



## IUPAC Technical Report

David G. Shaw\*, Ian Bruno, Stuart Chalk, Glenn Hefter, David Brynn Hibbert, Robin A. Hutchinson, M. Clara F. Magalhães, Joseph Magee, Leah R. McEwen, John Rumble, Gregory T. Russell, Earle Waghorne, Thomas Walczyk and Timothy J. Wallington

# Chemical data evaluation: general considerations and approaches for IUPAC projects and the chemistry community (IUPAC Technical Report)

<https://doi.org/10.1515/pac-2022-0802>

Received August 10, 2022; accepted June 15, 2023

**Abstract:** The International Union of Pure and Applied Chemistry has a long tradition of supporting the compilation of chemical data and their evaluation through direct projects, nomenclature and terminology work, and partnerships with international scientific bodies, government agencies, and other organizations. The IUPAC Interdivisional Subcommittee on Critical Evaluation of Data has been established to provide guidance on issues related to the evaluation of chemical data. In this first report, we define the general principles of the evaluation of scientific data and describe best practices and approaches to data evaluation in chemistry.

**Keywords:** Chemical data; critical evaluation; IUPAC; measurement uncertainty; metrology; standard property values.

---

\*Corresponding author: **David G. Shaw**, Dept. Chemistry and Institute of Marine Science, Univ. Alaska, Fairbanks, USA, e-mail: dgshaw@alaska.edu. <https://orcid.org/0000-0001-8914-2862>

**Ian Bruno**, Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK. <https://orcid.org/0000-0003-4901-9936>

**Stuart Chalk**, Dept. Chemistry, Univ. North Florida, Jacksonville, FL 32224, USA. <https://orcid.org/0000-0002-0703-7776>

**Glenn Hefter**, Chemistry Dept., Murdoch Univ., Murdoch, WA 6150, Australia. <https://orcid.org/0000-0001-9388-2783>

**David Brynn Hibbert**, School of Chemistry, UNSW Sydney, Sydney, NSW 2052, Australia. <https://orcid.org/0000-0001-9210-2941>

**Robin A. Hutchinson**, Dept. Chemical Engineering, Queen's University, Kingston, ON K7L 3N6, Canada. <https://orcid.org/0000-0002-0225-7534>

**M. Clara F. Magalhães**, School of Biological, Earth and Environmental Sciences, UNSW Sydney, Sydney, Australia; and LEAF – Linking Landscape, Environment, Agriculture and Food Research Center, Associated Laboratory TERRA, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal. <https://orcid.org/0000-0002-3359-850X>

**Joseph Magee**, National Institute of Standards and Technology, Applied Chemicals and Materials Division, Thermodynamics Research Center, Boulder, CO 80305-3328, USA. <https://orcid.org/0000-0002-9312-8593>

**Leah R. McEwen**, Physical Sciences Library, Cornell Univ., Ithaca, NY 14853, USA. <https://orcid.org/0000-0003-2968-1674>

**John Rumble**, R&R Data Services, Gaithersburg, MD, USA. <https://orcid.org/0000-0001-6705-5768>

**Gregory T. Russell**, School of Physical and Chemical Sciences, Private Bag 4800, Christchurch 8140, New Zealand. <https://orcid.org/0000-0002-6773-1428>

**Earle Waghorne**, School of Chemistry, Univ. College Dublin, Dublin, D04 V1W8, Ireland. <https://orcid.org/0000-0002-5271-0392>

**Thomas Walczyk**, Department of Chemistry, National University of Singapore, 117543 Singapore, Singapore. <https://orcid.org/0000-0002-9933-3483>

**Timothy J. Wallington**, Research & Advanced Engineering, Ford Motor Company, Dearborn, MI 48121, USA. <https://orcid.org/0000-0002-9810-6326>

## CONTENTS

<b>1</b>	<b>Introduction</b> .....	<b>1108</b>
<b>2</b>	<b>The nature and importance of chemical data evaluation</b> .....	<b>1109</b>
<b>3</b>	<b>Principal approaches to data evaluation</b> .....	<b>1110</b>
<b>4</b>	<b>Sources of uncertainty in chemical measurement</b> .....	<b>1111</b>
<b>5</b>	<b>Evaluation of measurement uncertainty in chemical data</b> .....	<b>1112</b>
<b>6</b>	<b>Disseminating data evaluation results</b> .....	<b>1114</b>
<b>7</b>	<b>Digital expression and accessibility of critically evaluated data</b> .....	<b>1114</b>
<b>8</b>	<b>Current IUPAC data evaluation activities</b> .....	<b>1115</b>
8.1	Standard Atomic Weights and isotopic abundances .....	1115
8.2	Atmospheric chemical kinetic data .....	1116
8.3	Solubility data .....	1116
8.4	Stability constant data .....	1116
8.5	Polymerization kinetics data .....	1117
<b>9</b>	<b>Conclusions</b> .....	<b>1117</b>
<b>10</b>	<b>Glossary of abbreviations and acronyms used in this guide</b> .....	<b>1118</b>
<b>11</b>	<b>Membership of sponsoring bodies</b> .....	<b>1118</b>
	<b>References</b> .....	<b>1119</b>

## 1 Introduction

At the time of writing, more than 204 million<sup>1</sup> characterized chemical substances have been identified in the CAS Registry [1], one of the world's largest substance databases in chemistry. Substances are characterized in a variety of ways by measurements that cover dozens of physical or chemical properties. With repeated measurements of the same property by various techniques over space and time, the number of measured property values in the peer-reviewed literature is vast and growing. Reported measurement results, however, may differ in quality, (defined as the “degree to which a set of inherent characteristics of an object fulfills requirements” [2]) and may not agree with one another. Moreover, the experimental information necessary to assess data quality [3] is incomplete or absent in many measurement reports. With many data for a given property to choose from and multiple sources of error in the underlying measurements, which are commonly difficult to identify for non-specialists, chemists and non-chemists alike depend on the critical evaluation of available data by experts for provision of preferred values for practical use.

To give guidance about how to design data evaluation projects, how to evaluate data for quality, and what needs to be considered to make such evaluations reliable and traceable over time, the Interdivisional Subcommittee on Critical Evaluation of Data (ISCED) was instituted in 2018 under the umbrella of IUPAC. This technical report is the first in a projected series providing guidance on the critical evaluation of chemical data for both preparers and users of such data, drawing on decades of experience gained from critical evaluations prepared under IUPAC auspices.

The primary audience for this report is any person or group performing critical evaluation of chemical data or wishing to do so. Individuals are usually part of a group having a thorough understanding of the methodology and measurement processes that have yielded the data to be evaluated, as well as the necessary expertise to identify potential sources of uncertainty in reported measurements for the assessment of the quality of measurement results. These groups may also include experts such as statisticians and data scientists involved in the data evaluation process. While special reference is made to IUPAC standards and procedures, guidelines should also be useful to groups working outside the IUPAC framework.

---

<sup>1</sup> This is the value as of December 31, 2022.

A secondary audience for this report includes people who interact with critically evaluated data such as (1) users of data who desire to better understand the characteristics and limitations of evaluated data, (2) producers of data who seek evaluation of their data, and (3) those who want to understand what constitutes a measurement of high quality, and what additional information concerning the measurement procedure needs to be provided for independent comparison and assessment of measurements.

Here we present general considerations and approaches to evaluation. Other planned papers in this series will provide a glossary of terms related to data evaluation, detailed approaches for evaluation of chemical data, dissemination and data standards, and an outline of strategies chosen in some current IUPAC data evaluation projects.

## 2 The nature and importance of chemical data evaluation

In this paper, the ISCED defines<sup>2</sup> chemical data as ‘data characterizing a property of a chemical substance or interactions of chemical substances.’ Thus, chemical data include quantity values revealed by measurement of the composition, structure, physical characteristics, changes, reactions and transformations of chemical elements and compounds.

Chemical data are of great interest and importance beyond the community that generates these data. Human desires to understand, predict, control, and manipulate materials and processes lead to application of chemical data in a wide array of activities, which range from teaching and research to trade and commerce to the protection of the health and well-being of the individual and environment.

For each of the millions of chemical substances that has been isolated and characterized, quantities measured include structure, reactivity, energetics, and many other properties. For many quantities measured, multiple techniques and methods are available and widely used. Thus, the number of measurement results is vast and growing.

Critical evaluation of data is a *post hoc* exercise, often relying solely on published reports that contain incomplete or poorly documented information. Therefore, evaluators need to determine whether there is sufficient information to accept some data, and to estimate measurement uncertainty that can be used, if necessary, in creating a consensus value and its uncertainty. With our present understanding of metrology (the science of measurement) [4], measurement results are required to have some estimate of their associated measurement uncertainties. Many data published have a limited or no assessment of uncertainty, which needs to be addressed in a critical evaluation. One of the tasks of critical data evaluation is to bring to the attention of data users (who might be unfamiliar with modern metrological concepts) the hazards and limitations of relying on data without measurement uncertainty estimates.

The evaluation of chemical data is an organized attempt to assist data users, both chemists and non-chemists, in understanding the strengths and weaknesses of reported measurements and selecting data best suited to their needs. For this paper, the ISCED defines evaluation of data as ‘assessment of documented and accessible chemical substance property data by a suitable authority, together with the assessment of methods by which these data have been obtained following agreed-upon concepts of how evaluations are performed.’ Thus, the evaluation of chemical data is the process of assessing the quality of a set of chemical measurement results for a specific quantity (property) based on the evaluation against pre-defined criteria to deliver a statement of that quality together with an expression of uncertainty. The quality of evaluated data is always limited by the quality of the measurement data on which the evaluation is based, and by the quality and completeness of the measurement report. This information includes the nature of the samples subjected to analysis to assess representativeness of the reported property data.

---

<sup>2</sup> In the preparation of this paper, the authors encountered terms for which definitions have not been recommended by IUPAC or another authoritative body. For such terms, interim definitions have been formulated and are presented here in single quotation marks; double quotation marks are used to indicate direct quotations. We will prepare a Recommendation for the definition of these interim terms in due course.

Data evaluation is important to chemical science and the users of chemical data for several reasons. Most obviously, a well-designed and skillfully executed critical data evaluation is of value to all data users and especially to users lacking the expertise necessary to evaluate data they wish to use. In favorable cases where metrologically modern measurements (*i.e.*, measurements with robust measurement uncertainty estimates) are evaluated, the evaluation provides a quantity value for a chemical property including an estimate of measurement uncertainty, thus giving the user the information necessary to evaluate the data's fitness for their intended use. In less favorable cases, where shortcomings of the original work make uncertainty estimation difficult or even impossible, users must be made aware of the situation.

Often evaluations are based on reports of measurements made by a variety of experimental techniques over a period of several decades and reported in units that are not always directly comparable. Harmonization and evaluation of such data sets can reveal the strengths and weaknesses of chemical measurement principles, methods, and procedures used and inform the measurement community about the kinds of experimental detail and contextual information (metadata) needed to fully evaluate reports of measurements. This can lead to improved choices of measurement techniques and methods and better data reporting practices.

### 3 Principal approaches to data evaluation

This section summarizes current thinking within ISCED about important considerations regarding the evaluation process. These points do not constitute a detailed protocol for critical evaluation but together suggest the range of considerations needed to produce high-quality evaluations of chemical data. As much as the quality of chemical data reflects the purpose for which, and the conditions under which, measurements were made, the approach used in an evaluation determines the quality of the resulting evaluation. Foremost, the quality of a critical data evaluation is determined by the quality of data evaluated. Other limiting factors to consider include the number of available measurements and the available human and financial resources to conduct the evaluation. Evaluators must therefore determine which quality level of the evaluation can be realistically achieved.

Categories of approaches to data evaluation are listed in order of increasing complexity and quality from A to D below:

- A. Selection and compilation of data from the literature based on a set of unified criteria for judgement of data quality defined by expert knowledge.
- B. Compilation and harmonization of data in the literature by standardizing reported measurement uncertainties, unit conversion or recalculation of reported quantity values by normalization to a common reference.
- C. Comparison of data compiled for a given property to decide on a consensus value and its uncertainty, either by selection of a single best measurement or by combining several measurements reported in the literature into preferred values and associated measurement uncertainty.
- D. Consideration of all sources of error including random and systematic error in reported measurement results to obtain a reference value with an expanded uncertainty range that includes the probable true property value with great certainty based on expert judgement.

Data compilations under Category A and B are not comparable in quality to Category C and D evaluations. Nevertheless, they are relatively easy to perform and provide an overview of existing data for a targeted user group, or they yield a database with measurements deemed fit for further evaluation based on expert judgement. Category C evaluations provide a single quantity value for a property but come with certain restrictions for use as stated uncertainty intervals are supposed to include the true value of the evaluated quantity or property but do not necessarily do so. Category D evaluations are the most desired as they come with the least restrictions for use. However, they are labor intensive which limits them to evaluation exercises where the required number of evaluations to be conducted is relatively small, and/or the importance of the results is relatively great, such as the Technical Reports of the IUPAC Commission on Isotopic Abundances and Atomic Weights (CIAAW) [5].

When publishing critically evaluated data, it is important that evaluators identify beforehand the potential user groups of their evaluated data and define requirements and measurement characteristics that measurement results must fulfill for their consideration. Once the evaluation is completed, evaluators must communicate clearly under which category the evaluation falls and any restrictions that apply to their use. This enables users to decide if the product of the evaluation is fit for their purposes. Certain applications, such as forensic investigations in court proceedings, may, depending on jurisdiction, require a property value to be adopted that should come with the smallest possible uncertainty interval that takes all identifiable sources into account. For other user groups, such as high school students in a laboratory class, uncertainty considerations might be of less concern. This requires that critically evaluated data be distributed together with the necessary information to permit users to make an informed decision about whether the evaluated data are fit for purpose, *i.e.*, fit for their intended use.

## 4 Sources of uncertainty in chemical measurement

Chemical data are produced by measurement. Measurements, however, can be conducted in different ways using the same or different measurement techniques or methods. Any measurement comes with an error, *i.e.*, a deviation of the measured quantity value from its true quantity value. The authoritative *Guide to the expression of uncertainty in measurement* (GUM) [6], makes clear that the true value of a measured quantity (property) will always remain unknown.

Sources of error in a measurement are multiple and can be random in nature as well as systematic (*i.e.*, biased), with both contributing to the overall measurement uncertainty. The extent to which sources of error in the measurement are understood and the reliability by which they are controlled limit measurement accuracy, which is defined as the “closeness of agreement between a measured quantity value and a true quantity value of a measurand” [4]. In an ideal situation, all sources of uncertainty in a measurement can be identified and their contribution to the overall measurement uncertainty quantified. In this situation, the combined measurement uncertainty, often reported as an expanded uncertainty ( $\pm$ ) interval, represents a range of quantity values in which the true value is supposed to lie with a stated probability.

The random component of measurement uncertainty can be estimated by repeated measurements and can be reliably deduced during data evaluation from the reported repeatability or reproducibility statement in a report or publication. This is different for systematic sources of error. Assessment of measurement bias ([4], entry 2.18) in reported data requires both considerable knowledge and expertise on the part of the evaluator and a detailed and accurate description of experimental procedures in the original measurement report. A primary requirement in every data evaluation exercise is therefore the competence of evaluators to identify possible sources of systematic error in reported measurements and to judge whether they have been appropriately accounted for in the reported uncertainty statement.

Ideally, for reasons of transparency, consistency and traceability, sources of systematic error to be assessed should be agreed upon before starting with the actual evaluation of measurements. During the evaluation process, however, additional error sources may be identified, requiring adjustment of the evaluation criteria. The following list is a non-comprehensive overview of sources of bias in chemical measurements that evaluators may consider in defining measurement characteristics to be assessed during data evaluation:

- Sampling (preparation, contamination, homogeneity, storage conditions, etc.).
- Representativeness of the object or material that has been analyzed (source, provider, sampling location).
- Availability of samples or reagents for independent reproduction of measurements.
- Control of measurement conditions (temperature, air humidity, etc.).
- Sample contamination during analysis.
- Quality control measures taken.
- Reference materials or calibrants used (certified, off-the-shelf, self-prepared).
- Assumptions made when conducting the measurement or in the evaluation of measurement results by the measurer.

- Algorithms and information used for data transformation.
- Method validation and validation criteria.
- Consistency of measurement results with accepted systematics and regularities.
- Appropriateness of applied statistical methods.
- Type of publication (peer reviewed publication, project report, written communication).

For making evaluations traceable over space and time, it is important that such procedural information (metadata, exemplified in the bulleted list above) is documented by the evaluators and made accessible together with the evaluation outcomes of a reported measurement. In reality, however, control of possible sources of bias is rarely comprehensive. Measurement reports do not contain all necessary information for judgement, or sources of bias are simply unknown. The term “dark uncertainty” was coined by Thompson and Ellison [7] to describe the extra uncertainty coming from differences in the quality by which measurement bias has been controlled between laboratories [8]. This indicates that even a careful bottom-up analysis following the GUM [6] cannot always encompass all sources of uncertainty. The evaluator may therefore decide to exercise expert judgement and expand the uncertainty of a reported or consensus value following GUM concepts for evaluation of measurement uncertainty. Some considerations for evaluating measurement uncertainty of compiled data are briefly introduced in the following section.

## 5 Evaluation of measurement uncertainty in chemical data

There are two aspects of measurement uncertainty that must be taken into consideration by the data evaluator.

First, the magnitude of the reported measurement uncertainty must be assessed in the evaluation process. This is as important as scrutiny of the reported value itself. With the known shortcomings in estimating measurement uncertainty, evaluators must decide beforehand about applicable rules and guidelines on how the reported uncertainty of a measurement shall be embedded into the evaluation process or how data that do not come with an uncertainty statement shall be dealt with. These rules and guidelines may include: (i) the outright rejection of data in the literature if reported without an uncertainty statement; (ii) the assignment of a conservative estimate of measurement uncertainty to such data based on expert judgement; (iii) the adoption of a measured property value and its measurement uncertainty as reported; or (iv) an adjustment of the reported uncertainty based on expert judgement.

None of these strategies is perfect. Each comes with certain disadvantages and caveats to be considered when deciding about a strategy for data evaluation. These aspects will be addressed in more detail in a forthcoming publication in this series of guidelines to assist evaluators in making strategic decisions for a data evaluation endeavor that is fit for the intended purpose.

Second, it must be decided how reported measurement uncertainties are used to determine a single consensus value and its uncertainty. In the conceptually most basic decisional approach, a single measurement is selected as the best measurement from various candidate measurements. By definition, the best measurement is thought to be of the highest quality in terms of control of random as well as systematic sources of error and the quality by which these sources of error in the measurement result have been quantified. Measurement reports, however, are not perfect and may not fully consider all sources of bias and their quantitative contribution to the uncertainty interval in which the true property value is supposed to lie with high certitude. Therefore, the reported uncertainty in a best measurement may be adjusted by expert knowledge based on careful analysis of the measurement report. This approach is in full agreement with GUM concepts, which explicitly permit the expansion of measurement uncertainties by means other than repeated measurement for assessing random sources of error (defined as Type A evaluation of measurement uncertainty [4], entry 2.28) to include all other sources of error (Type B evaluation of measurement uncertainty [4], entry 2.29). The quality of the resulting consensus value, however, depends largely on:

- (a) the level of detail at which the best measurement has been reported.
- (b) the experience of the evaluators in comprehensively identifying possible sources of bias in a measurement.

- (c) the evaluators' ability to estimate the magnitude of possible bias in a measurement that has not been accounted for in a measurement report; and finally.
- (d) there being a conservative attitude of the evaluators for avoiding an underestimation of the derived decisional uncertainty of the consensus value if it is based on a single measurement report only.

If conditions necessary to obtain a decisional consensus value and uncertainty cannot be fully met, it is possible to obtain a popular consensus value by calculating the mean value of available measurements weighted by a function of the reported measurement uncertainties (*e.g.*,  $1/u^2$ ). This approach, however, assumes rather optimistically that we are randomly drawing from a single Gaussian population of measurements that have a mean which is the true value of the measurand; this also requires the measurement object to be the same for all reported measurements. Even with adequate uncertainty budgets (see [4], entry 2.33, and [6]) for measurements spread over many years with different methods and procedures, this assumption almost certainly will underestimate the measurement uncertainty of the result. Under the usual assumption that we are drawing values at random from a normally distributed population, the consensus variance will become smaller as more results are added (the standard deviation of the mean tends to zero as the number of values being averaged increases) and can be made arbitrarily small by including more and more results. The laboratory random effects model addresses this shortcoming by including a between-laboratory variance that adds to individual laboratory variances [8], and which is possibly due to unforeseen systematic error.

Another possible approach is the hierarchical Bayesian procedure for which the consensus value is the mean of a probability distribution of the random variable  $m$  and the parameters of the laboratory effects model, given the experimental results. Although many evaluators would not consider themselves born-again Bayesians, methods of consensus building are offered in a website constructed by NIST at <https://consensus.nist.gov/app/nicob> [9, 10]. A method requiring few assumptions about the data, and which gives the greatest but perhaps most sensible uncertainty, is linear pooling [11]. An example of the use of this method (essentially, the overlaying of distributions) applied to the solubility of cadmium carbonate in aqueous systems is given by Hibbert, who showed the uncertainty from linear pooling was in accord with evaluators' suggestions despite the usual weighted method giving much smaller uncertainty [12].

An additional layer of complexity is added when taking into consideration that values for a chemical property of a given type or class of samples are not the same over space and time. Such variations can be negligibly small if a stoichiometrically well-defined substance of highest purity in monocrystalline form has been used in all measurements to be evaluated. The other extreme is the situation in which the chemical property value of the sample subjected to measurement depends highly on its origin. The isotopic composition of oxygen and its relative atomic mass may serve here as an example. In fact, every natural sample or reagent has a unique oxygen isotopic composition depending on its source due to natural isotope fractionation processes. Here, the concepts above for obtaining a consensus or reference value are not applicable and require either the tying of a chemical property value, such as the relative atomic mass of oxygen, to a specific, highly homogeneous object or substance, or the use of expert knowledge to suggest an interval which covers the range of chemical property values that the user may find in a certain object or chemical substance to high likelihood. This practice was adopted by the CIAAW in its 2009 Technical Report on the standard atomic weight of elements showing well measurable isotope abundance variations in nature [13]. With the recent revision of the IUPAC Standard Atomic Weights of the Elements [5], a total of 14 elements have been assigned an interval (minimum atomic weight, maximum atomic weight) rather than a value and uncertainty range as their Standard Atomic Weight.

The different possible approaches for identification of a consensus value and uncertainty for a given chemical property will be the subject of a forthcoming publication with relevant practical guidance. In any case, it is important that evaluators decide beforehand which approach they intend to use and agree on applicable rules and procedural guidelines to ensure that measurements are evaluated in a comparable manner and that data in the resulting dataset are internally consistent. As measurement reports are the sole source of information used, it is always possible to change evaluation strategies retrospectively and recalculate consensus values and uncertainties provided that evaluation concepts are clearly explained for transparency, decisions are laid down in

all necessary detail for traceability and strictly adhered to for consistency. It is therefore of importance that evaluators should document the approaches used to obtain any consensus results.

## 6 Disseminating data evaluation results

Evaluation of chemical data adds considerable value to a body of original measurement data compiled from available reports or the published chemical literature. Users benefit by having more understanding of and confidence in the data they are using. The measurement community benefits by knowledge of how their techniques can be improved. Chemistry in general benefits by having a firmer knowledge base. The widest possible dissemination of data evaluation results should be encouraged to achieve these benefits to the fullest.

A good report of evaluated chemical data has the following elements and characteristics:

- Reporting that is well-constructed, useful and understandable to user group(s).
- Full explanation of the process by which data were evaluated and of the criteria used in the evaluation.
- Well-defined statement of limitations to permit users to decide if the evaluation is fit for the intended purpose and to avoid overinterpretation of the evaluated data.
- Clear distinction between choices made by the evaluator(s) and the original authors.
- Reference to, and agreement with, accepted concepts in chemical metrology.
- Consistent use of standard terminology and nomenclature that is unambiguous and traceable in the chemical literature.
- An estimate of the uncertainty of the stated values together with explanation of its derivation, including the statistical procedure(s) used in the analysis, if appropriate referencing the software involved.
- A list of all original published data, and their sources, that have been considered, thereby allowing users to consult such data.
- For a stated quantity value that is not accompanied by an uncertainty estimate, the reasons should be given.

Decisions by the evaluator(s) should be documented and archived both as softcopies and as hardcopies to make them traceable over space and time, and to facilitate revision of previous decisions. This record-taking process should include: references for all original measurements considered in the evaluation (even if they have been excluded); all measurement methods used; and all measurement uncertainty estimates (termed experimental error in older reports), as reported by the measurer(s). See the following section.

Attribution of the individuals and organizations responsible for evaluated chemical data is important for a variety of reasons. Identifying the source or sponsorship of the evaluation effort gives a sense of authority and quality and enables the community to trace provenance as well as scientific recognition. For example, this is the case with projects of IUPAC, which is a leading authority in this space in chemistry. Many data evaluation projects in chemistry are team efforts, and assigning attribution to specific individuals can be difficult. At the same time, acknowledgment of individual evaluators is important for scientific credit, and care should be taken to be as inclusive as possible.

## 7 Digital expression and accessibility of critically evaluated data

The increasing prevalence of predictive models and other data-driven applications such as artificial intelligence and machine learning further emphasizes the need for quality data with clearly articulated uncertainty budgets. To enable distribution and use of critically evaluated data for a broad range of applications, ultimately data and associated descriptive information need to be machine-readable – that is, expressed in a form that can be processed by algorithms with an acceptable level of accuracy. Most downstream operations on these data would be expected to be managed by computers through software and online workflows, including publication and incorporation into databases. Note that:



- Proper machine-representation can present the data in almost any human-readable display desired.
- Data expression is not just about curation of the initial source but the need to transfer data with precise and consistent representation from system to system.

Precise transfer of critically evaluated data values and associated information is facilitated by well-defined digital data formats that fully articulate original measurements and meta-analyses involved in assessing uncertainty at the level of detail and nuance entailed in critical evaluation methods. Underlying such formats are metadata schemas that describe methods of determination, experimental conditions, and other relevant information (*e.g.*, reported measurement precision and sources of systematic error) to ensure consistent expression and systematic aggregation of measurement data parameters. Several IUPAC projects are underway to define standard metadata (experimental detail and contextual information) for different properties, including solubility, isotopic abundances, and magnetic resonance [14]. These IUPAC standards will also encapsulate criteria for automatically accessing and using these data via application programming interface calls and downstream applications.

Good (*i.e.*, well measured) data can be rendered useless or, even worse, misleading if poorly documented and communicated. Reported data often cannot be included in evaluation projects for lack of key parameters, such as the temperature at which a property was measured. Communication of chemical property measurements based on IUPAC guidelines for digital expression will facilitate FAIR (findable, accessible, interoperable, and reusable) data sharing [15] and will accommodate machine processing across data from multiple sources, including:

- Data/code curation (*i.e.*, data verification and validation).
- Use of standard reference data and data processing algorithms.
- Use of toolkits to verify correct processing with validation data.
- Assessment of fitness for purpose, data reuse and trust in the reusability of data.
- Future semantic applications (knowledge mining).

To ensure evaluated data and associated information are available for capture in digital machine-readable form, it is important to manage all data proactively throughout the evaluation process, from initial compilation of reported measurements to determination of preferred values, ideally through a shared database or repository for the project. At minimum, data should be tabulated; to the extent possible, data should be saved in files with open text-based formats (*e.g.*, comma separated values, CSV, or tab separated values, TSV) and, for IUPAC-sponsored projects, archived with the IUPAC Secretariat [16].

## 8 Current IUPAC data evaluation activities

IUPAC has a long history of performing data evaluations and several groups within IUPAC are now actively involved in evaluation of chemical data. These include the Commission on Isotopic abundances and Atomic Weights in the Inorganic Chemistry Division, the Task Group on Atmospheric Chemical Kinetic Data Evaluation in the Physical and Biophysical Chemistry Division, the Subcommittee on Solubility and Equilibrium Data in the Analytical Chemistry Division, which coordinates the Solubility Data Project and the Stability Constants Data Project, and the Subcommittee on Modeling of Polymerization Kinetics and Processes in the Polymer Division. In addition, the Committee on Publications and Cheminformatics Data Standards provides advice concerning digital expression and accessibility of evaluated data. The following sections present brief overviews of current data evaluation activities of members represented in ISCED.

### 8.1 Standard Atomic Weights and isotopic abundances

At its founding in 1919, IUPAC became the new home of the International Committee on Atomic Weights, which is now known as the IUPAC Commission of Isotopic Abundances and Atomic Weights (CIAAW). The Standard Atomic Weights of the elements are among the most fundamental data used by scientists and non-scientists alike to

connect the microscopic dimension of nature, *i.e.*, the number of atoms or molecules in a sample, with the macroscopic world, *i.e.*, the mass of a sample that is accessible by weighing. With the known mass of a sample of a chemical compound of known composition and purity, the amount of substance may be calculated using its molar mass, and the number of atoms and/or molecules in the sample can be calculated using their molar masses and the Avogadro constant. Such calculations are second nature for those engaged in physical sciences, life sciences and engineering; they are also essential in trade and commerce, *e.g.*, for the conversion of amount of substance (mole) into mass (kg) when converting concentration units.

Formally established in 1899, the CIAAW meets biannually to evaluate published data from isotope abundance measurements and atomic mass determinations, and therefore to regularly provide the scientific and non-scientific communities with numerical values for the Standard Atomic Weights [5, 13] and the natural isotopic composition of the elements [17]. Data are presented, together with uncertainty intervals in which their true values are supposed to lie with great certitude, for use by the international scientific, business and educational communities. Standard atomic weights preferred by CIAAW, together with related publications of past and present decisions as well as additional sources of information, can be found on the CIAAW website [17].

## 8.2 Atmospheric chemical kinetic data

The IUPAC Task Group on Atmospheric Chemical Kinetic Data Evaluation was founded in 1989 and provides evaluated atmospheric chemistry data used in models of climate change, stratospheric ozone depletion, urban and regional air pollution, and the formation and fate of persistent organic pollutants. Over 1400 reaction datasheets describing the preferred values are available to the global chemistry community on a searchable website [18] that provides access to the preferred values, grouped according to reaction phase and category. Preferred values have been published periodically in a series of articles (*e.g.*, [19]) in the journal *Atmospheric Chemistry and Physics*. Work is ongoing to extend the coverage of the database and to convert the reaction datasheets into machine-readable files. These files will facilitate more effective communication with the international chemistry community by allowing automated transfer of IUPAC preferred values into atmospheric models.

## 8.3 Solubility data

The Solubility Data Project (SDP) was established in the mid-1970s when a group of chemists and chemical engineers came together as the Solubility Data Commission within the Analytical Chemistry Division (ACD) of IUPAC. Since the reorganization of IUPAC in 2002 under the project system, the effort has continued under the direction of the ACD Subcommittee on Solubility and Equilibrium Data (SSED). The SDP works to exhaustively compile and critically evaluate reports of experimental measurements of solubility in the primary chemical literature. The results are organized as a series of volumes, *The Solubility Data Series*, each of which seeks to compile all published reports of solubility measurements for a group of chemically related systems. Where sufficient data of appropriate quality are available, they are critically evaluated in a transparent way. Details about the more than one hundred volumes published can be found on the IUPAC website [20]. In the early years of the SDP the current formulation of measurement uncertainty was not available and the evaluation was approached in other ways [20, 21]. Recent volumes incorporate estimates of measurement uncertainty of evaluated data where the quality of compiled data and accompanying metadata allow. Work is underway to convert all compilations and evaluations of the Solubility Data Series into machine-readable form.

## 8.4 Stability constant data

It has long been recognized that many researchers needing stability constant data for various purposes such as the ionization of organic and inorganic acids and bases in aqueous solution, the formation of coordination compounds, or for constructing data bases for chemical speciation modelling for scientific or engineering

calculations, have neither the time nor the expertise to distinguish among the numerous and often-conflicting reported values. To meet this need, stability constant data for homogeneous equilibria have been compiled and evaluated by IUPAC for almost 70 years. The Commission V.6 on Equilibrium Data of the Analytical Chemistry Division decided to undertake a critical evaluation of the existing data. Given the magnitude of such a task, this project was divided into smaller parts and volunteer experts were sought to do the work. Following the reorganization of IUPAC in 2002, the effort has continued under the direction of the ACD Subcommittee on Solubility and Equilibrium Data (SSED).

A list of the 37 publications resulting from this work can be found on the IUPAC website [22]. The early work in this area produced comprehensive compilations of *Stability Constants of Metal-Ion Complexes* published as monographs with several supplements (items 1 to 9 of [22]). These volumes were only intended to be exhaustive compilations of the chemical literature. In 2003 most of the collected data were incorporated (some after updating) into an electronic database (item 10 of [22]). The remaining 27 publications provided critical evaluations of various selected groups of stability constants, four of which were published as monographs in the *IUPAC Chemical Data Series* (items 11 to 14 of [22]) and 23 as Technical Reports in *Pure and Applied Chemistry*.

## 8.5 Polymerization kinetics data

The IUPAC Subcommittee on Modeling of Polymerization Kinetics and Processes [23] was established in 1987 after an earlier IUPAC Working Party had drawn attention to “the unsatisfactory state of published kinetic parameters for radical polymerization in particular ... where published values of allegedly the same kinetic parameters may vary by orders of magnitude” [24]. The most important of these kinetic parameters is the rate coefficient for (radical) propagation,  $k_p$ , as it describes the rate of incorporation of monomer into polymer, a process that is carried out on a scale of ca. 200 million tons each year worldwide. The Subcommittee has addressed this situation via a series of publications over the last three decades, commencing with styrene  $k_p$  in 1995 [25]. Recently this body of work, now amounting to  $k_p$  values over a temperature range of ca. 100 K for 13 vinyl monomers in all, was updated and summarized via a comprehensive reanalysis using a statistical model that better accounts for systematic interlaboratory variation [26]. Rather than the “deplorable state” [23] found in 1987, the data now show a previously unimaginable level of accuracy and precision (variance of about 10 % rather than orders of magnitude) as well as systematic variation from monomer to monomer [26] that is chemically sensible, both within and between families [27]. The principal reason for this utterly transformed situation is the emergence of the so-called PLP-SEC method for  $k_p$  determination [28], which is pulsed-laser polymerization (PLP) combined with size exclusion chromatography (SEC) analysis of the chain-length distribution of the resulting polymer. The Subcommittee has played a central role in popularizing this method and critically evaluating the data thus obtained, which is primarily achieved via built-in consistency checks [25, 29]. All older data has been excluded from IUPAC benchmark data sets, which consist entirely of  $k_p$  values determined by PLP-SEC [26]. Work is ongoing to extend the coverage of the database and to convert the data into machine-readable form [30].

## 9 Conclusions

Chemical data evaluation has been at the heart of activities within IUPAC since its founding in 1919. The IUPAC Interdivisional Subcommittee on Critical Evaluation of Data was instituted in 2018 to advance best practices. As a guide for future IUPAC projects and for the broader chemical community we have presented here the general considerations behind, approaches to and examples of chemical data evaluation activities within IUPAC until the end of 2022. Future papers in this series will provide a guide to measurement uncertainty, including in consensus values; a glossary of terms related to data evaluation; detailed approaches for evaluation of chemical data; a guide for preparation and dissemination of digital data; and an outline of data evaluation strategies employed in some current IUPAC data evaluation projects.

## 10 Glossary of abbreviations and acronyms used in this guide

---

ACD	Analytical Chemistry Division (of IUPAC) ( <a href="https://iupac.org/body/500/">https://iupac.org/body/500/</a> )
CIAAW	Commission on Isotopic Abundances and Atomic Weights ( <a href="https://www.ciaaw.org/">https://www.ciaaw.org/</a> )
GUM	Guide to the Expression of Uncertainty in Measurement [6]
ISCED	Interdivisional Subcommittee on Critical Evaluation of Data ( <a href="https://iupac.org/body/505/">https://iupac.org/body/505/</a> )
IUPAC	International Union of Pure and Applied Chemistry ( <a href="https://iupac.org">https://iupac.org</a> )
NIST	National Institute of Standards and Technology ( <a href="https://www.nist.gov/">https://www.nist.gov/</a> )
PLP-SEC	pulsed laser polymerization-size exclusion chromatography [25]
SDP	Solubility Data Project (of SSED)
SSED	Subcommittee on Solubility and Equilibrium Data ( <a href="https://iupac.org/body/502/">https://iupac.org/body/502/</a> )
VIM	International Vocabulary of Metrology – Basic and General Concepts and Associated Terms [4]

---

## 11 Membership of sponsoring bodies

This Technical Report was prepared by the Interdivisional Subcommittee on Critical Evaluation of Data (<https://iupac.org/body/505>) with participation by members of the Analytical Chemistry Division, the Inorganic Chemistry Division, the Physical and Biophysical Chemistry Division, the Polymer Division and the Committee on Publications and Cheminformatic Data Standards. During the period 2022–2023 the composition of these bodies was as follows:

**Interdivisional Subcommittee on Critical Evaluation of Data:** **Chair:** D. Shaw (USA); **Members:** I. Bruno (UK), S. Chalk (USA), A. Davies (UK), D. B. Hibbert (Australia), R. A. Hutchinson (Canada), M.C. Magalhães (Portugal), J. Magee (USA), L. McEwen (USA), I. Perminova (Russia), J. Rumble Jr. (USA), G. T. Russell (New Zealand), E. Waghorne (Ireland), T. Walczyk (Singapore), T. Wallington (USA).

**Analytical Chemistry Division:** **President:** D. Shaw (USA), **Vice President:** D. Craston (UK), **Secretary:** L. Torsi (Italy), **Past President:** Z. Mester (Canada), **Titular Members:** R. Apak (Turkey), V. Baranovskaia (Russia), J. Barek (Czech Republic), I. Kuselman (Israel), T. Takeuchi (Japan), S. Wiedmer (Finland); **Associate Members:** F. Emmerling (Germany), E. Flores (Brazil), I. Leito (Estonia), H. Li (China/Beijing), A. Tintaru (France), E. Waghorne (Ireland); **National Representatives:** R. Burks (USA), H. R. Byon (South Korea), O. Chailapakul (Thailand), J. Labuda (Slovakia), C. Lucy (Canada), M. C. Magalhães (Portugal), T. Pradeep (India), M. Ramalingam (Malaysia), R. Sha’Ato (Nigeria), D. van Oevelen (Netherlands).

**Inorganic Chemistry Division:** **President:** L. Armelao (Italy), **Secretary:** D. Rabinovich (USA), **Past President:** L. Öhrström (Sweden); **Titular Members:** E. Bouwman (Netherlands), J. Colón (Puerto Rico), M. C. Gimeno (Spain), P. Knauth (France), M. H. Lim (South Korea), J. Meija (Canada), T. Walczyk (Singapore); **Associate Members:** F. Abdul Aziz (Malaysia), M. Diop (Senegal), R. Macaluso (USA), K. Sakai (Japan), A. Sanson (Italy), X. K. Zhu (China/Beijing); **National Representatives:** H. Cohen (Israel), P. Harding (Thailand), M. Hasegawa (Japan), R. Hocking (Australia), P. Karen (Norway), L. Krivosudský (Slovakia), A. Logsdail (UK), O. Metin (Turkey), N. Ngobiri (Nigeria).

**Physical and Biophysical Chemistry Division:** **President:** P. Metrangolo (Italy), **Vice President:** F. Separovic (Australia), **Past President:** T. Wallington (USA), **Secretary:** A. Császár (Hungary), **Titular Members:** M. Fall (Senegal), J. Martins de Faria (Portugal), Z. Shuai (China/Beijing), I. Voets (Netherlands), A. Wilson (USA), M. Witko (Poland); **Associate Members:** K. Chong (Malaysia), T. Frankcombe (Australia), L. Montero-Cabrera (Cuba), I. Schapiro (Israel), H. Tokoro (Japan), V. Tsakova (Bulgaria); **National Representatives:** J. Frey (UK), T. C. Kurtén (Finland), L. C. Ngozi-Olehi (Nigeria), R. Orinakova (Slovakia), V. Parasuk (Thailand), M. Štěpánek (Czech Republic).

**Polymer Division:** **President:** C. Luscombe (Japan), **Vice President:** I. Lacik (Slovakia), **Secretary:** P. Topham (UK), **Past President:** G. T. Russell (New Zealand), **Titular Members:** C-H Chan (Malaysia), T. Junkers (Australia), P. Mallon (South Africa), J. B. Matson (USA), Y. Men (China/Beijing), M. Peeters (UK), P. Théato (Germany);

**Associate Members:** C. Fellows (Australia), D. N. Haase (USA), R. A. Hutchinson (Canada), J. Merna (Czech Republic), A. I. Ricardo (Portugal), M. H. Yoon (South Korea); **National Representatives:** R. Adhikari (Nepal), J-T Chen (China/Taipei), S. Guillaume (France), J. E. Imanah (Nigeria), A. Kishimura (Japan), G. Mechrez (Israel), S. Ramakrishnan (India), G. Raos (Italy), M. Tasdelen (Turkey), J. van Hest (Netherlands).

**Committee on Publications and Cheminformatic Data Standards: Chair:** L. McEwen (USA), **Secretary:** J. Liu (USA), **Titular Members:** G. M. Banik (USA), I. Bruno (UK), S. Chalk (USA), C. Nitsche (USA), K. Oisaki (Japan), E. Rios-Orlandi (Puerto Rico), C. Steinbeck (Germany), D. Vanderwall (USA); **Associate Members:** J. Frey (UK), S. Hannongbua (Thailand), G. M. Jones (UK); **Ex Officio:** H. D. Burrows (Portugal), R. M. Hartshorn (New Zealand), C. Humphris (UK), F. Meyers (USA), L. Soby (USA).

**Acknowledgment:** The authors gratefully acknowledge the support and funding for this work from IUPAC through project 2018-009-2-500 (<https://iupac.org/project/2018-009-2-500>) and comments from the reviewers which clarified and improved this Technical Report.

## References

- [1] <https://www.cas.org/cas-data/cas-registry> (accessed Dec 31, 2022).
- [2] D. B. Hibbert, E.-H. Korte, U. Örnemark. *Pure Appl. Chem.* **93**, 997 (2021), <https://doi.org/10.1515/pac-2019-0819>.
- [3] A. Bazyleva, J. Abildskov, A. Anderko, O. Baudouin, Y. Chernyak, J.-C. de Hemptinne, V. Diky, R. Dohrn, J. R. Elliott, J. Jacquemin, J.-N. Jaubert, K. G. Joback, U. R. Kattner, G. M. Kontogeorgis, H. Loria, P. M. Mathias, J. P. O'Connell, W. Schröer, G. J. Smith, A. Soto, S. Wang, R. D. Weir. *Pure Appl. Chem.* **93**, 253 (2021), <https://doi.org/10.1515/pac-2020-0403>.
- [4] Joint Committee for Guides in Metrology. *JCGM 200:2012, International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM)*, BIPM, Sèvres, France (2012), Most recent version available from, <https://www.bipm.org/en/publications/guides/> (accessed Jan 23, 2023).
- [5] T. Prohaska, J. Irrgeher, J. Benefield, J. K. Böhlke, L. A. Chesson, T. B. Coplen, T. Ding, P. J. H. Dunn, M. Gröning, N. E. Holden, H. A. J. Meijer, H. Moossen, A. Possolo, Y. Takahashi, J. Vogl, T. Walczyk, J. Wang, M. E. Wieser, S. Yoneda, X.-K. Zhu, J. Meija. *Pure Appl. Chem.* **94**, 547 (2022), <https://doi.org/10.1515/pac-2019-0603>.
- [6] Joint Committee for Guides in Metrology. *JCGM 100:2008, Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement (GUM)*, BIPM, Sèvres, France (2008), Most recent version available from, <https://www.bipm.org/en/publications/guides/> (accessed Jan 23, 2023).
- [7] M. Thompson, S. Ellison. *Accred. Qual. Assur.* **16**, 483 (2011), <https://doi.org/10.1007/s00769-011-0803-0>.
- [8] B. Toman, A. Possolo. *Accred. Qual. Assur.* **14**, 553 (2009), <https://doi.org/10.1007/s00769-009-0547-2>.
- [9] National Institute of Standards and Technology. *NIST Consensus Builder*, National Institute of Standards and Technology, Gaithersburg, Maryland (2017), <https://consensus.nist.gov/app/nicob> (accessed Jan 23, 2023).
- [10] A. Koepke, T. Lafarge, A. Possolo, B. Toman. *NIST Consensus Builder—User's Manual*, Technical Report (2016), [https://consensus.nist.gov/app\\_direct\\_i/nicob/\\_/NISTConsensusBuilder-UserManual.pdf](https://consensus.nist.gov/app_direct_i/nicob/_/NISTConsensusBuilder-UserManual.pdf) (accessed Jan 23, 2023).
- [11] F. Dietrich, C. List. In *The Oxford Handbook of Probability and Philosophy*, A. Hájek, C. Hitchcock (Eds.), pp. 519–542. Oxford University Press, Oxford (2017).
- [12] D. B. Hibbert. *J. Chem. Thermodyn.* **133**, 152 (2019), <https://doi.org/10.1016/j.jct.2019.02.007>.
- [13] A. M. H. van der Veen, J. Meija, A. Possolo, D. B. Hibbert. *Pure Appl. Chem.* **93**, 629 (2021), <https://doi.org/10.1515/pac-2017-1002>.
- [14] IUPAC Digital Standards, <https://iupac.org/what-we-do/digital-standards/> (accessed Jan 23, 2022).
- [15] I. Bruno, S. Coles, W. Koch, L. McEwen, F. Meyers, S. Stall. *Chem. Int.* **43**, 12 (2021), <https://doi.org/10.1515/ci-2021-0304>.
- [16] <https://iupac.org/projects/information-for-task-group-chairs/#outcomes-data-digital-outputs> (accessed Jan 23, 2023).
- [17] IUPAC Commission on Isotopic Abundances and Atomic Weights, <https://ciaaw.org> (accessed Jan 23, 2023).
- [18] IUPAC Task Group on Atmospheric Chemical Kinetic Data Evaluation, <https://iupac.aeris-data.fr/en/home/> (accessed Jan 23, 2023).
- [19] A. Mellouki, M. Ammann, R. A. Cox, J. N. Crowley, H. Herrmann, M. E. Jenkin, V. F. McNeill, J. Troe, T. J. Wallington. *Atmos. Chem. Phys.* **21**, 4797 (2021), <https://doi.org/10.5194/acp-21-4797-2021>.
- [20] IUPAC Solubility Data Series, <https://iupac.org/what-we-do/databases/solubility-data-series/> (accessed Jan 23, 2023).
- [21] H. Gamsjäger, J. W. Lorimer, M. Salomon, D. G. Shaw, R. P. T. Tomkins. *Pure Appl. Chem.* **82**, 1137 (2010), <https://doi.org/10.1351/PAC-REP-09-10-33>.
- [22] IUPAC Publications of Compilations and Critical Evaluations of Stability Constants, <https://iupac.org/wp-content/uploads/2022/02/502-Stability-constants-Refs-1.pdf> (accessed Jan 23, 2023).
- [23] IUPAC Subcommittee on Modeling of Polymerization Kinetics and Processes, [https://iupac.org/who-we-are/divisions/division-details/?body\\_code=428](https://iupac.org/who-we-are/divisions/division-details/?body_code=428) (accessed Jan 23, 2023).

- [24] G. C. Eastmond. *Makromol. Chem., Macromol. Symp.* **10/11**, 71 (1987), <https://doi.org/10.1002/masy.19870100105>.
- [25] M. Buback, R. G. Gilbert, R.A. Hutchinson, B. Klumperman, F.-D. Kuchta, B. G. Manders, K. F. O'Driscoll, G. T. Russell, J. Schweer. . *Macromol. Chem. Phys.* **196**, 3267 (1995), <https://doi.org/10.1002/macp.1995.021961016>.
- [26] S. Beuermann, S. Harrisson, R. A. Hutchinson, T. Junkers, G. T. Russell. *Polym. Chem.* **13**, 1891 (2022), <https://doi.org/10.1039/D2PY00147K>.
- [27] R. A. Hutchinson, S. Beuermann. *Pure Appl. Chem.* **91**, 1883 (2019), <https://doi.org/10.1515/pac-2018-1108>.
- [28] O. F. Olaj, I. Bitai, F. Hinkelmann. *Makromol. Chem.* **188**, 1689 (1987), <https://doi.org/10.1002/macp.1987.021880716>.
- [29] R. A. Hutchinson, M. T. Aronson, J. R. Richards. *Macromolecules* **26**, 6410 (1993), <https://doi.org/10.1021/ma00076a017>.
- [30] J. Van Herck, S. Harrisson, R. A. Hutchinson, G. T. Russell, T. Junkers. *Polym. Chem.* **12**, 3688 (2021), <https://doi.org/10.1039/d1py00544h>.