An Assessment of the Reliability of Atomic Weight Data Based on Successive Evaluations Between 1967 and 1993

P. De Bièvre, H.H. Ku[a], and H.S. Peiser[a]

Institute for Reference Materials and Measurements, Commission of the European Communities, Joint Research Centre, B-2440 Geel, Belgium

Received January 12, 1994; revised manuscript received March 9, 1994

The reliability of the (standard) atomic-weight (mean atomic mass) values is quantified based on the analysis of changes made from published improved measurements as they become available for each periodic evaluation. This numerical analysis tests the evaluation procedures and the current reliability of the tables published by the International Union of Pure and Applied Chemistry. One instance of faulty evaluation is high-lighted. The *post facto* test of past performance of evaluations might under appropriate conditions also serve for other data sets such as the fundamental constants of physics or property data for pure elements.

Key words: atomic weights; box plot; chemical measurements; data evaluation; IUPAC; reliability of data sets; uncertainty of values.

Contents

List of Tables

1. Introduction

Most scientists would agree that during the past few decades the careful evaluation of experimentally determined data sets has contributed greatly to progress in science and technology. Yet, successive evaluations, even when based on rather uniform methods of analysis, have often revealed previously unsuspected or inadequately considered errors that have led to changes in values in excess of the previously estimated uncertainties.

The recently published Guide to the Expression of Uncertainty in Measurement[1] is a very welcome innovation as it attempts to introduce hitherto lacking uniformity of treatments. It discourages vague safety factors, and encourages experimentalists to attach "standard" uncertainties (single standard deviation) to their measurements, although specified expanded uncertainties may be chosen. No doubt the evaluation of data sets in future will depend on the expectation of wide use of this Guide.

This paper attempts to answer whether conclusions can be drawn on the reliability of some data sets from successive evaluations. The authors will here examine the example of the Standard[b] Atomic Weights Tables. Similar evaluations could perhaps be attempted on the set of physical constants, the property data of pure elements, and other data sets, if and only if, they are subjected to repeated evaluations under rather uniform procedures.

---

[a] As Guest Scientists at the Institute in 1991. Correspondence to: 638 Blossom Drive, Rockville MD 20850, U.S.A.

Reprints available from ACS; see Reprints List at back of issue.

[b] The designation of "standard" for the atomic-weight values of the IUPAC Tables was introduced in 1979. It refers to the best published values with uncertainty in tabulation to include the range of atomic weights in all their unaltered terrestrial occurrences. The same concept, though previously not so carefully defined, has applied to the IUPAC Tables ever since atomic weights have no longer been regarded as constants of nature.

2. The IUPAC Table of Standard Atomic Weights

The Commission for Atomic Weights and Isotopic Abundances (CAWIA) of the International Union of Pure and Applied Chemistry (IUPAC) evaluates the published relevant scientific literature by methods described in "Element by Element Review of their Atomic Weights"[2]. Based on these procedures CAWIA publishes every two years a Table of Standard Atomic Weights, A_r(E)[c] for element E and their (absolute) uncertainties U(E), (see for example[3]). Entered in these Tables are the atomic weights (relative mean atomic masses) of the chemical elements in their terrestrial isotopic composition, unaltered by artificial means. The indicated range of values for each element is determined by:

a) the variability of isotopic compositions of the elements in their natural terrestrial occurrences;

b) the uncertainties in the assessment of that variability; and

c) the combined experimental uncertainties in the best published atomic-weight determinations. (When more than one determination is used, it is good practice to compare the samples directly and precisely so that a true difference between them is not added to the experimental uncertainty.)

If highly anomalous geological occurrences exist in which the element exhibits an atomic weight outside the indicated range, the Commission in specific instances may decide to indicate such an anomaly by a footnote, rather than by assigning a larger range for the standard atomic weight.

For most elements the IUPAC Table currently does not indicate whether the uncertainty is principally determined by experimental accuracy or by variability in terrestrial sources. Most elements still have uncertainties of their standard atomic weights limited by relevant experimental determinations. However, with improving measurement accuracies and also with added knowledge from mineral surveys, the number of elements is growing which have their standard atomic weights limited in tabulated precision by natural variabilities.

3. Changes in Standard Atomic Weights

The optimum conditions for drawing conclusions from successive evaluations of data sets arise if the data reviews are made regularly under invariant procedures, and if the measurement methods and potential error sources are numerous and independent. For the IUPAC Table these conditions have applied quite well, except that some of the evaluation procedures have actually changed slightly and therefore need to be discussed below. The one general principle that has never been altered is the aim for the IUPAC Table of (Standard) Atomic Weights with their indicated ranges to pass on to the user the best, reliable (as judged by CAWIA), fullest possible, published knowledge of the A_r(E) values in terrestrial materials (unaltered in isotopic composition). The all-important condition that has to be met is that the true A_r(E) value of any specific sample (unaltered artificially in isotopic composition) is expected with very high reliability to lie in the uncertainty range indicated in the Table. CAWIA has neither claimed nor permitted a statement to be made on its behalf regarding any statistical interpretation of "high reliability".

In the biennial Commission reviews all elements and all new relevant literature are included. All earlier literature is open to re-analysis. The very high specialization of the subject and the infrequency of attempts to measure any one atomic weight at or near the highest attainable accuracy, make this all-encompassing review possible with expectation of near completeness. However, one seldom has available even two new fully independent measurements to confirm A_r(E) and U(E) values for individual elements. It is then not unusual for the Commission in the evaluation process to find itself giving strong preference to a single highly reliable measurement over several inferior determinations. Thus, a statistical evaluation of the reliability of individual atomic-weight values is not possible. We therefore pose the question whether it is possible retrospectively to arrive at an estimate of reliability of the entire data set based on the changes over the years in the IUPAC Atomic Weights Tables.

It is certainly true and widely understood that the IUPAC atomic-weight tables since 1969 have been free from violent changes. and free from the clustering around recent, well believed, but erroneous values. This phenomenon has plagued some other data sets. All the atomic-weight changes recommended by IUPAC in the period under review were smaller than each sum of the uncertainties immediately before and after the change. For a meaningful, more discriminating assessment of this undoubtedly high reliability one must evidently submit the data to more stringent tests.

4. The Chart of A_r(E) Changes

In Table 1 each of a total of 64 IUPAC changes, made in the period of 1969 to 1993, is given by the element's symbol and the year of change in rows ordered by the ratio of the magnitude of the difference, D(E), in A_r(E) divided by the uncertainty U(E) indicated for the A_r(E) immediately before the change.

Excluded from Table 1 are changes made for monoisotopic elements, whose A_r(E) values are equal to their atomic masses (of individual isotopes) which are known to high accuracy. Besides, the reliability of atomic masses is not of direct concern to the IUPAC Commission, but these values are published with encouragement by the International Union of Pure and Applied Physics.

Included in Table 1 are not only changes in A_r(E), but also 15 instances of changes in U(E) alone, because such U(E) changes are recommended by the Commission only when they are based on new published determinations or re-assessments of previous work.

Altogether recorded in Table 1 are all changes, as defined above. Of these 27 changes increased, 22 decreased, 14 involved no change in A_r(E) value, and for 1 the new A_r(E) value had an additional zero digit in the last given decimal, so that D(E) and D(E)/U(E) were zero, too.

The chart would be even more meaningful if one could claim that every change was based on a new and independent

---

[c] For a specific element, E is replaced by its chemical symbol.

method of measurement. This is probably true, for instance, for the changes in germanium and thallium, but quite untrue, for changes in helium and magnesium. Nevertheless, one can search in the chart for valid evidence of reliability.

5. The Uncertainties of the Standard Atomic Weights

In 1967 one of the authors of this note (HSP) began to assess the uncertainties of A_r(E) values of all elements. At that time, most elements did not have an assigned U(E) value. The surprising reason given widely was that uncertainties, if printed, tend to reduce the confidence of users in the values themselves. A senior Commission member explained further: "when we give an atomic weight without explicitly stating an uncertainty, it implies that we confidently believe that the value is better than it would be with the last digit reduced or increased by one". A superior alternative for the Commission might have been to have adopted Eisenhart's interpretation of a number given without an uncertainty when systematic error and imprecision are negligible[4]: "...this number [is] accurate within $\pm \frac{1}{2}$ unit in the last significant figure given, that is, it will be understood that its inaccuracy before rounding was less than ± 5 units in the next place." CAWIA chose not to employ Eisenhart's convention because the above preconditions do not really apply to atomic-weight data.

At its 1969 meeting in Cortina d'Ampezzo, the Commission conceded that its own rule implied a variable accuracy dependent on the next suppressed (rounded off) digit. This consideration in addition to other inconsistencies in the then current Tables of Atomic Weights led to a revision of the entire Table giving each element the 1969 values, A_r(E) with their U(E). These are the base values from which our analysis of subsequent changes begins[5].

The Commission's conversion to a uniform system of uncertainty assessments has been gradual. The following rules have so far been applied:

a) Uncertainties must be stated by a single digit. Additionally from 1969 to 1985 only two U(E), values, ± 1 or ± 3 in the last place, were used so that the precision changed by a roughly constant factor of 3 between all successive options for tabulation. It should be noted, however, that in recent years the uncertainty could be expressed as any single-digit number. Thus the uncertainty steps can vary between succeeding options in A_r(E) by up to one order of magnitude depending only on the value of the last digit in A_r(E). If at some future time CAWIA were to view this situation as undesirable, a change to a two digit uncertainty would correct this problem.

b) Symmetric ranges, $\pm U$(E), must always be used. Consequently the implied range of the tabulated A_r(E) values always equals $2U$(E). This rule is still used, although for several elements, hydrogen and sulfur are examples, the most probable value is known not to be the mean in the range. It is quite likely that eventually, greater knowledge of natural variability in some A_r(E) values will commend the use of unequal uncertainties in the positive and negative directions.

Between 1969 and 1985 about half the elements had U(E) $= \pm 1$ so that any change D(E) in A_r(E) (without change in the number of digits) had to have D(E)/U(E) equal or greater than 1! This is the reason why changes for Ni, Xe, Zn, and Cd with D(E)/U(E) = 1 had to be recorded. They are just the result of the system of rules that unduly restrict the precisions of recorded uncertainties. At least four instances of D(E)/U(E) = 1 were to be expected from the rules and in no way suggest less than perfect reliability of the earlier IUPAC A_r(E) values paired with their U(E) values. In the chart (table I) the change for potassium at D(K)/U(K) = 1.33 alone stands out as a likely error of judgment by the Commission. Closer examination revealed a cause described in the next section.

6. Mass Spectrometry with its Potential Accuracy Supplants the Best Chemical Methods for the Determination of Atomic Weights

Up to about 1960 chemical determinations of A_r(E) values of elements with many stable isotopes were comparable in evaluated uncertainties with the best mass-spectrometric measurements. At about that time improvements in instruments and methodology turned very much in favor of mass-spectrometric determinations[6]. It was natural under those circumstances for chemists to redouble their attempts to show that in some instances chemical determinations could still be judged superior to mass-spectrometric measurements. Perhaps wishful thinking caused chemists to overstate their case. This may have been the circumstance by which the unfortunate change in A_r(K) was made on the recommendation by one of the present authors (HSP). The basis for the change was Bates and Wichers: Precise Intercomparison of Acids by Differential Potentiometric Titration with Hydrogen Electrodes[7]. It was beautiful work, but, as Murphy has pointed out much later[7], the Commission should have increased the estimated uncertainty on this work because:

a) No other atomic-weight determination had been made by a similar method and proven to be accurate;

b) The experimental measurements were carried out for purposes other than the determination of A_r(K) which it yielded as an afterthought.

It may also be worth mentioning that in future virtually all standard atomic weights will depend on mass-spectrometry, but that by no means should be interpreted as meaning the elimination of chemical procedures and measurements. Without excellent chemical work no absolute mass-spectrometric measurement or mineral survey will be credible.

7. The Reliability of A_r(E) Values Established

A glance at the chart (Table 1) confirms that conservative judgment was used in making changes in the Standard Atomic Weight values. This conclusion is corroborated by use of the box plot (see, for instance[9]). We have so analysed the data, by first taking into account the distinction between positive and negative D(E) values. We then find that the mode, mean, and median values are very close to zero. The box containing 50 % of the data is then constructed (indicated by shading the last values, F_u and F_l, of 0.30 and−0.25. Ideally, then the upper and lower cut-offs should be at D(E) values 1.12 and−1.08. The one potassium value again shows up as an obvious outlier.

The conclusion of a very high reliability can be put to one more test which, however, is necessarily based only on the very few elements with more than one atomic weight change during the 1969 − 1993 period. Successive D(E) values of opposite sign (7 instances) have occurred less frequently than D(E) values of same sign (9 instances). For highly reliable 1969 estimates, an excess of opposite signs would have been more probable. The reason for that asymmetry is the following: The probability of finding the true or a better value is generally symmetric about any determination, such as an original 1969 A_r(E) value. The same would apply also to the first redetermination (though it might have a narrower distribution) provided it is taken on its own. Any remaining validity of the original determination must distort the probability of finding the true value and the next redeterminations of A_r(E) on the same side of the original A_r(E) value.

8. Towards a Reliability Factor

One might wish to give some quantitative estimate of that reliability. Perhaps one could define d, a computed standard change in A_r(E) by:

$$d^2 = [1/(n - 1)] \Sigma [D_i/U_i(E)]^2$$

There d, would be a single value derived from the entire IUPAC/CAWIA data set of recommended D(E) values. Thus d could serve as an index of reliability. Perhaps CAWIA could make itself an aim to watch the d value over time and to increase caution if ever d rises above an agreed danger value, say, 0.50, as compared with the value from Table 1 of $d = 0.45$, despite the potassium outlier. The inherent difficulty is of course that any policy change in the degree of caution or discretion will only very gradually affect the d value which is a reliability measure principally of earlier A_r(E) values.

9. Conclusion

The analysis presented shows that if the uncertainties, U(E), in the atomic-weight data had corresponded to a single standard deviation from a mean between the highest and lowest true value for that element in its natural isotopic compositions, and if succeeding measurements had been truly independent, the spread of the D(E)/U(E)-value data would have been substantially greater than is found in the chart (Table 1). The fact that the D(E)/U(E) spread in Table 1 is small is the more remarkable since changes in estimated A_r(E) ranges of terrestrial sources substantially add to the changes of D(E) that are recorded. This result implies much more than that the designated uncertainties were made large so as to protect against later improved but discordant measurements. Above all, we claim to have demonstrated that CAWIA by careful evaluation of published results shielded the user of the standard atomic-weight data set from the many measurements which appear in the literature with new, but not fully reliable, values, or with erroneous or unduly tight uncertainties.

Although succeeding measurements were not always independent, the data suggest reliability far greater than would correspond to single standard deviations. The aim of IUPAC has feen fulfilled by which any chemist − taking any natural sample from research, industry, or commerce can confidently expect his or her true sample atomic weight to lie within the tabulated range with a probability far in excess of 95 %.

In the opinion of the authors, IUPAC should not change either the aim for the standard atomic weights or the evaluation principles . Users of other data sets must judge whether they require such very high reliability, or prefer truncated values in data sets that are subject to more probable and frequent future adjustments. The application of the numerical test, that is here proposed and that at best can give limited, long delayed, retrospective insights, is conditioned on periodic re-evaluations under closely similar policies and also implies a large diversity of data, measurement methods, and uncommon error sources.

10. Acknowledgments

All members of the Commission of the International Union of Pure and Applied Chemistry with responsibility for recommending atomic-weight values have during the past 25 years spared no effort in assuring the reliabilty of these values and their relevant uncertainties. We owe each and all Commission members profound thanks for their effort and expertise. Our collaboration with them over those years has been a great personal pleasure and privilege. Our colleagues on the Commission have encouraged us in the analysis here presented and have discussed some of the issues involved. IUPAC has given significant financial support to the Commission, but most financial support has been donated by individuals and institutions of many nations. Most prominent by far have been the National Bureau of Standards (renamed National Institute of Standards and Technology in 1988) and the Institute for Reference Materials and Measurements (formerly the Central Bureau for Nuclear Measurements) under whose auspices the work here presented was carried out with personal encouragement by its Director, Dr. W. Müller.

11. References

[1] Guide to the Expression of Uncertainty in Measurement, International Organization for Standardization, ISO, Geneva (1993).

[2] Element by Element Review of their Atomic Weights, Pure and Applied Chemistry, **56**, 695–768 (1984).

[3] Atomic Weights of the Elements 1991, Pure and Applied Chemistry, **64**, 1519–1534 (1992).

[4] C. Eisenhart, *Science*, **160**, 1201 (1968).

[5] Atomic Weights of the Elements 1969, Pure and Applied Chemistry, **21**, 93–108 (1970).

[6] J.R. De Laeter, P. De Bièvre, H.S. Peiser, Mass Spectrometry Reviews, **11**, 193–245 (1992).

[7] R.G. Bates and E. Wichers, J. Res. Natl. Bur. Standards, **59**, 9–17 (1957).

[8] T.J. Murphy, Private Communication.

[9] H.H. Ku, Statistical Concepts in Metrology − Postscript on Statistical Graphics, NBS Special Publication, **747**, 32–35 (1988).

TABLE 1. Chart of changes in IUPAC recommended atomic-weight values in relation to the prior estimates of their uncertainties

Order	(E)	A_r Before change	Associated U	Year of change	New A_r	D(E)	D(E)/U
1	K	39.102	0.003	1971	39.098	−0.004	−1.333
						Lower cut-off	**−1.075**
2	Ni	58.70	0.01	1979	58.69	−0.01	−1
3	Xe	131.30	0.01	1979	131.29	−0.01	−1
4	Ti	47.90	0.03	1979	47.88	−0.02	−0.667
5	Fe	55.847	0.003	1993	55.845	−0.002	−0.667
6	Ce	140.12	0.01	1985	140.115	−0.005	−0.5
7	Ti	47.88	0.03	1993	47.867	−0.013	−0.433
8	Sm	150.4	0.1	1979	150.36	−0.04	−0.4
9	H	1.0080	0.0003	1971	1.0079	−0.0001	−0.333
10	Ni	58.71	0.03	1973	58.70	−0.01	−0.333
11	Ba	137.34	0.03	1975	137.33	−0.01	−0.333
12	W	183.85	0.03	1991	183.84	−0.01	−0.333
13	Pt	195.09	0.03	1979	195.08	−0.01	−0.333
14	Cl	35.453	0.001	1985	35.4527	−0.0003	−0.3
15	Ba	137.33	0.01	1985	137.327	−0.003	−0.3
16	Lu	174.97	0.01	1977	174.967	−0.003	−0.3
						Lower fourth	**−0.25**
17	Ca	40.08	0.01	1983	40.078	−0.002	−0.2
18	In	114.82	0.01	1991	114.818	−0.002	−0.2
19	Pa	231.0359	0.0001	1985	231.03588	−0.00002	−0.2
20	Si	28.086	0.003	1975	28.0855	−0.0005	−0.167
21	Ir	192.22	0.03	1993	192.217	−0.003	−0.1
22	U	238.029	0.001	1979	238.0289	−0.0001	−0.1
23	Li	6.941	0.003	1983	6.941	0	0
24	Ne	20.179	0.003	1979	20.179	0	0
25	Mg	24.305	0.001	1985	24.3050	0	0
26	K	39.0983	0.0003	1979	39.0983	0	0
27	Ar	39.948	0.003	1979	39.948	0	0
28	Ga	69.723	0.004	1987	69.723	0	0
29	Mo	95.94	0.03	1975	95.94	0	0
30	Ru	101.07	0.03	1983	101.07	0	0
31	Ag	107.8682	0.0003	1985	107.8682	0	0
32	Xe	131.29	0.03	1985	131.29	0	0
						Median	**0.00**
33	La	138.9055	0.0003	1985	138.9055	0	0
34	Lu	174.967	0.003	1981	174.967	0	0
35	Hf	178.49	0.03	1985	178.49	0	0
36	Ta	180.9479	0.0003	1979	180.9479	0	0
37	Hg	200.59	0.03	1989	200.59	0	0
38	Re	186.2	0.1	1973	186.207	0.007	0.07
39	B	10.81	0.01	1983	10.811	0.001	0.1
40	K	39.098	0.003	1975	39.0983	0.0003	0.1
41	Cr	51.996	0.001	1983	51.9961	0.0001	0.1
42	Cd	112.41	0.01	1985	112.411	0.001	0.1
43	He	4.00260	0.00001	1983	4.002602	0.000002	0.2
44	Pd	106.4	0.1	1979	106.42	0.02	0.2
45	Ag	107.868	0.001	1981	107.8682	0.0002	0.2
46	Sb	121.75	0.03	1989	121.757	0.007	0.2333
47	Ga	69.72	0.01	1983	69.723	0.003	0.3
48	Os	190.2	0.1	1991	190.23	0.03	0.3
						Upper fourth	**0.3**
49	Tl	204.383	0.001	1985	204.3833	0.0003	0.3
50	V	50.9414	0.0003	1977	50.9415	0.0001	0.3333
51	Zn	65.37	0.03	1971	65.38	0.01	0.3333
52	Ni	58.69	0.01	1989	58.6934	0.0034	0.34
53	H	1.0079	0.0001	1981	1.00794	0.00004	0.4
54	N	14.0067	0.0001	1985	14.00674	0.00004	0.4
55	Zr	91.22	0.01	1983	91.224	0.004	0.4
56	Tl	204.37	0.03	1979	204.383	0.013	0.4333
57	Eu	151.96	0.01	1985	151.965	0.005	0.5
58	S	32.06	0.01	1983	32.066	0.006	0.6
59	Ge	72.59	0.03	1985	72.61	0.02	0.6667
60	Sn	118.69	0.03	1983	118.710	0.02	0.6667
61	Ne	20.179	0.001	1985	20.1797	0.0007	0.7
62	Zn	65.38	0.01	1983	65.39	0.01	1
63	Cd	112.40	0.01	1975	112.41	0.01	1
64	Sb	121.757	0.003	1993	121.760	0.003	1
						Upper Cut-off	**1.125**