BOLLETTINO DI GEOFISICA TEORICA ED APPLICATA VOL. 42, N. 1-2, PP. 19-39; MAR.-JUN. 2001

Probabilistic seismic hazard analysis: zoning free versus zoning methodology

S. MOLINA[1], C.D. LINDHOLM[2] and H. BUNGUM[2]

[1] *Dpto Ciencias de la Tierra y del Medio Ambiente, Universidad de Alicante, Spain*
[2] *NORSAR, Norway*

(Received February 20, 2001; accepted June 9, 2001)

Abstract - A Kernel method for probabilistic seismic hazard computation, without definition of source zones, has been compared with the classical Cornell-McGuire method. The Kernel method used here amalgamates statistical consistency with an empirical knowledge basis (earthquake catalogue), and incorporates parameters describing the structured character of the earthquake distribution. Statistical Kernel techniques are used to compute probability density functions for the size and location of future events, and linear trends based on geological or seismic information, and uncertainties in magnitude and epicentral location for each earthquake are incorporated in the statistical estimates. The Kernel method has been explored in terms of the available parameterisation, particularly aimed at understanding how tectonic knowledge of a region and expert judgement can be used in reasonable and statistically significant ways. The two computation methods were compared with synthetic data and with seismicity catalogues (from Norway and Spain). When using real data it was found that the Kernel method generally yields lower hazard results than the Cornell-McGuire approach. More specifically, it is found that the difference between the two methods increases with increasing deviation in the catalogue from the self-similarity assumption implied by the Gutenberg-Richter relationship. The Kernel method has features that circumvent some of the simplification drawbacks of the conventional zoning methods, and it has potentials to develop into a feasible alternative for hazard computation.

Corresponding author: S. Molina; Facultad de Ciencias, Dpto Ciencias de la Tierra, Universidad de Alicante, 03690 Alicante, Spain, phone: +34 965903987; fax: +34 965907552; e-mail: sergio.molina@ua.es

1. Introduction

The foundations of engineering seismic hazard analysis were established by Cornell (1968) who recognized the need for seismic hazard to be based on a method accounting for the intrinsic uncertainty associated with earthquake occurrence. The method, incorporating magnitude-frequency distributions and variability of ground motion relations, is still the basis for most current hazard computations.

Computers were the ideal tools for making statistical computations, and with Algermissen and Perkins (1973) and McGuire (1976) hazard computation was formalized into a scheme based on zones with distinct seismic activity parameters. Only shortly after, McGuire (1978) also published a computer program to model seismic hazard from faults in a statistical similarly consistent way. Notwithstanding the development made over the last 20 years, and the large number of computer codes addressing seismic hazard problems, many of the basic assumptions and algorithms have remained virtually unchanged. In particular, we should draw attention to the definition of source (seismicity) zones, which are usually defined with static activity parameters (in itself a gross simplification), and to the integration principle where contributions from both near and far are added to constitute the hazard. The approaches based on these principles are efficient and they satisfy the need for statistical consistency, however, some of their underlying assumptions may need to be discussed.

Anyone with experience in probabilistic seismic hazard assessment (PSHA) knows that the definition of source zones to some extent is a matter of expert judgement, reflecting that it is often difficult to delineate the zone borders. It is also widely recognized that seismic activity is not homogeneous within a zone, but it is accepted as a fair approximation and simplification for computational purposes. Furthermore, in low activity areas, the activity parameters (e.g. the Gutenberg-Richter a and b-values, and M_{max} as well) can be very difficult to evaluate properly, and in such areas the estimation of hazard at low probabilities (say, below 10^{-4}/ year) becomes very challenging, if not impossible.

The Cornell-McGuire probabilistic approach has its strength in the systematic parameterisation of seismicity and the way in which also epistemic uncertainties are carried through the computations, and into the end results. The simplicity of the computational model has, however, also motivated new approaches.

More recently, alternative PSHA approaches have been suggested, such as through extensions of the zonation method (Frankel, 1995; Frankel et al., 1996, 2000; Perkins, 2000) where multiple source zones, parameter smoothing and quantification of geology and active faults have been successfully applied. The Frankel et al. (1996) method applied a Gaussian function to smooth a-values from each zone, thereby being a forerunner for the later zonation-free approaches of Woo (1996), which we have focused on in this paper. Jackson and Kagan (1999) have also worked along similar lines when they developed a non-parametric method with a continuous rate-density function (computed from earthquake catalogues) used in earthquake forecasting.

Any seismic hazard analysis is a forecast down to low probabilities, and in some cases very low. The results are, therefore, sensitive to the balance and interaction between data, models

and methods, and one tool that (to some extent) helps delineating such effects is deaggregation. An area where such problems have been discussed extensively is southern California, where the earthquake hazard is high, but the historical record relatively short, and this has called for multidisciplinary approaches where in particular geologic and geodetic data are important (Ward, 1994, 1998; Field et al., 1999). One particular problem recently addressed has been the finding by WGCEP (1995), namely that the region is now subject to a significant earthquake (seismic moment) deficit. This deficit was later explained by Stein and Hanks (1998) as being due to a combination of an undercounting of historical earthquakes and an overprediction by the model. If so, the result is a significantly lower hazard, demonstrating (once again) how deceiving results can be when based on biased or deficient data or models.

The seismic hazard computation method presented by Woo (1996) tries to amalgamate statistical consistency with the empirical knowledge of the earthquake catalogue (with its fractal character) into the computation of seismic hazard. The aim of the present paper is to highlight some important properties of the Kernel method in view of the traditional Cornell-McGuire approach.

2. Computation principles of the Kernel method

The geometry of earthquake epicentres hardly ever satisfies the basic constraint of spatial uniformity presumed by the standard zonation (area source) methods, but rather a spatial distribution with a structured, fractal character. It follows that a seismic source model is most conveniently constructed by means of the epicentres of the earthquake catalogue itself, whenever possible supplemented with geological information on active faults. Inevitably there are, however, uncertainties in the catalogued magnitude and epicentral locations, moreover complicated by the fact that the size and location of future events never fully repeat the past seismicity pattern. In order to incorporate this variability, a smoothing operation has to be performed on the data, rendering statistical smoothing techniques as attractive. Jackson and Kagan (1999) applied a rate density function as a basis for their forecasts, which in turn was based on a spatial smoothing kernel. Woo (1996) likewise applied the kernel function of Vere-Jones (1992), but extended it to also include shape parameters, in the procedure focused on below (the KERFRACT computer code).

The standard probabilistic seismic hazard assessment procedure is based on the computations of the mean number of annual exceedances of a selected ground motion at the site, assuming that the occurrence of earthquakes is a Poisson process. The mean number of annual exceedances of a selected ground motion is then expressed as:

$$\nu(z) = \sum_{i,j} \lambda(M_i)\, P(r_j \mid M_i)\, G(z \mid M_i, r_j)$$

where $\lambda(M_i)$ is the mean number of events per unit of year of magnitude M_i, $P(r_j \mid M_i)$ is the probability that the distance to the site is r_j given an event of magnitude M_i, and $G(z \mid M_i, r_j)$ is

the probability that the ground motion level z will be exceeded, given an event of magnitude M_i at a distance r_j from the site. The Cornell-McGuire procedure of this probability function is applied to each area and fault source n with given parameters S_n, with the exceedances being added up over all the areas and fault sources.

In the KERFRACT procedure the expected annual number of events of magnitude M occurring at a given location is obtained by establishing a spatial activity rate density function from the earthquake catalogue. The kernel can be established in a grid, covering, say, 200 x 200 km, centred on the site, incorporating events from a complete catalogue covering a larger area. By summing over all events in the catalogue, the cumulative activity density is computed for each magnitude range, and a gridded activity rate density is obtained.

A computation based solely on the historical record will often yield a lower bound hazard estimate because of incomplete data, and to counter this, some background seismicity can be added. Background activity can be added at any magnitude level, both to fill in (apparent) holes in a Gutenberg-Richter distribution and to extend the activity between the maximum observed magnitude and the maximum credible magnitude for the area. Geological information (e.g., an active fault) can likewise be incorporated through synthetic events that mimic the seismogenic potential inferred from sources other than the earthquake catalogue.

The smoothing operation (see Woo, 1996 and Vere-Jones, 1992 for details) involves the introduction of a kernel $K(M,x)$, which is a magnitude and distance-dependent multivariate probability density function in which the contribution of each event is inversely weighted by its effective return period, thereby normalizing it to one year. The concept of effective return period (T_i) for an event i of magnitude M with given epicentre is such that at any time in the historical past, there is a likelihood that this event of magnitude M with given epicentre would have been observed. For example, an earthquake with magnitude 4.0 in an inhabited area has a probability of 1.0 of having been observed in the last 100 years, but this probability will decrease if the location of the earthquake is in offshore or uninhabited regions. Consequently, the catalogue is divided in time periods of equal length (e.g. 100 years) and for each time period the probability of detection (p_i) is evaluated. The effective return period for any event is the sum of all time periods multiplied with the detection probability of the period, that is:

$$T_i = 100 \cdot \sum_i p_i$$

The expression of the activity rate density is:

$$\lambda\,(M, x) = \sum_{i=1}^{N} \frac{K\,(M, x - x_i)}{T_i}$$

where $(x\text{-}x_i)$ is the site - earthquake distance for the i'th earthquake, T_i is the effective return period of the magnitude of the i'th earthquake and N is the number of earthquakes in the

catalogue. Various kernel functions can be chosen. The one preferred by Woo (1996) was an anisotropic kernel based on the isotropic kernel suggested by Vere-Jones (1992) and expressed as:

$$K(M, r) = \frac{P}{\pi} \frac{L - 1}{H^2(M)} \frac{1 + \delta}{1 + (\delta / 2)} \frac{c}{o} \frac{o}{s^2} \frac{\phi}{} \left(1 + \left(\frac{r}{H(M)}\right)^2\right)^{-PL}$$

where PL is the power law index ($PL > 1$) also called the "kernel fractal scaling index" which scales the falloff of the kernel probability density function. Recommended values for this parameter are between 1.5 and 2.0, corresponding to a cubic or quadratic decay of the probability density function with epicentral distance. The parameter ϕ is the angle of anisotropy, δ the degree of anisotropy, $H(M)$ is the bandwidth function detailed below and r is the radial separation distance.

A situation frequently observed is that smaller events are spatially clustered, whereas the larger infrequent earthquakes are spatially more disperse, and the bandwidth function $H(M)$ is designed to reflect this degree of spatial clustering of the catalogued earthquakes in a given magnitude range. An exponential representation $H(M) = c \cdot e^{(d \cdot M)}$ was suggested by Woo (1996), as used also in this study, however other representations may also be found appropriate for certain distributions. The c and d parameters are obtained by sorting events in magnitude bins. For all magnitude bins the shortest distance to other events in the same bin is evaluated, and a regression between shortest distance (H) and magnitude (M) within a magnitude bin can be performed. Low values of $H(M)$ for a certain magnitude range implies that these earthquakes are highly clustered, while high values implies less clustering, implying that events may migrate from one position to another more easily. It is shown below how $H(M)$ has a decisive influence on the kernel.

When investigating the seismicity distribution of an area, one may sometimes, on one hand, be confronted with a somewhat amorphous spatial distribution of earthquakes in the catalogue, and on the other hand be tempted to relate the seismicity spatially and generically to mapped faults or fault systems. Through the anisotropy parameter the kernel can be shaped to reflect a degree of correlation with the fault distribution. The parameter ϕ is an angle used to give a higher probability to the movement of the earthquake in a certain lineament direction, while δ is a parameter that modulates the degree of anisotropy. A δ value of zero implies full isotropy (no preferred directivity in the spatial distribution) while a value of 10 or more implies significant anisotropy.

In standard PSHA the earthquake catalogue, which carries the primary empirical information, is representing each earthquake with one single point in space with an assigned magnitude. With the highly varying uncertainty in epicentre location, depth and magnitude assessment, this use of the earthquake catalogue is at its best a simplification, which is useful and necessary. In other situations it is an over-simplification, which may be seriously misleading, and in the Kernel method used here, these uncertainties are incorporated. By working directly on a catalogue where each earthquake entry has an assigned uncertainty in epicentre and magnitude, one needs to assess realistic uncertainties, and these are consistently

carried through the computations and reflected in the shape of the probability density function.

3. Analysis and comparison of methods

The characteristics of an algorithm like the present KERFRACT code are sometimes conveniently evaluated with synthetic data, which is our first approach. In a next step we focus on real data.

Firstly, a synthetic catalogue with a pseudo random epicentre distribution was generated based on a doubly truncated Gutenberg-Richter relationship. The pseudo random catalogue was generated with a spatially inhomogeneous distribution, firstly since inhomogeneous distribution is more in accordance with real earthquake distribution, and secondly because a spatial homogeneous catalogue would yield very similar results for the two PSHA methods used in the testing. With this catalogue, it was possible to compare the results provided by the Cornell-McGuire method with the Kernel method at several sites, and the sensitivity of the results to the input parameters for the kernel function was studied. Synthetic catalogues were computed without and with anisotropy to simulate fault behaviour.

Secondly, the seismic hazard was assessed using the KERFRACT procedure at several sites based on historical seismicity from Norway and Spain, with a short historical earthquake catalogue (from 1657 to 1995, although the majority of the earthquakes occurred over the

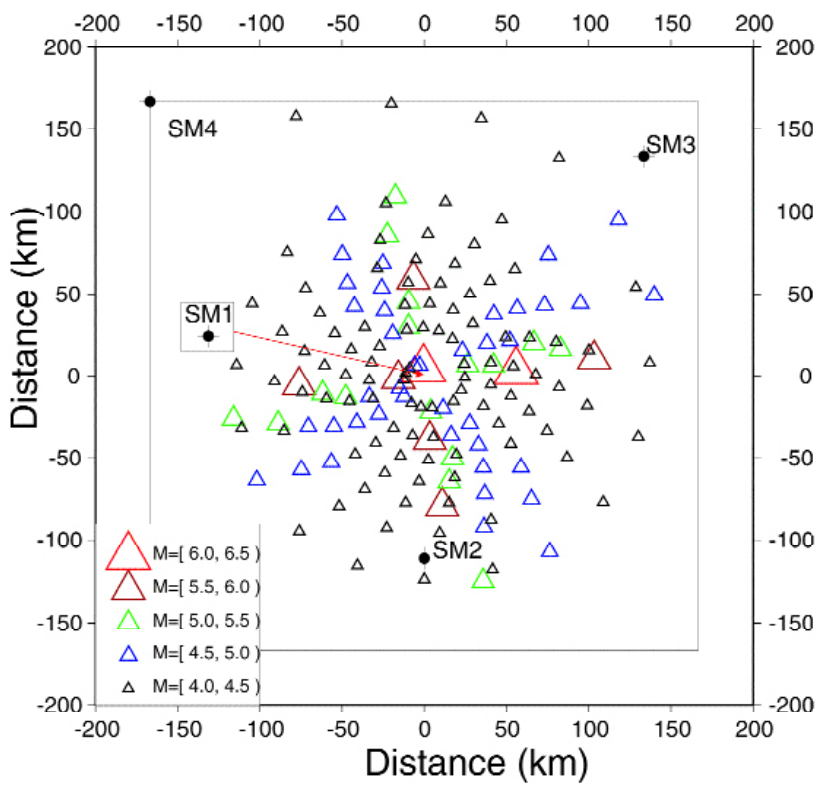

Fig. 1 - The synthetic catalogue generated for testing purposes and the four sites (SM1-SM4) at which seismic hazard was computed. The rectangle indicates the area source boundary used for NPRISK. The catalogue generation values and KERFRACT parameters are shown in Table 1, and the hazard results are shown in Fig. 2.

last 200 years), and a long one (from 1300 to 1998), respectively. The results obtained were compared with previous results (Molina, 1998; Bungum et al., 2000) obtained at these sites using the Cornell-McGuire methodology. One obvious difference between the two catalogues is that the Norwegian catalogue satisfies the Gutenberg-Richter distribution better than the Spanish one (see details below and insets in Figs. 6 and 9), but it was not possible to determine to which extent this may be related to quality problems such as in the intensity-magnitude conversion.

3.1. Synthetic earthquake catalogues

Fig. 1 shows the synthetic earthquake catalogue (165 earthquakes with magnitude ranging from 4.0 to 6.2) generated in accordance with parameters in Table 1, reflecting a doubly truncated Gutenberg-Richter magnitude distribution. Note that epicentres are distributed as $N(0,\sigma)$ around the centre. A source area defined for the subsequent hazard computation with the Cornell-McGuire method (computer code NPRISK; woo, 1994) is indicated together with four sites (SM1-SM4) for which hazard was computed with both KERFRACT and NPRISK. Note

Table 1 - Parameters used to generate the pseudo random seismicity catalogue shown in Fig. 1 (left); the critical parameters used in the KERFRACT hazard computation (right). The Cartesian coordinates of the test sites (SM1-SM4) and the PGA ground motion relation used in the computations are also included.

Catalogue generation parameters		**KERFRACT parameters**	
a	0.33	Fractal scaling (PL)	1.5
b	0.88	Anisotropy	$\delta = 0.0$; $\phi = 0$
Depth	5.0 km	Bandwidth H(M)	12.8 km
M_{low}	4.0	Effective return period	500 years
M_{high}	6.2	Epicenter uncertainty	25 km
Epicenter uncertainty	25 km	Magnitude uncertainty	0.1
Magnitude uncertainty	0.1		
Sites for hazard computation			
x = 0.0; y = 0.0	SM1	x = 133.0; y = 133.0	SM3
x = 0.0; y = -111.0	SM2	x = -167.0; y = 167.0	SM4
PGA relation (cm/s^2); unpublished for U.K.			
$\log (PGA) = 5.72 + 0.59\ M - 1.26 \log (R + 2.13 \exp (0.25\ M))$			

that in this computation the sites were defined only with Cartesian coordinates.

3.2. KERFRACT versus NPRISK results

The seismicity distribution with a concentration around a point is sometimes found similarly in nature, and in a hazard context one possible zonation could be the box enclosing all of the data in Fig. 1. Therefore, this simple data set is useful to illustrate some features of the two PSHA methods, and Fig. 2 shows the hazard results obtained with the KERFRACT and with

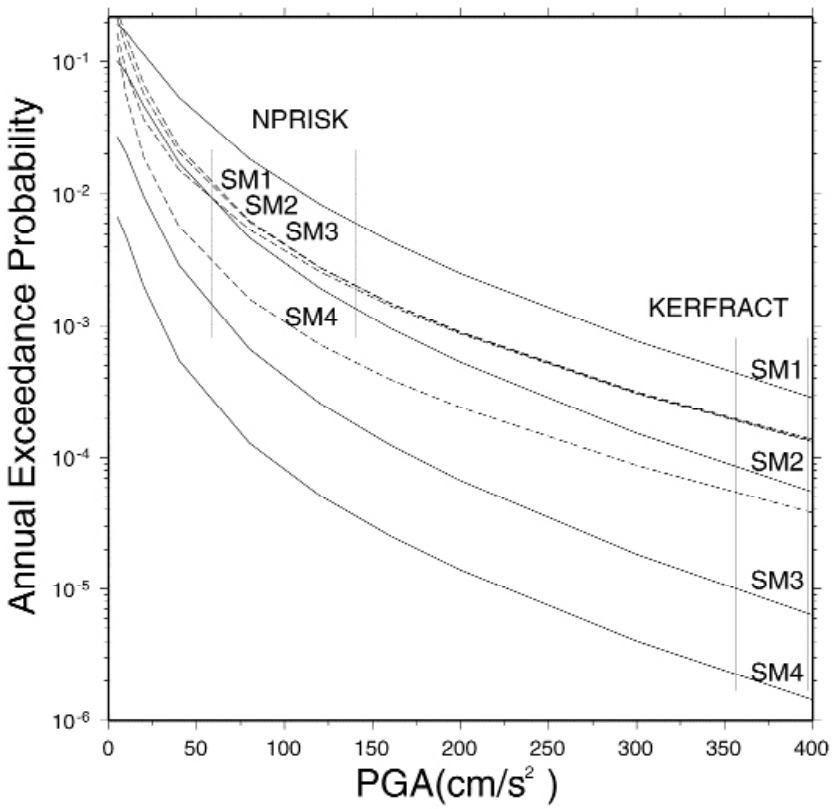

Fig. 2 - Comparison of the seismic hazard results obtained by KERFRACT (solid lines) and the results obtained by NPRISK (dotted lines) for the synthetic data and sites presented in Fig. 1.

NPRISK for the four sites in Fig. 1.

The most apparent observation is how the KERFRACT hazard results reflect the distance between the site and the earthquake "centre of gravity", and how poorly the NPRISK results reflect the same. For this case, Fig. 2 demonstrates that KERFRACT provides a more realistic (obviously referred to the synthetic data) hazard distribution than NPRISK, naturally because the non-uniformity of the seismicity in this case violates the homogeneity requirement of the Cornell-McGuire method. The behaviour of NPRISK is, as expected, with constant area parameters, so that the seismic hazard drops only when very close to the zone border. Outside the source zone, the NPRISK hazard results decline with distance, as expected. As substantiated by the real data computations (below).

3.3. Sensitivity of the results to three kernel-shaping parameters

In a PSHA study the practicing seismologist will usually apply his experience on parameter sensitivity, and his specific seismotectonic knowledge of the region. To this end KERFRACT provides the user with three parameters, through which the expert knowledge can be quantified with direct impact on the computed hazard, namely the fractal scaling parameter, the bandwidth

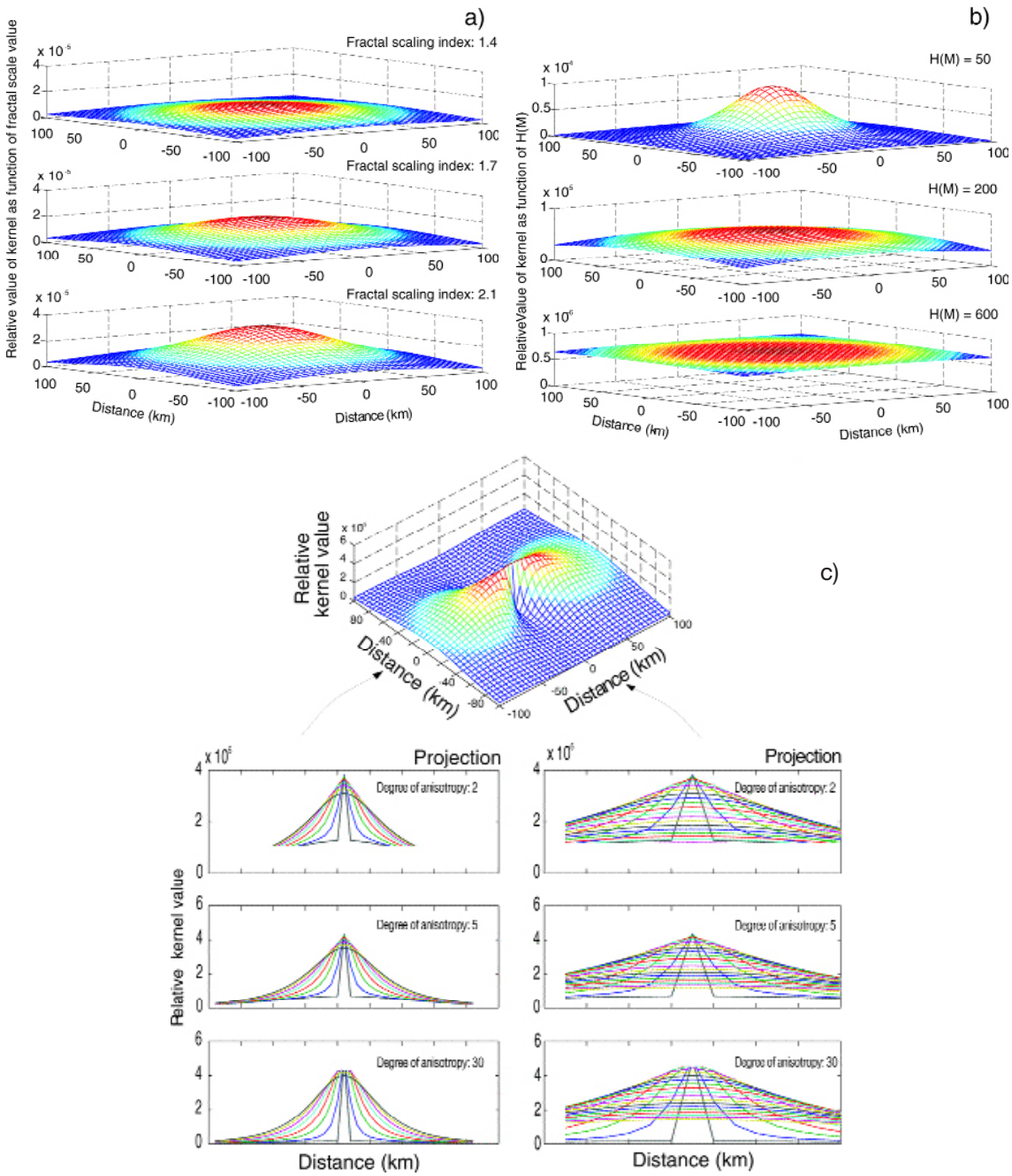

Fig. 3 - Sensitivity of the Kernel function to the three shaping parameters: a) fractal scaling *PL*; b) the bandwidth function *H*(*M*); and c) anisotropy degree δ. See also Table 2 and observe in particular the ratios, which are difficult to evaluate directly from the figures.

function and the degree (weight) of anisotropy (Fig. 3 and Table 2).

In addition to the epicentre uncertainty (which is individually assigned to each earthquake in the catalogue) the fractal scaling parameter is important for the shape of the kernel, defining the dynamics of the system. High values concentrate the probability function around historical

Table 2 - Variation of three KERFRACT parameters used to quantify the users' seismotectonic concept (see also Fig. 3). The Max/Min ratio refers to the ratio between maximum and minimum kernel values.

Parameter	Values	Max/Min ratio
Fractal scaling (PL)	1.4; 1.7; 2.1	4.7; 6.6; 10.0
Bandwidth function (km) (H(M))	50.0; 200.0; 600.0	44.5; 2.0; 1.1
Anisotropy degree (δ)	2.0; 5.0; 30.0	3.1; 6.9; 29.8

seismicity and vice versa. Fig. 3a and Table 2 show that the effect is significant with a peaking kernel for higher values of fractal scaling. The degree of earthquake clustering is very high in some plate margin environments and usually lower in intraplate regions. The bandwidth function is designed to reflect this degree of clustering into the kernel shape, and designed such that also variation in clustering with magnitude can be introduced too. We used an exponential form, but other functions may also be applied. As seen from Table 2, (the Max/Min ratio), and Fig. 3b the influence on the kernel can be quite strong, and through the differentiation with magnitude it is, for example, possible to generate a kernel that reflects a high degree of clustering of the smaller earthquakes and a lower degree of clustering for the larger and rare

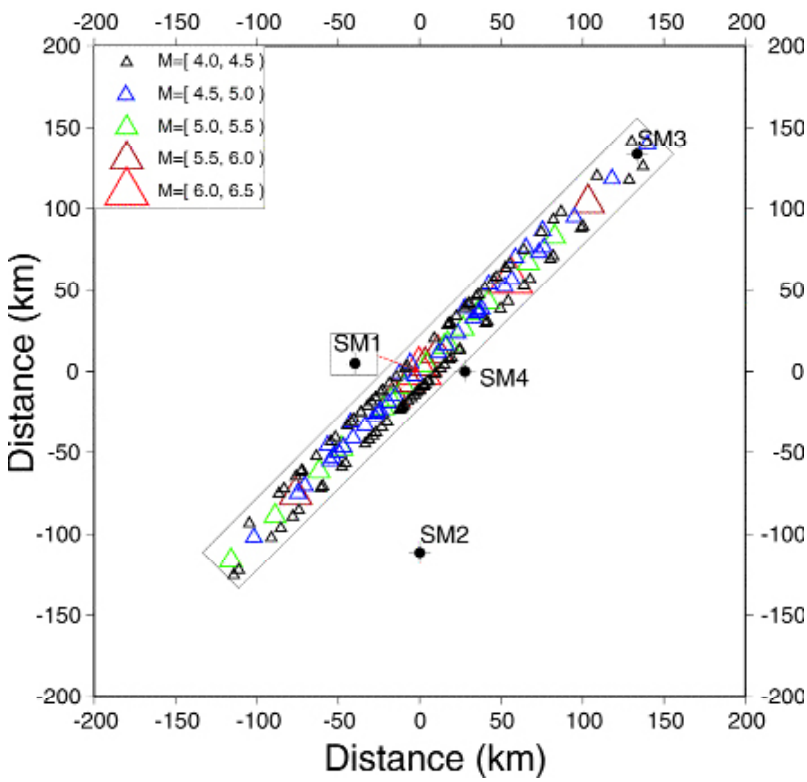

Fig. 4 - The synthetic catalogue simulating fault behaviour and the four sites (SM1-SM4) at which seismic hazard was computed. The rectangle indicates the area source boundary used by NPRISK. The catalogue generation values and KERFRACT parameters are shown in Table 3, and the hazard results are shown in Fig. 5.

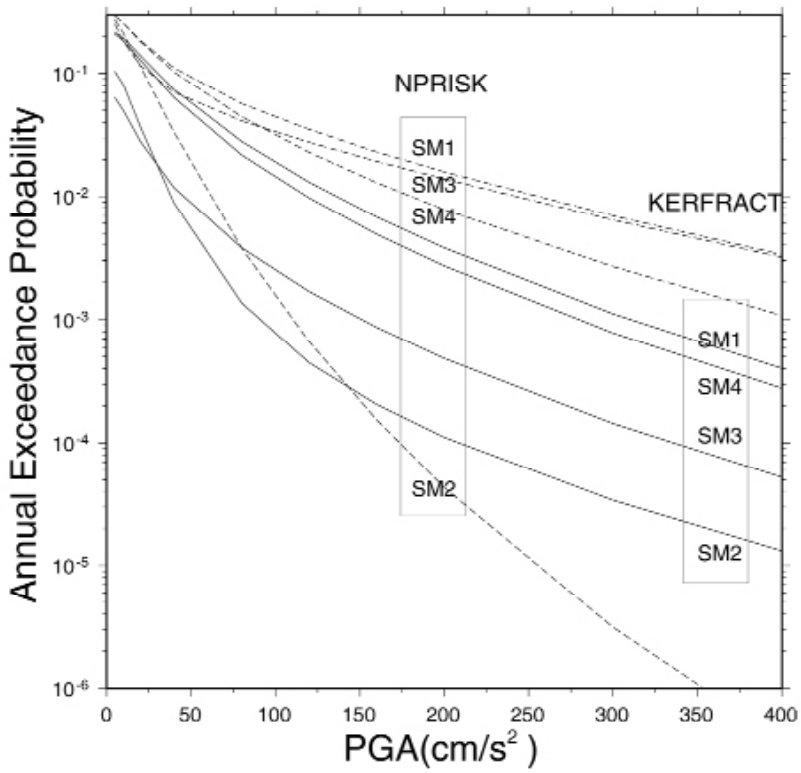

Fig. 5 - Comparison of the seismic hazard results obtained by KERFRACT (solid lines) and the results obtained by NPRISK (dotted lines) for the synthetic data and sites presented in Fig. 4.

events.

The anisotropy parameter (δ) is introduced to depict lineations in seismicity when active faults are assumed to be the earthquake sources. Quite frequently there may be reasons to associate certain earthquakes with certain mapped faults, albeit with varying degree of confidence. To this end, each earthquake in the catalogue can be assigned an alignment (corresponding to fault alignment) and a corresponding weight (degree of confidence), where a high value will shape the kernel to a higher probability along the strike direction of the fault. Fig. 3c shows profiles of the kernel across and along the fault trace for varying degrees of confidence (weights) displayed in Table 2. The higher confidence makes the kernel into a sharply peaking function along the fault trace as observed from Fig. 3c with maximum/minimum ratios ranging from 3 to 30 (Table 2).

Well defined seismicity often follows mapped structures or faults. Fig. 4 shows an example of a synthetic catalogue that reflects activity along a 380 km long and 30 km wide structure, where the synthetic data has a higher activity rate per unit area than in Fig. 1, but still has its concentration at the centre with lower activity at the ends. The associated generation parameters and the KERFRACT hazard parameters are shown in Table 3. Again the seismic hazard was computed at four sites (SM1 - SM4) with both NPRISK and KERFRACT, and the results are shown in Fig. 5. It is seen that the basic features of Fig. 2 are repeated, namely that KERFRACT provides relative hazard estimates in accord with distance from the activity centre, contrary to NPRISK. The observation that NPRISK yields higher hazard values for the sites SM1, SM2 and

SM4 than KERFRACT reflects the higher overall activity rate per unit area. Clearly decisive for the results is the very different bandwidth function (scaling from 2 to 74 km for magnitudes of 4.0 to 6.2 respectively) and the applied anisotropy, which implies a higher probability for migration of the high magnitudes along the mapped structure.

Table 3 - Parameters used to generate the pseudo random seismicity catalogue shown in Fig. 4 (left); the critical parameters used in the KERFRACT hazard computation (right). The Cartesian coordinates of the test sites (SM1-SM4) and the PGA ground motion relation used in the computations are also included.

Catalogue generation parameters		**KERFRACT parameters**	
a	0.33	Fractal scaling (PL)	1.5
b	0.88	Anisotropy	$\delta = 11.0$; $\phi = 45°$
Depth	5.0 km	Bandwidth H(M)	$0.0041*e^{1.58*M}$
M_{low}	4.0	Effective return period	500 years
M_{high}	6.2	Epicenter uncertainty	25 km
Epicenter uncertainty	25 km	Magnitude uncertainty	0.1
Magnitude uncertainty	0.1		
Sites for hazard computation			
x = 0.0; y = 0.0	SM1	x = 133.0; y = 133.0	SM3
x = 0.0; y = -111.0	SM2	x = 27.0; y = 0.0	SM4
PGA relation (cm/s^2); unpublished for U.K.			
log (PGA) = 5.72 + 0.59 M – 1.26 log (R + 2.13 exp (0.25 M))			

In conclusion, both the synthetic datasets demonstrate that KERFRACT yield results that are comparable with the results from NPRISK, and in addition better reflect hazard reduction when the distance to the activity centre increases. The two tests with the spatially inhomogeneous distributions (Figs. 1 and 4) also indicate that practically equal results were to be expected for spatially homogeneous activity distributions.

3.4. Real earthquake catalogues

SOUTHERN NORWAY. - Fig. 6 shows the earthquake distribution in southern Norway from 1657 to 1995 for magnitudes greater than or equal to the threshold magnitude (4.0). With a magnitude uncertainty of 0.4, also earthquakes with magnitude down to 3.6 were included. A regression on this catalogue was made for establishing the bandwidth function (see Table 4), which is also shown in an inset in Fig. 6 together with the frequency-magnitude distribution and the cumulative Gutenberg-Richter relation showing a good fit of the data to the obtained relation. All dependent events have been removed along with explosions.

To obtain the effective return period, the earthquakes were classified in magnitude intervals of 0.5 up to magnitude 5.4, while the earthquakes above 5.4 were studied one by one. The probability of being detected in each one of the centuries was evaluated and assigned according to their epicentral location and the temporal distribution of the magnitudes. The obtained

probabilities were subsequently used to estimate the effective return period for each event in the catalogue as stated previously.

The NPRISK hazard study used for comparison was by NORSAR and NGI (1998) (see also Bungum et al., 2000), where the peak ground acceleration (PGA) relation of Ambraseys et al. (1996) was used. Therefore, the same ground motion relation was used also in this study. Three sites were selected for comparative hazard estimates with KERFRACT (NW1-NW3) as shown in Fig. 6, and to stabilize the estimates, the seismic hazard was computed at the site and at additional eight surrounding sites within a 400 km^2 square centred at the site. The results in Fig. 7 are the mean values of the nine estimates around the site. In order to assess the seismic hazard, only earthquakes with epicentral distance less than 1000 km were used, assuming that earthquakes at greater distances have no influence on the estimated hazard.

For each site three KERFRACT computations were conducted. The first, K(1), used only the earthquakes in the catalogue, whereas the others also included some background seismicity. Since the kernel computation is strictly based on the historic earthquake catalogue, any lack of completeness or erroneous estimate of the effective return period will immediately be reflected in the hazard results. A short catalogue will inevitably show gaps (compared to the ideal continuous Gutenberg-Richter distribution) in the higher magnitudes and particularly between the largest observed and the largest possible earthquake. To compensate for these 'gaps' the

Table 4 - Parameters describing the real catalogues and the hazard computation parameters. Also the range of background activity (magnitudes) used in each KERFRACT computation (K(1)-K(3)) is included.

	W. Norway	**S. Spain**
Bandwidth $H(M) = c*e^{dM}$	c=0.048; d=1.55	c=1.340; d=0.6
G-R b-value	1.05	0.7
Focal depth	5 - 25 km	5 - 25 km
Comparative study	NORSAR and NGI (1998)	Molina (1998)
Site location	NW1: 5.5°, 61.5° NW2: 6.0°, 59.5° NW3: 9.0°, 61.0°	SP1: -4.0°, 37.0° SP2: -0.8°, 38.0° SP3: -2.5°, 38.0°
Mmax	7.0 (all the sites)	7.0(SP1); 7.1(SP2); 6.0(SP3)
K(1): Background activity	No	No
K(2): Background activity	M=5.9 - 7.0 (all the sites)	M=6.6 - 7.0 (SP1) M=7.0 - 7.1 (SP2) M=5.1 - 5.9 (SP3)
K(3): Background activity	As K(2) and "holes" above M=5.1 (all the sites)	As K(2) and "holes" above M=6.1(SP1&SP2); M=5.1(SP3)
PGA relations		
W. Norway	$\log (PGA) = -1.48 + 0.266\ M - 0.922 \log (R)$ PGA in units of gravity and R in km (Ambraseys et al., 1996)	
S. Spain	$\ln (PGA) = 2.564 + 0.834\ M - 0.737 \ln (R) - 0.006\ R$ PGA in cm/s^2 and R in km (Molina, 1998)	

KERFRACT computation is augmented with background activity at two levels, the K(2) level adds background activity in the magnitude range from the largest observed to the estimated M_{max}, and the K(3) level additionally adds background activity at magnitudes (above 5.0) where the historical record is void. The background seismicity was obtained by means of a

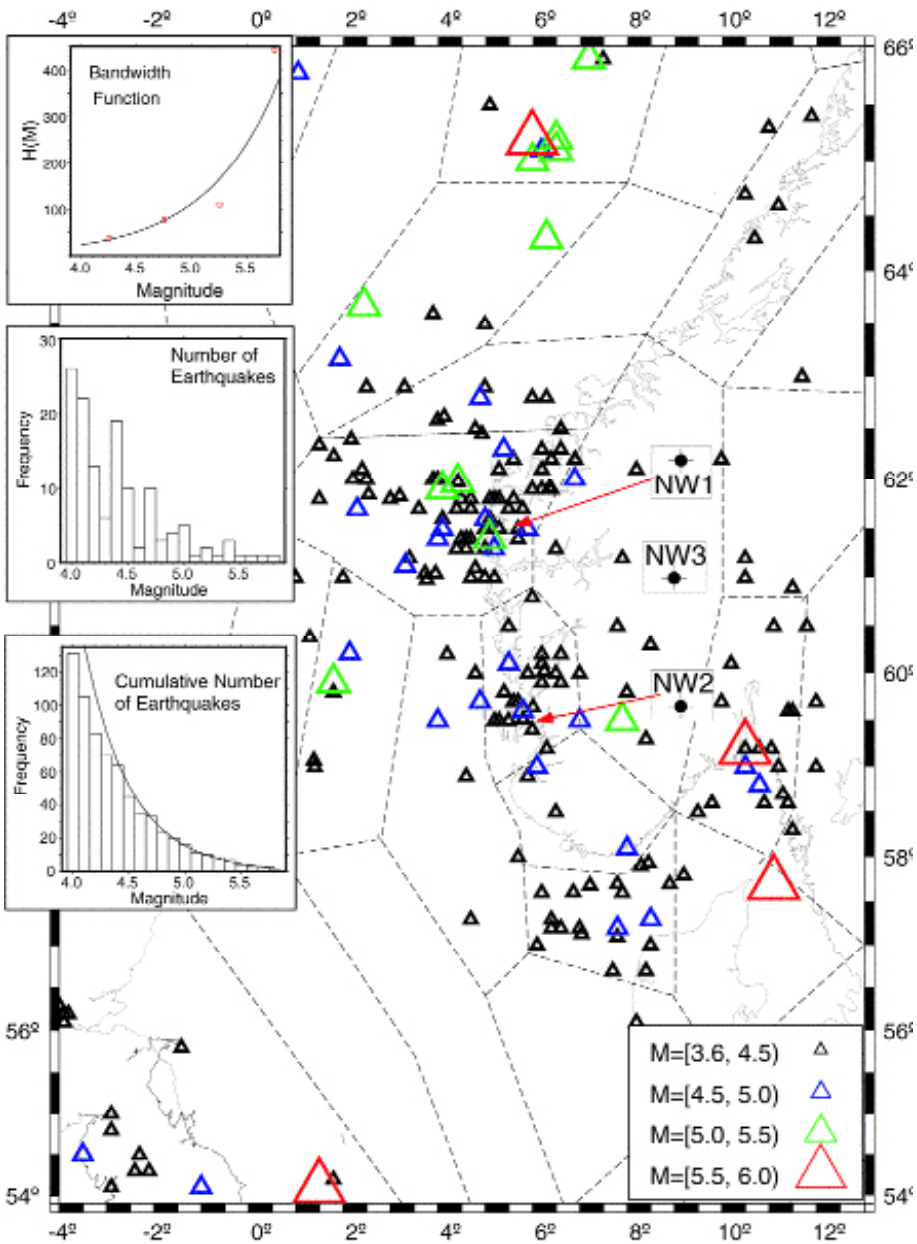

Fig. 6 - Earthquake distribution in a selected region of southern Norway (1657 - 1995) and with magnitudes greater than or equal to 3.6. Three test sites are indicated (NW1-NW3) where the seismic hazard was computed. The dotted lines on the map are the seismic sources used to apply Cornell-McGuire approach. The hazard results and reference results from the Cornell-McGuire approach are shown in Fig. 7. Insets: The bandwidth function $H(M)$ estimated from the data (upper), the magnitude-frequency distribution (middle) and the cumulative magnitude-frequency distribution (lower).

truncated Gutenberg-Richter relationship with a fixed b-value of 1.05 and calibrated to activity rates found in NORSAR and NGI (1998). The hazard results, obtained with the KERFRACT program, and the comparison with NPRISK results (NORSAR and NGI, 1998) are shown in Fig. 7 for each of the aforementioned sites.

As shown in Fig. 7, the hazard results obtained with K(1) using only the earthquakes

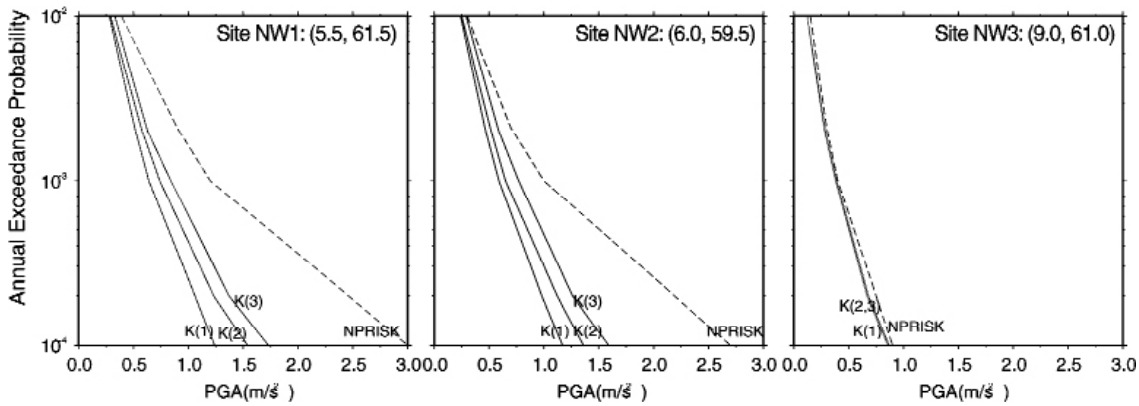

Fig. 7 - Comparison of the seismic hazard results obtained by KERFRACT (solid lines) and the results obtained by NPRISK (dotted lines) for the three test sites in Fig. 6. The KERFRACT hazard parameters are shown in Table 4.

included in the catalogue are much lower than the NPRISK results for sites with high seismic activity (NW1 and NW2), but quite similar for the low activity site (NW3). Furthermore, the effect of adding background activity is demonstrated through the considerable increase in the seismic hazard, however without reaching the NPRISK hazard levels. The observed discrepancy in hazard levels may be attributed to conservative estimates in the NORSAR and NGI (1998) study, possibly caused by the extrapolation of the log-linear earthquake distribution up to $M_{max} = 7.0$ (even though the sensitivity to this M_{max} value is very low due to the combination of low a-values and low weights). As in similar shelf areas, the data from Norway are insufficient to evaluate the character of the recurrence relationship for magnitudes above ~5.5, rendering this evaluation to be based on data from similar areas worldwide and on expert opinion of the seismic potentials in the region. Another and possibly more likely explanation for the observed hazard discrepancy is that the effective return periods, which are very different in on shore and off shore areas, have been overestimated, leading to under-conservative KERFRACT results.

A deaggregation analysis of the K(2) results for the NW1 site has been done for three ground motion levels, and Fig. 8 shows the contribution from each event in the earthquake catalogue (bar height is proportional to contribution). The figure demonstrates one of the main features of the Kernel method (valid for all PSHA methods), namely that at low ground motion levels the local, frequent earthquakes are most important. However, at higher ground motion levels the larger, infrequent (and more distant) earthquakes contribute increasingly, reflecting also that the probability density function is different for high and low magnitude events. This difference is directly related to the fractal scaling and to the bandwidth function, and it will also depend on the ground motion relation used. The influence of the location and magnitude

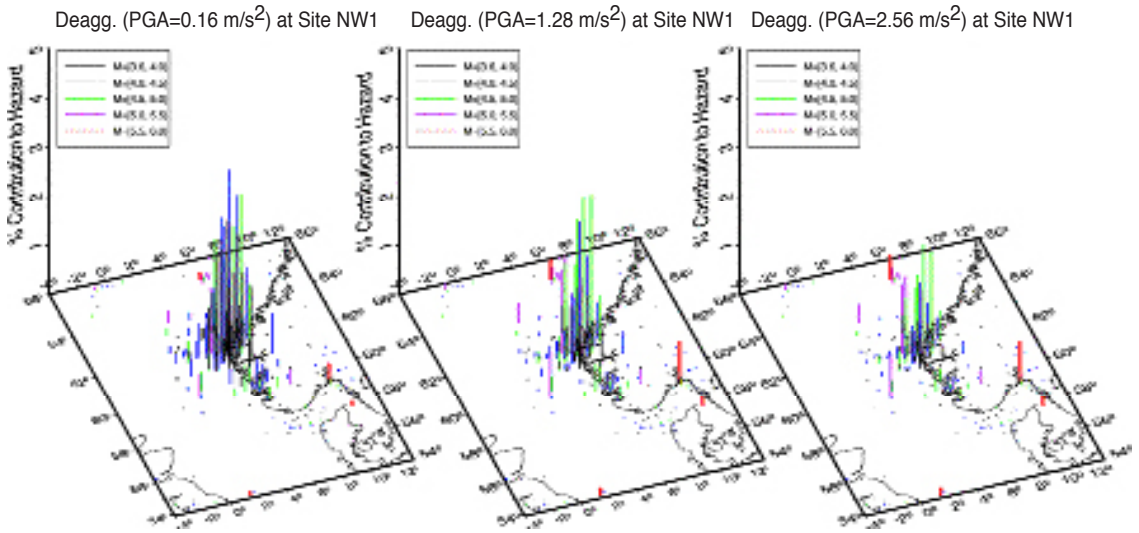

Fig. 8 - Deaggregation of the seismic hazard results, $K(1)$, for a target ground motion of 0.16 m/s^2 (left), 1.28 m/s^2 (centre) and 2.56 m/s^2 (right) for site NW1 in Fig. 6. The bars are placed at epicentral locations, and the height is proportional to the hazard contribution.

uncertainty on the results was investigated, and was found to be weak within reasonable location uncertainty intervals.

SOUTHERN SPAIN. - Fig. 9 shows the epicentral distribution in southern Spain for earthquakes with magnitudes greater than or equal to the threshold magnitude (4.0), including (as for Norway) also earthquakes from 3.6, due to the magnitude uncertainty. It is seen how the largest earthquakes are clustered on the north coast of Algeria, and on the Iberian Peninsula the seismicity follows the coastal mountain range (Betic Cordillera) where also the two largest earthquakes of 6.5 and 6.9 are found. Only shallow earthquakes (depth < 30 km) were included in the catalogue, and all dependent events were removed along with explosions assuring independence between all the events. Table 4 and the inset in Fig. 9 shows the bandwidth function obtained by means of a nearest-neighbourhood regression, yielding a more structured seismicity pattern (also for the higher magnitudes) than found for southern Norway. To obtain the effective return period, the earthquakes were grouped in magnitude intervals of 0.5 up to 5.9, while larger earthquakes were studied one by one, according to their epicentral location.

In this study a catalogue with 328 earthquakes ranging from magnitude 3.6 to 7.5 in the period 1300-1998 was used, with a magnitude estimation for historical earthquakes based on methodologies and relationships developed for the Ibero-Maghrebian region (Lopez Casado et al., 2000b). Epicentral location uncertainty from the reporting agency was used whenever available, and was otherwise set to 150 km for offshore events and 50 km for inland earthquakes before 1800. The epicentral location errors have decreased with time to around 10 km for most of the off shore events and to around 5 km for the inland earthquakes for the last 30 years and although they have an statistical sense instead of a physical sense we believe that the most scientific alternative to quantify the uncertainty of the epicentre in this catalogue is to use the

formal location error instead of using some estimation combining formal uncertainties and personal beliefs. Similarly, the magnitude uncertainty has been assessed to decrease from 0.7 for non-instrumental earthquakes to 0.3 for the last 20 years of instrumental recording.

Fig. 9 also shows the magnitude-frequency distribution (inset) of earthquakes from the threshold magnitude (4.0) to the maximum magnitude (7.5). When compared to the Norwegian data, the magnitude distribution seems less homogeneous and there are gaps of seismicity between large and smaller magnitudes. This may indicate that a characteristic earthquake behaviour fits the data better than the log-linear Gutenberg-Richter relation which also has a low *b*-value, but it may also reflect a quality problem with the estimation of historical earthquake magnitudes.

The comparative Cornell-McGuire study of Molina (1998) used a modified version of NPRISK with source zones defined for southern Iberia and northern Africa, and activity parameters and maximum magnitude for each source area was estimated with the application of various statistical methods (Utsu, 1966; Weichert, 1980; Kijko and Sellevoll, 1989). The same attenuation relations (Molina, 1998; Lopez Casado et al., 2000a) were used in the KERFRACT computations below. Three sites (SP1-SP3) were selected for hazard computation as shown in

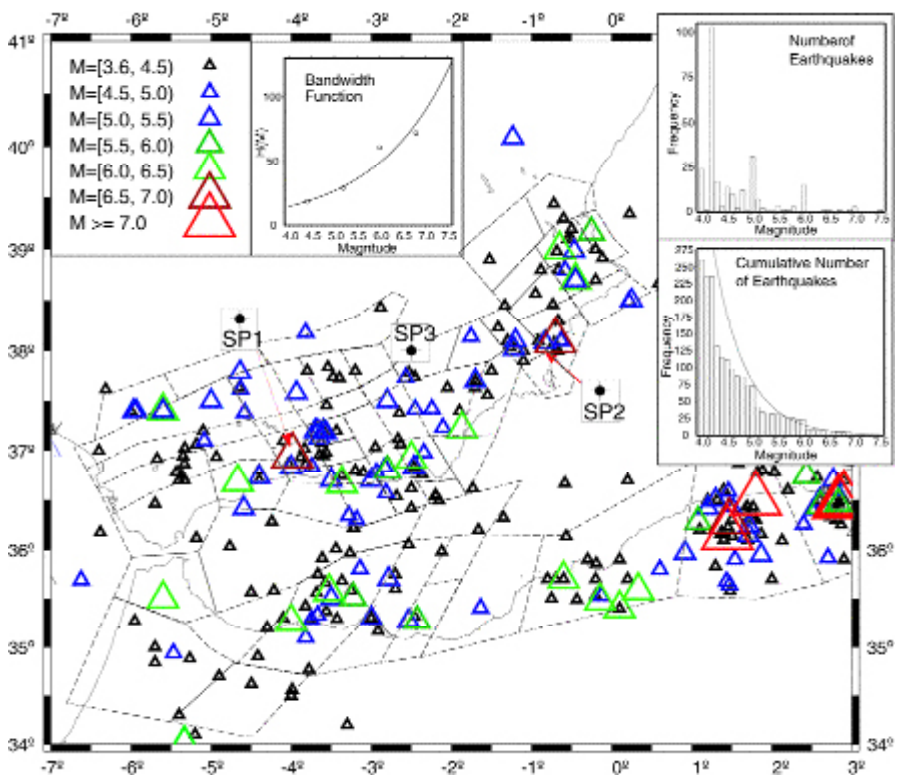

Fig. 9 - Earthquake distribution of a selected region of Spain (1300 - 1998) and magnitudes greater than or equal to 3.6. Three test sites are indicated (SP1-SP3) where the seismic hazard was computed. The dotted lines on the map are the seismic sources used to apply the Cornell-McGuire approach. The seismic hazard results are shown in Fig. 10 along with reference results from the Cornell-McGuire approach. Insets: The bandwidth function $H(M)$ estimated from the data (upper left), the magnitude-frequency distribution (upper right) and the cumulative magnitude-frequency distribution (middle right).

Fig. 9, and as for Norway, the results presented were computed as the mean value of the hazard at the site and eight surrounding sites. Three hazard computations were done for each of the sites following the same procedure as detailed above for Norway, where *K*(*1*) uses only the catalogue, *K*(*2*) uses some added background seismicity and *K*(*3*) adds background seismicity more extensively, by filling 'holes' in the Gutenberg-Richter relation. The background seismicity was estimated through the Gutenberg-Richter relation, and in order to facilitate comparison between the results obtained from both methodologies (the Kernel method and the Cornell-McGuire method), the maximum magnitude for each site was selected by following the Molina (1998) approach. If the kernel is established on a grid, covering, say 200 x 200 km, centered around the site, we look for the maximum magnitude proposed by Molina (1998) in the whole area of that grid. That is to say, maximum expected magnitudes were estimated (Molina, 1998) to 7.0, 7.1 and 6.0 around sites SP1, SP2 and SP3 respectively.

Fig. 10 shows the obtained hazard curves (PGA) along with the hazard curves from the comparative studies. The figure shows the same basic features as found for the Norwegian catalogue, with differences (but less than for Norway) between the Cornell-McGuire results (Molina, 1998) and the KERFRACT results at sites with high seismic activity, but with similar results for the low seismicity site (SP3). The significant gaps of magnitudes observed in the Spanish catalogue (for example around site SP2 within a 200 by 200 km squared region there is a gap between M = 5.2 and M = 6.0, and another gap between M = 6.0 and M = 6.9), are main factors responsible for the difference in the results between the two methodologies, and as expected the hazard increases when background seismicity is included.

A deaggregation analysis of the K(2) results for the three sites was done for three ground

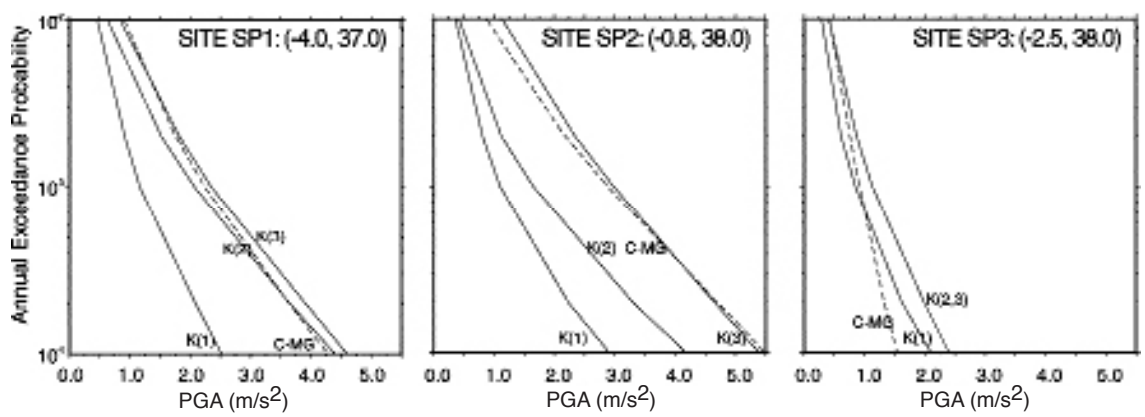

Fig. 10 - Comparison of the results provided by KERFRACT (solid lines) and the results provided by NPRISK (dotted lines) for the three test sites in Fig. 9. The KERFRACT hazard parameters are shown in Table 4.

motion levels as shown in Fig. 11 for site SP2. As for the previous examples from Norway, the distant larger earthquakes become important contributors at the higher ground motion levels, and as for the Norwegian catalogue the influence of alternative (but realistic) uncertainties in epicentre and magnitude were not leading to significant changes in the results.

4. Discussion and conclusions

We have documented above, extensive tests of a kernel based method for seismic hazard computation (KERFRACT), performed both with synthetic and real data. A Cornell-McGuire PSHA code (NPRISK) has been used for comparisons.

The tests with synthetic data yielded results comparable with results from the classical Cornell-McGuire methodology. These tests also demonstrated a well known (but sometimes forgotten) weakness of zonation based methods, namely that the assumption of a near uniform activity rate within a zone is a poor representation of the true activity rate. The Kernel method exhibited a hazard decay with increasing distance to the seismicity cluster which is more in accordance with intuitive expectations. However, the question as to which extent this decay is "correct", is not easy to resolve, also because this is related to the appropriateness of the applied attenuation relation.

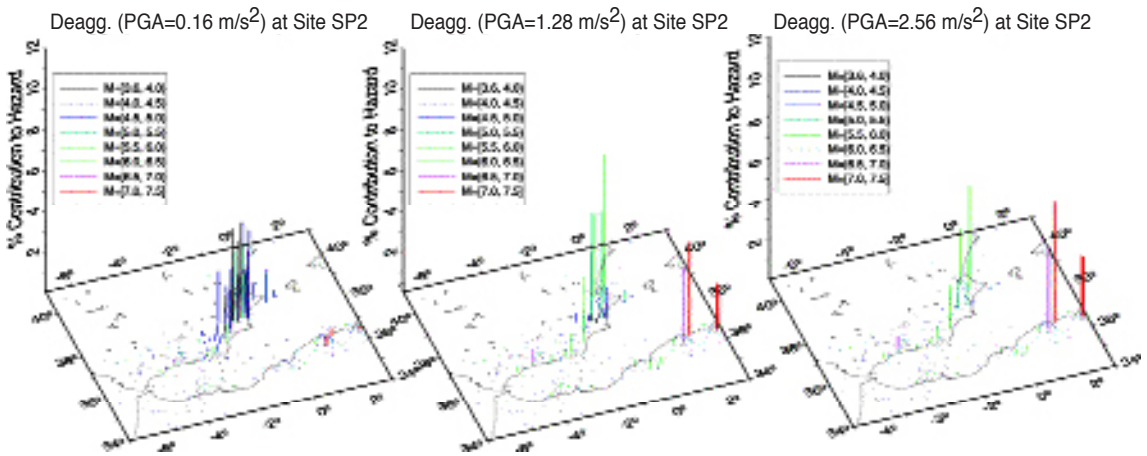

Fig. 11 - Deaggregation of the seismic hazard results, $K(1)$, for a target ground motion of 0.16 m/s2 (left), 1.28 m/s^2 and 2.56 m/s^2 (right) for site SM2 in Fig. 9. The bars are placed at epicentral locations, and the height is proportional to the hazard contribution.

The KERFRACT approach allows the user to quantify some of the basic seismotectonic concepts through parameters that describe the spatial decay, the degree of clustering as function of magnitude and the degree of anisotropy in the epicentre distribution. The experience with these parameters indicates that the spatial cluster parameter ($H(M)$) has the most significant impact on the results, while the general decay parameter (the fractal scaling PL) has a weaker impact in orders of magnitude. The anisotropy parameter (δ) will be useful mainly in cases when the epicentral distribution does not reflect some lineation which would be reasonable to assume on the basis of structural geological indications.

The Kernel method applied on real catalogues and compared with results from Cornell-McGuire based computations revealed that the KERFRACT yielded systematically lower seismic hazard results, and in particular for low probabilities in Norway. However, the KERFRACT results based on the Spanish catalogue were more in accord with the

corresponding Cornell-McGuire results, particularly when some background activity was added to the historical record. These results are not surprising since the Spanish catalogue covers a time window of nearly 700 years, and also covers a more active region in which the largest earthquakes in the historical record are closer to the maximum expected magnitude (from seismotectonic considerations) than is the case for Norway. The discrepancy for Norway between the two methods may also relate to an overestimation of the effective return period.

In conclusion, we have found the following characteristic features of the Kernel method.

- The results are based more directly on the historical record, which entails tighter connections to the empirical data. In fact, the results critically depend on a long, high quality, earthquake record.
- While there are fewer possibilities to include 'expert opinion', the possibilities that are facilitated are more explicit, thereby supporting a higher degree of transparency of the model and the results.
- When working with an historical record without adding background seismicity, the results may be regarded as 'lower bound' results, provided that there are no positive biases in the historical magnitudes.

While a deaggregation analysis is found useful for all PSHA computations, it is regarded as an absolutely necessary post-analysis that adds transparency to the KERFRACT results.

The results reported on here do indicate that the Kernel method is attractive in situations where a good historical record exists, but in regions with a poor or short historical earthquake record the classical Cornell-McGuire approach may still be preferable.

The investigations reported on in this paper are not exhaustive, and we recommend that more testing on real data from different tectonic environments is conducted before the Kernel approach can be used with the same confidence as the Cornell-McGuire method has been used for a long time. Through more experience it is also to be expected that new features can be added and developed.

Acknowledgments. This work was supported by NORSAR and the Conselleria de Cultura, Educacion y Ciencia of the Generalitat Valenciana, (POST00-01-66). We are very grateful to Dr. G. Woo for generously providing the computer code and for many useful comments and suggestions. Three anonymous reviewers are also thanked for their constructive comments.

References

Ambraseys N.N., Simpson K.A. and Bommer J.J.; 1996: *Prediction of horizontal response spectra in Europe*. Earth. Eng. Struct. Dyn., **25**, 371-400.

Algermissen S.T. and Perkins D.M.; 1973: *A technique for seismic zoning; general consideration and parameters*. In: NOOA (ed), Contributions to seismic zoning, Rockville, United States, 1785-1974.

Bungum H., Lindholm C., Dahle A., Woo G., Nadim F., Holme J.K., Gudmestad O.T., Hagberg T.S. and Karthigeyan K.; 2000: *New seismic zoning maps for Norway, the north Sea and UK*. Seism. Res. Lett., **71**, 687-697.

Cornell C.A.; 1968: *Engineering seismic risk analysis*. Bull. Seism. Soc. Am., **58**, 1583- 1606.

Field E.H., Jackson D.D. and Dolan J.F.; 1999: *A mutually consistent seismic-hazard source model for Southern California*. Bull. Seism. Soc. Am., **89**, 559-578.

Frankel A.; 1995: *Mapping seismic hazard in the central and eastern United States*. Seism. Res. Lett., **66**, 8-21.

Frankel A., Mueller C., Barnhard T., Perkins D., Leyendecker E., Dickman N., Hanson S. and Hopper M.; 1996: *National seismic hazard maps*. Open- File-Report 96-532, U.S.G.S., Denver, 110 pp.

Frankel A., Mueller C., Barnhard T., Leyendecker E., Wesson R., Harmsen S., Klein F., Perkins D., Dickman N., Hanson S. and Hopper M.; 2000: *USGS National seismic hazard maps*. Earthquake Spectra, **16**, 1-19.

Jackson D. and Kagan Y.; 1999: *Testable earthquake forecasts for 1999*. Seism. Res. Lett., **70**, 393-403.

Kijko A. and Sellevoll M.; 1989: *Estimation of earthquake hazard parameters from incomplete datafiles. Part 1: Utilization of extreme and complete catalogs with different threshold magnitudes*. Bull. Seism. Soc. Am., **79**, 645-654.

Lopez Casado C., Molina S., Delgado J. and Pelaez J.A.; 2000a: *Attenuation of the intensity with epicentral distance in the Iberian Peninsula*. Bull. Seism. Soc. Am., **90**, 34-47.

Lopez Casado C., Molina S., Giner J.J. and Delgado J.; 2000b: *Magnitude-intensity relationships in the Ibero-Maghrebian region*. Nat. Hazards, **22**, 271-300.

McGuire R.K.; 1976: *FORTRAN computer program for seismic risk analysis,* Open-File Report 76-67, U.S.G.S., Denver.

McGuire R.K.; 1978: *FRISK: Computer program for seismic risk analysis using faults as earthquake sources*. Open File Report No 78-1007, U.S.G.S., Denver.

Molina S.; 1998: *Sismotectonica y peligrosidad sismica del area de contacto entre Iberia y Africa*. Ph D. Thesis. Universidad de Granada (in Spanish).

NORSAR and NGI; 1998: *Seismic zonation for Norway*. Report prepared for Norwegian Council for Building Standardization (NBR).

Perkins D.; 2000: *Fuzzy sources, maximum likelihood and the new methodology.* In: Lapajne J.K. (ed), Seismicity modelling in seismic hazard mapping, Geophysical Survey of Slovenia, Ljubljana, pp. 67-75.

Stein R.S. and Hanks T.C.; 1998: *M > = 6 earthquakes in southern California during the twentieth century: no evidence for a seismicity or moment deficit*. Bull. Seism. Soc. Am., **88**, 635-652.

Utsu T.; 1966: *A statistical significance test of the difference in b-value between two earthquake groups*. Journal Phys. Earth, **14**, 37-40.

Vere-Jones D.; 1992: *Statistical methods for the description and display of earthquake catalogs*. In: A.T. Walden and P. Guttorp (eds), Statistics in the Enviromental and Earth Sciences, Edward Arnold, London, pp. 220-246.

Ward S.N.; 1994: *A multidisciplinary approach to seismic hazard in southern California*. Bull. Seism. Soc. Am., **84**, 1293-1309.

Ward S.N.; 1998: *On the consistency of earthquake moment rates, geological fault data, and space geodetic strain: the United States*. Geophys. J. Int., **134**, 172-186.

Weichert D.; 1980: *Estimation of earthquake recurrence parameters for unequal observations periods for different magnitudes*. Bull. Seism. Soc. Am., **70**, 1337-1347.

WGCEP (Working Group on California Earthquake Prediction); 1995: *Seismic hazards in southern California: probable earthquakes, 1994 to 2024*. Bull. Seism. Soc. Am., **85**, 379-439.

Woo G.; 1994: *NPRISK Seismic hazard computation algorithm based on Cornell-McGuire principles*. Code developed at NORSAR.

Woo G.; 1996: *Kernel estimation methods for seismic hazard area source modeling*. Bull. Seism. Soc. Am., **86**, 1-10.