Tectonic Deformation Estimation using Stream Gradients: Nonparametric Function Estimation from Difference Data using Splines and Conjugate Gradients

Tim C. Hesterberg, Derek C. Stanford, and Dorothy J. Merritts

Hesterberg and Stanford:
Research Department, MathSoft, Inc.
1700 Westlake Ave., Seattle, WA 98109
timh@statsci.com, stanford@statsci.com
www.statsci.com/Hesterberg
Merritts: Department of Geosciences,
Franklin & Marshall College Lancaster, PA 17604
D_Merritts@FandM.edu

Abstract

In 1811-1812 three great (8.0+) earthquakes occurred near New Madrid, Missouri. We estimate coseismic deformation in this area using stream elevation data from topographic maps. Streams have a natural profile, the gradient of which depends on the resistance of underlying sediment and the volume of stream flow. If tectonic processes elevate the upstream end of a segment a different amount than the downstream end, the stream will attempt to return to its natural gradient by incising, aggrading, or altering its sinuousity. This adjustment takes time, so deviations from the natural gradient may indicate geologically recent deformation.

We use penalized regression splines to estimate the natural stream profile and the deformation of the ground surface Estimation of the natural profile and deformation is based on nonparametric regression of the form $y_2 - y_1 = f(x_2) - f(x_1)$. This may be formulated as a linear regression problem, potentially with millions of parameters when estimating large-scale surfaces; the system may be solved using conjugate gradient methods.

Keywords: B-splines, Geomorphology, Regression splines, SL index, Tectonics, Tensor product, Ridge Regression.

1 Introduction

A number of different features of the landscape have been used to estimate coseismic deformation that has occurred prior to recent geodetic surveys, including fault-bounded mountain fronts, offset stream channels, or uplifted coastlines along active plate margins. Use of such features has been unsuccessful in continental interiors because the terrain in such areas typically has little relief and faults rarely rupture the surface. For example, our study area is within the New Madrid Seismic Zone (NMSZ), in an area mantled by more than 1 km of alluvial sediment that obscures crustal displacement.

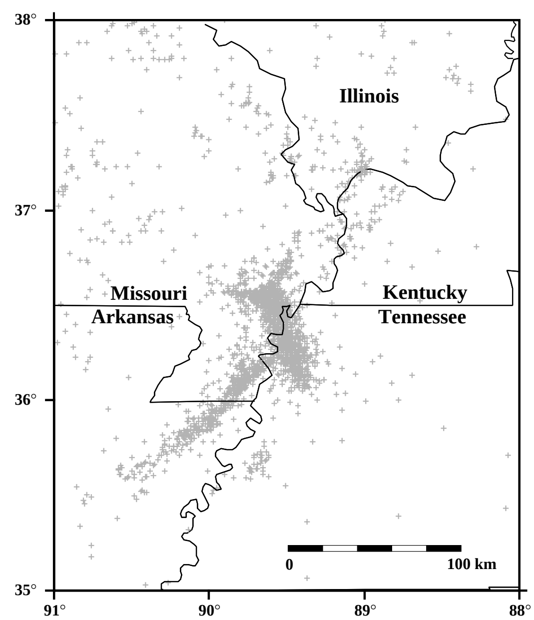

Figure 1: Earthquake epicenters of New Madrid region from 1974 to 1991.

Our purpose in this article is to describe a method of using topographic information for stream networks to identify crustal deformation. Stream channels have natural gradients that primarily depend on resistance to flow

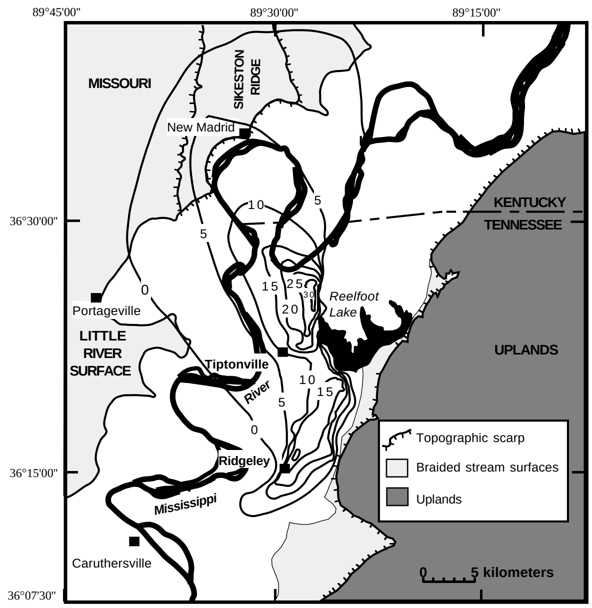

Figure 2: Surface contours (feet) in the Lake County Uplift Region, from Russ (1982).

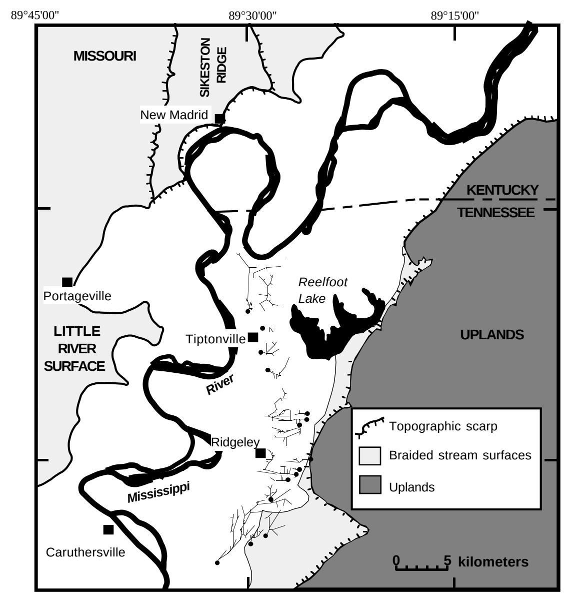

Figure 3: Watersheds analyzed in this study.

(dependent on bedrock or sediment type and properties) and volume of water and sediment flowing in the stream (Schumm, 1977, 1983). The gradient is steeper near the headwaters, and less steep downstream where the flow volume is greater. If tectonic processes change the gradient of an alluvial stream segment by elevating the upstream end of a segment more or less than the downstream end, the stream will attempt to return to its natural gradient, by incising, aggrading, or altering its sinuosity (Ouchi 1985). However, that process takes time, so that deviations of streams from their natural gradients could indicate recent tectonic deformation. Merritts (1987) and Merritts & Vincent (1989) found that stream networks are sensitive to slow rates of surface uplift and tilt, even where faults have not ruptured the surface. The focus of this article is on statistical methodology; further discussion of previous results is in Merritts & Hesterberg (1994), and new results in a companion article.

We begin with a description of our study area in Section 2. In Section 3 we describe a procedure for estimating the natural profile of streams in a study area, regardless of tectonic disturbances. In Section 4 we describe how to estimate the natural profile and surface deformation simultaneously. This regression procedure can extract a subtle signal from noisy data.

The procedures in Section 3 and 4 are based on using nonparametric regression from difference data, in which the response variable cannot be observed directly; instead the difference in values of the response variable (change in elevation over the course of a segment) is observed for pairs of values of the explanatory variable(s) (stream channel length and location of the start and end of each stream segment).

2 Lake County Uplift Area

In 1811 and 1812 the three largest historic earthquakes (magnitude 8.1-8.3) in the North American continental plate occurred in the NMSZ. Figure 1 shows the location of recent small earthquakes in this area, recorded by St. Louis University and the Center for Earthquake Research and Information; these may be aftershocks of the great earthquakes. Our study area is is the Lake County Uplift Area (LCU), a region within the NMSZ where surface deformation is known to have occurred. Russ (1982) used warped stream levies and other evidence to estimate the uplift pattern shown in Figure 2.

We digitized topographic data for 16 small watersheds that drain the LCU, shown in Figure 3, obtaining map coordinates (x and y), elevation z, and arclength of each segment l. We worked from 1:24,000 scale topographic maps obtained from the U.S. Geological Survey, which have a contour interval of five feet, using custom software *DigiStream* which is available from Merritts.

One of the smaller watersheds is shown in Figure 4.

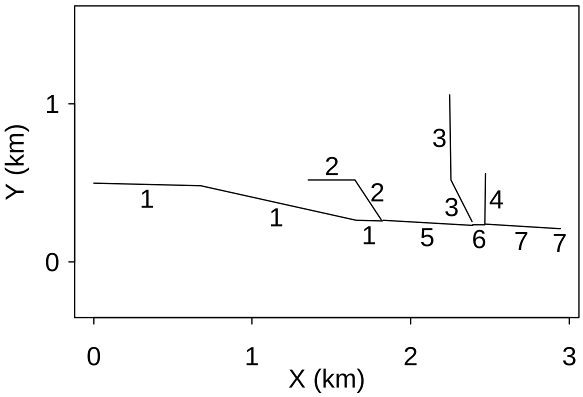

Figure 4: Example watershed with seven segments.

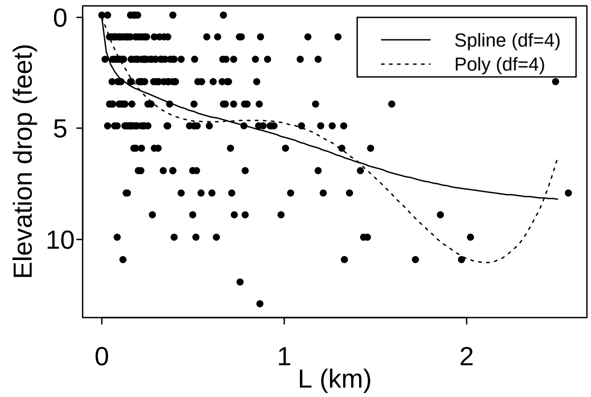

Figure 5: Elevation drops for feeder segments, with two estimates of the natural profile. L is the length of the segment.

There are seven segments in this watershed, Segment 7 is the outlet of the watershed.

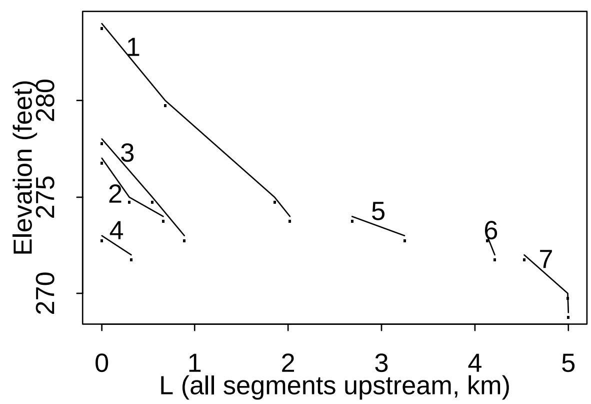

Figure 6: Longitudinal view of all segments in example watershed.

3 Natural Profile Estimation

We describe in this section a procedure for estimating the natural profile of streams in a study area, using nonparametric regression. We assume that the flow at any point on a stream network is a smooth function of (but not necessarily proportional to) the length of all segments upstream of the point. For comparison, the SL index (Hack 1973), the standard procedure for analyzing stream gradients, assumes that flow is a function of the length of the single longest sequence of segments upstream of a point. The SL index also assumes that the profile is given by a parametric curve $z = a - b\log(L)$. A third assumption, that flow is proportional to the drainage area upstream of any point, is attractive from physical considerations, but drainage area is more difficult to obtain.

Figure 5 shows a scatterplot of the elevation drop of feeder segments vs the length from the head of the segment. Feeder segments are those with nothing flowing into them, e.g. segments 1-4 in Figure 4.

Each segment is represented by plotting the elevation drop vs the length of the segment. It is evident that there is considerable variability in the elevation drops, but nonparametric regression, or scatterplot smoothing, may be used to obtain an estimate of the profile. Two such estimates are shown on the plot. A polynomial of degree 4, fitted by least-squares with no intercept (so the curve passes through the origin), produces an unacceptable estimate of the profile. A much better curve is obtained by a (cubic) smoothing spline with the equivalent of four degrees of freedom; see Hastie & Tibshirani (1990) for a discussion of this and other smoothing procedures, and of equivalent degrees of freedom. See also de Boor (1978) for a more complete discussion of splines. We also perform a square-root transformation on L prior to the smooth, and use an artificial point at the origin with weight 1000 to ensure that the curve passes near the origin (no such artificial point is needed with the regression splines discussed below).

The curves in Figure 5 are estimated using only feeder segments. Adding downstream segments results in longitudinal profiles like the one shown in Figure 6. The beginning elevation of segment 5 matches the end elevation of segments 1 and 2, and the value of L at the upstream end of segment 5 matches the sum of lengths of segments 1 and 2. We can move the individual segments up or down without changing the gradients. In Figure 7 the segments of the sample watershed are translated vertically so that they begin on the estimated profile from Figure 5. For this particular watershed the early segments are less steep, and the later segments steeper, than the natural profile estimated using feeder segments from all watersheds.

In order to incorporate information from downstream

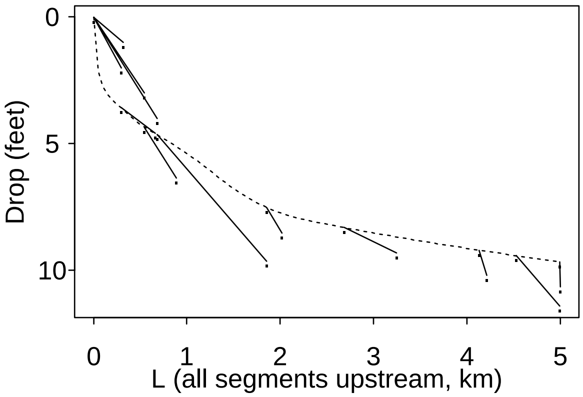

Figure 7: Longitudinal profile of example watershed, curve estimated from feeder segments from all watersheds.

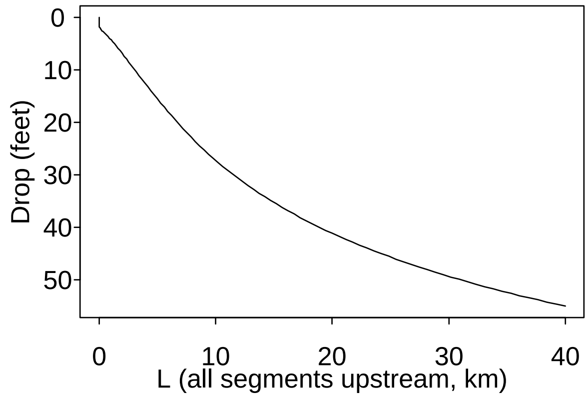

Figure 8: Estimated natural profile, using all segments and nonparametric estimation on difference data.

segments in estimating a natural profile we must abandon simple scatterplot smoothing, because downstream segments consist of pairs of observations which may be translated vertically arbitrarily. The natural profile should reproduce the observed elevation drops as closely as possible for all combinations of L_1 and L_2, the values of L at the upstream and downstream ends of the segment. We estimate the profile f using a model of the form

$$\text{drop} = z_1 - z_2 = f(L_1) - f(L_2) + \epsilon \qquad (1)$$

where f is a smooth nonparametric function (with $f(0) = 0$), and z_1 and z_2 are the elevations at the ends of the segment.

One way to fit this model would be to include each segment in the regression as two observations, one for the upstream and one for the downstream end, with a single dummy variable to allow for vertical translation, giving a model of the form

$$z_{(i,j)} = f(L_{(i,j)}) + \sum_{k=1}^{n} \beta_k D_k + \epsilon \qquad (2)$$

where $j = 1, 2$ for the upstream and downstream ends of a segment, and D_i is a dummy variable which is one for observations $(i, 1)$ and $(i, 2)$ and zero otherwise, $i = 1, \ldots, n$. In a dataset with n segments, this would require a regression on $2n$ observations using n dummy variables. This would be computationally expensive to fit using standard linear regression procedures if n is large, but might be feasible using a backfitting procedure (Hastie & Tibshirani 1990).

Our approach is to use regression splines, in particular B-splines. We use cubic B-splines s_j, which form a basis for the set of cubic splines, i.e. any cubic spline can be represented in the form

$$f(x) = \beta_0 + \sum_{j=1}^{J} \beta_j s_j(x). \qquad (3)$$

Regression splines are described in Hastie & Tibshirani (1992), and implemented in *S-Plus* (Chambers & Hastie 1992) as well as in other software such as *Matlab*. We choose the degrees of freedom J and the knots in advance, and transform L prior to computing the splines, e.g. by a square root transformation. Then (1) reduces to a regression of the form

$$z_1 - z_2 = \sum_{j=1}^{J} \beta_j (s_j(\sqrt{L_1}) - s_j(\sqrt{L_2})) + \epsilon \qquad (4)$$

which can easily be estimated by ordinary linear regression; the (i, j) element of the design matrix is the difference $s_j(\sqrt{L_{(i,1)}}) - s_j(\sqrt{L_{(i,2)}})$. There is no intercept in the regression. For comparison, in fitting a polynomial $f(x) = \sum_{j=0}^{J} \beta_j x^j$ to difference data, the regression would be of the form

$$\Delta z \equiv z_1 - z_2 = \sum_{j=1}^{J} \beta_j (x_2^j - x_1^j) + \epsilon \qquad (5)$$

We also use a dummy variable, which is 1 for feeder segments and 0 for other segments. This allows the estimated curve to drop sharply near $L = 0$. The resulting curve is shown in Figure 8. The gradient is very steep at first, then gradually levels off.

4 Deformation Estimation

Deformation of the earths surface during the 1811-1812 earthquakes caused the upstream ends of some segments

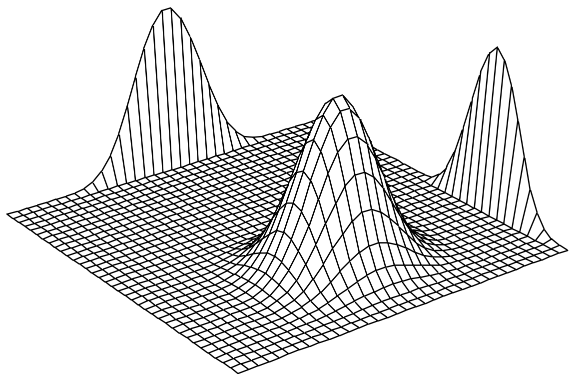

Figure 9: Tensor product B-splines. The product spline is $g_{m,k}(x,y) = s_m(x)S_k(y)$.

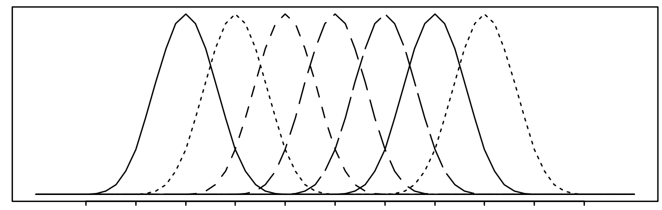

Figure 10: Basis for splines

to be elevated a different amount than the downstream ends, so those segments became more or less steep than their natural profile. We assume that streams have not had time to adjust their channels to return to their natural profile; this is more likely to be true for smaller streams, such as we analyze, than for larger streams. Then the actual drop of any segment should consist of the drop determined by the natural profile, plus the differential elevation gain from tectonic deformation. Our model for this is

$$\Delta z = f(L_1) - f(L_2) + g(x_1,y_1) - g(x_2,y_2) + \epsilon \quad (6)$$

where (x_1,y_1) and (x_2,y_2) are the coordinates (longitude and latitude) of the upstream and downstream ends of a segment, and g represents the deformation surface. We represent g using tensor product splines — if $s_m(x), m = 1,\ldots,M$ and $S_k(y), k = 1,\ldots,K$ are B-splines in x and y, then the tensor product splines are $g_{m,k}(x,y) = s_m(x)S_k(y)$. One such product spline is shown in Figure 9. Then g is of the form

$$g(x,y) = \sum_{m,k} \beta_{m,k} s_m(x) S_k(y). \quad (7)$$

Incorporating this into the regression results in a model of the form

$$\begin{aligned} \Delta z &= \textstyle\sum_{j=1}^{J} \beta_j(s_j(\sqrt{L_1}) - s_j(\sqrt{L_2})) + \eta I(\text{feeder}) \\ &\quad + \textstyle\sum_{m,k} \beta_{m,k}(g_{m,k}(x_1,y_1) - g_{m,k}(x_2,y_2)) + \epsilon \quad (8) \end{aligned}$$

In our application we are willing to assume that seismic deformation is zero outside of a study area, so use as components of the bases only functions $s_m(x)$ and $S_k(y)$ that are zero outside of a limited range. We also use equally-spaced knots, and use only basis components that are translates in the x and/or y directions of any single product spline. An example is shown in Figure 10. This provides a full basis for all splines only in the range inside the fourth knot from each border; if necessary we extend the knot grid so that this range includes the desired region. Use of these repetitive basis functions for modeling surfaces simplifies many calculations considerably.

The number of parameters in the model can be very large. If the knots in x and y are placed in a reasonably fine grid then the number of product splines, and degrees of freedom in the regression, can be very large; for example a 1000 by 1000 grid results in approximately 10^6 parameters. There are typically problems with colinearity or near-colinearity. Some product splines (9) are located in regions of the map with no segment ends; the corresponding columns in the linear model design matrix would be identically 0. In Hesterberg and Merritts (1995) we discuss ad-hoc remedies for these problems. Here we discuss a more systematic solution.

Our remedy is to used penalized regression, where the penalty reflects the lack of smoothness of a fit; the objective function is

$$\sum_{i=1}^{n} \hat{\epsilon}_i^2 + \lambda_f \int (f'')^2 + \lambda_g \int (g'')^2 \quad (9)$$

where $\hat{\epsilon}_i$ are the residuals and $(g'')^2$ is shorthand for $\left(\frac{\partial^2 g}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 g}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 g}{\partial y^2}\right)^2$.

These penalties are analogs of the penalties commonly applied when smoothing non-difference data. For simple scatterplot smoothing, for example, the objective $\sum_{i=1}^{n} \hat{\epsilon}_i^2 + \lambda_f \int (f'')^2$ has smoothing splines as a global solution; regression splines may be viewed as an approximation to smoothing splines. Similarly for smoothing against bivariate data, the global solution involves thin-plate splines.

The amount of penalty can be varied, to produce smoother and flatter surfaces at one extreme ($\lambda \to \infty$), or unflattened and unstable surfaces at the other ($\lambda \to 0$).

Each integral in 9 is a quadratic function of the parameters,

$$\inf(f'')^2 = \sum_{j_1}\sum_{j_1}\beta_{j_1}\beta_{j_2}\mathbf{M}f, j_1, j_2 = \beta_f'\mathbf{M}_f\beta_f$$

$$\inf(g'')^2 = \sum_{u_1}\sum_{u_2}\beta_{u_1}\beta_{u_2}\mathbf{M}g, u_1, u_2 = \beta_g'\mathbf{M}_g\beta_g \quad (10)$$

$$(11)$$

where $\mathbf{M}_f$ and $\mathbf{M}_g$ are positive definite matrices (not strictly positive definite; eigenvalues may be zero), $\beta_f = \{\beta_j\}$ for $j = 1, \ldots, J$, $\beta_g = \{\beta_u\}$ for $u = (m, k)$ with $m = 1, \ldots, M$ and $k = 1, \ldots K$. See Section 6 for further detains on these matrices.

The penalized least-square solution is:

$$\hat{\beta} = \left(\mathbf{X}'\mathbf{X} + \mathbf{M}\right)^{-1}\mathbf{X}'\mathbf{Y} \quad (12)$$

where $\mathbf{Y}$ is the response (vector of values of Δz), $\mathbf{X}$ is the usual design matrix, with columns corresponding to unpenalized terms (indicator variable for feeder segments) as well as f and g, β is the vector of all parameters (unpenalized, f, and g), and

$$\mathbf{M} = \begin{pmatrix} 0 & & 0 \\ 0 & \lambda_f\mathbf{M}_f & 0 \\ 0 & 0 & \lambda_g\mathbf{M}_g \end{pmatrix} \quad (13)$$

Use of the normal equations (12) is a computationally-inaccurate method of solving regression problems. An alternate approach is to augment $\mathbf{X}$ and $\mathbf{Y}$ with artificial observations,

$$\begin{pmatrix} \mathbf{Y} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} & \mathbf{X} & \\ 0 & \mathbf{Q}_f & 0 \\ 0 & 0 & \mathbf{Q}_g \end{pmatrix}\beta + \begin{pmatrix} \epsilon \\ \epsilon_f \\ \epsilon_g \end{pmatrix} \quad (14)$$

where $\mathbf{Q}^{(f)}$ and $\mathbf{Q}_g$ are square roots of $\lambda_f\mathbf{M}_f$ and $\lambda_f\mathbf{M}_g$, e.g. $\mathbf{Q}^{(f)\prime}\mathbf{Q}^{(f)} = \lambda_f\mathbf{M}_f$, ϵ is the vector of ordinary residuals, and ϵ_f and ϵ_g the artificial residuals corresponding to smoothness penalties for f and g. Note that $\int(f'')^2$ and $\int(g'')^2$ are represented as sums of squares of the artificial residuals. This system may be solved with any numerically-accurate procedure.

However, both (12) and (14) are impractical if the number of parameters is large.

5 Conjugate Gradient Solution

We turn to conjugate gradients methods for a solution. This method allows us to avoid creating large matrices; calculations involving the large matrices ($\mathbf{X}$ and $\mathbf{M}$) are done without actually creating the matrices, in a way that makes use of the sparsity of these matrices that arises from the properties of the spline bases.

Conjugate gradient is an iterative method, that for quadratic problems with p parameters obtains an exact solution (modulo numerical error) in p steps. In practice it may converge to a desired tolerance faster, particularly in well-conditioned problems. Let ϕ be the objective function to be minimized, in this case (9). Let β_i be the estimated parameter vector at iteration i, and then the steps in each iteration are:

Step 1 Find the gradient G. First calculate the residuals $\epsilon_i = \Delta z - \mathbf{X}\beta_i$, then

$$G_i = -2\mathbf{X}'\epsilon_i + 2\mathbf{M}\beta_i \quad (15)$$

The products $\mathbf{X}\beta_i$, $\mathbf{X}'\epsilon_i$ and $\mathbf{M}\beta_i$ are calculated without forming $\mathbf{X}$ and $\mathbf{M}$.

Step 2 Compute the conjugate direction H_i. The first step uses $H_1 = G_1$. For $i > 1$

$$H_i = G_i + \frac{||G_i||^2}{||G_{i-1}||^2}G_{i-1} \quad (16)$$

where $||v||^2 = \sum v_j^2$.

Step 3 Compute α, the ratio of the first directional derivative to the second directional derivative,

$$\alpha_i = \frac{\nabla\phi H_i}{H_i'\nabla^2\phi H_i} = \frac{G_i'H_i}{2(\mathbf{X}H_i)'(\mathbf{X}H_i) + 2H_i'(\mathbf{M}H_i)} \quad (17)$$

The $\mathbf{X}H$ and $\mathbf{M}H$ terms are calculated without forming $\mathbf{X}$ and $\mathbf{M}$. Note that the denominator is written in a way which avoids calculation of $\mathbf{X}'\mathbf{X}$.

Step 4 Update the estimate of B using

$$\beta_{i+1} = \beta_i + \alpha_i H_i \quad (18)$$

6 Sparse Matrix Computations

A key feature of the conjugate gradient calculations is that all calculations of the form $\mathbf{X}v$ and $\mathbf{M}v$, for any vector v, are performed without explicitly forming $\mathbf{X}$ or $\mathbf{M}$. The implementation makes use of the sparsity of these matrices, due to that fact that the basis functions we use for cubic splines are nonzero in only four inter-knot intervals, so that any particular value (x, y, or $\sqrt{L}$) results in nonzero values for four of the basis functions. Consider first the matrix $\mathbf{X}$. Each row is nonzero in at most 8 of the columns corresponding to f (4 each for L_1 and L_2), and is nonzero in at most 32 of the columns corresponding to g (16 for each end of the stream segment).

Products of $\mathbf{X}$ and v need sum only over the nonzero entries.

Next consider the penalty matrices $\mathbf{M}_f$ and $\mathbf{M}_g$. The integral $\int (f'')^2$ can be decomposed into intervals,

$$\int_R (f'')^2 = \sum_j \int_{R_j} (f'')^2 \tag{19}$$

where R_j is the interval between knots $j-1$ and j. Since $f''(l) = \sum_j \beta_j s_j''(l)$ involves a summation over at most four nonzero (and adjacent) terms for all l in any interval R_j, $\int_{R_j} (f'')^2 = \sum_{j_1} \sum_{j_2} \beta_{j_1} \beta_{j_2} \int_{R_j} s_{j_1}''(l) s_{j_2}''(l)$ with $|j_1 - j_2| \leq 3$ for the nonzero terms in any region. The result is that $\mathbf{M}_f$ is a banded matrix, with 3 nonzero entries above and below the diagonal.

Similarly, for $\mathbf{M}_g$, the integrals can be decomposed into rectangles,

$$\int_A ||g''||^2 dA = \sum_u \int_{A_u} ||g''||^2 dA_u \tag{20}$$

with u indexing all grid regions (in two dimensions). Within each grid region A_u, at most 16 spline products are nonzero. Through some algebraic manipulations, it can be shown that the smoothness term can be written in as $(\beta' \mathbf{M}_g \beta)$, where $\mathbf{M}_g$ is a sparse multiply-banded matrix containing integrals of of products of spline basis functions and their first and second derivatives. The repetitive nature of the component basis functions and regular spacing of the knots causes the nonzero terms to have only a few distinct values, which can be computed once and reused as needed. By using the structure of $\mathbf{M}$, $\mathbf{M}v$ can be computed quite efficiently; the time required increases linearly with the number of parameters (length of v).

For a fixed number of observations, as the grid for f or g becomes finer, the total running time is $O(p^2)$ where p is the total number of parameters, if the algorithm is run for the full p iterations. In contrast, methods that invert or multiply $p \times p$ matrices are $O(p^3)$.

It may be possible to reduce the number of iterations required for convergence (to within a desired accuracy) using preconditioning, using a wavelet transformation of the parameters corresponding to f and a bivariate wavelet transformation of parameters corresponding to g.

7 Limitations of the Procedure

The conjugate gradient procedure produces only parameter estimates, not various quantities useful in statistical inference, such as standard errors for parameter estimates. Our current remedy is to bootstrap the whole procedure. This typically requires 50 bootstrap replications to estimated standard errors with reasonable accuracy.

However, this bootstrap procedure is still under development. The simple bootstrap treats observations as independent, but there appears to be dependence between adjacent upstream and downstream segments. This might be handled using bootstrap methods that take dependence into account, but we have not done so yet.

Nor are the observations identically distributed; the residual variance is larger for feeder segments and other segments with relatively small values of L_1 than for those with large values of L_2. This may be handled using a weighted regression. This may be implemented by in the conjugate gradient method by straightforward modifications of (15) and (17).

Another current limitation is that we have not found an efficient way to determine the equivalent degrees of freedom for the f and g terms (see Hastie and Tibshirani 1990), for arbitrary values of the smoothing parameters λ_f and λ_g. That would provide a way for an analyst to choose the smoothing parameters. We currently choose smoothing parameters interactively, viewing the natural profile estimates and surfaces produced by different combinations, choosing values that appear to give reasonable results.

The procedure is unfinished; it does not currently support estimation of non-penalized terms, f, and g simultaneously. Results below are based on combinations of the normal scores and augmentation approaches.

8 Results

In earlier work (Merritts and Hesterberg 1994) we focused on the Lake County region of the NMSZ, where the 1811-1812 earthquakes occured. Results from that work are shown in contour plot in Figure 11. Uplift is generally greatest along the middle (from west to east) of the study area, with low areas about one-quarter and one-half the distance from north to south. The pattern exhibits four distinct high areas, three of which generally coincide with areas of uplift determined by Russ (1982). However, fine details might result from the inherent randomness in stream channel elevations or the limitations of working from maps with 5-foot contour intervals in a region of very little relief. The contours are less accurate on the edges of the study area. The overall amount of uplift shown is less than determined by Russ (1982), probably because of the flattening effect of the ridge regression.

That work was based on a simpler form of ridge regression, using a diagonal penalty matrix in place of $\mathbf{M}_g$; the result is to shrink the deformation estimate toward zero everywhere, rather than penalize roughness.

A bootstrap analysis (not described here) shows that the deformation estimate is highly statistically signifi-

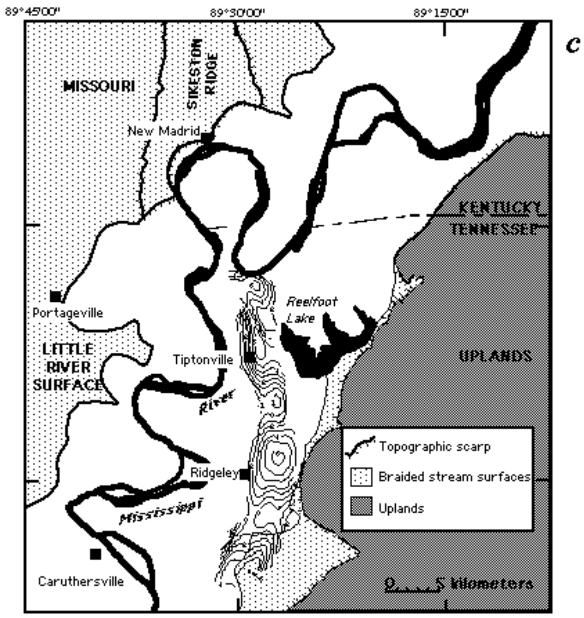

Figure 11: Estimate of surface deformation g (feet).

cant. Comparison of the deformation estimate with the topography shows that the deformation estimation procedure is not simply reproducing the contours of the land. Thus it appears that the procedure is finding some real effect in the landscape, and the agreement with the contours of Russ suggests that it is the recent tectonic deformation.

In newer work we have redigitized the Lake County Uplift region, using a newer version of the *DigiStream* software, and collected data in three additional, larger, regions within the NMSZ. A short summary of the areas is found in Table 1 Note that the other regions include larger areas and larger numbers of stream segments than the LCU. Furthermore, while earlier work involved only stream segments (a segment is the part of a stream between intersections with other streams), newer work also involves subsegments (a segment is broken into subsegments where it crosses a contour line); these more extensive data offer the possibility of greater accuracy. The combination of larger datasets, larger regions, and the desire to obtain greater accuracy by using a finer grid, motivates the development of estimation procedures that scale up better.

8.1 Fort Pillow

Fort Pillow is a control area, an area of considerable topographic relief (relative to most of the NMSZ), where it is believed that the relief is not caused by co-seismic deformation.

Figure 12 shows the stream network in this area.

Table 1: Areas within the New Madrid Seismic Zone

	LCU	Benton	FP	Crowley
Segments	294	2463	839	
Subsegments	396	9536	2981	76122
range(x) (km)	3.0	20	28	
range(y) (km)	8.5	21	14	
range(z) (m)	18	79	68	
mean(length)	479	94	502	
mean(Δz)	.97	2.9	9.0	

Summary of data collected in four areas within the NMSZ — Lake County Uplift region, Benton Hills, Fort Pillow, and Crowley's Ridge. It is known that there is seismic uplift in the LCU and suspected in Benton Hills and Crowley's Ridge; Fort Pillow is a control region. A segment is the part of a stream between intersections; this is broken into subsegments when the segment crosses a contour line. The last two rows refer to the average length and elevation drop for segments. text

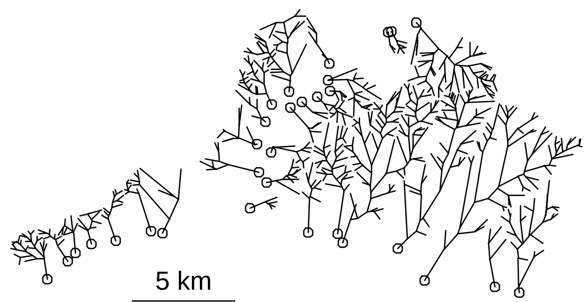

Figure 12: Fort Pillow Stream Network

There are 31 watersheds; the outlets of each are shown with small circles. The arm to the left is an area of considerable relief, as are upstream areas in the rest of the region.

Figures 13 and 14 are obtained using a model of the form (8), using a relative coarse (1 km) grid for estimating deformation, with penalty terms that result in the equivalent of 3.1 and 103 degrees of freedom for estimating f and g, respectively (the latter is a considerable reduction from the 352 degrees of freedom that would be used for g in the absence of any penalty). Parameters were chosen to minimize a weighted sum of squares, with smaller weights for feeder segments; the weights are inversely proportional to the residual variances from the two groups obtained from a model excluding g.

Results shown here are preliminary. Exploratory data analysis and interactive model building have suggested some ways in which the model fit might be improved,

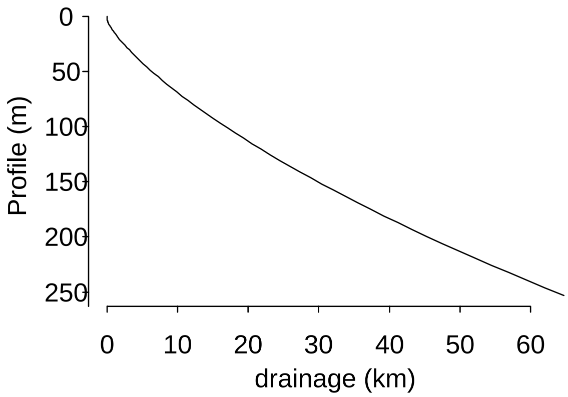

Figure 13: Fort Pillow-Estimated Natural Profile

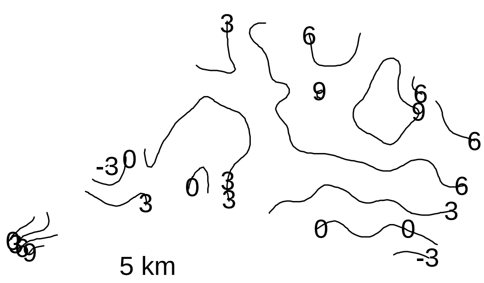

Figure 14: Fort Pillow-Estimated Deformation

while raising additional scientific questions.

9 Summary

This procedure can be extended in a number of ways. Differential flow resistance could be estimated by adding a spatial term to the model (6), of the form $(L_2 - L_1)h(x, y)$. Because the procedures described here are based on linear regression they can be made robust in the usual manner, using iterative weighted least squares to reduce the weights on observations with large residuals.

Acknowledgments

We wish to thank Harry Haddon for creating the *DigiStream* software used during the digitization process, and Denise Muriceak, Ed Stoner, Vasilios Lambrinos, and Spiro Gavaris for collecting the data and other assistance, T. Hastie for suggesting regression splines for estimating the natural profile of streams, and several geologists with expertise in the NMSZ for advice, suggestions, and information, including R. Wheeler, E. Schweig, M. Ellis, P. Bodin, and A. Crone. This work was supported by U.S. Geological Survey under award numbers 1408-93-G-2230 and 1434-95-G-2624, and by the National Science Foundation under award number EAR-9418681.

References

R. A. Becker, J. M. Chambers, and A. R. Wilks (1988), *The New S Language*, Pacific Grove, CA: Wadsworth and Brooks/Cole.

Chambers, J. M. and T. J. Hastie (1992), *Statistical Models in S*, Pacific Grove, CA: Wadsworth and Brooks/Cole.

de Boor, C. (1978), *A Practical Guide to Splines*, New York: Springer Verlag.

Hack, J. T. (1978), "Stream-Profile analysis and stream-gradient index," *U.S. Geological Survey Journal of Research" v. 1, p. 421-429.*

Hastie, T. J. and R. J. Tibshirani (1990), Generalized Additive Models *Chapman & Hall, New York, NY.*

Hesterberg, T. C. and Merritts, D. J. (1995), Tectonic Deformation Estimation using Stream Gradients: Non-parametric Function Estimation from Difference Data, Proceedings of the Section on Physical and Engineering Sciences, *American Statistical Association, 85–90.*

Merritts, D. (1987), Geomorphic response to late Quaternary tectonism, coastal northern California, Mendocino triple junction region," Ph.D. Thesis, *Univ. of AZ, Tucson.*

Merritts, D. and Hesterberg, T. C. (1994), Stream Networks and Long-Term Surface Uplift in the New Madrid Seismic Zone, Science **265**, *1081–1084.*

Merritts, D. and K. R. Vincent (1989), "Geomorphic response of coastal streams to low, intermediate, and high rates of uplift, Mendocino triple junction region, northern California," Geol. Soc. Am. Bull. *101, 1373.*

Russ, D. P. (1982), in Investigations of the New Madrid, Missouri, earthquake region *, F. A. McKeown and L. C. Pakiser, Ed., U.S. Geol. Surv. Prof. Pap., 1236,95-113 (1982)*

Schumm, S. A. (1977), The fluvial system *Wiley, New York.*

S. A. Schumm (1983), "Alluvial river response to active tectonics," in Active Tectonics, *National Academy of Sciences Press, Washington, DC, p. 80-94.*

Statistical Sciences Inc. (1991); S-PLUS Reference Manual, *Version 3.0, Seattle: Statistical Sciences, Inc.*