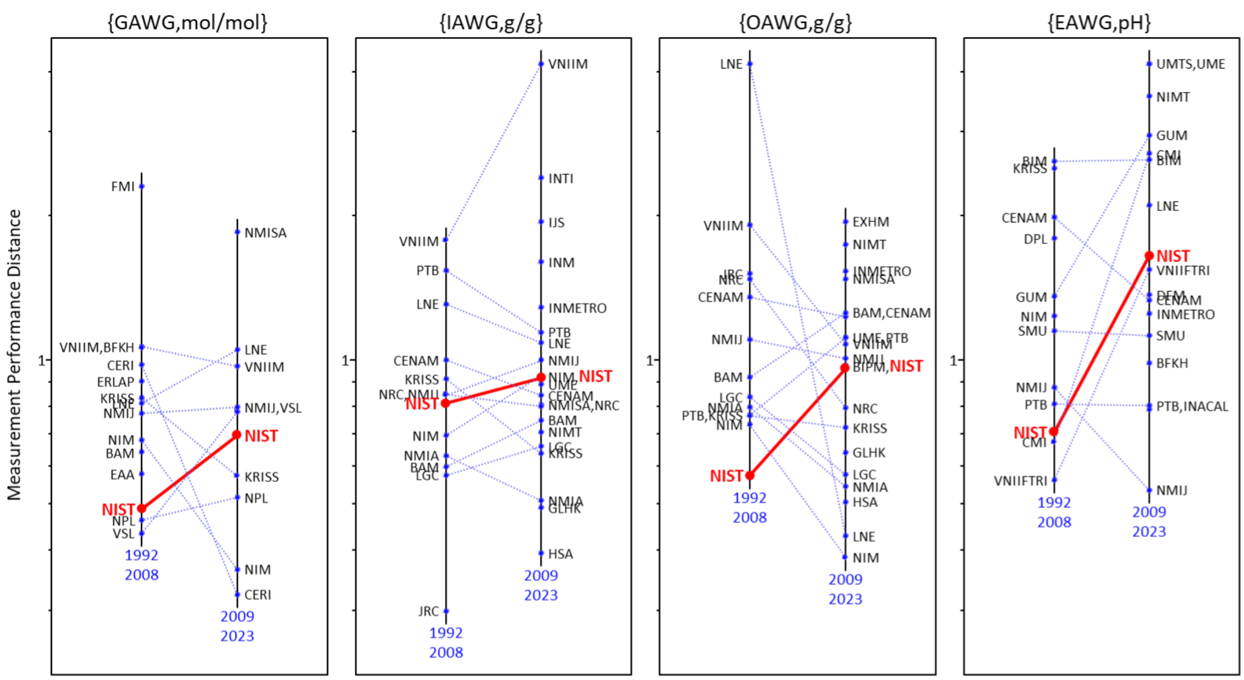
NIST Internal Report
NIST IR 8542

NIST's Engagement with CCQM Studies from 1992 to 2023: Peers, Early and Recent

David L. Duewer

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8542



NIST
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

(This page intentionally blank.)

NIST Internal Report
NIST IR 8542

NIST's Engagement with CCQM Studies from 1992 to 2023: Peers, Early and Recent

David L. Duewer
Chemical Sciences Division
Material Measurement Laboratory
National Institute of Standards and Technology

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8542

October 2024

U.S. Department of Commerce
Gina M. Raimondo, Secretary

National Institute of Standards and Technology
Laurie E. Locascio, NIST Director and Under Secretary of Commerce for Standards and Technology

NIST IR 8542
October 2024

Publication History
Approved by the NIST Editorial Review Board on 2024-10-01

How to Cite this NIST Technical Series Publication
Duewer DL (2024) NIST's Involvement with CCQM Studies from 1992 to 2023: Peers, Early and Recent. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Internal Report (IR) NIST IR 8542. https://doi.org/10.6028/NIST.IR.8542

NIST Author ORCID iDs
Duewer DL 0000-0002-3924-3064

Contact Information
david.duewer@nist.gov

Abstract

This report documents the performance of the National Institute of Standards and Technology (NIST) relative to that of its international peer organizations in studies coordinated by the Consultative Committee for Amount of Substance: Metrology in Chemistry and Biology (CCQM). The data and graphical tools used to analyze and document performance are part of the *CCQM_Retrospectoscope* analysis system, https://doi.org/10.18434/mds2-2952.

Keywords

Consultative Committee for the Amount of Substance: Metrology in Chemistry and Biology (CCQM); Electrochemical Analysis Working Group (EAWG); Gas Analysis Working Group (GAWG); Inorganic Analysis Working Group (IAWG); Key Comparison (KC); Organic Analysis Working Group (OAWG).

Table of Contents

List of Tables

List of Figures

Acknowledgments

I thank my colleagues Michele Schantz, Christina E. Cecelski, Katrice A. Lippa, Michal Máriássy (Slovak Institute of Metrology), and Kenneth W. Pratt for their comments and suggestions.

1. Introduction

An earlier report documented various aspects of NIST's measurement performance in interlaboratory comparison studies conducted by the Consultative Committee for the Amount of Substance: Metrology in Chemistry and Biology (CCQM) from 1992 to 2023 [1]. This report uses results reported in CCQM studies to identify organizations that are NIST's closest metrological peers based on several criteria. These identifications are made using the database and data analysis tools provided by the *CCQM_Retrospectoscope* [2].

Unlike the earlier report which addressed NIST's total engagement with the CCQM, the data analysis tools used in this report require relatively large numbers of results. Such numbers are only available for the analysis of gas mixtures, inorganics, organics, pH, electrolytic conductivity, thin film thickness, and isotopic delta-scale ratios. Only the first four of these chemical arenas have sufficient data to support identifying peer group changes between studies conducted during 1992 through 2008 ("Early" interval) and 2009 through 2023 ("Recent" interval).

The remainder of this report is organized as follows.

- Section 2, Definitions, briefly describes the metrological and *CCQM_Retrospectoscope* jargon used in this report. More complete descriptions are provided in [1,3].
- Section 3, Analysis Scope, describes the data used to identify peer organizations and why the end of 2008 is used as the boundary between Early and Recent studies.
- Section 4, Metrological Responsibility Peers, identifies organizations with interests and metrological responsibilities similar to NIST's history.
- Section 5, Participation Peers, identifies organizations with CCQM participation histories similar to NIST's history.
- Section 6, Coordination Peers, identifies organizations with CCQM coordination histories similar to NIST's history.
- Section 7, Measurement Similarity Peers, identifies organizations with measurement results similar to NIST's history.
- Section 8, Measurement Performance Peers, identifies organizations with measurement results similar to reference criteria, based on results in datasets in which NIST has results. This also enables assessing changes in NIST's performance between the Early and Recent intervals.
- Section 9, Power Law Scaling Function Peers, identifies organizations with power law scaling functions similar to NIST's history in the gas mixture, inorganic, and organic areas where results are expressed in either mole-fraction (mol/mol) or mass-fraction (g/g).

- Section 10, Blind Analysis Peers, identifies organizations with a history of blinding their measurement results to information provided by their coordination of studies similar to NIST's history.
- Section 11, Summary, provides a series of tables summarizing the results of the various peer analyses.

Acronyms and symbols used in the text of this document are defined in Appendix A.

Many sections of this document identify organizations using the codename used in the *CCQM_Retrospectoscope*. See Appendix B for the country, name, and metrological status of these organizations.

The graphical data analysis tools used to produce this report are identified in Appendix C.

2. Definitions

2.1. International Metrology Organizations

Metrology at the international level is highly hierarchical, with defined areas of authority and responsibility.

2.1.1. General Conference on Weights and Measures (CGPM)

The General Conference on Weights and Measures (CGPM: Conférence générale des poids et mesures) is the intergovernmental organization through which member states act on matters related to metrology and measurement standards. The CGPM is made up of delegates of the governments of the member states.

2.1.2. International Committee for Weights and Measures (CIPM)

The International Committee for Weights and Measures (CIPM: Comité international des poids et mesures) has the responsibility for promoting worldwide uniformity in units of measurement. It consists of eighteen individuals elected by the CGPM.

The CIPM established the Mutual Recognition Arrangement (MRA: Arrangement de reconnaissance mutuelle) as the framework for the mutual acceptance of national measurement standards and for recognition of the validity of calibration and measurement certificates issued by national metrology organizations [4,5].

2.1.3. Regional Metrology Organizations (RMOs)

Regional Metrology Organizations are regional associations of national metrology organizations with responsibilities that include proposing CIPM studies, carrying out RMO studies, and reviewing the quality systems and claimed measurement capabilities of their member states [6].

2.1.4. International Bureau of Weights and Measures (BIPM)

The International Bureau of Weights and Measures (BIPM: Bureau International des Poids et Mesures) is the Secretariat of the CIPM, providing the technical and logistical assistance for the preparation and conduct of the various CGPM, CIPM, and associated meetings and sessions [7]. The BIPM hosts the web-presence of the CIPM, including the Key Comparison Data Base (KCDB) that facilitates public access to the results of CIPM-sponsored interlaboratory studies [8]. The BIPM has conference and some laboratory facilities based at the Pavillon de Breteuil in Saint-Cloud, France, (adjacent to Sèvres). This site is considered international territory.

BIPM staff participate in and/or coordinate several types of chemistry-related international comparison studies, including gas mixture and organic purity analysis. The BIPM also conducts an ongoing key comparison of ozone reference photometers.

2.1.5. Consultative Committees (CCs)

The CIPM has consultative committees (CCs) to assist its work in various measurement science domains. The president of each CC is usually a member of the CIPM. Membership in most of the CCs is open to representatives from the national metrology organizations of CGPM member states. The CCs "are charged with planning and execution of key comparisons, and affirming the validity of the results" [9]. The CIPM currently has ten CCs.

2.1.6. Consultative Committee for the Amount of Substance: Metrology in Chemistry and Biology (CCQM)

The Consultative Committee for the Amount of Substance: Metrology in Chemistry and Biology (CCQM: originally the Comité consultatif pour la quantité de matière) is responsible for "developing, improving and documenting the equivalence of national standards (certified reference materials and reference methods) for chemical and biological measurements. It advises the CIPM on matters related to chemical and biological measurements including advice on the BIPM scientific programme activities" [10].

Within the CCQM, responsibilities for the diverse types of chemical and biological measurands are spread among various Working Groups.

2.1.7. CCQM Working Groups (WGs)

The CCQM currently has nine Working Groups (WGs) that conduct international studies. These WGs have the responsibility to "carry out Key Comparisons, and where necessary pilot studies, to critically evaluate and benchmark claimed competences for measurement standards and capabilities" and to "assist in identifying and establishing inter-laboratory work to improve the SI traceability" of measurement results [10].

2.1.7.1. Working Group on Cell Analysis (CAWG)

The Working Group on Cell Analysis (CAWG) responsibilities include the identification and quantification of intact cells and cell properties indicative of function as a result of emergent behavior in complex matrices and mixtures [10].

2.1.7.2. Working Group on Electrochemical Analysis (EAWG)

The Working Group on Electrochemical Analysis (EAWG) responsibilities include "pH, electrolytic conductivity measurements and coulometry" [10]. The EAWG is closely allied with the IAWG.

2.1.7.3. Working Group on Gas Analysis (GAWG)

The Working Group on Gas Analysis (GAWG) responsibilities include "gas composition (including binary and multicomponent mixtures); gas/liquid mixture composition;

nanoparticle and aerosol concentration; isotope ratio measurement; [and] concentration of dissolved gases in liquid or solid matrices” [10].

2.1.7.4. Working Group on Inorganic Analysis (IAWG)

The Working Group on Inorganic Analysis (IAWG) responsibilities include “amount of substance fraction or mass fraction measurements of the elements; cations and anions; inorganic compounds; and organo-metallic compounds. Matrices to be covered include pure materials, calibration solutions and complex samples such as those used for matrix reference materials.” [10].

2.1.7.5. Working Group on Isotope Ratios (IRWG)

The Working Group on Isotope Ratio (IRWG) responsibilities currently include definition of delta scales, relative isotope ratio measurements, C and N isotope ratio measurement, and understanding calibration modalities used in metal isotope ratio characterization [11].

2.1.7.6. Working Group on Nucleic Acid Analysis (NAWG)

The Working Group on Nucleic Acid (NAWG) responsibilities currently include the analysis of nucleic acid polymer sequences, their modifications, and their abundance [12].

2.1.7.7. Working Group on Organic Analysis (OAWG)

The Working Group on Organic Analysis (OAWG) responsibilities include “well-defined organic molecular entities”, excluding “gaseous compounds, organometallic compounds, and large bio-molecules” [10].

2.1.7.8. Working Group on Protein Analysis (PAWG)

The Working Group on Protein Analysis (PAWG) responsibilities include “proteins and peptides” [10].

2.1.7.9. Working Group on Surface Analysis (SAWG)

The Working Group on Surface Analysis (SAWG) responsibilities include “spatially resolved chemical surface analysis at the micro and nanoscale” [10].

2.2. Study Types

The CCQM conducts or supports RMOs in conducting the three types of interlaboratory measurement comparisons recognized in the CIPM MRA: Key Comparisons (KCs), Supplementary Comparisons (SCs), and pilot studies (PSs).

2.2.1. Key Comparisons (KCs)

Officially, CCQM-sponsored Key Comparisons (KCs) "test the principal techniques and methods in the field" and when sponsored by one of the RMOs they "extend the coverage of CIPM key comparisons regionally" [13]. Since there are few (if any) national measurement standards for chemical and biological measurands, in practice KCs for these measurands test the measurement capabilities of the participants. Results from completed KCs are published in the KCDB with full attribution [8] and are used to support the Calibration and Measurement Capabilities (CMCs) claims [14] of the participants.

2.2.2. Supplementary Comparisons (SCs)

Supplementary Comparisons (SCs) are intended to "meet needs not covered by key comparisons" [13]. SCs are sponsored by RMOs and typically address measurement issues primarily of importance to the smaller and less experienced national metrology organizations within the RMO's geographical region of responsibility. Like KCs, results from completed SCs are published in the KCDB with full attribution and are used to support CMCs.

2.2.3. Pilot studies (PSs)

Pilot studies (PSs) "establish measurement parameters for a 'new' field or instrument, or as a training exercise" [13]. Unlike KCs and SCs, results from PSs need not be made publicly available. The results from some PSs are held confidential and have been disclosed only to the study's participants. However, with the agreement of all participants (and the sponsoring body) results can be published. There is no standard format for published pilot studies (PPS); the information disclosed ranges from anonymous summaries to complete and fully attributed results.

In the remainder of this document, "PS" refers to unpublished pilot studies.

2.3. Study Participants

Participation in CIPM key and supplementary comparisons is restricted to organizations within the national or international metrology systems of signatories to the CIPM MRA. Direct participation in CCQM comparisons is further restricted to organizations with recognized experience and expertise with the subject matter measurements; these typically are from the larger national economies. Less experienced organizations participate through RMO-sponsored studies.

2.3.1. National Metrology Institutes (NMIs)

National metrology institutes (NMIs) have the responsibility for all of a nation's measurement infrastructure not specifically delegated to another organization.

2.3.2. Designated Institutes (DIs)

Designated institute (DIs) have the responsibility for specified aspects of a national or international measurement infrastructure.

2.3.3. International Agencies

There are three international agencies that have participated in CCQM studies. The European Commission – Joint Research Centre (JRC), once known as the Institute for Reference materials (IRMM), has interests and expertise in many fields; it can be regarded as an “international NMI.” The International Atomic Energy Agency (IAEA) participates in selected IRWG and IAWG studies; it can be regarded as a DI. The World Meteorological Organization (WMO) sponsors laboratories to participate as DIs in GAWG studies relevant to their expertise.

2.4. *CCQM_Retrospectoscope* Jargon

2.4.1. BaseUnit

A “BaseUnit” is the unit used to report a measurement value stripped of prefixes: M, d, c, m, μ, n, p, and f. With apologies to the CGPM, the keepers of the Système international d'unités (SI), “g” is therefore used rather than “kg” as the BaseUnit of mass.

The BaseUnits of greatest interest in this report are mass fraction (g/g), mole fraction (mol/mol), and hydrogen ion activity (pH).

2.4.2. Dataset

A *CCQM_Retrospectoscope* dataset consists of the results (values and associated uncertainties) for one measurand reported by a study’s participants, a unique name that identifies the measurand, the participant names, a reference value and its associated uncertainty, the BaseUnit for the results, the measurement year of the study, and the WG that conducted the study. Many studies provide results for multiple measurands; therefore one study may generate multiple datasets.

2.4.3. Measurand

“Measurand” is formally defined as the “quantity intended to be measured” [15§2.3]. In chemical practice, this requires specifying the name of the substance of interest (the analyte), the units of measurement, the sample matrix, and perhaps other information relevant to defining the intended quantity such as the measurement processes used in making the measurement. Note that “analyte” is metrologically just “the name of a chemical substance or one of its components” specified in a measurand [16§1.7]. Within the *CCQM_Retrospectoscope* system, the measurand is specified by the analyte, the BaseUnit, and the WG. The dataset name typically identifies the sample matrix and may provide other information, but this information is not otherwise associated with the measurand.

2.4.4. Measurement Year

The “measurement year” is nominally the year in which participant measurements were made. Since many studies take place over more than one calendar year, the measurement year of a dataset is the date by which the study’s “results must be reported by.” While specified in the study protocol, this date is frequently extended when there are external delays.

2.4.5. Study Identifiers

All CCQM studies are provided by the BIPM with a multi-part code of the form X.QM-YZ where X identifies the organizing body; Y identifies the study type (“K” for Key Comparisons, “S” for Supplementary Comparisons, and “P” for pilot studies), and Z is a one-to-three-digit study number followed by a variety of version designators as needed. While unambiguous, the official codes do not distinguish between published and confidential pilot studies, and the variability in the number of study digits complicates sorting.

The CCQM_Retrospectoscope database uses a standardized version of the official code. The variable length X (organizing body) is reduced to two letters, “P” is used as the study type only for confidential PSs, “Q” is used as the study type for published PSs, and the variable number of digits in Z (the study number) is always three digits with leading zeros as needed. For example:

- CCQM-K1.a, CCQM’s first Key Comparison, is transformed to CCQM-K001.a,
- CCQM-K10.2018 is transformed to CCQM K010.2018,
- CCQM-P1, the first CCQM pilot study but completed without a publicly available report, is transformed to CCQM-P001,
- CCQM-P55.2.2018 is transformed to CCQM Q055.2.2018,
- BIPM.QM-K1 is transformed to BIQM K001, and
- EUROMET.QM-S3 is transformed to EUQM-S003.

3. Analysis Scope

To ensure that results reflecting measurement capabilities can be independently verified, all analyses of quantitative results reported herein use only information from KC, SC, and PPS reports that were publicly available at the time of this document's publication. Information from non-public PSs is limited to the fact of an organization's participation and not the quantitative quality of that participation.

3.1. Available Data

Beginning in 1992 and as of this document's publication date, NIST has reported quantitative results in 492 public datasets from 165 CCQM-related KCs, SCs, or PPSs as well as 273 datasets from 78 PSs. Table 1 summarizes these studies and datasets by {WG, BaseUnit}, study type, and sponsoring organization. The proportions of datasets by {WG, BaseUnit} are displayed in Fig. 1.

Table 1. Number of Published Studies and Datasets by Focus Area and Study Type.

	Number Datasets					Number Studies				
{WG, BaseUnit}	All	KC	SC	PPS	PS	All	KC	SC	PPS	PS
{IAWG, g/g}	239	115		15	109	81	41		6	34
{OAWG, g/g}	188	81		15	92	65	33		7	25
{GAWG, mol/mol}	172	132	29	9	2	56	48	1	6	1
{EAWG, pH}	39	37			2	11	9			2
{SAWG, m}	26	8		18		3	2		1	
{PAWG, g/g}	22	1		3	18	5	1		1	3
{SAWG, a.u.}	16				16	1				1
{IRWG, ‰}	13			10	3	4			2	2
{EAWG, S/m}	10	3			7	7	3			4
{IRWG, n/n}	8	8				1	1			
{NAWG, bp}	8				8	1				1
{NAWG, n/L}	8				8	2				2
{OAWG, mol/mol}	5				5	1				1
{SAWG, mol/mol}	4	4				1	1			
{NAWG, g/L}	3	2			1	2	1			1
{IRWG, g/mol}	2				2	1				1
{NAWG, n/n}	2	2				1	1			
Total	765	393	29	70	273	243	141	1	23	78

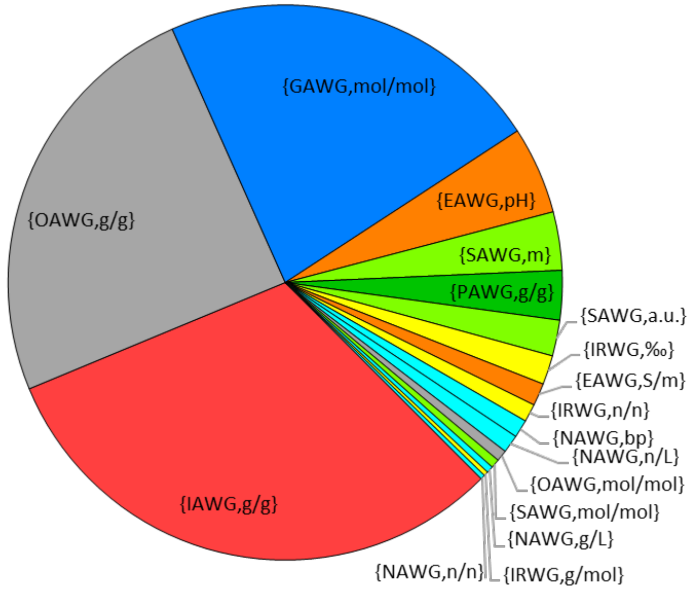


Fig. 1. Proportion of NIST's CCQM Datasets by {Working Group, BaseUnit} 1992 –2023.

3.2. Comparison Intervals

Documenting changes in measurement performance over time requires that information be partitioned into approximately equally representative sets. Only {GAWG, mol/mol}, {OAWG, g/g}, {IAWG, g/g}, and {EAWG, pH} have quantitative datasets from enough public studies to support "Early" versus "Recent" partitions. The cumulative distributions of the number of public studies and datasets over time are displayed in Fig. 2.

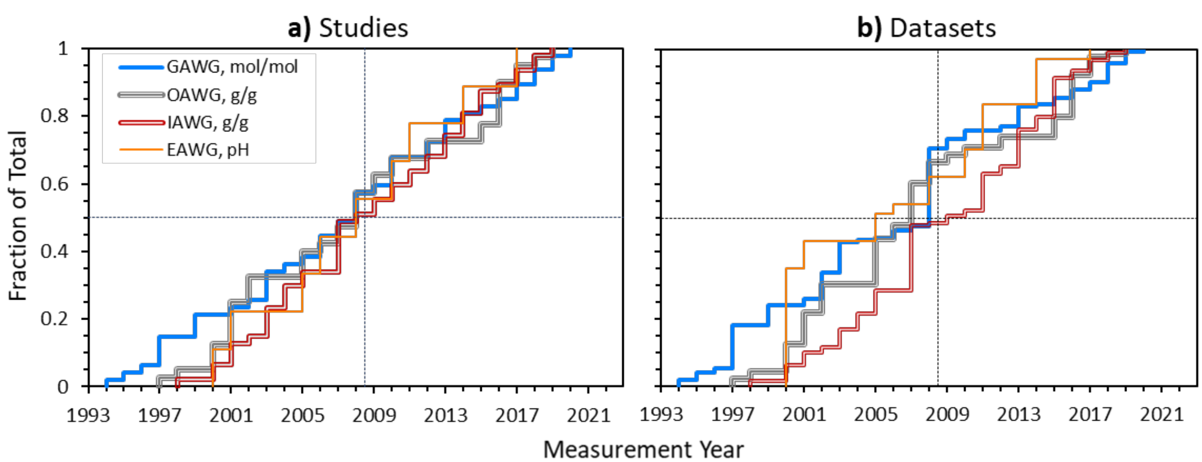


Fig. 2. Cumulative Distribution of NIST Participations in CCQM Studies and Datasets.

a) Cumulative distribution of {GAWG, mol/mol}, {IAWG, g/g}, {OAWG, g/g}, and {EAWG, pH} finalized studies that NIST participated in, 1992 through 2023. b) Cumulative distribution of {GAWG, mol/mol}, {IAWG, g/g}, {OAWG, g/g}, and {EAWG, pH} publicly accessible datasets containing quantitative NIST results, 1992 through 2023. Horizontal lines represent the 0.5 fraction of total studies or datasets; vertical lines represent the measurement year used to define the "Early" and "Recent" partitions (offset by -0.5 year for graphical clarity.)

No measurement year provides perfectly balanced measurement partitions for the four {WG, BaseUnit}s. The first publicly accessible results for the different WGs were in measurement years ranging from 1992 (GAWG) to 2000 (EAWG) and the most recently completed studies have measurement years ranging from 2017 (EAWG) to 2020 (GAWG). However, measurement year 2009 divides the four {WG, BaseUnit}s into approximately decade-long Early and Recent groups having roughly equal numbers of studies. The Early group (1992 through 2008) contains between 50 % (IAWG) and 70 % (GAWG) of the currently public datasets, but there are enough Recent (2009 through 2020) datasets for analysis.

The proportions of the NIST's publicly accessible datasets and studies in the Early and Recent intervals are displayed in Fig. 3.

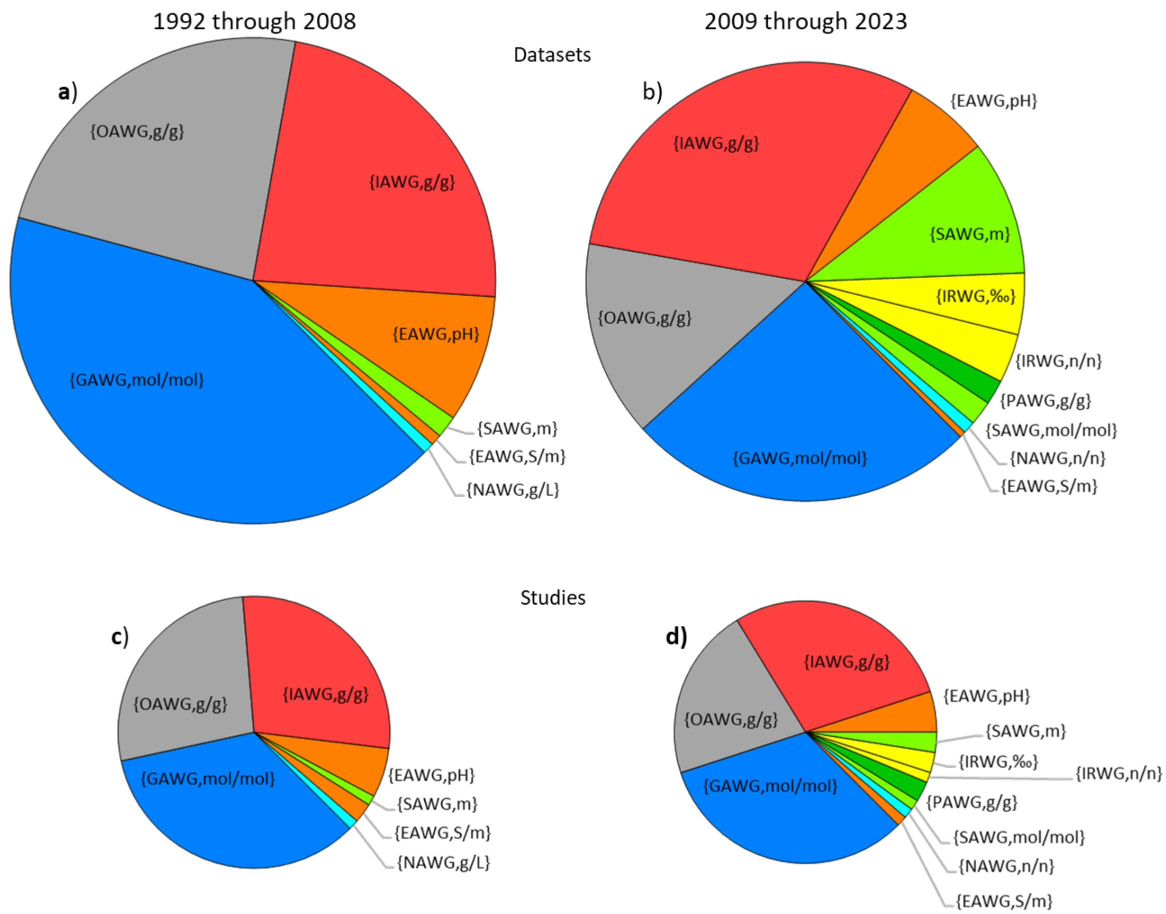


Fig. 3. Proportion of NIST's CCQM Early and Recent Datasets and Studies by {WG, BaseUnit}.

a) Proportions of the 271 publicly-accessible datasets NIST contributed to during measurement years 1992 – 2008.
b) Proportions of the 221 publicly-accessible datasets NIST contributed to during measurement years 2009 – 2023.
c) Proportions of the 85 publicly-accessible finalized studies NIST participated in during measurement years 1992 – 2008.
d) Proportions of the 80 publicly-accessible finalized studies NIST participated in during measurement years 2009 – 2023.
Areas of the pies are proportional to the total number of datasets (panels a and b) or studies (panels c and d).

4. Metrological Responsibility Peers

The NMIs, DIs, and international organizations that participate in CCQM studies vary widely in how they are organized and in their range of metrological responsibilities. While an NMI typically has broad responsibilities for its nation's metrological infrastructure, some or all of the responsibility for chemical and biological metrology may be delegated to DI(s).

Regardless of their names or metrological status, the studies in which organizations participate identify peers with chemical/biological responsibilities similar to NIST's. The radar plots of Fig. 4 include participations in PSs as well as KCs, SCs, and PPSs.

A participation rate for a given organization, i, in a given WG, $p_{\mathrm{WG},i}$,is the ratio (expressed as a percentage) of the number of studies in which the particular organization participated, n_i, relative to the number of studies conducted by the WG, n_{WG}:

$$p_{\mathrm{WG},i} = 100\, n_i / n_{\mathrm{WG}} \,. \tag{1}$$

The similarity of the i^{th} organization's participation rates to NIST's, $p_{\mathrm{WG,NIST}}$, is estimated as a composite distance, $d_{\mathrm{MR},i}$, across all WGs that conducted studies during the interval

$$d_{\mathrm{MR},i} = N_{\mathrm{unshared}} + \sqrt{\sum_{WG=1}^{N_{\mathrm{shared}}} \left(p_{WG,i} - p_{WG,\mathrm{NIST}}\right)^2} / N_{\mathrm{shared}} \;\; i \tag{2}$$

where N_{unshared} is the number of WGs in which either NIST participates and organization i does not or organization i participates and NIST does not and N_{shared} is the number of WGs in which both NIST and organization *i* participate. This definition gives greater weight to the number of different responsibilities than to differences within shared responsibilities.

During the Early interval, only NIM exactly matched NIST's participation in all eight of the eight WGs that conducted publicly available studies. NMIJ and KRISS participated in seven of the eight of the WGs. PTB, JRC, BAM, NMISA, NPL and NIMT participated in six of the eight (although NIMT participated in very few of the studies within each WG). Other organizations differed from NIST's pattern of participation by three or more WGs.

During the Recent interval, PTB, KRISS, VNIIM, NIM, and UME matched NIST's participation in all nine of the WGs conducting studies. INMETRO, NMIJ, and BAM participated in eight of the nine WGs and CENAM participated in seven of the nine. Other organizations differed from NIST's pattern of participation by three or more WGs.

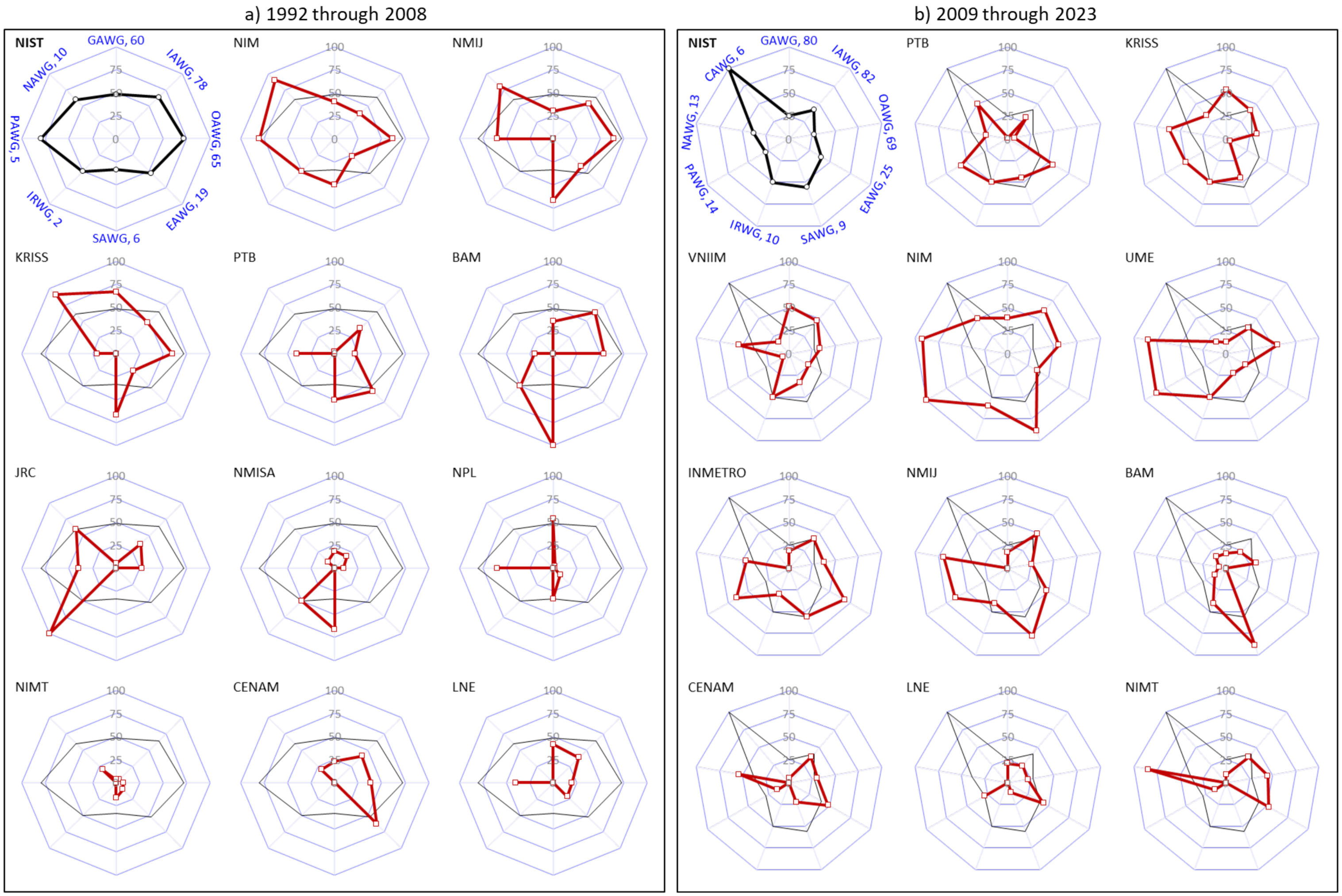


Fig. 4. NIST's Early and Recent Metrological Responsibility Peers.

a) Working Group (WG) participation rates for NIST and 11 organizations during the Early interval. b) WG participation rates during the Recent interval. The text in the upper-left corner of each segment is the organization's code name (see Appendix B for full descriptions). The vertices of the outer polygon represent participation in every available study. The inner polygons denote WG participation rates of 0.75, 0.50, and 0.25. The center denotes no participation. The segment to the top left of each multiplot displays NIST's rate as open black circles connected with a thick black line; the thin black lines in the other segments echo this pattern. The rates for other organizations are displayed as open red squares connected by a thick red line. The vertices of the top-left segments identify the WG and the number of studies the WG conducted during the interval.

5. Participation Peers

While the pattern of participation in the various WGs' studies identifies organizations with a similar range of national responsibilities, participation in similar proportions of studies conducted by a given WG identify organizations with similar levels of interest in a given chemical or biological arena.

5.1. {GAWG, mol/mol}, {IAWG, g/g}, {OAWG, g/g}, and {EAWG, pH} Participations

The participants in {GAWG, mol/mol}, {IAWG, g/g}, {OAWG, g/g}, and {EAWG, pH} studies during the Early and Recent intervals are depicted in Fig. 5. Each panel displays all organizations that participated in a) at least 20 % of the number of studies of the most active organization and b) at least three studies. The number of such organizations increased with time in all four arenas, but the pattern of changes amongst the most similar to NIST are arena-specific:

- {GAWG, mol/mol}
  NIST participated in roughly the same number of studies during the Early and Recent intervals. KRISS, VNIIM, VSL, NPL, NIST, LNE, and NIM were the most active organizations in the Early interval. KRISS, VNIIM, NPL, VSL, NIM, NMISA, and NIST were the most active organizations in the Recent interval. The only participation-peer group differences are the marked increased participation by NMISA and the somewhat decreased participation by LNE.
- {IAWG, g/g}
  The IAWG conducted more KCs and many fewer pilot studies (PSs and PPSs) during the Recent interval than in the Early. NIST, BAM, NMIJ, and KRISS were the most active organizations in KCs during the Early interval. NIM, NMIJ, VNIIM, INMETRO, NIST, KRISS, CENAM, and UME were then most active during the Recent interval. The major participation-peer group differences are increased participation by INMETRO and UME and decreased participation by BAM.
- {OAWG, g/g}
  As with the IAWG, the OAWG conducted more KCs and many fewer pilot studies during the Recent interval than in the Early. NIST, NMIJ, NIM, KRISS, and LGC were the most active organizations in KCs during the Early interval. While NIST, LGC, and NMIJ all participated in more KCs in the Recent interval than in the Early, they participated in only about half as many studies as the most active organizations: NIM and UME.
- {EAWG, pH}
  The EAWG studies on pH deal with essentially a single analyte in a limited set of aqueous buffer systems. As such, most of the organizations having national responsibility for pH measurements have participated in most of the relatively few studies during both intervals.

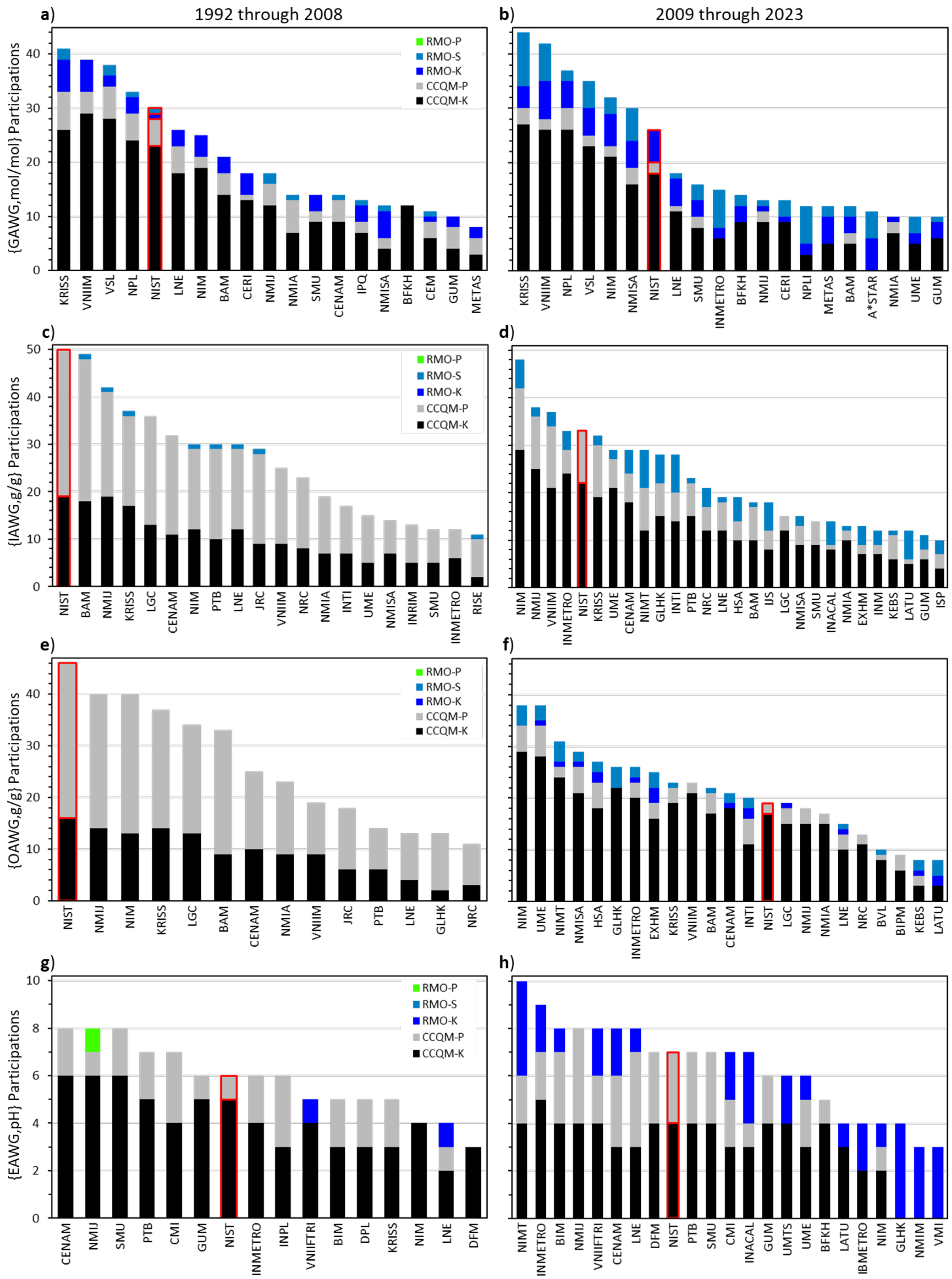


Fig. 5. Early and Recent Participations in GAWG, IAWG, OAWG, and {EAWG, pH} Studies.

Panels a, c, e, and g identify the participants and their number of participations in finalized {GAWG, mol/mol}, (IAWG, g/g}, {OAWG, g/g}, and {EAWG, pH} studies during the Early interval; panels b, d, f, and h display the participants and their number of participations during the Recent interval. Study types are identified in the panel legends. NIST's location within each panel is outlined in red.

5.2. Participation Peers in Other Working Groups

The participants in {EAWG, S/m} conductivity studies and in all SAWG, IRWG, PAWG, NAWG, and CAWG studies from 1992 through 2023 are depicted in Fig. 6. Each panel displays all organizations that participated in at least three studies. NIST has not been among the most active organizations in the EAWG conductivity, SAWG, IRWG, PAWG, or NAWG studies. It has been the most active in CAWG studies.

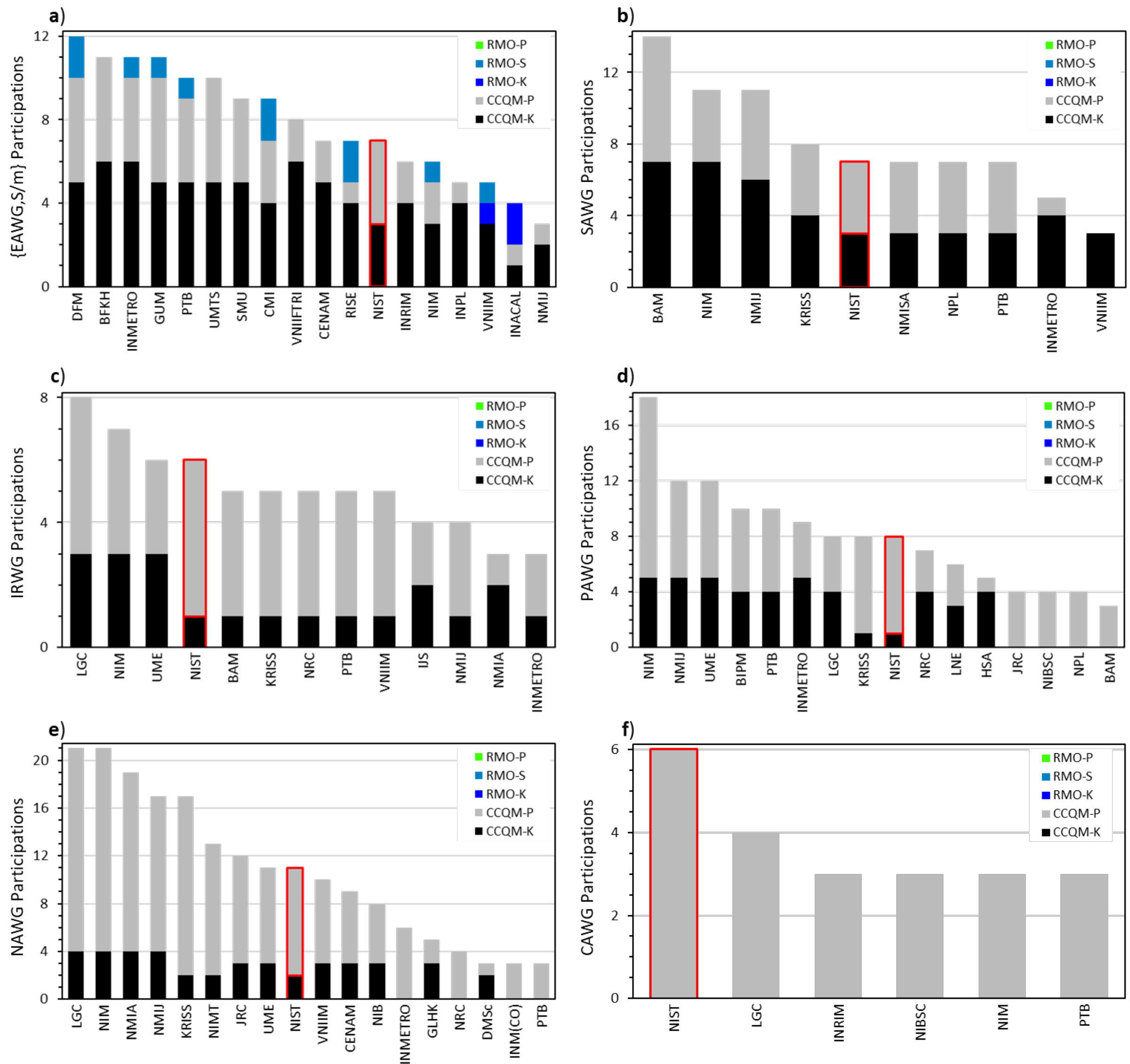


Fig. 6. Total Participations in {EAWG, S/m}, SAWG, IRWG, PAWG, NAWG, and CAWG Studies.

a) Participants and their number of participations in finalized {EAWG, S/m} studies from 1992 to 2023,
b) Participants and their number of participations in finalized SAWG studies with all types of BaseUnit from 1992 to 2023,
c) Participants and their number of participations in finalized IRWG studies with all types of BaseUnit from 1992 to 2023,
d) Participants and their number of participations in finalized PAWG studies with all types of BaseUnit from 1992 to 2023,
e) Participants and their number of participations in finalized NAWG studies with all types of BaseUnit from 1992 to 2023,
f) Participants and their number of participations in finalized CAWG studies with all types of BaseUnit from 1992 to 2023.
Study types are identified in the panel legends. NIST's location within each panel is outlined in red.

6. Coordination Peers

While participation in studies conducted by a given WG identifies organizations with metrological responsibilities in a given chemical or biological arena, coordination of those studies indicates acknowledged expertise and leadership.

Note: Unlike participations, coordinations can be and often are fractional counts. When two or more organizations co-coordinate a study, the *CCQM_Retrospectoscope* system divides one “coordination credit” equally among the organizations [3]. The way participations and coordinations are counted also differs when the same measurement protocol and sample materials are used in both a KC and a “parallel” PS. Organizations that participate in the KC may also participate in the PS, reporting values obtained using different measurement procedures and/or analysts; these count as two participations. Parallel PSs do not (generally) create much additional work for the coordinating organizations and so do not contribute to the coordination totals.

6.1. {GAWG, mol/mol}, {IAWG, g/g}, {OAWG, g/g}, and {EAWG, pH} Coordinations

The coordinators of {GAWG, mol/mol}, {IAWG, g/g}, {OAWG, g/g}, and {EAWG, pH} studies during the Early and Recent intervals are depicted in Fig. 7. Each panel displays all organizations that coordinated (or co-coordinated) any study during the interval. The number of such organizations increased slightly with time in all four arenas, but the change patterns are arena-specific:

- {GAWG, mol/mol}
  The same seven organizations coordinated most of the studies in both intervals: VSL, NPL, KRISS, NMIJ, VNIIM, BIPM, and NIST. NIST coordinated more studies in the Recent interval than in the Early interval.
- {IAWG, g/g}
  NIST, JRC, BAM, LGC, SMU, NMIJ, NRC, NIM, and EMPA coordinated most of the IAWG KCs during the Early interval. JRC and EMPA have not been active in the CCQM during much of the Recent interval and the NRC has focused on supporting RMO studies. While BAM has remained active, (KRISS, GLHK, VNIIM, and PTB) join (NIM, NMIJ, LGC, NIST, and SMU) as the most frequent KC coordinators during the Recent interval.
- {OAWG, g/g}
  NIST, LGC, KRISS, BAM, CENAM, NMIA, and NRC were the only coordinators of OAWG KCs during the Early interval. BIPM, NIST, NIM, UME, GLHK, and LGC have been the most frequent KC coordinators during the Recent interval.
- {EAWG, pH}
  PTB, SMU, and NIST were the only {EAWG, pH} KC coordinators during the Early interval. PTB, NIST, NMIJ, and VNIIFTRI have been the only KC coordinators during the Recent interval.

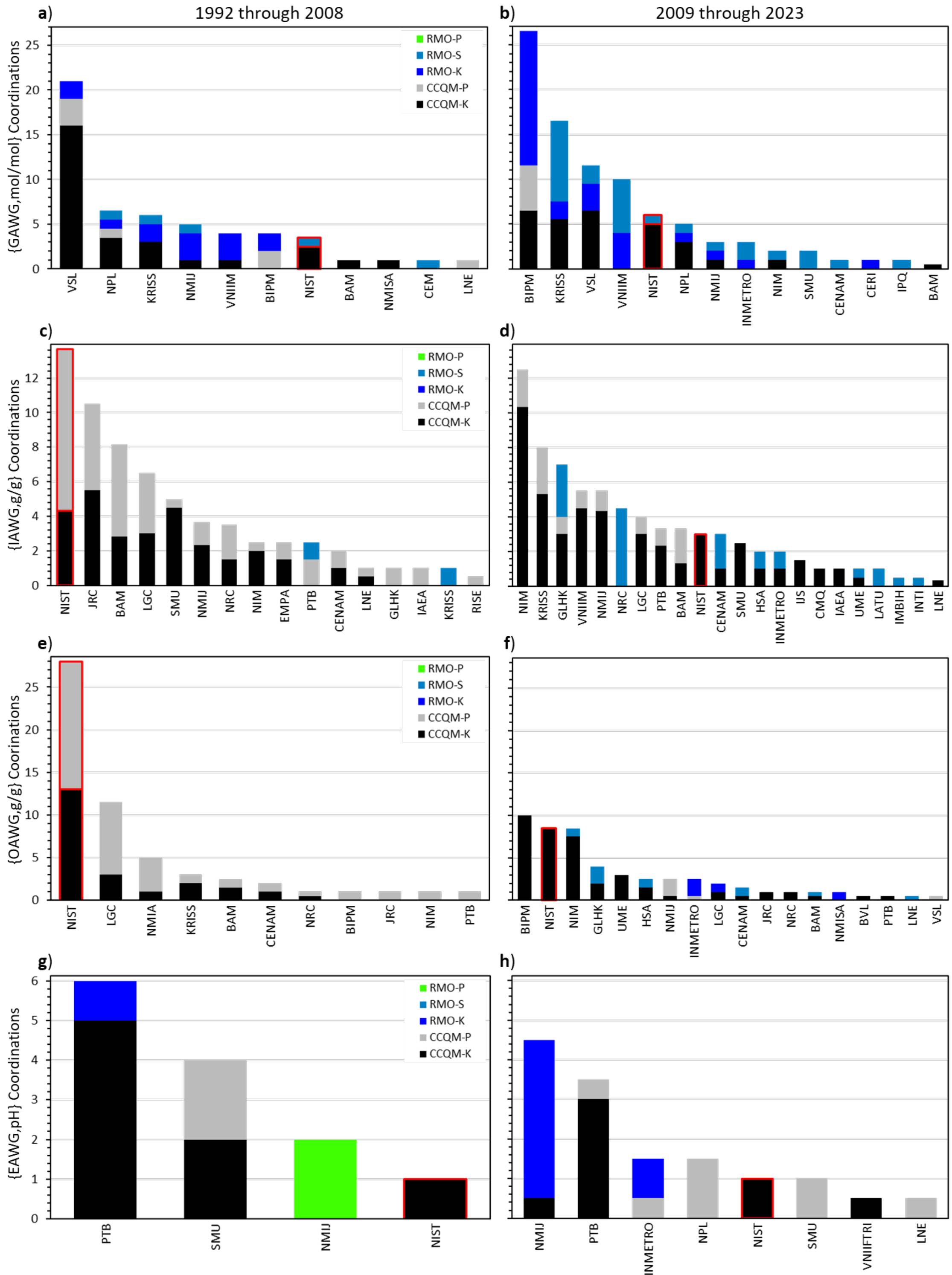


Fig. 7. Early and Recent Coordinations of GAWG, IAWG, OAWG, and {EAWG, pH} Studies.

Panels a, c, e, and g identify the coordinators and their number of coordinations in finalized {GAWG, mol/mol}, (IAWG, g/g}, {OAWG, g/g}, and {EAWG, pH} studies during the Early interval; panels b, d, f, and h display the coordinators and their number of coordinations during the Recent interval. Study types are identified in the panel legends. NIST's location within each panel is outlined in red.

6.2. Coordination Peers in Other Working Groups

The coordinators of {EAWG, S/m} conductivity studies and all SAWG, IRWG, PAWG, NAWG, and CAWG studies from 1992 through 2023 are depicted in Fig. 8. Each panel displays all organizations that coordinated any study. NIST has not coordinated any {EAWG, S/m} study nor has it coordinated any SAWG, IRWG, or PAWG KC. (NIM, JRC, LGC, NIST, and GLHK) have coordinated most of the NAWG studies. (NIST, PTB, and NIBSC) have been the most frequent coordinators of CAWG pilot studies.

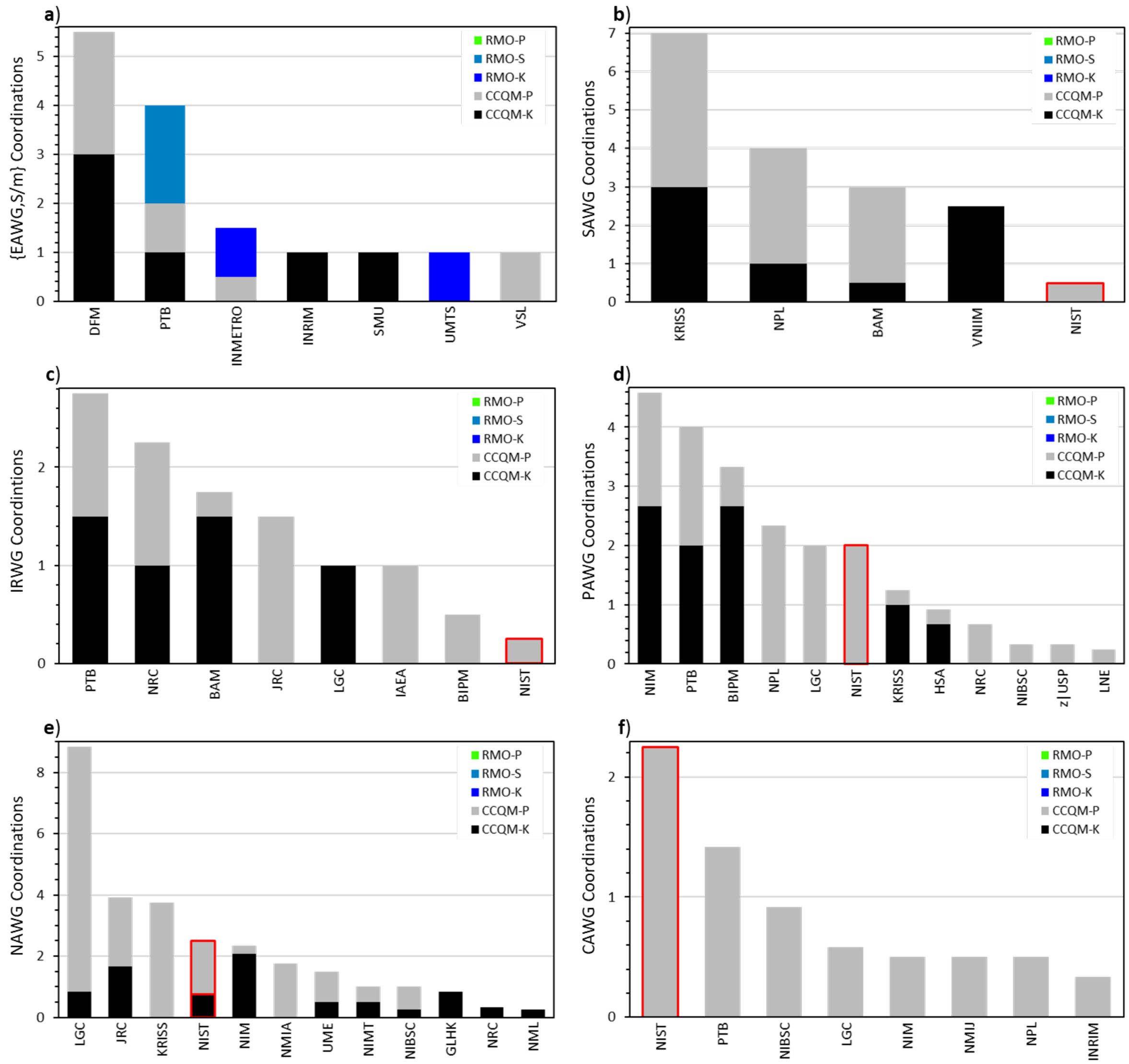


Fig. 8. Total Coordinations in {EAWG, S/m}, SAWG, IRWG, PAWG, NAWG, and CAWG Studies.

a) Coordinators and their number of coordinations in finalized {EAWG, S/m} studies from 1992 to 2023,
b) Coordinators and their number of coordinations in finalized SAWG studies with all types of BaseUnit from 1992 to 2023,
c) Coordinators and their number of coordinations in finalized IRWG studies with all types of BaseUnit from 1992 to 2023,
d) Coordinators and their number of coordinations in finalized PAWG studies with all types of BaseUnit from 1992 to 2023,
e) Coordinators and their number of coordinations in finalized NAWG studies with all types of BaseUnit from 1992 to 2023,
f) Coordinators and their number of coordinations in finalized CAWG studies with all types of BaseUnit from 1992 to 2023.
Study types are identified in the panel legends. NIST's location within each panel is outlined in red. Note: NIST never coordinated an {EAWG, S/m} conductivity study.

7. Measurement Similarity Peers

Similar metrological responsibilities and participation rates do not measure the degree of similarity between two organizations' measurement processes. However, quantitative estimates of measurement similarity are provided by comparing co-participant results from the same suite of datasets.

Measurement similarity estimates become more representative the larger the number of in-common datasets. However, the larger the suite the fewer organizations for which these estimates can be calculated. Defining the minimum size of the suite to be 25 % of the datasets containing a NIST result is an empirical compromise that provides estimates for a tractable number of organizations using a reasonable diversity of datasets.

7.1. Metrics

7.1.1. Median Relative Difference to NIST Result

Estimating an organization's typical measurement difference relative to NIST measurements using results from a diverse collection of datasets requires that the results be normalized to have a common location and scale. The summary transform used here is related to zeta-score (ζ-score) standardization [17§9.6]

$$X_{\mathrm{NIST},i} = \mathrm{Median}_{25\%}\left\{|x_i - x_{\mathrm{NIST}}|/\sqrt{u^2(x_i) + u^2(x_{\mathrm{NIST}})}\right\} \tag{3}$$

where x_i is a result from organization *i* in a given dataset, $u(x_i)$ is the standard uncertainty of that result, x_{NIST} is the dataset reference value, $u(x_{\mathrm{NIST}})$ is the standard uncertainty of that value, and $\mathrm{Median}_{25\%}\{\cdot\}$ is the function "take the median of the transformed values for organizations contributing to at least 25 % of the datasets containing a NIST result." Organizations that most of the time report values essentially equal to NIST's will have $X_{\mathrm{NIST},i}$ values located near zero while organizations routinely reporting values very different from NIST's (either smaller or larger) will likely have $X_{\mathrm{NIST},i}$ greater than two combined standard uncertainties distant from NIST's.

7.1.2. Median Uncertainty Relative to NIST Uncertainty

Assessing an organization's typical uncertainty assessment capability relative to NIST's likewise requires that standard uncertainties from the dataset collection be normalized to have a common origin and scale. The summary transform used here is:

$$Y_{\mathrm{NIST},i} = \mathrm{Median}_{25\%}\{u(x_i)/u(x_{\mathrm{NIST}})\}\ . \tag{4}$$

Organizations that most of the time report standard uncertainties essentially equal to NIST's will have $Y_{\mathrm{NIST},i}$values located near one. Given the rule-of-thumb that conscientious uncertainty estimates may differ by about a factor of two, organizations routinely reporting uncertainties much smaller than NIST's will likely have $Y_{\mathrm{NIST},i}$ somewhat less than 0.5 and organizations routinely reporting uncertainties much larger than NIST's will likely have $Y_{\mathrm{NIST},i}$ somewhat greater than 2.

7.1.3. Similarity Distance to NIST

The $X_{\text{NIST},i}$ and $Y_{\text{NIST},i}$ summary estimates can be combined into a Euclidean distance, $d_{\text{NIST},i}$, that summarizes similarity between the measurements reported by NIST and organization *i*:

$$d_{\text{NIST},i} = \sqrt{\left(X_{\text{NIST},i}\right)^2 + \left(\log_2\{Y_{\text{NIST},i}\}\right)^2} \quad (5)$$

where $\log_2\{\cdot\}$ is the binary logarithm function. This logarithm transforms the $Y_{\text{NIST},i}$ to have numerical values compatible with those of the $X_{\text{NIST},i}$: the reference $Y_{\text{NIST},i}$ value of 1 becomes zero ($\log_2\{1\} = 0$) and the factor-of-two scale becomes ± 1 ($\log_2\{0.5\} = -1$ and $\log_2\{2\} = 1$).

Based on the usual interpretation that about 95 % of "good" measurement difference values should be separated by no more than two combined standard uncertainty units and that most conscientious standard uncertainty estimates are "good" to about a factor of two, NIST's measurement peers are expected to have $d_{\text{NIST},i}$ of less than about two of the composite distance units.

However, the $X_{\text{NIST},i}$ and $Y_{\text{NIST},i}$ are related through the equality

$$\sqrt{u^2(x_i) + u^2(x_{\text{NIST}})} = u(x_{\text{NIST}})\sqrt{1 + \left(\frac{u(x_i)}{u(x_{\text{NIST}})}\right)^2} \,. \quad (6)$$

In consequence, for a given NIST-normalized bias, $|x_i - x_{\text{NIST}}|/u(x_{\text{NIST}})$, the $d_{\text{NIST},i}$ for $u(x_i)$ greater than $u(x_{\text{NIST}})$ by a factor are smaller than the $d_{\text{NIST},i}$ for $u(x_i)$ less than $u(x_{\text{NIST}})$ by the same factor. Examples of this asymmetry are displayed in Fig. 9, where $|x_i - x_{\text{NIST}}|/u(x_{\text{NIST}})$ is fixed at 2 or 4 and $u(x_i)/u(x_{\text{NIST}})$ ranges from 1/8 to 8/1 (i.e., -3 to +3 when $\log_2$-transformed). The asymmetry becomes more pronounced the larger the normalized bias.

The $d_{\text{NIST},i}$ values thus provide indicative estimates of similarity but need to be interpreted with care.

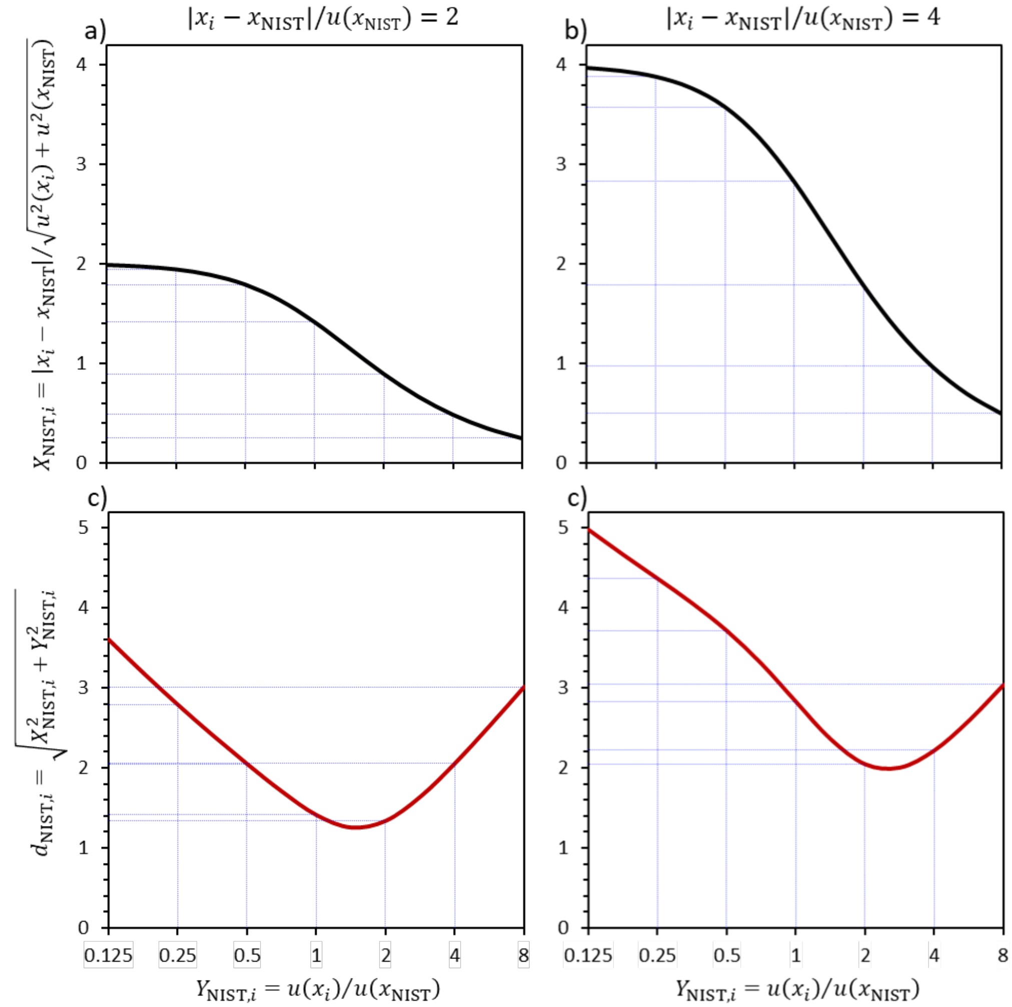


Fig. 9. Normalized Bias and Similarity Distance as Functions of Relative Uncertainty.

a) Normalized bias for NIST-normalized bias of 2 and NIST uncertainty of 1 as a function of relative uncertainty,
b) Normalized bias for NIST-normalized bias of 4 and NIST uncertainty of 1 as a function of relative uncertainty,
c) Similarity distance for NIST-normalized bias of 2 and NIST uncertainty of 1 as a function of relative uncertainty,
d) Similarity distance for NIST-normalized bias of 4 and NIST uncertainty of 1 as a function of relative uncertainty.
Thin guidelines connect factor-of-two relative uncertainties along the *x*-axes with the associated *y*-axis values.

7.2. {GAWG, mol/mol}

Table 2 lists the (non-continuous) GAWG-related studies used to identify NIST's gas mixture measurement peers. This list includes the 48 KC, SC, or PPS studies that produced the 156 quantitative {GAWG, mol/mol} datasets that were available as of this document's publication date. Fourteen datasets from seven continuous BIPM-K1 ozone reference photometer comparisons are not included.

Table 2. {GAWG, mol/mol} Studies in Which NIST Participated.

Study	Description	$\#_{set}$	Year	Coordinator
CCQM-K001.b	Carbon dioxide in nitrogen	3	1994	VSL
CCQM-K001.a	Carbon monoxide in nitrogen	3	1995	VSL
CCQM-K001.c	Nitric oxide (NO) in nitrogen	2	1996	VSL
CCQM-K001.d	Sulfur dioxide in nitrogen	2	1997	VSL
CCQM-K001.e	Natural gas type I	6	1997	VSL
CCQM-K001.f	Natural gas type II	6	1997	VSL
CCQM-K001.g	Natural gas type III	6	1997	VSL
CCQM-K003	Automotive emission gases in nitrogen	3	1999	VSL
CCQM-K004	Ethanol in air	1	1999	NPL
CCQM-K007	BTX in nitrogen	5	1999	NIST
CCQM-K010	BTX in nitrogen	3	2001	NIST,NPL
CCQM-K016.a	Natural gas type IV	12	2002	BAM,VSL
CCQM-K015	SF_6 and CF_4 in nitrogen	2	2003	KRISS
CCQM-K022	VOCs in air	8	2003	NMIJ
CCQM-Q041.1	Methane and CO_2 in air: Measurement capability	2	2003	VSL
CCQM-Q041.2	Methane and CO_2 in air: PSM comparison	2	2003	VSL
CCQM-K026.a	Nitric oxide (NO) in nitrogen	1	2004	NPL
CCQM-K041	Hydrogen sulfide in nitrogen	1	2005	NIST
CCQM-Q028	Ozone	2	2005	BIPM
CCQM-K052	Carbon dioxide in air	1	2006	VSL
CCQM-K054	n-Hexane in methane	1	2006	VSL
CCQM-Q073	Nitric oxide (NO) in nitrogen	1	2006	BIPM
CCQM-K046	Ammonia in nitrogen	1	2007	VSL
CCQM-K053	Oxygen (O_2) in nitrogen	1	2007	KRISS
CCQM-K051	Carbon monoxide in nitrogen	1	2008	NMISA
CCQM-K068	Nitrous oxide (N_2O) in synthetic air	1	2008	KRISS
CCQM-K071	Multi-component gas stack emissions	5	2008	VSL
EUQM-S003	Volatile organic compounds (VOCs) in air	29	2008	NPL
CCQM-K066	Methane purity	4	2009	NMIJ
CCQM-K074	Nitrogen dioxide (NO_2) in nitrogen	1	2010	BIPM
CCQM-K076	Sulfur dioxide in nitrogen	1	2010	NIST
CCQM-Q110.B1	NO_2 in nitrogen by FT-IR spectroscopy	1	2010	BIPM
CCQM-Q110.B2	NO_2 in nitrogen using synthetic spectra	1	2010	BIPM
CCQM-K084	Carbon monoxide in synthetic air at ambient level	1	2012	KRISS
CCQM-K093	Ethanol in nitrogen or air	1	2012	NPL
CCQM-K082	Methane in air: PSM comparison	2	2013	BIPM
CCQM-K083	Halocarbons in dry whole air	6	2013	NIST
CCQM-K101	Oxygen (O_2) in nitrogen	1	2013	NIM
CCQM-K111	Propane in nitrogen	1	2014	VSL
CCQM-K113	Noble gas mixture	3	2015	KRISS
CCQM-K121	Monoterpenes in nitrogen	4	2016	NIST

Study	Description	$\#_{set}$	Year	Coordinator
CCQM-K120.a	Carbon dioxide at background level	1	2017	BIPM,NIST
CCQM-K120.b	Carbon dioxide at urban level	2	2017	BIPM,NIST
CCQM-K010.2018	BTEX in nitrogen	7	2018	NIST
CCQM-K137	Nitric oxide (NO) in nitrogen	2	2018	BIPM
CCQM-K003.2019	Automotive emission gases in nitrogen	4	2019	VSL
CCQM-K117	Ammonia in nitrogen	1	2019	VSL
CCQM-K068.2019	Nitrous oxide (N_2O) in air at ambient levels	1	2020	BIPM,KRISS

The $X_{\mathrm{NIST},i}$ and $Y_{\mathrm{NIST},i}$ values for organizations with results in at least 25 % of the datasets with measurement dates within the relevant Early (1992 through 2008) and Recent (2009 through 2023) intervals are displayed in Fig. 10.

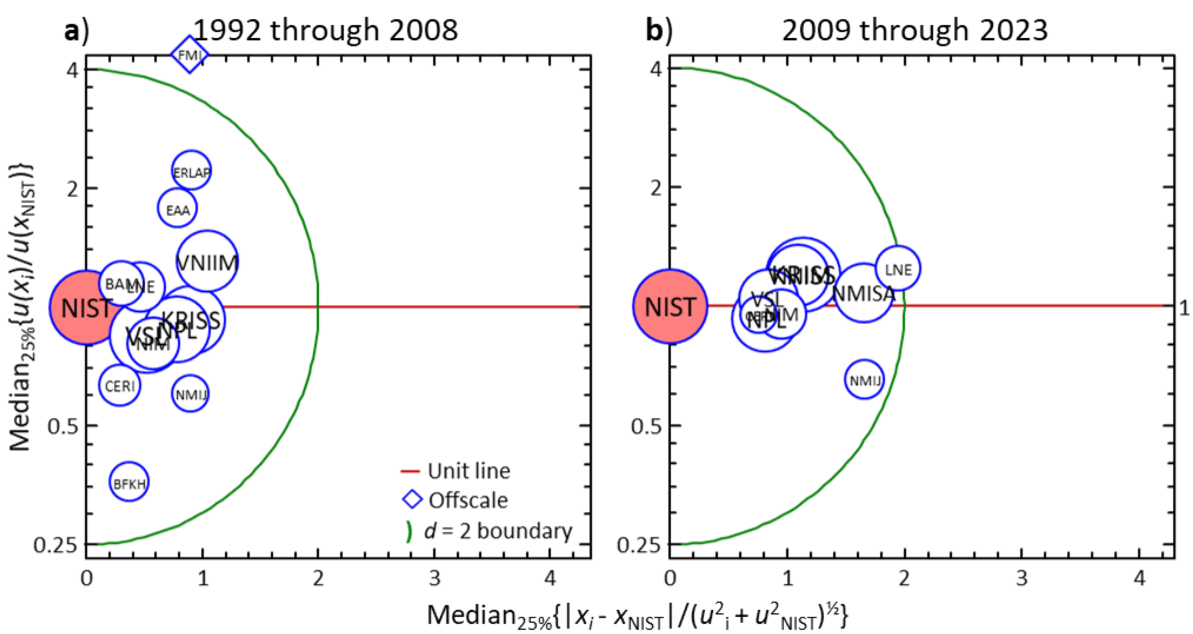


Fig. 10. {GAWG, mol/mol} Measurement Similarity Peers, Early and Recent.

a) Median measurement differences and uncertainties relative to NIST values for organizations that have results in at least 25 % of the {GAWG, mol/mol} datasets that NIST contributed to during the Early interval, 1992 through 2008.
b) Median measurement differences and uncertainties relative to NIST values for organizations that have results in at least 25 % of the {GAWG, mol/mol} datasets that NIST contributed to during the "Recent" interval, 2009 through 2023.
The normalized differences are plotted against the *x*-axis, where $x = 0$ is no difference. The uncertainty ratios are plotted against the *y*-axis, where $y = 1$ marks where the measurement uncertainties are equal. The symbol for NIST is located at the chart origin, $\{x = 0, y = 1\}$. The size of the symbols is related to the number of datasets in common with NIST. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to NIST.

All but one of the 13 organizations with results in at least 25 % of the 111 Early NIST datasets have $d_{\mathrm{NIST},i}$ within 2 composite distance units of NIST's summary estimates: BAM, LNE, VSL, NIM, CERI, NPL, KRISS, VNIIM, EAA, NMIJ, ERLAP, and BFKH. All nine of the organizations with results in at least 25 % of the 45 Recent NIST datasets have $d_{\mathrm{NIST},i}$ within 2 distance units: CERI, NPL, VSL, NIM, VNIIM, KRISS, NMISA, NMIJ, and LNE.

7.3. {IAWG, g/g}

Table 3 lists the IAWG-related studies used to identify NIST's inorganic analysis measurement peers. This list includes the 47 KC or PPS studies that produced the 130 quantitative {IAWG, g/g} datasets that were available as of this document's publication date. Several of the studies were coordinated jointly by the EAWG and IAWG; since the results from these studies are for inorganic analytes expressed as mass fractions, they are attributed to the IAWG.

Table 3. {IAWG, g/g} Studies in Which NIST Participated.

Study	Description	$\#_{set}$	Year	Coordinator
CCQM-K002	Cadmium and lead in natural water	2	1998	JRC
CCQM-K008	Monoelemental calibration solutions of Al, Cu, Fe, Mg	4	2000	EMPA,LNE
CCQM-K013	Amount content of cadmium and lead in sediment	2	2000	JRC
CCQM-K024	Cadmium in rice	1	2001	JRC,NMIJ
CCQM-Q029	Cadmium and zinc in rice	2	2001	JRC,NMIJ
CCQM-Q032	Anion calibration solutions	2	2001	EMPA
CCQM-Q013	Metals in synthetic food digest	2	2002	LGC
CCQM-K014	Calcium in human serum	1	2003	JRC
CCQM-K028	Tributyltin (TBT) in sediment	2	2003	LGC,NRC
CCQM-K031	Arsenic in shellfish	1	2003	NIST
CCQM-Q039	Toxic metals and methymercury in tuna	3	2003	JRC
CCQM-K033	Minor elements in steel	4	2004	NMIJ,NIST,BAM
CCQM-K034 [a]	Potassium hydrogen phthalate (KHP) purity	1	2004	SMU
CCQM-K035	Sulfur in diesel fuel	1	2004	NIST
CCQM-K042	Trace and minor elements in aluminum alloy	5	2005	BAM
CCQM-K043	Organo-mercury in salmon	4	2005	JRC
CCQM-K043.1	As, Hg, Se, methyl mercury in marine fish	1	2007	NMIJ
CCQM-K049	Toxic and essential elements in bovine liver	7	2007	NIST
CCQM-K056	Trace elements in whole-fat soybean powder	4	2007	NIM
CCQM-K057	Chemical composition of clay	5	2007	CENAM
CCQM-K058	Nitrogen and trace elements in silicon nitride power	4	2007	NMIJ,BAM
CCQM-K059	Nitrite and nitrate in calibration solutions	1	2007	SMU,NRC
CCQM-Q086	Total Se and selenomethionine in pharmaceuticals	3	2007	LGC,NRC
CCQM-K048 [a]	Potassium chloride (KCl) purity	1	2008	NIM
CCQM-K073 [a]	HCl purity	1	2009	NIST,CENAM
CCQM-K075	Toxic metals in algae	2	2009	IAEA
CCQM-K070	Mercury in natural water: low levels	1	2010	PTB,BAM,LNE
CCQM-K088	Lead in lead-free solder containing silver and copper	1	2010	NMIJ,NIM,KRISS
CCQM-K087	Monoelemental calibration solutions of Cr, Co and Pb	9	2011	PTB
CCQM-K089	Trace and essential elements in *Herba ecliptae*	5	2011	GLHK
CCQM-K096 [a]	Potassium dichromate ($K_2Cr_2O_7$) purity	1	2012	SMU,KRISS
CCQM-K097	Arsenobetaine in standard solution and in tuna tissue	2	2012	NMIJ,NIM
CCQM-K072	Zinc purity	7	2013	BAM
CCQM-K107	Elements and selenium speciation in human serum	5	2013	LGC
CCQM-K108	Total As, As species and Cd in brown rice flour	2	2013	NMIJ
CCQM-K048.2014 [a]	Chloride (Cl^-) in potassium chloride (KCl)	1	2014	NIM
CCQM-K123	Trace elements in biodiesel fuel	1	2014	NMIJ,NIST
CCQM-Q149	Zinc purity	3	2014	BAM
CCQM-K108.2014	Total arsenic and arsenic species in brown rice flour	3	2015	NMIJ
CCQM-K124	Trace elements and Cr speciation in drinking water	5	2015	NMIJ,GLHK

Study	Description	$\#_{set}$	Year	Coordinator
CCQM-K127	Toxic and trace elements in soils	7	2015	CENAM,IJS
CCQM-K125	Iodine and other elements in infant formula	3	2016	GLHK
CCQM-K034.2016 [a]	Amount of acid in a solid weak acid	1	2017	NIM
CCQM-K139	Elements in human serum	3	2017	HSA
CCQM-K143	Copper calibration solutions	1	2018	NIST
CCQM-K145	Toxic and essential elements in bovine liver powder	2	2018	NIM
CCQM-K034.2016.1 [a]	Amount of acid in a solid weak acid	1	2019	NIM

a) Study coordinated jointly by the EAWG and IAWG.

The $X_{\mathrm{NIST},i}$ and $Y_{\mathrm{NIST},i}$ values for organizations with results in at least 25 % of the datasets with measurement dates within the relevant Early (1992 through 2008) and Recent (2009 through 2023) intervals are displayed in Fig. 11.

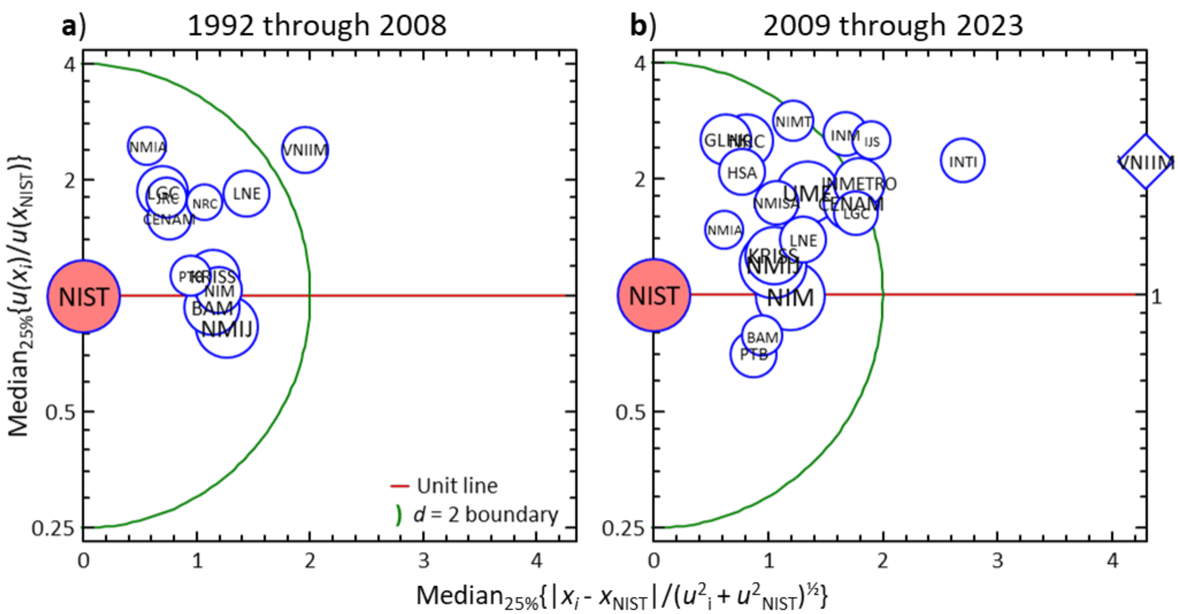


Fig. 11. {IAWG, g/g} Measurement Similarity Peers, Early and Recent.

a) Median measurement differences and uncertainties relative to NIST values for organizations that have results in at least 25 % of the {IAWG, g/g} datasets that NIST contributed to during the Early interval, 1992 through 2008.
b) Median measurement differences and uncertainties relative to NIST values for organizations that have results in at least 25 % of the {IAWG, g/g} datasets that NIST contributed to during the "Recent" interval, 2009 through 2023.
The normalized differences are plotted against the *x*-axis, where x = 0 is no difference. The uncertainty ratios are plotted against the *y*-axis, where *y* = 1 marks where the measurement uncertainties are equal. The symbol for NIST is located at the chart origin, {*x* = 0, *y* = 1}. The size of the symbols is related to the number of datasets in common with NIST. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to NIST.

All but one of the 12 organizations with results in at least 25 % of the 63 Early NIST datasets have $d_{\mathrm{NIST},i}$ within 2 composite distance units of NIST's results: PTB, CENAM, JRC, LGC, BAM, KRISS, NIM, NMIJ, NRC, NMIA, and LNE. Fifteen of the 20 organizations with results in at least 25 % of the 67 Recent NIST datasets have $d_{\mathrm{NIST},i}$ within 2 distance units: NMIA, PTB, BAM, NMIJ, KRISS, NIM, HSA, NMISA, LNE, GLHK, NRC, UME, LGC, CENAM, and NIMT.

7.4. {OAWG, g/g}

Table 4 lists the OAWG-related studies used to identify NIST's organic analysis measurement peers. This list includes the 40 KC or PPS studies that produced the 96 quantitative {OAWG, g/g} datasets that were available as of this document's publication date.

Table 4. {OAWG, g/g} Studies in Which NIST Participated.

Study	Description	$\#_{set}$	Year	Coordinator
CCQM-Q002	p,p'-DDE in isooctane	2	1997	LGC
CCQM-Q006	Cholesterol in human serum	2	1998	NIST
CCQM-K005	p,p'-DDE) in fish oil	2	2000	LGC
CCQM-Q009	Creatinine in human serum	2	2000	NIST
CCQM-Q017	PCBs in sediments	4	2000	NRC,NIST
CCQM-K006	Cholesterol in human serum	2	2001	NIST
CCQM-K011	Glucose in human serum	1	2001	NIST
CCQM-K012	Creatinine in human serum	2	2001	NIST
CCQM-K021	p,p'-DDT in fish oil	2	2001	LGC
CCQM-Q008	Glucose in human serum	2	2001	NIST
CCQM-K025	PCBs in sediments	5	2002	NIST,NRC
CCQM-K027.a	Ethanol in aqueous matrix	2	2002	LGC,BAM
CCQM-K027.b	Ethanol in aqueous matrix	1	2002	LGC,BAM
CCQM-K038	PAHs in solution	5	2005	NIST
CCQM-K039	Organochlorine pesticides in solution	4	2005	NIST
CCQM-K040	PCB congeners in solution	4	2005	NIST
CCQM-K047	VOCs in solution	4	2006	CENAM,NIST
CCQM-K050	PAHs in soil and particulates	10	2007	CENAM,BAM
CCQM-Q020.e	Theophylline purity	2	2007	BIPM,LGC
CCQM-K062	Nutrients in infant/adult formula	3	2008	NIST
CCQM-K063.a	Cortisol in human serum	1	2008	NIST
CCQM-K063.b	Progesterone in human serum	1	2008	NIST
CCQM-Q020.f	Digoxin purity	1	2008	BIPM,LGC
CCQM-K055.a	17-β-Estradiol purity	1	2009	BIPM
CCQM-K080	Creatinine in human serum	1	2009	NIST
CCQM-K055.b	Aldrin purity	1	2010	BIPM
CCQM-K079	Ethanol in aqueous matrix	1	2010	BAM,NIST
CCQM-K055.c	L-(+)-Valine purity	1	2012	BIPM
CCQM-K095	Organochlorine pesticides in tea	2	2012	GLHK,NIM
CCQM-K102	PBDEs in sediment	3	2015	JRC
CCQM-K132	Vitamin D metabolites in human serum	3	2015	NIST
CCQM-K055.d	Folic acid purity	1	2016	BIPM
CCQM-K095.1	PAHs in tea	2	2016	NIST
CCQM-K109	Urea and uric acid in human serum	4	2016	HSA
CCQM-K131	PAHs in acetonitrile	3	2016	NIST
CCQM-K142	Urea and uric acid in human serum or plasma	2	2016	HSA,NIST
CCQM-K078.a	Multi-component amino acids in dilute HCl solution	4	2017	BIPM
CCQM-K147	Niacin (vitamin B_3) in milk powder	1	2017	NIST,CENAM
CCQM-K146	Benzo[*a*]pyrene in olive oil	1	2018	NIM
CCQM-K148.a	Bisphenol-A purity	1	2019	BIPM

The $X_{\mathrm{NIST},i}$ and $Y_{\mathrm{NIST},i}$ values for organizations with results in at least 25 % of the datasets with measurement dates within the relevant Early (1992 through 2008) and Recent (2009 through 2023) intervals are displayed in Fig. 12.

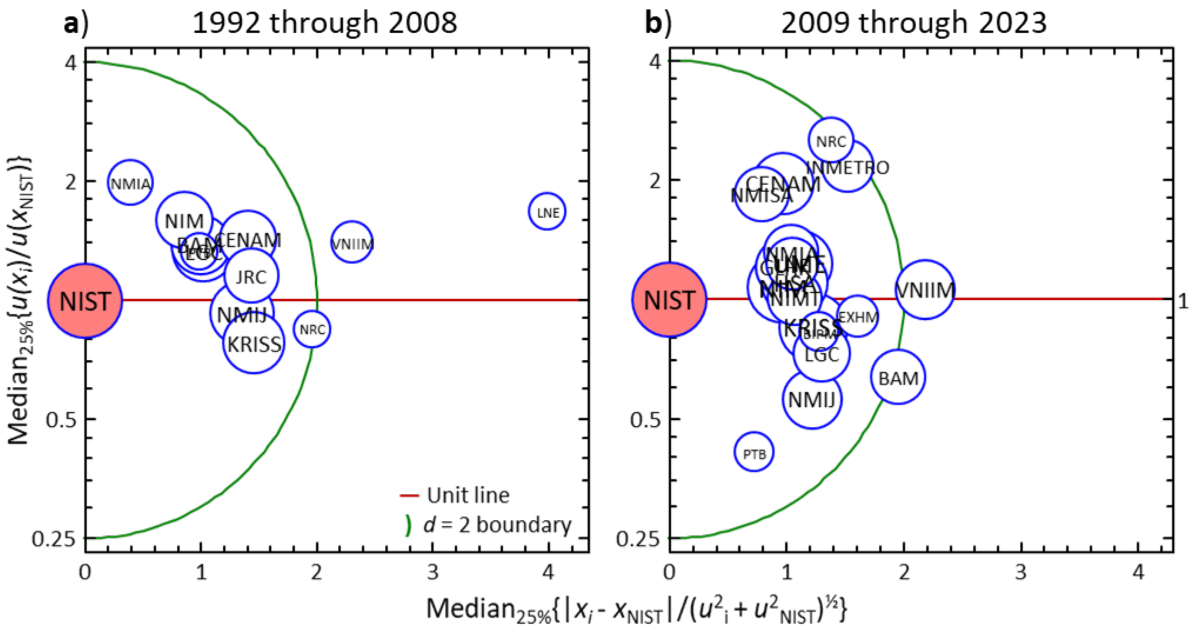


Fig. 12. {OAWG, g/g} Measurement Similarity Peers, Early and Recent.

a) Median measurement differences and uncertainties relative to NIST values for organizations that have results in at least 25 % of the {OAWG, g/g} datasets that NIST contributed to during the Early interval, 1992 through 2008.
b) Median measurement differences and uncertainties relative to NIST values for organizations that have results in at least 25 % of the {OAWG, g/g} datasets that NIST contributed to during the “Recent” interval, 2009 through 2023.
The normalized differences are plotted against the *x*-axis, where $x = 0$ is no difference. The uncertainty ratios are plotted against the *y*-axis, where $y = 1$ marks where the measurement uncertainties are equal. The symbol for NIST is located at the chart origin, $\{x = 0, y = 1\}$. The size of the symbols is related to the number of datasets in common with NIST. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to NIST.

Ten of the 12 organizations with results in at least 25 % of the 64 Early NIST datasets have $d_{\mathrm{NIST},i}$ within 2 composite distance units of NIST’s results: NMIA, PTB, NIM, BAM, LGC, NMIJ, JRC, CENAM, KRISS, and NRC. Seventeen of the 19 organizations with results in at least 25 % of the 32 Recent NIST datasets have $d_{\mathrm{NIST},i}$ within 2 distance units: NIM, GLHK, NIMT, HSA, LNE, NMIA, UME, NMISA, KRISS, BIPM, CENAM, LGC, NMIJ, PTB, EXHM, INMETRO, and NRC.

7.5. {EAWG, pH}

Table 5 lists the EAWG-related studies used to identify NIST's pH measurement peers. This list includes the 9 KC studies that produced the 37 quantitative {EAWG, pH} datasets that were available as of this document's publication date.

Table 5. {EAWG, pH} Studies in Which NIST Participated.

Study	Description	$\#_{set}$	Year	Coordinator
CCQM-K009	pH of phosphate buffer	13	2000	PTB
CCQM-K017	pH of phthalate buffer	3	2001	PTB
CCQM-K019	pH of borate buffer	3	2005	PTB
CCQM-K018	pH of carbonate buffer	1	2006	SMU
CCQM-K020	pH of tetroxalate buffer	3	2008	NIST
CCQM-K019.1	pH of borate buffer	3	2010	PTB
CCQM-K091	pH of an unknown phthalate buffer	5	2011	PTB
CCQM-K099	pH of phosphate buffer	5	2014	PTB
CCQM-K018.2016	pH of carbonate buffer	1	2017	NIST

The $X_{\mathrm{NIST},i}$ and $Y_{\mathrm{NIST},i}$ values for organizations with results in at least 25 % of the datasets with measurement dates within the relevant Early (1992 through 2008) and Recent (2009 through 2023) intervals are displayed in Fig. 13.

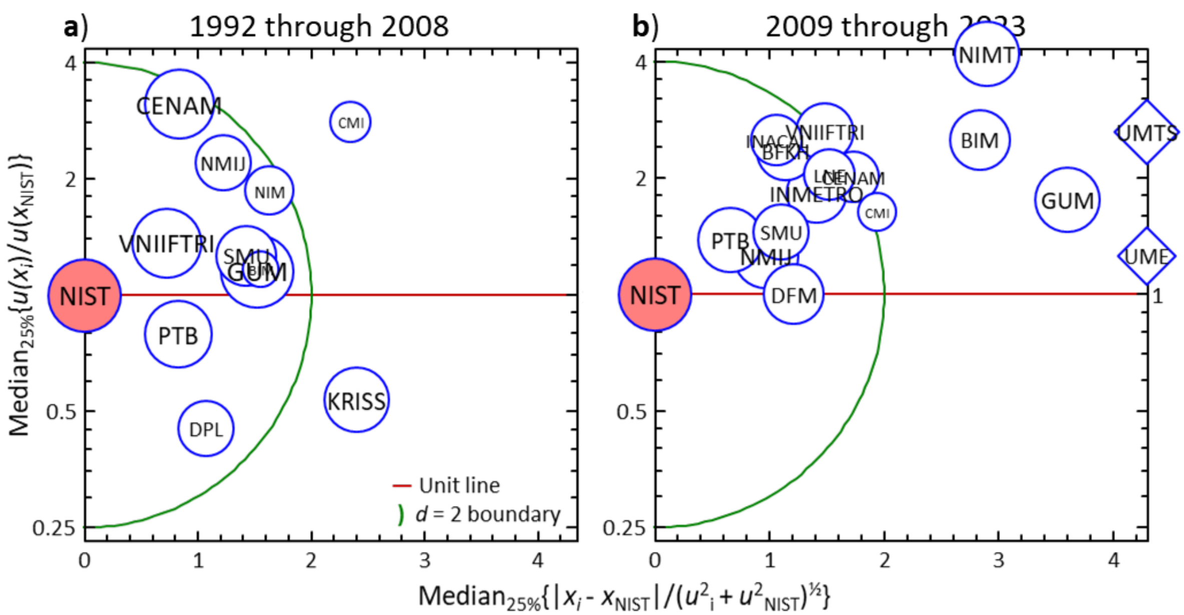


Fig. 13. {EAWG, pH} Measurement Similarity Peers, Early and Recent.

a) Median measurement differences and uncertainties relative to NIST values for organizations that have results in at least 25 % of the {EAWG, pH} datasets that NIST contributed to during the Early interval, 1992 through 2008.
b) Median measurement differences and uncertainties relative to NIST values for organizations that have results in at least 25 % of the {EAWG, pH} datasets that NIST contributed to during the "Recent" interval, 2009 through 2023.
The normalized differences are plotted against the *x*-axis, where $x = 0$ is no difference. The uncertainty ratios are plotted against the *y*-axis, where $y = 1$ marks where the measurement uncertainties are equal. The symbol for NIST is located at the chart origin, $\{x = 0, y = 1\}$. The size of the symbols is related to the number of datasets in common with NIST. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to NIST.

Nine of the 11 organizations with results in at least 25 % of the 23 Early NIST datasets have $d_{\mathrm{NIST},i}$ within 2 composite distance units of NIST's results: VNIIFTRI, PTB, SMU, GUM, DPL, BIM, NMIJ, CENAM, and NIM. Eight of the 16 organizations with results in at least 25 % of the 14 Recent NIST datasets have $d_{\mathrm{NIST},i}$ within 2 distance units: PTB, NMIJ, DFM, SMU, INMETRO, BFKH, INACAL, and LNE.

7.6. {EAWG, S/m}

Table 6 lists the EAWG studies used to identify NIST's electrolytic conductivity measurement peers. This list includes the three KC studies that produced the three publicly accessible quantitative {EAWG, S/m} datasets in which NIST participated. NIST stopped making electrolytic conductivity measurements in 2014 after developing and validating a solution preparation "recipe" that enables users to prepare their own artifact standards.

Table 6. {EAWG, S/,m} Studies in Which NIST Participated.

Study	Description	$\#_{set}$	Year	Coordinator
CCQM-K036.a	Electrolytic conductivity (0.5 S/m)	1	2005	DFM
CCQM-K036.b	Electrolytic conductivity (5 mS/m)	1	2005	DFM
CCQM-K092	Electrolytic conductivity (0.05 S/m and 20 S/m)	1	2011	SMU

The $X_{\mathrm{NIST},i}$ and $Y_{\mathrm{NIST},i}$ values for organizations that also participated in all three of these studies are displayed in Fig. 14.

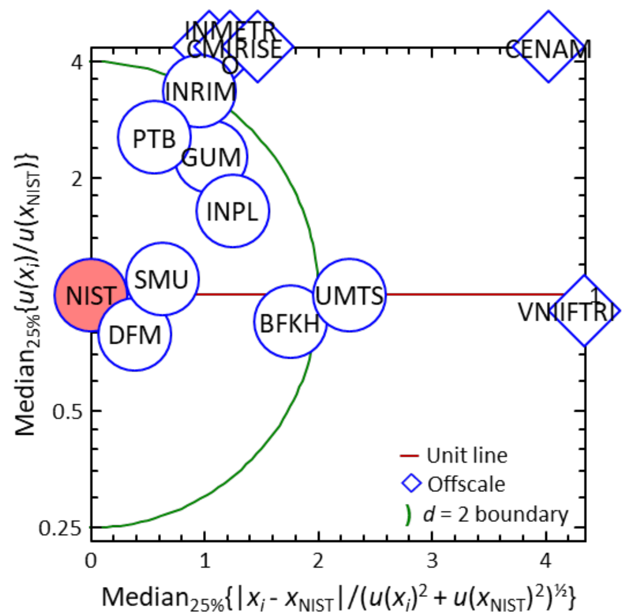


Fig. 14. {EAWG, S/m} Measurement Similarity Peers.

Median measurement differences and uncertainties relative to NIST values for organizations that participated in the same studies as NIST. The normalized differences are plotted against the *x*-axis, where *x* = 0 is no difference. The uncertainty ratios are plotted against the *y*-axis, where *y* = 1 marks where the measurement uncertainties are equal. The symbol for NIST is located at the chart origin, {*x* = 0, *y* = 1}. The size of the symbols is related to the number of datasets in common with NIST. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to NIST.

Six of the 13 organizations with results in the three datasets have $d_{\mathrm{NIST},i}$ within 2 composite distance units of NIST's results: DFM, SMU, INPL, PTB, GUM, and BFKH.

7.7. {SAWG, m}

Table 7 lists the SAWG-related studies used to identify NIST's film thickness measurement peers. This list includes the three KC or PPS studies that produced the 26 quantitative {SAWG, m} datasets that were available as of this document's publication date.

Table 7. {SAWG, m} Studies in Which NIST Participated.

Study	Description	$\#_{set}$	Year	Coordinator
CCQM-K032	Thickness of silicon dioxide on silicon film	4	2005	NPL
CCQM-Q190	Thickness of hafnium oxide film	18	2017	KRISS
CCQM-K157	Thickness of hafnium oxide film	4	2021	KRISS

The $X_{\mathrm{NIST},i}$ and $Y_{\mathrm{NIST},i}$ values for organizations with results in at least 25 % of the datasets are displayed in Fig. 15.

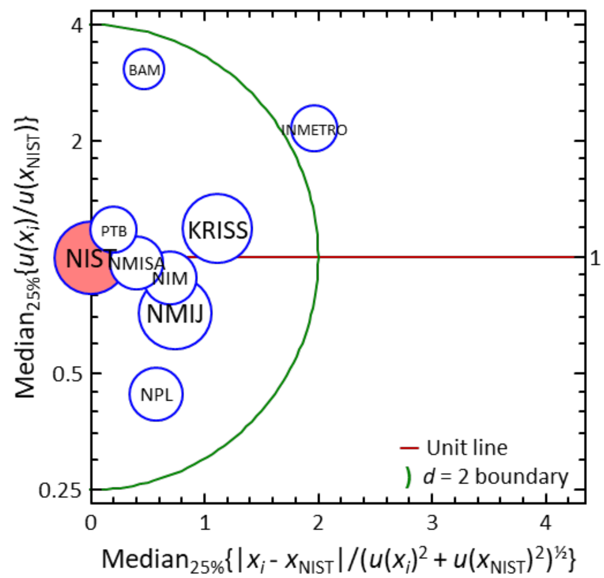


Fig. 15. {SAWG, m} Measurement Similarity Peers.

Median measurement differences and uncertainties relative to NIST values for organizations that have results in at least 25 % of the {SAWG, m} datasets that NIST contributed to. The normalized differences are plotted against the *x*-axis, where *x* = 0 is no difference. The uncertainty ratios are plotted against the *y*-axis, where *y* = 1 marks where the measurement uncertainties are equal. The symbol for NIST is located at the chart origin, {*x* = 0, *y* = 1}. The size of the symbols is related to the number of datasets in common with NIST. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to NIST.

Seven of the 8 organizations with results in these 26 datasets have $d_{\mathrm{NIST},i}$ within 2 composite distance units of NIST's results: PTB, NMISA, NIM, NMIJ, KRISS, NPL, and BAM.

7.8. {IRWG,‰}

Table 8 lists the IRWG-related studies used to identify NIST's isotopic δ-scale measurement peers. The two PPS studies provide the ten quantitative {IRWG,‰} datasets that were available as of this document's publication date.

Table 8. { IRWG,‰} Studies in Which NIST Participated.

Study	Description	$\#_{set}$	Year	Coordinator
CCQM-Q204	CO_2 isotope ratios ($\delta^{13}C$ and $\delta^{18}O$) in pure CO_2	8	2022	BIPM,IAEA
CCQM-Q213	Copper isotope ratios in high purity copper	2	2022	NRC,NIST,BAM,PTB

The $X_{\text{NIST},i}$ and $Y_{\text{NIST},i}$ values for organizations with results in at least 25 % of the datasets are displayed in Fig. 16.

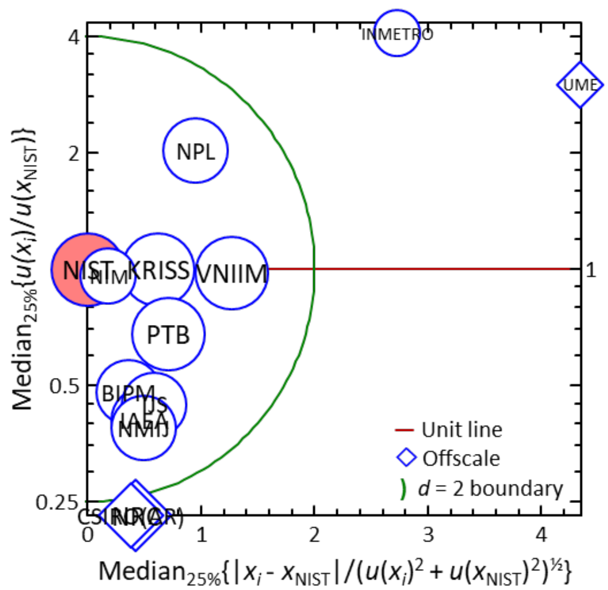


Fig. 16. {IRWG,‰} Measurement Similarity Peers.

Median measurement differences and uncertainties relative to NIST values for organizations that have results in at least 25 % of the {IRWG,‰} datasets that NIST contributed to. The normalized differences are plotted against the *x*-axis, where *x* = 0 is no difference. The uncertainty ratios are plotted against the *y*-axis, where *y* = 1 marks where the measurement uncertainties are equal. The symbol for NIST is located at the chart origin, {*x* = 0, *y* = 1}. The size of the symbols is related to the number of datasets in common with NIST. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to NIST.

Nine of the 13 organizations with results in these ten datasets have $d_{\text{NIST},i}$ within 2 composite distance units of NIST's results: NIM, KRISS, PTB, BIPM, VNIIM, IJS, NPL, IAEA, and NMIJ.

8. Measurement Performance Peers

The estimates of measurement similarity presented in Section 7 do not necessarily indicate the quality of the measurements. However, estimates of measurement performance are provided by comparing an organization's results to values that are characteristic of the datasets themselves rather than to just another participant.

Limiting the comparisons to datasets in which NIST has a result and organizations that have results in a minimum of 25 % of those datasets ensures that the performance estimates are for organizations with interests similar to NIST.

8.1. Metrics

8.1.1. Median Relative Difference to Reference Values

The relative difference from the dataset reference value, $X_{\mathrm{ref},i}$, is:

$$X_{\mathrm{ref},i} = \mathrm{Median}_{25\%}\left\{|x_i - x_{\mathrm{ref}}|/\sqrt{u^2(x_i) + u^2(x_{\mathrm{ref}})}\right\} \tag{7}$$

where x_{ref} is the dataset's reference value, $u(x_{\mathrm{ref}})$ is one-half of the 95 % expanded uncertainty associated with the reference value, and the other terms are as described in Section 7.1.1. For the few datasets without an assigned reference value, the *CCQM_Retrospectoscope* assigns x_{ref} as the median of the technically valid results in the dataset and $u(x_{\mathrm{ref}})$ as the Q_n robust standard deviation divided by the square root of the number of those results [3].

8.1.2. Median Uncertainty Relative to All-Participant Median Uncertainty

The uncertainty relative to the participants' median uncertainty, $Y_{\mathrm{set},i}$, is:

$$Y_{\mathrm{set},i} = \mathrm{Median}_{25\%}\{u(x_i)/\mathrm{Median}\{u(x_{\mathrm{set}})\}\} \tag{8}$$

where $\mathrm{Median}\{u(x_{\mathrm{set}})\}$ is the median of all the uncertainties associated with the dataset's valid results and the other terms are as described in Section 7.1.2. Note: $u(x_{\mathrm{ref}})$ is not used as the normalization parameter since it is not necessarily related to the expected uncertainty associated with a measurement result.

8.1.3. Performance Distance to Dataset Characteristic Values

The $X_{\mathrm{ref},i}$ and $Y_{\mathrm{set},i}$ summary estimates can be combined into a Euclidean distance, $d_{\mathrm{set}.i}$, that summarizes organization *i*'s measurement performance:

$$d_{\mathrm{set},i} = \sqrt{\left(X_{\mathrm{ref},i}\right)^2 + \left(\log_2\{Y_{\mathrm{set},i}\}\right)^2} \tag{9}$$

where $\log_2\{\cdot\}$ is the binary logarithm function as described in Section 7.1.3.

Although less well defined, the relationship between $X_{\mathrm{ref},i}$ and $Y_{\mathrm{set},i}$ is likely similar to that between $X_{\mathrm{NIST},i}$ and $Y_{\mathrm{NIST},i}$; the $d_{\mathrm{set}.i}$ estimates likewise need to be interpreted with care.

8.2. {GAWG, mol/mol}

The studies used to identify NIST's gas mixture measurement performance peers are listed in Table 2. The $X_{\mathrm{ref},i}$ and $Y_{\mathrm{set},i}$ values for organizations with {GAWG, mol/mol} results in at least 25 % of datasets with NIST results and measurement dates within the relevant Early (1992 through 2008) and Recent (2009 through 2023) intervals are displayed in Fig. 17.

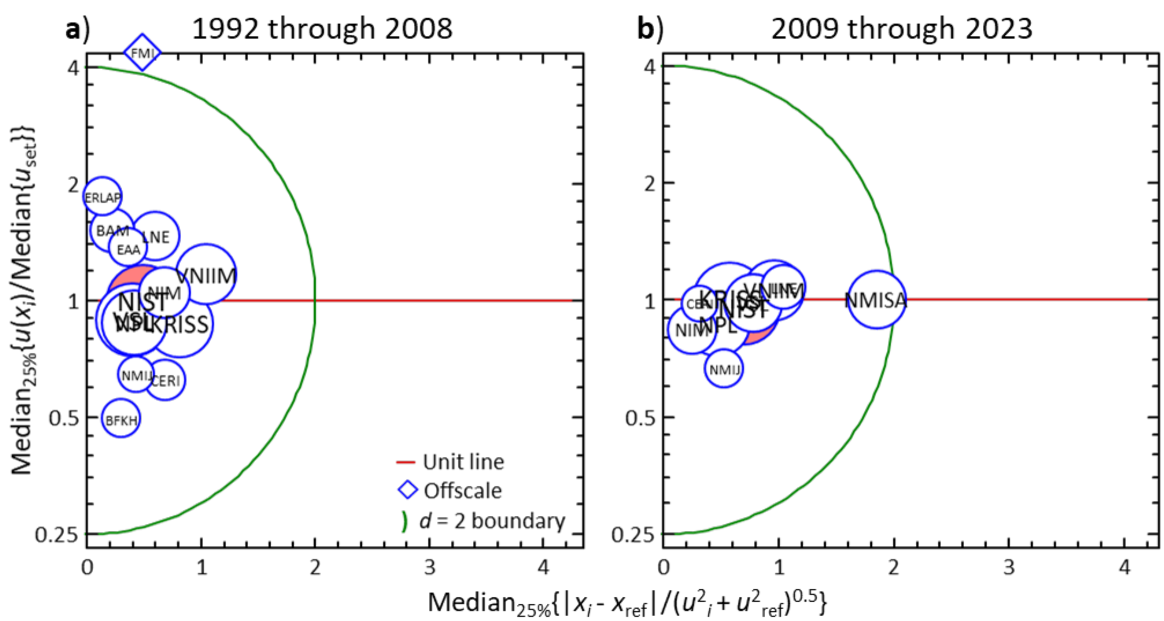


Fig. 17. {GAWG, mol/mol} Measurement Performance Peers, Early and Recent.

a) Median measurement differences and uncertainties relative to dataset performance characteristics for organizations that participated with NIST in at least 25 % of the {GAWG, mol/mol} datasets that NIST contributed to during the Early interval, 1992 through 2008.
b) Median measurement differences and uncertainties relative to dataset performance characteristics for organizations that participated with NIST in at least 25 % of the {GAWG, mol/mol} datasets that NIST contributed to during the Recent interval, 2009 through 2023.
The normalized differences are plotted against the *x*-axis, where *x* = 0 is no difference. The uncertainty ratios are plotted against the *y*-axis, where *y* = 1 marks where the measurement uncertainties are equal. The size of the symbols is related to the number of datasets in common with NIST. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to NIST.

All but one of the 14 organizations with results in at least 25 % of the 111 Early NIST datasets have $d_{\mathrm{set},i}$ within 2 composite distance units of the dataset's characteristic parameters: VSL, NPL, NIST, EAA, BAM, NIM, NMIJ, LNE, KRISS, ERLAP, CERI, BFKH, and VNIIM. All ten organizations with results in at least 25 % of the 45 Recent NIST datasets have $d_{\mathrm{set},i}$ within 2 distance units: CERI, NIM, NPL, KRISS, NIST, VSL, NMIJ, VNIIM, LNE, and NMISA.

The performance distances between the two intervals are compared in Fig. 18.

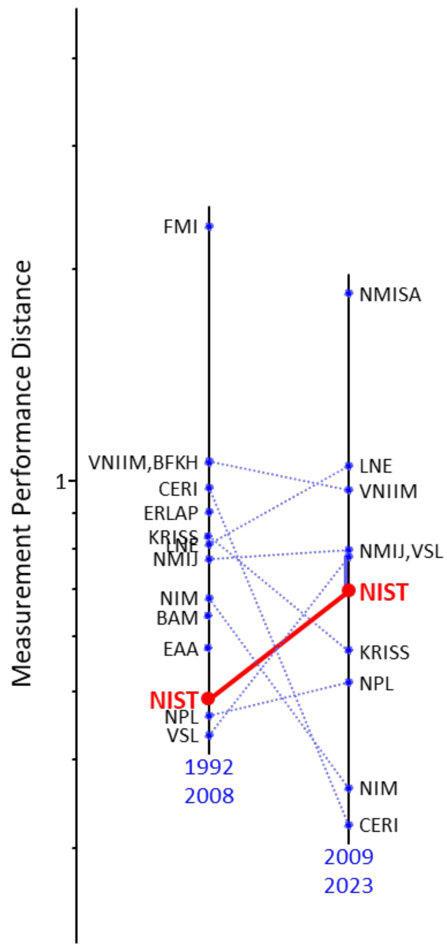


Fig. 18. Change in NIST {GAWG, mol/mol} Measurement Performance Between Intervals.

Each closed circle represents one organization's composite distance relative to the dataset's characteristic parameters. The vertical line with log_{10}-spaced tic marks to the left defines the distance axis. Results for the Early (1992 through 2008) interval are shown on the middle vertical line. Results for the Recent (2009 through 2023) interval are shown on the vertical line to the right. Lines connecting dots highlight the change in measurement performance between the intervals. NIST's results are highlighted in red.

8.3. {IAWG, g/g}

The studies used to identify NIST's inorganic analysis performance peers are listed in Table 3. The $X_{\mathrm{ref},i}$ and $Y_{\mathrm{set},i}$ values for organizations with {IAWG, g/g} results in at least 25 % of datasets with NIST results and measurement dates within the relevant Early (1992 through 2008) and Recent (2009 through 2023) intervals are displayed in Fig. 19.

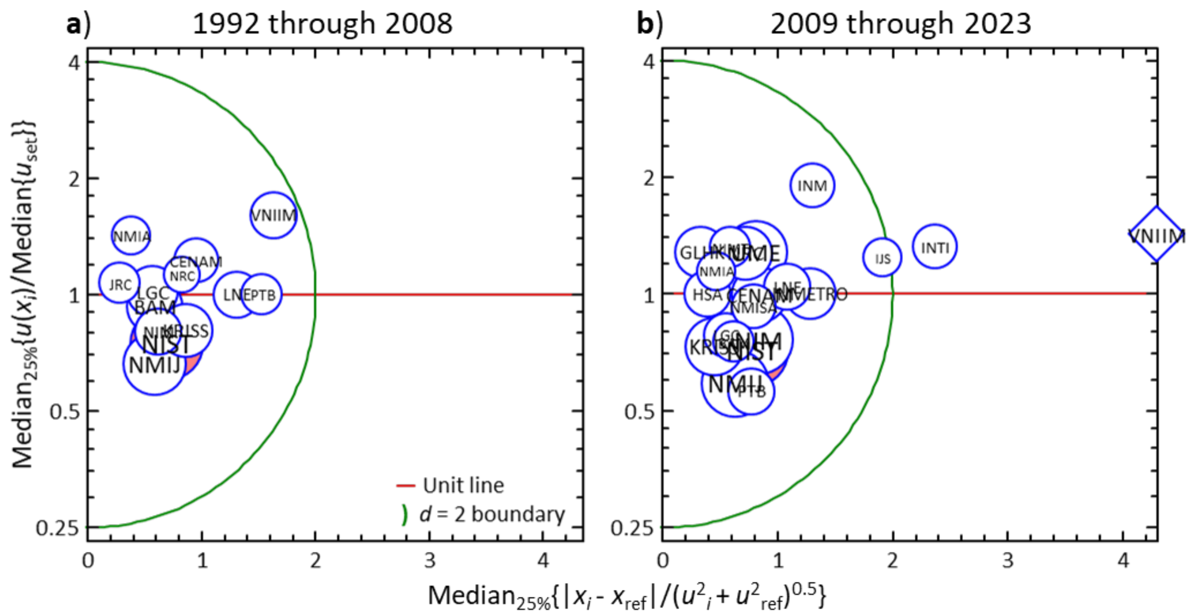


Fig. 19. {IAWG, g/g} Measurement Performance Peers, Early and Recent.

a) Median measurement differences and uncertainties relative to dataset performance characteristics for organizations that participated with NIST in at least 25 % of the {IAWG, g/g} datasets that NIST contributed to during the Early interval, 1992 through 2008.
b) Median measurement differences and uncertainties relative to dataset performance characteristics for organizations that participated with NIST in at least 25 % of the {IAWG, g/g} datasets that NIST contributed to during the Recent interval, 2009 through 2023.
The normalized differences are plotted against the *x*-axis, where *x* = 0 is no difference. The uncertainty ratios are plotted against the *y*-axis, where *y* = 1 marks where the measurement uncertainties are equal. The size of the symbols is related to the number of datasets in common with NIST. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to NIST.

All of the 14 organizations with results in at least 25 % of the 63 Early NIST datasets have $d_{\mathrm{set},i}$ within 2 composite distance units of the dataset's characteristic parameters: JRC, LGC, BAM, NMIA, NIM, NIST, NMIJ, NRC, KRISS, CENAM, LNE, PTB, and VNIIM. Nineteen of the 21 organizations with results in at least 25 % of the 67 Recent NIST datasets have $d_{\mathrm{set},i}$ within 2 distance units: HSA, GLHK, NMIA, KRISS, LGC, NIMT, BAM, NRC, NMISA, CENAM, UME, NIST, NIM, NMIJ, LNE, PTB, INMETRO, INM, and IJS.

The performance distances for the two intervals are compared in Fig. 20.

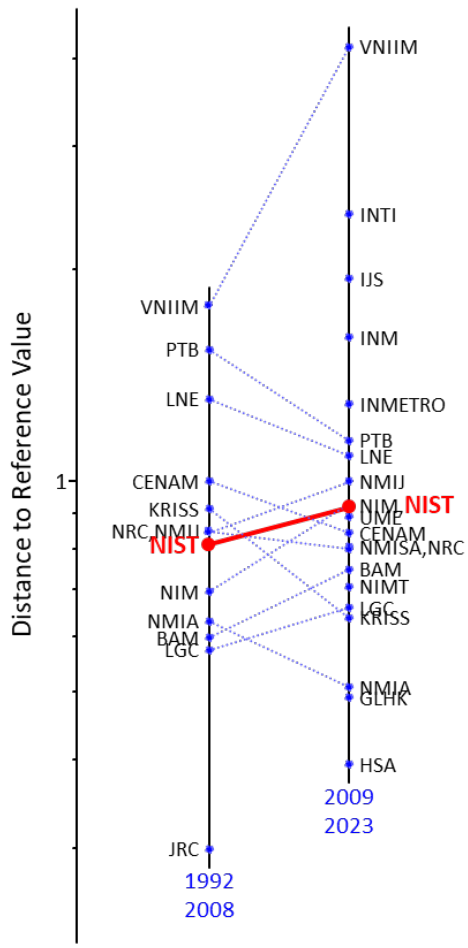


Fig. 20. Change in NIST {IAWG, g/g} Measurement Performance Between Intervals.

Each closed circle represents one organization's composite distance relative to the dataset's characteristic parameters. The vertical line with log_{10}-spaced tic marks to the left defines the distance axis. Results for the Early (1992 through 2008) interval are shown on the middle vertical line. Results for the Recent (2009 through 2023) interval are shown on the vertical line to the right. Lines connecting dots highlight the change in measurement performance between the intervals. NIST's results are highlighted in red.

8.4. {OAWG, g/g}

The studies used to identify NIST's organic analysis performance peers are listed in Table 4. The $X_{\mathrm{ref},i}$ and $Y_{\mathrm{set},i}$ values for organizations with {OAWG, g/g} results in at least 25 % of datasets with NIST results and measurement dates within the relevant Early (1992 through 2008) and Recent (2009 through 2023) intervals are displayed in Fig. 21.

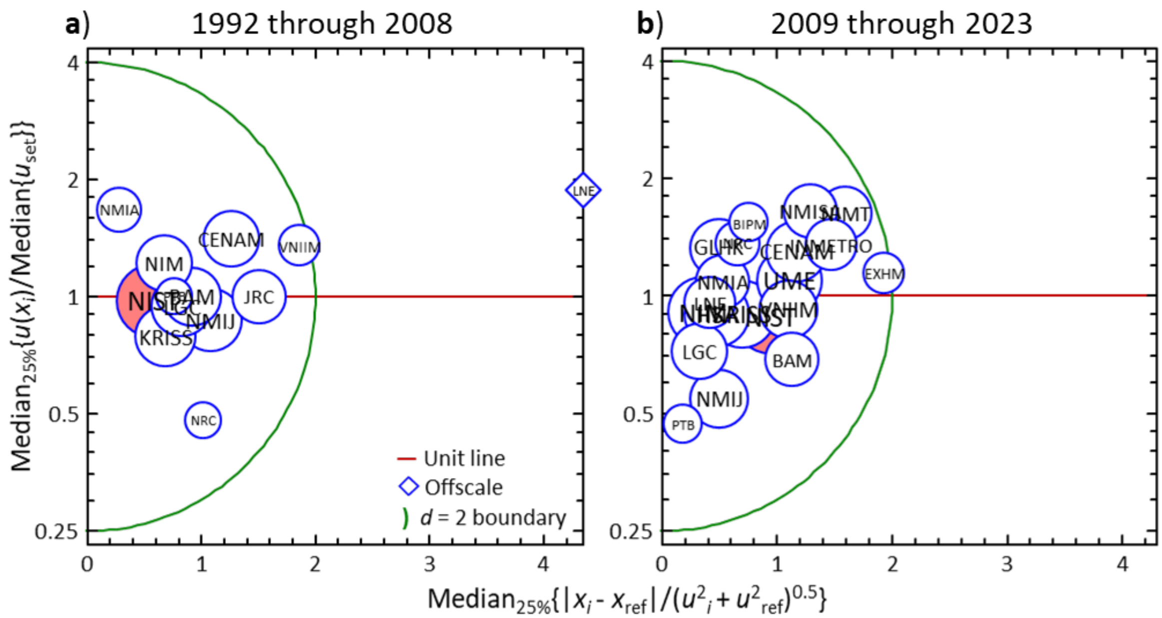


Fig. 21. {OAWG, g/g} Measurement Performance Peers, Early and Recent.

a) Median measurement differences and uncertainties relative to dataset performance characteristics for organizations that participated with NIST in at least 25 % of the {OAWG, g/g} datasets that NIST contributed to during the Early interval, 1992 through 2008.

b) Median measurement differences and uncertainties relative to dataset performance characteristics for organizations that participated with NIST in at least 25 % of the {OAWG, g/g} datasets that NIST contributed to during the Recent interval, 2009 through 2023.

The normalized differences are plotted against the *x*-axis, where *x* = 0 is no difference. The uncertainty ratios are plotted against the *y*-axis, where *y* = 1 marks where the measurement uncertainties are equal. The size of the symbols is related to the number of datasets in common with NIST. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to NIST.

All but one of the 14 organizations with results in at least 25 % of the 64 Early NIST datasets have $d_{\mathrm{set},i}$ within 2 composite distance units of the dataset's characteristic parameters. All 20 of the organizations with results in at least 25 % of the 32 Recent NIST datasets have $d_{\mathrm{set},i}$ within 2 distance units: NIM, LNE, HSA, NMIA, LGC, GLHK, KRISS, NRC, NIST, BIPM, NMIJ, VNIIM, PTB, UME, CENAM, BAM, NMISA, INMETRO, NIMT, EXHM.

The performance distances for the two intervals are compared in Fig. 22.

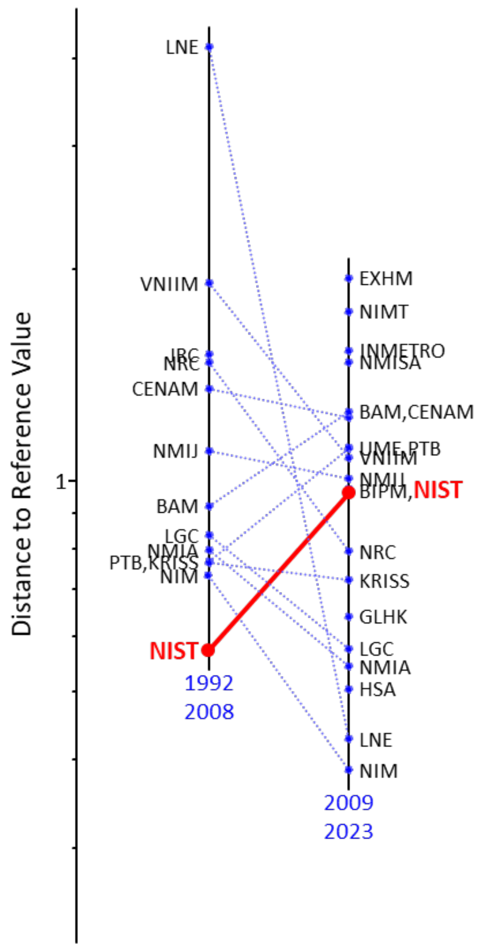


Fig. 22. Change in NIST {OAWG, g/g} Measurement Performance Between Intervals.

Each closed circle represents one organization's composite distance relative to the dataset's characteristic parameters. The vertical line with $\log_{10}$-spaced tic marks to the left defines the distance axis. Results for the Early (1992 through 2008) interval are shown on the middle vertical line. Results for the Recent (2009 through 2023) interval are shown on the vertical line to the right. Lines connecting dots highlight the change in measurement performance between the intervals. NIST's results are highlighted in red.

8.5. {EAWG, pH}

The studies used to identify NIST's pH measurement performance peers are listed in Table 5. The $X_{\mathrm{ref},i}$ and $Y_{\mathrm{set},i}$ values for organizations with {EAWG, pH} results in at least 25 % of datasets with NIST results and measurement dates within the relevant Early (1992 through 2008) and Recent (2009 through 2023) intervals are displayed in Fig. 23.

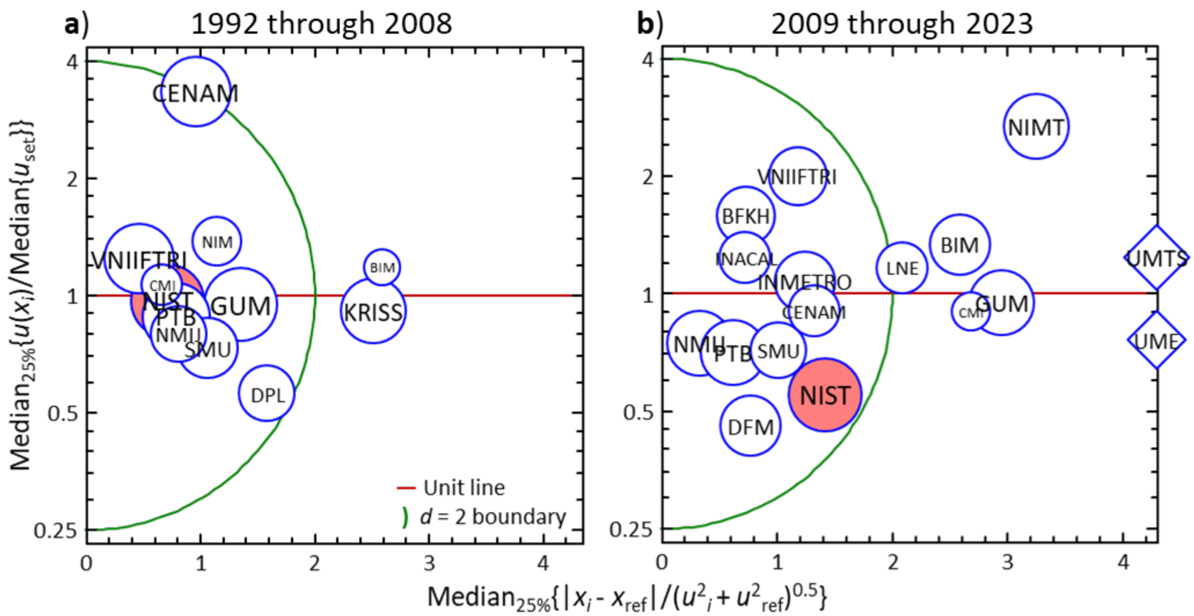


Fig. 23. {EAWG, pH} Measurement Performance Peers, Early and Recent.

a) Median measurement differences and uncertainties relative to dataset performance characteristics for organizations that participated with NIST in at least 25 % of the {EAWG, pH} datasets that NIST contributed to during the Early interval, 1992 through 2008.
b) Median measurement differences and uncertainties relative to dataset performance characteristics for organizations that participated with NIST in at least 25 % of the {EAWG, pH} datasets that NIST contributed to during the Recent interval, 2009 through 2023.
The normalized differences are plotted against the *x*-axis, where *x* = 0 is no difference. The uncertainty ratios are plotted against the *y*-axis, where *y* = 1 marks where the measurement uncertainties are equal. The size of the symbols is related to the number of datasets in common with NIST. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to NIST.

Ten of the 12 organizations with results in at least 25 % of the 23 Early NIST datasets have $d_{\mathrm{set},i}$ within 2 composite distance units of the dataset's characteristic parameters: VNIIFTRI, CMI, NIST, PTB, NMIJ, SMU, NIM, GUM, DPL, and CENAM. Ten of the 17 organizations with results in at least 25 % of the 14 Recent NIST datasets have $d_{\mathrm{set},i}$ within 2 distance units: NMIJ, INACAL, PTB, BFKH, SMU, INMETRO, CENAM, DFM, VNIIFTRI, and NIST.

The performance distances for the two intervals are compared in Fig. 24.

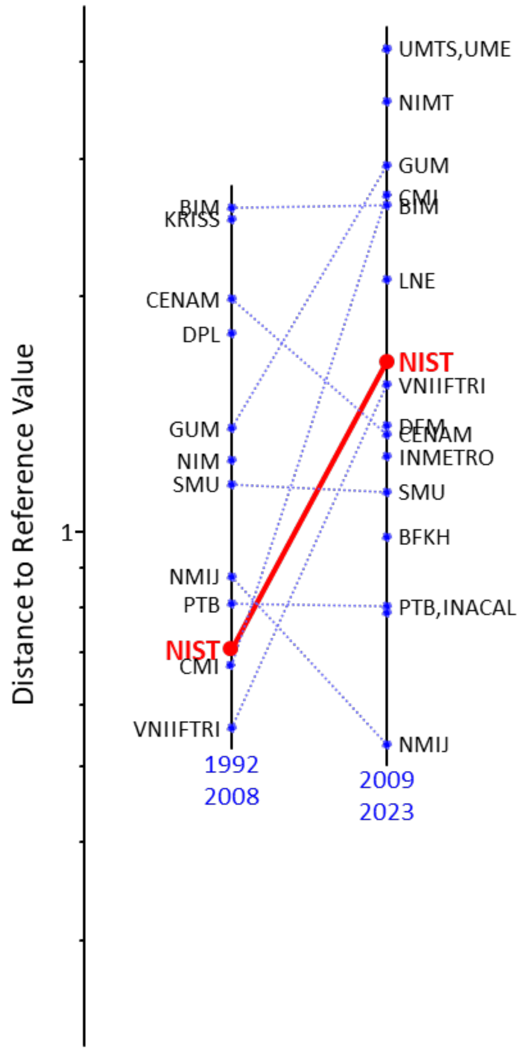


Fig. 24. Change in NIST {EAWG, pH} Measurement Performance Between Intervals.

Each closed circle represents one organization's composite distance relative to the dataset's characteristic parameters. The vertical line with $\log_{10}$-spaced tic marks to the left defines the distance axis. Results for the Early (1992 through 2008) interval are shown on the middle vertical line. Results for the Recent (2009 through 2023) interval are shown on the vertical line to the right. Lines connecting dots highlight the change in measurement performance between the intervals. NIST's results are highlighted in red.

8.6. {EAWG, S/m}

The studies used to identify NIST's electrolytic conductivity measurement performance peers are listed in Table 6. The $X_{\mathrm{ref},i}$ and $Y_{\mathrm{set},i}$ values for organizations that co-participated with NIST in {EAWG, S/m} studies are displayed in Fig. 25.

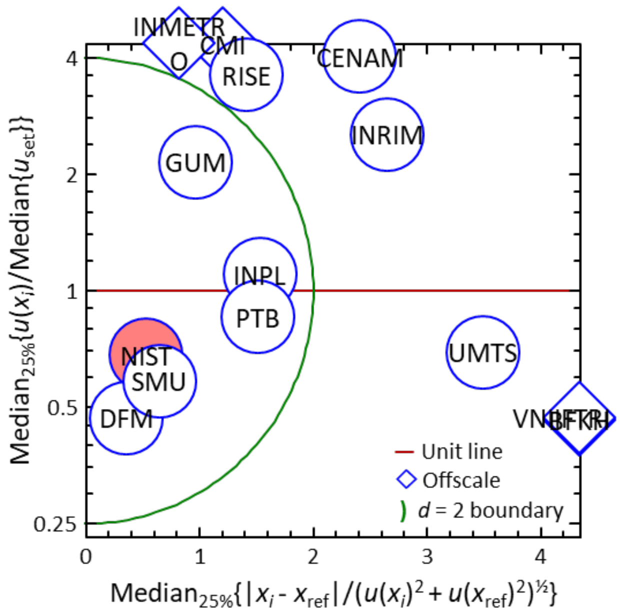


Fig. 25. {EAWG, S/m} Measurement Performance Peers.

Median measurement differences and uncertainties relative to dataset performance characteristics for organizations that participated with NIST in the same {EAWG, S/m} studies. The normalized differences are plotted against the *x*-axis, where *x* = 0 is no difference. The uncertainty ratios are plotted against the *y*-axis, where *y* = 1 marks where the measurement uncertainties are equal. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to the dataset performance characteristics.

Six of the 14 organizations with results in the three datasets have $d_{\mathrm{set},i}$ within 2 composite distance units of the dataset's characteristic parameters: NIST, SMU, DFM, GUM, PTB, and INPL.

8.7. {SAWG, m}

The studies used to identify NIST's thin-film measurement performance peers are listed in Table 7. The $X_{\mathrm{ref},i}$ and $Y_{\mathrm{set},i}$ values for organizations that co-participated with NIST in {SAWG, m} studies are displayed in Fig. 26.

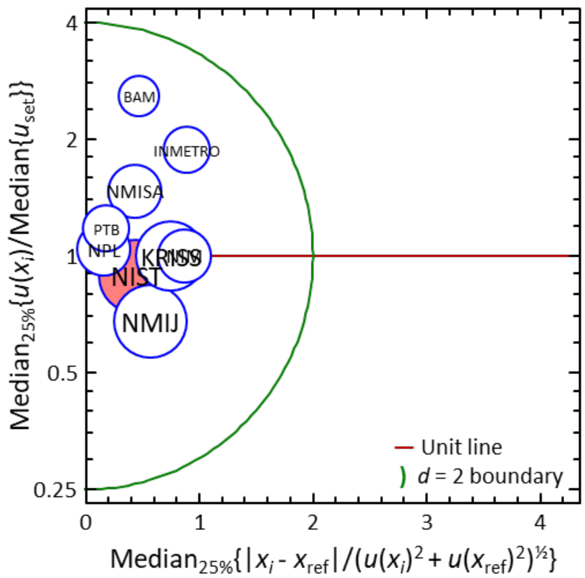


Fig. 26. {SAWG, m} Measurement Performance Peers.

Median measurement differences and uncertainties relative to dataset performance characteristics for organizations that participated with NIST in at least 25 % of the {SAWG, m} studies that NIST participated in. The normalized differences are plotted against the *x*-axis, where *x* = 0 is no difference. The uncertainty ratios are plotted against the *y*-axis, where *y* = 1 marks where the measurement uncertainties are equal. The size of the symbols is related to the number of datasets in common with NIST. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to NIST.

All nine of the organizations with results in these 26 datasets have $d_{\mathrm{set},i}$ within 2 composite distance units of the dataset's characteristic parameters: NPL, PTB, NIST, NMISA, KRISS, NMIJ, NIM, INMETRO, and BAM.

8.8. {IRWG,‰}

The studies used to identify NIST's isotopic δ-scale measurement performance peers are listed in Table 8. The $X_{\mathrm{ref},i}$ and $Y_{\mathrm{set},i}$ values for organizations that co-participated with NIST in {IRWG,‰} studies are displayed in Fig. 27.

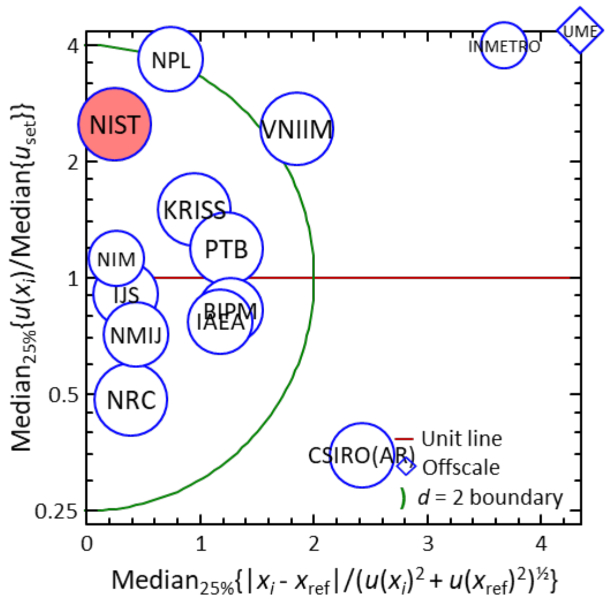


Fig. 27. {IRWG,‰} Measurement Performance Peers.

Median measurement differences and uncertainties relative to dataset performance characteristics for organizations that participated with NIST in at least 25 % of the {IRWG,‰} studies that NIST participated in. The normalized differences are plotted against the *x*-axis, where *x* = 0 is no difference. The uncertainty ratios are plotted against the *y*-axis, where *y* = 1 marks where the measurement uncertainties are equal. The size of the symbols is related to the number of datasets in common with NIST. The horizontal line marks the unit uncertainty ratio. The semicircle bounds the region within 2 composite distance units nearest to NIST.

Nine of the 14 organizations with results in these ten datasets have $d_{\mathrm{set},i}$ within 2 composite distance units of the dataset's characteristic parameters: NIM, IJS, NMIJ, NRC, KRISS, IAEA, PTB, BIPM, and NIST.

9. Power Law Scaling Function Peers

In the early 1980s Horwitz demonstrated that when chemical concentration is expressed as mass- or mole-fraction, x, then interlaboratory reproducibility, σ_R, has about the same dependence on mean concentration, $\bar{x}$, for a very wide range of analytes [18,19]. Thompson later showed that Horwitz's function is equivalent to:

$$\sigma_R = 0.02\bar{x}^{0.8495} \tag{10}$$

and is applicable to concentrations over the "Horwitz interval" from about ($10^{-7.5}$ to 10^{-2}) mass- or mole-fraction [20, 21]. The relationships between reproducibility and concentration below and above this range have different forms [21, 22]. Thompson recently presented a plausible explanation for the function's form and limitations [23].

The Horwitz function is a particular example of the power law fluctuation scaling relationship:

$$s = \alpha\bar{x}^{\beta} . \tag{11}$$

relating measurement standard deviation to measurement mean value for a wide variety of precision metrics and measurement processes [24, 25]. Taking logarithms of both sides, this reduces to the linear equation:

$$y = a + bx \tag{12}$$

where y is $\log(s)$, a is $\log(\alpha)$, b is β, and x is $\log(\bar{x})$.

The linear relationships between the logarithms of NIST's reported {GAWG, mol/mol}, {IAWG, g/g}, and {OAWG, g/g} measurements and their associated standard uncertainties are displayed in Fig. 28. Along with the $\{x, u(x)\}$ pairs for results reported to finalized KCs, SCs, and PPSs, each panel displays three power law scaling functions: a two-parameter (a and b) best-fit, a one-parameter best-fit ($100a$, with b fixed at 1; i.e., a constant coefficient of variation (%CV) expressed as percent), and the Horwitz function. Within the Horwitz interval the best-fit power laws are well below but roughly parallel to the Horwitz function; the $\{x, u(x)\}$ much below this interval are better described by the constant %CV model.

The $\{x, u(x)\}$ pairs within the Horwitz interval display power law scaling, but with differing values for the α and β parameters. The $\{x, u(x)\}$ pairs reported by other participants in CCQM studies display similar behavior (data not shown).

Note: There are too few $\{x, u(x)\}$ currently available for reliable estimation of the scaling behavior of measurands studied by other CCQM WGs that are reported as mol/mol or g/g concentrations. Scaling functions for measurands not related to mole- or mass fraction are unlikely to be valid over the same ($10^{-7.5}$ to 10^{-2}) interval.

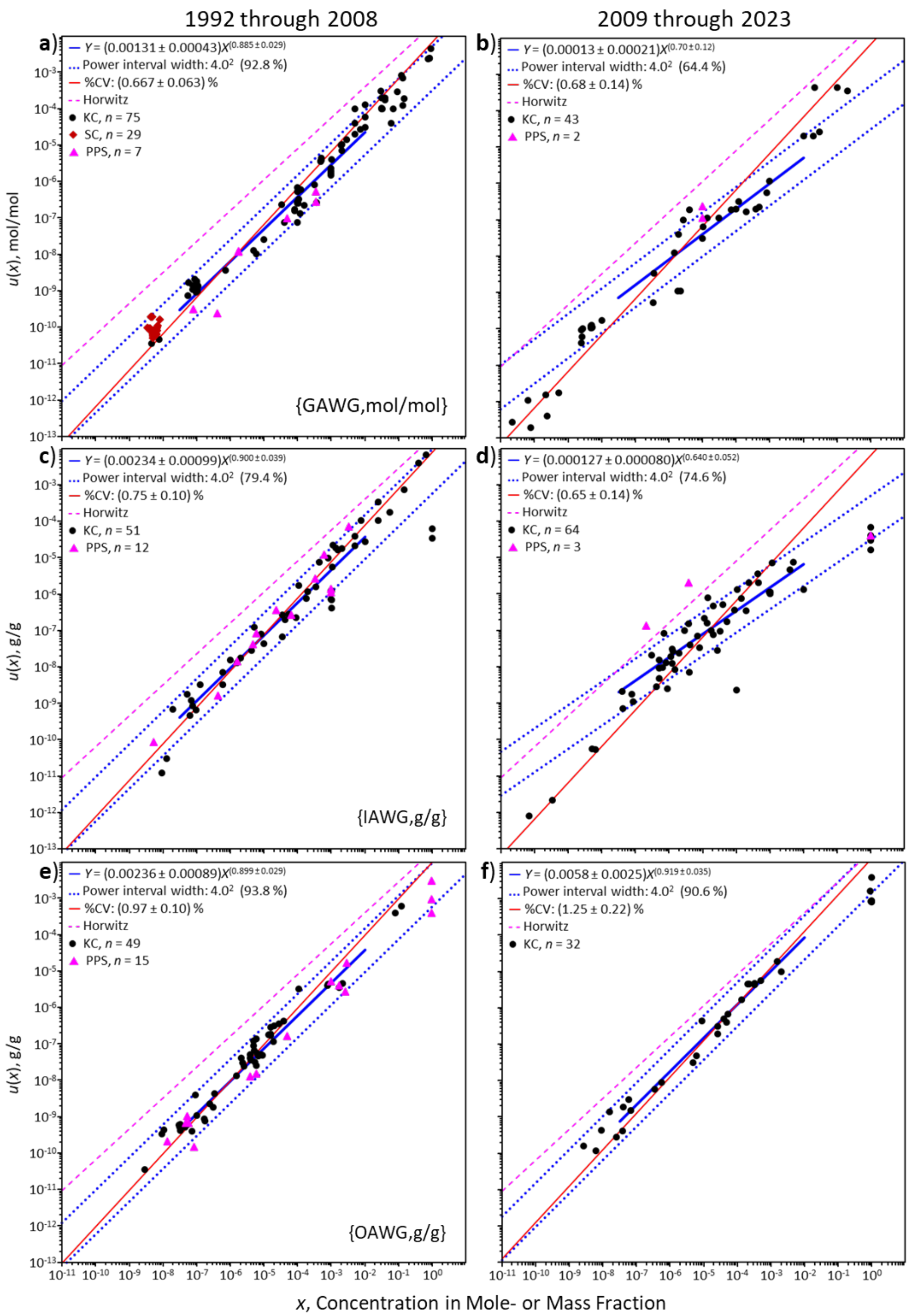


Fig. 28. NIST Uncertainties as Functions of Concentration, Early and Recent Intervals.

Each symbol represents a NIST result its standard uncertainty, $\{x, u(x)\}$. Black circles denote Key Comparisons (KCs), red diamonds supplementary comparisons (SCs), and magenta triangles published pilot studies (PPSs). Solid blue lines spanning the interval $10^{-7.5}$ to 10^{-2} denote best fit power laws, parallel dotted blue lines bound factor-of-four intervals on the best fit. Solid red lines denote best fit coefficients of variation. Dashed magenta lines denote the Horwitz function. Panels a) and b) present {GAWG, mol/mol} results; c) and d) {IAWG, g/g} results; and e) and f) {OAWG, g/g} results.

9.1. Comparing Coefficients

If the relationships between standard uncertainty and concentration values reported by CCQM participants are adequately summarized by the power law's α and β parameters, participants with similar measurement and uncertainty evaluation capabilities should have similar summary parameter values.

The CCQM_Retrospectoscope system estimates the *a* and *b* parameter values of Eq. 12 using Excel's linear regression function, LINEST. In addition to the parameter values, the regression system provides estimates of the parameters' standard uncertainties, $u(a)$ and $u(b)$. The correlation between the two regression parameters depends only on the distance the x values are from zero [26]. The correlation coefficient, r, is separately estimated as

$$r = \sum_i^n x_i / \sqrt{n \sum_i^n x_i^2} \ . \qquad (13)$$

where n is the number of x values. The five parameters (a, $u(a)$, b, $u(b)$, and r) plus a coverage factor, k, suffice to define an approximate 95 % joint coverage interval for a given (*a*,*b*) pair as shown in Fig. 29.

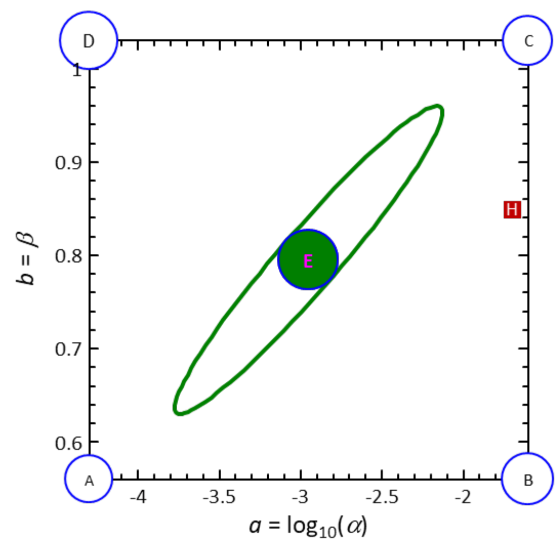


Fig. 29. Example Summary Parameter Display.

Each symbol represents exemplar {*a*,*b*} coefficients plotted in a particular region of possible coefficient values. The ellipse bounding "E" is an approximate 95 % confidence coverage interval. The small square labeled "H" represents the Horwitz function coefficients, {$\log_{10}$(0.02),0.8495}.

The relationships among the power law scaling ($\{a, b\}$) summary coefficient for organizations that contributed to at least 20 % of the {GAWG, mol/mol}, {IAWG, g/g}, or {OAWG, g/g} studies during the Early and Recent intervals are displayed in Fig. 30.

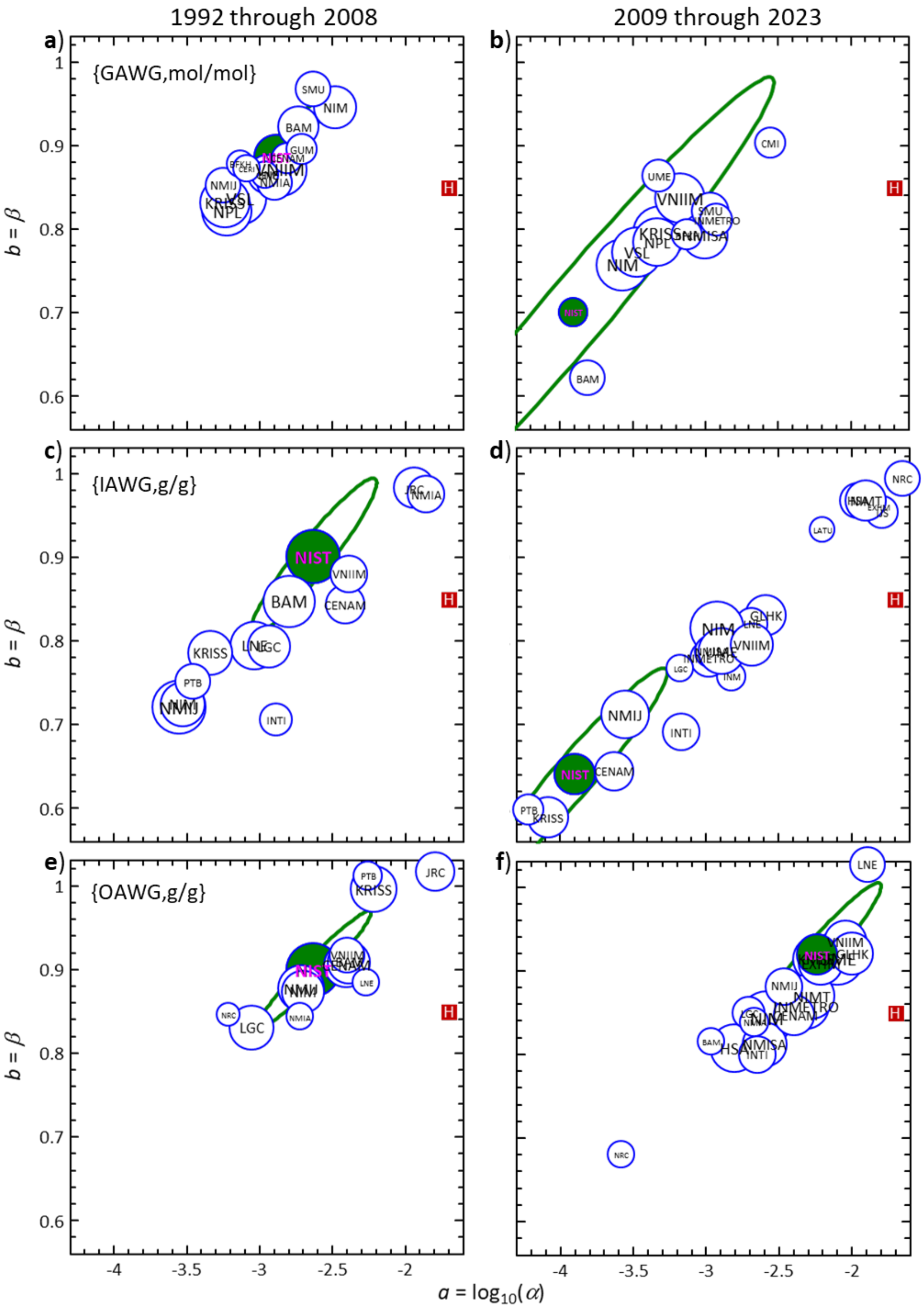


Fig. 30. GAWG, IAWG, and OAWG Power Law Scaling Coefficients, Early and Recent Intervals.
Each circle represents the {*a*,*b*} power law scaling coefficients for the results submitted by one organization during the Early or Recent intervals. The size of the circle is related to the number of results. Panels a) and b) display coefficients for {GAWG, mol/mol} results, c) and d) for {IAWG, g/g} results, and e) and f) for {OAWG, g/g} results. The ellipse in each panel is an approximate 95 % confidence coverage interval on the NIST coefficients (the ellipse in panel a) is hidden by symbols for other organizations.) The small squares labeled "H" represent the Horwitz function coefficients, {$\log_{10}(0.02)$,0.8495}.

9.2. Interpreting the Power Law Coefficients

The “A” to “D” symbols at the corners of Fig. 29 are at equal graphical distances from the central symbol “E”. However, as indicated by the ellipse in Fig. 29, the coefficients $\{a_\mathrm{A}, b_\mathrm{A}\}$ and $\{a_\mathrm{C}, b_\mathrm{C}\}$ have more in common with $\{a_\mathrm{E}, b_\mathrm{E}\}$ than do $\{a_\mathrm{B}, b_\mathrm{B}\}$ or $\{a_\mathrm{D}, b_\mathrm{D}\}$. The graphical distance between symbols is thus at best an incomplete guide to coefficient similarity. Rather than attempting to quantify small differences among similar results, the power law summary graphs are more suited to qualitative visual evaluation.

- {GAWG, mol/mol}
  The relationships between reported mole fractions and their associated uncertainties for the {GAWG, mol/mol} Early interval, Fig. 30a, are very similar for all organizations. NIST’s 95 % coverage ellipse encompasses nearly the same region as the symbols for other organizations and is (almost) completely hidden beneath them. NIST’s {GAWG, mol/mol} Early interval power law peers are thus: BAM, BFKH, CEM, CENAM, CERI, GUM, KRISS, LNE, NIM, NMIA, NMIJ, NPL, SMU, VNIIM, and VSL.

  The {GAWG, mol/mol} Recent interval power law coefficients, Fig. 30b, for NIST and at least two other organization (BAM and CMI) differ from the majority. However, because the uncertainties on the NIST coefficients are large the NIST 95 % joint confidence ellipse encompasses or closely abuts the symbols for the other organizations. For lack of discriminatory power, NIST’s {GAWG, mol/mol} Recent interval power law peers are thus: BAM, BFKH, CMI, INMETRO, KRISS, NIM, NMISA, NPL, SMU, UME, VNIIM, and VSL.
- {IAWG, g/g)
  The power law coefficients for the {IAWG, g/g) Early and Recent interval participants are considerably more diverse than those for the {GAWG, mol/mol} and {OAWG, g/g}. NIST peers during the Early interval include BAM, CENAM, LGC, LNE, and VNIIM. Peers during the Recent Interval include: CENAM, INTI, KRISS, LGC, NMIJ, PTB.
- {OAWG, g/g)
  The power law coefficients for all the {OAWG, g/g) Early and Recent interval participants are all fairly similar, with NIST’s coefficients with the clusters. NIST peers during the Early interval include all but JRC: BAM, CENAM, KRISS, LGC, LNE, NIM, NMIA, NMIJ, NRC, PTB, VNIIM. NIST peers during the Recent interval include all but NRC: BAM, CENAM, EXHM, GLHK, HSA, INMETRO, INTI, KRISS, LGC, LNE, NIM, NMIT, NMIA, NMIJ, NMISA, UME, VNIIM.

Note: Power law scaling coefficients were evaluated for all participants with results in at least 20 % of the datasets for the given {WG, BaseUnit} during the given Early (1992 through 2008) or Recent (2009 through 2023) interval, using all of those datasets. The organizations that participated in 20 % of the {WG, BaseUnit} datasets are not necessarily the same organizations that participated in 25 % of the datasets in which NIST also participated.

10. Blind Analysis Peers

Ideally, coordination of a study does not influence participation results. Practically, complete separation of the roles of analyst and coordinator can be difficult. This is especially challenging for organizations with limited staff. Some idea of how well coordinating organizations "blind" themselves is given in Fig. 31, plotting median measurement performance estimated using datasets for which organizations were not coordinators, $d_{\text{set:without},i}$, relative to their median performance using datasets for which they were coordinators, $d_{\text{set:with},i}$. Results are presented only for organizations that contributed results in at least four "with coordination" and at least four "without coordination" datasets. Given NIST's lower coordination rate during the Recent interval, these results perforce cover the entire 1992 through 2023 interval.

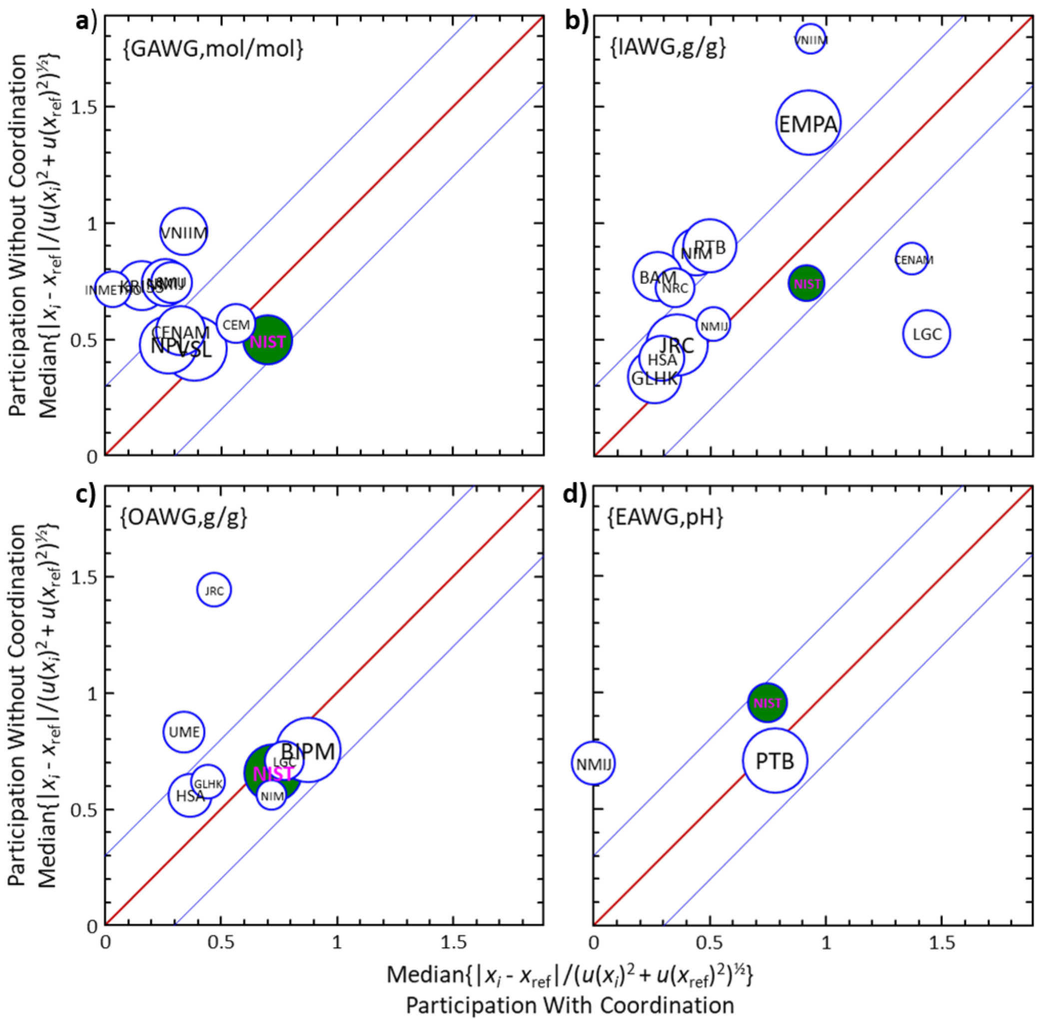


Fig. 31. GAWG, IAWG, OAWG, and EAWG Blind Analysis Peers.

Each circle represents the median measurement performance in datasets where the organization was not a coordinator relative to their performance where they were a coordinator. The size of the circle is related to the number of participations where they were a coordinator relative to the number where they were not a coordinator. Results in all panels are for the entire period of CCQM activity, 1992 through 2023. The NIST result is in green. Panels a) reports the "blind" performance in {GAWG, mol/mol}, b) {IAWG, g/g}, c) in {OAWG, g/g}, and d) {EAWG, pH} datasets.

The distance, d_{blind}, from a point $\{d_{\mathrm{set:with},i}, d_{\mathrm{set:without},i}\}$ to the diagonal equality line, $d_{\mathrm{set:without},i} = d_{\mathrm{set:with},i}$, is [27]

$$d_{\mathrm{blind}} = (d_{\mathrm{set:with}} - d_{\mathrm{set:without}})/\sqrt{2}\,. \quad (14)$$

The maximum NIST distance in the four panels of Fig. 31 is 0.15. Using twice this as an empirical definition of "close", NIST's closest "blind analysis" peers are

- {GAWG, mol/mol)
  While to the left of the equality line, the coordinates for CEM, VSL, NPL, and CENAM are within the ± 0.30 interval.
- {IAWG, g/g)
  While to the left of the equality line, the coordinates for NMIJ, GLHK, JRC, HSA, NRC, PTB, and NIM are within the ± 0.30 interval.

  The coordinates for CENAM and LGC are well to the right of the equality line, indicating that their measurement performance was better when not coordinating studies than when coordinating. One interpretation for this could be that participation was assigned to relatively new analysts while coordination was entrusted to more experienced staff.
- {OAWG, g/g)
  The coordinates for NIM, BIPM, LGC, GLHK, and HSA are within the ± 0.30 interval.
- {EAWG, pH)
  The coordinate for PTB is within the ± 0.30 interval.

11. Summary

The following tables list the organizations identified as NIST's peers in terms of participations (Section 5), coordinations (Section 6), measurement similarity (Section 7), measurement performance (Section 8), power law scaling behavior (Section 9), and blind analysis (Section 10). Checkmarks (✓) denote classification as a peer by the criteria of each section. Within each table organizations are (somewhat subjectively) ordered by the number and nature of peer classifications.

11.1. {GAWG, mol/mol}

Table 9. {GAWG, mol/mol} Peer Organizations, Early and Recent Intervals.

	1992 through 2008					2009 through 2023					All
Organization	Participation	Coordination	Similarity	Performance	Power Law	Participation	Coordination	Similarity	Performance	Power Law	Blind Analysis
CERI	✓		✓	✓	✓	✓	✓	✓	✓		
KRISS	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
NIM	✓		✓	✓	✓	✓	✓	✓	✓	✓	
NMIJ	✓	✓	✓	✓	✓	✓	✓	✓	✓		
NPL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
VNIIM	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
VSL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
BAM	✓	✓	✓	✓	✓	✓	✓			✓	
BFKH	✓		✓	✓	✓	✓				✓	
LNE	✓	✓	✓	✓	✓	✓		✓	✓		
NMISA	✓	✓				✓		✓	✓	✓	
A*STAR						✓					
BIPM		✓					✓				
CEM	✓	✓			✓						✓
CENAM	✓				✓		✓				✓
CMI										✓	
EAA			✓	✓							
ERLAP			✓	✓							
GUM	✓				✓	✓					
INMETRO						✓	✓			✓	
IPQ	✓						✓				
METAS	✓					✓					
NMIA	✓				✓	✓					
NPLI						✓					
SMU	✓				✓	✓	✓			✓	
UME						✓				✓	

CERI, KRISS, NIM, NMIJ, NPL, VNIIM, and VSL have been NIST's primary gas analysis peers during both the Early and Recent intervals. BAM, BFKH, and LNE were primary peers during the Early interval. NIMSA has become a primary peer in the Recent interval.

11.2. {IAWG, g/g}

Table 10. {IAWG, g/g} Peer Organizations, Early and Recent Intervals.

	1992 through 2008					2009 through 2023					All
Organization	Participation	Coordination	Similarity	Performance	Power Law	Participation	Coordination	Similarity	Performance	Power Law	Blind Analysis
BAM	✓	✓	✓	✓	✓	✓	✓	✓	✓		
CENAM	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
KRISS	✓	✓	✓	✓		✓	✓	✓	✓	✓	
LGC	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
LNE	✓	✓	✓	✓	✓	✓	✓	✓	✓		
NIM	✓	✓	✓	✓		✓	✓	✓	✓		✓
NMIA	✓		✓	✓		✓		✓	✓		
NMIJ	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓
NRC	✓	✓	✓	✓		✓	✓	✓	✓		✓
PTB	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓
JRC	✓	✓	✓	✓							✓
VNIIM	✓			✓	✓	✓	✓				
GLHK		✓				✓	✓	✓	✓		✓
HSA						✓	✓	✓	✓		✓
INMETRO	✓					✓	✓		✓		
IJS						✓	✓		✓		
INTI	✓					✓	✓			✓	
NIMT						✓		✓	✓		
NMISA	✓					✓		✓	✓		
UME	✓					✓	✓	✓	✓		
EMPA		✓									
EXHM						✓					
GUM						✓					
IAEA							✓				
IMBIH							✓				
INACAL						✓					
INM						✓			✓		
INRIM	✓										
ISP						✓	✓				
KEBS						✓					
LATU						✓					
RISE	✓	✓									
SMU	✓	✓				✓	✓				

BAM, CENAM, KRISS, LGC, LNE, NIM, NMIA, NMIJ, NRC, and PTB are NIST's primary inorganic analysis peers during both the Early and Recent intervals. JRC was primary during the Early interval. GLHK, HSA, INMETRO, NMISA, and UME have become primary peers in the Recent interval.

11.3. {OAWG, g/g}

Table 11. {OAWG, g/g} Peer Organizations, Early and Recent Intervals.

	1992 through 2008					2009 through 2023					All
Organization	Participation	Coordination	Similarity	Performance	Power Law	Participation	Coordination	Similarity	Performance	Power Law	Blind Analysis
BAM	✓	✓	✓	✓	✓	✓	✓		✓	✓	
CENAM	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
KRISS	✓	✓	✓	✓	✓	✓		✓	✓	✓	
LGC	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
NIM	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
NMIA	✓	✓	✓	✓	✓	✓		✓	✓	✓	
NMIJ	✓		✓	✓	✓	✓	✓	✓	✓	✓	
NRC	✓	✓	✓	✓	✓	✓	✓	✓	✓		
PTB	✓	✓	✓	✓	✓		✓	✓	✓		
JRC	✓	✓	✓	✓			✓				
BIPM		✓				✓	✓	✓	✓		✓
EXHM						✓		✓	✓	✓	
GLHK	✓					✓	✓	✓	✓	✓	✓
HSA						✓	✓	✓	✓	✓	✓
INMETRO						✓	✓	✓	✓	✓	
LNE	✓				✓	✓	✓	✓	✓	✓	
NIMT						✓		✓	✓	✓	
NMISA						✓	✓	✓	✓	✓	
UME						✓	✓	✓	✓	✓	
VNIIM	✓				✓	✓			✓	✓	
BVL						✓	✓				
INTI						✓				✓	
KEBS						✓					
LATU						✓					
VSL							✓				

BAM, CENAM, KRISS, LGC, NIM, NMIA, NMIJ, NRC, and PTB are NIST's primary organic analysis peers during both the Early and Recent intervals. JRC was a primary peer during the Early interval. BIPM, GLHK, HSA, INMETRO, LNE, NIMT, NMISA, and UME have become primary peers in the Recent interval.

11.4. {EAWG, pH}

Table 12. {EAWG, pH} Peer Organizations, Early and Recent Intervals.

	1992 through 2008				2009 through 2023				All
Organization	Participation	Coordination	Similarity	Performance	Participation	Coordination	Similarity	Performance	Blind Analysis
NMIJ	✓	✓	✓	✓	✓	✓	✓	✓	
PTB	✓	✓	✓	✓	✓	✓	✓	✓	✓
SMU	✓	✓	✓	✓	✓	✓	✓	✓	
VNIIFTRI	✓		✓	✓	✓	✓		✓	
CENAM	✓		✓	✓	✓			✓	
DPL	✓		✓	✓					
GUM	✓		✓	✓	✓				
NIM	✓		✓	✓	✓				
BFKH					✓		✓	✓	
DFM	✓				✓		✓	✓	
INACAL					✓		✓	✓	
INMETRO	✓				✓	✓	✓	✓	
LNE	✓				✓	✓	✓		
BIM	✓		✓		✓				
CMI	✓			✓	✓				
GLHK					✓				
IBMETRO					✓				
INPL	✓								
KRISS	✓								
LATU					✓				
NIMT					✓				
NMIM					✓				
NPL						✓			
UME					✓				
UMTS					✓				
VMI					✓				

NMIJ, PTB, SMU, and VNIIFTRI are NIST's primary pH measurement peers during both the Early and Recent intervals. DPL and GUM were primary peers during the Early interval. BFKH, DFM, INACAL, and INMETRO have become primary peers in the Recent interval.

11.5. {EAWG, S/m}, {SAWG, m}, and {IRWG, ‰}

Table 13. {EAWG, S/m}, {SAWG, m}, and {IRWG, ‰} Peer Organizations.

{EAWG, S/m}

Organization	Participation	Coordination	Similarity	Performance
BFKH	✓		✓	
DFM	✓	✓	✓	✓
GUM	✓		✓	✓
INPL	✓		✓	✓
PTB	✓	✓	✓	✓
SMU	✓	✓	✓	✓
CENAM	✓			
CM I	✓			
INACAL	✓			
INMETRO	✓	✓		
INRIM	✓	✓		
NIM	✓			
NMIJ	✓			
RISE	✓			
UMTS	✓	✓		
VNIIFTRI	✓			
VNIIM	✓			
VSL		✓		

{SAWG, m}

Organization	Participation	Coordination	Similarity	Performance
BAM	✓	✓	✓	✓
KRISS	✓	✓	✓	✓
NPL	✓	✓	✓	✓
NIM	✓		✓	✓
NMIJ	✓		✓	✓
NMISA	✓		✓	✓
PTB	✓		✓	✓
INMETRO	✓			✓
VNIIM	✓	✓		

{IRWG, ‰}

Organization	Participation	Coordination	Similarity	Performance
BIPM		✓	✓	✓
IAEA		✓	✓	✓
IJS	✓		✓	✓
KRISS	✓		✓	✓
NIM	✓		✓	✓
NMIJ	✓		✓	✓
NRC	✓	✓		✓
PTB	✓	✓	✓	✓
VNIIM	✓		✓	
BAM	✓	✓		
INMETRO	✓			
JRC		✓		
LGC	✓	✓		
NMIA	✓			
NPL			✓	
UME	✓			

BFKH, DFM, GUM, INPL, PTB, and SMU were NIST's primary peers until NIST stopped ceased active support of t electrolytic conductivity measurements in 2014.

BAM, KRISS, NPL, NIM, NMIJ, NMISA, and PTB are NIST's primary thin film thickness analysis peers.

BIPM, IAEA, IJS, KRISS, NIM, NMIJ, NRC, PTB, and VNIIM are NIST's primary isotopic δ-scale measurement peers.

11.6. PAWG, NAWG, and CAWG

Table 14. PAWG, NAWG, and CAWG Peer Organizations.

PAWG

Organization	Participation	Coordination
BIPM	✓	✓
HSA	✓	✓
KRISS	✓	✓
LGC	✓	✓
LNE	✓	✓
NIBSC	✓	✓
NIM	✓	✓
NPL	✓	✓
NRC	✓	✓
PTB	✓	✓
BAM	✓	
INMETRO	✓	
JRC	✓	
NMIJ	✓	
UME	✓	

NAWG

Organization	Participation	Coordination
GLHK	✓	✓
JRC	✓	✓
KRISS	✓	✓
LGC	✓	✓
NIM	✓	✓
NIMT	✓	✓
NMIA	✓	✓
NRC	✓	✓
UME	✓	✓
CENAM	✓	
DMSC	✓	
INM(CO)	✓	
INMETRO	✓	
NIB	✓	
NMIJ	✓	
PTB	✓	
VNIIM	✓	
NIBSC		✓

CAWG

Organization	Participation	Coordination
INRIM	✓	✓
LGC	✓	✓
NIBSC	✓	✓
NIM	✓	✓
PTB	✓	✓
NMIJ		✓
NPL		✓

All participants in the PAWG, NAWG, and CAWG studies are primary NIST peers.

References

[1] Duewer DL, Cecelski CE, Cleveland MH, Lin NJ, Lippa KA, Mann JL, Molloy JL, Nelson MA, Norris JE, Pratt KW, Rhoderick GC, Schantz MM, Semerjian HG, Sarkar S, Watters RL, Windover DA (2024) NIST's Involvement with CCQM Studies from 1992 to 2023: History and Performance. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Internal Report (NIST IR) 8524. https://doi.org/10.6028/NIST.IR.8524.

[2] Duewer DL (2023) CCQM_Retrospectoscope. https://doi.org/10.18434/mds2-2952. See also: Duewer DL, Nelson MA, Cecelski CE. CCQM_Retrospectoscope Quick Start: A beginner's guide to a data analysis system for visualizing results from CCQM studies (2023) NIST Internal Report (NIST IR) 8487. https://doi.org/10.6028/NIST.IR.8487.

[3] Duewer DL, Nelson MA, Cecelski CE. CCQM_Retrospectoscope Reference Manual. A detailed reference to the CCQM_Retrospectoscope, a suite of graphical tools for the meta-analysis of measurement results from CCQM studies (2023) NIST Internal Report (NIST IR) 8478. https://doi.org/10.6028/NIST.IR.8478.

[4] CIPM-MRA-2003. Mutual recognition of national measurement standards and of calibration and measurement certificates issued by national metrology institutes, Paris, 14 October 1999. Technical Supplement revised in October 2003 (pages 38-41). https://www.bipm.org/documents/20126/43742162/CIPM-MRA-2003.pdf. Accessed: 8/7/2024.

[5] Werhahn O, Olson DA, Kuanbayev C, Henson A (2023) The CIPM MRA – success and performance. Metrologia 60(4) 042001. https://doi.org/10.1088/1681-7575/ace191.

[6] CIPM MRA-P-12 (2021) Coordination within the CIPM MRA. Consultative Committees, Regional Metrology Organizations, JCRB. https://www.bipm.org/documents/20126/43742162/CIPM-MRA-P-12.pdf. Accessed: 8/7/2024.

[7] BIPM. Our mission and objectives. https://www.bipm.org/en/mission-objectives. Accessed 8/7/2024.

[8] Key Comparison Database (KCDB). https://www.bipm.org/kcdb/. Accessed 8/7/2024.

[9] BIPM. Consultative Committees of the CIPM. https://www.bipm.org/en/committees/cc. Accessed 8/7/2024.

[10] BIPM. Consultative Committee for Amount of Substance: Metrology in Chemistry and Biology (CCQM). https://www.bipm.org/en/committees/cc/ccqm. Accessed 8/7/2024.

[11] CCQM IRWG. Strategy for Rolling Programme Development (2021-2030). https://www.bipm.org/documents/20126/2071059/CCQM-IRWG Strategy document 2021-2030.pdf. Accessed 8/7/2024.

[12] CCQM-NAWG. Strategy Document 2021-2030. https://www.bipm.org/documents/20126/2071059/CCQM-NAWG Strategy document 2021-2030.pdf. Accessed 8/7/2024.

[13] CIPM MRA-G-11 (2021) Measurement comparisons in the CIPM MRA. Guidelines for organizing, participating and reporting. https://www.bipm.org/documents/20126/43742162/CIPM-MRA-G-11.pdf. Accessed: 8/7/2024.

[14] CIPM MRA-G-13 (2022) Calibration and measurement capabilities in the context of the CIPM MRA. https://www.bipm.org/documents/20126/43742162/CIPM-MRA-G-13.pdf. Accessed: 8/7/2024.

[15] JCGM/WG2 (2012. JCGM 200:2012 International vocabulary of metrology – Basic and general concepts and associated terms (VIM). https://www.bipm.org/en/publications/guides/vim.html.

[16] Hibbert DB (2023) Compendium of Terminology in Analytical Chemistry, IUPAC Orange Book, The Royal Society of Chemistry. https://doi.org/10.1039/9781788012881.

[17] International Organization for Standardization. (2022). Statistical methods for use in proficiency testing by interlaboratory comparison, ISO Standard 13528:2022. https://www.iso.org/standard/78879.html.

[18] Horwitz W, Kamps LR, Boyer KW (1980) Quality assurance in the analysis of foods and trace constituents. J Assoc Off Anal Chem 63:1344–654. https://doi.org/10.1093/jaoac/63.6.1344.

[19] Horwitz W (1982) Evaluation of Analytical. Methods Used for Regulation of Foods and Drugs. Anal Chem 54(1):67A-76A. https://doi.org/10.1021/ac00238a002.

[20] Thompson M, Lowthian PJ (1997) The Horwitz Function Revisited. J Assoc Off Anal Chem 80(3):676-679. https://doi.org/10.1093/jaoac/80.3.676.

[21] Thompson M, Sykes M, Mathieson K. Wood R (2022) Comparison of reproducibility precision on mass fraction in some interlaboratory studies of methods of food analysis. Anal Bioanal Chem 414:1105–1114. https://doi.org/10.1007/s00216-021-03736-3.

[22] Thompson M (2013) An emergent optimal precision in chemical measurement at low concentrations. Anal Meth 5:4518–9. https://doi.org/10.1039/C3AY40557E.

[23] Thompson M (2022) Towards an explanation of the Horwitz function. Anal Bioanal Chem 414:1671-1676. https://doi.org/10.1007/s00216-021-03791-w.

[24] Hanley QS (2016) Chemical Measurement and Fluctuation Scaling. Anal Chem 88(24):12036-12042. https://doi.org/10.1021/acs.analchem.6b02335.

[25] Holland R, Rebmann R, Williams C, Henley QS (2017) Fluctuation Scaling, Calibration of Dispersion, and Detection of Differences. Anal Chem 89(21):11568-11575. https://doi.org/10.1021/acs.analchem.7b02909.

[26] Analytical Methods Committee (2019) The correlation between regression coefficients: combined significance testing for calibration and quantitation of bias. Anal Methods 11:1845-1848. https://doi.org/10.1039/c9ay90041a.

27 Wikipedia. Distance from a point to a line. https://en.wikipedia.org/wiki/Distance_from_a_point_to_a_line. Accessed 6/14/2024.

Appendix A. List of Symbols, Abbreviations, and Acronyms

A.1. Abbreviations and Acronyms

BIPM	Bureau International des Poids et Mesures
CAWG	Cell Analysis Working Group
CC	consultative committee
CCQM	Originally "Consultative Committee for the Quantity of Matter"; now "Consultative Committee for the Amount of Substance: Metrology in Chemistry and Biology"
CGPM:	General Conference on Weights and Measures
CIPM MRA	Comité International des Poids et Mesures Mutual Recognition Arrangement
CMCs	Calibration and Measurement Capabilities
DI	designated institute, an organization having the responsibility for a specified aspect of a nation's measurement infrastructure. Only NMIs and DIs can participate in CCQM KCs.
EAWG	Electrochemical Analysis Working Group
GAWG	Gas Analysis Working Group
IAEA	International Atomic Energy Agency
IAWG	Inorganic Analysis Working Group
IRWG	Isotope Ratio Working Group
JRC	European Joint Research Centre
KC	Key Comparison
KCDB	Key Comparison Database
NAWG	Nucleic Acid Analysis Working Group
NIST	National Institute of Standards and Technology
NMI	national metrology institute, an organization having broad responsibility for a nation's measurement infrastructure not delegated to a DI. Only NMIs and DIs can participate in CCQM KCs.
OAWG	Organic Analysis Working Group
PAWG	Protein Analysis Working Group
PPS	Published pilot study
PS	pilot study
RMO	Regional Metrology Organization
SAWG	Surface Analysis Working Group
SC	Supplementary Comparison
WG	Working Group
WMO	World Meteorological Organization

A.2. Symbols

α	power law scale parameter
β	power law exponent parameter
σ_{R}	Interlaboratory reproducibility
a	intercept parameter for the linearize power law, equal to $\log_{10}(\alpha)$
b	slope parameter for the linearized power law, equal to β
d_{MR}	metrological responsibility distance (Eq. 2)
d_{NIST}	similarity-to-NIST distance (Eq. 5)
d_{set}	similarity-to-dataset-characteristics distance (Eq. 9)
d_{blind}	blind analysis distance (Eq. 14)
i	index
k	coverage factor, locally defined
n	number, locally defined
N	number, locally defined
p_{WG}	proportion of studies in which an organization participates (Eq. 1)
r	correlation coefficient
s	standard deviation
$u(\cdot)$	standard uncertainty
x	value of a variable
X	value of a variable used to estimate a distance
y	value of a variable
Y	value of a variable used to estimate a distance

A.3. Functions

$\log_2(\cdot)$	binary logarithm of a value
$\log_{10}(\cdot)$	decadic logarithm of a value
$\mathrm{Median}\{\cdot\}$	median value of a set of values
$\lvert\cdot\rvert$	absolute value of a value

Appendix B. Codenames for National Metrology and Designated Institutes

The following information is derived from the BIPM's "CIPM MRA participants" webpage, https://www.bipm.org/en/cipm-mra/participation. "Status" refers to the organization's CIPM MRA participation status: as a national Signatory/NMI (NMI), an international Signatory/NMI (IntOrg), or a designated institute of a signatory (DI).

Code	Country	Organization Name	Status
A*STAR	Singapore	National Metrology Centre, Agency for Science, Technology and Research	NMI
BAM	Germany	Bundesanstalt für Materialforschung und -prüfung	DI
BFKH	Hungary	Government Office of the Capital City Budapest	NMI
BIM	Bulgaria	Bulgarian Institute of Metrology	NMI
BIPM	CIPM	Bureau International des Poids et Mesures	IntOrg
BVL	Germany	Federal Office of Consumer Protection and Food Safety	DI
CEM	Spain	Centro Español de Metrología	NMI
CENAM	Mexico	Centro Nacional de Metrología	NMI
CERI	Japan	Chemicals Evaluation and Research Institute	DI
CMI	Czech Republic	Ceský metrologický institut	NMI
DFM	Denmark	Danish Institute of Fundamental Metrology	NMI
DMSc	Thailand	Department of Medical Science	DI
DPL	Denmark	RadiometerA/S	DI
EAA	Austria	Environment Agency Austria (Umweltbundesamt GmbH)	DI
EMPA	WMO	Swiss Federal Laboratories for Materials Science and Technology	DI
ERLAP	EU	European Reference Laboratory of Air Pollution	IntOrg
EXHM	Greece	National Laboratory of Chemical Metrology	DI
GLHK	Hong Kong	Government Laboratory (Hong Kong)	DI
GUM	Poland	Główny Urząd Miar	NMI
HSA	Singapore	Health Sciences Authority	DI
IAEA	UN	International Atomic Energy Agency	IntOrg
IBMETRO	Bolivia	Instituto Boliviano de Metrología	NMI
IJS	Slovenia	Jozef Stefan Institute	DI
IMBiH	Bosnia-Herzegovina	Institute of Metrology of Bosnia-Herzegovina	NMI
INACAL	Peru	Instituto Nacional de Calidad	NMI
INM	Romania	National Institute of Metrology	NMI
INM(CO)	Colombia	Instituto Nacional de Metrología de Colombia	NMI
INMETRO	Brazil	Instituto Nacional de Metrologia, Normalização e Qualidade Industrial	NMI
INPL	Israel	National Physical Laboratory of Israel	NMI
INRIM	Italy	Istituto Nazionale di Ricerca Metrologica	NMI
INTI	Argentina	Instituto Nacional de Tecnología Industrial	NMI
IPQ	Portugal	Instituto Português da Qualidade	NMI
ISP	Chile	Instituto de Salud Pública	DI
JRC	EU	Joint Research Centre - JRC - European Commission	IntOrg
KEBS	Kenya	Kenya Bureau of Standards	NMI
KRISS	Republic of Korea	Korea Research Institute of Standards and Science	NMI
LATU	Uruguay	Laboratorio Tecnológico del Uruguay	NMI
LGC	UK	formerly "Laboratory of the Government Chemist"	DI
LNE	France	Laboratoire National de Métrologie et d'Essais	NMI
METAS	Switzerland	Swiss Federal Office of Metrology and Accreditation	NMI
NIB	Slovenia	National Institute of Biology	DI
NIBSC	UK	National Institute for Biological Standards and Control	DI
NIM	China	National Institute of Metrology	NMI
NIMT	Thailand	National Institute of Metrology	NMI
NIST	USA	National Institute of Standards and Technology	NMI
NMIA	Australia	National Measurement Institute of Australia	NMI
NMIJ	Japan	National Metrology Institute of Japan	NMI
NMIM	Malaysia	National Metrology Institute of Malaysia	NMI
NMISA	South Africa	National Metrology Institute of South Africa	NMI

Code	Country	Organization Name	Status
NPL	UK	National Physical Laboratory	NMI
NPLI	India	National Physical Laboratory of India	NMI
NRC	Canada	National Research Council of Canada, Institute for National Measurement Standards	NMI
PTB	Germany	Physikalisch-Technische Bundesanstalt	NMI
RISE	Sweden	Research Institutes of Sweden AB	NMI
SMU	Slovakia	Slovenský metrologický ústav	NMI
UME	Türkiye	TÜBITAK Ulusal Metroloji Enstitüsü	NMI
UMTS	Ukraine	All-Ukrainian State Scientific and Production Center of Standardization, Metrology, Certification and Consumer' Rights Protection	DI
VMI	Vietnam	Vietnam Metrology Institute	NMI
VNIIFTRI	Russian Federation	Institute for Physical-Technical and Radiotechnical Measurements	DI
VNIIM	Russian Federation	D.I. Mendeleyev Institute for Metrology	DI
VSL	Netherlands	Van Swinden Laboratorium	NMI

Appendix C. Where the Figures and Tables Came From

All information presented in this document' is current as of its publication date.

C.1. Figures

Except for Fig. 9, figures in this document were generated using one of the following *CCQM_Retrospectoscope* subsystems:

- **Lab_Activity**: Pie chart representations of the relative proportions of datasets and studies in each WG (Fig. 1, Fig. 3). Output from this subsystem provided the information used to generate Fig. 2.
- **Lab_Priorities**: Radar chart representation of the proportion of participations in the CCQM WGs (Fig. 4).
- **WG_Participations**: Bar chart representation of the number of participations in CCQM KC and PS and RMO KC, SC, and PS studies (Fig. 5, Fig. 6).
- **WG_Coordinations**: Bar chart representation of the number of coordinations of CCQM KC and PS and RMO KC, SC, and PS studies (Fig. 7, Fig. 8).
- **Peer_Bilateral**: Scattergram display of median reported uncertainty relative to NIST uncertainty as a function of median scaled bias from NIST result (Fig. 10 to Fig. 16).
- **Peer_Unilateral**: Scattergram display of median reported uncertainty relative to median participant uncertainties as a function of median scaled bias from reference value (Fig. 17, Fig. 19, Fig. 21, Fig. 23, Fig. 25 to Fig. 27). Output from this subsystem provided the information used to generate Fig. 18, Fig. 20, Fig. 22, and Fig. 24.
- **Lab_Uncertainty**: Scattergram display of NIST's standard uncertainties as a function of reported concentrations (Fig. 28).
- **WG_Power**: Scattergram display of power law scaling parameters (Fig. 29, Fig. 30). **Peer_Global** was used to identify organizations with suitable data. **Lab_Uncertainty** is used to evaluate the power law scaling parameters.
- **WG_Diagonal**: Scattergram display of measurement performance in studies coordinated by the organization as a function of measurement performance in studies coordinated by other organizations (Fig. 31).

C.2. Tables

Except for Table 9 to Table 14, tables in this document were generated or informed by one of the following *CCQM_Retrospectoscope* subsystems:

- **Database_Checkup**: Summaries of the number of datasets and studies by sponsoring body, WG, study type, and BaseUnits (Table 1).
- **Dataset_Locate**: Number of finalized studies and their datasets (Table 2 to Table 8).