ESTIMATION OF LONG-TERM VOLCANIC HAZARD USING A COX PROCESS WITH A MULTIVARIATE POTENTIAL

OLIVIER JAQUET, CHRISTIAN LANTUÉJOUL AND JUNICHI GOTO
In2Earth Modelling Ltd, CH-4051 Basel , Switzerland.
Ecole Nationale Supérieure des Mines de Paris, F-77300 Fontainebleau, France.
Nuclear Waste Management Organisation of Japan, J-1080014, Tokyo, Japan.

ABSTRACT

Recent efforts for siting a geologic repository in Japan have stimulated the development of a specific stochastic model for improving uncertainty characterization. A Cox process, characterized by a multivariate potential, was developed to enable assimilation of geological information and geophysical data. This model accounts for the observed spatial patterns and since future activity is more likely to occur at past event locations, the simulation method was made conditional on these event sites. The description of uncertainty of future volcanic activity is improved using such a multivariate approach. The concepts and mathematical basis of this model are given and a methodological illustration is provided using a database of Quaternary Volcanoes of Japan, in this case the Quaternary Tohoku volcanic arc in northern Honshu.

INTRODUCTION

Long-term volcanic hazard is gaining relevance due to increasing societal demands on time scales of hundreds to hundreds of thousands of years as regards to the siting of critical facilities. For sites near volcanically active regions, long-term volcanic hazards often constitute the dominant source of uncertainty as input for risk assessments. Uncertainty is mainly related to imperfect knowledge of non-linear volcanic processes, to space-time variability of distribution and intensity for volcanic events and to a limited amount of information. For these reasons, the estimation of volcanic hazard based on a probabilistic formalism is internationally getting more frequently employed.

The siting of a geological high-level radioactive waste (HLW) repository in Japan, in principle, will consider regions that are not obviously excluded on the grounds of recent and current volcanism. Therefore, it is important for both the Nuclear Waste Management Organisation of Japan (NUMO) and the broader geosciences community to have a probabilistic methodology, along with more commonly used deterministic approach, for the assessment of long-term volcanic hazard at potential sites located in non-excluded regions.

The probabilistic methodology for volcanic hazard assessment, currently developed by NUMO (Apted et al., 2004; Chapman et al., 2008; Tsuchi et al., 2008) with an international team of geoscientists (Japan, UK, USA, Switzerland and France), addresses time frames of 10'000, 100'000 and 1 million years . As part of this methodology, several stochastic models were developed using specific geological conceptualisations based on various data sets and information related to past and current volcanic activity and their geophysical signature.

The theoretical basis and concepts of one of these models is given below and an illustration is provided using data from the Quaternary Tohoku volcanic arc in northern Honshu.

CONCEPTUALISATION

According to Tamura et al. (2002; 2009), magma productivity is not uniform along the Tohoku volcanic arc due to the 3D structure of the mantle wedge. This productivity may be controlled by locally developed hot regions within the mantle wedge beneath the volcanic arc having the shape of inclined fingers with an average width of 50 km. The hot finger model of Tamura (Figure 1) is based on spatial clustering of volcanic events, topographic data, seismic low-velocity zones in the mantle wedge (Hasegawa et al., 2005) and local gravity anomalies.

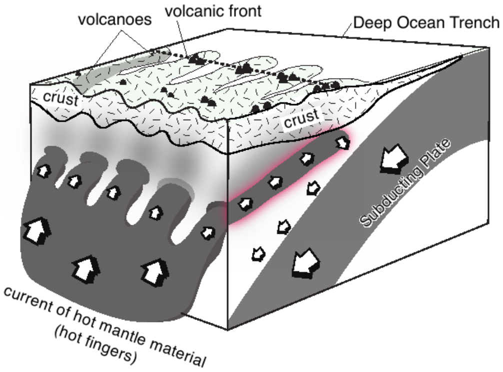

Figure 1: Hot finger model according to Tamura et al. (2009).

Due to the stability of the plate tectonic regime for the past several hundreds of thousands years (Kondo, 2009), characteristics of future volcanism by analogy with Quaternary volcanism are likely to remain similar for the next tens of thousands of years (up to ca. 0.1 million years).

Consequently, from a conceptual point of view, the following pieces of evidence and hypotheses have to be incorporated within our modelling perspective: (1) the spatial distribution of Quaternary volcanic events tends to be clustered, (2) the spatial distribution of these events is correlated to the geophysical signature of crustal and mantle structures underneath, and (3) future volcanic events are more likely located in zones of past activity and above hot regions within the mantle wedge.

Therefore, in order to integrate geological and geophysical information, we introduce the notion of potential of volcanism which represents the propensity of a given region to be affected by volcanic events. The potential of volcanism, being unknown, is considered as randomly structured within the context of the stochastic model. While the structured part represents the current geologic and geophysical knowledge, the random part describes the uncertainty associated with this information.

MODEL DEVELOPMENT

In the stochastic model, the volcanic events that occurred during the Quaternary period are considered as a particular realization of a (random) point process. Similarly, the potential of volcanism during this period is interpreted as a particular realization of a (positive) random intensity function Z . Partitioning the volcanic region of interest into small domains $\left(a_i ,\ i \in I \right)$, the potential of domain a_i is denoted by Z_i . In what follows, the two properties of the Cox process (Cox, 1955; Lantuéjoul, 2002) are assumed to be satisfied: (1) The number N_i of volcanic events that occurred in the domain a_i is Poisson distributed with random mean Z_i ; (2) Given the potentials Z_i , Z_j , ..., Z_k of the domains a_i , a_j , ..., a_k , the numbers of volcanic events N_i , N_j , ..., N_k are mutually independent. It should be pointed out that property (2) means that the numbers of volcanic events in pairwise disjoint domains are only conditionally independent. They are not independent because the potential conveys its own structure on the Cox process.

Property (1) can be formulated as:

$$P\left\{ N_i = n \right\} = E\left(e^{-Z_i} \frac{Z_i^n}{n!} \right) \qquad n = 0,\ 1,\ 2,\ \ldots \tag{1}$$

which implies that N_i is not Poisson distributed unless Z_i is deterministic. Note, however, that N_i and Z_i possess the same mean, and even the same conditional mean:

$$E\left\{ N_i \mid Z \right\} = Z_i \tag{2}$$

In other words, given the potential of all domains of interest, the mean number of events occurred in domain a_i is precisely the potential of that domain. Among all factors that contribute to the potential, the location of hot regions at depth is available using geophysical methods. This geophysical information is also interpreted as a particular realization of a random function $S=\left(S_i,\, i\in I\right)$. Owing to the integration of geophysics to the potential, the distribution of the number of volcanic events is not directly dependent on the geophysical information when the potential is known:

$$D\left(N_i \mid Z, S\right) = D\left(N_i \mid Z \right) \qquad (3)$$

A flexible way to model the dependence relationships between the potential and the geophysical information is to write that their respective transforms in Gaussian space are bigaussian. More precisely, it is assumed that Z and S can be written as:

$$Z_i = \varphi_Z \left(Y_i^Z\right) \;\; S_i = \varphi_S \left(Y_i^S\right) \qquad (4)$$

where φ_Z and φ_S are two monotonic increasing functions (i.e. Gaussian anamorphosis functions), and where Y^Z and Y^S are two standardized Gaussian random functions. Moreover, Y^Z and Y^S are related by the regression formula:

$$Y_i^Z = \rho\, Y_i^S + \sqrt{1-\rho^2}\; Y_i^R \;\; i\in I \qquad (5)$$

which involves the correlation coefficient ρ and a third standardized (residual) Gaussian random function Y^R independent of Y^S .

For the estimation of volcanic hazard, the proposed stochastic model (Cox process) is characterized by a multivariate potential of volcanism, since this potential presents dependencies with past volcanic activity as well as with geophysics.

Conditional simulation

The estimation of volcanic hazard is performed by simulating the distribution of volcanic events likely to occur during a certain period of time in the future within the region of interest. In addition, the simulation has to deliver volcanic events that are more likely to be located in zones of past activity. Therefore, the simulation requires to be conditioned to all available data corresponding to the location of past volcanic events. The idea is to simulate the potential of volcanism conditioned on the number of volcanic events and on the geophysical information known for the analysed region. The conditioning is achieved only on the numbers of past volcanic events n_i and the geophysical data s_i known in each domain a_i . In particular, the exact location of each volcanic event is not taken into account. The proposed algorithm for the conditional simulation of the potential is based on the Gibbs sampler (Geman and Geman, 1984). This is an iterative algorithm that comprises the following steps:

(i) generate $Y_i^R \sim \mathit{Gaussian}\left(0, 1\right)$ *for each* $i\in I$ *;*

(ii) select an index i *at random;*

```
(iii) generate $y_0 \sim D\left(Y_i^R \mid Y_j^R = y_j^R,\ j \neq i\right)$
(iv) compute $y_i^S = \varphi_S^{-1}\left(s_i\right)$ and $z_0 = \varphi_Z\left(\rho\, y_i^S + \sqrt{1-\rho^2}\, y_0\right)$
(v) generate $n_0 \sim Poisson\left(z_0\right)$ ;
(vi) if $n_0 = n_i$ ; then put $y_i^R = y_0$ and $z_i = z_0$ ;
(vii) go to (ii).
```

Step (i) is used to initialize the Gaussian residuals Y_i^R . This can be done either by generating them separately, as written in the algorithm, or jointly, by resorting to the standard techniques used for simulating Gaussian random functions such as Choleski decomposition (Wilkinson, 1965), circulant embedding (Dietrich and Newsam, 1997) or turning bands method (Lantuéjoul, 2002). The usual procedure for step (iii) is to sample y_0 from a Gaussian distribution. Its mean is the simple kriging estimate of Y_i^R starting from all y_j^R except y_i^R . Its variance is the corresponding kriging variance (Wackernagel, 2003). The correlation between the potential of volcanism and the geophysics is introduced at step (iv). In its design, the algorithm runs forever. In practice, it is stopped when each potential z_i has been effectively updated more than several hundred times.

At this stage, only the potential $\left(z_i,\ i \in I\right)$ of the past volcanic events has been generated. This potential is representative only of the period of time t_p from which all data originate. If the objective is to simulate the volcanic events that will occur during the future period of time t_f , then the future potential $\left(z_i^f,\ i \in I\right)$ is required. Provided that the potential varies very slowly through time, there is no inconvenience to assume that the past and future potentials are proportional. This leads to the following algorithm to simulate the future volcanic events n_i^f :

```
(i) compute $z_i^f = z_i \cdot \left(t_f / t_p\right)$ for each $i \in I$ ;
(ii) generate $n_i^f \sim Poisson\left(z_i^f\right)$ for each $i \in I$ .
```

The conditional simulation algorithm allows the estimation of volcanic hazard for each domain of the region of interest during the period of time considered. A Monte Carlo approach is performed using several thousands simulations in order to derive stable probability estimates:

$$P\left\{N_i^f \geq 1\right\} \approx \frac{1}{K_{sim}} \sum_{k=1}^{K_{sim}} 1_k\left(n_i^f \geq 1\right) \qquad (6)$$

where K_{sim} is the total number of simulations and $1_k\left(n_i^f \geq 1\right)$ equals 1 when the k^{th} simulation assigns the domain a_i one or more volcanic events, and 0 otherwise.

TOHOKU CASE STUDY

The issue of concern is the estimation of the probability for the formation of new polygenetic volcanoes over a proposed performance period of 0.1 million years. All Quaternary polygenetic volcanoes (those active within the last 1.8 million years) of the Tohoku region are considered as representing potentially active volcanic areas for the future. The Tohoku volcanic arc is located in northern Honshu (Figure 2), Japan and consists of more than 100 volcanic edifices erupted during the Quaternary. A volcanic edifice corresponds to the volcanic event considered for the Tohoku case study; i.e. a polygenetic volcano is composed of one or more edifices. Polygenetic volcanoes are characterized by a geomorphology and geologic structures that are created by many episodes of eruptive activity likely to affect broad areas. The formation of new polygenetic volcanoes can take place at locations up to tens of kilometres away from sites of previous eruptions. Some of the volcanoes of the Tohoku region remain active today. We illustrate our methodology using a database of Quaternary Volcanoes of Japan (details can be found in Mahony and Sparks, 2009).

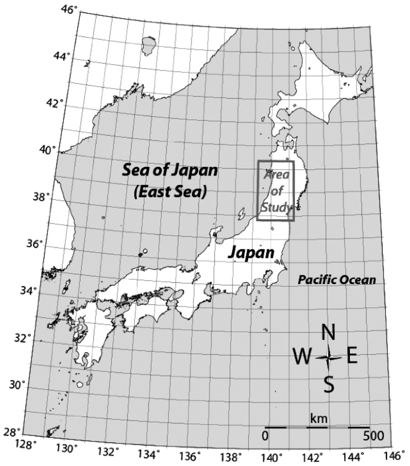

Figure 2: Location map of the study region, situated in northern Honshu, Japan.

The potential of volcanism for the Tohoku region is believed to assimilate geological data related to Quaternary volcanic events as well as seismic data indicating hot regions at depth. The latter consists of S-wave velocity perturbations along the inclined low-velocity zone in the mantle wedge of NE Japan (Hasegawa and Nakajima, 2004). Due to the generality of the developed model, other kind of geophysical data could be applied, if the presence of a statistical correlation can be revealed between volcanic events and the geophysics.

Estimation of volcanic hazard

The simulation of the potential requires the estimation of the following parameters: (1) the Gaussian transforms for the potential and the seismic data, (2) the correlation coefficient and (3) the variograms describing the randomly structured spatial patterns for the potential and the seismic. The details of the estimation procedure for each of these parameters are given in Jaquet and Lantuéjoul (2009).

With the Cox model, simulations of the number of future volcanic events were carried out over a period of 0.1 million years for the Tohoku region (Figure 3). A Monte Carlo approach was performed using 10’000 simulations in order to obtain stable probability estimates. The likelihood of future volcanic events was displayed in form of a hazard map for the period of interest related to the siting for critical facilities (Figure 4). The comparison with stochastic models with a deterministic potential, i.e. a deterministic intensity function, (Connor and Connor, 2009) shows that the integration of the uncertainty for the potential of volcanism causes the resulting probability estimates to have dissemination effects: i.e. zones of high probability are likely to display lower values and zones of low probability have a tendency toward increased values. These results reveal the importance of accounting for the uncertainty associated with the potential of volcanism when assessing long-term volcanic hazard.

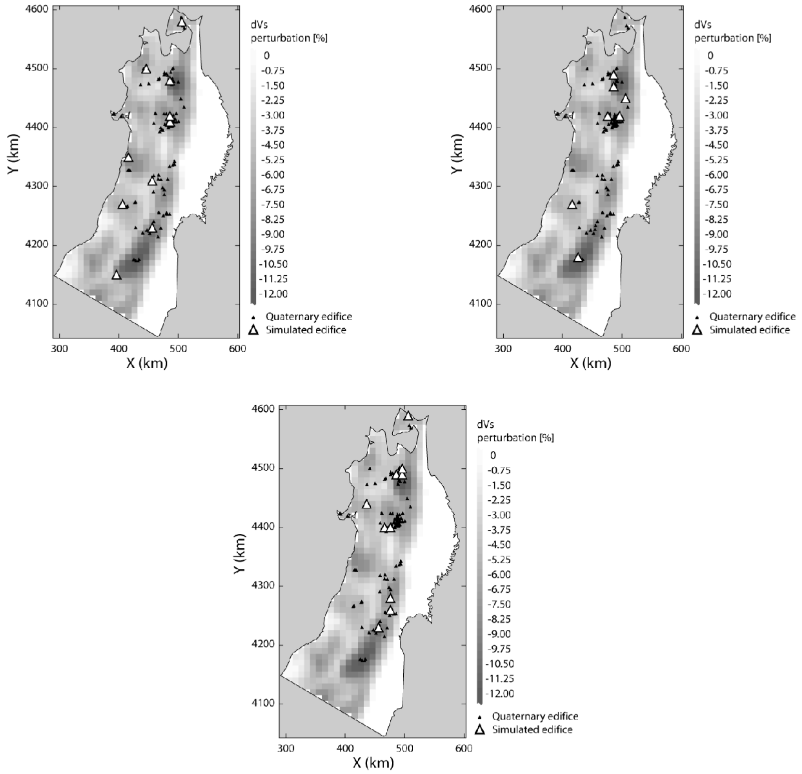

Figure 3: Three Cox simulations with a multivariate potential of volcanism. The simulated edifices are likely to be located in zones with past activity as well as in zones with seismic anomalies (i.e. low dVs values).

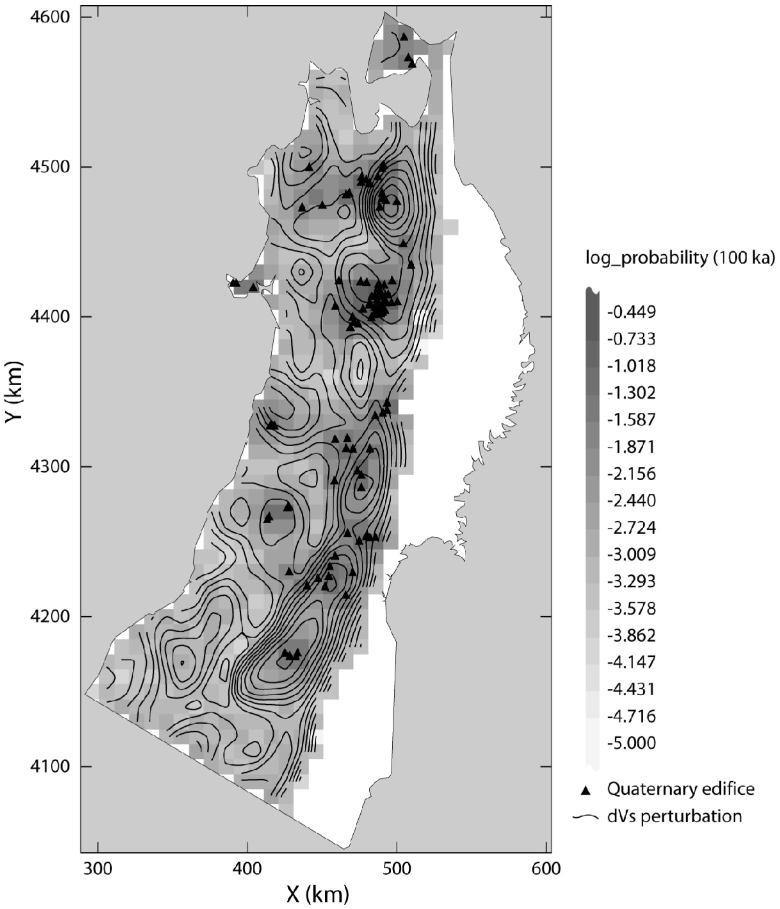

Figure 4: Volcanic hazard map for the Tohoku region and for the next 0.1 million years (domain scale is 10 km by 10 km).

CONCLUSIONS

A stochastic model is proposed for the long-term assessment of volcanic hazard in relation to the siting for critical facilities. By considering a Cox process with a multivariate potential of volcanism, the assimilation of Quaternary geological information and geophysical data becomes operational for hazard calculations. Such multivariate approach enhances the characterization of uncertainty when forecasting volcanic activity on the long term. In particular, the use of geophysical data, should improve hazard maps for regions located between clusters of past volcanic events.

This stochastic model is part of the development of a probabilistic methodology for the assessment of volcanic hazard for potential HLW repository sites in Japan;

it provides a valuable contribution for the evaluation of uncertainties. In particular, as several probabilistic models have been developed for this siting issue in Japan (Jaquet et al., 2008), they will contribute to the perception of the epistemic uncertainty (Woo, 1999). The comparison of hazard results relying on different concepts such as the non-homogeneous Poisson process (with a smooth deterministic potential) and the Cox model (using a random potential) should deliver some evidence with respect to the variability of this conceptual uncertainty and its possible reduction.

Non-homogeneous patterns for the recurrence rate of volcanic events (Smith and Keenan, 2005) need to be considered even when sufficient age data are lacking. Preliminary results have shown the existence of structured behaviours for the occurrence of volcanic events in time. The use of solely homogeneous Poisson models (in time) is by no means conservative when estimating volcanic hazard at long term. Further developments of the model will consider additional geological information (fault location) and geophysical data (e.g. gravity anomalies) that also correlate with the spatial distribution of volcanic events. Ultimately, Cox processes with non-stationary potential are likely needed for complex volcanic regions.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of the Nuclear Waste Management Organization of Japan in making this work possible. Useful discussions with H. Tsuchi, K. Kitayama, H. Kawamura, H. Kondo, Y. Tamura, C. Connor, S. Sparks, N. Chapman and L. Connor improved this manuscript.

REFERENCES

Apted M., Berryman K., Chapman N. A., Cloos M., Connor C. B., Kitayama K., Sparks R.S.J. and Tsuchi. H. (2004) *Locating a radioactive waste repository in the ring of fire*, Vol. 85, Nb. 45, pp. 465-480.

Chapman N. A., Goto J. and Tsuchi H. (2008) *Likelihood of Tectonic Activity Affecting the Geological Stability of a Repository in Japan: Development of NUMO's ITM Methodology*, proceedings of the Stability and Buffering Capacity of the Geosphere for Long-term Isolation of Radioactive Waste: Application to Crystalline Rock, Manchester, November 13-15, 2007, OECD Nuclear Energy Agency, Paris, in press.

Connor C. B. and Connor L. J. (2009) *Estimating spatial intensity with kernel methods,* In Volcanism, Tectonism, and Siting Nuclear Facilities, In Volcanism, Tectonism and the Siting of Nuclear Facilities, edited by C. B. Connor, N. A. Chapman and L. J. Connor, in press.

Cox D. R. (1955) *Some statistical methods connected with series of events,* Journal of the Royal Statistical Society, Series B (Methodological), XVII, No. 2, pp. 129-164.

Dietrich C. and Newsam G. (1997) *Fast and exact simulation of stationary Gaussian processes through the circulant embedding of the covariance matrix.* Siam Journal on Scientific Computing, 18(4):1088-1107.

Geman S. and Geman D. (1984) *Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images, IEE Trans. Pattern Anal. and Mach Intel.*, Vol. 6, pp. 721-741.

Hasegawa A. and Nakajima J. (2004) *Geophysical constraints on slab subduction and arc magmatism,* Geophysical monograph 150, IUGG, Vol. 19, pp. 81-93.

Hasegawa A., Nakajima J., Umino N. and Miura S. (2005) *Deep structure of the northeastern Japan arc and its implications for crustal deformation and shallow seismic activity,* Tectonophysics, 403, pp. 59-75.

Jaquet O., Connor C. B. and Connor L. J. (2008) *Probabilistic methodology for long-term assessment of volcanic hazards,* Journal of Nuclear Technology, Vol. 163, pp. 180-189.

Jaquet O. and Lantuéjoul C. (2009) *Cox process models for the estimation of long-term volcanic hazard,* In Volcanism, Tectonism, and the Siting of Nuclear Facilities, edited by C. B. Connor, N. A. Chapman and L. J. Connor, in press.

Kondo H. (2009) *Regional-scale volcanology in support of site-specific investigations,* In Volcanism, Tectonism and the Siting of Nuclear Facilities, edited by C. B. Connor, N. A. Chapman and L. J. Connor, in press.

Lantuéjoul C. (2002) *Geostatistical simulation: models and algorithms.* Springer, 256 pp.

Mahony S. and Sparks R.S.J. (2009) *Assessing lon-term volcanic hazards using volcano distribution databases,* In Volcanism, Tectonism and the Siting of Nuclear Facilities, edited by C. B. Connor, N. A. Chapman and L. J. Connor, in press.

Smith E.I. and Keenan D.L. (2005) *Yucca mountain could face greater volcanic threat, Eos*, Transactions, American Geophysical Union, Vol. 86, No. 35, pp. 317-321.

Tamura Y., Tatsumi Y., .Zhao D., Kido Y. and Shukuno H. (2002) *Hot fingers in the mantle wedge: new insights into magma genesis in subduction zones,* Earth and Planetary Sciences Letters, 197, pp. 105-116.

Tamura Y., Nakajima J., Kodaira S. and Hasegawa A. (2009) *Tectonic setting of volcanic centers in subduction zones: 3-D structure of mantle wedge and arc crust,* In Volcanism, Tectonism and the Siting of Nuclear Facilities, edited by C. B. Connor, N. A. Chapman and L. J. Connor, in press.

Tsuchi H., Goto J., Chapman N. A. and Kawamura, H. (2008) *Tectonic Hazard and the Siting of Japan's HLW Repository: Preliminary Results of the Tohoku Case Study in the ITM Project*, proceedings of the 2008 International High-Level Radioactive Waste Management Conference, Las Vegas, September 7-11, 2008, American Nuclear Society, La Grande Park, in press.

Wackernagel H. (2003) *Multivariate geostatistics.* Third edition, Springer-Verlag, Berlin, 387 pp.

Wilkinson J.H. (1965) *The algebraic eigenvalue problem,* Oxford University Press, New York, 1965, 662 pp.

Woo G. (1999) *The mathematics of natural catastrophes.* Imperial College Press, London, 292 pp.