

Geringe Unsicherheit und ein hoher Grad des Vertrauens – zum Abschied von Bernhard Heck

Maria Hennes

Geodätisches Institut, Karlsruher Institut für Technologie
E-Mail: maria.hennes@kit.edu

Zusammenfassung

Die Metrologie lehrt, dass eine Messgröße ohne Angabe der Messunsicherheit, die dem Messwert vernünftigerweise zuzuordnende Streubreite, wertlos ist. Dieser Beitrag zeigt den Mehrwert der Messunsicherheit unter Angabe des Grades des Vertrauens gegenüber der bisher in der Geodäsie gebräuchlichen empirischen Standardabweichung auf und schlägt den Bogen zur Zusammenarbeit mit meinem Kollegen Bernhard Heck.

1 Motivation

In der Bandbreite möglicher Typen von Festschriftbeiträgen zwischen wissenschaftlichem Fachbeitrag und persönlichem Gruß habe ich mich für eine Kombination entschieden, mit der ich zeigen möchte, dass sich gewisse, zunächst sehr fach-spezifisch erscheinende Zusammenhänge auch auf andere Aspekte übertragen lassen. Allem voran möchte ich mit diesem Beitrag meinem Kollegen Bernhard Heck meinen Dank aussprechen für die hervorragende Zusammenarbeit in mehr als 17 Jahren gemeinsamen Schaffens am GIK. „Unsicherheit“ und „Grad des Vertrauens“ sind Begriffe, die auch eng mit der Thematik Messunsicherheit verknüpft sind. Diese Thematik verbindet alle Disziplinen, die sich mit qualitativ hochwertiger Messung und Vermessung beschäftigen. Gleichwohl lassen sich auch Analogien in Bezug auf die Zusammenarbeit unter den Kollegen finden.

2 Das Wesen der Messunsicherheit

Qualität ist hier im Sinne der DIN-EN-ISO-9000:2015-11 (2015) zu verstehen, die Qualität als den Grad, in dem ein Satz inhärenter Merkmale eines

Objekts Anforderungen erfüllt, definiert. Im Kontext des (Ver-)Messens ist eine Messgröße oder Zielgröße qualitativ hochwertig, wenn sie die Anforderung hoher Genauigkeit in Bezug auf die inhärenten, also innewohnenden Merkmale erfüllt. Alle innewohnenden Merkmale tragen zu den Abweichungen der Messwerte vom idealen Wert, dem Sollwert, bei und führen zu mehr oder weniger variierenden bzw. streuenden Beobachtungen.

Bekanntlich ist die Messunsicherheit ein dem Messergebnis zugeordneter Parameter, der die Streuung der Werte kennzeichnet, die in vernünftiger Weise dem Messergebnis zugeordnet werden kann (nach Definition JCGM (2008), vgl. auch Krystek (2012)). Damit ist ein ganz wesentliches Merkmal der Messunsicherheit genannt, nämlich, dass die Messunsicherheit eine realistische und vertrauenswürdige Angabe der möglichen Schwankungs- bzw. Streubreite eines Messergebnisses oder einer Zielgröße darstellt, wobei alle zur Abweichung beitragenden Einflüsse berücksichtigt werden. Demzufolge kann sie in der Regel nicht vollständig empirisch aus stochastisch streuenden Messdaten berechnet werden. Dies kann zum einen durch den zu kleinen Stichprobenumfang verursacht sein. Zum

Erschienen bei KIT Scientific Publishing
Schriftenreihe des Studiengangs
Geodäsie und Geoinformatik 2018,1

DOI Einzelbeitrag:
10.5445/KSP/1000080218

Festschrift zur Verabschiedung von
Prof. Dr.-Ing. Dr. h.c. Bernhard Heck
(Schw)Ehre, wem (Schw)Ehre gebührt

DOI Festschrift:
10.5445/KSP/1000080324



Dieses Werk ist lizenziert unter einer
Creative Commons Namensnennung
- Weitergabe unter gleichen Bedin-
gungen 4.0 International Lizenz (CC BY-SA 4.0): <https://creativecommons.org/licenses/by-sa/4.0/deed.en>

anderen wird es kaum möglich sein, die Stichprobe messtechnisch so zu erfassen, dass sich während der Messung alle wirkenden Einflüsse in der Stichprobe in Form von charakterisierenden Schwankungen in der Messgröße selbst zeigen. Diese Einflüsse sind aber trotzdem (als unbekannte und manchmal sogar systematisch wirkende Abweichungen) vorhanden und beeinflussen die Belastbarkeit des Messergebnisses. Derartige Abweichungen vom tatsächlichen Wert sind heimtückisch, weil sie in der empirischen Datenerhebung nicht erkannt werden können, aber das Messergebnis trotzdem verfälschen – und zwar mehr, als die klassisch empirisch abgeleitete Standardabweichung vermuten lässt. Der Nutzer des Messergebnisses (oder der abgeleiteten Zielgröße) benötigt aber belastbares Wissen über die zu erwartende Schwankungsbreite der Zielgröße einschließlich derartiger potentieller unbekannter systematisch wirkender Einflüsse.

Derartige Abweichungen, die zusätzlich zu den in den (Wiederholungs-)Messungen enthaltenen zufälligen Streuungen die Qualität der Zielgröße charakterisieren, werden – im Gegensatz zur klassisch empirisch ermittelten Standardabweichung – durch die Messunsicherheit abgedeckt. Sie liefert – auch wenn sie mit Unsicherheit bezeichnet wird – eine Sicherheit, nämlich diejenige, dass der Metrologe (sei es nun der Wissenschaftler oder der Ingenieur) davon ausgehen kann, dass vernünftigerweise keine weiteren Beiträge zu erwarten sind, die die Qualität des Messergebnisses darüber hinaus verschlechtern. Dies ist von zentraler Bedeutung, da eine empirisch ermittelte Standardabweichung in der Regel – und naturgemäß – die Situation viel zu optimistisch beschreibt: unter Umständen gilt sie nur für Wiederholbedingungen und erlaubt noch nicht einmal eine Aussage über die Reproduzierbarkeit (die ebenfalls nicht zwingend alle verfälschenden Ursachen enthält).

Dies gilt auch für Standardabweichungen, die im Rahmen einer Ausgleichsrechnung erhalten werden, denn naturgemäß kann dieser Rechenprozess auch nur diejenigen Schwankungen berücksichtigen, die während der Datenerhebung in Form von Variationen der gemessenen Größe selbst aufgetreten sind. Denn der nachfolgende (!) Ausgleichsprozess garantiert

naturgemäß nicht, dass während der Messung selbst alle Abweichungen, die auf den Messwert wirken, in schwankender Form aufgetreten sind und sich in den Beobachtungen niedergeschlagen haben. Diese an sich triviale Feststellung wird allerdings oft übersehen. Die Messunsicherheit hingegen ist ein sicheres, verlässliches und belastbares Qualitätsmaß, wenn sie sachgerecht und nachvollziehbar ermittelt wird.

Der Leitfaden zur Angabe der Unsicherheit beim Messen (Guide to the expression of uncertainty in measurement, GUM, vgl. JCGM (2008)¹) sieht daher nicht nur die Berücksichtigung aller Einflüsse vor, sondern empfiehlt auch die Auflistung aller berücksichtigten Unsicherheitsquellen, um einen Überblick zu geben, inwieweit Vollständigkeit erreicht wurde. Die Formulierung dieses Leitfadens als praxisnahe Handreichung täuscht allerdings darüber hinweg, dass die dort dargelegte Vorgehensweise konsistent mit dem Bayes-Theorem ist, das es erlaubt, (Vor-)Wissen über eine Größe mit der ihm zugehörigen Wahrscheinlichkeit zu berücksichtigen. In Fall der Messunsicherheit besteht das Vorwissen in der Kenntnis des Betrags einer möglicherweise auftretenden Messabweichung und ihrer Wahrscheinlichkeit des Auftretens (vgl. auch Weiser und Wöger (1999)).

3 Zur Berechnung der Messunsicherheit

Die Messunsicherheit kann sowohl für unmittelbar gemessene Größen, als auch für aus Messungen abgeleitete Zielgrößen ermittelt werden. Im Folgenden werden die einzelnen Schritte dokumentiert.

- a) Entscheidend ist die vollständige Erfassung aller wirkenden Einflüsse, denn alles, was vernachlässigt und übersehen wird, schadet der Qualität der berechneten Messunsicherheit und damit der Beurteilbarkeit des Messergebnisses. Als Werkzeug für die möglichst vollständige Ermittlung aller Einflüsse eignet sich ein Ishikawa-Diagramm. Die Modifizierung für eine Anwendung des Vermessungswesens zeigt Hennes (2007).

¹Das Joint-Committee for Guides in Metrology (JCGM) hat den 1993 erschienenen korrespondierenden ISO-Guide mit redaktionellen Änderungen für nicht-kommerziellen Gebrauch publiziert. Die deutsche Übersetzung ist als DIN-Vornorm ENV V 13005 im Umlauf.

- b) Die jeweilige Schwankungsbandbreite der einzelnen Einflüsse mitsamt der statistischen Verteilung ist festzustellen. Die Standardunsicherheit kann bei Vorliegen von geeignetem Datenmaterial als Standardabweichung aus empirischen Messreihen berechnet werden. Bei ausreichendem Stichprobenumfang kann auch die statistische Verteilung ermittelt werden. Sollten lediglich die Grenzen der Schwankungsbreite bekannt sein, kann unter der Annahme einer plausiblen Verteilung die Transformation in die Standardabweichung bzw. die Standardunsicherheit erfolgen. Beispielsweise handelt es sich bei digital erfassten Werten um eine Gleichverteilung innerhalb des kleinsten Anzeigewertes (Quantisierungsfehler), so dass sich gemäß

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \quad (3.1)$$

mit $f(x) = 1$ die Standardabweichung σ_Q für den durch Quantifizierung eingebrachten Unsicherheitsbeitrag durch Division des kleinsten Quantisierungsschritts durch $\sqrt{12}$ ergibt.

Bei fehlendem statistischen Datenmaterial zur Ableitung von Standardabweichungen empfiehlt es sich also, Grenzen der möglichen Schwankungsbreite und die Verteilung innerhalb des Intervalls heranzuziehen, oder sogar die Standardabweichung² abzuschätzen, um eine verlässliche Gesamt-Messunsicherheit berechnen zu können. In der bisherigen Anwendung der Varianzfortpflanzung werden allerdings oft Einflüsse ignoriert, deren Beträge mangels geeigneter empirischer Datengrundlage abgeschätzt werden müssten. Eine derartige Vernachlässigung bekannter Variationen stellt jedoch die Verlässlichkeit (Vertrauenswürdigkeit) des Genauigkeitsmaßes weit mehr in Frage als eine vernünftige Abschätzung ihrer Variationsbandbreite (ohne empirische Datengrundlage). Um die Berücksichtigung von zusätzlichem Wissen über die Messwertvariation, das erst zu einem späteren Zeitpunkt verfügbar ist, zu vereinfachen, empfiehlt sich die tabellenartige Dokumentation aller Unsicherheitsbeiträge und der damit verknüpften Annahmen.

- c) Um die Wirkung der einzelnen Einflüsse auf die Zielgröße zu erkennen, müssen alle Einflüsse in das funktionale Modell der Zielgröße einfließen. Aufgrund des funktionalen Modells zwischen Messgröße(n) und Zielgröße ist auf das Vorliegen systematisch wirkender Einflüsse zu prüfen. Diese sind soweit möglich durch entsprechende Korrekturen im funktionalen (deterministischen) Modell zu beseitigen. Die Wirkung eines systematischen Einflusses, dessen aktuelle Größe unbekannt und dessen mögliche Streubreite bekannt ist, wird durch (ggf. zusätzliche) Parametrisierung im deterministischen Modell berücksichtigt und bei der Berechnung der Zielgröße im deterministischen Modell mit einem entsprechenden Wert belegt (wenn der Einfluss additiv wirkt, z. B. mit Null). Dies garantiert die wirkungsgerechte Berücksichtigung bei der Bestimmung der Messunsicherheit.
- d) Durch Bilden des totalen Differentials wird die Wirkung der einzelnen Einflüsse offensichtlich, oder anders gesagt, es wird deutlich, wie sensitiv die Zielgröße auf jeden einzelnen Einfluss reagiert. Jedes Differential wird daher als Sensitivitätskoeffizient bezeichnet. Die Analyse von Sensitivitätskoeffizienten kann vorteilhaft bei der Optimierung von Messanordnungen genutzt werden.
- e) Analog dem Varianzfortpflanzungsgesetz werden die einzelnen Einflüsse auf die Zielgröße fortgepflanzt: die Wurzel aus der Quadratsumme aller mit dem jeweiligen Differential gewichteten Standardunsicherheiten wird kombinierte Messunsicherheit genannt. Dieser Ansatz folgt einer seit langem in der Geodäsie angewendeten Methode, die bisher allerdings nur empirisch erhaltene Standardabweichungen aus normalverteilten Größen (oder was man dafür hielt) berücksichtigt. In der geodätischen Fachwelt stand man der Einbeziehung von nicht-normalverteilten stochastischen Größen in die Quadratsumme zunächst verhalten gegenüber, weil weder erkannt worden war, dass diese Vorgehensweise konform zur Berücksichtigung von Vorinformationen gemäß des Bayes-Theorems ist, noch dass sich nach dem zentralen Grenzwertsatz

²Zur Abgrenzung, zur empirisch aus einer Messreihe gewonnenen Standardabweichung, wird im Kontext der Messunsicherheitsermittlung jede Standardabweichung mit dem Begriff „Standardunsicherheit“ bezeichnet.

der Statistik (vgl. Koch (2004)) das Ergebnis einer Faltung von sehr vielen unterschiedlichen Verteilungen einer Normalverteilung annähert. Diese Annäherung ist in geodätischen Anwendungen meistens auch bei wenigen Komponenten ausreichend, weil sich hier viele der Einflüsse normalverteilt verhalten. Der durch nicht 100-prozentige Einhaltung des Grenzwertsatzes verbleibende Modellfehler ist in der Regel erheblich kleiner ist als die grundsätzliche Vernachlässigung von nicht-normalverteilten Komponenten im Gesamtbudget. Die Metrologie empfiehlt daher allen Anwendern, die ein belastbares Ergebnis vertreten müssen, eine vollständige Berücksichtigung aller Einflüsse einer unvollständigen Berechnung vorzuziehen. In Sonderfällen, in denen die Anwendbarkeit des zentralen Grenzwertsatzes fragwürdig ist, empfiehlt der Leitfaden GUM die Anwendung der Monte-Carlo-Methode.

f) Analog zur Varianzfortpflanzung kann eine vollbesetzte Kovarianzmatrix mitgeführt werden, um auftretenden Korrelationen Rechnung zu tragen. Diese sind empirisch aus Messreihen zu ermitteln oder aus bekannten Funktionalen abzuleiten. Die empirische Bestimmung kann nicht immer mit sinnvollem Aufwand für alle Prozesse durchgeführt werden, auch wenn von ausgeprägten Korrelationen auszugehen ist. Auch für die Funktionalen liegen nicht immer Parametrisierungen vor. Daher ist die Berücksichtigung von aus Erfahrung bekannten Korrelationen empfehlenswert, denn auch hier gilt, dass die Vernachlässigung von Korrelationen zu optimistische Standardabweichungen bzw. zu optimistische kombinierte Messunsicherheiten zur Folge hat (vgl. JCGM (2008), Abschnitt 5.2.5). Durch eine vernünftige Besetzung der Kovarianzmatrix verliert die über Varianzfortpflanzung berechnete „Varianz“ zwar an statistischer Exaktheit, gewinnt aber an höherer Qualität, weil die inhärenten Merkmale, nämlich die Korrelationen, so gut es eben geht berücksichtigt sind. Zur Abgrenzung wird diese Größe anstelle von Varianz mit „quadratischer kombinierter Messunsicherheit“ bezeichnet.

g) Die unter (f) erhaltene kombinierte Messunsicherheit beschreibt analog zur aus Varianzfortpflanzung erhaltenen Standardabweichung ein Intervall, in dem 68% der erwarteten Realisierungen der Zielgröße liegen. Anstelle des Begriffs Wahrscheinlichkeit wird der Begriff „Grad des Vertrauens“ verwendet, um von der rein auf Zufallsprozesse basierenden empirischen Ermittlung abzugrenzen. Die so genannte „erweiterte Messunsicherheit“ wird durch Multiplikation der kombinierten Messunsicherheit mit dem Faktor k erhalten, um Schwankungsbreiten angeben zu können, die anderen Wahrscheinlichkeiten entsprechen. Typischerweise wird $k=2$ bzw. $k=3$ verwendet, was einer Erhöhung des Grades des Vertrauens auf etwa 95% bzw. etwa 99% entspricht³.

4 Weitere Eigenschaften der Messunsicherheit

Aus den Berechnungsschritten aus Abschnitt 3 ergeben sich weitere Merkmale der Größe „Messunsicherheit“:

- 1) Die Messunsicherheit berücksichtigt alle beteiligten Einflüsse, nicht nur diejenigen, für die Standardabweichungen durch redundante Beobachtungen gewonnen werden können.
- 2) Das „Werkzeug“ Messunsicherheit betrachtet alle störenden Einflüsse und verschließt nicht die Augen vor schwieriger zu beschreibenden, die vom „normalen“ – also angenommenem normalverteilten – Verhalten abweichen, die gerne als normalverteilt unterstellt oder vernachlässigt werden.
- 3) GUM empfiehlt, nicht nur die Größe des Einflusses, sondern auch seine Wirkung zu betrachten, damit eine Messanordnung zielgerechter optimiert werden kann. Die vollständige Modellierung in einem deterministischen Modell, das alle Einflüsse enthält, liefert also die Grundlage für die Bewertung, wie heikel ein Störeinfluss ist – oder auch die Erkenntnis, dass zwar die Schwankungsbreite eines Einflusses groß ist, aber seine Wirkung aufgrund des kleinen Sensitivitätskoeffizienten vernachlässigbar klein ist. Demselben Prinzip folgt zwar auch die Varianzfortpflanzung, jedoch wurde in der Ver-

³Hier wird impliziert, dass annähernd Normalverteilung vorliegt und der Freiheitsgrad groß ist. Die hiermit verknüpften kleinen Modellfehler sind klein und im Hinblick auf die generelle Verlässlichkeit der Angabe tolerierbar.

gangenheit nicht viel Wert auf die vollständige Erfassung aller Einflüsse gelegt.

- 4) Durch die quadratische Zusammenfassung der Einzelbeiträge werden die betragsmäßig großen und damit kritischen (Unsicherheits-)Beiträge überproportional berücksichtigt. Dies ist konsistent zur klassischen Varianzfortpflanzung.
- 5) Wie in der Varianzfortpflanzung auch bleibt die Wirkrichtung der Einflussgröße (das Vorzeichen) irrelevant.
- 6) Durch den „Grad des Vertrauens“ ist eine intuitive Interpretationshilfe gegeben.

Mit der kombinierten Messunsicherheit liegt also eine Größe vor, die alle Einflüsse auf die Genauigkeit einer Zielgröße sachgerecht berücksichtigt. Im Gegensatz dazu deckt die über Varianzfortpflanzung erhaltene Standardabweichung einer Zielgröße nur selektiv das Augenfällige (mit empirischen Standardabweichungen leicht beschreibbare) ab, wobei in der Praxis nicht-normalverteilte Größen gerne vernachlässigt werden. Die Messunsicherheit enthält das gesamte Zusatzwissen über die Qualität der Mess- bzw. Zielgröße, und zwar nach den Empfehlungen des GUM nachvollziehbar dokumentiert. Damit ist sie eine sehr verlässliche Größe, die letztendlich Sicherheit verschafft, weil alle Unwägbarkeiten abgedeckt sind.

5 Korrelation

Abschließend soll die Korrelation zur Zusammenarbeit mit Bernhard Heck hergestellt werden. Ohne Frage ist er ein hervorragender Forscher und Hochschullehrer. Sich im Detail zur Forschung zu äußern, möchte ich den Fachkollegen überlassen; seine Lehrleistung wird durch zahlreiche erfolgreich tätige Alumni bestätigt. Neben herausragender Forschung und engagierter Lehre war Bernhard Heck stets um ein sehr gutes kollegiales Verhältnis zu allen Menschen in seinem Umfeld besorgt. So möchte ich den Bogen schlagen zu Bernhard Heck als Kollegen und auf die unter Abschnitt 4 genannten Eigenschaften eingehen.

Um es gleich vorweg zu nehmen: Bernhard Heck galt unter uns Kollegen und Kolleginnen als Vertrauensperson, die Sicherheit in die vielschichtigen Unsicherheiten des Wirkens an einer universitären Forschungseinrichtung mit sich bringt. Es steht außer Frage, dass jeder an Universitäten Tätige in sei-

nem Wirken manchmal eine gewisse Unsicherheit verspürt, beispielsweise im Dickicht der Verwaltungsgewohnheiten, bei Personal- und Finanzfragen, oder weil Planungssicherheit nicht immer gegeben ist. Dies führt zu Unsicherheit im Projektmanagement, hinzu kommt Unsicherheit durch universitäre Umstrukturierungsmaßnahmen und auch manchmal Unsicherheit in kniffligen fachlichen Fragen. Viele meiner Kollegen werden wohl meiner Erfahrung zustimmen, dass ein Gespräch mit Bernhard Heck immer wieder die Einschätzung von zur Unsicherheit beitragenden Einflussfaktoren erleichterte. Ich habe ihn als jemanden erfahren, der immer möglichst alle Beteiligten und Hintergründe im Blick hatte, (vgl. 1) in Abschnitt 4. Auch wenn es sich um heikle und komplexe Situationen handelte, bei denen schon die Beurteilung der Ausgangssituation aufgrund ihrer Besonderheit herausfordernd war, nahm er sich der Sache an, (vgl. 2). Insbesondere seine Beurteilung der Wirkung von einzelnen Vorfällen und Randbedingungen im Gesamtkontext war mir immer wieder wertvoll für die weitere Entscheidung, die ggf. auch zu einer begründeten (!) Änderung der Vorgehensweise führte, (vgl. 3). Bernhard Heck hat bei seinen Stellungnahmen stets das Kritische im Blick gehabt und entsprechend bewertet, (vgl. 4). Ganz besonders wichtig ist mir hervorzuheben, dass er stets ohne Vorurteile und ohne Bevorzugung einzelner Gruppen oder Strömungen Stellung bezogen hat, (vgl. 5). Durch den kollegialen Austausch gelang es oft, die von außen aufgeprägte Unsicherheit durch Benennung der einzelnen Komponenten und der Erkenntnis ihres Zusammenwirkens handhabbarer zu machen und Entscheidungen auf eine belastbarere Basis zu stellen.

Ich möchte es nicht versäumen, mich an dieser Stelle auch im Namen von Kollegen, Mitarbeitern und Studierenden zu bedanken, dass wir mit ihm als Kollegen mit einem hohen Grad des Vertrauens zusammen arbeiten konnten, weil er immer ein Gefühl der Sicherheit vermittelt hat. Wahrscheinlich bin ich nicht die Einzige, die ihn als Kollegen vermissen wird.

Literatur

- DIN-EN-ISO-9000:2015-11 (2015): Qualitätsmanagementsysteme - Grundlagen und Begriffe.
- Hennes, M. (2007): Konkurrierende Genauigkeitsmaße - Potential und Schwächen aus der Sicht des Anwenders. AVN 2007.
- JCGM (2008): Guide to the expression of uncertainty in measurement. Hrsg. von J. C. for Guides in Metrology. Joint Committee for Guides in Metrology. URL: http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf (besucht am 29. 11. 2017).
- Koch, K.-R. (2004): Parameterschätzung und Hypothesentests in linearen Modellen. 4. Aufl., S. 118.
- Krystek, M. (2012): Berechnung der Messunsicherheit - Grundlagen und Anleitung für die praktische Anwendung. Beuth.
- Weiser, K. und Wöger, W. (1999): Messunsicherheit und Messdaten. Wiley-VCH.