ON THE ANALYSIS OF RECURRENCE CHARACTERISTICS OF VARIABLE ACTIONS

Horea SANDI *)

ABSTRACT

The paper is devoted to some methodological problems raised by the analysis of hazards due to variable actions having implications for the risk of damage to structures. The basic recurrence model used is that of Poissonian stochastic processes. The techniques of calibration of specific recurrence characteristics are discussed, adopting a critical point of view versus statistical analyses relying exclusively on data like annual maxima. The adoption of some types of distributions is critically discussed, from the point of view of their compatibility with the Poissonian model referred to. Only the Gumbel and Fréchet distributions are accepted as adequate for the purpose adopted. Starting from their common properties an unbounded family of distributions is proposed. This family makes it possible to adopt calibrations providing an approximation of unlimited closeness to observation samples. The case of a pluri-dimensional characterization of the randomness of observation data is then tackled, considering as an illustrative case the directional statistical analysis of sequences of wind events. Some specific expressions are proposed for the directional analysis, leading to a good approximation of observation samples. A case study relying on the expressions is then presented.

Keywords: actions on structures, hazard, Poissonian processes, extreme value distributions

REZUMAT

Lucrarea este dedicată unor probleme metodologice ridicate de analiza hazardurilor datorite acţiunilor variabile care au implicaţii pentru riscul de avariere a structurilor. Modelul de bază utilizat este cel de proces stochastic poissonian. Sunt discutate tehnicile de calibrare a caracteristicilor de recurenţă specifice, adoptându-se o poziţie critică faţă de analizele statistice care au drept obiect exclusiv parametri ca maximele anuale. Este discutată critic adoptarea anumitor tipuri de distribuţii, prin prisma compatibilităţii lor cu modelul poissonian menţionat. Sunt reţinute (dintre distribuţiile clasice), drept modele adecvate, distribuţiile Gumbel şi Fréchet. Pornind de la proprietăţile comune ale acestora, este propusă în continuare o familie (nemărginită) de distribuţii care generalizează distribuţiile clasice menţionate, această familie permiţând o mulare oricât de fină pe ansamblurile de date de observaţie. În continuare, este analizat cazul unei caracterizări pluri-dimensionale a caracterului aleator al datelor de observaţie, considerându-se, ca exemplu ilustrativ, problematica analizei statistice direcţionale a succesiunii cazurilor de vânt. Sunt propuse, pentru analiza direcţională, expresii specifice, care permit o bună aproximare a ansamblurilor de date de observaţie. Este prezentat un studiu de caz, având la bază utilizarea distribuţiilor propuse.

Cuvinte cheie: acţiuni, hazard, procese poissoniene, distribuţii de valori extreme

1. INTRODUCTION

The paper is devoted to some important aspects of the characterization of variable actions and of corresponding hazards to structures. Cases of occurrence of these actions at levels of severity that are relevant for structural safety and risk are dealt with. Such cases are singled out and their characterization is dealt with. On the other hand, the recurrence of actions at various severity levels is dealt with too, in order to characterize corresponding hazards.

A case of occurrence of a certain action may be characterized in various terms, in a more or less detailed manner, depending upon the aspects of interest considered. This problem is briefly dealt with. The first approach is based on a scalar (or 1D) characterization. Thereafter a more general, *n*D, approach is dealt with and an illustration, corresponding to the consideration of wind speed, together with its azimuthal orientation, is presented.

The complementary problem dealt with is represented by the analysis of recurrence

*) Honorary researcher, Institute of Geodynamics „Sabba S. Ştefănescu" of the Romanian Academy & Member, Academy of Technical Sciences of Romania; e-mail: sandi@geodin.ro & horeasandi@yahoo.com

characteristics of actions. The use of stochastic processes as basic recurrence models is adopted. As it is well known, the Poissonian model is largely used in order to investigate the features of sequences of cases of occurrence of actions like the seismic or meteorological ones. Even if this model is in disagreement with some features of the sequences referred to and of features of phenomena or of originating processes, like tendency to annual periodicity of occurrence of meteorological actions or consequences of the phenomenon of energy accumulation lying at the basis of earthquake generation, it is used especially due to some hard facts. On one hand, the outcome of its use appears to be in many cases satisfactory in order to quantify hazards for relatively long time horizons. On the other hand, the calibration of more sophisticated models is quite seldom feasible, due to the limits to the observation data at hand, which often raise difficulties even for the calibration of the characteristics of Poissonian processes.

Due to the reasons mentioned, the acceptance of the Poissonian model lies at the basis of following developments. Their main objective is to propose some tools intended to contribute to a more consistent approach to hazard analyses, to be in agreement with the basic requirement of invariance of results with respect to some processing conventions.

The paper is based on the developments of [7], to which some additional developments and explanations are added.

2. ANALYTICAL DEVELOPMENTS

2.1. General

The model adopted in order to characterize one case of occurrence of a variable action corresponds to sequences of occurrence cases of durations that are short in comparison with exposure durations which are relevant for the analysis of the risk affecting structures. So, the formulations adopted correspond implicitly to an ideal situation, for which the duration of application during an occurrence case is negligible.

The space of parameters characterizing actions, q, considered for the beginning, is mono – dimensional, while a random variable Q, which can take various values q for the successive cases of occurrence, is used. Such spaces lie implicitly at the basis of most code provisions, where hazard characteristics are related to one single parameter, as a rule an intensity or amplitude type parameter.

It is nevertheless of interest to push this approach to a multi – dimensional characterization of the randomness of the parameters characterizing a case of occurrence. Some developments in this sense are presented in subsections 2.6, 2.7, 2.8.

2.2. Review of some basic characteristics and relations in case of 1D characterization of randomness

The basic recurrence characteristics and the basic relations between them, as appropriate for the case of Poissonian processes, are reviewed subsequently. They are to be used in the quantification of recurrence of actions under the hypotheses accepted. A (scalar, random) severity measure, Q, is used.

The basic recurrence characteristics used throughout the paper are:

expected occurrence frequency density for $Q = q$, denoted $n_Q(q)$,

- an *expected number of occurrence cases characterized by values $Q \geq q$*, during an observation, or exposure, time interval T, denoted $N_Q(Q \geq q, T)$, where

$$N_Q(Q \geq q, T) = T \int_q^\infty n_Q(q')\,\mathrm{d}q' \qquad (2.1)$$

(e.g. : $N_Q(Q \geq q, 1$ yr.) is the yearly expected occurrence frequency of values $Q \geq q$);

- an *expected occurrence* return *period*, $T_Q(q)$, given by the expression:

$$T_Q(q) = [\int_q^\infty n_Q(q')\,\mathrm{d}q']^{-1} = [N_Q(Q \geq q, T)/T']^{-1} =$$

$$= T / N_Q(Q \geq q, T) \qquad (2.2)$$

- *probabilities of non-occurrence and non - exceedance*, $P_{Q0}(q, T)$, or, more general,

- *probabilities of m-time occurrence or exceedance*, $P_{Qm}(q, T)$, for a time horizon T, where

$$P_{Q0}(q, T) = \exp[-N_Q(Q \geq q, T)] \qquad (2.3)$$

$$P_{Qm}(q,T) = \exp[-N_Q(Q \geq q,T)] \times$$
$$\times [N_Q(Q \geq q,T)]^m / m! \qquad (2.3')$$

(where $m = 0, 1, 2, \ldots$), with the obvious condition

$$\sum_m^{0,\infty} P_{Qm}(q,T) \equiv 1 \qquad (2.4)$$

Of course, the time related parameters T and $T_Q(q)$ will be expressed in terms of the same time units, usually in years.

Such a simple, *1D*, approach cannot account, of course, for the multi-dimensional variability of actions and it should be surpassed, at least from a theoretical viewpoint, in order to reach a more detailed characterization of randomness. An example of possible extension in this view is presented, as mentioned before, in subsections 2.6, 2.7, 2.8.

2.3. Some consequences for extreme value distributions

Two widely used, by now classical, distributions of maxima for random variables, are considered here as a starting point: the Gumbel and the Fréchet distributions. According to the author's views, they are the single distributions, widely used to date, that are compatible with a recurrence representation corresponding to the basic Poissonian process model. As a counterexample, the lognormal distribution, which is also used in this field, is not compatible with this basic model. The Gumbel and the Fréchet distributions will be considered here as non-occurrence and non-exceedance probabilities (which shall be explicitly related to some reference exposure horizon T, in order to make sense):

- the Gumbel distribution (or the Type I extreme value distribution) is:

$$F_Q^{(I)}(q,T) = P_{Q0}^{(I)}(q,T) = \exp\{-\exp[-\alpha(T)(q-u)]\} \qquad (2.5)$$

$\alpha(T) > 0$);

- the Fréchet distribution (or the Type II extreme value distribution) is:

$$F_Q^{(II)}(q,T) = P_{Q0}^{(II)}(q,T) = \exp\{-[q/u]^{-k(T)}\} =$$
$$= \exp\{-\exp[-k(T)(\ln q - \ln u)]\} \qquad (2.6)$$

(where $k(T) > 0$, $q > 0$).

Note that these two distributions are closely related in case one looks at the last variant of expression (2.6). This remark suggests an attempt of generalization, namely the proposal of a more general kind of distribution,

$$F_Q^{(S)}(q,T) = P_{Q0}^{(S)}(q,T) =$$
$$= \exp\{-\exp[-\alpha(T)(\varphi<q> - \varphi<u>)]\} \qquad (2.7)$$

(where $\alpha(T) > 0$, while $\varphi<q>$ is a monotonically increasing function of q). Note here also that in case of the distribution (2.5), one has an expression $\varphi<q> = q$, while in case of the distribution (2.6), one has an expression $\varphi<q> = \ln q$.

The consideration of such a more general kind of distribution provides considerably improved possibilities to fit observation data, as compared to the constraints created by the exclusive use of the classical distributions (2.5) and (2.6). A way to calibrate the newly introduced function deserves therefore to be examined. As an example, a simple possible solution of interest (which directly generalizes the Gumbel variant and is used in connection with the case study of Section 3) is

$$\varphi<q> = q^\beta \qquad (2.8)$$

(where $\beta > 0$). The higher the value β is, the less conservative the distribution (2.7) will be, i.e. the lower the probabilities of occurrence & exceedance of the most severe values q will be.

Comment: the formulation adopted, where two parameters explicitly depending on the exposure duration T, $\alpha(T)$ and $k(T)$ respectively, appear, anticipates some implications of subsequent developments.

Given the expression (2.3), the expected numbers of occurrence cases corresponding to the expressions (2.5), (2.6) and (2.7) will be respectively

$$N_Q^{(I)}(Q \geq q,T) = \exp[-\alpha(T)(q-u)] \qquad (2.9)$$

$$N_Q^{(II)}(Q \geq q,T) = [q/u]^{-k(T)} =$$
$$= \exp[-k(T)(\ln q - \ln u)] \qquad (2.10)$$

$$N_Q^{(S)}(Q \geq q,T) = \exp[-\alpha(T)(\varphi<q> - \varphi<u>)] \qquad (2.11)$$

and, more specifically, in case the expression (2.8) is adopted,

$$N_Q^{(S)}(Q \geq q, T) = \exp[-\alpha(T)\,(q^{\beta} - u^{\beta})] \quad (2.12)$$

Note that the well known Richter earthquake magnitude recurrence law [2], rewritten by using the notations introduced previously,

$$\lg N_Q^{(R)}(Q \geq q, T) = a(T) - bq \quad (2.13)$$

may be made equivalent to a Gumbel distribution. Note also the relation

$$a\,(T) = a\,(1) + \lg T, \quad (2.14)$$

which is of interest in case one wants to change the reference time interval of recurrence characteristics (e.g.: in case of passage from $T = 1$ yr. to $T = 100$ yrs.).

2.4. Discussion on the calibration of hazard characteristics

The developments presented are connected with methodological consequences. On this basis, it turns out that a reasonable way to calibrate the recurrence characteristics will consist of a direct calibration of the function $N_Q\,(Q \geq q, T)$. All other recurrence characteristics referred to in Section 2.2 can be easily determined on this basis. The results obtained in this way are based on the whole of observation data at hand and do not depend on an artificial convention, like e.g. exclusive consideration of *yearly* maxima, which is rather usual. The exclusive consideration of yearly maxima, widely applied due to the developments of [3], frequently referred to as Gumbel statistics, and implicitly accepted in some reference publications like [1] or [8], leads to questionable results. E.g., the consideration of statistics of monthly maxima would lead to different results. So, a natural basic requirement, namely that of invariance of the outcome on calibration of hazard characteristics, is not satisfied in this way.

2.5 Some additional consequences for the techniques of calibration of recurrence characteristics

In case of adoption of a distribution (2.7), two expressions that are useful in order to determine recurrence characteristics may be derived. In case one considers two reference values of q, q_1 and q_2 respectively (where usually $q_1 < q_2$) for which one knows the corresponding return periods, $T_1 = T_Q(q_1)$ and $T_2 = T_Q(q_2)$ (e.g. in case of wind 10 years and 50 years, respectively) and one wants to determine the return period $T_Q(q)$ for some other value, q, on the basis of these data, one obtains the relation (recommended primarily for interpolation)

$$\ln T_Q(q) = [(\varphi < q_2 > - \varphi < q >) \ln T_1 + (\varphi < q > - \varphi < q_1 >) \ln T_2] \,/\, (\varphi < q_2 > - \varphi < q_1 >) \quad (2.15)$$

In the opposite sense, in case one wants to determine the value q corresponding to a return period of interest T, one obtains the relation

$$\varphi < q > = [(\ln T_2 - \ln T)\,\varphi < q_1 > + (\ln T - \ln T_1)\,\varphi < q_2 >] \,/\, (\ln T_2 - \ln T_1) \quad (2.16)$$

2.6. Some steps towards a multi – dimensional randomness characterization in recurrence analysis

Returning to the closing remarks of Subsections 2.1 and 2.2, some developments oriented towards hazard characterization for actions for which a multi – dimensional randomness characterization is required, are presented. The author discussed this problem in a rather general, formal, frame in [4, 5, 6].

Assume now, instead of the 1D space considered in subsection 2.2, the definition and use of a space of a finite number of dimensions, having the coordinates q_j ($j = 1, \ldots, J$), which can be referred to as a space of *macroscopic action characteristics* q_j. This space is, of course, not really a vectorial space (among other, several coordinates that are proper to this space cannot take negative values). In this latter variant the expected frequency density $n_Q\,(q)$ of relations (2.1) etc. would be replaced by an expected frequency density $n_Q(q_j)$, where q_j are the coordinates of the space, while the expected frequency $N_Q\,(Q \geq q, T)$ would be replaced by an expected frequency of cases $(q_j) \in \Omega$, $N_Q(\Omega, T)$, where Ω is a (closed) domain of the space of coordinates q_j. Instead of the expression (2.1) one has now

$$N_Q((q_j)\in\Omega,T)=T\int_\Omega n_Q(q_j)\,\mathrm{d}\Omega\,,\quad(2.17)$$

while instead of the first variant of expression (2.2) one has, for the return period of occurrence of events with $(q_j) \in \Omega$

$$T_Q(q)=[\int_\Omega n_Q(q_j)\,\mathrm{d}\Omega]^{-1}=$$

$$=[N_Q((q_j)\in\Omega),T)/T]^{-1}=$$

$$=T/N_Q((q_j)\in\Omega),T)\qquad(2.18)$$

The expressions (2.3) and (2.3′) will be replaced in this case by

$$P_{Q0}((q_j)\in\Omega),T)=\exp[-N_Q((q_j)\in\Omega),T]\qquad(2.19)$$

and

$$P_{Qm}((q_j)\in\Omega),T)=$$

$$=\exp[-N_Q((q_j)\in\Omega),T]\times[N_Q((q_j)\in\Omega),T)]^m/m!\qquad(2.19)$$

while the condition (2.4) will become

$$\Sigma_m^{0,\infty} P_{Qm}((q_j)\in\Omega),T)\equiv 1\qquad(2.20)$$

2.7. A case of 2D recurrence analysis: the directional analysis of recurrence of wind cases

This time, a specific problem is dealt with, namely the consideration of wind directionality.

In this frame, a case of incidence of wind action will be characterized by two parameters:

- an intensity measure, q (usually, a reference wind speed value);
- a directionality measure, θ (usually, an azimuthal angle).

Due to pragmatic reasons, the parameter θ should be discretized as θ_k, correspondingly lumping observation data, since there will never exist data to support in a satisfactory way a continuous directional characterization. The number k could vary from 1 to 8 or from 1 to 16, depending upon the amount and accuracy of observation data.

Returning to Subsection 2.2, attention is to be paid, as a starting point, to the recurrence characteristics $n_Q(q)$ and $N_Q(Q \geq q, T)$. The corresponding characteristics to be used for a two – dimensional approach will be:

- an *expected directional occurrence frequency density*, denoted $n_Q{}^{(2)}{}_k(q)$;
- an *expected directional number of occurrence cases for a time interval T*, $N_Q^{(2)}{}_k(Q\geq q,T)$.

Two corresponding basic relations can represent a starting point for directional hazard analyses:

$$n_Q(q)=\Sigma_k n_Q{}^{(2)}{}_k(q)\qquad(2.21)$$

and

$$N_Q(Q\geq q,T)=\sum_k N_Q^{(2)}{}_k(Q\geq q,T)\qquad(2.22)$$

Further relations, analogous to those of Subsection 2.2, can be developed according to needs.

2.8. Additional methodological elements concerning the analysis of the recurrence process of wind cases, under consideration of the azimuthal direction of wind flow

The problem of selecting of a suitable type of expressions of the distributions to be used, to fit the requirements of statistical analysis of the recurrence process under consideration of the azimuthal direction of wind flow, is now examined. The basic adequacy criterion considered is to satisfy the relations (2.21) and (2.22) without allowing any residues. This condition is quite strongly restrictive.

A type of function corresponding to the requirement formulated is some power of q, q^k, and, consequently, also some linear combination of such functions. Keeping in view the general tendency observed while performing recurrence analyses, namely the monotonic decrease of functions $n_Q(q)$ and $N_Q(Q \geq q, T)$, the powers k should be strictly negative. A suitable solution, given its simplicity, is a linear combination,

$$N_Q(Q\geq q,T)=\Sigma_j c_j\, q^{-aj}\qquad(2.23)$$

and, consequently,

$$n_Q(q) = a\,\Sigma_j c_j\, j\, q^{-aj-1} \qquad (2.24)$$

where a is a positive constant (which can be equal to 1 in case there are no reasons to adopt a non-integer value), while the sequence of numbers k is a (relatively short) segment of the sequence of natural numbers.

A passage to directional recurrence analysis, combining the developments of relations (2.21) and (2.24), or (2.22) and (2.23) respectively, leads to double addition, with respect to the indices j and k.

3. A CASE STUDY

The case study presented concerns the recurrence of relatively strong wind occurrence cases at the location of the meteorological station of the City of Ploieşti, in Romania. The time interval for which data available were used is from 1966 to 2001, i.e. a total length of 36 years. This was the most appropriate interval from the point of view of homogeneity of the observation techniques used. The approach presented in previous section was applied. The results presented are related only to the analysis of recurrence irrespective of wind directionality.

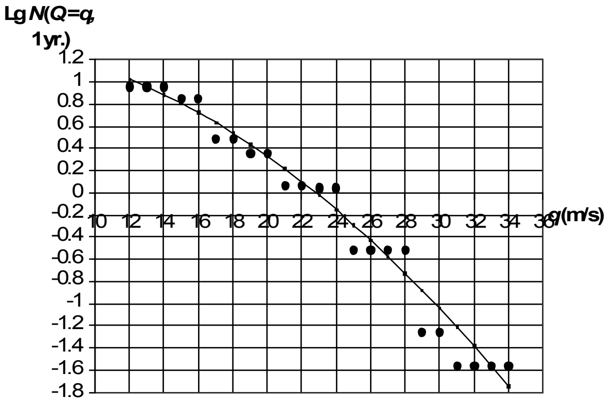

Fig. 3.1. Cumulated number of events (reduced to one year) and estimated recurrence characteristic lg N_Q ($Q \geq q$, *1 yr.*) for wind speed in Ploiesti

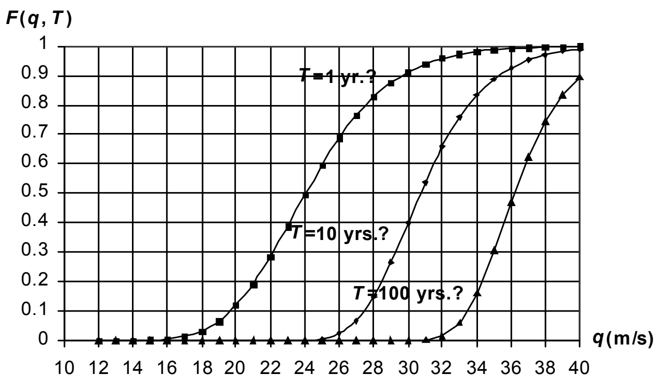

Fig. 3.2. Probability function of wind speed in Ploiesti, for various exposure durations T, $F_Q^{(S)}(q, T)$

A look at the data concerning the recurrence of peak gust wind velocities led to the conclusion that, at least for the domain covered by the data at hand, an expression (2.8) of φ <q> would be fairly appropriate, as shown subsequently.

The cumulative number of observed cases of occurrence, normalized, or reduced to a one year time interval, is presented in graphic terms in Fig. 3.1. As shown there, it turned out that a value β = 2.0, as adopted in order to derive the recurrence characteristic lg N_Q ($Q > q$, T), appears to be satisfactory. In order to determine the value of α, it was stated, on the basis of data at hand, that

$$N_Q(Q \geq 33\ \mathrm{m/s},\ 36\mathrm{yrs.}) \approx 1 \qquad (3.1)$$

and

$$N_Q(Q \geq 13\ \mathrm{m/s},\ 36\mathrm{yrs.}) \approx 329 \qquad (3.2)$$

which led to the condition

$$\alpha \times (33^2 - 13^2) = \lg(329/1) \approx 2.5172 \qquad (3.3)$$

and to the value

$$\alpha \approx 2.5172/(33^2 - 13^2) \approx 0.002736 \qquad (3.4)$$

So, the expression of lg N_Q ($Q \geq q$, T) became

$$\lg N_Q(Q > q,\ 1\ \mathrm{yr.}) =$$
$$= \alpha \times (33^2 - q^2) - \lg(36\ \mathrm{yrs.}) =$$
$$= 0.002736 \times (1089 - q^2) - 1.5563 \qquad (3.5)$$

i.e.

$$\lg N_Q(Q \geq q, T) =$$
$$= 0.002736 \times (1089 - q^2) - 1.5563 + \lg T, \qquad (3.5')$$

$$N_Q(Q \geq q, T) =$$
$$= 10 \uparrow [0.002736 \times (1089 - q^2) - 1.5563 + \lg T] \qquad (3.5'')$$

A plot of this outcome is given in Fig. 3.1.

The probability function corresponding to an exposure duration T, $F_Q^{(S)}(q, T)$, will be

$$F_Q^{(S)}(q,T) = \exp[-N_Q(Q \geq q, T)] =$$
$$= \exp\{-10 \uparrow [0.002736 \times (1089 - q^2) - 1.5563 + \lg T]\} \quad (3.6)$$

and is plotted in Fig. 3.2, alternatively for exposure durations T of 1, 10 and 100 years.

Obtaining the value β = 2.0 means that the distribution derived is less conservative than the classical Gumbel distribution (for which β = 1.0), i.e. that it leads to lower occurrence / exceedance probabilities for the most severe wind speed values.

4. CONCLUDING CONSIDERATIONS

The developments presented make it possible to derive following conclusions:

1. It is recommendable to carry out hazard analyses using as a starting point the structure of relations that are proper to Poissonian processes. The appropriate way to proceed is to calibrate the expected number of occurrence cases characterized by values $Q \geq q$, during an observation, or exposure, time interval T, denoted $N_Q(Q \geq q, T)$.

2. The adoption of the use of functions $\varphi < q >$, as introduced in Subsection 2.2, appears to be suitable. A calibration on the basis of observation data is easily feasible and leads to invariant results. Of course, the quality of results will depend on the quality of basic information data available.

3. There are situations requiring exceeding the 1D approach. A case where this requirement is obvious is that of analysis of recurrence of cases of wind, accounting for directional distribution too.

4. There exist, of course, also other situations in which a multi-dimensional approach of hazard characterization may be of interest. The developments of subsection 2.6 provide some guidelines in this sense.

REFERENCES

[1] Benjamin J. R. & Cornell C. A. – *Probability, statistics and decision for civil engineers*, McGraw-Hill, New York ,1970

[2] Esteva L.: Seismicity – *Ch. 6. in Seismic risk and engineering decisions,* (Ed,'s. C. Lomnitz & E. Roseblueth), Elsevier, Amsterdam-Oxford-New York, 1976

[3] Gumbel E. J. - *Statistics of Extremes*. Columbia University Press, New York, 1958

[4] Sandi H. – *Some analytical considerations on the concept of design accelerograms*, In Proc. 8-th US National Conference on Earthquake Engineering, San Francisco, EERI, 2006

[5] Sandi H. – *O privire critică asupra modelărilor utilizate în ingineria structurilor* (*A critical look on the modellings used in structural engineering*), Proc. 4th Scientific Session, "CIB 2008", organized by the University Transilvania Braşov, Braşov, Romania, November 2008

[6] Sandi H. – *Considerations on the updating of earthquake resistant design codes*, In Proc.10-th International Conference on Structural Safety and Reliability (ICOSSAR 09). Osaka, Balkema, 2009

[7] Sandi H. – *Towards a Gumbel – Fréchet type family of extreme value distributions*, Proc. International Symposium on Reliability Engineering and Risk Management, Shanghai, 2010, Kluwer

[8] Simiu E. & Scanlan R. H. – *Wind effects on structures: an introduction to wind engineering,* John Wiley and Sons, New York, 1978