Cella, Meroni and Zonno:

CZAR: An Expert System for seismic hazard analysis

Tiré à part des
Cahiers du Centre Européen de Géodynamique et de Séismologie.
Volume 6 - 1992

"CZAR: An Expert System for Seismic Hazard Analysis"

Cella F., Meroni F., Zonno G.

Istituto per la Geofisica della Litosfera - Consiglio Nazionale delle Ricerche
Via Ampère 56 - 20131 Milano - Italy

INTRODUCTION

In the field of risk evaluation and seismic hazard assessment, it is necessary to codify a great quantity of aspects of the so called knowledge and to supply an intelligent support for the not-well-defined problems (data uncertainty, lack of rigorous solution algorithms). The main feature of an expert system is to emulate effectively the behaviour of a human expert in a particular and defined field, enabling the final user to improve its decisional process and giving access to him to a knowledge base otherwise not clearly codified. From these general considerations the intention came to develop the prototype CZAR (Classificatore Zone A Rischio) that is an expert system reproducing the Italian seismic classification based on the definition of Seismic Hazard given by Progetto Finalizzato Geodinamica (PFG) of the Consiglio Nazionale delle Ricerche (CNR). The expert system built up on the commercial shell Nexpert Object is working on a personal computer through a graphic interface developed with the Graphical User Interface (GUI) of Window 3.0. This user friendly interface makes possible the choice of different procedures to estimate the hazard parameters and also allows the activation of the classification inferential process. The influence of different assumptions and strategies has been evaluated by a mathematical algorithm suggested by the general structure of the Bayes' theorem. In this paper the prototype of the expert system has been applied to the data relating to Toscana region (central Italy) and the interactive evaluation of the maps furnishes a relative measure for discrepancies on seismic classification in the 2nd seismic category.

BRIEFLY ON ITALIAN SEISMIC CLASSIFICATION

Being Italy a highly seismically active country, its various governments, administrating the territory all years along, started facing the problem of preserving people and property from earthquakes more than two centuries ago. In fact the first legislative measures were taken by the Bourbon government after the 1783 earthquakes in Calabria, which caused more than 30,000 death. Later on the choice of the sites suitable for the rebuilding, as well as construction standards, were based on the regulations issued by the Pope's state, just after the 1859 earthquake in Norcia. On the unification of Italy, any kind of regulation fell into disuse and consequently the Italian state was caught unprepared for the unexpected situation created by the 1883 earthquake, which badly ruined many villages on Ischia island. The quake, which destroyed Reggio Calabria and Messina on Dec. 28, 1908 causing 80,000 death, was one of the strongest events in the Italian peninsula over the last ten centuries. Soon afterwards, the national seismic classification had been promulgated: it consisted of a list of the municipalities, where technical regulations for building, defined by the Royal Decree, were applied. Subsequently the seismic classification has been updated after any destructive earthquake. This produced a few surprising elements from the point of view of the classification, the most glaring ones are those relating to the recent 1968 Belice earthquake, that ruined a not seismically classified area. This is due to the fact that, till then, the seismic classification was uncomplete because it was based on some particular events. In fact the only municipalities having suffered damage for earthquakes since after 1908 were classified and no scientific examination of the whole Italian territory seismicity had been done at all.

After the 1976 Gemona earthquake, a national project was started in Italy by the Progetto Finalizzato Geodinamica (PFG) granted by the National Research Council (CNR) which sponsored further multidisciplinary activities both on the basic study of the earth sciences and on the methodological study of seismic hazard and risk mitigation. The Irpinia earthquake focused a general effort by the scientific community to produce, as a synthesis of all available information, the seismic classification for new buildings, with the "Proposta di riclassificazione del territorio nazionale" [1]. This proposal was presented to the Italian government and translated by the Ministry of Public Works into a series of decrees which constitute the present seismic classification of all Italian territory.

PHILOSOPHY OF SEISMIC CLASSIFICATION

The seismic classification of the territory points out the areas where the constructions should be built up according to precise technical regulations, so that every building should attain an assigned protection level. The seismic classification of the territory means assigning to a region a defined security level, according to the seismicity degree of the zone. The problem is the choice of the parameter to adopt in order to compare sites with different seismicities and to look for a principle enabling the homogeneous determination of security levels relating to various sites.

The classification philosophy proposed by PFG [1] [2] , derived from the necessity to give a "rough" preliminary classification, for the zones that would be exposed to a seismic risk comparable with the estimated value for sites already classified as seismic regions after 1908.

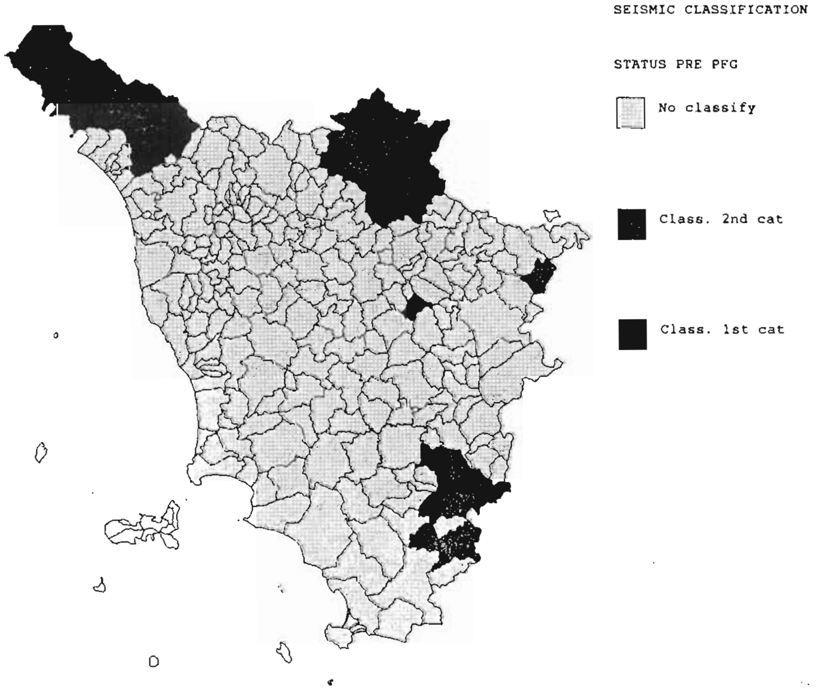


Fig. 1 Official seismic classification map of the Toscana region on the status pre PFG.

Unfortunately this was lacking a rational analysis to cluster seismic zones in two categories together with their corresponding security levels, synthetically represented by the design seismic coefficient C (C=0.1 for the 1st Category and C=0.07 for the 2nd Category).

So, the proposal of PFG classification suggested that any new zone should be simply included in the 2nd category, because the difference between the security levels for a classified and not-classified municipalities, is much more important than the difference between the corresponding levels for 1st and 2nd category. Then, a suitable classification has been obtained as a way to protect the new reinforce concrete buildings, since the variations of the security level, depending on the vulnerability class, are not considered. The seismic hazard value assigned to the entire administrative territory has been estimated on the territory area of each municipality.

Three of possible criteria to the sites classification, representing different philosophies of protection, are listed here below. The first two are referring to seismic hazard parameters; the last one is also a risk coefficient, being its determination dependent on typical elements of the risk analysis just like: economical costs, expected damage and expected number of dead. All of them depend, though in different ways, on frequency and earthquakes intensity:

- observed maximum intensity criterion: it is relating to the maximum observed macroseismic intensity (I_{max}) recorded in the past. Or better, it is the intensity we could have observed in that point of the territory, given the assumed models. Using this criterion we consider equally dangerous two sites which experienced the same maximum intensity, considered apart from the frequency of the happened events;
- return period criterion: it is based on the intensity assigned to a certain return period $\underline{T}$. More precisely, it is the probability that the intensity could exceed a fixed limit within a certain period. Unfortunately such a criterion moves an uncertainty into the model, being subordinate to the subjective choice of the return period. A comparison based only on the intensity, given a precise return period, takes care of the event frequency, but just in partial way. Indeed we consider equally dangerous two sites with the same value for $I(\underline{T})$, without taking into account that they could be characterized by different intensity values for other return periods;
- marginal cost for saved life criterion: the marginal cost gives a satisfactory answer to the question concerning the distribution of the supporting resources for the earthquake prevention on different sites. This parameter fully considers the frequency of the events, whose intensity values are greater than the first damage threshold. The coefficient C is calculated when a building is designed, so to ensure a certain elasticity degree at a shock. C_{rif} is the same coefficient but in this case it is calculated in reference to a particular site, the one showing the maximum hazard value.

EVALUATION OF SEISMIC CLASSIFICATION PARAMETERS

The seismic classification of Italy is mainly based on the analysis of the effects of historical earthquakes, and the macroseismic intensity was selected as ground-motion parameter for the seismic hazard assessment. This restricted the analysis essentially to macroseismic data and suitable probabilistic methods were used for the elaboration.

Let us study the assumptions made for a simplified model, which will enable us to calculate the three seismic classification parameters.

Seismic Hazard

We assume the hypotheses that the earthquake occurrence follows a stationary Poisson process with parameter ρ and that the relationship between the intensity decay and the distance as known. Then we can compute, using some methodology [3] (*), the local intensity distribution $F_s(I_s)$ at the site. If we restrict our interest in the case when the local intensity is not less than to a threshold I_{st} ($I_s \geq I_{st}$), then we can indicate the conditional distribution as:

$$1\text{-}F_s(I_s) = \Pr\,[I_s \geq i \mid I_s \geq I_{st}]$$

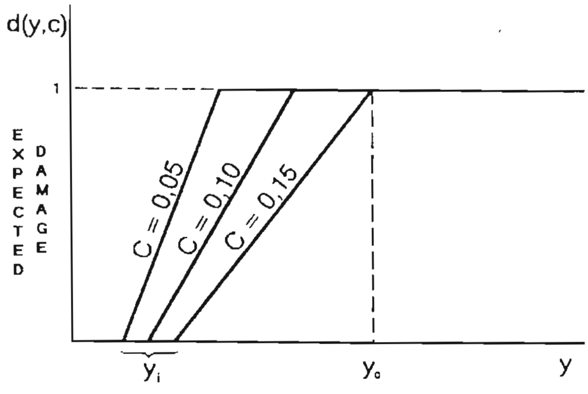
We assume to define the return period $T_s(I_{st})$ as the inverse of the annual probability that an event may occur with the intensity $I_s \geq I_{st}$. That is

$$T_s(I_{st}) = 1\,/\,[1 - \exp(-\,\rho\,\Pr\,(I_s \geq I_{st})] = 1\,/\,[1 - \exp(-\,\rho\,(1 - F_s(I_s)))]$$

Therefore, the seismic hazard relating to the considered site, is completely defined by the relationship between I_{st} and the return period T_s, being T_s the average intercurrence time between two events, given $I_s \geq I_{st}$. So, we may remark that on given assumptions, the earthquake occurrence of events $I_s \geq I_{st}$ follows a stationary Poisson process whatever is the value for I_{st}.

Seismic Risk

In case we consider that the earthquake effects on buildings are depending on a sole intensity parameter I_s, this value could also be transformed into the maximum acceleration parameter y, when this relation $\ln y = a\,(I_s) - b$ is assumed. Then, we get the distribution function $F_s(I_s)$ turned into $F_y(y)$. When an earthquake occurs, both the cumulative probability $F_y(y)$ and the density probability $f_y(y)$ are known and independent of the past and from the previous earthquake data.



$$d(y,c) = \begin{cases} 0 & y \leq y_i \\ \dfrac{y - y_i}{y_c - y_i} & y_i < y \leq y_c \\ 1 & y > y_c \end{cases}$$

Fig. 2 Function damage versus acceleration.

(*) The PFG used the Gumbel I type distribution for different periods of earthquake completeness: 1570-1969 ($I_0 \geq$ VIII) and 1770-1969 ($I_0 \geq$ V). The maximum values, at 10 year intervals, for each site obtained using an average attenuation law for the whole of Italy were used to estimate the parameters of the chosen distribution.

The buildings are classified: each class standing for a certain vulnerability degree, which is represented by the relationship between the acceleration y_c, leading to collapse, and the value of the design seismic coefficient C, which defines the intensity of the lateral forces that simulate the effect of the earthquake (see fig. 2).

The direct damage caused by an earthquake starts to be appreciable when y rises to over a certain value for y_i, and it is proportionally increasing as a linear function of y, gradually reaching the value 1, when $y = y_c$. Generally, the threshold y_i and y_c are a function of the design seismic coefficient C when the relations $y_i = \varepsilon + C$ and $y_c = \xi C + \zeta$ are assumed.

As for the risk to human life, it is assumed that life loss will occur just only when a building collapses. In this case, the number of victims will be proportional to the number of people living in that building. The risk to human life is then measured by the quantity V that is the expected number of victims per year per exposed person; this value is also proportional to the number of fallings down.

The total annual cost, including damage and expenses for seismic risk prevention, is given as shown here below:

$$D(C) = D_1 + D_2 + D_3 \qquad (1)$$

where:

$D_1 = Q (h C + \vartheta)$ is the additional cost for building, due to seismic design

$D_2 + D_3 = \psi Q \int_0^{\infty} d (y,C) \ \rho \ f_y(y) \ dy$ is the cost of all future damage

D_2	is the updated value of the direct future damage
D_3	is the updated value of the indirect future damage
Q	is the initial cost for the building per person in the case of design without seismic code
ψ	is a factor for the indirect damage
ρ	is the parameter of the Poisson process, i.e. the expected annual number of events.

It is assumed that the community accepts a 'social discount factor' γ, on whose basis it is possible to evaluate either the present value of a future cost or the splitting of a present cost into instalments.

Thinking in purely economic terms, an optimum value for the design seismic coefficient C can be defined minimizing the expression (1). To avoid absurd consequences it has been suggested [4] that it must be established the maximum price the community is willing to pay to save a human life. Then, the decision depends on the comparison between this maximum price and the marginal cost of a saved life.

The marginal cost of a saved life can be evaluated how illustrated in the figure 3. The increasing of the design seismic coefficient ΔC corresponds to a variation ΔD of the total monetary cost and to a reduction ΔV of the expected number of victims (that is consequently corresponding to a ΔL increase of the number of saved human life). The marginal cost for saved life $\mu(C)$ then can be defined by the ratio $\Delta D/\Delta L$ [4] [5].

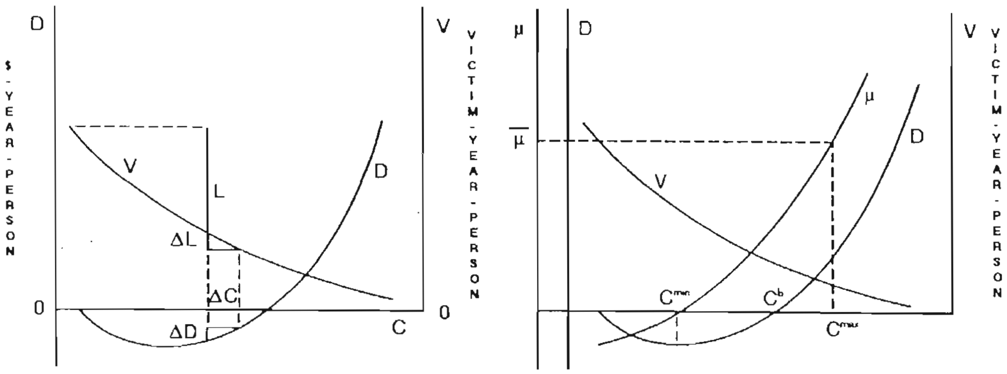


Fig. 3 Definition of the marginal cost of a saved life ΔD/ΔL (left side) and reference values of C for decision making (right side) (from Grandori G. , 1982 [4]) .

The value for μ, corresponding to a given C, is the amount of money necessary to save one more human life by increasing the seismic coefficient starting from the given value of C. It is possible to prove [6] that when the marginal cost for saved life is constant and the same for any site, you get the best distribution of funds for the seismical risk prevention on all the sites. If we suppose that the figure 3 (right side) is representative of a reference site, then the design coefficient must certainly be greater than C^{min}. Designing with $C = C^b$ means having the same monetary costs as if there was no aseismic code, but it would greatly reduce the number of expected victims. Lastly, without entering into details, on the hypotheses above described it is possible to define [4] the marginal cost by the following relation

$$\mu(C) = \frac{\Delta D}{\Delta L} = \frac{Q\left(h + \psi \int_{y_i}^{y_c} \frac{\delta}{\delta C}\, d(y,C)\,\rho\, f_y(y)\,dy\right)}{p\,\xi\,\rho\, f_y(y_c)} \qquad (2)$$

where:

p is a proportional coefficient (the computation of the ratio C/C_{rif} doesn't depend on the coefficient p)

(*) The PFG used the following data: accomodation of 25 m^2 per person; Italian seismic code; Italian market prices for 1980; 1 dollar = 1000 liras; capital investment (social discount rate) at 10 per cent for year; nominal life of the building >50 years. The additional construction cost ha been calculated as much as 7 per cent for C=0.1 The coefficient ψ, that accounts for indirect damage, is assumed $\psi = 1.5$.

THRESHOLD METHOD FOR THE SEISMIC CLASSIFICATION

Now, let's quickly analyse the considerations, which are basis for the evaluation of the threshold levels for the adopted three parameters. The problem solved by the PFG was to find a guideline to keep an homogeneous treatment of the entire Italian territory. Then being confirmed by a statistical analysis starting from the year 1000, the average of the observed maximum intensity on the seismic classified municipalities is equal to the value of I_{max} = VIII MCS and considering that such a value corresponds also to the start threshold for structural damage referring to MCS scale, it has been assumed that value as limit for a municipality to be inserted in 2nd seismic category. Before considering the return period criterion it was discussed about which was the return period to adopt. Finally T=500 was selected, considering that the range of the observed data of the seismic catalogue was based on the last 1000 years information. As all parameters have to produce the same effect, the estimate of the thresholds of I_{500} and C/C_{rif} was fixed in way to nearly approximate the number of the municipalities exceeding the limits with the number of those municipalities satisfying the maximum observed intensity criterion (i.e. see Tab.1).

The consideration and results of PFG make reference to "General Catalogue of Italian Earthquakes" [7] (in the following we refer to it as Catalogue Iaccarino-Carrozzo) because at that time it was the most complete among the official publications concerning the entire country. Now in this paper it has been necessary to refer to it, in order to analyse the seismic classification promulgated by PFG, but we have considered a revised and updated seismic catalogue "Catalogo dei terremoti italiani dall'anno 1000 al 1980" [8] (in the following we refer to it as Catalogue PFG-CNR). The revision of the historical intensity earthquakes was done for several regions, so that the total amount of events increased approximately from 20,000 to 30,000. We can see that the sample average of the maximum observed intensity on the seismic classified municipalities dropped down and consequently so did the thresholds of other two parameters evaluated as explained before. To allow a comparison the start threshold for structural damage I_{max} = VIII can be maintained fixed to compute the other two thresholds; in the following we are referring to that as Catalogue PFG-CNR with 'old' thresholds. In the following table, you can see them according to the said catalogues.

	Cat. Iaccarino-Carrozzo			Catalogue PFG-CNR			Catalogue PFG-CNR		
	I_{max}	I_{500}	C/C_{rif}	I_{max}	I_{500}	C/C_{rif}	I_{max}	I_{500}	C/C_{rif}
	('PFG' thresholds)			('old' thresholds)			('new' thresholds)		
threshold (*)	8.00	7.32	1.90	8.00	7.37	1.60	7.53	6.90	1.18
ther.-σ/4	7.69	6.94	1.55	7.69	6.73	1.24	7.23	6.26	0.82
	strategy VP1			strategy VP2			strategy VP3		

Tab. 1 Threshold values for the three parameters for the different strategies VP1, VP2 and VP3 - (*) The threshold values for C/C_{rif} are multiplied by 10.

For each of the three parameters it is possible to draw a map, where our country appears shared in several zones, in which the assumed values for the parameter are greater than or equal to a given threshold. Using one or the other of these parameters means to get not only conceptual differences, but also practical results considerably unlike each other. The PFG decided to use all the three parameters, considering them equally important.

The adopted methodology, reproduced by us and called 'threshold method', wants the municipalities to be included in the 2nd category, when *at least one of their three parameters is greater than or equal to a given threshold and one of the other two is superior to the threshold value decreased by a quarter of the standard deviation.*

Thresholds	T_{Imax}	T_{I500}	$T_{C/Crif}$	T_{Imax} -σ/4	T_{I500} -σ/4	$T_{C/Crif}$ -σ/4
Classification guide I_{max}	1	-	-	-	1	-
	1	-	-	-	-	1
Classification guide I_{500}	-	1	-	1	-	-
	-	1	-	-	-	1
Classification guide C/C_{rif}	-	-	1	1	-	-
	-	-	1	-	1	-

Tab. 2 Criteria of seismic classification on PFG methodology. The value equal to 1 means that the parameter is greater than or equal to the relative threshold.

As summarized in the Table 2, if, for instance, the value of the parameter I_{max} (classification guide I_{max}) is greater than or equal to the given threshold T_{Imax} and either I_{500} or C/C_{rif} exceed corresponding thresholds T_{I500} -σ/4 and $T_{C/Crif}$ -σ/4, the municipalities have to be classified in the 2nd seismic category.

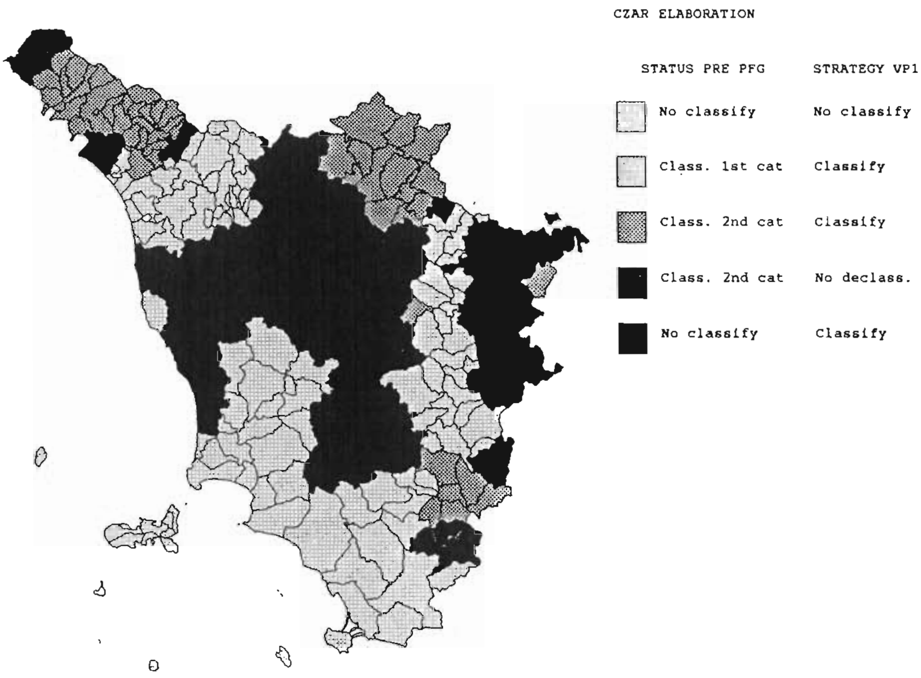


Fig. 4 Seismic classification map of the Toscana region using the criteria PFG with the strategy VP1. This is exactly the actual seismic classification as already promulgated by PFG [1].

It is quite simple to explain the reason for such a procedure: there is no evident justification to prefer one parameter to the other two, as no precise choice has been made about protection laws to be brought into operation in our country. So, we thought it is suitable to the classification every municipality pointed out by any parameter, *only when the values of the other two are not definitely clashing; this means that at least one of them cannot be inferior to the reduced threshold value, as previously said.*

From the logic point of view, each classification strategy can be schematically shared into three different 'blocks' which represent the subsequent steps of the elaboration:

- use of the seismic catalogue
- estimate of the hazard parameters
- assumption of the protection level.

Here below we are analysing three maps of seismic classification on the Toscana region, obtained using the same methodology for the hazard and risk estimate used by PFG, but different for the used catalogue and cut level values (i.e. different strategies). The classification methodology PFG (see Tab. 2) applied to the strategy VP1 (see Tab. 1) exactly reproduces the seismic classification on the Toscana territory (fig. 4) (as already promulgated by PFG [1]) . If it's applied to the strategies VP2 and VP3 (see Tab. 1) it enables two different possible classification ways for the Toscana region. They represent, then, two different protection philosophies.

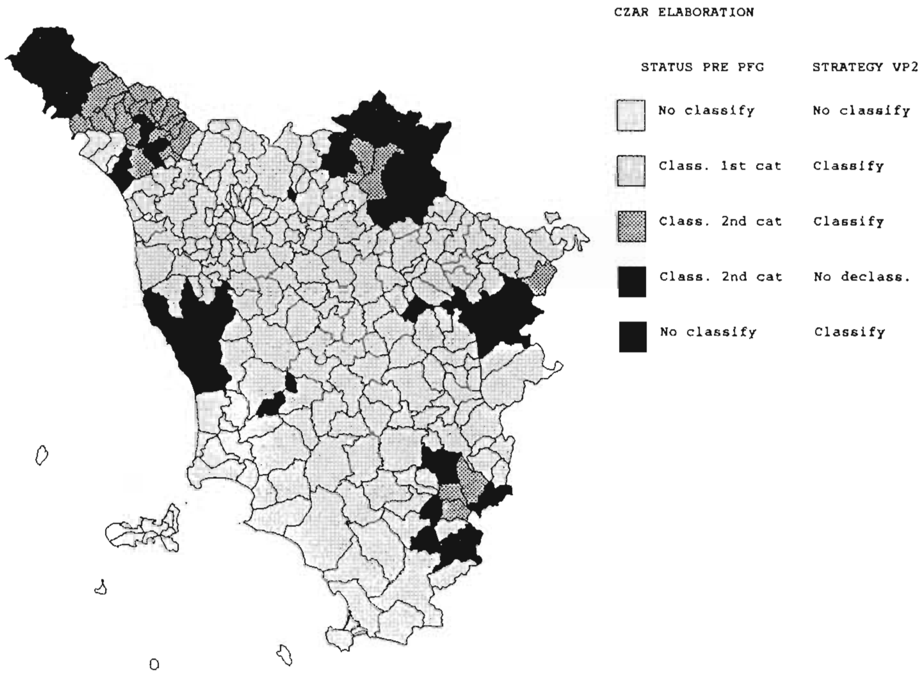


Fig. 5 Seismic classification map of the Toscana region using the criteria PFG with the strategy VP2.

Using a different catalogue and assuming to keep the threshold for I_{max}= VIII MCS as starting damage threshold (as already explained), the number of the municipalities to be classified is considerably reduced (see fig. 5). Whereas, using the cut levels calculated by PFG-CNR catalogue, the reduction is slightly inferior (see fig. 6). This aspect can also be put in evidence by the number of the municipalities, which were already classified in the 'pre-PFG status'; they are reconfirmed by the new methodology, according to the principle of the 'non-declassification' (i.e. a municipality, which is not classified by the PFG methodology but it was in the 'pre-PFG' status, must be anyway confirmed as belonging to the seismic category).

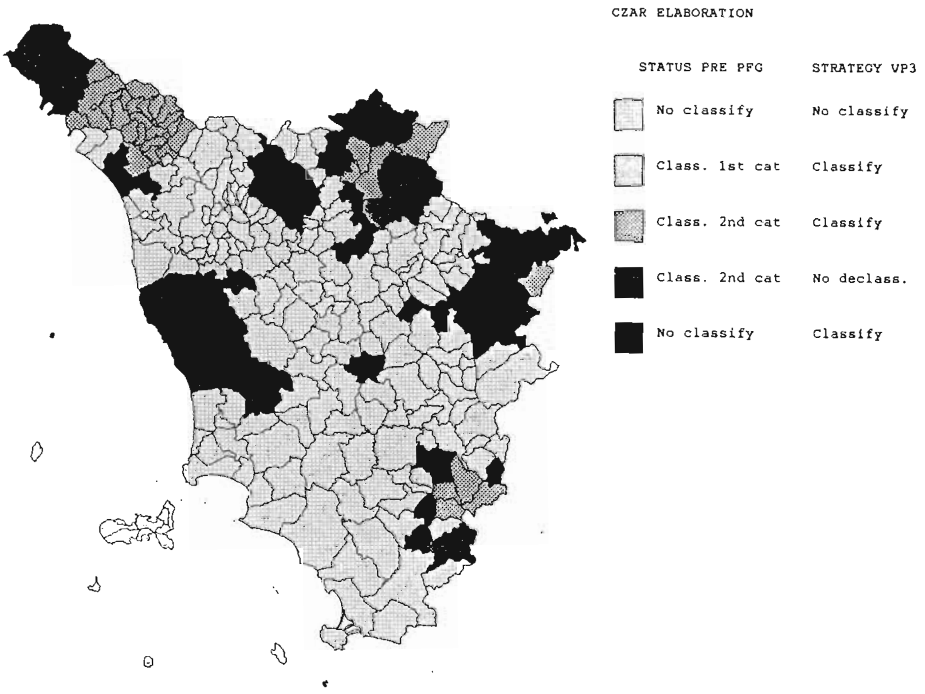


Fig. 6 Seismic classification map of the Toscana region using the criteria PFG with the strategy VP3.

GRADUAL METHOD FOR THE SEISMIC CLASSIFICATION

Through the threshold method we cannot actually quantify how much the municipalities are close to the classification threshold. A second strategy, called 'gradual method', adopted here, starts from the assumption of building up a gradual global insert function F. The used procedure, building up single weight insertion functions f for each parameter, I_{max} , I_{500} and C/C_{rif} is based on the definition of three constraints. The first one states that a weight equal to 1 must be assigned to the maximum intensity value (I = XII) on the scale MCS and at the same time a weight equal to 1 must be given to the value of the design seismic coefficient C corresponding to the municipality chosen as landmark (for Italy, such a municipality is Messina), that is, in other words, $C/C_{rif} = 1$. The 2nd constraint states an insertion weight P equal to 2/3 for threshold values attained by the three parameters (see Tab.1). The last constraint, then, imposes a weight 0 for null values of the three parameters.

We should note that using these above mentioned weight functions not only the information about the threshold value overcoming is produced but also a weight is given, representing how much the threshold value is exceeded. The three above said items enable the estimate of the coefficients of piecewise linear functions, used for insertions in the 2nd seismic category.

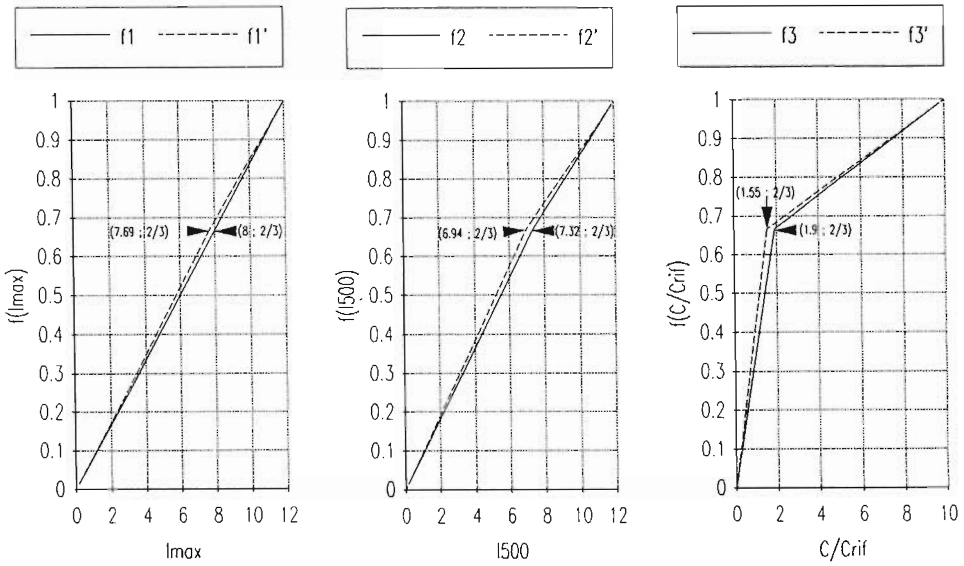


Fig. 7 Weight functions for the 2nd category insertion $f_1(I_{max})$, $f_2(I_{500})$, $f_3(C/C_{rif})$ and $f_1'(I_{max})$, $f_2'(I_{500})$, $f_3'(C/C_{rif})$ relating to full threshold values (continuous line) and to reduced threshold values lowered of $\sigma/4$ (dashed line) according to VP1 strategy.

Generically calling H one of the previously defined strategies (see Tab. 1), we can determine a global measure of classification F ($\underline{X}$, H) , evaluated on the municipality characterized by the vector $\underline{X} = (I_{max} , I_{500} , C/C_{rif})$, by composing the three insertion functions as follows:

$$F(\underline{X}, H) = \mathrm{MAX} \left| \begin{array}{cccccc} 4/6 & 0 & 0 & 0 & 1/6 & 1/6 \\ 0 & 4/6 & 0 & 1/6 & 0 & 1/6 \\ 0 & 0 & 4/6 & 0 & 1/6 & 1/6 \end{array} \right| \left| \begin{array}{c} f_1(I_{max}) \\ f_2(I_{500}) \\ f_3(C/C_{rif}) \\ f_1'(I_{max}) \\ f_2'(I_{500}) \\ f_3'(C/C_{rif}) \end{array} \right| =$$

$$= \mathrm{MAX} \left| \begin{array}{l} 4/6 * f_1(I_{max}) + 1/6 * f_2'(I_{500}) + 1/6 * f_3'(C/C_{rif}) \\ 1/6 * f_1'(I_{max}) + 4/6 * f_2(I_{500}) + 1/6 * f_3'(C/C_{rif}) \\ 1/6 * f_1'(I_{max}) + 1/6 * f_2'(I_{500}) + 4/6 * f_3(C/C_{rif}) \end{array} \right| \qquad (3)$$

The value of the classification function F($\underline{X}$,H) is given by the maximum of component of the vector shown in the relation (3) and if this value is greater than or equal to 2/3 threshold, the municipality can be classified in the 2nd seismic category, according to the adopted strategy. The meanings of the weight coefficients in the classification functions for the three parameters I_{max} , I_{500} and C/C_{rif} can be better explained, considering the functions f as 'classification guide' and the use of the functions f' as constraint to quantify the affirmation *'definitely clashing indication'*.

The adopted model assigns a weight 4/6 when the whole threshold value is overcome, whereas it assigns a weight 1/6 on the reduced threshold value overcoming. The final value of the classification function F ($\underline{X}$ ' , H) corresponding to the threshold value $\underline{X}' = (T_{Imax} , T_{I500} , T_{C/Crif})$, is then equal to 4/6*2/3+1/6*2/3+1/6*2/3 = 2/3; while for the minimum and maximum values of all three parameters the function F ($\underline{X}$, H) could be respectively 0 and 1 (really these values are not reached) (*).

The seismic classification derived from the application of the global insertion function F($\underline{X}$,H), not taking into account some elements included in the 'Seismic Reclassification Proposal' [1], as, f.i. , the principle of non-declassification. In the same way the global insertion function F($\underline{X}$,H), doesn't take into account the information extracted from seismotectonic models. These elements cannot be included in the insertion criteria, either in the PFG threshold method or in the gradual method we have adopted. In fact, these must be outward constraints, later called 'weak links', which, just as they are, will be included in the architecture of the expert system CZAR.

RELATIVE MEASURE FOR DISCREPANCIES ON SEISMIC CLASSIFICATION

A comparison among the classifications based on the 'information sheets' obtained from the previously listed strategies can give the relative measure for discrepancies on seismic classification. If we assume H and E as two different strategies (see Tab. 1), the functions F($\underline{X}$, H) and F($\underline{X}$, E) determine different weights to insert the same municipality in the 2nd seismic category. A comparison among the values at each municipality can give a first comparison between the two strategies effects.

For values in the range between (0 , 2/3), the municipality turns out not to be classified, that is H=0 (E=0); whereas for values between $2/3 \leq P < 1$ it appears as classified H=1 (E=1). We get just two hypotheses (1 = classified; 0 = not classified), mutually disjoint and exhausting the entire probability space. Therefore another way to measure the effects due to two different strategies is the evaluation of the regional conditional probabilities P(E=1 | H=1), P(E=1 | H=0) that can give a relative measure for the discrepancies on seismic classification. It is possible to do that on a regional scale, basing the estimate on frequency and using the formula of the conditional probability.

$$P(E=1 \mid H=1) = \frac{P(E=1 \text{ and } H=1)}{P(H=1)} \quad ; \quad P(E=1 \mid H=0) = \frac{P(E=1 \text{ and } H=0)}{P(H=0)} \qquad (4)$$

The effects of the second strategy E, had on all the municipalities of a region and compared with ones produced by the first strategy H, provide a characteristic average figure.

(*) In the previous paper [9] [10] , it has been used a simplified model, where the function weight doesn't take the reduced threshold values into account. For a proper and detailed representation of each municipality classification it is, however, necessary to use all six insertion functions (see fig. 7) so to better approach the PFG classification criteria (see Tab. 2) .

In practice the regional conditional probabilities are calculated (4) by computing the classified and not-classified municipalities, issued from the two different strategies.

The mathematical algorithm (in the following we refer to it as 'stability index'), proposed in this paper, furnishes a relative measure for discrepancies on seismic classification and it is suggested by the general structure of the Bayes' theorem.

The insertion measure of classification $F(\underline{X}, H)$ (in the following we refer to it as 'prior insertion measure'), issued from the first strategy H , just through the defined formulas (5) and (6), can be updated through the regional conditional probabilities $P(E=1 \mid H=1)$, $P(E=1 \mid H=0)$, $P(E=0 \mid H=0)$ and $P(E=0 \mid H=1)$.

$$F(\underline{X}, H=1 \mid E=1) = \frac{F(\underline{X}, H)\ P(E=1 \mid H=1)}{[1 - F(\underline{X}, H)]\ P(E=1 \mid H=0) + F(\underline{X}, H)\ P(E=1 \mid H=1)} \qquad (5)$$

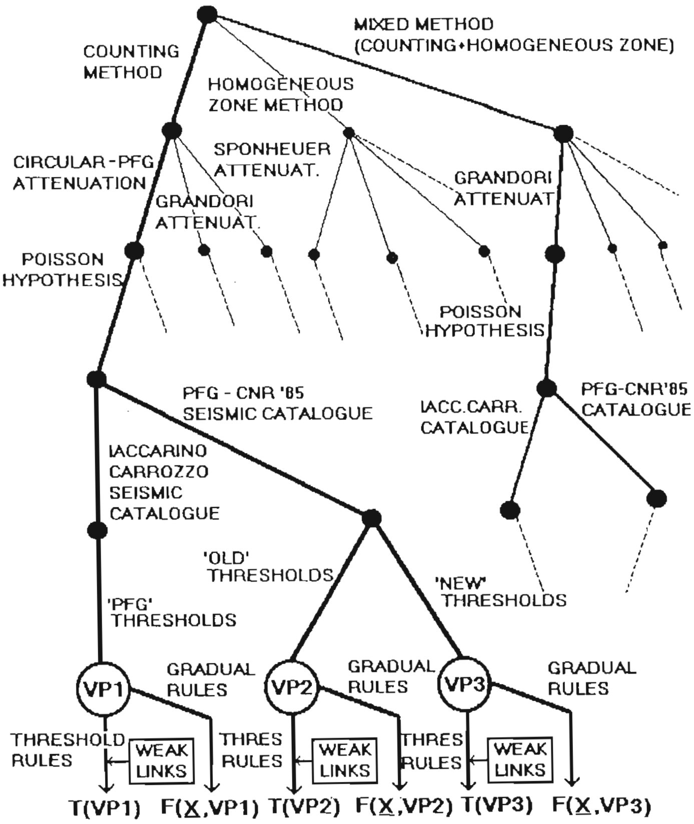
$$F(\underline{X}, H=1 \mid E=0) = \frac{F(\underline{X}, H)\ P(E=0 \mid H=1)}{[1 - F(\underline{X}, H)]\ P(E=0 \mid H=0) + F(\underline{X}, H)\ P(E=0 \mid H=1)} \qquad (6)$$

The functions $F(\underline{X}, H=1 \mid E=1)$ and $F(\underline{X}, H=1 \mid E=0)$ (in the following we refer to it as 'posterior insertion measure') are defined only when the regional conditional probability $P(E=1 \mid H=1)$ is greater than $P(E=1 \mid H=0)$ [i.e. $P(E=0 \mid H=0)$ is greater than $P(E=0 \mid H=1)$]. This hypothesis assumes a certain degree of *similarity* between the different strategies H and E, that must produce a minimum common set of classified municipalities.

SEISMIC CLASSIFICATION STABILITY OF THE TOSCANA REGION

The procedures to estimate Seismic Hazard and also the classification are not univocal [11] , as it has previously shown. It is important to compare the final result on the land and control the solutions through the above defined "stability index" of the seismic classification, varying the different procedures. In fig. 8 some of the possible procedures for the seismic hazard evaluation are represented in a tree structure, where their relating parameters I_{max} , I_{500} , C/C_{rif} are consequently estimated. The limits of the hazard assessment, upon which the Italian classification was based, were well known and, therefore, studies for better defining the regional hazard could be done introducing a seismogenic zoning and proper attenuation laws. As new methodology for the seismic hazard assessment, the Mixed Method (**) could be used, as proposed by Grandori et al. (1991) [12] , which seems more stable for seismic classification purposes. Such Mixed Method has actually been implemented on a GNDT package [13] with different option features, but in this paper we will only use the methodology coming from PFG [2] [3] , leading to the strategies VP1, VP2 and VP3. We will discuss here about our approaching through a short application on the Toscana region, without considering the effects of the 'weak links' above defined (fig. 8) .

(**) This procedure estimates the intensity distribution at the site given the intensity distribution of the seismic source and assuming the regional seismicity as uniformly distributed over that zone (method of "homogeneous zone"); on the other hand, by evaluating the intensity at the site for each event of the seismic source through the attenuation law, the number of events for unit time is estimated (method of the "counting").



$$T(H)=\begin{cases}0 & \textbf{No Classify}\\ 1 & \textbf{Classify}\end{cases}\qquad F(\underline{X},H)=\begin{cases}0<P<2/3 & \textbf{No Classify}\\ 2/3\leq P<1 & \textbf{Classify}\end{cases}$$

Fig. 8 Scheme of the possible procedures for the seismic classification. The boldface lines represent the part till now implemented in the expert system CZAR.

Numbers of the successes and failures			
Strategy VP1		Strategy VP3	
H = 1	H = 0	E = 1	E = 0
161	126	86	201

Numbers of coincidental events on VP1 - VP3		
	E = 1	E = 0
H = 1	74	87
H = 0	12	114

Tab. 3 The numbers of the successes H = 1 (E = 1) and failures H = 0 (E = 0) are relating to the seismic classification event, using the strategies VP1 and VP3. The number of the concordances and discordances on each municipality are showed on the right table, considering VP1 as first strategy.

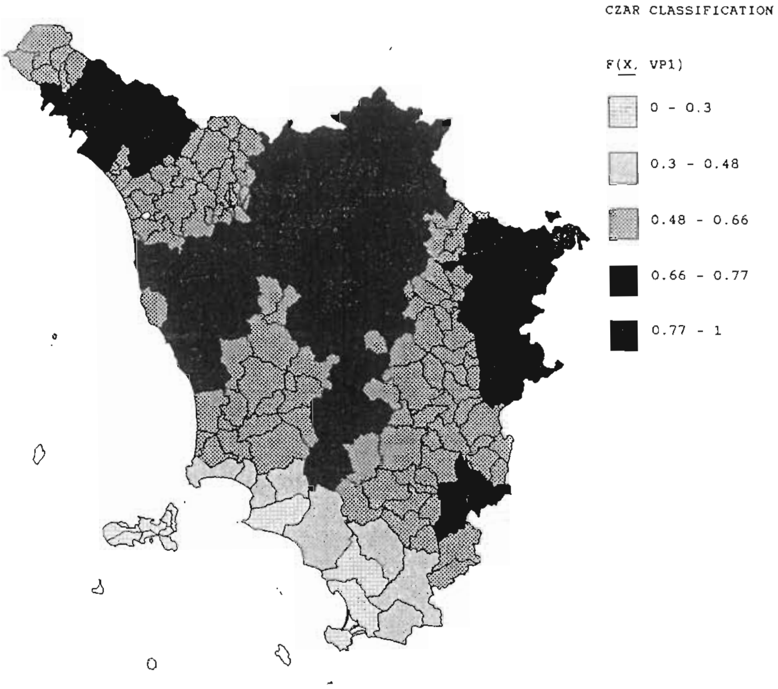


Fig. 9 Seismic classification map of the Toscana region using the function F(X ,VP1).

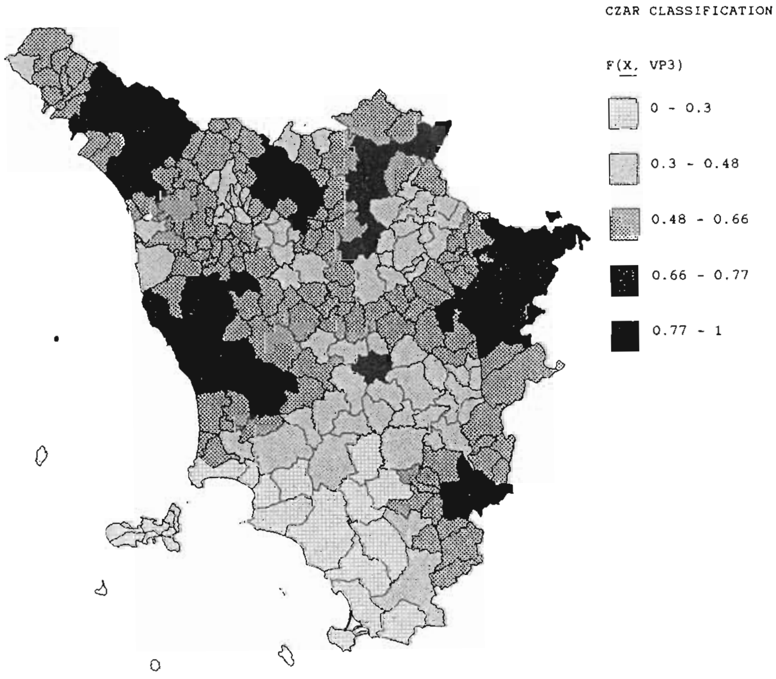


Fig. 10 Seismic classification map of the Toscana region using the function F(X ,VP3).

The proposed approach provides a practical tool to measure how the different procedures are critical on the seismic classification. Applying the 'stability index' to the Toscana region, we check the behaviour due to the use of two different catalogues and different levels of protection. If the two strategies VP1 (Iaccarino-Carrozzo Catalogue & 'PFG' thresholds) and VP3 (PFG-CNR Catalogue & 'new' thresholds) (see Tab.1) are used for the seismic classification, it is possible to estimate two different insertion measures F(X,VP1) (see fig. 9) and F(X,VP3) (see fig. 10) , which produce two different gradual maps of seismic classification.

	P(E=0 I H= 0)	P(E=0 I H=1)	P(E=1 I H=0)	P(E=1 I H=1)
VP1 - VP3	0.90476	0.54037	0.09524	0.45963
VP3 - VP1	0.56716	0.13953	0.43284	0.86047

Tab. 4 The regional conditional probabilities are computed (4) for the Toscana region, using two different combination of the strategies VP1 and VP3.

The computation of the regional conditional probabilities is given (see Tab. 3 and Tab. 4) by the steps previously defined. We can note that the hypothesis of *similarity* is satisfied because P(E=1 I H=1) > P(E=1 I H=0) [i.e. P(E=0 I H= 0) > P(E=0 I H=1)]. The insertion measure F(X , VP1) is considered as 'prior' measure, while the 'posterior' one, considered as the modified insertion measure F(X, VP1 I VP3=1) and F(X, VP1 I VP3=0), can be calculated through the relations (5) and (6) , using the conditional regional probabilities (see Tab. 4). A graphical representation is showed as follows:

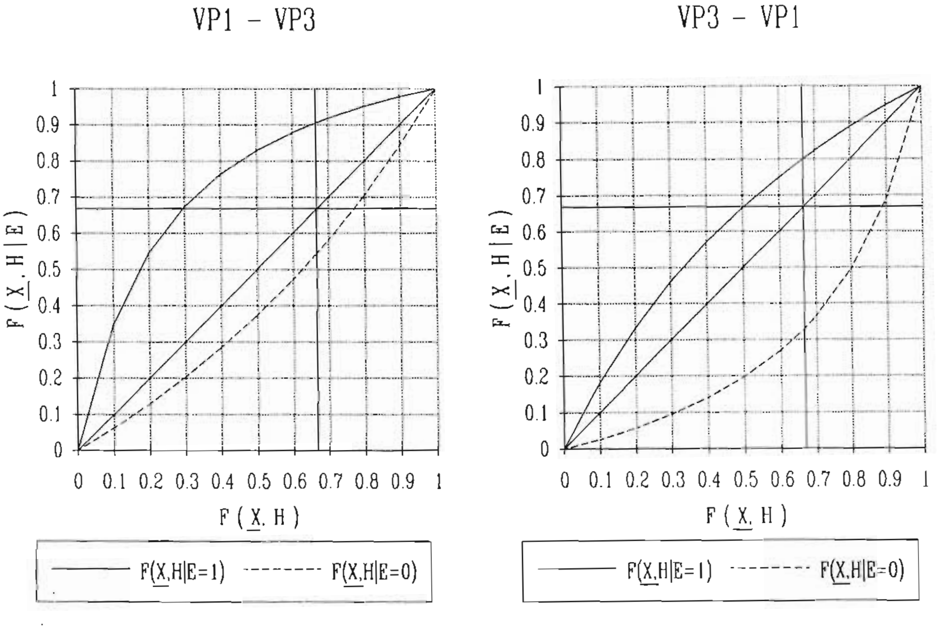


Fig. 11 Graphical representation of the relations F(X, VP1 I VP3=1) and F(X, VP1 I VP3=0).

The diagonal represents the case when the two functions are overlapped and the 'prior' measure of insertion is not modified. This particular situation is true when the following conditional probabilities have the same values : $P(E=1 \mid H=1) = P(E=1 \mid H=0)$ and $P(E=0 \mid H=0) = P(E=0 \mid H=1)$. This means that the two strategies determine the minimum common set of classified municipalities, which doesn't satisfy the hypothesis of *similarity* . If we consider this case as reference point, we can understand the meaning of the other two lines $F(\underline{X}, VP1 \mid VP3=1)$ and $F(\underline{X}, VP1 \mid VP3=0)$. The continuous line is representing the modified insertion (or 'posterior') when the second strategy E classifies the analysed municipality; on the contrary with the dashed line we are considering the non classification case.

We define the ' stability index ' through the composition of the relations $F(\underline{X}, VP1 \mid VP3=1)$ and $F(\underline{X}, VP1 \mid VP3=0)$ applied to the Toscana region, where the values $F(\underline{X}, VP1 \mid VP3=1)$ and $F(\underline{X}, VP1 \mid VP3=0)$ are assigned to a municipality, respectively when $F(\underline{X}, VP3 = 1)$ or $F(\underline{X}, VP3 = 0)$. In this case (fig. 12) we can see the 'stability index' represents the classification level through a poor graded scale; almost all municipalities are included in the range just under the classification level (0.48 - 2/3).

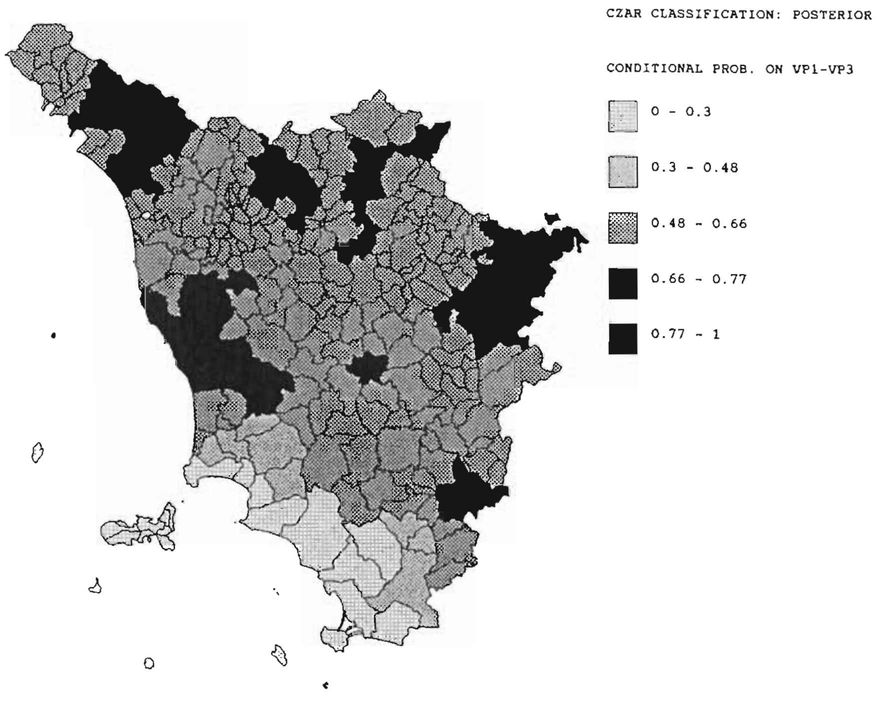


Fig. 12 The 'stability index' through the composition of the relations $F(\underline{X}, VP1 \mid VP3=1)$ and $F(\underline{X}, VP1 \mid VP3=0)$ applied to the Toscana region. The conditional probability are computed on VP1 - VP3.

A better contrasting effect can be noticed in the fig. 13, where the prior insertion measure $F(\underline{X}, VP3)$ (in fig. 10) and the conditional probabilities, are calculated on VP3 - VP1. Analysing this 'stability index', we can see the darker zoning as the area where the relative measure on seismic classification is greater and where, using the regional conditional probability (taking the information from the strategy VP1), the modified insertion measure is increasing in connection with the original 'prior measure', represented here by F(X , VP3).

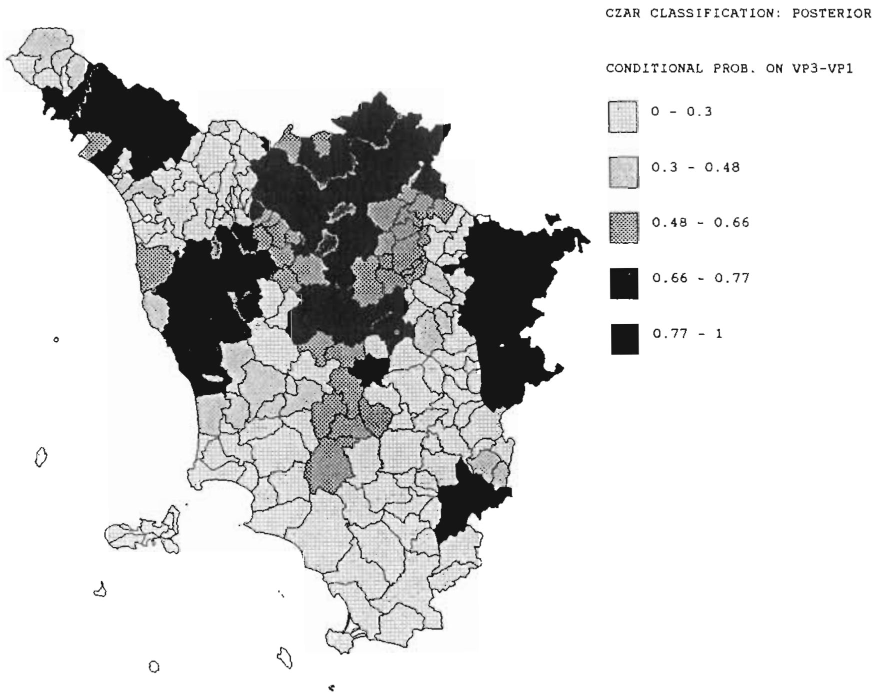


Fig. 13 The 'stability index' through the composition of the relations F(X, VP3 | VP1=1) and F(X, VP3 | VP1=0) applied to the Toscana region. The conditional probability are computed on VP3 - VP1 .

ARCHITECTURE OF EXPERT SYSTEM

The module of the expert system CZAR (Classificatore Zone A Rischio) is built using a commercial expert system shell Nexpert. Nexpert Object (Neuron Data Inc.) [14] is an expert system shell, which includes a set of operators, a rule-based programming language, a variety of editors for rule and data structures, and a sophisticated graphical means for displaying rules and data. Nexpert uses an object-oriented data structure, which gives it great power to represent complex relationships between data. A rule in Nexpert consists of conditions, a hypothesis, and a series of actions. If the conditions of a rule are all true, the hypothesis is set true and the actions are executed. The conditions and actions are operations, such as retrieving data, checking values of variables, and writing files. Another important feature of Nexpert is the powerful interface with external world at the inferential module. In fact from inside this shell it is possible to make a link with common useful data-base (in our case we have used the DBIII+ data base format) , maintaining large amount of static information outside the knowledge base. With efficient operation names 'Retrieve' and 'Write' , data from more commercial data-base and worksheet package are get in-out.

Knowledge base

The combination of the rule and object networks is called knowledge base. The order, in which the rules are executed, is determined by Nexpert's inference engine and is based on either backward and forward propagation strategies. This order is opportunistic and depends on information available at the time of evaluation for each rule. The Nexpert's object-oriented representation reproduces the external world using a formalism based on classes, objects and properties hierarchically linked.

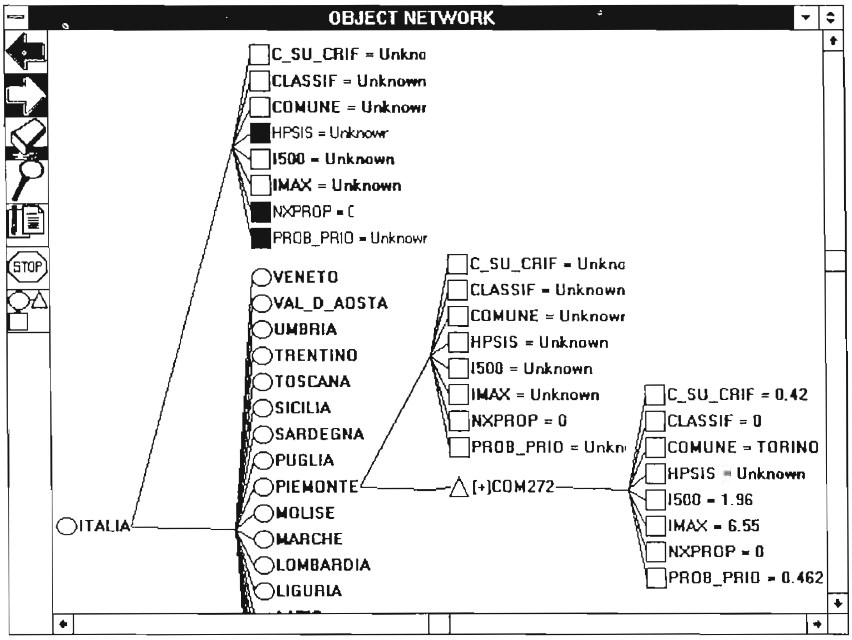


Fig. 14 The Nexpert data structure of Italian territory is represented by a network where the classes (circles) are the regions, the objects (triangles) are the municipalities and the properties (squares) are the parameters.

The graphical interface

The graphical interface offers an immediate and friendly interaction between the final user and the Expert System, where 'interaction' means to give the start to the Expert System and make possible an easy interpretation of the computed results.

Instead of showing numbers which is not allowing a good and immediate comprehension, we represent the relationships among results by a cartographical support through different coloured maps, enabling the final user to compare the various feasible procedures and finally to verify the 'stability index' of the yielded results. When the implemented classification strategies are varied, we find out that the updating activity on the maps reveals to be an useful tool to the achieved outcome estimate.

The graphical interface is developed in MS-Windows 3.0 environment with the support of Software Development Kit by MicroSoft Inc., carrying out an application with characteristics of multitasking environment on Personal Computer and giving the advantages of a Graphical User Interface (GUI) , such as, for instance, the management of multiple windows, menus, buttons, graphical libraries for design and the mouse use. This environment has not, therefore, been chosen only for its well known user friendliness, for its versatile properties and for being so widely diffused in the Personal Computer world, but also because the Nexpert Object user interface is developed under Ms-Windows package.

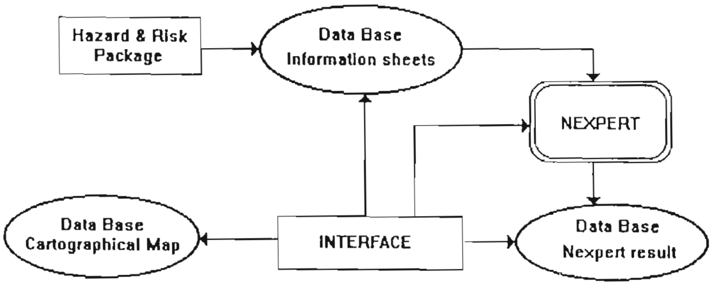


Fig. 15 General Software Architecture of the Expert System CZAR.

There are three Data Bases connecting the interface: the first one is a cartographical data base, in which there is a digitized map of Toscana region; in the second one we have seismic data calculated by the 'Hazard and Risk' package (information sheets containing the parameters I_{max} , I_{500} and C/C_{rif}). The third data base contains the results elaborated by the Expert System: the values of the classification function and the 'stability index', or the composed posterior insertion measure, relating to each municipality of the regional territory.

Besides, there are also some other information that cannot be graphically visualized: the interface enables the interrogation of single elements, the municipalities in our case, for a dynamic consultation action of these data. Selecting through the mouse the determined municipality point to be questioned, a retrieving function is started inside the data base, looking for additional information, such as the number of inhabitants, the geographical co-ordinates, the values extracted from a catalogue and the values elaborated by CZAR. These data appear inside an overlaying window at the screen bottom.
Another task of the interface is starting up the CZAR Expert System so to put under our control the whole working cycle: from the results generation to their own visualization. The activation of an Expert System originates from a menu and the parameters, required by the Expert System during the computation, are given by the user through Dialog-Boxes, guiding his choices about the strategy to be followed by the Expert System.

ACKNOWLEDGEMENTS

The authors are very grateful and heartily thank Prof. V. Petrini for his suggestions on this research; Dr. R. Rotondi for useful discussions had with her and both of them for the last reading of the work. This work is supported by grants from GNDT and partially from CNR.

REFERENCES

[1] Petrini V., Coord. (1980): "Proposta di riclassificazione sismica del territorio nazionale", publ. 361, ESA, Roma, 83 pp.

[2] Petrini V. (coordinator) (1981): "Carta della pericolosità sismica d'Italia", CNR-Progetto Finalizzato Geodinamica, publ. 361, ESA, Roma.

[3] Gruppo di Lavoro Scuotibilita' (1979): "Carte preliminari di scuotibilita' del territorio nazionale", CNR- Progetto Finalizzato Geodinamica, publ. 227, ESA, Roma, 25 pp.

[4] Grandori G. (1982) "Cost - benefit analysis in earthquake engineering", Proced. of the 7th European Conference on Earthquake Engineering, September 20-25-1982, Athens-Greece.

[5] Grandori G. & V. Petrini (1977): "Comparative analysis of the seismic risk in sites of different seismicity", Int. J. Earth Eng. and Struct. Dyn., vol.5.

[6] Grandori E. & Grandori G. (1977): "An application of decision theory to seismic zoning", Proced. 6th W.c.E.e., New Delhi.

[7] Carrozzo M.T., G. de Visentini, F. Giorgetti and Iaccarino E. (1973): "General catalogue of Italian earthquakes", report CNEN,RT/ Prot.(73).

[8] Postpischl D., 1985; "Catalogo dei terremoti italiani dall'anno 1000 al 1980", Quaderni de "La Ricerca scientifica", n. 114-Vol. 28, CNR.

[9] Grazioli A., Zonno G. (1988) "Nuove tecnologie per l'analisi dei parametri di Pericolosità Sismica: prime applicazioni", Atti del 7° Convegno del GNGTS, Roma, 30 novembre - 2 dicembre 1988.

[10] Meroni F., Zonno G. (1990) "Stato di implementazione del prototipo 'CZAR' in Nexpert", Atti del 2° Workshop *Informatica e Scienze della Terra,* Sarnano, ottobre 1990.

[11] Mayer-Rosa D., Schenk V., Ed. (1989): "TERESA", Special Issue of Natural Hazards, Vol. 2, Nos. 3 & 4.

[12] Grandori G., Garavaglia E. and Petrini V. (1991): "Analysis of some of the uncertainties in the statistical elaboration of historical data for local seismic hazard evaluation", Città del Messico, ICASP6, june 1991.

[13] Petrini V. (1991): "Software package GNDT for Seismic Hazard and Risk analysis", unpublished report.

[14] Neuron Data Inc. (1990): "Nexpert Object r.2 MS-DOS", manuals.