Free Software Development.
1. Fitting statistical regressions

L. Jäntschi

Faculty of Science and Engineering of Materials, Technical University of Cluj-Napoca, Romania,

lori@lori.east.utcluj.ro

Abstract

The present paper is focused on modeling of statistical data processing with applications in field of material science and engineering. A new method of data processing is presented and applied on a set of 10 Ni–Mn–Ga ferromagnetic ordered shape memory alloys that are known to exhibit phonon softening and soft mode condensation into a premartensitic phase prior to the martensitic transformation itself. The method allows to identify the correlations between data sets and to exploit them later in statistical study of alloys. An algorithm for computing data was implemented in preprocessed hypertext language (PHP) and a hypertext markup language interface for them was also realized and put onto comp.east.utcluj.ro educational web server and it is accessible via http protocol at the address http://comp.east.utcluj.ro/~lori/research/ regression/linear/v1.5/. The program running for the set of alloys allow to identify groups of alloys properties and give qualitative measure of correlations between properties. Surfaces of property dependencies are also fitted.

Keywords: Modeling, Analytical methods, Automat processing of data, Server side technologies.

1. Introduction

In field of statistical data processing it exist a large set of software to compute and fit the regressions, but few are free.

Even for free software, another problem it appear, operating system license and portability of the software. As example, to use well known Microsoft Excel software, you

must have Microsoft Windows and Microsoft Excel license and portability of Excel program is restricted to Windows platform. To import Excel files in another programs or platforms, conversion modules are needed, and conversion is not totally in most of the cases.

Platform independent and free software is a real alternative to this. First step to build totally free software is to install a free operating system. Unix-like operating systems are known to be free, but even here exist licensed software. In order to select a totally free Unix-like operating system, best offer come from BSD family.

The most secure and license check for installed software is NetBSD. The NetBSD detect so called "license agreement" and do not permit to install a software if the software contain unacceptable license agreement (different form free) and software can be installed only if the acceptance is explicitly stipulated by the user in configuration files of the system. Another advantage of NetBSD operating system is his huge portability under various hardware platforms from i386 family to Sun and Macintosh machines.

By another hand, most full featured operating system of BSD family is FreeBSD. One of the advantages of the FreeBSD operating system is his software portability. With adequate packages, under FreeBSD, can be executed DOS, Windows, Linux and Sun-OS programs. Another advantage of FreeBSD system comes from easiest to install and use it.

Once you have an operating system installed, the next step is to choose a proper programming language for software developing.

Here, some major questions must be answered. In terms of programming, portability of resulted program can be a problem. As example, if we are chose to implement the algorithm in Visual Basic, the execution of the program is restricted to Windows machines. If Perl is our choice, a Unix-based machine is necessary to run program.

Even if we chouse to implement the program in C language, we will have serious difficulties to compile the programs on machines running with different operating systems.

The complexity of program building is also a serious reason in language chousing. Is known that C based languages is poor in simplicity and necessity more time to expend for application building than other languages.

Other questions require an answer: We want a server based application or client based application? We want a server side application or a client side application?

As example, a client side application can have disadvantage of execution on client, and dependence of processing speed by power of client machine. If we prefer this variant, a java script or visual basic script is our programming language.

A server side application requires a web server installed. The area of web servers is also a large set, but few have multiplatform capability. If we want a full featured web server, Apache is our solution.

Under Apache, we have the possibility to execute programs already compiled in C, Fortran and Java, under Unix machines we can directly execute Perl programs, and, most important, under all operating system platforms we can execute PHP programs if we previously install php language and module binaries.

The advantage of PHP programs consist in his portability under most of operating system platforms and internal compilation feature that do not necessity the compilation "by hand" from the user. The disadvantage can be same internal compilation that consumes supplementary time in execution. But, this disadvantage can be partially eliminated through installing a PHP proxy, that store compiled programs and next execution of the unmodified program use this compiled binary. In terms of program developing PHP is easy to use, the language borrow syntax from C, Pascal, Basic and Perl, but do not borrow the complex declaration syntax from them. The pointer mechanism is absolute. Thus, a variable used as a string, can be exploited as an integer or real if the value represent a number such that. Class constructing is also available and PHP posses a strong library of database connectivity. Modular programming, recursivity and graphics are at home! Module loading of compiled programs in other languages and execution of binary programs is also available. System services such as mail service are easy to exploit in PHP scripts. A very easy mechanism to link PHP scripts to HTML scripts make PHP language to be one of the best. Shell executing commands make PHP a useful platform for system administration (PhpSysInfo, WebAdmin, PhpMyAdmin, PhpPgAdmin). As a conclusion, PHP is our choice!

PHP programs are putted on a web server data folder and executed by them using PHP module. The output of the PHP program is in HTML style and can be viewed by any web client (Microsoft Internet Explorer, Mozzilla, Opera, Netscape, Konkueror).

2. Theoretical considerations

Many statistical procedures for processing data are available.[1] Most of them offer a voluble set of possibilities and variants, but which one to consider them? That is not a easy question and the frequent answer is: that is choice of analyst.[2,3]

Data mining technology offer in this area of knowledge some answers, but not a complete answer.[4] By other hand, to interpret experiment results, data need to be well processed.[5] Structure investigations are frequently combined with statistical processing.[6] In

most of cases, best results are obtained with specific procedures in contrast to general numeric algorithms.[7,8]

Modeling of structure is benefit to property predictions.[9,10] Nonstandard statistical evaluation procedures then are helpful.[11]

The design of statistical processing program is depicted in fig. 1:
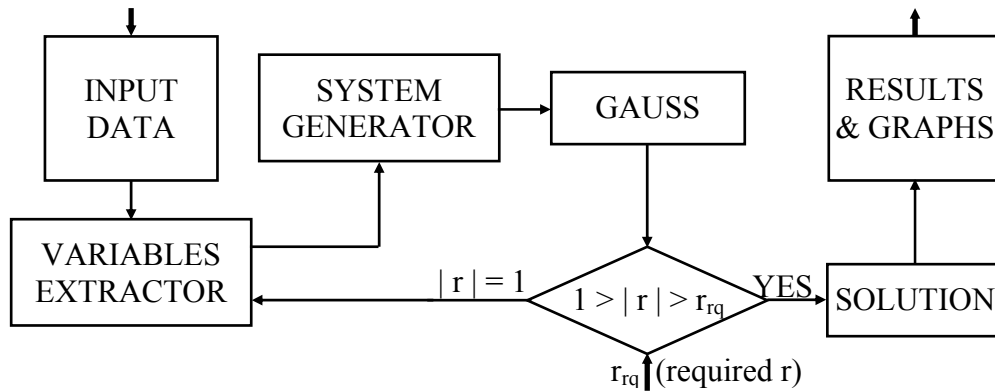

Fig.1 Program design scheme

The INPUT module read a text format data, process input data, split it into rows and columns. The rows represent the values of variables and columns represent the variables.

The INPUT module passes data in matrix form to the VARIABLES EXTRACTOR module. This module starts with pairs of columns from data matrix; continue with sets of three columns and so on, until entire set is browsed. For every set, pass it to the SYSTEM GENERATOR.

Another input in VARIABLES EXTRACTOR module comes from DECISION module. If the DECISION module detects a correlation of absolute coefficient 1 then it passes to the VARIABLES EXTRACTOR the first variable label from the regression equation and for the next extractions of variables the passed variable is marked as dependent and is not passed never again to the SYSTEM GENERATOR module.

The SYSTEM GENERATOR computes average means, needed by GAUSS module.

If name *n_rows* it assigned to number of rows, *n_cols* to number of columns, *data* to array of data, the output of module SYSTEM GENERATOR is computed by formulas:

$$M[i,j] = \frac{\sum\limits_{k=1}^{n\_rows} data[k][i] \cdot data[k][j]}{n\_rows} \; ; \; M[0,j] = \frac{\sum\limits_{k=1}^{n\_rows} data[k][i]}{n\_rows} \; , \; 1 \le i, j \le n\_cols \qquad (1)$$

Linear regression and PLS (partial least squares) are most used methods in statistical processing of data. Presented method (SYSTEM GENERATOR and GAUSS modules) uses them.

4

GAUSS module solves a linear system of equations in form:

$$\sum_{j=1}^{n\_cols} M[i,j] \cdot x_j = 1, \ 1 \le i \le n\_cols \tag{2}$$

If answer of algorithm solving is *undetermined system* and null variable is x_{n_cols} then GAUSS module solve determined system of *n_cols* order given by equation (3):

$$\sum_{j=1}^{n\_cols-1} M[i,j] \cdot x_j = M[0,j], \ 1 \le i < n\_cols \tag{3}$$

When system is solved a unique solution is found. Then, system extended matrix contain at column n_cols the coefficients of regression equation:

$$a_1 \cdot x_1 + ... + a_i \cdot x_i + ... + a_{n\_cols} \cdot x_{n\_cols} + a_{n\_cols+1} = 0 \tag{4}$$

where the coefficients a_{n_cols+1} and a_{n_cols+1} are resulted regression coefficients. Note that equation (4) is in implicit form; to obtain an explicit form is necessary to extract dependent variable from (4). The last coefficient is assigned to -1:

$$a_{n\_cols+1} = -1 \tag{5}$$

At the end of module SOLUTION it result an implicit linear regression equation between given variables through his values in columns (eq. 4). Equation 4 can be exploited to obtain explicit linear regression equations for each variable which has no null coefficient a_i:

$$\hat{x}_i = \left(\frac{a_1}{-a_i}\right) \cdot x_1 + ... + \left(\frac{a_{i-1}}{-a_i}\right) \cdot x_{i-1} + \left(\frac{a_{i+1}}{-a_i}\right) \cdot x_{i+1} + ... + \left(\frac{a_{n\_cols}}{-a_i}\right) \cdot x_{n\_cols} + \left(\frac{a_{n\_cols+1}}{-a_i}\right) \tag{6}$$

Sum of residues S_i can be now evaluated:

$$S_i = \left(\frac{a_{n\_cols+1}}{a_i} + \sum_{j=1}^{n\_cols} \frac{a_j}{a_i} \cdot x_i\right)^2 \tag{7}$$

To compare one equation to another, a calibration is required. Let to explicit this. If x_1 values (data[1] from input) are percents expressed in values from 0 to 100 and x_2 are premartensitic temperatures transformation expressed in K with values from 100 to 600, then also sum of residues are expressed square of same measurement units. To make independence of measurement unit and measure order, values S_i are divided with own sum of squares of variable measurements (M[i,i] from INPUT module, equation 1). Final equation, with substitution x_i = data[k,i], $1 \le k \le n\_rows$ and summing is:

$$Q_i = \sum_{k=1}^{n\_rows} \left(\frac{a_{n\_cols+1}}{a_i} + \sum_{j=1}^{n\_cols} \frac{a_j}{a_i} \cdot data[k,i]\right)^2 / M[i,i] \tag{8}$$

and express relative residues of variable x_i when variable x_i is assumed to be dependent of independent variables $x_1, ..., x_{i-1}, x_{i+1}, x_{cols}$. Note that the dependence and independence

statistical concept is hard to prove in practical situations, but will see later, can be decelerated. For a good correlation, Q_i should be smallest possible value. The value of Q_i is computed in SOLUTION module for every equation.

Another quantitative measure for a good correlation is correlation coefficient between measured data x_i and estimated values $\hat{x}_i$ from equation (6). Assuming that M is mean operator, for r is given by:

$$r(x_i, \hat{x}_i) = \frac{M(x_i \cdot \hat{x}_i) - M(x_i) \cdot M(\hat{x}_i)}{\left(M(x_i \cdot x_i) - M(x_i) \cdot M(x_i)\right)^{1/2} \cdot \left(M(\hat{x}_i \cdot \hat{x}_i) - M(\hat{x}_i) \cdot M(\hat{x}_i)\right)^{1/2}} \tag{9}$$

The absolute value of r must be high for a good correlation. The value of r is computed in DECISION module for implicit equation and in SOLUTION module for every explicit equation. More additionally tests are also available in other programs such as Microsoft Excel or Statsoft Statistica.

3. Implementation

A graphical interface was built in PHP with a TEXTAREA for input data and an INPUT SUBMIT button for submitting data to the server. The server is a Free BSD Unix based server (5.0 DP1 software version) with an Apache web server (1.3.26 software version) running on. The server is hosted in educational network of Technical University of Cluj-Napoca with address 193.226.7.211 and name comp.east.utcluj.ro. The PHP language was compiled with GDI (graphical device interface) and MySQL (database server) support and the PHP software version is 4.2.3. The MySQL database server is also installed and running on and his software version is 3.23.52. The input interface is presented in fig. 2:
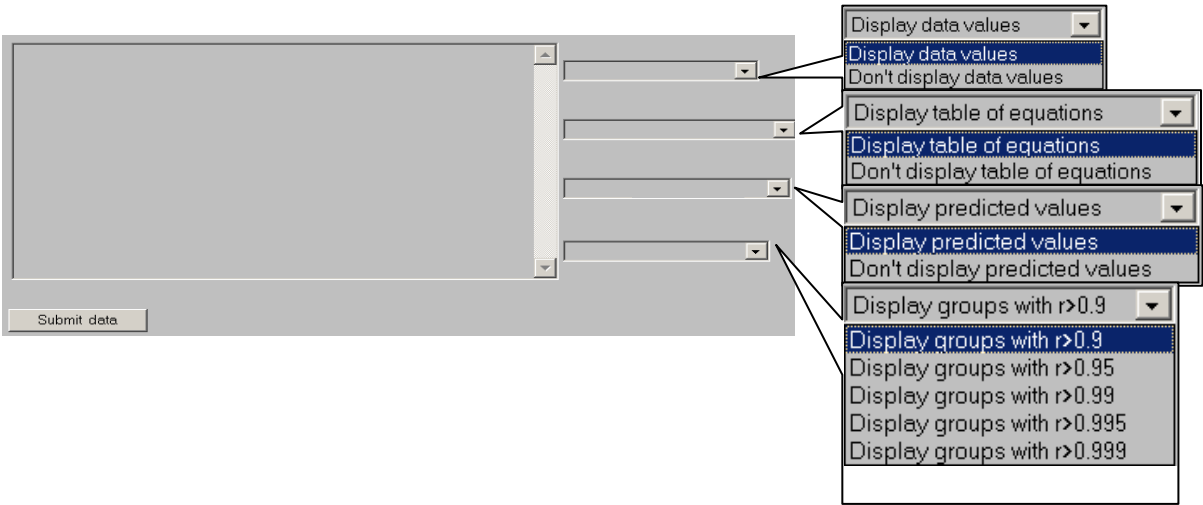

Fig.2 The output of http://193.226.7.211/~lori/research/regression/linear/v1.5/index.php

The output of the RESULTS & GRAPHS module is passed depending on user equation selection to an another PHP program that make graphical representation of regression:[12]
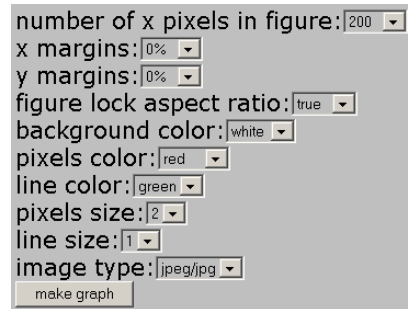

Fig.3 The output of http://193.226.7.211/~lori/research/regression/graphics/lin_plot/do.php

The program build have 21 subroutines and a main program, specified below:

- function af_ec($n,$coef,&$t) // display an equation;

- function af_mt($titlu,&$tabel,$n_r,$n_c) //display any matrix with a title;

- function af_rez($n_r,$n_c,&$d,&$m,&$c,&$t,$n_o,$pr) //list a table with best equations founded;

- function ch_ln($l1,$l2,&$cc,$r) // Gauss linear algebra method, change two lines in system extended matrix;

- function cnk($k,$n_r,$n_c,&$data,&$tab,$pr,&$dep,&$inv) //make recursive all possible combination c(n,k);

- function data_copy($n_r,$n_c,&$d,&$t,&$d_t,&$n_t) //extract a subset of data from entire set;

- function do_means(&$data,&$mean,$n_rows,$n_cols) //make all (xi, xi*xj) means;

- function ec_by($n,&$coef,$by,&$c_by) //calculate coefficients for explicit equation from implicit equation coefficients;

- function ec_val($n,&$valori,&$coef) //compute value of implicit equation in given point

- function estimare($n_r,$n_c,&$d,&$c,$x,&$x_est) //compute value for explicit equation in given point for given dependent variable;

- function im_ln($nr,$rw,&$cc,$r) //Gauss linear algebra method, make a unitary element into system extended matrix;

- function mx_rw($cl,&$cc,$r) //Gauss linear algebra method, find the best line for zeroes in system extended matrix;

- function n_to_s($nr) //format and display a real number;

- function r_stat($n,$k,&$d1,&$d2) //compute correlation coefficient r;

- function rd_gs($cc,$r,&$cf) //Gauss linear algebra iterative algorithm;

- function reg_lin_1($n_r,$n_c,&$d,&$t,$pr) //make linear regression if possible; return answer;

- function res(&$t,$n) //reset counter for recursive c(n,k);

- function sum_r($n_rows,$n_cols,&$data,&$coef,$cor) //calculate sum of residues;

- function ze_pd(&$cc,$r) //Gauss linear algebra method, make supdiag. zeroes in system matrix;

- function ze_sd($e,&$cc,$r) //Gauss linear algebra method, make subdiag. zeroes in system matrix;

- main program //input data and requested minimal correlation coefficient and display founded equations;

4. Results and discussion

The presented model make data preprocessing to a set of 10 Ni–Mn–Ga ferromagnetic ordered shape memory alloys that are known to exhibit phonon softening and soft mode condensation into a premartensitic phase prior to the martensitic transformation itself and is a extension added to the model presented in book [13]. The properties are described in table 2.[14] Previous versions of this program were reported in [15-18].

Table 1. Processed Data

Column	Property	Measurement unit
1	Alloy State (Poly- or Single-crystalline alloy)	1, -1 (PC, SC respectively)
2	e/a	Electron/atom ratio
3	Concentration of Ni	%
4	Concentration of Mn	%
5	Concentration of Ga	%
*6	T_1 (rows 1-7), T_M' (rows 8-10)	K
7	T_M, premartensitic temperature transformation	K

*temperatures: T_1= martensitic transformation; T_M'=intermartensitic transformation in Group III alloys.

The input data can be copied and pasted into TEXTAREA from figure 2. By pressing submit button the input data are processed. By chousing to display data values, the input data are also displayed. In table 2 are presented the input data, such as are showed by PHP program:

Table 2. Input data values (output by PHP program)

0	1	2	3	4	5	6	7
1	1	7.35	49.6	21.9	28.5	4.2	178
2	1	7.36	47.6	25.7	26.7	4.2	152
3	-1	7.45	49.7	24.3	26.0	183	218
4	1	7.50	50.9	23.4	25.7	113	224
5	-1	7.51	49.2	26.6	24.2	184	240
6	1	7.56	47.7	30.5	21.8	227	240
7	1	7.57	51.1	24.9	24.0	197	248
8	-1	7.78	53.1	26.6	20.3	417	379
9	-1	7.83	51.2	31.1	17.7	443	415
10	-1	7.91	59.0	19.4	21.6	633	517

Program computes and output the regression equations. With an $r_{rq} = 0.9$ the program found over 60 different implicit equations of linear regression with $r > r_{rq}$, almost impossible to obtain by hand or in some program with statistics kernel. If value of r_{rq} is increased to $r_{rq} = 0.99$, number of implicit equation founded is reduced at 29. For $r_{rq} = 0.999$ number of implicit equation founded is 12.

That is a large set! If we are interested to study dependence between two variables form set, then we select the proper table from output of the program. Best result is displayed in table 4, and it correlates the temperatures T_1 and T_M:

Table 4. Linear regression between martensitic, premartensitic and intermartensitic temperatures

x_0	x_1	x_2	x_3	x_4	x_5	x_6	Equation	Residue	Correlation
0	0	0	0	0	1	1	$+x_5*1.00-x_6*1.80=-2.73*10^2$	0.37	0.98369
0	0	0	0	0	1	1	$-x_5*0.55+x_6*1.00=+1.50*10^2$	0.21	0.98369
0	0	0	0	0	1	1	$+x_5*3.66*10^{-3}-x_6*6.62*10^{-3}=-1.00$	0.42	0.98369

If we are looking for totally dependent variables (and here exists, sum of concentrations is 100%), the program finds it and also eliminate one of them from set. In table 5 is showed program response for correlating variables x_2, x_3 and x_4 (dependent variable: x_2):

Table 5. Founded dependent variable in group of concentrations of Ni(x_2), Mn(x_3) and Ga(x_4)

x_0	x_1	x_2	x_3	x_4	x_5	x_6	Equation	Residue	Correlation
0	0	1	1	1	0	0	$+x_2*1.00+x_3*1.00+x_4*1.00=+1.00*10^2$	0.00	1.00000
0	0	1	1	1	0	0	$+x_2*1.00+x_3*1.00+x_4*1.00=+1.00*10^2$	0.00	1.00000
0	0	1	1	1	0	0	$+x_2*1.00+x_3*1.00+x_4*1.00=+1.00*10^2$	0.00	1.00000
0	0	1	1	1	0	0	$-x_2*1.00*10^{-2}-x_3*1.00*10^{-2}-x_4*1.00*10^{-2}=-1.00$	0.00	1.00000

If we are looking for dependences between e/a and concentrations, simply select the founded equations from program output. In table 6 is showed the dependence of *e/a* by concentration of Ni and Mn:

Table 6. Dependence of *e/a* by Ni(x_2) and Mn(x_3) expressed by explicit and implicit equations

x_0	x_1	x_2	x_3	x_4	x_5	x_6	Equation	Residue	Correlation
0	1	1	1	0	0	0	$+x_1*1.00-x_2*7.02*10^{-2}-x_3*3.99*10^{-2}=+2.98$	$7.35*10^{-4}$	0.99995
0	1	1	1	0	0	0	$-x_1*1.42*10^1+x_2*1.00+x_3*0.56=-4.25*10^1$	$1.55*10^{-3}$	0.99997
0	1	1	1	0	0	0	$-x_1*2.50*10^1+x_2*1.75+x_3*1.00=-7.47*10^1$	$5.43*10^{-3}$	0.99992
0	1	1	1	0	0	0	$-x_1*0.33+x_2*2.34*10^{-2}+x_3*1.33*10^{-2}=-1.00$	$1.86*10^{-3}$	0.99995

If we are looking to express one of the temperatures depending by concentrations, then the following equations are useful (table 7).

Table 7. Dependence of T_M (premartensitic temperature) by Mn(x_3) and Ga(x_4) concentrations

x_0	x_1	x_2	x_3	x_4	x_5	x_6	Equation	Residue	Correlation
0	0	0	1	1	0	1	$+x_3*0.56+x_4*1.00+x_6*2.36*10^{-2}=+4.45*10^1$	$4.88*10^{-2}$	0.99300
0	0	0	1	1	0	1	$+x_3*2.37*10^1+x_4*4.23*10^1+x_6*1.00=+1.88*10^3$	0.16	0.99012

The equations that contain maximum of independent terms (without one concentration) is given at the end of output file of the program (table 8). Note that the variable x_2 are missing from regression equation (are dependent to the x_3 and x_4 variables).

Table 8. Most comprehensive multi-linear dependence in data set

| x_0	x_1	x_2	x_3	x_4	x_5	x_6	Equation	Residue	Correlation
1	1	0	1	1	1	1	$-x_0*6.77*10^{-4}+x_1*1.00+x_3*2.79*10^{-2}$ $+x_4*6.61*10^{-2}+x_5*5.27*10^{-6}-x_6*1.09*10^{-4}=+9.82$	$4.12*10^{-4}$	0.99999
1	1	0	1	1	1	1	$-x_0*2.42*10^{-2}+x_1*3.57*10^{1}+x_3*1.00$ $+x_4*2.36+x_5*1.88*10^{-4}-x_6*3.92*10^{-3}=+3.51*10^{2}$	$4.36*10^{-3}$	0.99995
1	1	0	1	1	1	1	$-x_0*1.02*10^{-2}+x_1*1.51*10^{1}+x_3*0.42$ $+x_4*1.00+x_5*7.97*10^{-5}-x_6*1.65*10^{-3}=+1.48*10^{2}$	$1.98*10^{-3}$	0.99999
1	1	0	1	1	1	1	$+x_0*6.17-x_1*9.11*10^{3}-x_3*2.54*10^{2}$ $-x_4*6.03*10^{2}-x_5*4.80*10^{-2}+x_6*1.00=-8.95*10^{4}$	$9.44*10^{-2}$	0.99668
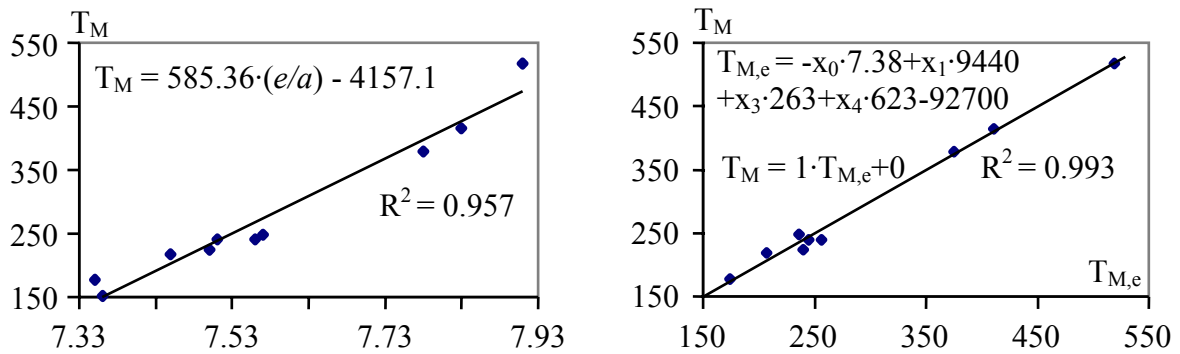
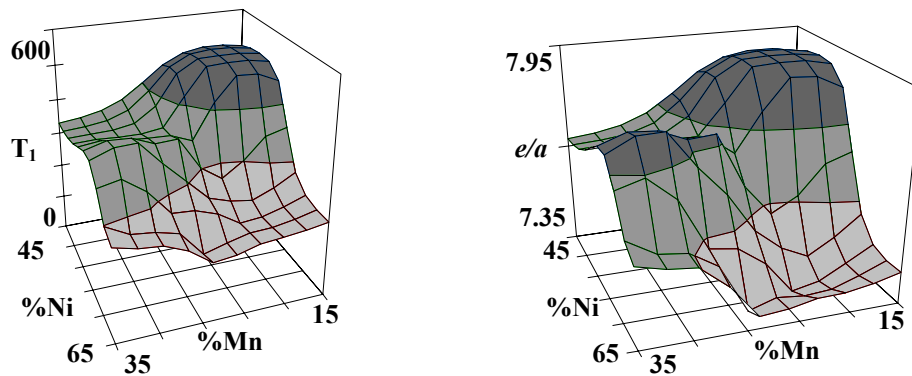Fig. 3. (a) Regression between T_M and e/a; (b) T_M and (Alloy state, *e/a*, %Mn, %Ga)

Fig. 4. (a) surface plot of T_1 and (b) e/a by composition (%Ni,%Mn)

In figure 3 are plotted some selected dependences from data set. For fitting, the suggested selections from PHP program were used. Figure 3a show a mono-variable dependence between T_M and *e/a*, and figure 3b show a multi-linear variable dependence between T_M and alloy state (codified by -1 and 1) electron/atom ratio and composition (%Mn and %Ga).

Figures 4b and 4c show surface dependences of T_1 and respectively *e/a* of composition (%Ni, %Mn).

5. Conclusions

Considering the advantages of implemented software technology (machine and operating system portability, graphical interface and database connectivity features, easiest of programs developing, free type license agreement, http capability) the programming language and the program itself is the one of the best choice now available.

Looking to the output sums of residues from tables, is easy to observe now that the properties: type of alloy, and his martensitic, intermartensitic and premartenistic temperatures are interrelated together; these properties have the same order of sum residues in global equation, that is also expected conclusion. Very small same order of sum residues for concentrations suggest a strong interrelation between them, that is also true, because %Ni+%Mn+%Ga = 100. This conclusion lead to consider the 3D plots fitted in figure 3 (b and c) of electron/atom ratio and T_1 temperature dependencies by concentration (%Ni,%Mn). The figure 3a prove good correlation between T_1 and e/a.

6. Acknowledgments

References

1. H. Nascu, Lorentz Jäntschi, T. Hodisan, C. Cimpoiu, G. Câmpan, Some Applications of Statistics in Analytical Chemistry, Reviews in Analytical Chemistry, Freud Publishing House, XVIII(6), 1999, 409-456.

2. Lorentz Jäntschi, M. Unguresan, MathLab. Statistical Applications, International Conference on Quality Control, Automation and Robotics, May 23-25 2002, Cluj-Napoca, Romania, work published in volume "AQTR Theta 13" (ISBN 973-9357-10-3) at p. 194-199.

3. Lorentz Jäntschi, H. Nascu, Numerical Description of Titration (MathCad application), International Conference on Quality Control, Automation and Robotics, May 23-25 2002, Cluj-Napoca, Romania, work published in volume "AQTR Theta 13" (ISBN 973-9357-10-3) at p. 259-262.

4. Lorentz Jäntschi, M. Unguresan, Parallel processing of data. C++ Applications, Oradea University Annals, Mathematics Fascicle, 2001, VIII, 105-112, ISSN 1221-1265.

5. L. Marta, I. G. Deac, I. Fruth, M. Zaharescu, Lorentz Jäntschi, Superconducting Materials: Comparision Between Coprecipitation and Solid State Phase Preparation, International Conference on Science of Materials, August 20-22, 2000, Cluj-Napoca, Romania, work published in volume 2 (ISBN 973-686-058-2) at p. 627-632.

6. L. Marta, I. G. Deac, I. Fruth, M. Zaharescu, Lorentz Jäntschi, Superconducting Materials at High Temperature: Thermal Treatment and Pb Addtition, International Conference on Science of Materials, August 20-22, 2000, Cluj-Napoca, Romania, work published in volume 2 (ISBN 973-686-058-2) at p. 633-636.

7. C. Cimpoiu, Lorentz Jäntschi, T. Hodisan, A New Method for Mobile Phase Optimization in High-Performance Thin-Layer Chromatography (HPTLC), Journal of Planar Chromatography - Modern TLC, Springer, 11(May/June), 1998, 191-194, ISSN 0933-4173.

8. C. Cimpoiu, Lorentz Jäntschi, T. Hodisan, A New Mathematical Model for the Optimization of the Mobile Phase Composition in HPTLC and the Comparision with Other Models, Journal of Liquid Chromatography and Related Technologies, Dekker, 22(10), 1999, 1429-1441, ISSN 1082-6076.

9. M. Diudea, Lorentz Jäntschi, A. Graovac, Cluj Fragmental Indices, Math/Chem/Comp'98 (The Thirteenth Dubrovnik International Course & Conference on the Interfaces among Mathematics, Chemistry and Computer Sciences), June 22-27, 1998.

10. Lorentz Jäntschi, G. Katona, M. Diudea, Modeling Molecular Properties by Cluj Indices, Commun. Math. Comput. Chem. (MATCH), 2000, 41, 151-188, ISSN 0340-6253.

11. C. Sârbu, Lorentz Jäntschi, Statistic Validation and Evaluation of Analytical Methods by Comparative Studies. I. Validation of Analytical Methods using Regression Analysis (in Romanian), Revista de Chimie, Bucuresti, 49(1), 1998, 19-24, ISSN 0034-7752.

12. Lorentz Jäntschi, Free Software Development. 2. Statistical regressions plotting with GDI interface on Apache web server and PHP module, Leonardo Electronic Journal of Practices and Technologies, Cluj-Napoca, in press.

13. Mircea Diudea, Ivan Gutman, Lorentz Jäntschi, Molecular Topology, Nova Science, Huntington, New York, 2001, 332 p., ISBN 1-56072-957-0.

14. V.A. Chernenko, J. Pons, C. Seguý′, E. Cesari, Premartensitic phenomena and other phase transformations in Ni–Mn–Ga alloys studied by dynamical mechanical analysis and electron diffraction, Acta Materialia, 50, 53-60, 2002, Elsevier, USA.

15. Lorentz Jäntschi, Real Time Property Investigation in Sets of Alloys, Second International Conference on Advanced Materials and Structures, 19-21 September 2002, Timişoara, Romania, p.189-194, ISBN 973-8391-50-4.

16. Lorentz Jäntschi, Automat Server Side Processing of Statistical Data, UNITECH'02 International Conference, 21-22 November 2002, Gabrovo, Bulgaria.

17. Lorentz Jäntschi, Property Investigations with an Automat Correlation Routine and Applications for a Set of Alloys, MATEHN'02 International Conference, 12-14 September 2002, Cluj-Napoca, Romania.

18. Lorentz Jäntschi, Web Server Routine for Materials Properties Correlation Study. Theory and Applications for a Set of Alloys, Acta Materialia, Elsevier, USA, in press.