

Günther Meinrath

## Extended traceability of pH: an evaluation of the role of Pitzer's equations

Received: 16 January 2002 / Revised: 12 August 2002 / Accepted: 12 August 2002 / Published online: 8 October 2002  
© Springer-Verlag 2002

**Abstract** The 2002 IUPAC recommendation on pH (provisional) has taken its own philosophy to provide a basis for comparable and traceable assignment of a value, from a measurement, to the quantity pH. Whereas the substituted 1983 IUPAC recommendation relied heavily on precisely prescribed experimental techniques and procedures, the current recommendation defines a hierarchical relationship between references for comparison (primary and secondary standards) and objective criteria on the comparison of measurements with these standards. The recommendation aims at a traceability chain from the national metrological institution (NMI) level down to field and laboratory measurements. Currently, however, the traceability chain is developed to the level of certified reference materials (CRM), namely the above mentioned primary and secondary standards. To complete the traceability chain, several theoretical and practical aspects have to be pondered. In part, the methods for comparative assessment of different options have yet to be developed. As an illustrating example of the complexity of issues to be considered in a further extension of the traceability chain is estimation of the doubt associated with Pitzer coefficients. The Pitzer equations for activity coefficient modelling are explicitly mentioned in the 2002 IUPAC recommendation on pH (provisional) as enabling possible improvement in the ionic strength extrapolations to zero ionic strength. An assessment of uncertainty of ternary Pitzer coefficients is given for the first time.

**Keywords** Metrology · pH measurement · Pitzer coefficients · Uncertainty analysis · Latin hypercube sampling · Bootstrap

### Introduction

pH is one of the most frequently determined quantities inside and outside chemistry. pH is, in fact, often measured routinely and understood as a matter of course. On the other hand, the quantity pH is difficult to convey and difficult to determine as soon as the problems resulting from its definition are fully considered. The quantity pH is, for instance, the only quantity given in the IUPAC Green Book immeasurable by its definition. In the 1983 IUPAC “compromise” recommendation, these important features of pH were hidden behind the two mutually exclusive operational definitions that always led to a value from a pH measurement – even though, in most situations, to different values for the same sample [1, 2, 3].

The new 2002 IUPAC recommendation on pH (provisional) no longer concentrates on measurement procedures. To measure means to compare [4]. A measurement is “a set of operations having the object of determining the value of a quantity” [5]. Hence, values are compared with other values. To enable comparisons, a hierarchical relationship among reference materials must exist. Reference materials carry accepted values; national reference materials carry values that by the Mutual Recognition Agreement (MRA) [6, 7] are mutually accepted by the national metrological institutes (NMI) in 38 economically developed countries. The hierarchical order is established by objective criteria against which measured values are interpreted. By applying criteria and reference materials to the comparisons (measurements) a traceability chain is established relating the value from one measurement to all other values related to the same reference by an unbroken chain of comparisons each with stated uncertainty [5, 8].

A major advantage of metrological concepts over operational definitions based on specific experimental methods is that assigned values are traced back to SI units, i.e. they are independent of the method used. Currently, operations leading to a value of the quantity pH are based on electrochemical measurements [9] which are not entirely free from conventions. Hence assigned values are not fully traced back to SI units with uncertainties mainly originat-

G. Meinrath (✉)  
RER Consultants Passau,  
Schiessstattweg 3a, 94032 Passau, Germany  
e-mail: rer@panet.de

ing from insufficient insight into electrolyte solution theory. The uncertainty of  $\text{pH}(\text{S})$  of primary standards with respect to the SI system has been estimated to be  $U(\text{pH}(\text{S}))=0.04$  at the present state of the art [9]. There is an explicit statement in the 2002 IUPAC recommendation on  $\text{pH}$  (provisional) that this uncertainty is expected to be minimized by application of the Pitzer theory of electrolyte solutions, particularly for solutions of high ionic strength. So far, however, this theory has been left for further investigation, mainly because of lack of the data necessary to perform the respective calculations, for example dissociation constants of buffer materials. Further uncertainties are involved in potentiometric determinations, e.g. by glass combination electrodes, the preferred experimental methods of  $\text{pH}$  determination, for example by the occurrence of junction potentials.

There are other techniques, e.g. photochemical methods [10, 11] that might have advantages in certain environments. Traceability to SI units of these methods remains to be established, to incorporate  $\text{pH}$  values thus obtained into the traceability chain mentioned above. A clear advantage of photometric over potentiometric methods is the absence of junction potentials. The potential advantages of the traceability concept realized in the 2002 IUPAC recommendation on  $\text{pH}$  (provisional) however are not yet fully realized. There are several issues of mostly statistical nature to be clarified as discussed elsewhere [12]. To make the advantages work practically, the traceability chain must be available from the national reference laboratories down to laboratory and field measurement level. Currently, the 2002 IUPAC recommendation on  $\text{pH}$  (provisional) considers mainly primary and secondary standards provided by NMI and/or accredited laboratories. Currently, analysts rely on commercial calibration material that is not yet traceable to the primary standards at the NMI. The 2002 IUPAC recommendation leaves the detailed way open to future discussions on how to establish traceability between primary/secondary standards and calibration solutions at the working level. A series of options must be considered and a wide range of interests accommodated. Establishing a traceability chain requires references and criteria against which these options and interests can be tested. Without the 2002 IUPAC recommendation of  $\text{pH}$  (provisional), any future activity would lose direction. It must be emphasized that to make its promises work it must be completed by future discussions [12, 13]. Some elements of these discussions relating to statistical considerations of the assignment of a value from a calibration–estimation procedure to a measurement have been given previously [12]. Some further aspects focusing on statistical properties of Pitzer activity coefficient estimation will be provided in the following discussion.

---

### The role of statistics in chemistry

The search for objective criteria to assess the information content of a collection of data is the task of statistics.

Knowledge of statistics is certainly not a strength of chemists in general. Statistics do not have a significant share in chemical education. This apparent negligence in statistical education is partly paralleled by the negligence in information science education for chemists.

The comparatively minor share of analytical chemistry in some academic curricula is certainly not improving the situation [14].

Analysis of the statistical content of papers published in two renowned journals of analytical chemistry in 1994 concluded that in 8% of the papers the statistics were simply incorrect, in 40% the authors had trouble understanding, defining, or using confidence limits, and in 25% of the papers the design was inappropriate or weak, so statistics could not be used legitimately to summarize results [15]. The establishment of traceability chains, evaluation of uncertainty statements for given samples of experimental data, and comparison of given experimental data with previously obtained information or data from the literature requires a working knowledge in statistics and numerical computation.

Natural sciences direct attention to the elucidation of truth assumed to be present in nature (i.e. the true value of a quantity) and hidden only by human limits to control equipment, environment, and experiment. However, there is nothing like “true uncertainty” in nature. Uncertainty expresses doubt and is an essential part of communicating scientific results. The true value is supposed to be known to nature; the uncertainty is a part of our limited abilities to extract this truth. A reasonable estimate of uncertainty is an indispensable part of each experimentally measured value of a quantity [16, 17]. Statistical techniques offer objective means of obtaining such estimates. Because an uncertainty is not a quantity found in nature but is a measure of (human) doubt, different statistical methods provide different aspects instead of one value that might be taken for granted. There is no single point of view. However, any analysis is better than no analysis.

Much focus in teaching statistics to chemists is given to the normal distribution, hypothesis testing (sometimes called “Fisherian statistics” because of the dominating contribution of the biologist R.G. Fisher), and frequentist arguments (sometimes named “NPW” statistics for the statisticians J. Neyman, K. Pearson, and A. Wald). Presented in this way, statistics can be presented in a mathematically formal language. A fundamental justification for this presentation can be found in the central-limit theorem which states that variates which are sums of many independent effects tend to be normally distributed as the number of effects becomes large. This theorem can be proven. The proof is given in most fundamental textbooks on probability and mathematical statistics. The central-limits theorem, however, does not provide an explanation for other observations that are a part of the every-day laboratory experience of most experimenters, i.e. outliers, correlations, and dependent data [18, 19]. The emphasis on mathematical presentation, the lack of experience in handling the results of a statistical analysis, the variability of results depending on the statistical method used, the

readiness by which modern instruments present large numbers of numerical values of analytical quantities [20] and – in part – an uncritical attitude towards computers might be a part of the reason students of chemistry consider statistics a “curse” [15].

Values of chemical quantities are often the result of complicated procedures in which common progression-of-error analysis cannot provide satisfactory results. Correlation of different levels, non-linearity and non-normality are some of the most often encountered nuisance factors contributing to the variability of results in repeated determinations of a quantity [21, 22]. Contemporary statistics in combination with modern high-speed computers are powerful statistical tools for handling these situations [23, 24, 25]. Efron has mentioned in 1990 that a today’s data analyst can afford to expend more computation on a single problem than the world’s yearly total of statistical computation in 1920. Since 1990 the CPU clock speed of an average desktop computer has increased by another factor of approximately 1000.

The still increasing relevance and impact of chemical measurements, e.g. in international trade (hormones in food, BSE, foot and mouth disease, pesticides in green tea, are just a few key words capturing public attention) demands dependable and defensible criteria for assessment of the quality of a measurement [26]. International agreements like MRA and ISO standards (e.g. ISO 17025) [26, 27] underpin these necessities. It would be unwise to ignore modern statistical tools and their power in supporting our need to assess the reasonable doubt to be associated with a value from chemical measurement. A possible barrier to accepting these techniques might result from the observation that chemists have tended to overestimate the accuracy of their measurements [25, 26, 28, 29] or have ignored uncertainty in the subsequent uses of a value [27, 29].

In summary, analytical chemistry will need metrology to express and communicate the quality of the information produced. Metrology in turn needs thorough statistical analysis of the experimental data obtained. It seems that the share of statistics in chemical curricula must be extended to meet metrological challenges.

---

### Approach to single-ion activity coefficient by use of splitting conventions

When selecting a model to interpret experimental data, two extreme positions can be taken – a model with many parameters being sufficiently flexible to model all kinds of data at the expense of poor predictive power, or a model with few parameters that is not as flexible but might have better predictive abilities. Between these both extreme positions, the transition is fluent and the abundance of ionic strength corrections available in literature mirrors this situation. In Section 5.4 of the 2002 IUPAC recommendation on pH (provisional), the Pitzer virial approach to ionic strength correction is explicitly mentioned. Hope is expressed that these equations might over-

come some limitations in the Bates–Guggenheim convention [9, 30].

All elements of the preceding argumentation can be found in the following discussion of Pitzer coefficient uncertainty. Even though widely used, only a single study dealing with Pitzer coefficient uncertainty is available [31]. The possibility that activity coefficients calculated on basis of the Pitzer equations, at least in certain circumstances, might give unreliable results, has not occurred.

The Pitzer equations (in the single-ion activity formulation of Møller, Harvie, and Weare [32]) are given in the Appendix. These authors also have shown the descriptive power of the Pitzer equations when applied to multi-component electrolyte systems of high ionic strength. To describe a binary electrolyte system, three terms,  $\beta^{(0)}$ ,  $\beta^{(1)}$ , and  $C^\phi$ , are required (the term  $\beta^{(2)}$  will not be considered). In ternary systems additional term, the mixing terms  $\Psi$  and  $\theta$ , are needed. The Pitzer virial equations are empirical equations based on statistical thermodynamics. Nevertheless these terms represent interactions between ions of like and opposite charge but without physical meaning. It is, hence, futile to discuss and evaluate possible numerical relationships within these terms.

The Pitzer equations have been fitted to many binary and ternary electrolyte systems [32, 33, 34, 35, 36]. Interestingly, Pitzer coefficient data are based on almost the same basic electrolyte data sources. In several instances the electrolyte data (commonly osmotic coefficients,  $\phi$ ) are smoothed or selected from a larger amount of available data before interpretation [37]. It has, furthermore, been observed that the fitting criteria have a considerable effect on the optimum value of a value obtained from such a fitting exercise [31].

In the framework of pH determination in aqueous solutions, ionic strength corrections play an important role at two critical points. First point is the extrapolation to zero ionic strength currently done via the Bates–Guggenheim convention. The second point is the derivation of single-ion activity coefficients of  $H^+$  at higher ionic strengths, e.g. for predicting liquid junction potentials.

The quantity pH is immeasurable because its definition involves a single-ion activity. Such single-ion activities cannot be experimentally determined by use of thermodynamically valid electrochemical cells. A feasible route to single-ion activities is a suitable splitting convention in which, for a reference ion pair, the assumption is made that the mean activity coefficient, is split into a cationic and an anionic part. Several splitting conventions have been proposed in the past. Some relevant conventions are listed below:

$$\ln \gamma_{Cl} = \ln \gamma_{HCl}^\circ(I) : \text{hydrochloric acid convention (16 m)} \quad (1)$$

$$\ln \gamma_{Cl} = \ln \gamma_{NaCl}^\circ(I) : \text{sodium chloride convention (6.0 m)} \quad (2)$$

$$\ln \gamma_{Cl} = \ln \gamma_{KCl}(I) : \text{extended MacInnes convention (4.6 m)} \quad (3)$$

$$\ln \gamma_{Cl} = -A \frac{\sqrt{I}}{1 + 1.5\sqrt{I}}: \quad (4)$$

Bates – Guggenheim convention (low ionic strength)

The current definition of the quantity pH is based on the Bates–Guggenheim convention. Because the true single-ion activity coefficient is unknown, any decision to split mean activity coefficients is necessarily conventional. Because of the electrical neutrality requirement, any difference between true, but unknown, cationic single activity coefficients is offset by the corresponding difference in the anionic single activity coefficient. It is therefore necessary to combine only single-ion activity coefficients derived exclusively under the same splitting convention. The splitting convention itself must be established separately for each ionic strength. Thus, deciding on a splitting convention limits the complete system of single-ion activities by the solubility of the salt on which the splitting convention is based. The solubilities of the different salts are given in parentheses in Eqs. (1), (2), (3), and (4).

The Bates–Guggenheim convention is not connected with a specific salt, and is limited to low ionic strengths. Extension of any of the splitting conventions given above to ionic strengths higher than 0.1 mol L<sup>-1</sup> introduces inconsistency because the extrapolation to zero ionic strength is based on a system different from extrapolation to higher ionic strength. It should be emphasized that such an inconsistency is not necessarily a disadvantage as long as the uncertainty contributed by this inconsistency to the total uncertainty of the quantity pH can be assessed and as long as this uncertainty is lower than the contribution resulting from the uncertainty in the splitting convention chosen to represent extrapolation to higher ionic strength. Again, considerable work will be necessary to identify the possible choices and to establish a dependable framework in which these uncertainties can be assessed objectively.

### Analysis of the uncertainty in binary Pitzer coefficients

The suitability of the Pitzer interaction model for description of complex electrolyte systems at high ionic strengths has been successfully demonstrated [32, 36]. These demonstrations have contributed to the popularity of the Pitzer model in several fields of electrolyte solution chemistry. As mentioned earlier [40], most applications ignore the fact that Pitzer coefficients are derived from experimental measurements and, hence, are affected by uncertainty. The complexity of the equations (cf. Appendix) seems to make assessment of uncertainty prohibitively demanding. In a pioneering effort, however, Marshall et al. [31] have performed statistical analysis on Pitzer coefficients by ordinary linear regression (OLS) based on the observation that Pitzer's equations are linear.

Ordinary linear regression analysis with the familiar optimisation criterion of least sum of residual squares is based on the conditions:

1. a linear function is a correct expectation function;
2. the response from experiment is the expectation function plus random disturbance;
3. the disturbance is independent of the expectation function;
4. each disturbance has a normal distribution;
5. each disturbance has a zero mean;
6. the disturbances have equal variances; and
7. the disturbances are distributed independently.

These limitations of ordinary linear regression result in unsatisfactory estimates if dependent observations, correlated residuals, non-normally and/or residuals following different distributions, etc., occur, as is true for most experimental data [38].

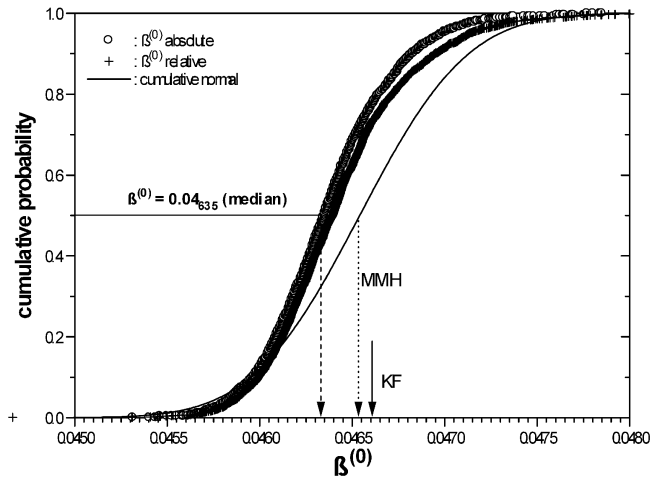
The limitations of ordinary linear regression are well known in the statistical literature [39, 40, 41] and can be overcome by computer-intensive statistical methods, e.g. the bootstrap [23, 24, 25]. These methods replace sophisticated mathematical treatment by computing power and can be applied to complex models that are difficult or even impossible to treat otherwise. These methods are widely applied in social and medical sciences and are becoming increasingly used in chemistry [21, 22, 42].

Pitzer equations are commonly calculated from the parameters evaluated, and have been tabulated by Kim and Frederick [33, 34]. These parameters are derived from determinations of osmotic coefficients compiled by Hamer and Wu [37], Goldberg [43, 44], Goldberg and Nuttall [45], Goldberg, Nuttall and Staples [46], Robinson and Stokes [47] and others. In the compilation of Hamer and Wu, the various procedures used to combine activity coefficient data from different sources and different methods used to arrive at "best" values are discussed in some detail. It is, furthermore, shown, by presenting activity coefficient data for KCl as an example, that different assessments will arrive at slightly different "best" values.

In the following, the 27 KCl activity coefficient data given Table 28 in the compilation of Hamer and Wu [37] are taken as a basis of discussion. These data are of interest because the extended MacInnes splitting convention is based on KCl. In addition, 30 NaCl activity coefficient data from Table 16 [37] are used to obtain CDF (cumulative distribution functions) for NaCl which were subsequently used in conjunction with 20 activity coefficients in the ternary system Na–K–Cl from [48] (electrode C data only) to obtain uncertainty estimates for ternary coefficients.

The computer-intensive methods applied for statistical analysis were bootstrapping and jackknifing. Latin hypercube sampling in combination with bootstrap statistics has been used to estimate a CDF for ternary Pitzer coefficients conditional on the CDF for the binary systems KCl and NaCl. These are the first uncertainty estimates for ternary Pitzer coefficients.

The theory of the statistical methods is well established [23, 24]. A bootstrap analysis is a computer-intensive resampling technique creating subsamples from *n* experimental data by drawing randomly *n* data points from the original data with replacement. Hence, in a subsample a

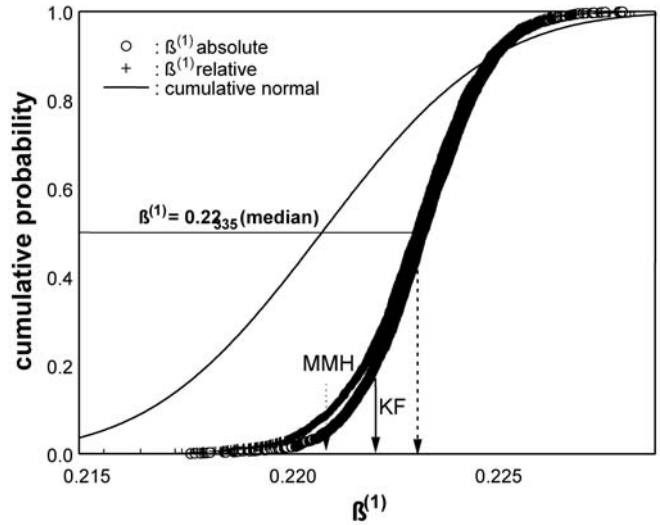


**Fig. 1** CDF of Pitzer coefficient  $\beta^{(0)}$  for KCl compared with results from normal distribution analysis by Marshall, May, and Hefter [31] and the value given in the compilation of Kim and Frederick [33]. The close agreement between the three distributions (OLS, bootstrap with absolute residuals, and bootstrap with relative residuals) is apparent and might indicate the effect of data preprocessing and smoothing [37]

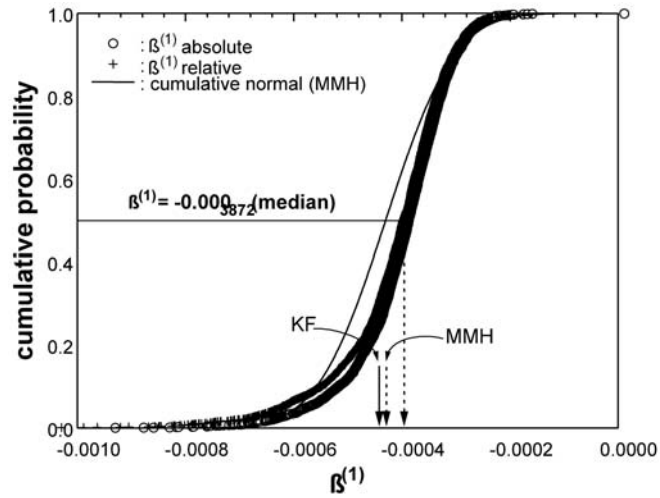
certain data point might appear several times whereas another data point is not represented in a subsample. The process of subsample creation and subsequent interpretation by the theoretical model is repeated a large number of times, where large number means  $m=1000$  and more. Thus for each model parameter  $m$  realisations are obtained. These realisations can be ordered. By giving a weight of  $1/m$  to each of the realisations, the empirical cumulative distribution functions (CDF) are obtained. The empirical CDF is a maximum likelihood estimate of the true but unknown distribution function. There is no need to presume a CDF like the normal distribution, etc. Inference to empirical CDF can be made by nonparametric statistical techniques such as Kolmogorov–Smirnov tests or Wilcoxon–Mann–Whitney tests [49]. Application of these tests requires either familiarity with statistical programs (i.e. SPSS or S plus) and/or computational ability. An account on these techniques, with examples, is given elsewhere [25, 49]. No further account will be given.

Results from bootstrapping the KCl data are given in Figs. 1, 2, and 3 for terms  $\beta^{(0)}$ ,  $\beta^{(1)}$ , and  $C^\phi$ . (these terms are defined in the Appendix). The cumulative distribution functions (CDF) are compared with the normal distributions of Marshall, May and Hefter [31] (MMH) and mean values reported by Kim and Frederick [33] (KF). For the bootstrap analysis two different CDF have been obtained. Both distributions result from different optimisation criteria applied in the fitting procedure. The absolute criterion minimizes the sum of squared residuals whereas the relative criterion takes the magnitude of the data into account by dividing the residuals by the experimental values before squaring.

Only for  $\beta^{(1)}$  does the OLS distribution deviates from the bootstrap result. There is no obvious reason for this deviation. For the other terms, however, the three distri-



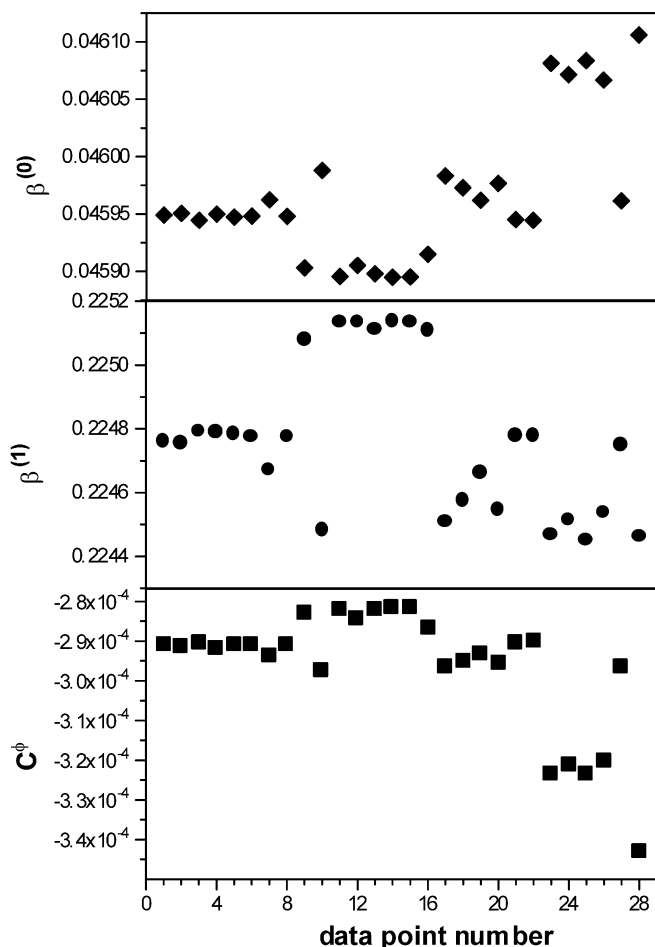
**Fig. 2** CDF of Pitzer coefficient  $\beta^{(1)}$  for KCl compared with results from normal distribution analysis of Marshall, May, and Hefter [31] and the value given in the compilation of Kim and Frederick [33]



**Fig. 3** CDF of Pitzer coefficient  $C^\phi$  for KCl compared with results from normal distribution analysis of Marshall, May, and Hefter [31] and the value given in the compilation of Kim and Frederick [33]

butions and the four location estimates (mean values in case of KF and MMH, medians for the bootstrap estimates) are close together. This can be assumed to be a result of data preprocessing and smoothing.

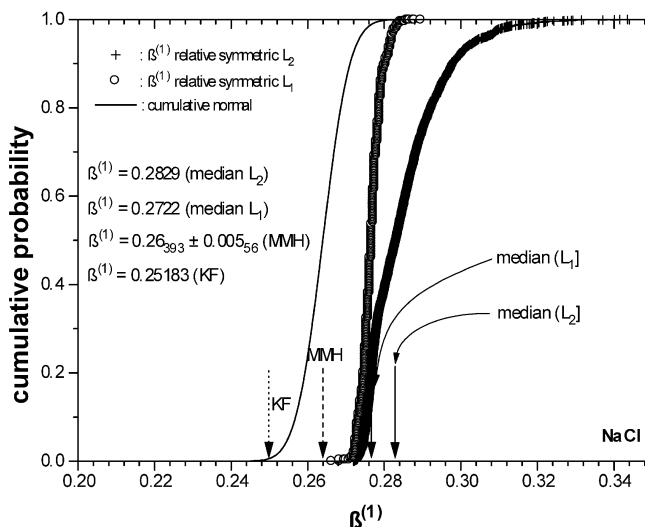
The correlation among the data cannot be easily removed by smoothing. Correlation is graphically visualized in Fig. 4, in which the result of a jackknife analysis is given for the binary coefficients of KCl. In a jackknife analysis the effect of successively omitting one of the data points after the other from the fitting procedure is calculated. The binary estimates thus produced vary in such a way that the values of  $C^\phi$  and  $\beta^{(1)}$  are high if the values for  $\beta^{(0)}$  are low, and vice versa. In addition, the strong effect



**Fig. 4** Variability of the mean value of binary Pitzer coefficients  $\beta^{(0)}$ ,  $\beta^{(1)}$ , and  $C^\phi$  calculated by jackknife analysis. The negative correlation between  $\beta^{(1)}$  and both  $\beta^{(0)}$  and  $C^\phi$  is readily visible as is the strong effect of the data points at higher ionic strength (higher data point no.)

of the data points at higher ionic strength (data point nos. 22–28) are clearly visible. The strong shift especially in the values for  $\beta^{(0)}$  and  $C^\phi$  on omission of data point no. 28 is suspicious – the results are almost outlying. Because of the sensitivity of OLS to extraneous observations, the wide confidence limits given by OLS for  $\beta^{(1)}$  might be the result of the large effect of the data point for the highest concentration. None of the values varies sufficiently, however, to enable inclusion of the mean values reported by KF ( $\beta^{(0)}=0.04661$ ,  $\beta^{(1)}=0.22341$ ,  $C^\phi=-0.00044$ ) into the diagrams of Fig. 4.

The KCl example shows that none of the four location measures provided by the different approaches is significantly deviating on both the bootstrap CDF and the normal distribution when the .16 and .84 percentiles of the CDF (corresponding to  $1\sigma$  in the normal case) are taken as the confidence limits. Four legitimate location estimates are, nevertheless, available – none of which can claim exclusive consideration – and there are further arguments leading to other values! It is common to choose the sum of squared residuals as the optimization criterion – either ab-



**Fig. 5** Comparison of fitting results of  $\beta^{(1)}$  for the binary electrolyte system NaCl. The  $L_1$  and  $L_2$  criteria are applied in the bootstrap analysis, whereas MMH uses the  $L_2$  criterion in combination with normal distribution assumptions

solute or relative. This criterion is commonly termed the  $L_2$  criterion in the statistical literature. Its ubiquity results from maximum likelihood considerations of data sets for which the conditions 1–7 given above are valid. In this case the optimum value is the maximum likelihood estimate of the quantity of interest. If these conditions are not valid (and they are not for almost all experimental data sets) the  $L_2$  criterion has lost its special position. It is then natural to sum the absolute differences, the so-called  $L_1$  criterion. Consequences are shown in Fig. 5 for  $\beta^{(1)}$  for NaCl. In this case the location measures deviate significantly on basis of the parameter distributions. The KF mean value is below the lower .95 percentile of the MMH cumulative normal distribution. The CDF also differ slightly but both are above the upper .95 percentile of MMH cumulative normal distribution. Please note that these differences do not favour or reject one of the location measures but require the user of the data to be appropriately sceptical about the quality of information extractable from the experimental data available. The analysis shows, for example, that already the second decimal place of the value for  $\beta^{(1)}$  is doubtful; it is inappropriate to tabulate the value to five decimal places ( $\beta^{(1)}=0.25183$ ) as given in KF.

It is an illustrative example of the flexibility of the Pitzer model itself that even grossly deviating values do not result in fitting problems; at least no such example has been indicated. The Pitzer model undoubtedly has high descriptive power, partly a consequence of the many terms involved in the fitting process. For a ternary system, i.e. KCl–NaCl the six binary system terms  $\beta^{(0)}_{\text{KCl}}$ ,  $\beta^{(1)}_{\text{KCl}}$ ,  $C^\phi_{\text{KCl}}$ ,  $\beta^{(0)}_{\text{NaCl}}$ ,  $\beta^{(1)}_{\text{NaCl}}$ , and  $C^\phi_{\text{NaCl}}$  and the ternary system properties  $\psi_{\text{K-Na}}$  and  $\theta_{\text{K-Na-Cl}}$  are needed. The large number of variables results in high fitting ability but the uncertainty introduced by the many uncertainty-affected variables is large. Consequently, the predictive power of the Pitzer equations is restricted.

## An account of the estimation of Pitzer coefficients of ternary electrolyte solutions

Ternary electrolyte solutions, i.e. the NaCl–KCl system, require as additional Pitzer coefficients the mixing coefficients  $\theta_{ij}$  and  $\psi_{ijk}$ . The virial coefficients  $\theta_{ij}$  account for interactions of ions of equal sign whereas the virial coefficients  $\psi_{ijk}$  account for cation–cation–anion and cation–anion–anion triple interactions. These coefficients can be calculated from osmotic coefficients of mixtures of the ternary systems of interest of different composition. In evaluating these coefficients the binary coefficients  $\beta^{(0)}$ ,  $\beta^{(1)}$ , and  $C^\phi$  are necessary input data. It is a relatively simple procedure to put mean values into an OLS regression procedure and to extract the mean ternary coefficients conditional on the binary solutions and the experimental data. The procedure is less straightforward if the empirical distributions replace the mean values. There is also no algorithm available for estimating confidence regions conditional on binary coefficient distributions as is the Gauss–Legendre algorithm in the binary coefficient example where only data variability had to be accounted for.

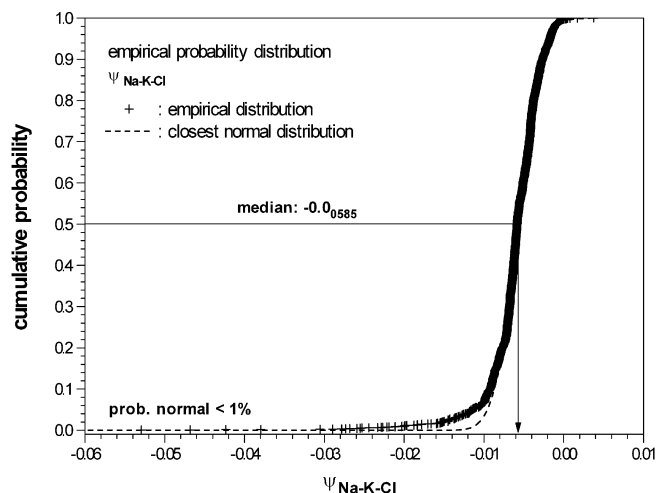
It is possible to perform a Monte Carlo study in which binary coefficients are randomly drawn from the binary coefficient empirical distributions. Many cumulative normal distributions would result. There is no reason to assume the central-limit theorem would be valid for six contributing distributions.

The approach taken here is a combination of Latin hypercube sampling (LHS) [49, 50] and bootstrapping. Twenty LHS samples of the six binary coefficients (three for each binary electrolyte system) have been generated from the empirical CDF. Each sample is subsequently used as input

vector for a bootstrap analysis with 2500 replicates. The advantage of the LHS is its efficiency in homogeneously representing all sections of a distribution (especially the tails of a distribution) in a limited number of repetitions. Twenty LHS samples were selected on the basis of a variance analysis – the variance of repeated calculations stabilized at approximately 12 to 15 LHS samples. The procedure was implemented in a higher computer language and the total of 50,000 bootstrap replicates took 16 h on a 700 MHz CPU. The bootstrap results were pooled and a result for  $\psi_{K-Na-Cl}$  is given in Fig. 6 on basis of 2500 regularly spaced points of 50,000 data points. The long tail is evident. The median is not different from the value  $\psi_{Na-K-Cl} = -0.0098$  given by KF – a consequence of the almost normal distributions of the binary coefficients and smoothing of the ternary data in Ref. [48]. A Kolmogorov–Smirnov test results in a probability of less than 1% that the CDF is normal, however. The maximum deviation between the closest fitting normal distribution and the CDF does not occur in the long tail but almost in the centre of the distribution. As is observed with many ternary parameters of univalent-univalent electrolyte systems, the hypothesis that these parameters are actually zero and the observed small values are random effects would not be easy to reject.

Analysis of uncertainties associated with Pitzer coefficients thus indicates that these uncertainties are not negligible. The uncertainties associated with the Bates–Guggenheim convention are quite narrow. It is highly doubtful whether replacement by a Pitzer equation approach will actually result in an improvement.

Last but not least it must be remembered that this analysis is by far not exhaustive. Possible correlation with the data, e.g. because of smoothing of the data tabulated in the literature, correlation of the coefficients, and non-linearity effects are not discussed. With reference to the hope expressed in the 2002 IUPAC recommendation on pH (provisional) that the Pitzer model might help improve the extrapolation to low ionic strength it must be remembered that the power of the Pitzer model lies at high ionic strengths. It is, however, necessary to stay within one splitting convention when considering extrapolation both to high ionic strength by use of a Pitzer virial model and to zero ionic strength by use of the Bates–Guggenheim convention. As already mentioned in the 2002 IUPAC recommendation on pH (provisional), the electrolyte data available are not suitable for estimation of activity coefficients for more unusual salts. The discussion has, furthermore, been limited to univalent–univalent electrolytes. In addition, the problem of estimating activity coefficients from multidimensional uncertainty distributions is not an easy task. Computer-based techniques on basis of Bayesian statistics are available but are neither easy to implement nor easy to explain [51, 52].



**Fig. 6** Cumulative empirical probability distribution function of the ternary coefficient  $\psi_{Na-K-Cl}$ . The CDF is obtained conditional on the CDF of the binary parameters by a combination of LHS with a bootstrap scheme. From the in total 50,000 estimated values of  $\psi_{Na-K-Cl}$  an interpolation of 2500 regularly spaced data points is shown. The probability of the resulting CDF being normal is assessed to be less than 1% from a Kolmogorov–Smirnov test

## Reflections on the future development of pH measurement

Because of the ubiquitous importance of the medium water in many fields inside and outside chemistry, the quan-



tity pH is of fundamental relevance. One aspect often not given its due importance is that pH measurement and its quality also affects the opinion of non-chemists about chemistry itself. These users rely on the validity of protocols and procedures and expect reliability. To disappoint these expectations has the potential to damage the reputation of chemistry. Thus claimed progress in pH measurement must persist under scrutinizing investigation. It is in the central focus of metrology in chemistry to create objective measures against which such critical inquiries can be performed. The 2002 IUPAC recommendation on pH (provisional) is a first stage in this task for pH. Further stages must follow. Several subjects can be identified in which detailed options must be measured critically and as quantitative as possible [13]:

- completion of the traceability chain down to the laboratory and field level;
- extension to a wider range of ionic strengths;
- extension to a wider range of media; and
- derivation of calibration protocols enabling measurement of traceable values for different analytical techniques

None of these four topics (the list is not complete by far) has a simple solution. The challenge is in part the need to develop objective criteria and methods to appraise possible options comparatively. The analysis of Pitzer coefficient uncertainty has been given in some detail to give an idea of the complexity of the issue. The rôle of glass electrode calibration has been given elsewhere [12]. A more general approach to UV–visible spectroscopy is also to be found elsewhere [22]. Surely, alternative options will turn up and extensions will become necessary.

---

## Conclusions

The tenor of the preceding sections can be summarized in three points. To establish information of defined quality, we need:

- references and standards;
- chains of comparison and evaluation (traceability chain); and
- objective criteria.

References and standards enable us to measure (=to compare) equal things. The traceability chain gives an indication of the quality of our activities and a measure against which to judge progress and improvement. The objective criteria provide the openness of the system towards future developments, new demands, and alternative methods. At present progress and improvement can easily be claimed but is very difficult to assess or even to prove, i.e. in the event of conflicting interest. Within the system the hierarchical structure is the stabilizing factor.

The instrumental revolution in analytical chemistry from 1970 to the present and the enormous increase in electronic calculation power have resulted in computer-based equipment capable of obtaining large amounts of data in a

short time. Chemical information is enclosed in these data heaps together with random noise, scatter, and other nuisance effects. Objective criteria to judge data are developed by statistical science. The amounts of data obtained during electrode calibration and interpretation of sample potentials can be handled easily. But with the intention of avoiding overinterpretation and underestimation, statistical techniques can aid rational discussion.

An example has been given on the assignment of uncertainty to Pitzer model virial coefficients. These coefficients are widely used in literature but almost always given without an associated estimate of uncertainty. Because Pitzer model equations are multi-parameter equations their modelling capacity is great, as already amply demonstrated in literature. Conversely, their predictive power is low, because each experimentally determined value contributes measurement uncertainty from its determination, as shown by estimating this uncertainty by computer-intensive statistical methods. Comparison with an OLS approach by Marshall, May, and Hefter [31] for the binary electrolytes indicates that within the limitations of each of the different methods the assigned uncertainty estimates do not highly diverge substantially.

Future work in the field of pH measurement must be directed toward extension of the new IUPAC recommendation on pH to a wider range of applications and media. Several options are possible. Each option, however, needs detailed and unprejudiced analysis. Past experience indicates that such an analysis is time-consuming and tedious, but unavoidable.

The determination of pH is largely based on electrochemical measurements. For most situations, the focus on electrochemical methods is justified by their efficiency in time and costs. The new IUPAC recommendation on pH also indicates the problems that result from our incomplete theoretical understanding of electrolyte solutions. It is therefore necessary to investigate other techniques, i.e. photometric pH measurement in the same manner and with the same scrutiny as for electrochemical methods. Alternative traceable methods for study of pH are urgently needed.

The still increasing influence of chemical measurements, e.g. in international trade, demand dependable criteria for appraisal of a value for a quantity from chemical measurements that is defensible even in the event of conflicting interests. This development gains further momentum from international agreements and standards, e.g. the BIPM “Guide to the Expression of Uncertainty”, the “Mutual Recognition Agreement” and the already widely recognized ISO standards, e.g. ISO 17025. The changes and requirements emerging from this development are difficult to foresee. The procedure(s) of assigning values to the quantity pH is, however, a pioneering example of the challenges that face other areas of chemistry, i.e. chemical thermodynamics with its wide fields of applications.



## Appendix

The Pitzer equations:

Equations (5), (6), (7), and (8) describe summations of interactions between different ions as a function of ionic strength,  $I$ , on the molality scale, where  $\phi$  gives the osmotic coefficient and  $\gamma$  a single-ion activity coefficient based on the appropriate splitting convention. Indices  $M$ ,  $c$ , and  $c'$  denote cations, indices  $X$ ,  $a$ , and  $a'$  denote anions.  $N$  and  $n$  represent uncharged species. Molal concentrations are represented by the symbol  $m$ .

$$\begin{aligned} \sum_i m_i(\phi - 1) = & 2 \frac{-A^\phi \sqrt{I}^3}{1 + 1.2\sqrt{I}} + \sum_{c=1}^{N_c} \sum_{a=1}^{N_a} m_c m_a (B_{ca}^\phi + ZC_{ca}) \\ & + \sum_{c=1}^{N_c-1} \sum_{c'=c+1}^{N_c} m_c m_{c'} (\Phi_{cc'}^\phi + \sum_{a=1}^{N_a} m_a \Psi_{cc'a}) \\ & + \sum_{a=1}^{N_a-1} \sum_{a'=a+1}^{N_a} m_a m_{a'} (\Phi_{aa'}^\phi + \sum_{c=1}^{N_c} m_c \Psi_{aa'c}) \\ & + \sum_{n=1}^{N_n} \sum_{a=1}^{N_a} m_n m_a \lambda_{na} + \sum_{n=1}^{N_n} \sum_{c=1}^{N_c} m_n m_c \lambda_{nc} \end{aligned} \quad (5)$$

$$\begin{aligned} \ln \gamma_M = & z_M^2 F + \sum_{a=1}^{N_a} m_a (2B_{Ma} + ZC_{Ma}) + \sum_{c=1}^{N_c} m_c (2\Phi_{Mc} + \sum_{a=1}^{N_a} m_a \Psi_{Mca}) \\ & + \sum_{a=1}^{N_a-1} \sum_{a'=a+1}^{N_a} m_a m_{a'} \Psi_{aa'M} + |z_M| \sum_{c=1}^{N_c} \sum_{a=1}^{N_a} m_c m_a C_{ca} + \sum_{n=1}^{N_n} m_n (2\lambda_{nM}) \end{aligned} \quad (6)$$

$$\begin{aligned} \ln \gamma_X = & z_X^2 F + \sum_{c=1}^{N_c} m_c (2B_{cX} + ZC_{cX}) + \sum_{a=1}^{N_a} m_a (2\Phi_{Xa} + \sum_{c=1}^{N_c} m_c \Psi_{Xac}) \\ & + \sum_{c=1}^{N_c-1} \sum_{c'=c+1}^{N_c} m_c m_{c'} \Psi_{cc'X} + |z_X| \sum_{c=1}^{N_c} \sum_{a=1}^{N_a} m_c m_a C_{ca} + \sum_{n=1}^{N_n} m_n (2\lambda_{nX}) \end{aligned} \quad (7)$$

$$\ln \gamma_N = \sum_{c=1}^{N_c} m_c (2\lambda_{nc}) + \sum_{a=1}^{N_a} m_a (2\lambda_{na}) \quad (8)$$

The charge of a species is given by  $z$ . The Debye–Hückel coefficient  $A^\phi$  is given by Eq. (9);  $A^\phi=0.392$  at 25°C.

$$A^\phi = \frac{1}{3} \sqrt{\left( \frac{2\pi N_0 d_w}{1000} \right) \left( \frac{e^2}{DKT} \right)^3} \quad (9)$$

where  $N_0$  is the Loschmidt number,  $d_w$  and  $D$  are the density and static dielectric constant, respectively, of water, and  $e$  represents the electron charge. Temperature  $T$  is given in Kelvin. The symbol  $k$  is the Boltzmann constant.

$C$  and  $C^\phi$  are related according to Eq. (10):

$$C_{MX} = \frac{C_{MX}^\phi}{2\sqrt{|z_M z_X|}} \quad (10)$$

where  $Z$  is defined by Eq. (11):

$$Z = \sum_i |z_i| m_i \quad (11)$$

and  $F$  is given by Eq. (12):

$$\begin{aligned} F = & -A^\phi \left( \frac{\sqrt{I}}{1 + 1.2\sqrt{I}} + \frac{2}{1.2} \ln(1 + 1.2\sqrt{I}) \right) + \sum_{c=1}^{N_c} \sum_{a=1}^{N_a} m_c m_a B_{ca}' + \\ & \sum_{c=1}^{N_c-1} \sum_{c'=c+1}^{N_c} m_c m_{c'} \Phi_{cc'}' + \sum_{a=1}^{N_a-1} \sum_{a'=a+1}^{N_a} m_a m_{a'} \Phi_{aa'}' \end{aligned} \quad (12)$$

Coefficients  $B$  are understood as functions of ionic strength,  $I$ , and described by Eq. (13). Functions  $g(x)$  and  $g'(x)$  are calculated from Eq. (14):

$$\begin{aligned} B_{MX}^\phi = & \beta_{MX}^{(0)} + \beta_{MX}^{(1)} e^{-\alpha_{MX}\sqrt{I}} + \beta_{MX}^{(2)} e^{-12\sqrt{I}} \\ B_{MX} = & \beta_{MX}^{(0)} + \beta_{MX}^{(1)} g(\alpha_{MX}\sqrt{I}) + \beta_{MX}^{(2)} g(12\sqrt{I}) \\ B_{MX}' = & \beta_{MX}^{(1)} \frac{g'(\alpha_{MX}\sqrt{I})}{I} + \beta_{MX}^{(2)} \frac{g'(12\sqrt{I})}{I} \end{aligned} \quad (13)$$

Constant  $\alpha$  is 2.0 for electrolyte solutions with at least one monovalent ion. For 2:2 ion pairs or even higher charged electrolytes  $\alpha=1.4$ . For electrolytes with monovalent ions,  $\beta^{(2)}$  is almost zero and is not considered in this implementation.

$$g(x) = 2 \frac{1 - (1+x)e^{-x}}{x^2} \quad (14)$$

$$g'(x) = -2 \frac{1 - (1+x + \frac{x^2}{2})e^{-x}}{x^2}$$

with

$$x_{MN} = 6z_M z_N A^\phi \sqrt{I} \quad (15)$$

Mixing terms  $\Phi$  of asymmetric electrolytes depend on ionic strength,  $I$ , and are made up from terms  $\theta$ ,  $E_\theta$ , and  $E_\theta'$ :

$$\begin{aligned} \Phi_{ij}^\phi = & \theta_{ij} + {}^E \theta_{ij}(I) + I {}^E \theta'_{ij}(I) \\ \Phi_{ij} = & \theta_{ij} + {}^E \theta_{ij}(I) \\ \Phi'_{ij} = & {}^E \theta'_{ij}(I) \end{aligned} \quad (16)$$

$$\begin{aligned} {}^E \theta_{MN}(I) = & \frac{z_M z_N}{4I} (J_0(x_{MN}) - \frac{1}{2} J_0(x_{MM}) - \frac{1}{2} J_0(x_{NN})) \\ {}^E \theta'_{MN}(I) = & \frac{z_M z_N}{8I^2} (J_1(x_{MN}) - \frac{1}{2} J_1(x_{MM}) - \frac{1}{2} J_1(x_{NN})) - \frac{{}^E \theta_{MN}}{I} \end{aligned} \quad (17)$$

Functions  $J_0$  and  $J_1$  are integrals according to Eq. (18). Suitable approximations are given by Pitzer.

$$J_0(x) = \frac{1}{4}x - 1 + \frac{1}{x} \int_0^\infty \left[ 1 - e^{-\frac{x}{y}} e^{-y} \right] y^2 dy \quad (18)$$

$$J_1(x) = \frac{1}{4}x - \frac{1}{x} \int_0^\infty \left[ 1 - \left( 1 + \frac{x}{y} e^{-y} \right) e^{-\frac{x}{y}} e^{-y} \right] y^2 dy$$

## References

1. Covington AK, Bates RG, Durst RA (1983) *Pure Appl Chem* 55:1467
2. Baucke FGK, Naumann R, Alexander-Weber C (1993) *Anal Chem* 65:3244
3. Naumann R, Alexander-Weber C, Baucke FGK (1994) *Fresenius J Anal Chem* 349:603
4. Price G (2001) *Measurement* 29:293–305
5. Deutsches Institut für Normung (1994) (ed) *International vocabulary of basic and general terms in metrology (VIM)*. Beuth, Berlin, Germany
6. Bureau International des Poids et Mesures (1999) (ed) *Mutual recognition agreement*. Bureau International des Poids et Mesures, Sèvres, France
7. Richter W (1999) *Fresenius J Anal Chem* 365:569–573
8. De Bièvre P (2000) *Accred Qual Assur* 5:224–230
9. Buck RP, Rondinini S, Covington AK, Baucke FGK, Brett CMA, Camoes MF, Milton MJT, Mussini T, Naumann R, Pratt KW, Spitzer P, Wilson GS (2001) *The measurement of pH – definition, standards and procedures, report of the working party on pH, IUPAC Provisional Recommendation*. IUPAC
10. Clayton TD, Byrne RH (1993) *Deep-Sea Res* 140:2115
11. Yamazaki H, Sperline RP, Freiser H (1992) *Anal Chem* 64:2720
12. Meinrath G, Spitzer P (2000) *Mikrochim Acta* 135:155
13. Spitzer P, Meinrath G (2002) (eds) *Importance of traceable pH measurements in science and technology*, Proc 166th PTB seminar. PTB, Braunschweig, Germany
14. Felber H (2001) *Accred Qual Assur* 6:439
15. Thompson M (1994) *Analyst* 119:127 N
16. Bureau International des Poids et Mesures (1993) *Guide to the expression of uncertainty in measurement*. Bureau International des Poids et Mesures, Sèvres, France
17. Eurachem/CITAC (1995) *Quantifying uncertainty in analytical measurement*. LGC, Teddington, UK
18. Beckman RJ, Cook RD (1983) *Technometrics* 25:119
19. Dixon WJ (1950) *Ann Math Stat* 2:488
20. Chalmers RA (1993) *Talanta* 40:121
21. Meinrath G (2000) *Fresenius J Anal Chem* 368:574
22. Meinrath G, Lis S (2002) *Anal Bioanal Chem* 372:333–340
23. Efron B, Tibshirani R (1991) *Science* 253:390
24. Efron B, Tibshirani RJ (1993) *An introduction to the bootstrap*. monographs on statistics and applied probability 57. Chapman and Hall, London, UK
25. Meinrath G (2000) *Chemometrics Intell Lab Syst* 51:175
26. De Bièvre P (2001) *Accred Qual Assur* 6:451
27. ISO (1999) *ISO 17025, General requirements for the competence of testing and calibration laboratories*. ISO, Genève, Switzerland
28. Meinrath G (2001) *Fresenius J Anal Chem* 369:690
29. Wold S (1993) *Chemometrics Intell Lab Syst* 30:109
30. Covington AK, Ferra MIA (1994) *J Solution Chem* 23:1
31. Marshall SL, May PM, Hefter GT (1995) *J Chem Eng Data* 40:1041
32. Harvie CE, Møller N, Weare JH (1984) *Geochim Cosmochim Acta* 48:723
33. Kim HT, Frederick WJ (1988) *J Chem Eng Data* 33:177
34. Kim HT, Frederick WJ (1988) *J Chem Eng Data* 33:278
35. Pitzer KS (1991) *Ion interaction approach: theory and data correlation*. In: Pitzer KS (ed) *Activity coefficients in electrolyte solutions*, 2nd edn, CRC Press, Boca Raton, USA, chap 3, p 76
36. Harvie CE, Weare JH (1980) *Geochim Cosmochim Acta* 44:981
37. Hamer WJ, Wu Y-C (1972) *J Phys Chem Ref Data* 1:1047
38. Beran J (1992) *Stat Sci* 7:404
39. Meinrath G, May P (2002) *Mine Water Environ* 21:24–35
40. Meinrath G (1998) *J Radioanal Nucl Chem* 232:179
41. Frank IE, Friedman JH (1993) *Technometrics* 35:109
42. Roy T (1994) *J Chemometrics* 8:37
43. Goldberg RN (1979) *J Phys Chem Ref Data*. 8:1005
44. Goldberg RN (1981) *J Phys Chem Eng Data* 10:671
45. Goldberg RN, Nuttall RL (1978) *J Phys Chem Ref Data* 7:263
46. Goldberg RN, Nuttall RL, Staples BR (1979) *J Phys Chem Ref Data* 8:923
47. Robinson RA, Stokes RH (1965) *Electrolyte solutions*. Butterworths, London, UK
48. Huston R, Butler JN (1969) *Anal Chem* 41:1695
49. Meinrath G, Ekberg C, Landgren A, Liljenzin JO (2000) *Talanta* 51:231
50. McKay MD, Beckman RJ, Conover WJ (1979) *Technometrics* 21:239
51. Casella G, George EI (1992) *Am Stat* 46:167
52. Lecoutre B, Lecoutre M-P, Poitevineau J (2001) *Int Statist Rev* 69:399