausschließlich somatische Coliphagen gemessen wurden, wird sich auf diesen Parameter beschränkt. Zu Vergleichszwecken wird neben der Auswertung auf Basis realer Messwerte, zudem ein fiktives Szenario betrachtet, in dem angenommen wird, dass es sich nicht nur bei

5, sondern bei allen 7 genommenen Messwerten um Negativbefunde gehandelt hätte. Darüber hinaus wird in der Ergebnisdarstellung zusätzlich die empirische Auswertungsvariante nach Badegewässerrichtlinie ergänzt. Dargestellt werden ausschließlich die Punktschätzer bzw. Punktschätzer + 95 % Konfidenzintervalle für das 10. Perzentil.

Für die Bestimmung der Zulaufverteilung wird der komplette Datensatz für die beiden Betriebszustände bei $\Delta\text{SAK} = 34\%$ und $\Delta\text{SAK} = 47\%$ verwendet, sodass für den Zulauf 24 Proben zur Verfügung stehen. Für die Zulaufverteilung wurde eine Lognormalverteilung gewählt. Wie in 5.5.4.1 beschrieben wurde, wird für Ablaufverteilung eine negative Binomialverteilung gewählt, da diese für ganzzahlige, überdisperse Werte geeignet ist, und die 0 als Messwert einschließt.

5.6.4 Ergebnisse (Versuchsphase 2)

Die Ergebnisse der Validierung auf Basis der Anreicherungsdaten und Szenarien sind in Abbildung 15 dargestellt. Die resultierenden Ablaufverteilungen der beiden Anreicherungszenarien sind in Abbildung 14 dargestellt. Die Ergebnisse zeigen, dass für das Szenario in dem angenommen wurde, dass es sich bei 7 von 7 Proben um Negativbefunde gehandelt hätte, eine Validierung quantitativ möglich gewesen wäre, da das komplette Konfidenzintervall, und somit auch die untere Toleranzgrenze oberhalb des zu validierenden Zielwerts liegt. Dies unterstreicht die generelle Eignung der Anreicherungs- und Auswertungsmethode. Den Effekt der beiden realen Positivnachweise auf das Validierungsergebnis ist ebenfalls dargestellt. Es wird deutlich das aufgrund der beiden Positivnachweise, sich die Ablaufverteilung deutlich verbreitert und somit sowohl der Punktschätzer des 10. Perzentils als auch die untere Toleranzgrenze unterhalb des Validierungsziels liegen. Die Anlage wäre folglich noch nicht validiert. Zu bemerken ist jedoch das, im vorliegenden Fall die Validierung quantitativ erfolgte. Im Vergleich zum empirischen Ansatz der Badegewässerrichtlinie bei dem die Quantifizierung der Log_{10} -Reduktion durch die niedrigen Zulaufkonzentrationen limitiert ist, zeigen die beiden Anreicherungszenarien somit, dass eine quantitative Validierung nun zumindest theoretisch möglich ist.

Ablaufverteilung somatischer Coliphagen nach Anreicherung

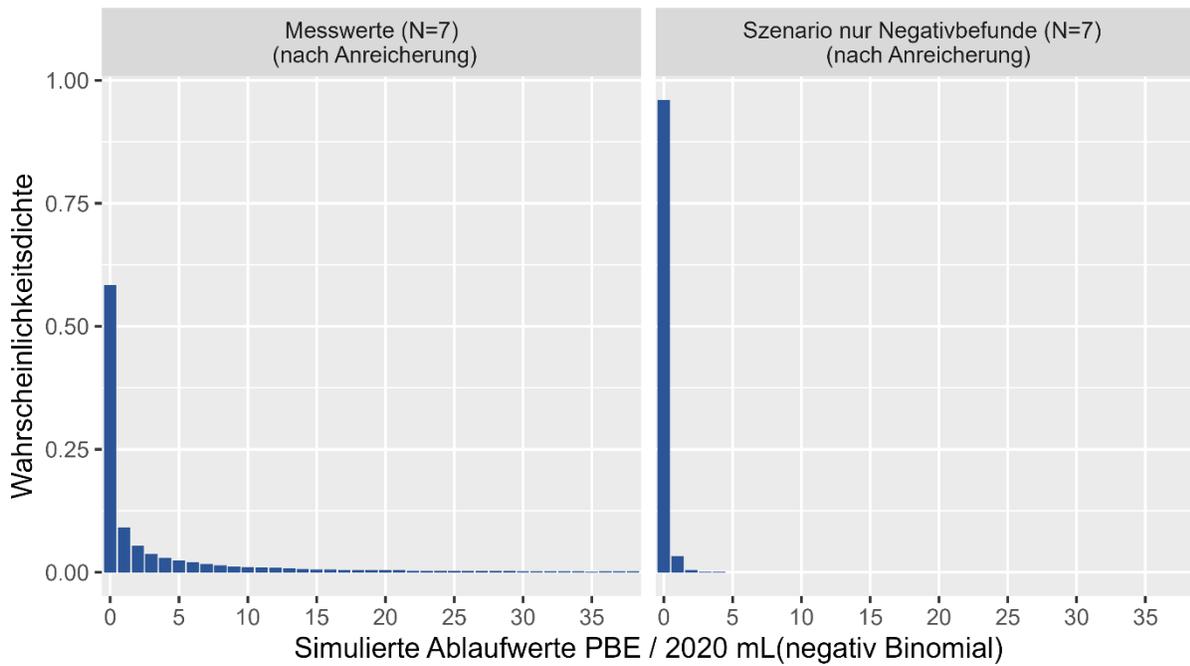


Abbildung 14: Simulierte Ablaufverteilung auf Basis realer Messwerte, sowie unter der Annahme von ausschließlichen Negativbefunden im Ablauf der Anlage

10. Perzentil der Log_{10} -Reduktionsverteilung vor und nach Anreicherung

Auswertung bezieht sich auf Reduktion von somatischen Coliphagen

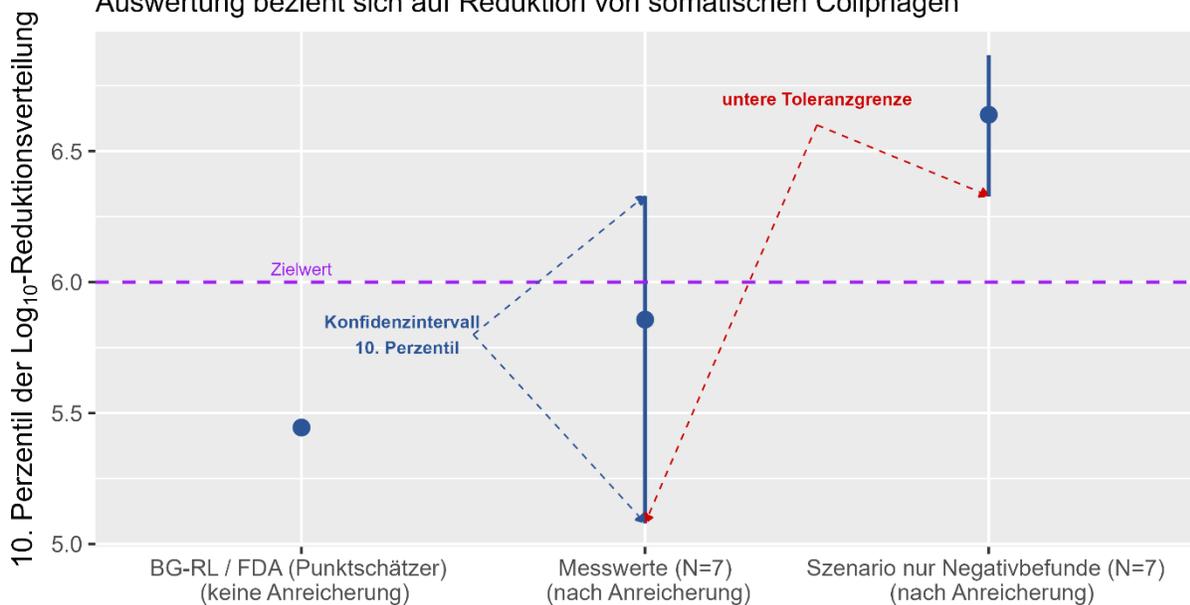


Abbildung 15: Darstellung des Punktschätzers des 10. Perzentils auf Basis der empirischen Methode nach Badegewässerrichtlinie auf Basis nicht-angereicherter, gepaarter Werte für den Betriebszustand $\Delta\text{SAK} = 47\%$, und den berechneten Konfidenzintervallen für das 10. Perzentil auf Basis angereicherter Proben, wobei die untere Grenze des Konfidenzintervalls der unteren Toleranzgrenze entspricht.

5.7 Zusammenfassung, Diskussion und Handlungsempfehlungen

5.7.1 Zielparameter

In den vorangegangenen Kapiteln wurden verschiedene Auswertungsalternativen aufgezeigt. Es wurde gezeigt, dass sich für eine Validierung sowohl das 10. Perzentil der Verteilung der Log_{10} -Reduktionswerte, als auch die Erfolgsrate r , der erfolgreich durchgeführten Validierungsmessungen als Zielparameter eignen. Das 10. Perzentil wird grundsätzlich als der geeignetere Zielparameter angesehen, da es eine bessere Beschreibung der vorhandenen Information darstellt, als eine berechnete Erfolgsrate. Für die Berechnung der Erfolgsrate macht es keinen Unterschied, ob in einem Validierungsversuch eine Log_{10} -Reduktion von 6,1 oder 7,1 bestimmt wurde. Der Versuch geht in beiden Fällen als „Zielwert eingehalten“ in die Auswertung ein. Hierdurch geht die Information verloren, dass im zweiten Fall, eine um den Faktor 10 höhere Entfernung ermittelt wurde. Durch die Nutzung eines parametrischen 10. Perzentils wird diese Information berücksichtigt.

Die Validierung über die Erfolgsrate als Zielparameter kann jedoch unter bestimmten Bedingungen und Voraussetzungen eine interessante Alternative bzw. Notwendigkeit darstellen, nämlich falls:

1. die Vorgaben der EU-Verordnung 2020/741 direkt umgesetzt werden, und Werte $< \text{BG}$ im Ablauf als „Erfolg“ gewertet werden dürfen (notwendige Voraussetzung),
2. die Konzentration an Indikatororganismen im Zulauf der Anlage nicht ausreichend hoch ist, um die Zielwerte numerisch zu validieren, UND
3. die Wahrscheinlichkeit eines Misserfolgs sehr gering ist, da jeder Misserfolg mit zusätzlichen Proben kompensiert werden muss.

Tatsächlich kann für den Fall, dass Voraussetzung 1 gilt und gleichzeitig mit hoher Sicherheit davon ausgegangen werden kann, dass alle Werte im Ablauf der Anlage $< \text{BG}$ liegen, theoretisch auf die Analyse des Zulaufs verzichtet werden. Auf diese Weise kann der analytische Aufwand verringert und die Wirtschaftlichkeit des Ansatzes verbessert werden. Ein Minimum von 25-28 Proben wäre jedoch noch immer notwendig.

5.7.2 Quantifizierung der statistischen (Un-)Sicherheit

Für eine statistische Validierung sollte der jeweilige Zielparameter mit ausreichender statistischer Sicherheit oberhalb der gesetzten Zielwerte liegen. Die statistische Sicherheit von Zielparametern wird oft über Unsicherheitsintervalle quantifiziert. Unsicherheitsintervalle für statistische Parameter werden auch Konfidenzintervalle genannt. Liegt z.B. das 95%-Konfidenzintervall eines Parameters in Gänze oberhalb eines bestimmten Zielwertes, kann dies in der Regel mit einem statistisch signifikanten Unterschied zum Konfidenzniveau 95% gleichgesetzt werden. Konfidenzintervalle lassen sich sowohl für das 10. Perzentil, als auch für die Erfolgsrate r berechnen. Für eine statistische Validierung über Konfidenzintervalle sollte dementsprechend das jeweilige Konfidenzintervall in Gänze oberhalb der zu validierenden Zielwerte liegen. Für die Erfolgsrate sollte das Konfidenzintervall folglich in Gänze oberhalb von 90% liegen, für das 10. Perzentil oberhalb des zu validierenden Log_{10} -Reduktionswerts, (6 / 5 / 4) für (Coliphagen / *E. coli* / Sporen von *Clostridium perfringens* Sporen).

Die Spannweite des Konfidenzintervalls und somit auch die Anforderung an die Validierung hängen wesentlich vom Grad der geforderten statistischen Sicherheit ab. Es lassen sich beispielsweise 80%, 90%

und 95% Intervalle konstruieren. Je breiter das Intervall, desto mehr Aufwand ist erforderlich um zu zeigen, dass es in Gänze oberhalb des Zielwerts liegt. Die Breite des Intervalls nimmt mit der Anzahl an Stichproben ab. In der wissenschaftlichen Literatur wird in der Regel ein Konfidenzniveau von 95%, bzw. $\alpha = 0.05$ angesetzt. Dies entspricht einer tolerierten falsch-positiv Rate von 5%, es wird also toleriert das 1 von 20 Kläranlagen als „validiert“ bewertet wird, obwohl dies in Wahrheit nicht der Fall ist. Wird das Konfidenzniveau verringert, erhöht sich diese Toleranzschwelle entsprechend.

Als weiteres Unsicherheitsintervall wurde das Toleranzintervall betrachtet, insbesondere die einseitige untere Toleranzgrenze. Diese ist eng verwandt mit dem Konfidenzintervall für das 10. Perzentil, da die untere Grenze eines einseitigen unteren Toleranzintervalls zum Konfidenzniveau $(1-\alpha)$ und $P = 0,9$, mit der unteren Grenze eines Konfidenzintervalls des entsprechenden Perzentils (hier 10. Perzentil) gleichgesetzt werden kann. Liegt also beispielsweise die untere Grenze eines einseitigen Toleranzintervalls mit $P = 0,9$ und $\alpha=0,05$ oberhalb eines zu validierenden Zielwertes, kann hierdurch gezeigt werden, dass das einseitige untere Konfidenzintervall des 10. Perzentils mit $\alpha = 0,05$ ebenfalls oberhalb des Zielwerts liegt.

Für eine statistische Validierung, dass 90% der Validierungsmessungen die Zielwerte einhalten, also größer als diese sind, wird die Verwendung der unteren Grenze des einseitigen unteren Toleranzintervalls, bzw. des einseitigen unteren Konfidenzintervalls des 10. Perzentils als die richtige Wahl angesehen. Dabei erscheint ein statistisches Konfidenzniveau von 95%, vor allem wegen seiner großen Akzeptanz und seiner Stellung als „quasi Standard“ in der wissenschaftlichen Fachliteratur zunächst als geeigneter Wert für das zu definierende Vertrauensniveau. Es ist zu betonen, dass Anpassungen auf höhere oder niedrigere Werte (99%, 90%, 80%) grundsätzlich denkbar sind und die Verwendung von 95% oft auf Konvention beruht. In der Versuchsplanung wird beispielsweise zur Abschätzung der benötigten Stichproben oft nur ein Konfidenzniveau von 80% herangezogen.

5.7.3 Erforderliche Anzahl an Stichproben

Bei der hier empfohlenen Nutzung eines einseitigen unteren Toleranzintervalls ($\alpha = 0,05$, $P = 0,9$) ist es grundsätzlich nicht notwendig, die Anzahl an Stichproben im Vorfeld vorzugeben, bzw. zu begrenzen, da die Spannweite des Intervalls bei kleinen Stichproben aufgrund der statistischen Unsicherheit sehr weit ist (Abbildung 13). In den aufgeführten Beispielen wurde gezeigt, dass in Fällen, in denen die Differenz zwischen geforderter und installierter Log_{10} -Reduktion sehr hoch ist, bereits eine geringe Anzahl an Stichproben ($N = 6-10$) für eine Validierung ausreichen kann. Ist die Differenz jedoch klein, wird eine höhere Anzahl an Stichproben benötigt. Dass für sehr leistungsfähige Anlagen, die die Vorgaben der Verordnung 2020/741 deutlich übererfüllen, eine geringere Anzahl an Stichproben für die Validierung benötigt wird, ist konsequent und der Umstand das dies durch die Nutzung von Toleranzintervallen berücksichtigt wird, stellt einen wesentlichen Vorteil der Nutzung von Toleranzintervallen dar. Die benötigte Anzahl an Stichproben, hängt neben der Differenz zwischen installierter und geforderter Reinigungsleistung auch von der Streuung der erhobenen Messwerte ab. Abbildung 16 zeigt den Zusammenhang zwischen der Differenz zwischen der mittleren Reduktionsleistung und dem geforderten Zielwert, der Streuung der gewonnenen Log_{10} -Reduktionswerte und der Anzahl der benötigten Stichproben für den Fall normalverteilter Log_{10} -Reduktionswerte. Aus der Abbildung wird deutlich, dass bei einer Differenz von zwei logarithmischen Stufen, (Beispiel: Zielwert = 5 und mittlere Reduktion = 7 für *E. coli*, oder Zielwert = 6 und mittlere Reduktion = 8 für Coliphagen), eine Validierung mit weniger als 10 Stichproben möglich ist solange die Standardabweichung der Log_{10} -Reduktionswerte kleiner ist als 0,8.

Beträgt die Differenz nur eine logarithmische Stufe, wäre eine Validierung mit weniger als 10 Stichproben bis zu einer Standardabweichung von 0,4 möglich und mit weniger als 20 Stichproben bis zu einer Standardabweichung der Log_{10} -Reduktionswerte von 0,5. Eine Erhöhung der Stichprobenanzahl > 20 führt nur zu geringfügigen Änderungen und wird als nicht praktikabel angesehen. Die Abhängigkeit der benötigten Stichprobenanzahl von der Größe der Standardabweichung verdeutlicht eine Empfindlichkeit der Methode gegenüber vereinzelt, sehr hohen oder sehr niedrigen Werten, da diese die Standardabweichung stark erhöhen können. Treten solche Fälle auf, sollten extreme Werte über Ausreißertest gegebenenfalls entfernt werden. Als Beispiel sei auf die Messwerte der Indikatorzulaufkonzentrationen in Abbildung 4 verwiesen. Für KA1 zeigt sich, dass die Messwerte für *C. perfringens* im Zulauf um die 10^5 KBE/100mL liegen, jedoch in einer Probe ein Wert von 10^2 KBE/100mL gemessen wurde. Berechnet man die Standardabweichung aus allen Messwerten ergibt sich ein Wert von 1,1. Entfernt man den niedrigen Messwert, als Ausreißer, ergibt sich ein Wert von 0,06, was die Validierung stark vereinfacht. Eine Entfernung des niedrigen Werts scheint in diesem Fall geboten, da er die Auswertung unverhältnismäßig stark negativ beeinflusst.

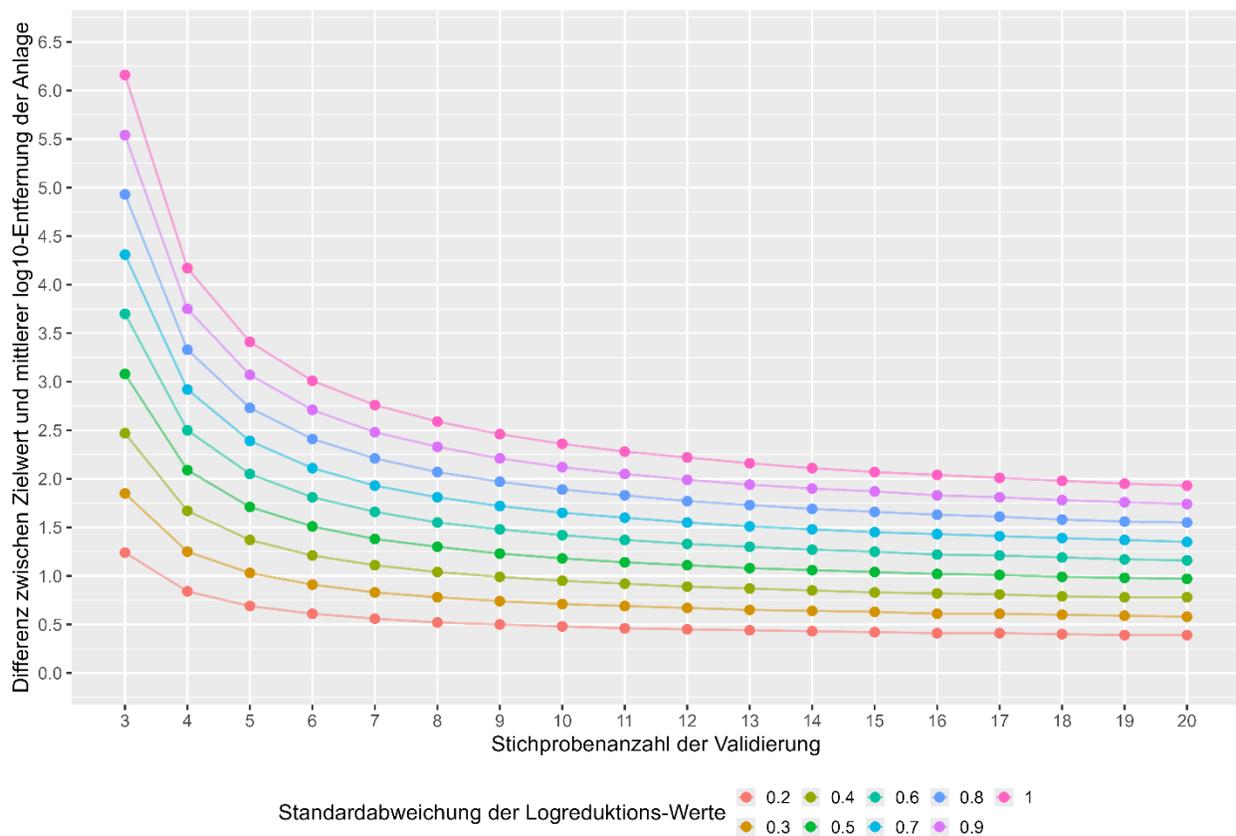


Abbildung 16: Zusammenhang zwischen mittlerer Log_{10} -Reduktion, Streuung der Log_{10} -Reduktion und der Anzahl benötigter Stichproben.

Eine Mindestmenge an Stichproben ist vor allem für Verfahren notwendig, die ausschließlich auf Punktschätzern beruhen und die statistische Parameterunsicherheit somit nicht berücksichtigen. Dies ist bei der Methode der europäischen Badegewässerrichtlinie und auch den Anforderungen an die mikrobiologische Qualität von Bewässerungswasser in den US FDA zur Schätzung von Perzentilgrenzen der Fall. In beiden Verfahren bleibt die statistische Unsicherheit unberücksichtigt. Um eine gewisse statistische

Sicherheit zu gewährleisten, wird eine Mindestanzahl von 16 (Badegewässerrichtlinie) bzw. 20 Stichproben (US FDA) gefordert.

Warum auf die Nutzung von Toleranzintervallen verzichtet wird, bleibt in diesem Zusammenhang unklar. Denkbar wäre eine einfachere Berechnungsmethode. Hier bleibt jedoch festzustellen, dass bei einer vorgegebenen Stichprobenanzahl, sich die Berechnungsmethode ausschließlich im numerischen Wert des Intervallfaktors k unterscheiden (siehe Gleichung (5-8)). Für den Punktschätzer ist dieser stets $k = 1,282$, für die untere Grenze des einseitigen Toleranzintervall wäre dieser für 16 Stichproben stets 2,03 und für 20 Stichproben 1,93. Eine entsprechende Anpassung wäre demzufolge leicht umzusetzen. Aus diesem Grund wird letztendlich kein wesentlicher Grund gesehen, die Berechnung von Punktschätzern, der Berechnung von unteren Toleranzgrenzen vorzuziehen. Die erhöhte Flexibilität eines Toleranzintervalls, bei dem die Anzahl der notwendigen Stichproben nicht vorgegeben wird, sondern von der installierten und beobachteten Reinigungsleistung abhängt, wird sowohl als methodischer als auch praktischer Vorteil gesehen, da sehr leistungsfähige Anlagen auch mit weniger als 16 oder 20 Stichproben validiert werden können.

5.7.4 Ungepaarte Auswertung

Ungepaarte Auswertungsverfahren als auch nicht-parametrische Toleranzintervalle erscheinen vor allem dann als nützlich, wenn ein größerer Anteil der Konzentrationswerte im Klärwerksablauf unterhalb der Bestimmungsgrenze liegt (5-95%). In solchen Fällen ist es wahrscheinlich, dass die Verteilung berechneter gepaarter Log_{10} -Reduktionswerte nicht normalverteilt ist. Dies kann über Standardtests, wie beispielsweise Shapiro-Wilk, oder Anderson Darling Tests geprüft werden. In solchen Fällen ist eine Berechnung von Toleranzgrenzen, die auf der Annahme normalverteilter Werte beruht aus statistischer Sicht anfechtbar, da die Basisannahmen nicht mehr gegeben sind.

Bei ungepaarten Auswertungsalternativen über Monte-Carlo Verfahren, wird die Verteilung der Log_{10} -Reduktionswerte über die *Verteilung der Messwerte* im Zulauf und Ablauf über statistische Simulationen ermittelt.

Annahmen über die *Verteilung der Log_{10} -Reduktionswerte* müssen somit nicht mehr getroffen werden. Werden Messwerte unterhalb der Bestimmungsgrenze als 0 / Volumeneinheit angegeben, steht mit der negativen Binomialverteilung eine geeignete Verteilung zur Verfügung, um ganzzahlige, überdispertierte Werte abzubilden, und die „0“ als Messwert einschließt. Ungepaarte Auswertungsalternativen haben im Vergleich zu nicht-parametrischen Ansätzen den Vorteil, die sie potenziell mit weniger Stichproben durchgeführt werden können. Nachteilig ist, dass die Anwendung ungepaarter Auswertungsmethoden über Monte-Carlo Verfahren methodisch anspruchsvoller ist als über nicht-parametrische Verfahren. Die Vorteile der ungepaarten Auswertung über Monte-Carlo-Verfahren überwiegen aus wissenschaftlicher Sicht deutlich.

Tabelle 12. Vergleich verschiedener Auswertungsalternativen

Methode	Geeignet wenn:	Empfohlen wenn:	Nachteile	Vorteile
(1) Analytisches, klassisches Toleranzintervall über gepaarte Log ₁₀ -Reduktionswerte	Log ₁₀ -Reduktionswerte normalverteilt (nur vereinzelte oder fast alle Werte im Ablauf < BG)	Annahme über Normalverteilung der Log ₁₀ -Reduktionswerte gehalten werden kann	<ul style="list-style-type: none"> Gleiche Anzahl von Zu- und Ablaufwerten (potenzieller Informationsverlust) 	<ul style="list-style-type: none"> Einfach (analytische Lösung + Excelsheet) Methodisch richtig + flexibel berücksichtigt Parameterunsicherheit
(2) Bayes'sches Toleranzintervall über Monte-Carlo-Verfahren und ungepaarte Auswertung	Immer	Annahme über Normalverteilung der Log ₁₀ -Reduktionswerte nicht gehalten werden kann.	<ul style="list-style-type: none"> methodisch aufwendiger 	<ul style="list-style-type: none"> Sehr flexibel Berücksichtigt Parameterunsicherheit Ungleiche Anzahl im Zu- und Ablauf möglich
(3) Nicht-parametrisches Toleranzintervall gepaarter Log ₁₀ -Reduktionswerte	Immer	Werte < BG im Ablauf bei zu niedriger Zulaufkonzentration als „Erfolg“ bzw. „hoher Rang“ gewertet werden dürfen UND Die Auslegung der Anlage Grund zur Annahme geben, dass „Misserfolge“ als extrem unwahrscheinlich anzusehen sind.	<ul style="list-style-type: none"> Hohe Anzahl an Stichproben notwendig Einzelnen Fehlversuche erfordern zusätzliche Stichproben 	<ul style="list-style-type: none"> Potenziell ohne Volumenerhöhung im Klärwerksablauf möglich. Einfache Auswertung
(4) Binomialansatz (Erfolgsrate)	Immer			
(5) Punktschätzer mit vorgegebenen N (analog zu BG-RL, US-FDA)	Statistische Unsicherheit keine Rolle spielen soll	Nicht empfohlen	<ul style="list-style-type: none"> Statistische Unsicherheit wird nicht berücksichtigt (k-Wert stets 1,282) Vorgegebenes N macht Methode unflexibel Informationsverlust bei ungleichem N in Zu- und Ablauf kein offensichtlicher Mehrwert im Vergleich zu (1) 	Akzeptanz, da bereits in anderen Richtlinien zur Berechnung von Perzentilen verwendet
Empfohlen	Empfohlen unter bestimmten Bedingungen		Nicht empfohlen	

6 Schlussfolgerung und Handlungsempfehlungen

Im vorliegenden Leitfaden wurde detailliert auf die Planung, Durchführung und Auswertung eines Validierungsmonitorings nach Verordnung 2020/741 eingegangen. Im Leitfaden wurden die gesetzlichen Anforderungen zusammengefasst, und Empfehlungen zu Auswahl von Validierungsparametern, der Versuchsdurchführung, Probenanreicherung und Datenauswertung erarbeitet. Es können die folgenden Schlussfolgerungen gezogen und Empfehlungen gegeben werden:

Validierungsparameter:

1. Die Validierung sollte auf Basis der etablierten Indikatororganismen für *E. coli*, *C. perfringens* (Sporen) und Coliphagen (gesamt) und genormter Analyseverfahren durchgeführt werden.
2. Verdünnungen und Methodenauswahl sollten so erfolgen, dass die untere Bestimmungsgrenze bei Messungen im Ablauf stets < 1 pro Probenvolumen beträgt.

Notwendigkeit der Probenanreicherung

3. Falls durch Zusatzregelungen auf nationaler Ebene eine quantitative Validierung auch bei zu niedrigen Zulaufkonzentrationen gefordert wird, müssen größere Probenvolumina im Ablauf des Klärwerks genommen und angereichert werden, um eine quantitative Validierung durchführen zu können.
4. Die Parameter *E. coli* und *Clostridium perfringens* Sporen liegen in der Regel in ausreichend hoher Konzentration im Zulauf von Kläranlage vor, um die geforderte Log₁₀-Reduktion von 5 (*E. coli*) und 4 (*Clostridium perfringens* Sporen) in mindestens 90% der Validierungsmessungen quantitativ validieren zu können.
5. Coliphagen (gesamt) als Summe aus somatischen und F-spezifischen Coliphagen liegen in der Regel nicht in ausreichend hoher Konzentration im Zulauf von Kläranlagen vor, um eine Log₁₀-Reduktion von 6 in 90% der Validierungsmessung quantitativ validieren zu können. Hier ist eine Anreicherung in der Regel notwendig.
6. Auf Basis der Auswertungen von Daten aus neun Klärwerken in Deutschland kann geschlossen werden, dass in Deutschland die Filtration und Anreicherung von 15 -25 L ist mit sehr hoher Wahrscheinlichkeit ausreichend hoch ist, um eine quantitative Validierung für somatische Coliphagen sowie Coliphagen (gesamt) durchführen zu können. Aufgrund der niedrigen Trübungs- und AFS-Werte können solche Volumina vergleichsweise einfach über Ultrafiltrationsmembranen filtriert werden.
7. Da bei einer Anreicherung über Ultrafiltrationsmembranen mehrere Indikatororganismen parallel angereichert werden können, und pro Parameter nur ein Teil des Eluats für die Analyse benötigt wird, sollte in Betracht gezogen werden, das angereicherte Eluat auch für die analytische Validierung der Parameter *E. coli* und *C. perfringens* Sporen zu verwenden. Hierdurch entsteht praktisch kein analytischer Mehraufwand, jedoch kann die Anzahl der notwendigen Stichproben durch eine Erhöhung der beobachtbaren Differenz zwischen Zu- und Ablauf reduziert werden.

Datenauswertung

8. Bei der Datenauswertung sollte die untere einseitige Toleranzgrenze mit $P = 0,9$ berechnet und für die Validierung herangezogen werden.
9. Für das benötigte Vertrauensmaß $(1-\alpha)$ wird in der Wissenschaft in der Regel ein Wert von 0,95 herangezogen ($\alpha = 0,05$). Dieser Wert kann theoretisch an die vorliegende Problemstellung angepasst werden, jedoch wird ein Wert $> 0,8$ als sinnvoll erachtet, um der konzeptionellen Idee eines Toleranzintervalls hinreichend Rechnung zu tragen.
10. Eine minimale Stichprobenanzahl wird nicht empfohlen, da die vorgeschlagene Methodik die stichprobenabhängige Parameterunsicherheit bereits berücksichtigt.
11. Die Berechnung der unteren einseitigen Toleranzgrenze über ungepaarte Bayes'sche Ansätze stellt die flexibelste Methodik für die Bestimmung der unteren Toleranzgrenze dar, da sie sowohl mit nicht-normalverteilten Log_{10} -Reduktionswerten, einer ungleichen Anzahl von Zu- und Ablaufwerten als auch mit Negativbefunden im Ablauf, also 0 / Volumeneinheit methodisch schlüssig umgehen kann. Wesentlicher Nachteil ist die methodische Umsetzung, die Erfahrung in statistischer Simulation erfordert.
12. Für gepaarte Log_{10} -Reduktionswerte, die der Annahme einer Normalverteilung nicht widersprechen, kann die untere Toleranzgrenze über klassische, analytische Annäherungsverfahren erfolgen. Diese kann als einfache Excel-basierte Umsetzungen implementiert werden.
13. Bei der gepaarten Auswertung sollte bei Unterschreiten der unteren Bestimmungsgrenze der dekadische Logarithmus der Zulaufkonzentration als validierte Log_{10} -Reduktion angesetzt werden. Bei einer untern BG von <1 führt dies zu identischen Ergebnissen wie die Verwendung der unteren Bestimmungsgrenze.

Dokumentation

14. Die Durchführung des Validierungsmonitorings ist vor dessen Umsetzung in einem detaillierten Validierungsprotokoll festzuhalten.
15. Das Validierungsprotokoll sollte mindestens die folgenden Punkte enthalten:
 - a. Beschreibung des Aufbereitungsprozesses
 - b. Definition der Zeiträume in denen Wasser zu Bewässerung aufbereitet werden soll
 - c. Beschreibung der zu validierenden Betriebsfenster und Betriebszustände
 - d. Probennahmenplan mit genauen Probennahmezeitpunkten, ggf. mit Zwischenauswertungen nach Chargen von 5, 10, 15, 20 Stichproben.
 - e. Beschreibung des Umgangs mit Betriebsstörungen
 - f. Beschreibung der Art Probennahme und Probenkonservierung inklusive des Transports zum Labor.

- g. Bestimmung und Nennung unabhängiger und zertifizierter Probennehmer und für die Analysen zuständigen Labore.
- h. Probenbegleitschein in dem Ort, Zeitpunkt, Messparameter und Analysenmethode dokumentiert sind.
- i. Ggf. Beschreibung der Anreicherungsmethode, inklusiver Auswertungsverfahren
- j. Beschreibung der Datenauswertung
- k. Rollen der beteiligten Personen während der Validierung, u.a. Anlagenbetrieb, Probenahme, Analyselabor, Datenauswertung

Offene Punkte und Vereinfachung der Implementierung

1. Bei der Probenanreicherung und Auswertung über UF-Hohlfasermembranen ist die parameterspezifische Wiederfindung eine wichtige Größe. Für F-spezifische Phagen lagen diese in den durchgeführten Experimenten sehr hoch, was die Belastbarkeit der Ergebnisse in Frage stellt. Um dieses Anreicherungsverfahren weiter zu etablieren, sollte die Wiederfindung für die Validierungsparameter besser quantifiziert werden.
2. Wie beschrieben stellt die Berechnung von Toleranzintervallen mit Hilfe ungepaarter Bayes'scher Verfahren das flexibelste statistische Verfahren dar. Sein wesentlicher Nachteil liegt in der komplizierteren praktischen Umsetzung. Da dieses statistische Verfahren, in den meisten Fällen angewendet werden kann, sollte eine Umsetzung in Form eines Auswertungstools in Betracht gezogen werden.
3. Im Allgemeinen kann es bei der Umsetzung der beschriebenen statistischen Methoden zu Fehlern kommen. Um diese Fehlerquelle auszuschalten, sollten für die statistische Auswertung standardisierte Auswertungsskripte, oder Excelsheets zur Verfügung gestellt werden.

7 Literatur

- DWA 2024 Merkblatt DWA-M 1200-2 Wasserwiederverwendung für landwirtschaftliche und urbane Zwecke in Deutschland - Teil 2: Anforderungen an die weitergehende Wasseraufbereitung
Gelbdruck, Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V.
- Gelman, A., Hill, J. and Vehtari, A. (2020) *Regression and Other Stories*, Cambridge University Press, Cambridge.
- Hahn, G.J., Meeker, W.Q. and Escobar, L.A. 1991 *Statistical Intervals: A Guide for Practitioners*.
- Hill, V.R., Polaczyk, A.L., Kahler, A.M., Cromeans, T.L., Hahn, D. and Amburgey, J.E. 2009. Comparison of hollow-fiber ultrafiltration to the USEPA VIRADEL technique and USEPA method 1623. *J Environ Qual* 38(2), 822-825.
- Kort, A., Petzoldt, H., Hügler, M. and Thronicker, O. 2017 *Entwicklung einer Methodik zur verbesserten Identifizierung und Bewertung von Eintragsquellen mikrobiologischer Belastungen in der Trinkwasserprozesskette*. Forschung, D. (ed), DVGW Deutscher Verein des Gas- und Wasserfaches e. V.
- Krishnamoorthy, K. and Mathew, T. (2009) *Statistical Tolerance Regions: Theory, Applications, and Computation*, Wiley.
- Shi, H., Pasco, E.V. and Tarabara, V.V. 2017. Membrane-based methods of virus concentration from water: a review of process parameters and their effects on virus recovery. *Environmental Science: Water Research & Technology* 3(5), 778-792.
- WHO 2016 *Quantitative Microbial Risk Assessment: Application for Water Safety Management*, World Health Organisation, Geneva, Switzerland.

Anhang

I. Vorversuch zur Bestimmung des Korrekturfaktors „repräsentatives Ablaufvolumen V_r “

Im folgenden Anhang wird auf die Ermittlung des Korrekturfaktors „repräsentatives Ablaufvolumen“ durch einen Vorversuch eingegangen. Wie für die allgemeine Validierung erfolgt eine solche Ermittlung über die einseitige untere Toleranzgrenze gewonnener Messdaten, nach:

$$TI_u = MW - k_1 SD$$

Für den Vorversuch wird eine Probenanzahl von 10-15 empfohlen, da für eine weitere Erhöhung der Präzision eine unverhältnismäßig hohe Anzahl an Proben genommen werden müsste.

Die exakte Bestimmung des Intervallfaktors k_1 erfolgt über die inverse kumulierte Verteilungsfunktion der nicht-zentralen Student-t Verteilung. Diese Funktion ist leider in der Standardanwendung von MS Excel nicht verfügbar, sondern erfordert spezialisierte Software wie R oder Python.

Dennoch lässt sich der Faktor wie folgt analytisch annähern bestimmen, sodass Unterschiede in der Praxis für $N > 6$ zu vernachlässigen sind:

$$a = 1 - \frac{z_\alpha^2}{2(N-1)} \quad (7-1)$$

$$b = z_p^2 - \frac{z_\alpha^2}{N} \quad (7-2)$$

$$k_1 = \frac{z_p + \sqrt{z_p^2 - ab}}{a} \quad (7-3)$$

$$TI_U = MW - k_1 SD \quad (7-4)$$

Mit:

z_α : z-Wert des Konfidenzniveaus α (für $\alpha = 0.05 \rightarrow z_\alpha = 1,65$)

z_p : z-Wert des zu schätzenden Perzentils (hier $P = 0,9 \rightarrow z_p = 1,282$)

N: Anzahl der Stichproben

MW: Mittelwert der logarithmierten Zulaufwerte

SD: Standardabweichung der logarithmierten Zulaufwerte

TI_u : einseitiges unteres Toleranzintervall

Werte für k_1 , für $\alpha = 0.95$ und $P = 0,9$, sind für verschiedene Werte für N im Anhang (Kapitel II) gelistet.

Liegt Tl_u oberhalb des Zielwerts kann mit einer Sicherheit von *alpha* von einer hinreichen hohen Zulaufkonzentration ausgegangen werden. Eine Validierung kann direkt, ohne weitere Aufkonzentrierung, mit den in Kapitel 5 dargestellten Methoden erfolgen.

Eine Aufkonzentrierung nach Kapitel 4.6 kann dennoch sinnvoll sein, wenn für die installierte Aufbereitungskette eine deutlich höhere Logreduktion erwartet wird, als dies durch gemessenen Werte validiert werden kann. Liegen die gemessenen Zulaufwerte beispielsweise im Bereich zwischen 10^6 und 10^7 , jedoch wird für die Anlage eine Logreduktion > 7 erwartet, kann durch eine Aufkonzentrierung trotz hinreichend hoher Zulaufwerte, die Anzahl der benötigten Stichproben je nach Auswertungsverfahren deutlich reduziert werden.

Liegt Tl_u unterhalb des zu validierenden Zielwertes, muss davon ausgegangen werden, dass mehr als 10% der berechneten Logreduktionen den Zielwert unterschreiten und nur als „ $> \log_{10}(C_{Zulauf})$ “ angegeben werden können. In diesen Fällen ist eine Aufkonzentrierung erforderlich. Der Korrekturfaktor V_r der für die Schätzung des minimal zu filtrierenden Volumens benötigt wird kann wie folgt berechnet werden:

$$KF = 10^{(Zielwert - Tl_u)}$$

$$V_r = V_{analysiert} \cdot KF$$

II. Korrektur k Werte Toleranzintervall

N	k approx (alpha = 0.95)	k exact (alpha = 0.95)	k (alpha = 0.5) (z-Wert)	Kommentar
5	3.38	3.41	1.28	
6	2.96	3.01	1.28	
7	2.71	2.76	1.28	
8	2.54	2.58	1.28	
9	2.42	2.45	1.28	
10	2.32	2.35	1.28	
11	2.24	2.28	1.28	
12	2.18	2.21	1.28	
13	2.13	2.16	1.28	
14	2.09	2.11	1.28	
15	2.05	2.07	1.28	
16	2.01	2.03	1.28	Badegewässerrichtlinie
17	1.98	2	1.28	
18	1.96	1.97	1.28	
19	1.93	1.95	1.28	
20	1.91	1.93	1.28	US FDA Bewässerungswasser
21	1.89	1.91	1.28	
22	1.87	1.89	1.28	
23	1.86	1.87	1.28	
24	1.84	1.85	1.28	
25	1.83	1.84	1.28	
26	1.81	1.82	1.28	
27	1.8	1.81	1.28	
28	1.79	1.8	1.28	
29	1.78	1.79	1.28	
30	1.77	1.78	1.28	
31	1.76	1.77	1.28	
32	1.75	1.76	1.28	
33	1.74	1.75	1.28	
34	1.73	1.74	1.28	
35	1.72	1.73	1.28	
36	1.72	1.72	1.28	
37	1.71	1.72	1.28	
38	1.7	1.71	1.28	
39	1.7	1.7	1.28	
40	1.69	1.7	1.28	
41	1.68	1.69	1.28	
42	1.68	1.69	1.28	
43	1.67	1.68	1.28	
44	1.67	1.67	1.28	
45	1.66	1.67	1.28	
46	1.66	1.66	1.28	
47	1.65	1.66	1.28	
48	1.65	1.65	1.28	
49	1.64	1.65	1.28	
50	1.64	1.65	1.28	