

STATISTICAL MODEL OF SYSTEMATIC ERRORS: AN ASSESSMENT OF THE BA-CU AND CU-Y PHASE DIAGRAM

E.B. Rudnyi

Chemistry Department

Moscow State University

119899 Moscow

Russia

e-mail rudnyi@comp.chem.msu.su

homepages <http://www.chem.msu.su/~rudnyi/welcome.html>

Abstract

In a series of experimental measurements, the difference among values obtained in several experiments (between-errors) quite often are greater than the reproducibility scatter within an experiment (within-errors). This is typically explained in terms of systematic experimental errors. The application of a method devised in mathematical statistics, i.e., The Estimation of Variance Components, allows us to treat this problem. In the present work, the use of this method is considered for the non-linear thermodynamic model pertaining to the assessment of the BaCu and Cu-Y phase diagrams. The linear error model comprising a reproducibility error with the shift and tilt systematic errors was employed to describe the scatter observed. Special attention is paid to visualizing the quality of the fit.

Keywords

random factor, mixed model, estimation of variance components, maximum likelihood, systematic errors, phase diagram, thermodynamic modeling

Introduction

For a given system, the more experiments that are conducted, the more reliable the results, especially when the experiments are done by different authors. Nobody seems to disagree with this statement. Along the way, however, one has to deal with systematic errors

because the results of the distinct experiments usually differ more between each other than the reproducibility error in a single experiment. Formally speaking, there is a statistically significant difference between distinct experiments, i.e., the ratio of the corresponding sum of squares is more than the Fisher's criterion allows.

In order to proceed further in such a situation, mathematical statistics suggest to us random and mixed models [1] wherein deviations from the theoretical model are simulated as several random effects, with different components due to measurement errors (experimental noise) and systematic errors (laboratory or experimental factors). Extensive use of such models can be found in meta-analysis [2, 3].

The phrase 'statistical description of the systematic errors' implied above may look somewhat contradictory at first. An IUPAC Report [4] even defines systematic errors as those for which statistics cannot be used:

"Uncertainties which are known to exist or are inferred but for which insufficient information is available for a statistically valid treatment to be used, so that their effect upon the results is dependent on the experience and judgment of the author are usually called systematic errors."

A middle ground between the IUPAC definition and statistics is still quite possible. To this end, it is necessary to introduce a hierarchy of errors. If there are results of only a single experiment, mathematical statistics can say nothing about systematic errors without doubt. Yet, when there are results from several experimenters, it is conceivable that different experimenters have made the same systematic errors by chance, and if so, statistics could say something about them. In the latter case, there are two levels of random errors: 1) the reproducibility errors within a single experiments and 2) the systematic errors that are constant in the single experiment and that change randomly from one experiment to another.

It is possible to go further and say that in all the experiments there is the same error, e.g., all the experimenters employed the same sample that was not of high purity. Statistics does not work with such an error again until new experiments indicate that the latter error has become random. A new level of random errors can be introduced and the analysis repeated.

In a previous paper [5], a mixed model comprising a non-linear theoretical model with two-level random effects to describe between- and within-errors has been considered. The laboratory effect was composed of a random shift and a random "tilt", that is, a random

regression slope shared by all measurements from the same series. Such an error model (linear error model) is quite simple and practical, but powerful enough to be used in many important applications. The example of simultaneously processing 26 different experiments on vaporization thermodynamics of potassium chloride under the linear error model has been given, and the results obtained have been compared with those after the traditional least square method.

Two examples of the application of the linear error model to the assessment of the phase diagram (Ba-Cu and Cu-Y binaries) are given in the present work. The important feature of the current case is that it was necessary to treat the results from qualitatively different experiments. The implications for the data processing only will be discussed; a discussion of the results from a thermodynamics point of view is given elsewhere [6, 7].

Linear error model

It is assumed that there are results from M experimental series and each series is comprised of N_i pairs of experimental points

$$\{y_{ij}, x_{ij}\}; i = 1, \dots, M; j = 1, \dots, N_i \quad (1)$$

where x_{ij} is the input and y_{ij} is the response. Note that y_{ij} and/or x_{ij} in distinct series may be completely different physical quantities. The theoretical relationship between input and response in each series is assumed to be known, except for the unknown parameters (vector Θ)

$$\varepsilon_{ij} = f_i(x_{ij}, \Theta) + \varepsilon_{ij} \quad (2)$$

where ε_{ij} is the total error in y_{ij} and the notation f_i implies that the theoretical model may be different in different series. The only imperative is that the unknown parameters are the same inside all the theoretical relationships.

The following linear error model has been previously introduced: [5]

$$\varepsilon_{ij} = \varepsilon_{r,ij} + \varepsilon_{a,i} + \varepsilon_{b,i}(x_{ij} - x_i) \quad (3)$$

where

$$x_i = (\sum_j x_{ij})/N_i$$

It consists of the reproducibility error $\varepsilon_{r,ij}$ (experimental noise) and two systematic errors, $\varepsilon_{a,i}$ and $\varepsilon_{b,i}$. The former accounts for the shift systematic error and the latter for the tilt laboratory factor (tilt systematic error). Note that the linear error model is a special case of the general error model described in [1].

Equation (3) implies that the experimental points in a series are randomly scattered over the line that shifted and tilted from the theoretical model because of the systematic errors. Fig 1 presents the supposed deviate behavior in a single series and allows us to write down a geometrical interpretation of equation (3)

$$FC = FE + DC + ED \quad (4)$$

It should be mentioned that the series scattered over the theoretical model without any shift and tilt is desirable. However, this is usually impossible because there are other series besides that one displayed in Fig. 1 and the distance between the series is more than the measurement noise.

In order to make statistics work, it is assumed that the errors $\varepsilon_{r,ij}$, $\varepsilon_{a,i}$ and $\varepsilon_{b,i}$ are independent random quantities with a distribution that is not too far from normal. The expected values of errors are assumed to be zero and the variances are $\sigma_{r,i}^2$, $\sigma_{a,i}^2$ and $\sigma_{b,i}^2$ accordingly.

The task is to determine the unknown parameters (vector Θ) in the theoretical models and unknowns variances $\sigma_{r,i}^2$, $\sigma_{a,i}^2$ and $\sigma_{b,i}^2$ (variance components) from the experimental values (1). To this end, applying the maximum likelihood method, i.e., maximizing the function

$$L = \ln \det \mathbf{D}(\varepsilon) - \varepsilon' \mathbf{D}(\varepsilon)^{-1} \varepsilon \quad (5)$$

has been suggested. In Eq. (5), the unknown parameters (vector Θ) are inside the vector ε consisting from all the deviates in all the series, and the unknown variance components, $\sigma_{r,i}^2$, $\sigma_{a,i}^2$ and $\sigma_{b,i}^2$ are in the dispersion matrix $\mathbf{D}(\varepsilon)$ of the vector ε . More details on the linear error model including the algorithm for maximizing (5) and the comparison with the least squares are given elsewhere [5].

Thermodynamic model

Chemical thermodynamics states that if the molar Gibbs energies of all the phases in the system are known, one can compute all the thermodynamic properties of the system as well as the complete phase diagram [8] (see also publications in the journal CALPHAD). This allows us to establish only the molar phase Gibbs energies as unknowns. All other measurable quantities then will be some function of these unknowns.

In the case of the Ba-Cu binary, it is necessary to determine Gibbs energies of the three phases. The two phases are stoichiometric compounds, BaCu and BaCu₁₃ with Gibbs energy given by Eqs (16-18) in Annex and the third is a solution, liquid alloy Ba-Cu with Gibbs energy given by Eqs. (19-20). The coefficients a_{mn} , b_{mn} , c_{mn} in Eq. (18) and the interaction parameters A_p , B_p , C_p in Eq. (20) are the unknowns and all of them compose the vector Θ in Eq. (2). In the case of the Cu-Y binary, the Gibbs energies of the five stoichiometric compounds, Cu₆Y, Cu₄Y, Cu₇Y₂, Cu₂Y, CuY and the binary liquid alloys of Cu-Y are to be determined. In both binaries, the Gibbs energies of pure components as functions of temperature were taken from SGTE compilation [9].

In order to determine unknowns in the Gibbs energies of the stoichiometric compounds and the liquid alloys, the available experimental values are to be processed. The equations corresponding to the function f_i in Eq. (2) for those experiments that were available for the Ba-Cu and Cu-Y binaries are given below.

There are three kinds of calorimeter experiments. Solution calorimetry of the stoichiometric compounds produces the enthalpy of formation at some temperature (usually 298 K). Under the model given by Eq. (18), the relationship between the Gibbs energy and enthalpy (Eq. 21) leads to the equation

$$\Delta_r H_{mn,ij} = a_{mn} - c_{mn} T_i + \varepsilon_{ij} \quad (6)$$

Mixture calorimetry allows us to measure the enthalpy of mixing. From Eqs. (20) and (21), one can write

$$\Delta_{mix} H_{ij} / x_{1,ij} x_{2,ij} = \sum_{p=0}^q (A_p - C_p T_i)(x_{2,ij} - x_{1,ij})^p + \varepsilon_{ij} \quad (7)$$

The enthalpy of mixing is divided by the product of $(x_{1,ij} x_{2,ij})$ in order to make the variances of reproducibility errors homogeneous. This means that in the case of Eq. (7) the hypothesis that $D(\varepsilon_{r,ij}) = \sigma_{r,i}^2$ is acceptable.

If the experimental partial enthalpies of mixing are given, then Eqs. (22) and (23) lead to the following theoretical models

$$\Delta H'_{1,ij} = \sum_{p=0}^q x_{2,ij}^2 (A_p - C_p T_i)(x_{2,ij} - x_{1,ij})^{p-1} (x_{2,ij} - x_{1,ij} - 2p x_{1,ij}) + \varepsilon_{ij} \quad (8)$$

$$\Delta H'_{2,ij} = \sum_{p=0}^q x_{1,ij}^2 (A_p - C_p T_i)(x_{2,ij} - x_{1,ij})^{p-1} (x_{2,ij} - x_{1,ij} + 2p x_{2,ij}) + \varepsilon_{ij} \quad (9)$$

Drop calorimetry yields the change in enthalpy $H_T - H_{298}$ for some composition of the system as a function of temperature. In the subsolidus area we have

$$H_T - H_{298}^{ij} = H_s(T_{ij}, x_{2,i}) - H_s(298, x_{2,i}) + \varepsilon_{ij} \quad (10)$$

where H_s is the enthalpy of the solid phases at some composition. Provided that the mole fraction $x_{2,i}$ of the second component describing the composition is located between the two stoichiometric compounds A_mB_n and A_rB_s the mole fractions of which are equal to $x_{2,mn} = n/(m+n)$ and $x_{2,rs} = s/(r+s)$, respectively, the following holds true:

$$H_s = \{H_{mn}(x_{2,rs} - x_{2,i}) + H_{rs}(x_{2,i} - x_{2,mn})\}/(x_{2,rs} - x_{2,mn})$$

Here H_{mn} and H_{rs} stand for the enthalpies of the two stoichiometric compounds, which include the unknown parameters and which may be obtained from the Gibbs energies given by Eq. (16) by the means of Eq. (21).

When drop calorimetry is performed for the melts, Eq. (10) changes to

$$H_T - H_{298}^{ij} = H_l(T_{ij}, x_{2,i}) - H_s(298, x_{2,i}) + \varepsilon_{ij} \quad (11)$$

where H_l is the enthalpy of the melt that in turn can be obtained from Eq. (19) by employing the same Eq. (21). Eq. (10) contains the unknown parameters c_{mn} and c_{rs} from Eq. (18) and Eq. (11) additionally holds the parameters A_p and C_p from the Gibbs energy of the melt (Eqs. 19-20).

Besides the thermodynamic data, there are results of thermal analysis: the coordinates of the monovariant (liquidus) and nonvariant (eutectics and peritectics) equilibria. In the first case, the temperature of the liquidus is a function of the melt composition, i.e., the mole fraction of the second component in the melt, x_2 . Thus, it can be written

$$T_{mono,ij} = T_{mono,ij}^{calc}(x_{2,ij}, \Theta) + \varepsilon_{ij} \quad (12)$$

where $T_{mono,ij}^{calc}$ is to be calculated at the given values of the unknown parameters from the equilibrium criterion (24). Equation (24) cannot be solved in the closed form and the $T_{mono,ij}^{calc}$ is to be found by numerical means.

In the case of nonvariant equilibrium, Eq. (12) should be changed to a pair of equations

$$T_{non,ij} = T_{non,ij}^{calc}(\Theta) + \varepsilon_{ij} \quad (13)$$

$$x_{2,non,ij} = x_{2,non,ij}^{calc}(\Theta) + \varepsilon_{ij} \quad (14)$$

because the temperature and the composition no longer depend on each other. To find the coordinates of the non-variant equilibrium, the equilibria criterion in the form of a system of

Eqs. (25) and (26) should be employed. The latter must be solved by numerical analysis at any given set of the unknown parameters.

Finally, it should be mentioned that not only those listed above, but all the experimental equilibrium values will be some functions of the unknowns in Eqs. (18) and (20) due to the thermodynamics laws.

Estimating unknown parameters and variance components

All available results on the Ba-Cu and Cu-Y systems found in the literature were at first subjected to expert analysis, and some works and experimental points that can clearly be classified as "wrong" were thrown out [6, 7]. Tables 1 and 2 list the literature data that passed the expert analysis and were employed for the statistical procedure. Shown are the code ascribed to the series, the experimental quantities measured with a reference to the theoretical model, the number of experimental points and the code of the paper in which the results were obtained. The experimental points were separated into the distinct series shown in Tables 1 and 2 because

- 1) they were measured by different authors, or
- 2) they were measured in the same work but the author clearly indicated that he made several distinct experiments, or
- 3) they were measured in the same experiment but they are described by different equations.

The only exception to these rules was made in the Cu-Y system with the results of the thermal analysis since each author gave just one point for the monovariant or non-variant equilibrium. This forced us to combine the values obtained by different authors in a single series for a given phase equilibrium. In all, there were 178 experimental points in 18 series pertaining to the Ba-Cu binary and 185 points in 35 series for the Cu-Y system.

The total error ε_{ij} in each series was modeled by the linear error model (3). In some series where there was no an input variable (non-variant equilibria and the enthalpy of the reaction 6), the tilt systematic error is not applicable and it was set to zero. After the expert analysis, the series were combined into groups that are separated by the solid lines in Tables 1 and 2, and the series in such a group were assumed to possess the same quality. From a statistical point of view, it means that the distribution of the reproducibility error $\varepsilon_{r,ij}$, the

distribution of the shift systematic error $\varepsilon_{a,i}$, and the distribution of the tilt systematic errors $\varepsilon_{b,i}$ for all the series in a single group can be considered to be the same, respectively. Let the index α enumerate the groups of the series. Then for all series included in the group α the following was assumed to hold:

$$D(\varepsilon_{r,ij}^\alpha) = \sigma_{r,\alpha}^2, D(\varepsilon_{a,i}^\alpha) = \sigma_{a,\alpha}^2, D(\varepsilon_{b,i}^\alpha) = \sigma_{b,\alpha}^2 \quad (15)$$

The small difference between grouping the series in the Ba-Cu and Cu-Y binaries can be seen in the case of the liquidus and nonvariant temperatures (compare Tables 1 and 2). Both nonvariant and monovariant (liquidus) temperatures are determined from the kinks obtained on the cooling or heating curve. However, if several thermal curves are measured at different compositions in the same region of the phase diagram this gives rise to several liquidus temperatures but to the same nonvariant temperature. Thus, the final precision of the nonvariant temperature should be higher than for the monovariant one. Because of that consideration, the nonvariant and monovariant temperatures were put in the two different groups in the Ba-Cu system (see Table 1). Yet, in the Cu-Y binary all these temperatures were combined in a single group because the number of measurements was not enough to separate them.

In the Ba-Cu system, the application of the hypothesis (15) and the division of the experimental series into four groups (see Table 1) gave us a total of nine unknown variance components (see Table 4). The tilt systematic error is not applicable to the nonvariant temperatures (group N) and the enthalpy of reaction (17) (group F). Also, in the original paper on the enthalpy of reaction (17) there were no primary experimental enthalpies but the mean enthalpy with its standard deviation. This forced us to set the values of $\sigma_{r,i}^2$ in the group F as a given constant *a priori*. The same considerations led us to 12 unknown variance components in the case of the Cu-Y binary (see Table 5).

The maximum likelihood function (Eq. 5) was maximized to determine the unknowns in the Gibbs energies and the unknown variance components. The number of the unknowns in the Gibbs energy should be determined during the data processing since it is impossible to say *a priori* how many terms in the Gibbs energy of mixing are necessary and whether all of three parameters in the temperature function (Eqs. 8 and 10) are necessary to fit the experimental values adequately. To this end, the simplest approach was applied. Consecutive maximization

of the function (5) has been performed starting with a small number of the unknown parameters in the Gibbs energies and then increasing their number.

The recommended number of unknowns in the Gibbs energies (see Table 3, all other parameters were set equal to zero) was chosen based upon several considerations. First, further increase in the number of unknowns did not lead to a sharp increase in the value of function L (Eq. 5) and to a decrease in the values of variance components. At the same time, the columns of the Jacobean became almost linearly dependent and the entire task ill-behaved.

Second, thermodynamic constraints were taken into account. When the parameters b_{mn} and c_{mn} were made free, this improved the description of the liquidus and non-variant temperatures but led to the prediction that the stoichiometric phases are thermodynamically unstable at room temperature. Such phase behavior ("pendent phase") occurs rather rarely, and it was assumed to be an artifact of insufficient precision in the available experimental values. Note that in similar systems both the heat capacity and the entropy of reaction (7) are close to zero [10]. Thus, the solution with the parameters b_{mn} and c_{mn} set equal to zero was preferred. A further increase in the number of terms in the Gibbs energy of mixing, in turn, led to the prediction that the liquid alloys of Ba-Cu and Cu-Y have a miscibility gap with a lower critical point that is also unlikely in such systems.

The main difference in our data processing as compared with other "phase diagram optimizations" lies in introducing the shift and tilt systematic errors in the error model (see Eq. 3). For the sake of comparison, another solution that will be referenced as ML1 has also been found. Here, the variance components corresponding to the shift and tilt systematic errors were set to zero *a priori*. In the latter case, the dispersion matrix $\mathbf{D}(\boldsymbol{\varepsilon})$ becomes diagonal, and this makes it quite close to the weighted least squares (WLS) with the difference that in WLS the weights should be assigned *a priori* and in the maximum likelihood they are obtained during the maximization procedure.

Both solutions for both binaries are presented in Tables 3 to 5.

Visualizing the quality of the fit

The first test of whether parameters found are reliable is the quality of the fit by the model of the primary experimental values. To this end, three types of graphs that can help in visualizing the quality of the fit are discussed in this section.

The first plot is the usual type of graph when the final fitted curves are displayed along with the original experimental values (see Fig. 2 and 3 in the case of the Ba-Cu binary, and Fig. from 6 to 8 for Cu-Y system). The problem in analyzing these graphs is related to the low scale that makes it difficult to follow small differences between the model and the experimental values. Differences of more than several percents only can be clearly seen on such graphs. Another problem is that the comparison between the descriptions of different heterogeneous data is quite difficult. For example, compare Fig. 6 to 8 and say what the type of values was described in the best way.

The second type of graph is presented by Fig. 4 for the Ba-Cu system and Fig. 8 and 9 for Cu-Y. On these graphs, the ordinate shows the value of the deviate divided by the reproducibility standard deviation of the corresponding series. Plotting the deviates by themselves allowed us to enhance the scale and making them dimensionless permitted us to combine the deviates for heterogeneous data. In Fig. 4, 8, and 9 much more subtle effects of the fitting are seen and the direct comparison of the heterogeneous data is possible. The behavior of the deviates clearly indicates that the difference between distinct series is more than the reproducibility noise, i.e., that there are systematic errors within the series (compare with Fig. 1). Also, it is possible to state that the linear error model (3) is applicable in the cases of the Ba-Cu and Cu-Y binaries.

The third type of graph addresses the problem of the numerous experimental points. Because of this, it is still difficult to compare the distinct series, especially at first glance. As discussed elsewhere [5], it is possible to draw such a graph when a whole series will be presented by a single point (see Fig. 5 and 11). The idea is that at any given set of parameters each series is characterized by the shift error, $\varepsilon_{a,i}$ and by the tilt error, $\varepsilon_{b,i}$. Then, plotting $\varepsilon_{b,i}$ vs. $\varepsilon_{a,i}$ values gives us a very useful graph that indicates the extent of overall agreement among all the series. Fig. 1 allows us to suggest a graphic interpretation of these errors, i.e., the shift error is the AB interval and the tilt error is equal to the ratio of DE to BD. The formulas for estimating the numerical values of $\varepsilon_{a,i}$ and $\varepsilon_{b,i}$ are given in Ref. [5].

The tilt and shift errors have different meanings and in Ref. [5] it was suggested to plot the value of $\varepsilon_{b,i}(P_i/N_i)^{1/2}$ instead of the $\varepsilon_{b,i}$ where

$$P_i = \sum_j (x_{ij} - x_i)^2$$

Besides having the same dimension, the product $\varepsilon_{b,i}(P_i/N_i)^{1/2}$ has the same meaning as the $\varepsilon_{a,i}$. The latter shows the mean difference between the total ε_{ij} and the reproducibility $\varepsilon_{r,ij}$ errors in the given series due to the shift, and the product $\varepsilon_{b,i}(P_i/N_i)^{1/2}$ displays the similar difference due to the tilt.

Finally, in order to make it possible to compare heterogeneous series between each other, the values of $\varepsilon_{a,i}$ and $\varepsilon_{b,i}(P_i/N_i)^{1/2}$ were divided by the reproducibility standard deviation of the i -th series. Thus, if a point in Fig. 5 and 11 is shifted from the center along the x-axis by the unit value it means that the shift of this series is equal to the mean reproducibility error in this series. If a point went along the y-axis by the unit value, this says that the given series has a tilt that led to the mean difference between ε_{ij} and $\varepsilon_{r,ij}$ equal to its mean reproducibility error.

Each of the graphs described above has its advantages and drawbacks, and the best results can be achieved by analyzing the three types of the graphs simultaneously.

Discussion

The approach shown in the present paper comprises two steps - the expert analysis and the statistical procedure. The role of the former should be emphasized - the final values of the recommended parameters are based heavily upon the expert analysis, i.e., in spite of the advanced statistics treatment a lot of subjectivity still remains. Actually, the subjectivity can never be completely excluded in practical applications. Hope for the "golden" algorithm that would just take raw data and produce the true answer automatically is ungrounded because before data processing we must always postulate some hypothesis that cannot be proved empirically in that treatment.

In this study, there are three statements that affect the results most. First concerns the choice of the functions for the Gibbs energies and the number of unknowns parameters inside them. It is necessary to understand that the true behavior of the Gibbs energy probably differs from the chosen functional dependency and the error produced by that fact is ignored in the discussion as follows. What can be said is just that the precision of the available experimental data is not enough to enhance the number of unknowns at the present time.

The second major hypothesis is about the linear error model (3). Figures 2 to 4 and 6 to 10 clearly display the fact that the distinct series have systematic errors and that Eq. (3)

seems to be sensible but the question whether the concrete form of Eq. (3) is the "true" one still remains.

Finally, it is implied that all the series inside the chosen group have the same quality (Eq. 15). The division of the series into the groups is likely the most subjective hypothesis. Another expert may well say that a particular work is the "right" one and other data should be ignored.

The application of the maximum likelihood method permitted the qualitative expert information only to be employed. After the hypothesis (15), maximizing the likelihood function yielded the estimates of unknown parameters and of unknown variance components concurrently. This simplifies the analysis considerably as compared with the least squares method when the dispersion matrix $\mathbf{D}(\epsilon)$ must be supplied up to the constant, i.e., there the expert must express his/her recommendations as concrete numerical values (quantitative information).

The numerical values of the variance components, namely the standard deviation estimates (square root of the variance), for the recommended solution ML are given in Tables 4 and 5. The reproducibility standard deviation, $\sigma_{r,\alpha}$ shows the scatter associated with the noise for the series in the group α . The shift $\sigma_{a,\alpha}$ and tilt $\sigma_{b,\alpha}$ standard deviations display the order of magnitude of the shift and the tilt, respectively, for the series in the group α . For example, the values of the variance components for the eight series of liquidus temperatures, L1 to L8, in the Ba-Cu system are as follows, $\sigma_{r,\alpha} = 8.2$ K, $\sigma_{a,\alpha} = 22$ K, and $\sigma_{b,\alpha} = 120$ K. This means that while the experimental noise was about 8 K the eight series are shifted from the calculated liquidus by about 22 K, and the slope of the series is different from the slope of the model by about 120 K. In the example above the slope has the same dimension as the intercept because mole fraction is a dimensionless quantity.

The mixed models do not enjoy widespread use in physical chemistry. More often, the error model includes the reproducibility error only. To this end, as was mentioned in the previous section, the solution ML1 was found when the shift and tilt errors were eliminated from the error model. The hypothesis that there is no statistically significant difference between distinct series is incorrect. It can be tested by any statistical criterion (for instance, the Fisher test), and actually can be seen from Fig. 2 to 4 and 6 to 10 without any statistics. Still,

it is interesting to compare the solutions ML and ML1 to see what the consequences of neglecting the systematic errors are.

The main difference between solutions ML and ML1 lies in the structure of the dispersion matrix $\mathbf{D}(\epsilon)$. It is non-diagonal in solution ML due to the laboratory effects being in the linear error model (3) and diagonal in solution ML1 when all the variances of the systematic errors have been set equal to zero *a priori*. Thus, the solution ML1 underestimates the systematic errors, and as a result, it overestimates the reproducibility errors (see Table 4 and 5). This means that in the circumstance displayed in Fig. 1 solution ML1 explains all the difference between experimental points and the theoretical model as reproducibility noise only.

The wrong dispersion matrix, however, gives reasonable values of the unknown parameters (see Table 3) and fits the experimental values (see Fig. 2 to 4 and 6 to 10) well. The bigger difference concerns the standard deviations of the parameters, and thus, the standard deviation of the fit. The standard deviations estimated in solution ML1 are smaller those from solution ML by factor from 1.5 to 2 and seems to be too optimistic. The reason for smaller standard deviations estimated in solution ML1 comes again from underestimating the systematic errors. Increasing a number of the experimental points in a single series reduces the uncertainty in the parameters produced by the reproducibility error but cannot affect the uncertainty coming from systematic errors. Solution ML separates the total deviation into the reproducibility noise and the systematic errors and can take into account the different influences of the number of the points on the reproducibility and systematic errors. Solution ML1 treats the total deviation as the reproducibility error only, and thus, overestimates the influence of the number of the experimental points.

This can be clearly seen in the Ba-Cu system where the main controversy is a difference between the barium partial enthalpies measured by two different groups (see series H1 and H2 in Fig. 2, 4 and 5); they differ even in sign. Yet, there is no reason to prefer one of the series to another and my expert conclusion is that the enthalpy of mixing in the Ba-Cu melts should be not far from zero. It should be noted that when both series have been excluded from the data processing it leads to a high predicted enthalpy of mixing. So, both series were put in and this happened to be a good example of how the number of points influences the results in solutions ML and ML1.

The number of the points in the series H1 is 68, rather high, and in the series H2 it is just five. The solution ML has ascribed large systematic errors to the partial enthalpies (see Table 4), and as a result, the liquidus temperatures were approximated better than the partial enthalpies (see Fig. 2 to 5). The opposite is the case in the solution ML1 where, because of equating systematic errors to zero, the high total number of points in partial enthalpies has pulled the predicted solution from the liquidus temperatures.

One may say that the series H1 is more consistent with the phase diagram data in both solutions, that can be seen in Fig. 2 to 5, and this can be a reason to neglect the series H2. This would be true if we could be sure that there is one term only in the Gibbs energy of mixing, as accepted in the present model. However, this is a hypothesis only, at present we can say nothing more without new additional experimental measurements.

In the Cu-Y system, it is possible to see that series Ts1 to Ts8 and Tl1 to Tl8 (all of them are measured in the same work, 89Qi) are tilted upward. This can hardly be seen from Fig. 7 but can clearly be detected in Fig. 10 and 11. This is an indication that either more parameters should be included in the model for the Cu-Y system or there were some inherent problems in the experimental setup of 89Qi work. The attempt to include the heat capacities in the Gibbs energies of the solid phases and the melt did not improve the situation much. Also note the difference between the series listed above and series Tl9 to Tl11 measured by different authors. So, the final expert conclusion was again that it is impossible to resolve the problem without new experiments.

Conclusions

Qualitatively different experimental results pertained to the Ba-Cu and Cu-Y systems have been processed simultaneously in order to achieve the best description of the primary experimental values, and thus, to obtain the most reliable Ba-Cu and Cu-Y phase diagrams.

The analysis of deviates have showed that, as expected, the differences between different experiments was greater than the reproducibility errors, and that the liner error model (3) was appropriate in order to describe systematic errors existing in the original works.

Three different types of plots have been employed to visualize the quality of the fit. Advantage of each plot has been discussed.

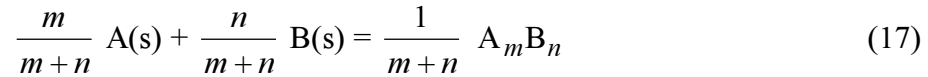
The treatment under the linear error model has been compared with a traditional approach when systematic errors are simply discarded. The result have showed that the traditional treatment may lead to acceptable values of unknown parameters but with seriously reduced standard deviations.

Annex. Thermodynamic relationships.

The consideration below is limited to an A-B binary system. Let us denote the molar Gibbs energies of the two pure components in the solid and liquid states as G_{1s} , G_{2s} , G_{1l} and G_{2l} accordingly. The Gibbs energy of the stoichiometric compound A_mB_n (point phase) is tied to the Gibbs energies of the pure solid components:

$$G_{mn} = m G_{1s} + n G_{2s} + (m + n) \Delta_r G_{mn} \quad (16)$$

where $\Delta_r G_{mn}$ is the Gibbs energy of the reaction



The temperature behavior of the Gibbs energy is usually expressed as follows

$$\Delta_r G_{mn} = a_{mn} + b_{mn} T + c_{mn} T \ln T \quad (18)$$

under an assumption that the heat capacity is constant.

The molar Gibbs energy of the binary melt, G_l is connected with the Gibbs energies of the liquid components

$$G_l = x_1 G_{1l} + x_2 G_{2l} + \Delta_{mix} G \quad (19)$$

where x_1 and x_2 are the component mole fractions (note that $x_1 + x_2 = 1$). The Gibbs energy of mixing, $\Delta_{mix} G$, can be described by a polynomial in composition (see for example [11])

$$\Delta_{mix} G = x_1 RT \ln x_1 + x_2 RT \ln x_2 + \sum_{p=0}^q x_1 x_2 (A_p + B_p T + C_p T \ln T) (x_2 - x_1)^p \quad (20)$$

The enthalpy can be calculated from the Gibbs energy as follows

$$H = -RT^2 \left(\frac{\partial(G/T)}{\partial T} \right) \quad (21)$$

The relationships between the partial enthalpies of mixing and the integral enthalpy of mixing are given by a pair of equations

$$\Delta H'_1 = \Delta_{mix} H + x_2 \left\{ \frac{\partial(\Delta_{mix} H)}{\partial x_1} - \frac{\partial(\Delta_{mix} H)}{\partial x_2} \right\} \quad (22)$$

$$\Delta H'_2 = \Delta_{mix} H + x_1 \left\{ \frac{\partial(\Delta_{mix} H)}{\partial x_2} - \frac{\partial(\Delta_{mix} H)}{\partial x_1} \right\} \quad (23)$$

In the case of the equilibrium $A_m B_n$ -L between the liquid and the given compound, the following must be held

$$G_{mn}(T_{mono,ij}^{calc}) = m \mu_1(T_{mono,ij}^{calc}, x_{2,ij}) + n \mu_2(T_{mono,ij}^{calc}, x_{2,ij}) \quad (24)$$

where μ_1 and μ_2 are the chemical potentials (the partial Gibbs energies) of the two components in the melt. These functions can be derived from the Gibbs energy of the melt (Eq. 19) by the use of the equations analogous to Eqs. (22) and (23).

When two compounds and liquid are under equilibrium, $A_m B_n$ -L- $A_r B_s$, the equilibrium criterion leads to a system of two equations with two unknowns

$$G_{mn}(T_{non,ij}^{calc}) = m \mu_1(T_{non,ij}^{calc}, x_{2,non,ij}^{calc}) + n \mu_2(T_{non,ij}^{calc}, x_{2,non,ij}^{calc}) \quad (25)$$

$$G_{rs}(T_{non,ij}^{calc}) = r \mu_1(T_{non,ij}^{calc}, x_{2,non,ij}^{calc}) + s \mu_2(T_{non,ij}^{calc}, x_{2,non,ij}^{calc}) \quad (26)$$

Acknowledgments

The research described in this publication was made possible by Grant MRL300 from the International Science Foundation and Russian Government. I greatly appreciate the invitation of the Computers in Chemistry Division of the American Chemical Society to present this work at the ACS spring meeting in 1996, and I thank the Petroleum Research Fund for sponsoring my trip.

References

- [1] C.R. Rao; J. Kleffe. Estimation of variance components and applications. North-Holland, Amsterdam, (North-Holland Series in Statistics and probability, v.3), 1988.
- [2] G.V. Glass, B. McGaw, M.L Smith. Meta-analysis in social research. Beverly Hills, CA: Sage, 1981.
- [3] L.V. Hedges, I. Olkin. Statistical methods for meta-analysis. New York: Academic Press, 1985.
- [4] G. Olofsson, S. Angus, G.T. Amstron, A.N. Kornilov. J. Chem. Thermodynamics. 13 (1981) 603.
- [5] E.B. Rudnyi. Chemometrics Intell. Lab. Sys. Accepted for publication, 1996.
- [6] E.B. Rudnyi. Zh. Fiz. Khim. 70 (1996) 986, (English translation in Rus. J. Phys. Chem.).
- [7] E.B. Rudnyi. Zh. Fiz. Khim. 70 (1996) 1007, (English translation in Rus. J. Phys. Chem.).

- [8] L. Kaufmann, H. Bernstein. Computer calculation of phase diagrams. Academic Press, New York and London, 1970.
- [9] A.T. Dinsdale. CALPHAD. 15 (1991) 317.
- [10] D.Sh. Tsagareishvili. Methods for estimating thermal and deformation properties of the crystalline neorganic substances (In Russian). Tbilisi, Metsniereba, 1977.
- [11] H.L. Lukas, J. Meiss, E.T. Honig. CALPHAD. 6 (3) 229.

Table 1. Experimental results used for the assessment of the Ba-Cu system

Exp. code	Set of values	Eq. N	Comments	N_i	Paper code
F1	$\{ \Delta_r H_{mn,ij} \}$	6	$\Delta_r H_{\text{BaCu}}$ (298 K)	1	93Kon
H1	$\{ \Delta H'_{1,ij}, x_{2,ij} \}$	8	$\Delta H'_1$ (1400 K)	68	89Nik
H2	$\{ \Delta H'_{1,ij}, x_{2,ij} \}$	8	$\Delta H'_1$ (1500 K)	6	93Use
H3	$\{ \Delta H'_{2,ij}, x_{2,ij} \}$	9	$\Delta H'_2$ (1150 K)	43	89Nik
L1	$\{ T_{mono,ij}, x_{2,ij} \}$	12	Ba-L	3	71Bru
L2	$\{ T_{mono,ij}, x_{2,ij} \}$	12	Ba-L	1	93Kon
L3	$\{ T_{mono,ij}, x_{2,ij} \}$	12	BaCu-L	5	71Bru
L4	$\{ T_{mono,ij}, x_{2,ij} \}$	12	BaCu-L	1	93Kon
L5	$\{ T_{mono,ij}, x_{2,ij} \}$	12	BaCu ₁₃ -L	2	59Bra
L6	$\{ T_{mono,ij}, x_{2,ij} \}$	12	BaCu ₁₃ -L	1	71Bru
L7	$\{ T_{mono,ij}, x_{2,ij} \}$	12	Cu-L	13	59Bra
L8	$\{ T_{mono,ij}, x_{2,ij} \}$	12	Cu-L	2	71Bru
N1	$\{ T_{non,ij} \}$	13	Ba-L-BaCu	8	71Bru
N2	$\{ T_{non,ij} \}$	13	Ba-L-BaCu	1	93Kon
N3	$\{ T_{non,ij} \}$	13	BaCu-L-BaCu ₁₃	15	59Bra
N4	$\{ T_{non,ij} \}$	13	BaCu-L-BaCu ₁₃	5	71Bru
N5	$\{ T_{non,ij} \}$	13	BaCu-L-BaCu ₁₃	2	93Kon
N6	$\{ T_{non,ij} \}$	13	BaCu ₁₃ -L-Cu	1	93Kon

The references to the original publications are elsewhere [6].

Table 2. Experimental results used for the assessment of the Cu-Y system

Exp.	Set of values	Eq.	Comments	N_i	Paper code
F1	$\{ \Delta_r H_{mn,ij} \}$	6	$\Delta_r H_{\text{CuY}}$ (298 K)	1	89Sid
H1	$\{ \Delta_{mix} H_{ij}, x_{2,ij} \}$	7	$\Delta_{mix} H$ (1410 K)	13	83Sud
H2	- " -	7	- " -	13	83Sud
H3	- " -	7	$\Delta_{mix} H$ (1373 K)	5	84Wat
H4	- " -	7	- " -	3	84Wat
H5	- " -	7	- " -	2	84Wat
H6	- " -	7	- " -	1	84Wat
H7	- " -	7	- " -	3	84Wat
H8	- " -	7	- " -	4	84Wat
H9	- " -	7	- " -	3	84Wat
H10	- " -	7	- " -	5	84Wat
H11	- " -	7	$\Delta_{mix} H$ (1963 K)	8	90Sid
Ts1	$\{ H_T - H_{298}{}_{ij}, T_{ij} \}$	10	solid, $x_Y=0.093$	6	89Qi
Ts2	- " -	10	solid, $x_Y=0.200$	7	89Qi
Ts3	- " -	10	solid, $x_Y=0.278$	5	89Qi
Ts4	- " -	10	solid, $x_Y=0.333$	7	89Qi
Ts5	- " -	10	solid, $x_Y=0.341$	5	89Qi
Ts6	- " -	10	solid, $x_Y=0.434$	4	89Qi
Ts7	- " -	10	solid, $x_Y=0.500$	6	89Qi
Ts8	- " -	10	solid, $x_Y=0.670$	5	89Qi
Tl1	- " -	11	melt, $x_Y=0.093$	4	89Qi
Tl2	- " -	11	melt, $x_Y=0.200$	4	89Qi
Tl3	- " -	11	melt, $x_Y=0.278$	5	89Qi
Tl4	- " -	11	melt, $x_Y=0.333$	4	89Qi
Tl5	- " -	11	melt, $x_Y=0.341$	3	89Qi
Tl6	- " -	11	melt, $x_Y=0.434$	3	89Qi
Tl7	- " -	11	melt, $x_Y=0.500$	6	89Qi
Tl8	- " -	11	melt, $x_Y=0.670$	5	89Qi
Tl9	- " -	11	melt, $x_Y=0.200$	1	84Wat
Tl10	- " -	11	melt, $x_Y=0.333$	1	84Wat
Tl11	- " -	11	melt, $x_Y=0.500$	1	84Wat

Table 2 (continued). Experimental results used for the assessment of the Cu-Y system

Exp.	Set of values	Eq.	Comments	N_i	Paper code
L1	$\{T_{mono,ij}, x_{2,ij}\}$	12	Cu ₄ Y -L	4	61Dom 61Hae 66Bea 89Qi
L2	- " -	12	Cu ₂ Y -L	4	61Dom 61Hae 89Qi
L3	- " -	12	CuY-L	4	61Dom 61Hae 89Qi
N1	$\{T_{non,ij}\}$	13	Cu-L-Cu ₆ Y	6	57Daa 61Dom 61Hae 75Fed 88Dui 89Qi
N2	- " -	13	Cu ₆ Y -L-Cu ₄ Y	3	61Dom 61Hae 66Bea
N3	- " -	13	Cu ₄ Y -L-Cu ₇ Y ₂	2	61Hae 66Bea
N4	- " -	13	Cu ₇ Y ₂ -L-Cu ₂ Y	3	61Dom 61Hae 89Qi
N5	- " -	13	Cu ₂ Y -L-CuY	4	61Dom 61Hae 89Qi
N6	- " -	13	CuY-L-Y	4	61Dom 61Hae 60Lov 89Qi
X1	$\{x_{2,non,ij}\}$	14	Cu-L-Cu ₆ Y	2	57Daa 61Dom
X2	- " -	14	Cu ₆ Y -L-Cu ₄ Y	1	61Dom
X3	- " -	14	Cu ₇ Y ₂ -L-Cu ₂ Y	2	57Daa 61Dom
X4	- " -	14	Cu ₂ Y -L-CuY	1	61Dom
X5	- " -	14	CuY-L-Y	2	61Dom 60Lov

The references to the original publications are elsewhere [7].

Table 3. Thermodynamics parameters (Gibbs energy in J/mol)

Ba-Cu			Cu-Y		
	ML	ML1		ML	ML1
a_{BaCu}	-2594±240	-2858±120	$a_{\text{Cu}_6\text{Y}}$	-12610±560	-12550±340
$a_{\text{BaCu}_{13}}$	-925±87	-1040±37	$a_{\text{Cu}_7\text{Y}_2}$	-16560±680	-16480±410
-	-	-	$a_{\text{Cu}_4\text{Y}}$	-17280±710	-17200±420
-	-	-	$a_{\text{Cu}_2\text{Y}}$	-20100±790	-20060±470
-	-	-	a_{CuY}	-20180±740	-20240±460
A_0	-7191±2080	-3713±480	A_0	-82460±3250	-80330±1940
-	-	-	A_1	38480±4180	41950±2300
B_0	3.36±3.60	-2.166±0.96	B_0	14.74±1.70	12.73±0.95
-	-	-	B_1	-6.61±4.80	-10.58±2.70

Table 4. Variance components estimated in the system Ba-Cu

group	$\sigma_{r,\alpha}$		$\sigma_{a,\alpha}$		$\sigma_{b,\alpha}$	
	ML	ML1	ML	ML1	ML	ML1
	kJ/mol		kJ/mol		kJ/mol	
F	3.9*	3.9*	0	0*	n/a	n/a
H	1.8	3.9	7.7	0*	41	0*
	K		K		K	
L	8.2	26.8	22	0*	120	0*
N	4.0	4.1	0.21	0*	n/a	n/a

The values were kept constant during the maximization of (5).

Table 5. Variance components estimated in the system Cu-Y

group	$\sigma_{r,\alpha}$		$\sigma_{a,\alpha}$		$\sigma_{b,\alpha}$	
	ML	WLS	ML	WLS	ML	WLS
	kJ/mol		kJ/mol		kJ/mol	
F	2.0*	2.0*	3.3	0*	n/a	n/a
H	5.6	14.3	11	0*	160	0*
	J/mol		J/mol		J/K/mol	
Ts,Tl	370	1210	1100	0*	14	0*
	K		K		K	
L,N	14	16	7.1	0*	0	0*
	dimensionless		dimensionless		dimensionless	
X	0.013	0.02	0.015	0*	n/a	n/a

The values were kept constant during the maximization of (5).

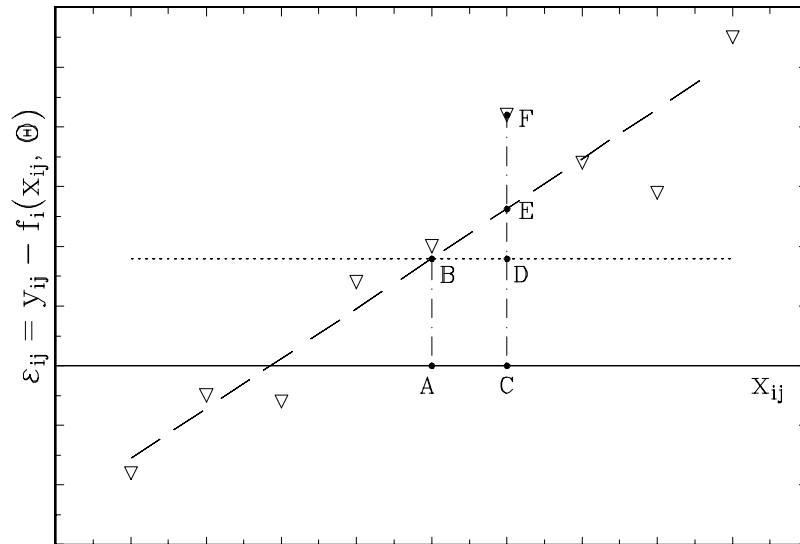


Fig. 1. Typical behaviour of the total deviates ε_{ij} in the i -th series as related to a theoretical model. The symbol ∇ shows the deviates, the long dashed line displays the best fit for the deviates, the small dashed line indicates the best fit that is parallel to the theoretical model. The point A corresponds to the mean of x_{ij} in this series.

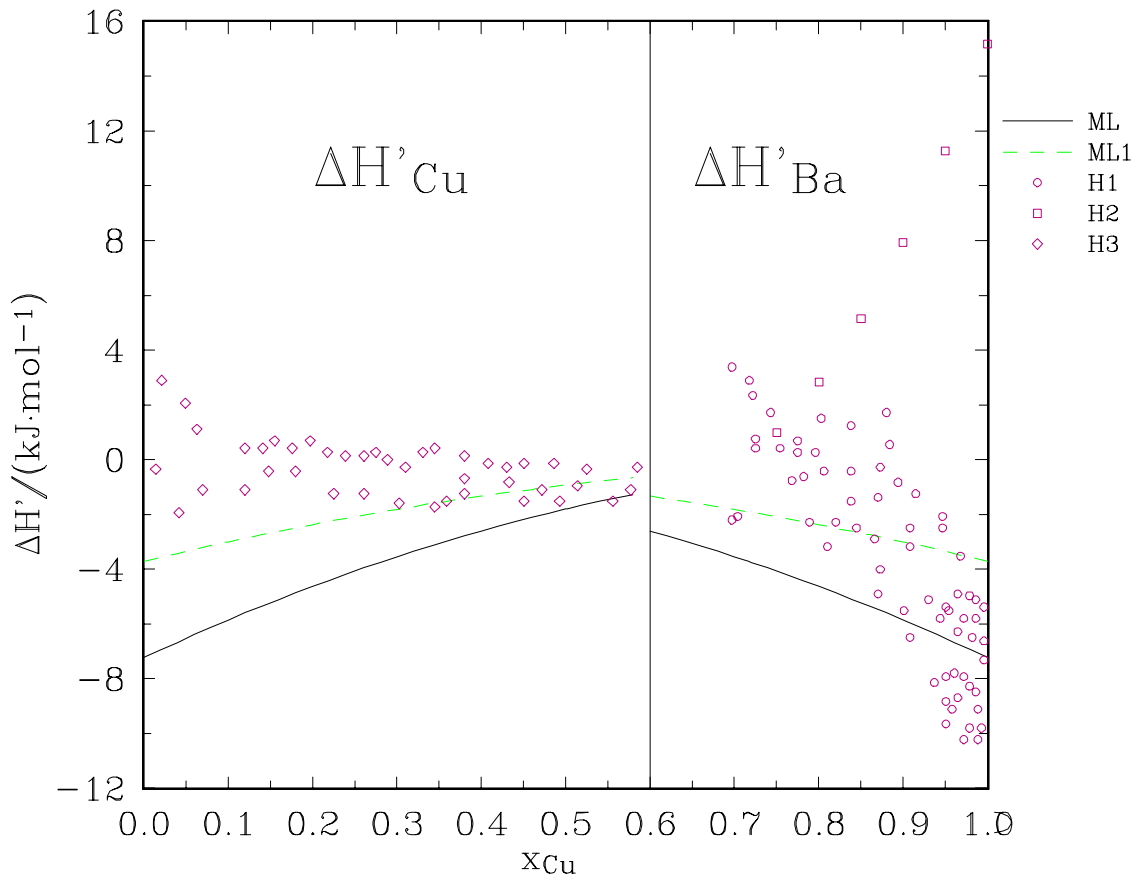


Fig. 2. Partial enthalpies in the Ba-Cu liquid alloys.

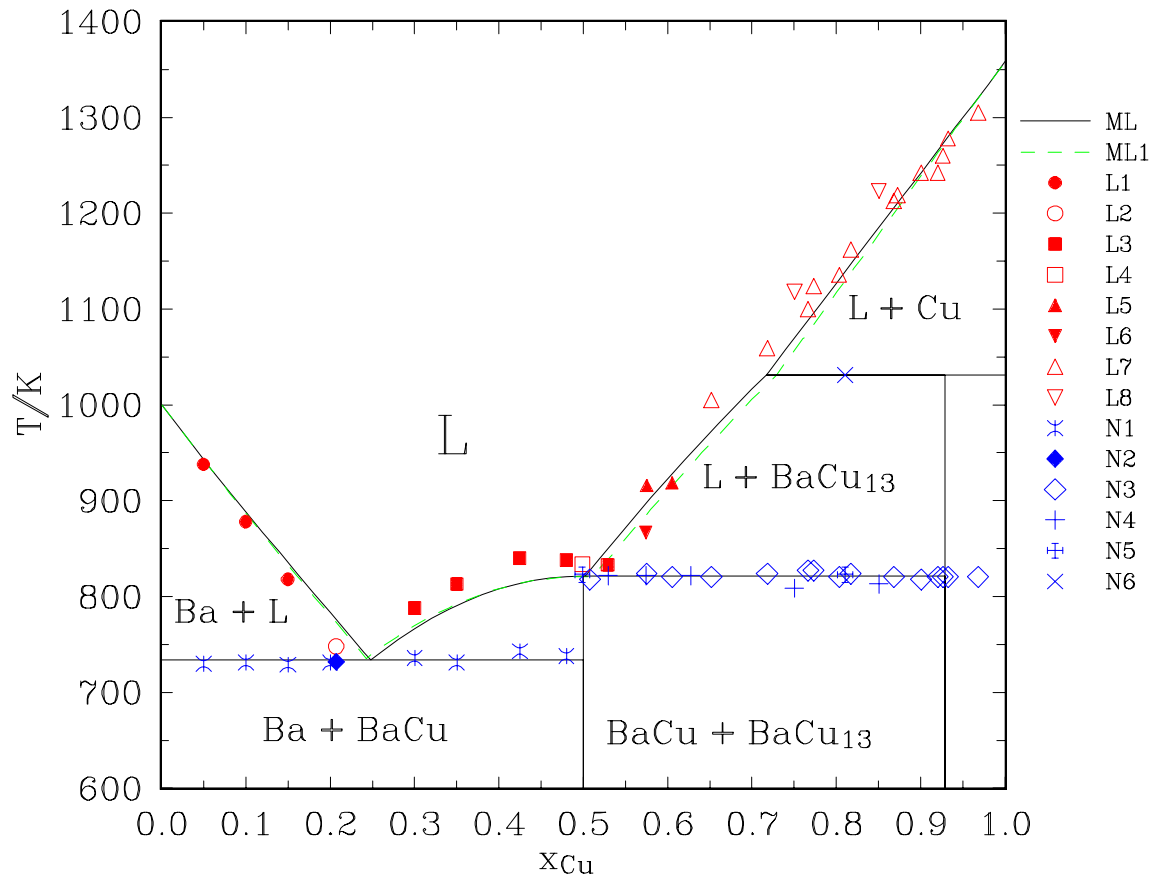


Fig. 3. The Ba-Cu phase diagram.

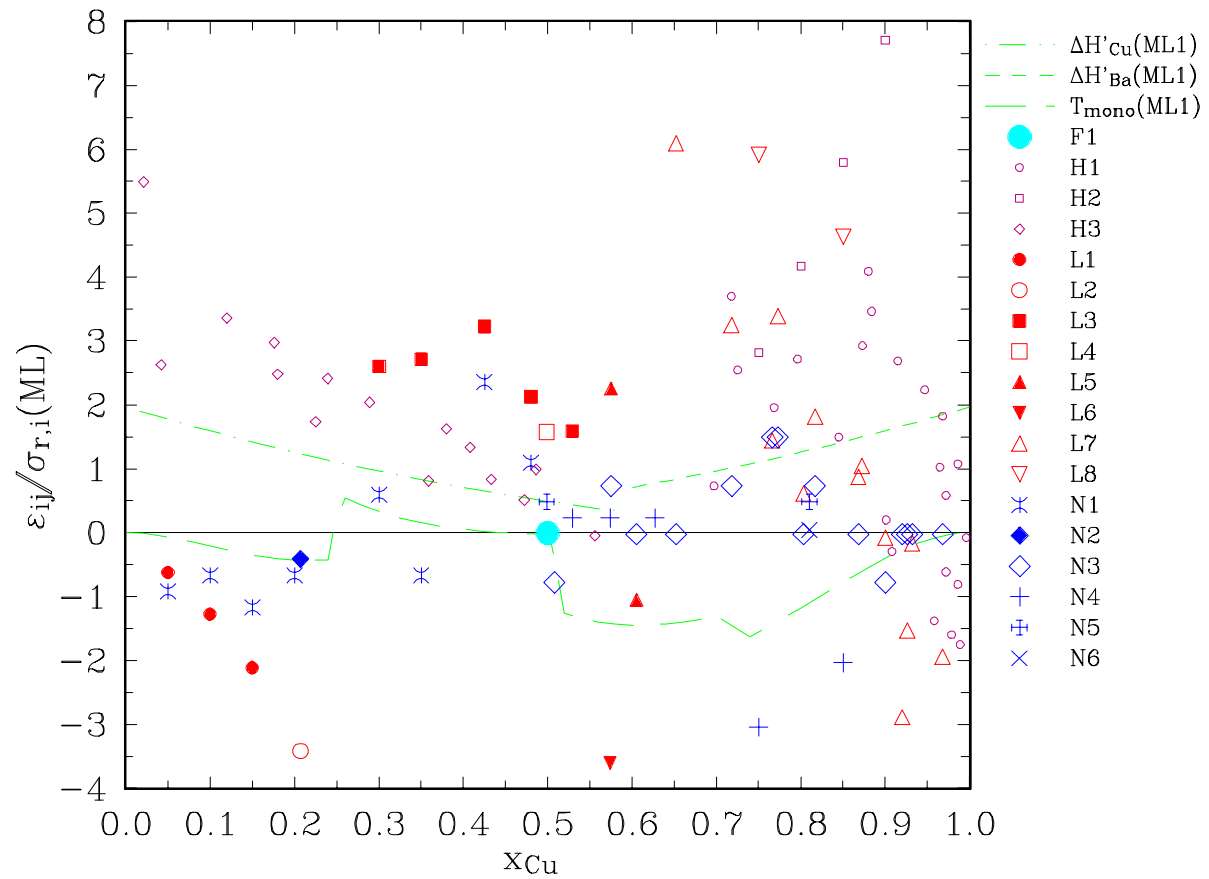


Fig. 4. The deviates of the experimental points in the Ba-Cu system as related to the solution ML.

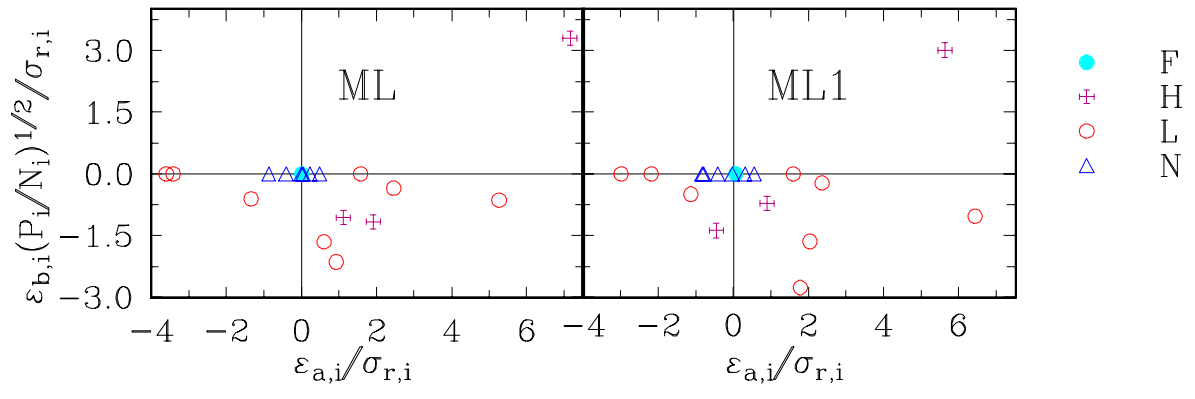


Fig. 5. Tilt systematic errors vs. shift errors in the Ba-Cu system.

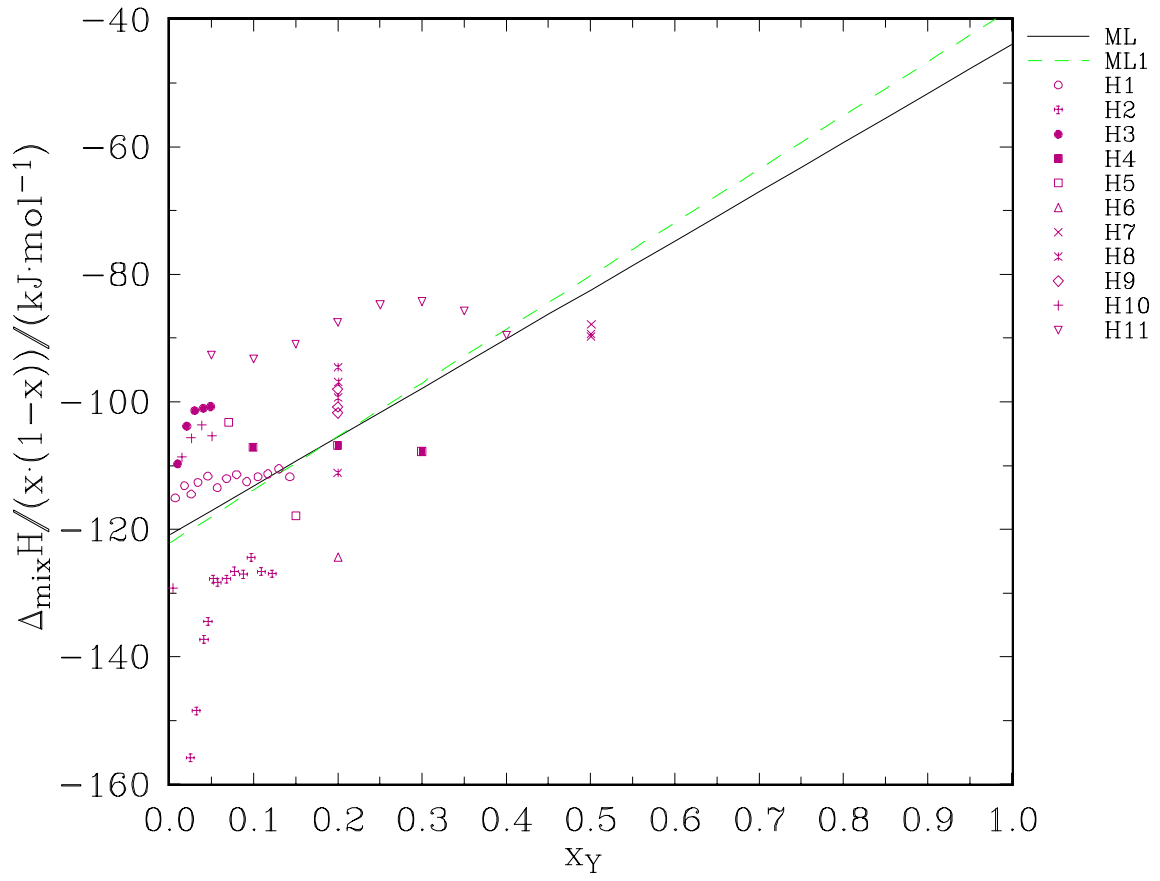


Fig. 6. The enthalpy of mixing of the Cu-Y liquid alloys.

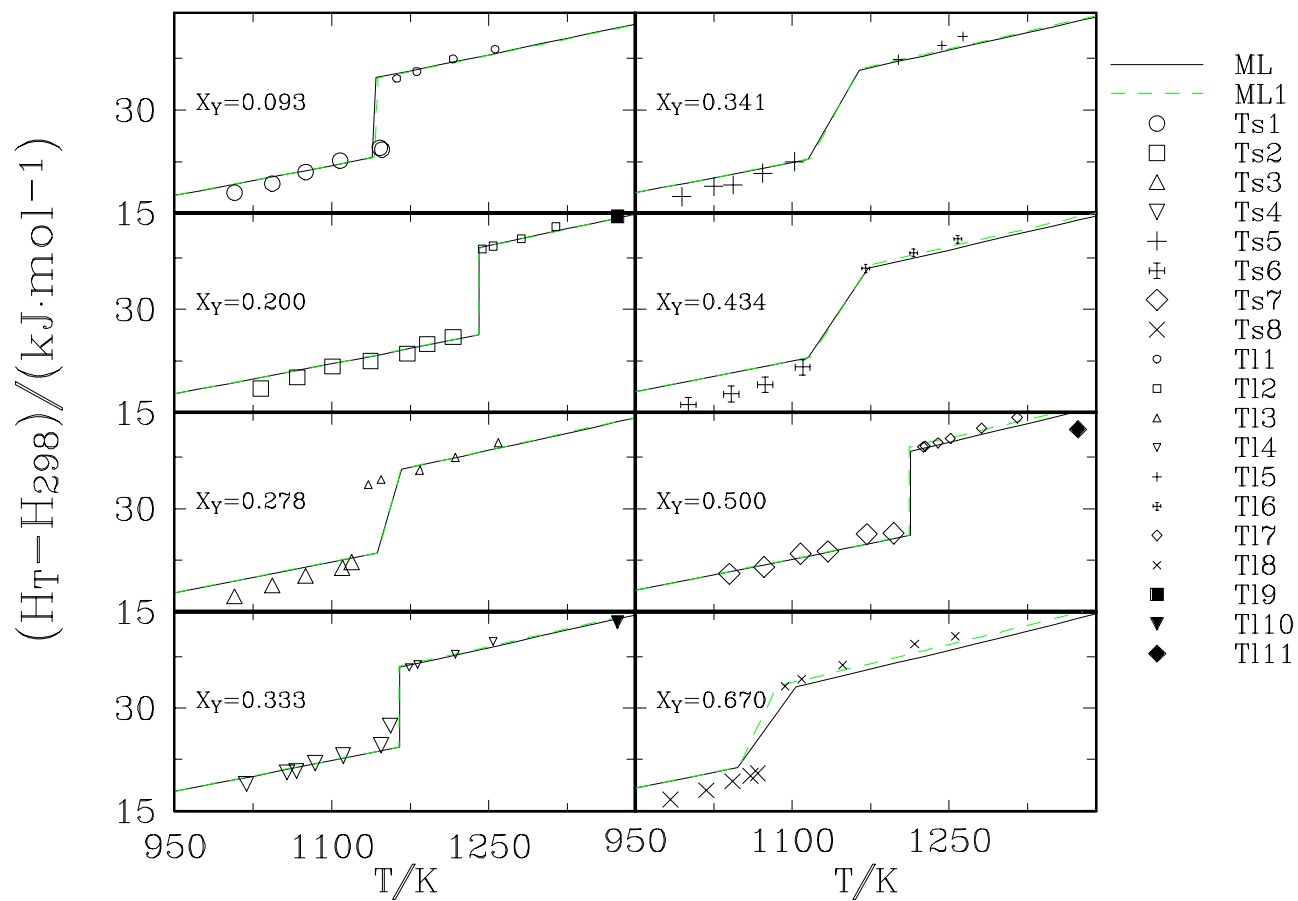


Fig. 7. The temperature dependence of enthalpy in the Cu-Y system.

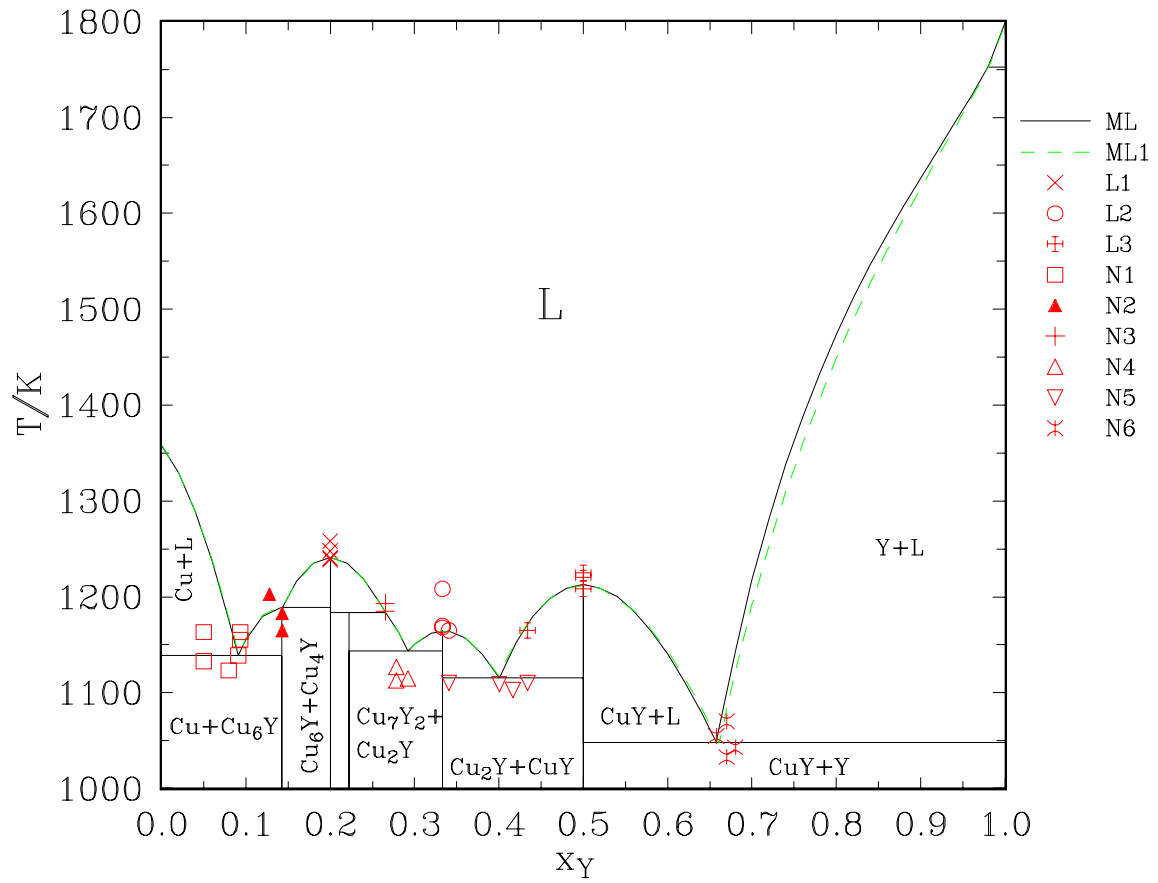


Fig. 8. The Cu-Y phase diagram.

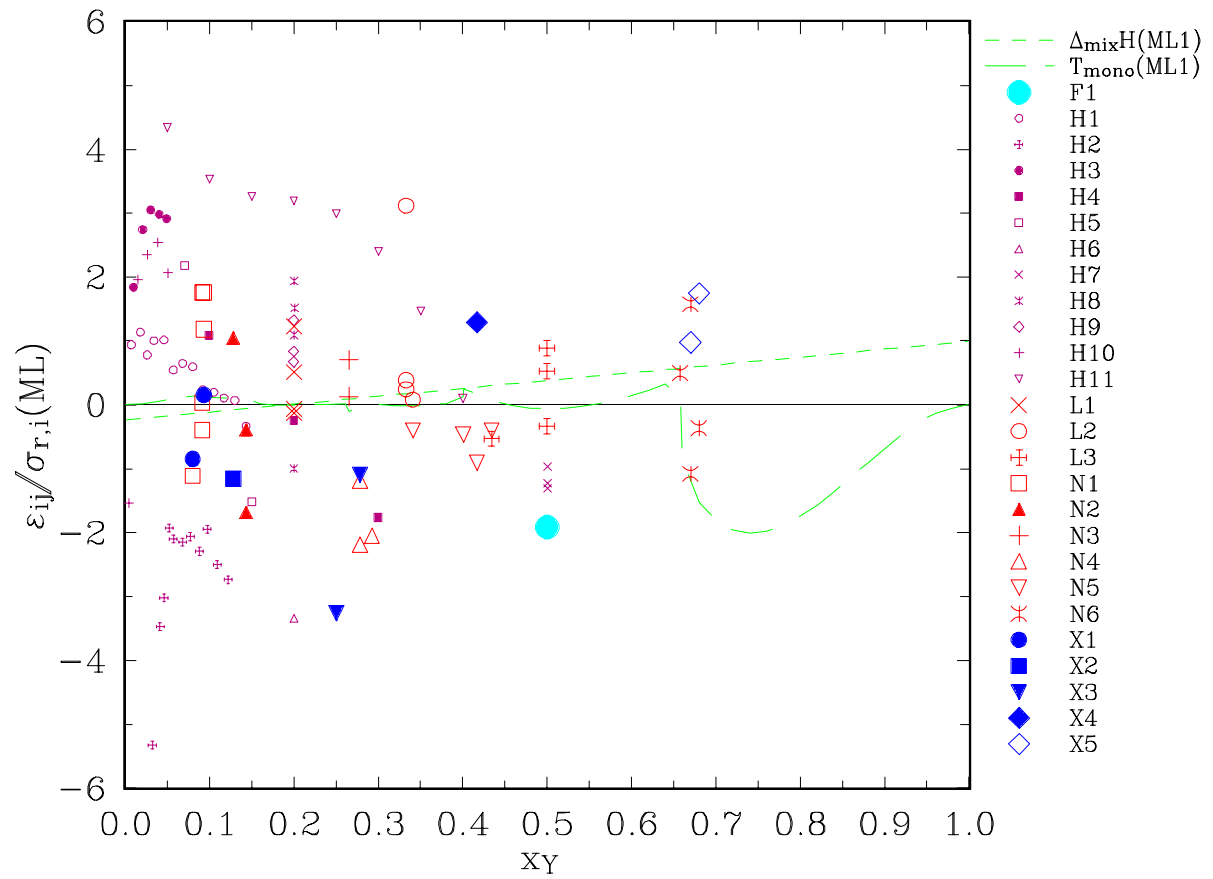


Fig. 9. The deviates of the experimental points in the Cu-Y system as related to the solution ML (part 1).

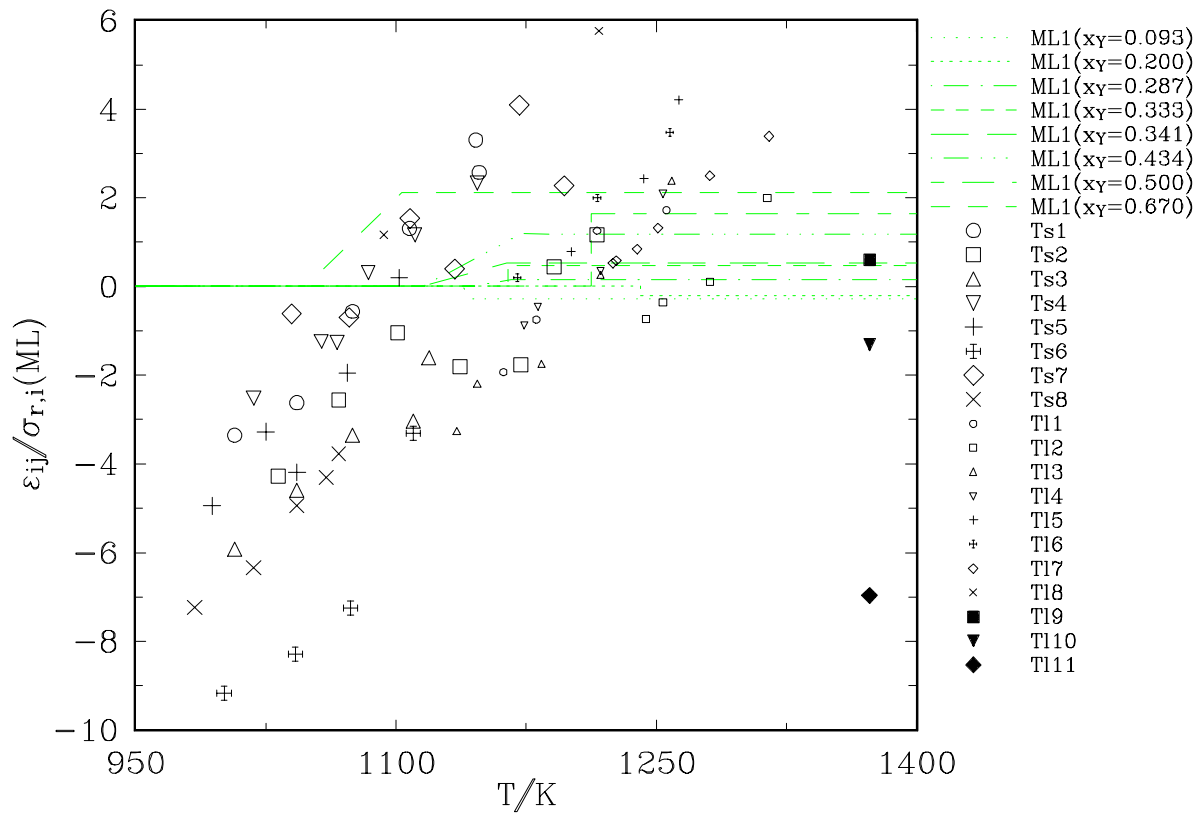


Fig. 10. The deviates of the experimental points in the Cu-Y system as related to the solution ML (part 2).

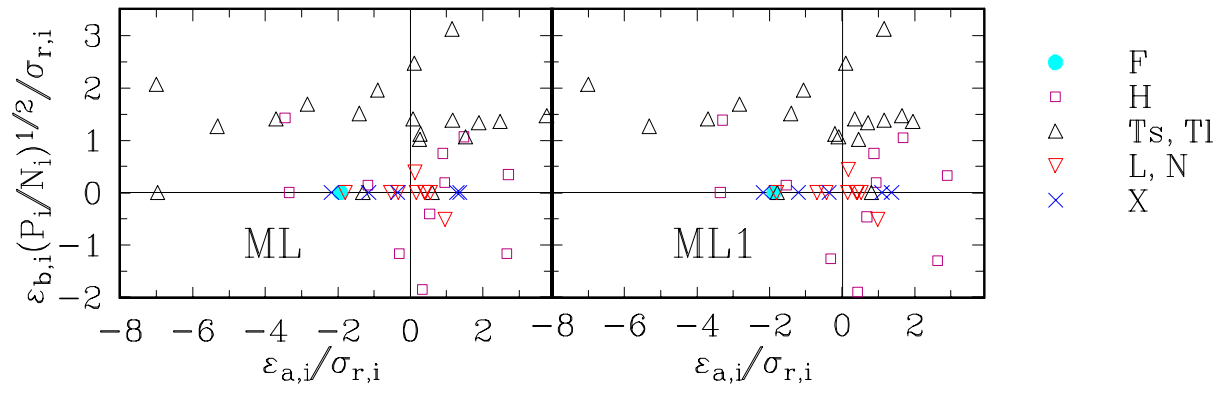


Fig. 11. Tilt systematic errors vs. shift errors in the Cu-Y system.