AUTOMATIC METHODS FOR GENERATING SEISMIC INTENSITY MAPS

DAVID R. BRILLINGER, [1] *University of California*
CHANG CHIANN [2] AND PEDRO A. MORETTIN, [2] *University of São Paulo*
RAFAEL A. IRIZARRY, [3] *Johns Hopkins University*

Abstract

For many years the Modified Mercalli (MM) scale has been used to describe the observed effects of sizeable earthquakes on buildings and people. Initially isoseismal lines of the effects were added to maps by hand. Some objective methods have been proposed eg. DeRubeis et al. [14], Brillinger [2, 3, 4, 6], Wald et al.[33], Pettenati et al. [25]. The work presented here develops such methods further. In particular the ordinal character of the MM scale is specifically taken into account. Numerical smoothing is basic to the approach and methods involving splines, local polynomial regression and wavelets are illustrated. The approach presented also allows the inclusion of explanatory variables, for example site effects. The procedure is implemented for data from the 17 October 1989 Loma Prieta event.

Keywords: Isoseismals; Mercalli scale; ordinal data; smoothing; local polynomials; splines; wavelets

AMS 1991 Subject Classification: Primary: 62M30
Secondary: 86A15

1. Introduction

It has been common for seismic researchers concerned with earthquake effects to classify damage that has occurred on an ordinal scale, see for example Reiter [27]. Working with such a scale has the advantage that values may sometimes be inferred for historical events, for example Justo and Salwa [20] do such a study for the 1531 Lisbon earthquake. Such data may then be used to develop risk estimates covering hundreds of years for a region of concern.

In many cases the damage information is summarized via isoseismal contours superposed on a map. These lines are the loci of points that separate areas of equal seismic intensity and prove useful to quantify the shaking pattern and to

[1]Postal address: Department of Statistics, University of California, Berkeley, CA 94720. Email address: brill@stat.berkeley.edu.

[2]Postal address: Department of Statistics, University of São Paulo, SP 05315-970, Brazil

[3]Postal address: Department of Biostatistics, Johns Hopkins, University, Baltimore, MD 21205

understand the damage, see Bullen and Bolt [9], Reiter [27]. A principal difficulty has been that the curves are drawn subjectively by hand.

One often used intensity scale is the *Modified Mercalli Intensity* (MMI) Scale, *ibid.* It is still used nowadays even though many precise instrumental recordings are often available, eg. the three component traces of ground motion. The MMIs still provide important supplementary information, see Wald et al. [33] and because of the direct relation of intensity values to damage they are in fact often what engineers desire the most.

The intentions of the analysis to be presented is to further develop automatic displays of earthquake damage effects. The statistical methods employed include the generalized additive model and local smoothing as provided by splines, polynomials and wavelets. The layout of the paper is the following. After this introduction, Section 2 gives some background on seismology and the Loma Prieta event. Section 3 provides the statistical methods to be used in the analysis and Section 4 some background on ordinal data. The statistical model is discussed in Section 5 and the results in Section 6. The final Section includes discussion and conclusions.

2. Background on Seismology and Loma Prieta Event

After a sizeable earthquake observations are made of its effects on structures and people. The observations are often recorded on a descriptive scale. One such is that of Modified Mercalli (MM) intensities. It has 12 ordinal levels of increasing severity. For example the description of MM_{VIII} reads

> "Damage slight in specially designed structures; considerable in ordinary substantial buildings, with partial collapse; great in poorly built structures. Disturbs persons driving motor cars. Fall of chimneys,..."

while MM_{III} includes

> "Felt quite noticeably indoors, especially on upper floors of buildings, but many people do not recognize it as an earthquake ..."

The complete scale may be found in Bullen and Bolt [9], Perkins and Boatwright [24].

When intensity data are examined, there is found to be a general fall-off in severity of effect with distance from the earthquake source. Figure 1 presents the observations for the Loma Prieta event of 17 October 1989. (The maps use arabic rather than roman numerals.) Intensities 0, II - IX were observed in the event. The epicenter of the earthquake is marked by a large dot. A general description of the event is provided in Bolt [1]. The event had magnitude 6.9, duration 10 seconds, depth of 19km, and led to 63 deaths, 1300 buildings destroyed and 5.9 billion dollars damage. There were 921 observations of MM intensity in all. The data displayed and analyzed are those employed in Stover et al. [29]. In the figures one notes some intensities of levels VIII and IX at some distance from the epicenter. These were a result of local site and building conditions. Figure 2 is a plot of MMI value against distance from the epicenter. (The intensity values have been jittered on the plot to make the distinct cases more apparent.)

Perkins and Boatwright [24] list some factors on which MM intensities depend. These are: the size of the earthquake, the distance of the site from the earthquake source, the focusing of the earthquake energy and the geologic material underlying the site.

Disadvantages of the use of MMI values include: a) no measurement may be available if no person or damageable objects were present in an area and b) local site conditions are not taken into account.

3. Background on Local Polynomial, Spline and Wavelet Models

3.1. Generalized Linear and Additive Models

The classical linear model $Y = \beta'\mathbf{x} + \varepsilon$ postulates that ε is (normally) distributed with mean 0 and variance σ^2. However in many situations the form of the data makes this model inappropriate. For example, it is clear that for the data discussed in this paper $\mu = E(Y|\mathbf{x})$ is nonnegative. To handle such situations we may generalize the linear model by assuming that Y follows some exponential family distribution, not necessarily normal, and that the dependence of the mean μ on the covariate, $\mathbf{x}$, is given by a link function $h(\mu) = \beta'\mathbf{x}$, see McCullagh and Nelder [22]. We can further generalize this model by keeping the additive structure but relaxing the linear assumption. The model can be written as $g\{E(Y|x)\} = \sum_{j=1}^{p} f_j(x_j)$ with the x_j the covariates in $\mathbf{x}$ and the f_j arbitrary smooth functions. This is usually called a generalized additive model, see Hastie and Tibshirani [19]. The function `gam()` in S-Plus fits this model by a socalled *local scoring procedure*. In order to obtain smooth estimates of the f_js this proceder uses splines and local polynomials, which are described in the next two sections. Wavelets are another choice and are described in Section 3.4.

3.2. Local polynomials

A method for smoothing is to fit local polynomials. For a data set (x_k, y_k), $k = 1, \cdots, n$, the fitted value, $\hat{y}_j$, at x_j is the value of, say, a dth degree polynomial fit to the data using weighted least squares. The weight for (x_k, y_k) is large if x_k is close to x_j and small if it is not. For example, we may consider the symmetric triweight function

$$h(u; d) = \begin{cases} (1- \mid u/d \mid^3)^3, & \text{if } \mid u \mid < d, \\ 0, & \text{otherwise,} \end{cases}$$

with d a *window size* or *span*. For each (x_j, y_j), weights $h_j(x_k)$ are defined for all x_k, $k = 1, \cdots, n$, by

$$h_j(x_k; d_j) = h(|x_j - x_k|; d_j).$$

We may define d_j as the distance from x_j to its qth nearest neighbor of x_j. It can be convenient to choose $q = \lfloor pn \rfloor$, where $0 < p < 1$. Note that as p increases the neighborhood of influential points increases leading to a smoother fit.

This procedure is easily extended to the case of fitting a bi-dimensional surface to data $(\mathbf{x}_k, y_k)$, with $\mathbf{x} \in R^2$. Once a distance is defined for points in R^2 the weight function $h(u; d)$ is assigned in exactly the same way, and a local polynomial surface may be computed. Cleveland [11], Cleveland et al. [12] describe the functions `loess()` and `lo()` and indicate some methods for choosing p and d in practice. Hastie and Tibshirani [19] describe how this procedure, involving a weighted least squares criterion, is extended to the more generalized, non-normal, case.

3.3. Splines

Fitting splines is another useful approach to smoothing (Wahba [32]). One defines a finite dimensional linear space of functions by considering piecewise polynomials. The regions that define the pieces are separated by a sequence of knots or breakpoints $t_1, \ldots, t_K$ and it is permitted to put constraints on the behavior of the functions at these break points. For example, the spline function `bs()` in S-Plus produces cubic smoothing splines which are piecewise cubic polynomials with continuous first and second derivatives at the knots. The smoothing splines approach (Silverman [28]) defines the linear space of cubic spline functions with knots at the unique values of the predictor measurements and chooses the function in this space that minimizes a penalized least squares criterion. This procedure is extended to the generalized additive model approach by consdiering a penalized likelihood criterion instead of a least squares one. The procedures described can be directly extended to the case where predictor measurements are two-dimensional. A tensor spline function for two predictors is a product of two one-dimensional basis functions - one for each predictor. This permits the construction of the model matrix in S-Plus for a pair of covariate vectors x_1, x_2 through bs(x_1)*bs(x_2). See Hastie and Tibshirani [19] for further details.

3.4. Wavelets

Another approach is via wavelet technology. From two basic functions, the *scaling function* $\phi(x)$ and the *wavelet* $\psi(x)$ one defines infinite collections of translated and scaled versions, $\phi_{jk}(x) = 2^{j/2}\phi(2^j x - k),\ \psi_{jk}(x) = 2^{j/2}\psi(2^j x - k),\ j, k \in Z$ often set up so that $\{\phi_{\ell k}(\cdot)\}_{k \in Z} \cup \{\psi_{jk}(\cdot)\}_{j \geq \ell; k \in Z}$ forms an orthonormal basis of $L_2(R)$, for some ℓ. A key point (Daubechies [13]) is that it is possible to construct compactly supported ϕ and ψ that generate an orthonormal system and have space-frequency localization, which allows parsimonious representations for wide classes of functions in wavelet series.

In the process of estimating the model of the paper one needs two-dimensional wavelet bases. A socalled multiresolution analysis (MRA) of $L_2(R^2)$ is obtained through the tensor product of two 1-dimensional MRA's on R. Define the bivariate scaling functions as $\Phi(x,y) = \phi(x)\phi(y)$ and the wavelets by $\Psi^h(x,y) = \phi(x)\psi(y)$, $\Psi^v(x,y) = \psi(x)\phi(y)$ and $\Psi^d(x,y) = \psi(x)\psi(y)$. Further define $\mathbf{V}_j = V_j \otimes V_j$ and its orthogonal complement in $\mathbf{V}_{j+1}$, $\mathbf{W}_j$, by $\mathbf{W}_j = \overline{\text{span}}\{\Psi^{\mu}_{j\mathbf{k}}(x,y) : \mathbf{k} = (k_1,k_2), j, k_1, k_2 \in Z, \mu = h, v, d\}$, which consists of three different wavelets (horizontal, vertical and diagonal). One can write

$$L_2(R^2) = \mathbf{V}_\ell \bigoplus\nolimits_{j\geq \ell} \mathbf{W}_j$$

and so any function $f \in L_2(R^2)$ can be expanded as

$$f = \sum_{\mathbf{k}} \sum_{\mu=h,v,d} c_{l\mathbf{k}} \Phi_{l\mathbf{k}}(x,y) + \sum_{j=l}^{\infty} \sum_{\mathbf{k}} \sum_{\mu=h,v,d} d_{j\mathbf{k}} \Psi^{\mu}_{j\mathbf{k}}(x,y), \tag{1}$$

with the wavelet coefficients given by

$$c_{l\mathbf{k}} = \int_{R^2} f(x,y)\Phi_{l\mathbf{k}}(x,y)dxdy, \quad d_{j\mathbf{k}} = \int_{R^2} f(x,y)\Psi^{\mu}_{j\mathbf{k}}(x,y)dxdy. \tag{2}$$

Wavelet bases often entertained are the Haar, Shannon, Meyer, Franklin and the compactly supported Daubechies. These can be used to generate the 2-d bases via tensor products. See Bruce and Gao [8] for details. In practice, the sums in (1) run from $j = l$ to J and $k = 0$ to 2^{j-1}, where J is the largest j such that $d_{j\mathbf{k}} \neq 0$. An instance of the use of two-dimensional wavelets as described here is Chiann and Morettin [10].

4. Some Background on Ordinal Data

Ordinal data refers to response variables whose values are categories that are ordered. Characteristics include that it does not make sense to talk of "distance" between categories and that adjacent categories may be sensibly merged with the ordinality remaining. General references are McCullagh and Nelder [**?**] and Fahrmeir and Tutz [17]. Brillinger [5] gives an example of the analysis of an ordinal-valued time series.

A convenient model leading to ordinal-valued data is based on the category boundaries or threshold approach. It is assumed that the observable response Y is a categorized version of a latent continuous variable ζ. If $1, 2, \ldots, k$ are the categories of the response Y, then

$$Y = i \iff a_{i-1} < \zeta \leq a_i, \quad i = 1, 2, \ldots, k, \tag{3}$$

where $a_0 = -\infty < a_1 < \ldots < a_k = +\infty$. The $\{a_i\}$ are cut values or thresholds. If $\mathbf{U}$ is a vector of explanatory variables they are introduced by assuming that

$$\zeta = -\mathbf{U}'\beta + \varepsilon, \tag{4}$$

for $\beta = (\beta_1, \ldots, \beta_p)$ and ε is a random variable.

These considerations lead to models based on cumulative response probabilities

$$\gamma_i = \text{Prob}(Y \leq i|\mathbf{U}) = F(a_i + \mathbf{U}^{'}\beta), \tag{5}$$

rather than the category probabilities $\text{Prob}(Y = i)$. Here F is the distribution function of ε. The model (3) is called a threshold model.

Several choices of F are possible. If F is the logistic distribution function, one has the so-called proportional odds model. If $F(x) = 1 - \exp\{-\exp\{x\}\}$, the extreme-minimal-value distribution, for some constants α_i, one has

$$\log(-\log(1 \ - \ \text{Prob}(Y > i|\mathbf{U}))) = \alpha_i + \mathbf{U}^{'}\beta, \quad i = 1, \ldots, k-1, \tag{6}$$

called the *grouped Cox model*. This means that this extreme value distribution leads to a generalized linear model with the complementary log-log link. Below a motivation is given for the use of a latent variable and the extreme-value distribution in the earthquake case.

5. Statistical Methods

5.1. The Seismic Intensity Model

If (x_j, y_j) is the location of the $j-$th measurement and I_j the observed MM intensity, then one can consider the data as a realization of a spatial marked point process $\{(x_j, y_j), I_j\}$. A basic fact to be incorporated into the modelling of this circumstance is that the intensities are ordinal-valued. A convenient model leading to ordinal-valued data was given in the preceding section.

Turning to the present situation, with (x, y) denoting location, consider a latent variable ζ such that

$$\zeta_j = g(x_j, y_j) + \varepsilon_j, \quad j = 1, \ldots, J, \tag{7}$$

where ε_j has an extreme value distribution. Here J is the number of locations with measured intensity and $g(\cdot)$ is some (smooth) function of location. Now suppose that the intensity I_j at the location j is i if $a_{i-1} < \zeta_j \leq a_i$, for some cut values $\{a_i\}$. Then

$$\text{Prob}\{I_j = i\} = \text{Prob}\{a_{i-1} < \zeta \leq a_i\} = \pi_{ij}, \tag{8}$$

in this case leading to

$$\log(-\log(1 - \text{Prob}\{I_j > i \mid (x_j, y_j)\})) = \alpha_i + g(x_j, y_j), \tag{9}$$

for some constants α_i. In the present situation the additivity of the model corresponds to the effect of location at a given site being the same for all intensities. Such grouped continuous models were considered in McCullagh [21] and McCullagh and Nelder [22]. Turning to the data, by conditioning one may act as if the successive cells are independent, see Pregibon [26], and fit the model via the usual glm algorithms. This is what is done in the examples below.

The spatial dependence is introduced here via the dependence of the function $g(\cdot)$ on location (x, y). The motivation for the state variable ζ and the extreme value distribution is that ζ of (7) represents the strength of the earthquake effect at the location (x_j, y_j). The extreme value distribution is employed because the intensity recorded is the maximum observed by an observer looking around the site, Reiter [27].

Explanatory variables $\mathbf{U}_j$ may be included via assuming a form

$$\zeta_j = \beta^{'} \mathbf{U}_j + g(x_j, y_j) + \varepsilon_j. \tag{10}$$

If the ε_j have c.d.f. $F(\cdot)$, then (8) has the form

$$\pi_{ij} = F(a_i - g(x_j, y_j)) - F(a_{i-1} - g(x_j, y_j)) \tag{11}$$

in the case of no explanatories.

Each variable I_j corresponds to a multinomial distribution, with π_{ij} given by (8), hence the likelihood may be written

$$\prod_{j=1}^{J} \prod_{i=0}^{12} \pi_{ij}^{Y_{ij}}, \tag{12}$$

where

$$Y_{ij} = \begin{cases} 1, & \text{if } I_j = i \\ 0, & \text{otherwise.} \end{cases}$$

The unknowns will include $\{a_i\}, g(\cdot)$ and the parameters in the distribution of ε. All the values $i = 0, ..., 12$ (or 0 to XII) may not appear in the data set.

5.2. Choices of the function $g(\cdot)$.

One can anticipate two situations: $g(\cdot)$ of the same level of smoothness throughout the region and $g(\cdot)$ having different scales at different places. The first case may be studied, for example, by local regression or splines. The second may be approached via wavelet methods.

A further aspect of the use of wavelets is the replacement of the fitted values $\hat{\beta}^{\mu}_{j,\mathbf{k}}$ by shrunken values $\hat{\beta}^{\mu}_{j,\mathbf{k}} w(\hat{\beta}^{\mu}_{j,\mathbf{k}}/s_{j,\mathbf{k}})$, with s the estimated standard error of $\hat{\beta}$ and w a shrinkage function. This can lead to improved estimates and reduces the need to estimate J. See Donoho and Johnstone [15], Brillinger et al. [7], Chiann and Morettin [10].

5.3. Uncertainty estimation

To compute a shrunken estimate one needs to estimate the s_{jk}. Also one will wish to display uncertainty somehow and to infer whether particular explanatories need to be in the model.

In the case of maximum likelihood estimates classical large sample expressions are available. The functions `glm()` and `gam()` provide standard errors. In another seismological context Musmeci [23] proposes the use of a bootstrap procedure. The boostrap and an alternate procedure, the jackknife, are discussed in Efron and Tibshirani [16].

6. Results

In the first set of computations the model (10) was fit by the functions `gam()` and `lo()` of S-Plus, see Hastie [18].

Figure 3 gives the estimate of $g(\cdot)$ obtained. This function shows the relative estimated effects of location. The contours of level -4 correspond to high intensities. One notes again the occurrence of high damage far from the epicenter of the event evidenced in Figure 1.

Figure 4 gives estimates of the α_i of model (10) and $\pm$ 2 standard error limits. Except for the case of $MMI = 0$ (corresponding to no effect observed) one sees a steady increase in the values with intensity. The value corresponding to $MM\ IX$ is poorly estimated.

Figure 5 gives the surface when approximated by a spline surface.

[I am trying to understand and improve this estimate.] This differs from Figure 3 in in that the contours are less detailed, yet it is based on 84 degrees of freedom, while Figure 3 was based on about 40.

Figure 6 gives the result of an approximation via wavelets. The shannon wavelet was used with $j = 0, 1, 2$. This differs from Figures 3 and 5 in that it shows more local detail. Notice in particular how a small area near the bay-area appears with high intensities.

Figure 7 gives the shrinkage results. The effect of using shrinkage is that some of the details present in Figure 6 are removed. However, Figure 7 still shows more local detail than Figures 3 and 5.

7. Discussion and Conclusions

The computations were carried out using widely available statistical functions.

The complimentary loglog link employed in the computations resulted from physical conditions

The model may be extended to include covariates directly.

The intentions have been ...

We have seen how the methods described in this paper permit the automatic construction of seismic intensity maps. Three methods were presented that provide similar results. Local polynomials and splines provide a smooth map, while wavelets provide a map with more local details. The local details may be attenuated via the use of shrinkage.

In Brillinger [2] it is shown how relations obtained by analyzing MM intensities may be used in computing premiums for earthquake insurance. In particular using assumed loss ratio values for buildings of some type of interest, one may estimate the expected loss for such a building situated a given distance from an earthquake source.

Acknowledgements

It is a pleasure to acknowledge the seminal work that David Vere-Jones has done in statistical seismology. He introduced contemporary point process methods into the field, Vere-Jones [30]. He prepared important reviews, Vere-Jones and Smith [31]. He was the first to set down many important models. Happy Birthday David. Somewhere there will be a notable earthquake today, but hopefully not at Wellington.

Lindy Brewer of the United States Geological Survey provided the MM intensity data for the Loma Prieta event.

This research was supported by the NSF Grants INR-9600251 and DMS-9971309 and the CNPq grant 910011-96-6.

References

[1] BOLT, B.A. (1993). *Earthquakes*. Freeman, New York.

[2] BRILLINGER, D.R.(1993). Earthquake risk and insurance. *Environmetrics* 4, 1-21

[3] BRILLINGER, D. R. (1994a). Trend analysis: time series and point process problems. *Environmetrics* 5, 11-19.

[4] BRILLINGER, D.R. (1994b). Examples of scientific problems and data analyses in demography, neurophysiology and seismology. *J. Computational and Graphical Statistics* 3, 1-22.

[5] BRILLINGER, D.R. (1996). An analysis of an ordinal-valued time-series. In *Athens Conference on Applied Probability and Time Series.* Springer Lecture Notes in Statistics 115, pp. 74-87

[6] BRILLINGER, D.R. (1997). Random process methods and environmental data: the 1996 hunter lecture. *Environmetrics*, 8, 269-281.

[7] BRILLINGER, D.R., MORETTIN, P.A., IRIZARRY, R.A. and CHIANN, C. (2000). Some wavelet-based analyses of Markov chain data. To appear, *Signal Processing*

[8] BRUCE, A. and GAO, H.-Ye (1996). *Applied Wavelet Analysis with S-Plus.* Springer, New York.

[9] BULLEN, K.K. and BOLT, B.A. (1985). *An Introduction to the Theory of Seismology.* Cambridge University Press, Cambridge.

[10] CHIANN, C. AND MORETTIN, P.A. (2000). Estimation of time-varying linear systems. To appear, *Statistical Inference for Stochastic Processes*.

[11] CLEVELAND, W.S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* 74, 829-836.

[12] CLEVELAND, W.S., GROSSE, E. and SHYU, W.M. (1992). Local regression models. In Chambers, J.M. and Hastie, T.J. (eds), *Statistical Models in S*, Wadsworth, Pacific Grove, pp. 309-376

[13] DAUBECHIES, I. (1992). *Ten Lectures on Wavelets.* SIAM, Philadelphia.

[14] DE RUBEIS, V., GASPARINI, C., MARAMAI, A., MURRU, M. and TERTULLIANI, A. (1992). The uncertainty and ambiguity of isoseismal maps. *Earthquake Eng. Structural Dynamics* 21, 509-523.

[15] DONOHO, D.L. and JOHNSTONE, I.M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* 81, 425-455

[16] EFRON, B. and TIBSHIRANI, R. (1993). *An Introduction to the Bootstrap.* Chapman and Hall, London.

[17] FAHRMEIR, L. and TUTZ, G. (1994). *Multivariate Statistical Modelling Based on Generalized Linear Models.* Springer-Verlag, New York

[18] HASTIE, T.J. (1992). Generalized additive models. In Chambers, J.M. and Hastie, T.J. (eds), *Statistical Models in S*, Wadsworth, Pacific Grove, pp. 249-308.

[19] HASTIE, T.J. and TIBSHIRANI, R. (1990). *Generalized Additive Models.* Chapman and Hall, London.

[20] JUSTO, J.L. and SALWA, C. (1998). The 1531 Lisbon earthquake. *Bull. Seismol. Soc. America* 88, 319-328.

[21] McCULLAGH, P. (1980). Regression models for ordinal data. *J. Royal Statistical Soc. B* 42, 109-127.

[22] McCULLAGH ,P. and NELDER, J.A. (1989). *Generalized Linear Models*, Second Edition. Chapman and Hall, New York.

[23] MUSMECI, F. (1984). A method for drawing confidence bounds on seismic contour maps. Italian National Agency for Nuclear and Alternative Energy Sources (ENEA), Rome.

[24] PERKINS, J. B. and BOATWRIGHT, J. (1995). *On Shaky Ground.* ABAG, Oakland, California.

[25] PETTENATI, F., SIROVICH, L. and CAVALLINI, F. (1999). Objective treatment and synthesis of macroseismic intensity data sets using tesselation. *Bull. Seismol. Soc. America*, 89, 1203-1213.

[26] PREGIBON, D. (1980). Discussion of paper by P. McCullagh. *J. Royal Statistical Society, Series B* 42, 109-127.

[27] REITER, L. (1990). *Earthquake Hazard Analysis.* Columbia University Press, New York.

[28] SILVERMAN, B. W. (1985) Some aspects of the spline smoothing aproach to non-parametric regression curve fitting. *J. Royal Statistical Society, Series B* 47, 1–21.

[29] STOVER, C.W., REAGOR, B.G., BALDWIN, F. and BREWER, L.R. (1990). *Preliminary isoseismal map for the Santa Cruz (Loma Prieta), California, earthquake of October 18, 1989 UTC.* OpenFile Report 90-18, National Earthquake Information Center, Denver.

[30] VERE-JONES, D. (1970). Stochastic models for earthquake occurrence (with discussion). *J. Royal Statistical Society B* 32, 1-62.

[31] VERE-JONES, D. and SMITH, E.G.C. (1981). Statistics in seismology. *Commun. Statist.-Theor. Meth.* A10, 1559-1585.

[32] WAHBA, G. (1990). *Spline Models for Observational Data.* CBMS-NSF Regional Conference Series, SIAM. Philadelphia.

[33] WALD, D.J., QUITORIANO, V., DENGLER, L.A. and DEWEY, J.W. (1999). Utilization of the internet for rapid community intensity maps. *Seismological Research Letters* 70, 680-697.

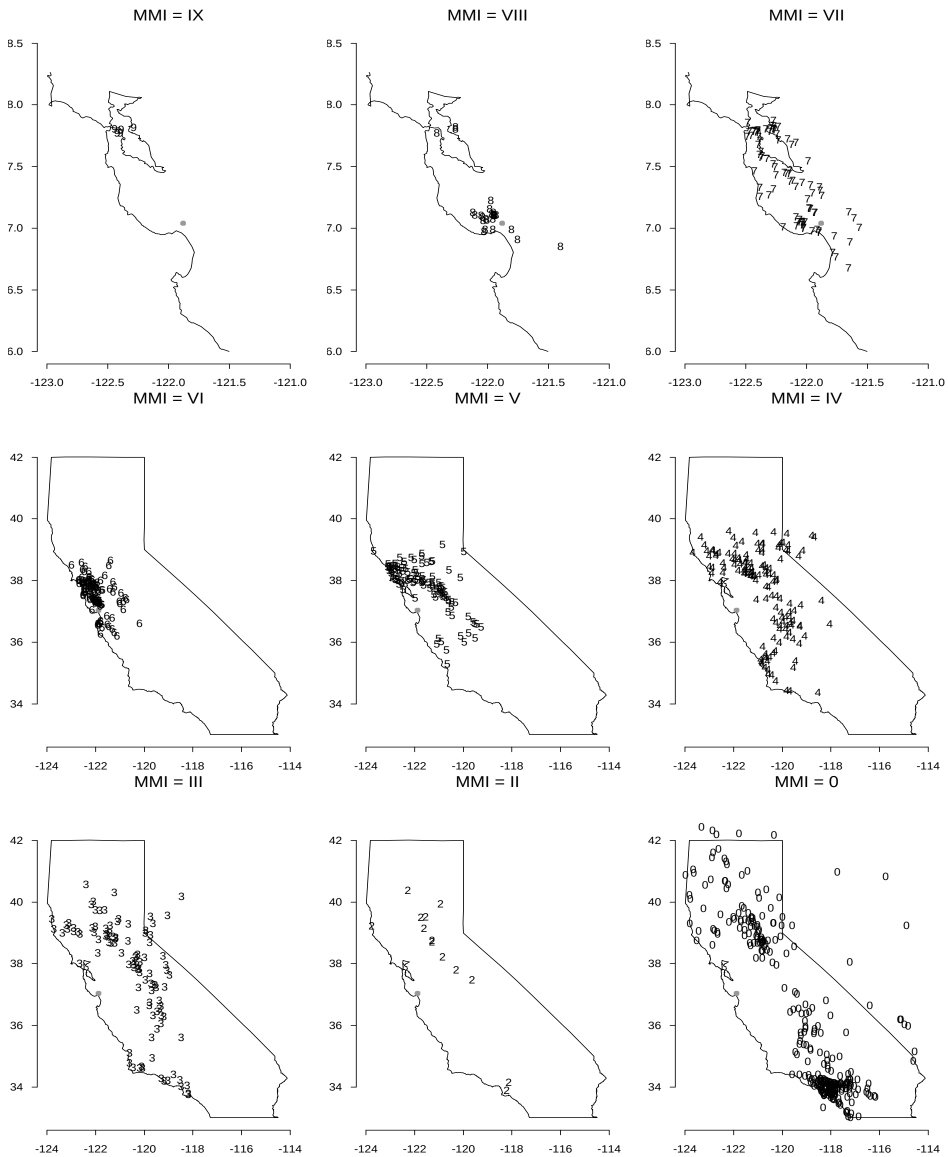

Figure 1: Locations of MM intensities.

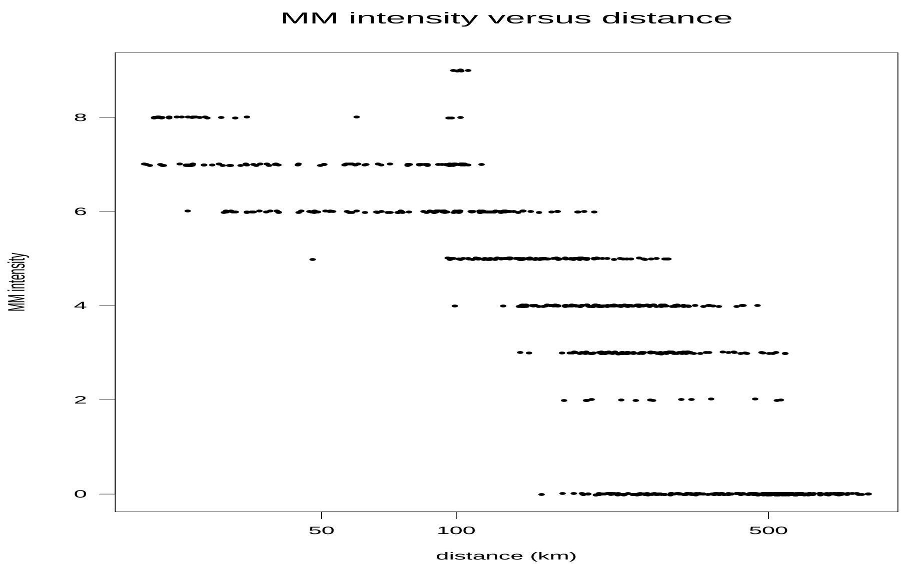

Figure 2: MM intensities versus distance from the source.

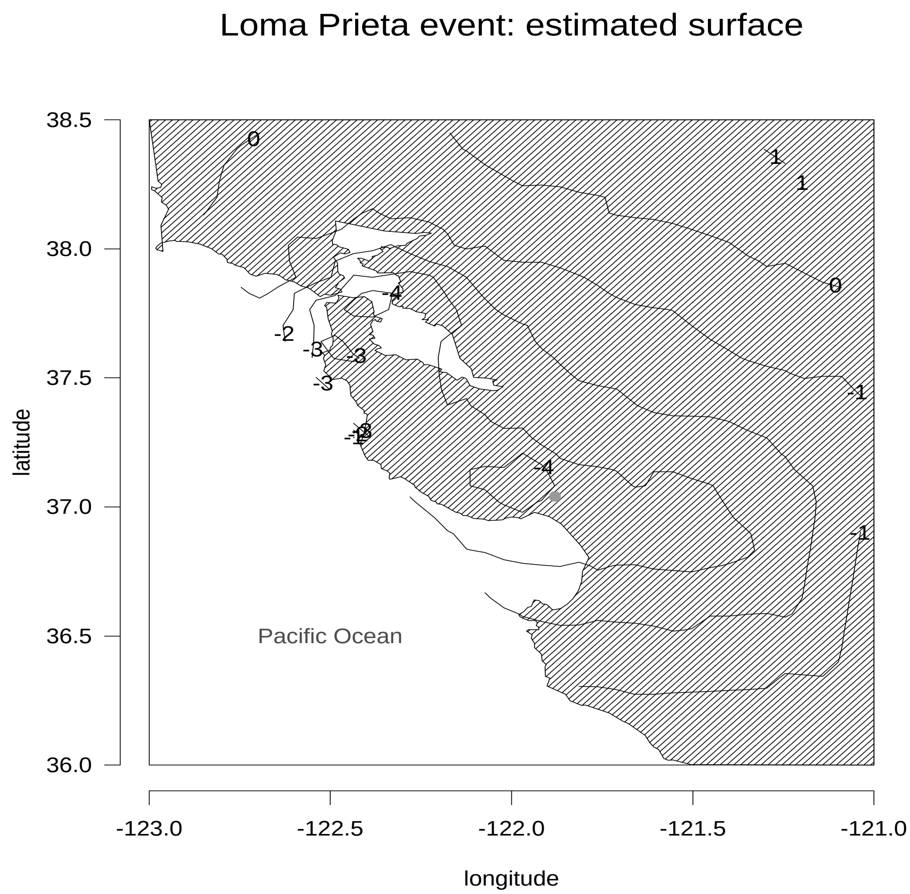

Figure 3: Estimated $g(x, y)$ of the model (10) using `lo()`.

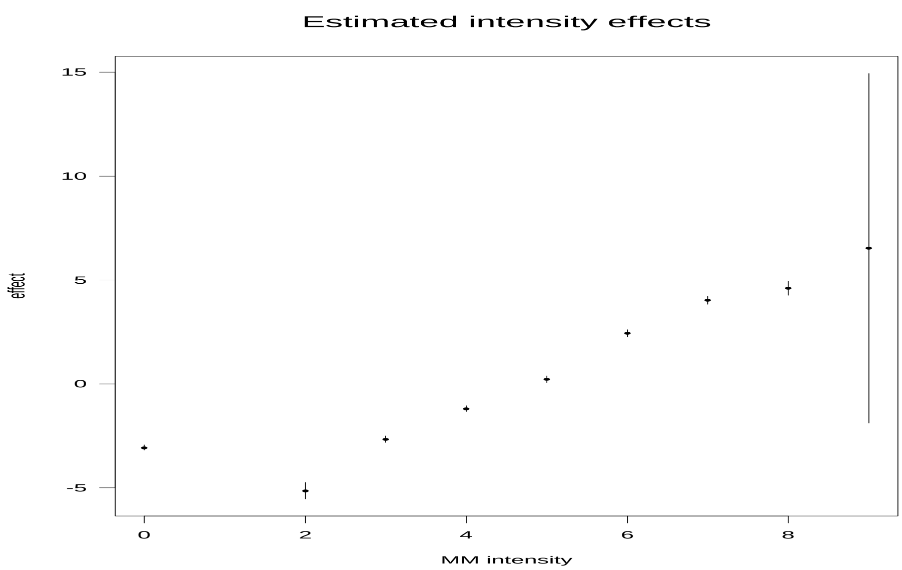

Figure 4: Estimates of the α_i of the model (10).

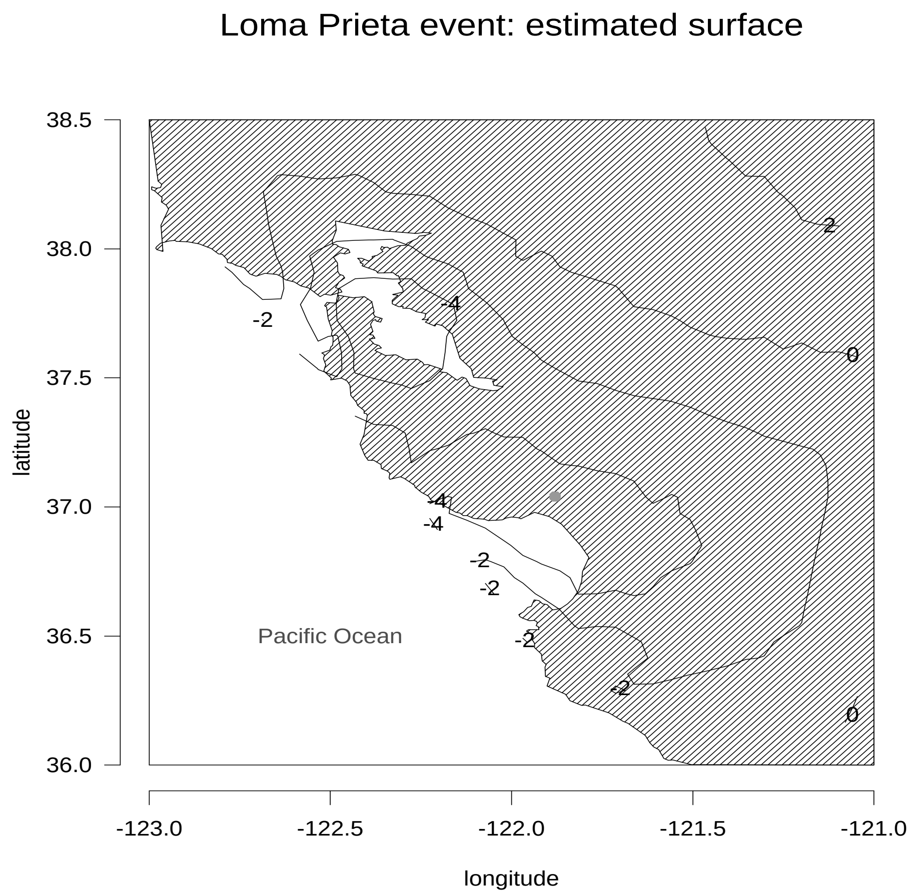

Figure 5: Estimated $g(x, y)$ of the model (10) using $bs(\cdot)$.

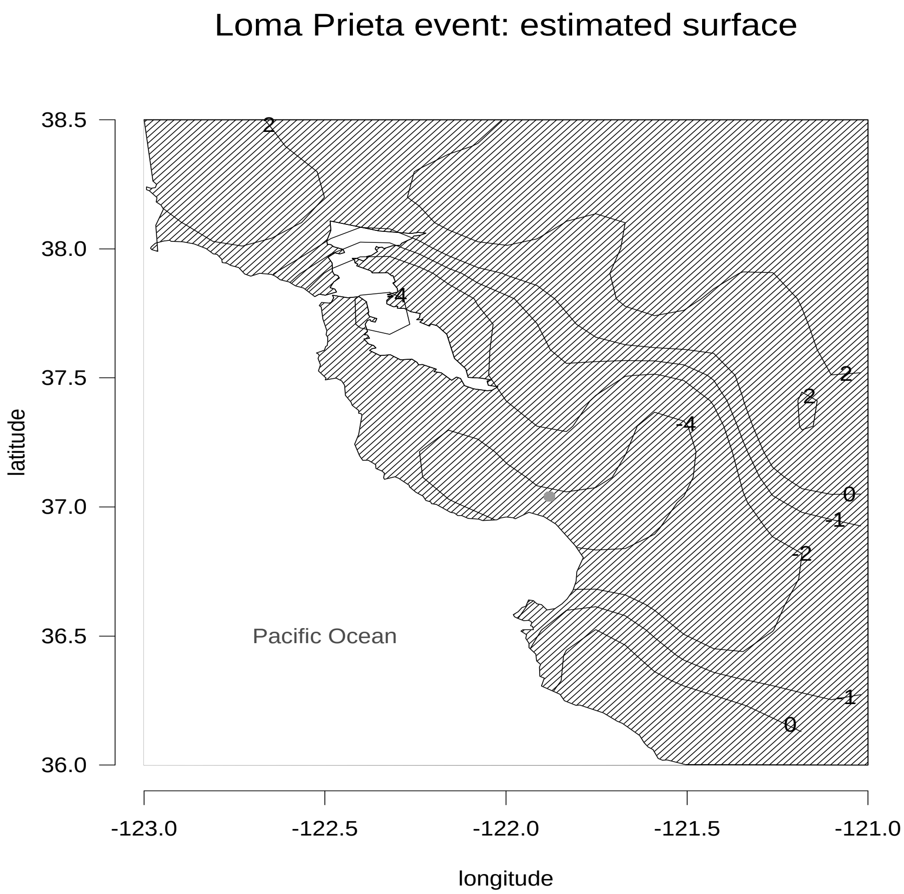

Figure 6: Estimated $g(x, y)$ of the model (10) using Shannon wavelets.

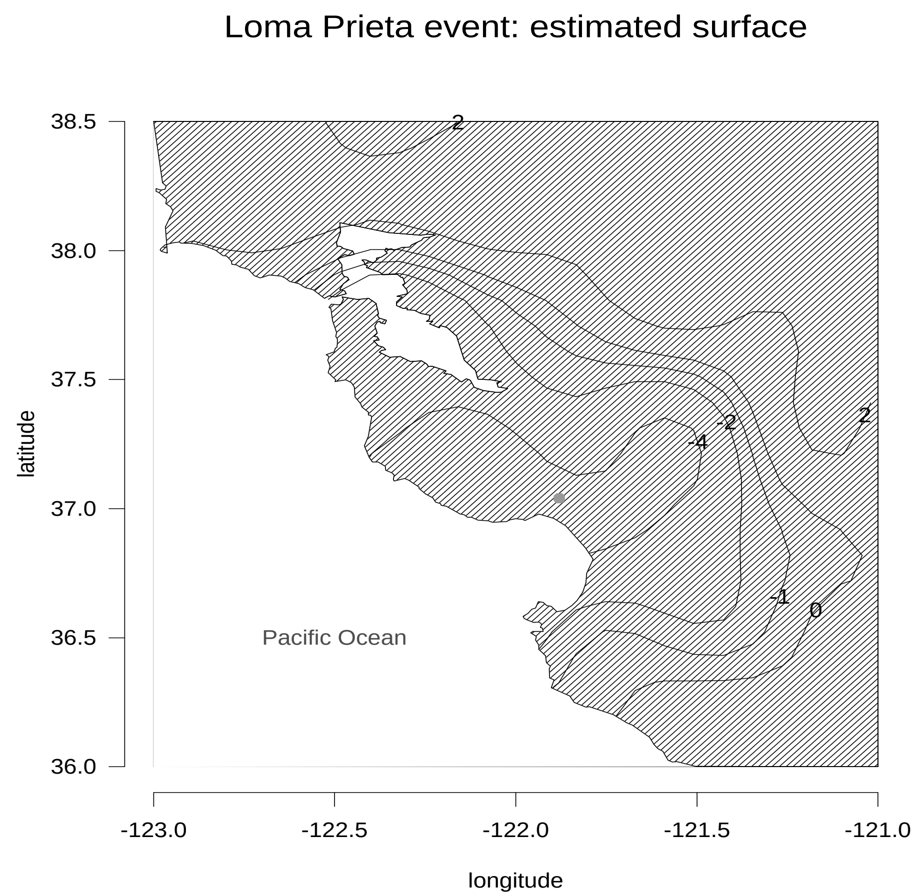

Figure 7: Estimated $g(x, y)$ of the model (10) using Shannon wavelets with shrinkage.