STUDIA UNIVERSITATIS BABES-BOLYAI, CHEMIA, LIV, 1, 2009

Dedicated to Professor Liviu Literat, at his 80th anniversary

CASE STUDY OF STRUCTURAL SAFETY BASED ON ARTIFICIAL INTELLIGENCE

CĂLIN IOAN ANGHEL[a], IOAN LAZĂR

ABSTRACT. The paper provides a new insight into safety assessment using artificial intelligence methods. In safety index assessment, beside of the traditional *FORM/SORM* methods, a *minimax probability machine* approach was implemented. The core of the procedure based on a binary classification approach was developed according to a novel type of support vector machine in a *minimax* manner. The procedure, involving a link between artificial intelligence and structural reliability methods was developed as a user-friendly computer program in MATLAB language. For simplicity, only safety index approaches was presented. A numerical example replicating some previous related works reveals the opportunity of this approach in safety analysis.

Keywords: *risk assessment, safety index, limit state function, support vector machine, minimax approach, pressure vessel.*

INTRODUCTION

To avoid failure or major incidents technological systems must to be safe. Basically, safety can be thought as a state, a perceived state or a quality. Approaches of system safety presume planned, disciplined and systematic assessments to prevent or reduce accidents throughout the life cycle of a technological system. Structural safety related to process equipment is a part of the global safety of technological system. Because safety is a state, it cannot be quantified directly. Rather, assessing and controlling risk assure this state. The most usually definitions for risk in engineering applications are [1]:

- a combination of the likelihood and the consequences of a future event;
- the failure probabilities for a number of different scenarios;
- the product between the probability of occurrence and the quantified consequence of a future event.

[a] *Universitatea Babeş-Bolyai Cluj-Napoca, Facultatea de Chimie şi Inginerie Chimică, Str. M. Kogălniceanu Nr. 1, RO-400084 Cluj-Napoca, canghel@chem.ubbcluj.ro*

The equipment for process industries, are hardly affected by: fluctuations and variations in service loading and operating conditions, scatters in material properties and manufacturing process, uncertainties regarding the analytical models, continuous chemical corrosion and so on. Many engineering decisions concerning the safety of technological equipment for process industries are usually deterministic. These deterministic models provide a difficulty in handling variations and uncertainties in service conditions or variability concerning the main variables. To avoid these drawbacks, probabilistic methods became well-established tools in the risk assessment. The structural reliability procedures [2, 3] based on safety index and failure probability are still of greater importance in engineering applications. There are various methods to compute the safety index. Sampling-based methods (Monte Carlo Simulation, Latin Hypercube Sampling, etc.), analytical approximation methods like the Response Surface, the First-and Second-Order Reliability (*FORM/ SORM*) methods, advanced second moment method (ASM) or asymptotic techniques are among the most common methods used to estimate safety index or the failure probability. In general all these methods are time consuming, have limitations and an approximate nature. Despite of the recent important advances these traditional methods cannot fulfil in a satisfactory manner all the demands. The variability and uncertainty included in parameters of the model can determine major changes in the safety of the system. Within this context over the past decade intense effort has been devoted to bring ideas from the artificial intelligence field into engineering problems related to structural safety [4-7].

The aim of this paper is to present a procedure suitable for engineers in the stage of structural risk analysis. For simplicity it is based on the previous second mentioned definition of the risk. To compute the safety index the procedure focuses on the reliability concept developed in an artificial intelligence manner based on the support vector machine (SVM-a primarily two-class classifier) in a *minimax* approach. Some advantages of the support vector machine in *minimax* approach are: (a) provides an explicit direct upper bound on the probability of misclassification of new data, without making any specific distribution assumptions and (b) obtains explicit decision boundaries based on a global information of available data. The support vector machine and *minimax* approach, named *minimax probability machine classification*, has become an active research topic [8-11]. The problem of finding the safety index will be reported as a binary classification problem according to the traditional reliability methods better known by their acronyms *FORM/SORM*. The binary classification is applied to samples obtained by simulating a performance function or to values obtained by experimental analyses. Basically in structural safety methods the boundary between the two domains *S*-safe and *F*-unsafe is called the failure limit hyper surface and corresponds to a limit state function. This limit state function is tantamount to a decision function that classifies the samples. In principle the problem becomes a binary classification one and the decision function becomes a function whose sign represents the class assigned to any sampled point. The proposed procedure named *minimax decision procedure*, involving a link between artificial intelligence and reliability methods was developed as a user-friendly computer program in MATLAB language. The main parts of the paper are: (1) introduction, (2) theoretical overview for safety index assessment in *minimax* approach, (3) numerical examples implemented in the *MATLAB* package and (4) conclusions. Numerical example is reported to the risk of failure for a thin-wall pressure vessel during its serviceable life. It is associated with structure's strength, corrosion effects and the presumption of miss regulated of heating/cooling closed loop system.

Safety Index Assessment in a Minimax Approach

Safety and safety index assessment of systems/structures are well established. Basic knowledge exists and many papers are available in the literature [2,3]. The paper focuses only on the methods better known by their acronyms *FORM/SORM*. Instead of these traditional methods, the safety index approach will be obtained based on a decisional procedure using a binary classification procedure. Because basic fundamentals and principles were presented elsewhere [6,7] only the main principle of implemented procedure will be presented. A starting point is the establishment of the performance or response function, which gives the relation between the inputs of the system and the chosen performance. This performance function depends on the set of governing input parameters representing vectors of all random variables, $\mathbf{x} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_d}) \in \boldsymbol{R}^d$. If *PF* represents performance function and PF_{CR} represents a critical value of particular interest of this performance function, the limit state function (*LSF*) is the locus of points of the performance functions that can be defined as:

$$LSF = PF_{CR} - PF(\mathbf{x}) = 0 \quad \text{or} \quad LSF = PF_{CR}/PF(\mathbf{x}) = 1\,. \tag{1}$$

The traditional *FORM/SORM* algorithms work into an equivalent standard normal $\boldsymbol{U}^d$ space to estimate the safety index and the failure probability. The transformation of the basic random variables $\mathbf{x} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_d}) \in \mathbf{R}^{\mathbf{d}}$ into equivalent independent and uncorrelated variates $\mathbf{u} = (\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_d}) \in \boldsymbol{U}^{\boldsymbol{d}}$ should be made using the Rackwitz-Fiessler or Rosenblatt-Nataf or other similar isoprobabilistic transformations [2]. Therefore the limit state function is also transformed into equivalent standard normal $\boldsymbol{U}^d$ space by:

$$LSF(\boldsymbol{u}) = PF_{CR} - PF(\boldsymbol{u}) = 0 \quad \text{or} \quad LSF(\boldsymbol{u}) = PF_{CR}/\,PF(\boldsymbol{u}) = 1 \tag{2}$$

$$\boldsymbol{F} = \left\{ \boldsymbol{u} \in \boldsymbol{U}^{\boldsymbol{d}} \,\middle|\, \boldsymbol{LSF}(\boldsymbol{u}) \le 0 \quad \text{or} \quad \boldsymbol{LSF}(\boldsymbol{u}) \le 1 \right\} \tag{3}$$

In practice the state of any system can be divided into two regions: (*F*)-the failure domain, a region where combinations of system parameters lead to an unacceptable or unsafe system response-defined by the existence of a performance function whose non-positive values establish the non-reliability domain (eqn.3) and (*S*)-the safety domain, a region where system response is acceptable, defined by the existence of a performance function whose positive values establish the safety domain:

$$S = \left\{ \mathbf{u} \in \mathbf{U}^{d} \middle| LSF(u) > 0 \ \text{ or } \ LSF(u) > 1 \right\}. \tag{4}$$

Basically *LSF* is a boundary between the safe and failure regions. Geometrically *LSF* represents a failure limit hyper surface. The safety index $\boldsymbol{\beta}$ is reported to the point (often referred as the most probable point *MPP*) that lies closest to the origin of the system (**u*** = ($\boldsymbol{u_1^*, u_2^*,\dots,u_d^*}$)) in transformed reduced space $\boldsymbol{U^d}$ and belongs to the limit state surface. There is a direct relationship between the safety index and the probability of failure:

$$\boldsymbol{P_f} = \text{Prob}\left\{ \boldsymbol{u} \in \boldsymbol{U}^{\boldsymbol{d}} \middle| \boldsymbol{LSF}(\boldsymbol{u}) \le 0 \ \text{ or } \ \boldsymbol{LSF}(\boldsymbol{u}) \le 1 \right\} = \Phi(-\beta) \tag{5}$$

where Φ(…) is the one-dimensional standard Gaussian (normal) cumulative density function. The eqn. (5) represents the first order approximation for the failure probability. Basically, this relationship is only approximate, but in the unique case of a linear limit state function of Gaussian distributed random variables the relationship is exact.

To improve this drawback our procedure *estimates the safety index* based on artificial intelligence methods. The safety index will report to an equivalent hyper plane (equivalent linear form of the limit state function) into a *binary minimax classification manner*. The proposed procedure estimates the location of the most probable point (*MPP*) and calculates the safety index $\boldsymbol{\beta}$. Once the location of the *MPP*, in the reduced standard normal space has been found and safety index $\boldsymbol{\beta}$ was calculated, the failure probability of the system may be achieved. Into a binary classification problem, a pattern may be given through the sets of labelled data-points in $\boldsymbol{R}^n$ (or in other space) as: {*x1;x2,…xN*} and {*y1,y2…yM*}. Usually the label of a training pattern represents the category or the class to which the pattern belongs, in our case C1-safety domain and C2-failure domain. The demanding problem is to find a function $f : \boldsymbol{R}^n \to \boldsymbol{R}$ which is positive on the first set and negative on the second, such as $f(x_i) > 0,\ i = 1,\dots N$ and $f(y_i) < 0,\ i = 1,\dots M$. If the inequalities hold we say that "*f*" or its 0-level set $\{x \,/\, f(x) = 0\}$, separates or classifies the two sets of points. This function "*f*" sometimes viewed as a decision function is often named as the classifier. The classifier is built in a random manner on a training (learning) set and is validated on a test set (testing). Because training data is labelled, the classifier transforms these labels to new data (test data) as long as the feature space sufficiently distinguished each label as well. A heuristic approach to approximate linear discrimination (when the two sets of points cannot be perfectly linearly separated) is based on support vector classifiers. In the light of *SVM* a simplified reformulation for separating the set of *m* training vectors belonging to two separate classes, with a hyper plane, $\mathbf{w}^T \cdot z - b = 0$, is:

$$\left\{ \left(\vec{\mathbf{z}}_i, y_i\right) \middle| \vec{\mathbf{z}}_i \in R^n \ , \ y_i \in \{-1,+1\}, \ i = 1,..,m \right\}, \tag{6}$$

where y_i ($y_i \approx \mathrm{sign}\left(\mathbf{w}^T \cdot z_i - b\right)$) is the associated "truth" given by a trusted source. The underlying problem of interest is to establish a decision function $f : \mathbf{R}^n \to \{\pm 1\}$ using input-output training labelled data from eqn. (6). In principle a decision function is a function *f(z)* whose sign represents the class assigned to data points ***z***. If the points are linearly separable, in a two-class classification, then there exists an *n*-vector **w** and a scalar *b***.** To optimally separate the sets of vectors (eqn. 6) is equivalent to maximising the separation margin or distance between two parallel hyper planes $\mathbf{w}^T \cdot z = b + 1$ and $\mathbf{w}^T \cdot z = b - 1$. Minimising the probabilities that data vectors fall on the wrong side of the boundary we can establish the classifier design. A possible way is the attempt to control the misclassification probabilities in a worse case by setting as minimising the worst case-maximum probability of misclassification of future data points. This is the *minimax* approach [12,13]. Basically, as was stated in a *minimax* approach reported as a binary classification problem, $\mathbf{z_1}$ and $\mathbf{z_2}$ denote random vectors data from each of two classes as $\mathbf{z}_1 \in$ Class C1 and $\mathbf{z}_2 \in$ Class C2. Thus to estimate the location of the most probable point a hyper plane that separates the two classes of points with maximal probability must be determined (eqn. 7).

$$H(\mathbf{w}, b) = \left\{ \mathbf{z} \middle| \mathbf{w}^T \cdot \mathbf{z} = b \right\}, \ \text{where } \mathbf{w} \in \mathbf{R}^n \setminus \{0\} \text{ and offset } b \in \mathbf{R} \tag{7}$$

To overcome intricate non-linear classification problems, the kernel trick [8-13] is used to map the input data points into a high-dimensional '*d*' feature space $\boldsymbol{R}^d$, where a linear classifier corresponds to a non-linear hyper plane in the original space. In a *minimax* approach the classifier must minimise the misclassification probability by an optimal separating hyper plane that separates the two classes of points with maximal probability. In our safety index approach this optimal separating hyper plane is an equivalent failure limit hyper surface and represents a boundary between the safe and failure regions. Because **w** is the normal to the hyper plane and $\|\mathbf{w}\|$ is the Euclidean norm of **w**, according to well-known statements [14], $|b|/\|\mathbf{w}\|$ is the perpendicular distance from the optimal separating hyper plane to the origin and then it can be identified as the reliability-safety index $\boldsymbol{\beta}$.

$$\beta \cong |b|/\|\mathbf{w}\| \tag{8}$$

For non-linear classification problems in terms of a kernel function satisfying the Mercer's condition, the decision is transferred from the reduced original $\boldsymbol{U}^d$ space into reduced high dimensional feature space. For the output, it is not **w** that is returned, but instead of it the weights of the decomposition of **w** in the span of the data points are obtained [12,13]. The basis vectors can be interpreted as co-ordinate axes in the subspace and the "weights" of the basis vectors determine the corresponding "co-ordinate values" of the point. In our procedure there are some favourable circumstances where the linear discrimination works properly in the safety index assessment: (1) all projections of a standard normal distribution are also normal and generally linear functions of normally distributed data result also in normal distributions; (2) because the variates **u** are a set of standard uncorrelated variates the axes of the subspace $\boldsymbol{U}^d$ defined by these variates are orthogonal. Because the elements of kernel matrix were reduced to linear inner products and the classifier is provided in a linear configuration into a feature space, according to previous mentioned circumstances the general prerequisites for an exact assessment of eqn. (5) are fulfilled. The simulations are performed cyclic by multiple random trials. The performance of the procedure was evaluated based on test set accuracy (percentage of well-classified test data) and lower bound on the probability of correct classification of future data. Long random trials do not get improved accuracy or more reliable predictions, thus we considered appropriate to obey recent statements [11] and to work with a reduced learning set and to limit the random cyclic trials to k = 50…100.

RESULTS AND DISCUSSIONS

Numerical application presents an example based on the already published paper [16]. The numerical analysis has been carried out for a thin-wall pressure vessel working into a technological process of natrium salicilate. The risk of failure is based on traditional permissible stress methods and serviceability limit state. It is reported to the structural condition beyond which the service criteria specified for the component are no longer met [17,18]. The limit state function (*LSF*) was established based on the presumption of miss regulated of multi-fluid heating/cooling closed loop system. This situation can conduct to an uncontrolled value of pressure through the reactor jacket. Under these circumstances coupled with variations, uncertainties in service conditions, variability concerning the main variables and continuous chemical corrosion, it is possible that the pressure vessel cannot fulfil all the strength demands.

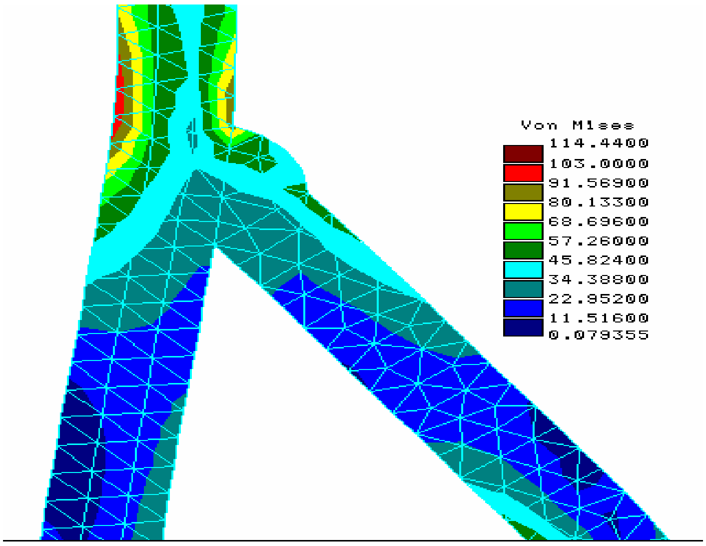
The main dimensional parameters of the pressure vessel, material characterization and operating technological conditions, are shown in Table1. Full statistic characterization of all the parameters on which the limit state function depends is not available. These parameters were set according to values known in practice or to values reported in literature [15,16].

Table 1. The main parameters and working conditions of the case study

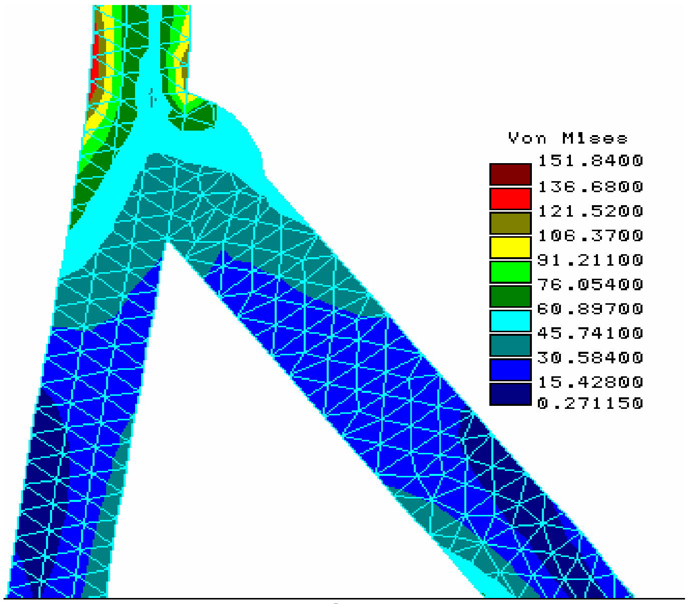
Variable/Parameter	**Symbol**	**Nominal value**	**Statistical distribution**
Design thickness [mm]	So	6	LogNormal $\mu = 6$; $\sigma = 0.60$
Corrosion rate [mm/year]	Vc	0.1	Normal $\mu = 6$; $\sigma = 0.60$
Young Modulus [MPa]	E	$2.12*10^5$	Normal $\mu = 6$; $\sigma = 0.60$
Yield technical limit [MPa]	$\sigma_{0.2}$	255	Normal $\mu = 6$; $\sigma = 0.60$
Admissible stress [MPa]	σ_{ad}	170	Nominal value
Outer nominal diameter [mm]	D	445	LogNormal $\mu = 6$; $\sigma = 0.60$
Equivalent length for outer operating pressure [mm]	L	250	LogNormal $\mu = 6$; $\sigma = 0.60$
T_U - service life [year]	Tu	12	Normal $\mu = 6$; $\sigma = 0.60$
Operating pressure [MPa]	Poe	1.2	Normal $\mu = 6$; $\sigma = 0.60$
▪ Technological mixture -- NaOH and phenol; ▪ Steel K41.2b STAS 2883/2-80			

The state of stresses in various operating technological conditions obtained by simulations with the finite element method (FEM) using the professional package *COSMOS/M Designer II* is presented in Figure 1. The higher stresses were determined in the area of welded reactor jacket. Because stresses don’t exceed permissible stress and stress concentrations are moderate, the structure is one that will not be expected to fail, thus a real structural safety may be suggested. Because uncertainties and variability concerning the main variables, in fact any structure has a risk, even though it can be extremely small. It follows that even though a structure is safe there is still the possibility of failure.

By the proposed procedure we can express a global risk of failure for the mentioned pressure vessel. According with main dimensional parameters of the pressure vessel, material characterization and technological conditions (Table 1) the LSF system vectors are obtained by simulations of deterministic relationships (Table 2). These simulations might substitute real data as inputs to consequence models and illustrate the variability and uncertainty of the process.



a.



b.

Figure 1. Equivalent Von Misses stresses in the area of welded reactor jacket
a. Von Misses stress at the beginning of the life cycle (So = 6 mm)
b. Von Misses stress along the life cycle (So = 4.8 mm)

The experiments are performed cyclic „*k*” times by multiple random trials. In each experiment the data was randomly split into (50%…90%) training and (50%…10%) test set. A number of different training and test sets are randomly chosen at each cycle. In each cycle the safety index $\boldsymbol{\beta}$ is evaluated and stored. The results of the proposed procedure are presented in Table 3. The performance of the procedure based on the test set accuracy (percentage of well-classified test data) and on the lower bound of correct classification of future data reveals a reasonable level. In every experiment the lower bound of correct classification of future data is smaller than the test set accuracy as was stated by [12,13]. Thus the lower bound is not violated and the linear approach is a robust one.

Table 2. Deterministic-explicit relations

Outer operating pressure in reactor jacket	***Deterministic Numerical Values***	
	Tu=0 [year]	Tu=12 [years]
***Ellipsoidal* closure** $D^2/H*s \le E/2*\sigma_{0.2} \rightarrow 0.0269 < 0.25$ – Plastic domain **Outer critical operating pressure** $P_{cr} = 8*s*\sigma_a*H/(D+(2H/D)*s)*D$	4.54 MPa	3.33 MPa
Cylindrical shell $1.5(2*s/D)^{0.5} \le L/D \le (D/2*s)^{1/3} \rightarrow$ $0.2463 < 0.55955 < 6.089$ - Short shell $s/D > 1.1*10^{-2}(L^2*\sigma^{t2}_{0.2}/10^{-6}*D2*E^2) \rightarrow$ $1.3483*10^{-2} > 0.8524*10^{-2}$ **Outer critical operating pressure** $P_{cr} = 2*E^t*s/D*K1$ $K1=(1.5*E^t/\sigma_{0.2})+1.39*\sigma^t_{0.2}*D*L^2(C_{s2}-1)/ E^t*s^3$ $C_{s2}=1.5(1+0.93(C_{s1}-1.5)*(L*\sigma^t_{0.2}/ E^t*s)^2*D/s)$ where $C_{s1} = 2.6$	4.36MPa	3.067MPa
Original explicit limit state functions *LSF* $\Rightarrow$ ***LSF* = P_{cr} – *Poe***		
Original explicit kernel function A polynomial kernel with a unit offset $K(z_i, z_j) = (z_i \cdot z_j + 1)$		

Table 3. Results for safety index assessment based on averaged values

Safety index $\boldsymbol{\beta}$	Probability of failure	LBCPFD [%]	TSA [%]
2.8131	0.00245	0.896	0.990
k = 100 cycles *LBCPFD* = lower bound on correct classification of future data *TSA* = test set accuracy (percentage of well-classified test data)			

The reported results are averaged over the entire random cyclic trials, as was stated [12,13]. Due to uncontrolled combinations of the variations in service loads, uncertainties and variability of any parameters, during service elapsed time, damage can occur. In fact, the mentioned pressure vessel has a risk, even though it is small. Extending Mc.Leods and Plewes's conversion scale to structural safety, this probability of failure $P_f = 0.00245$ is placed on the scale of risk in the range between $10^{-2}...10^{-4}$. This risk is characterized as a reduced one, when the failure is possible. It follows that the pressure vessel is still possibility of failure even though it is relative safe under a pure deterministic analysis.

CONCLUSIONS

The paper presents another insight into safety assessment and structural risk analysis using artificial intelligence methods. The proposed procedure offers a greater reliability in the safety prediction and reduction the risk of failure. High values for safety factor β lead to low values for the risk of failure. This approach improves some drawback of the traditional reliability methods, reduces the need for excessive safety margins in design or additional cumbersome experimental-analytical approaches. At the same time it focuses on uncontrolled combinations between the variations in service loads, uncertainties and variability of parameters, during service elapsed time. This is a key factor that could have a profound impact on the risk assessment. Based on numerical application we can highlight that using a corrosion decay model it is possible to establish a safer working life of the pressure vessel. These types of studies become recommended and necessary for engineers, especially for chemical engineers to work out optimal safety decisions, inspection and maintenance schedules.

REFERENCES

1. D.G. Elms, *Struct. Safety*, **1999**, *21*, 311.
2. R.E. Melchers, "Structural reliability analysis and prediction", John Wiley & Sons, New York, **1999**.
3. O. Ditlevsen, O., H.O. Madsen, "Structural reliability", John Wiley&Sons. New York, **1996**.
4. J.E. Hurtado, D.A. Alvarez, *Journal of Structural Engineering*, **2003**, *129*(8), 1141.
5. J.E. Hurtado, *Struct.Safety*, **2004**, *26(3)*, 271.
6. C.I. Anghel, Al. Ozunu, *Chem. Pap*. **2005**, *59(6b)*, 469.
7. C.I. Anghel, *Int Jnl of Pressure Vessels and Piping*, **2009**, (in press).
8. V.N. Vapnik, "The nature of statistical learning theory", 2nd edition, Springer, New York, **2000**.

9. B. Scholkopf, C. Burges, A. Smola, (Eds.), “Advances in Kernel Methods. Support Vector Machines”, MA MIT Press Cambridge, **1998**.
10. C.J.C. Burges, *Data Min. Knowl. Disc,* **1998**, *2*, 121.
11. A. Bordes, S. Ertekin, J. Weston, L. Bottou, *Journal of Machine Learning Research*, **2005**, *6*, 1579.
12. G.R.G. Lanckriet, EL. Ghaoui, C. Bhattacharyya, M.I. Jordan, in: T.G. Dietterich, S. Becker, Z. Ghahramani (Eds.), *Advances in Neural Information Processing Systems*, MA. MIT Press Cambridge, **2002**, *14*, 801 (http://robotics.eecs.berkeley.edu/~gert/)
13. G.R.G. Lanckriet, EL. Ghaoui, C. Bhattacharyya, M.I. Jordan, *Journal of Machine Learning Research*, **2002**, *3*, 555.
14. C. I. Anghel, I. Lazar, R.I. Iatan, “Proceedings – Conferința Natională de Echipamente, Instalații şi Inginerie de Proces”**, 2004**, 253.
15. Al. Pavel, D. Popescu, *Rev. Chim.*, **1998**, *49, 2*, 128.
16. *** ASME - Boiler And Pressure Vessel Code, Section VIII, The American Society of Mechanical Engineers; New York, 1**995**.
17. *** Eurocode pr.EN 13445-3, Unfired Pressure Vessels – Part 3: Design, **1999**.
18. *** Eurocode pr.ENV 1991-1, Eurocode 1: Basis of design and actions on structures, Part 1 Basis of Design, **1994**, (CEN/TC 250).