Reproducing Earth's Kernel: Uncertainty of the Shape of the Core-Mantle Boundary from PKP and PcP Travel-Times

By

Philip B. Stark

Nicolas W. Hengartner

Technical Report No. 322
February 1992
(revised August 1992)

Department of Statistics
University of California
Berkeley, California 94720

Reproducing Earth's Kernel: Uncertainty of the Shape of the Core-Mantle Boundary from PKP and PcP Travel-Times

Philip B. Stark

Nicolas W. Hengartner

Department of Statistics

University of California at Berkeley

August 7, 1992

Abstract

Within ray theory, traveltimes sample a finite set of points on the core-mantle boundary (CMB). Only if the CMB is smooth can we estimate its shape from traveltimes. The assumed degree of smoothness affects the uncertainty of the estimate. We find lower bounds on the uncertainty of the CMB shape by assuming (i) traveltime errors are independent and identically distributed (iid) with zero mean and known variance, (ii) linearization does not compromise accuracy, (iii) event locations are exact, (iv) radial velocity follows the IASP91 model, and (v) the only three-dimensional structure is a low-degree spherical harmonic perturbation to the CMB. We compute

confidence intervals and simultaneous confidence envelopes for the CMB with a new technique based on reproducing kernels. Simultaneous confidence envelopes have a specified probability of containing the entire CMB, rather than just one point on the CMB, and thus allow inferences about the CMB shape. Suppose the standard deviation of traveltime errors is $0.25sec$ (it is probably closer to $1.0sec$), and that the CMB has a degree 4 spherical harmonic expansion. Then the width of a 95% confidence envelope around the least-squares estimate of the CMB using 776,958 PKP phases and 84,947 PcP phases ranges from $0.32km$ under Europe and Tonga to a maximum of $1.06km$ under the mid-Pacific, with a median width of $0.58km$. If the CMB has structure up to degree 20, the envelope width ranges from $8.14km$ to $245.0km$. It is wider than $40.5km$ for 50% of the CMB, and wider than $67.3km$ over 25% of the CMB. We construct a degree 20 CMB model extending almost $93km$ into the lower mantle beneath the mid-Pacific that changes traveltimes by only $0.1sec$ rms.

Introduction.

In recent years there have been many depictions of three-dimensional structure in Earth's deep interior using ISC traveltime data [*Bulletin of the International Seismological Centre*, 1967-1989] , for example, *Clayton and Comer* [1983], *Creager and Jordan* [1986], *Doornbos and Hilton* [1989], *Dziewonski and Woodhouse* [1987], *Dziewonski* [1984], *Dziewonski et al.* [1977], *Gudmundsson* [1989], *Inoue et al.* [1990], *Morelli and Dziewonski* [1987], *Pulliam et al.* [1992], and *Sengupta and Toksöz* [1976]. These models include three-dimensional structure in the mantle, topography of the core-mantle boundary (CMB) and anisotropy of the inner core.

We present a method to find optimistic uncertainty estimates as a function of position

in linear and linearized inverse problems, subject to smoothness constraints on the unknown model. We apply the method to an idealized problem to find uncertainties in the CMB from ISC data. Smoothness constraints are essential in ray-theoretic tomography, since rays sample zero-volume paths in the medium and thus the data have no information about off-raypath structure. (*Stark and Nikolayev* [1992] study the sensitivity of traveltimes to off-raypath structure using the Born approximation.) However, if velocity varies smoothly, velocity near a raypath resembles velocity on the raypath, and we can estimate more of the structure. Current methods smooth by regularizing and by parametrizing velocity structure with truncated spherical harmonic expansions or pixels with finite dimensions. Spatial uncertainty estimates are quite sensitive to the assumed degree of smoothness, since only the smoothness controls off-raypath structure.

In our idealization,

A1 Seismic velocity varies only with depth, except that the CMB may have a three-dimensional perturbation

A2 The raypaths given by the IASP91 model [*Kennett and Engdahl*, 1991] are accurate

A3 The CMB is smooth in a sense made precise below

A4 Source relocations in the IASP91 model are perfect

A5 Traveltimes are accurate to rms σ *sec.*

Complications including more general three-dimensional structure, uncertainties in the IASP91 model, nonlinearity, lack of compelling information about the smoothness of the CMB, and inaccuracies in source locations make the uncertainty in the real tomographic

reconstruction problem larger. Thus we find a lower bound on the difficulty finding the CMB from traveltime data. The theory holds for three-dimensional reference models as well, but the cost of ray tracing is prohibitive at present.

Consider the following measures of uncertainty of the CMB as a function of angular position $\hat{r}$:

U1 The size of a bump on the CMB at $\hat{r}$ whose effect on the data is indistinguishable from noise

U2 The largest difference at $\hat{r}$ between two CMB models that produce nearly the same data

U3 The largest difference at $\hat{r}$ between two CMB models that both fit the data within the noise level

U4 The width of a confidence interval for a least-squares estimate of the CMB at $\hat{r}$

U5 The width of a confidence interval for the best possible estimate of the CMB at $\hat{r}$

U6 The width at $\hat{r}$ of a simultaneous confidence envelope around the least-squares estimate of the entire CMB.

These measures of uncertainty have the same spatial pattern as a function of $\hat{r}$; they are equal up to a constant scaling factor. *Donoho* [1989] establishes the equivalence of measures U2 (the "modulus of continuity") and U5 quite generally for linear inverse problems with Gaussian noise. The other equivalences follow from additional linear structure in the idealized problem A1-A5. The easiest measure to compute is U1, which we show below can be found simultaneously for all $\hat{r}$ by constructing a reproducing kernel [*Aronszajn*, 1950] for the space of CMB models.

The most relevant measure is U6, since all current CMB estimates are based on least-squares, and simultaneous confidence envelopes tell us about the shape of the CMB. The question of "continents on the core" [*Creager and Jordan*, 1986] has to do with the presence and magnitude of "bumps" on the CMB. If we naively construct a confidence interval for each point on the CMB, the confidence level of the set of intervals taken together is not as large as the confidence level of the individual confidence intervals. The difference between U4 and U6 is the number of points we wish to estimate:

Suppose we wish to compare the height of the CMB at 10 points. If the estimates of CMB topography at these points were independent, and we had a 95% confidence interval for each of the points, the simultaneous confidence level for the set of intervals would be $0.95^{10} < 0.6$. To have simultaneous 95% confidence, we need to lengthen the individual intervals so that their confidence level is $0.95^{1/10} \approx 0.995$. The individual confidence level required to maintain a fixed simultaneous confidence level goes up as the number of intervals goes up. If estimates of CMB topography were independent at each point, we would need infinitely wide confidence intervals at each point to be able to say with 95% confidence that the CMB was within all the intervals.

However, if the CMB is smooth, least-squares estimates at different points are not independent, and it is possible to find a nontrivial "simultaneous confidence envelope" for the entire CMB—a pair of functions such that our confidence that the CMB lies between those functions is 95%. The assumed smoothness permits nontrivial confidence intervals at every point on the CMB without their simultaneous coverage probability going to zero. A simultaneous confidence envelope for the whole CMB allows us to test hypotheses such as the existence of continents on the core, because we can with confidence delineate the shape

of the CMB. Normalization constants for simultaneous confidence envelopes as well as the other measures of uncertainty are tabulated below.

Using the reproducing kernel approach we explore the dependence of the uncertainty on the assumed smoothness and clarify the dependence of the uncertainty on the source-receiver geometry. The method can be applied to mantle tomography (Pulliam and Stark, in prep.), exploration tomography, and other linear or linearized inverse problems with smoothness constraints. Extreme smoothness is necessary for an infinite-dimensional space to possess a reproducing kernel (equivalently, for the point-evaluator to be a bounded linear functional). We do not know any reason the CMB should be smooth, but we ignore that issue and proceed.

The Model Problem.

Mathematical Model for the Observations. We observe an n-vector of traveltime data $\boldsymbol{\tau}$ related to the unknown velocity model v_0 via a vector of nonlinear functionals $\mathbf{F}$, with additive noise $\boldsymbol{\epsilon}$:

$$\boldsymbol{\tau} = \mathbf{F}v_0 + \boldsymbol{\epsilon}, \tag{1}$$

where

$$\boldsymbol{\tau}, \boldsymbol{\epsilon} \in \mathbf{R}^n, \tag{2}$$

$$\mathbf{F} : \mathsf{V} \rightarrow \mathbf{R}^n, \tag{3}$$

$$v_0 \in \mathsf{V}, \tag{4}$$

and V is a linear vector space of Earth models. We assume (1) can be linearized about a reference model v_1 (e.g., IASP91):

$$\boldsymbol{\tau} = \mathbf{F}v_1 + \mathbf{G}(v_0 - v_1) + \boldsymbol{\sigma} + \boldsymbol{\epsilon}, \tag{5}$$

where $\mathbf{G} : \mathsf{X} \to \mathbf{R}^n$ is a vector of linear functionals depending on the reference model v_1, X is a Hilbert space of perturbations, and $\boldsymbol{\sigma}$ is the vector of linearization errors depending on v_0 and v_1. Define

$$x_0 = v_0 - v_1 \tag{6}$$

and

$$\boldsymbol{\delta} = \boldsymbol{\tau} - \mathbf{F}v_1. \tag{7}$$

The linearized data relations are

$$\boldsymbol{\delta} = \mathbf{G}x_0 + \boldsymbol{\sigma} + \boldsymbol{\epsilon}. \tag{8}$$

The constraint that the CMB is smooth is

$$x_0 \in \mathsf{C}, \tag{9}$$

where the "constraint set" C is a convex subset of the model space X. We assume for simplicity (the theory applies more generally) that the perturbation x_0 is just a perturbation to the CMB, and write $x_0(\hat{r})$ as a function on the unit sphere $\{\hat{r} : \|\hat{r}\| = 1\}$. Let X be the space of square-integrable CMB models:

$$\mathsf{X} \equiv \left\{ x(\hat{r}) : \int_{\|\hat{r}\|=1} |x(\hat{r})|^2 < \infty \right\}. \tag{10}$$

Then the spherical harmonics $\{Y_l^m\}_{l=0}^{\infty}{}_{m=-l}^{l}$ comprise a basis for X. The norms on the model space and data space will be denoted $\|\cdot\|$. We assume a bound on the two-norm of the observational errors:

$$\|\boldsymbol{\epsilon}\|_2 \leq \chi. \tag{11}$$

The choice of χ is discussed in detail below. It depends on the statistics of the traveltime errors and on which measure U1-U6 we choose. We neglect the linearization errors, which make the uncertainty larger, so our results remain optimistic.

Uncertainties for Linear Functionals.

Suppose we wish to know how well a linear functional $L : \mathbf{X} \to \mathbf{R}$ (such as the point-evaluator $Lx = x(\hat{s})$) can be determined from data $\boldsymbol{\delta} = \mathbf{G}x_0 + \boldsymbol{\epsilon}$, the constraint $x_0 \in \mathsf{C}$, where C is convex, and the knowledge that $\|\boldsymbol{\epsilon}\| \leq \chi$. If two models x and y both satisfy the data and lie in C, how different can Lx and Ly be? I.e., what is

$$U(\chi, L, \mathbf{G}, \mathsf{C}, \boldsymbol{\delta}) \equiv \sup\{|Lx - Ly| : \|\mathbf{G}x - \boldsymbol{\delta}\| \leq \chi, \|\mathbf{G}y - \boldsymbol{\delta}\| \leq \chi \text{ and } x, y \in \mathsf{C}\}? \quad (12)$$

(This is uncertainty measure U3.) To determine the uncertainty (12) before making measurements (in analogy to a confidence interval whose length is fixed before the observations are known), we must find the largest $U(\chi, L, \mathbf{G}, \mathsf{C}, \boldsymbol{\delta})$ that could arise from any $\boldsymbol{\delta} = \mathbf{G}z + \boldsymbol{\epsilon}$ when $z \in \mathsf{C}$ and $\|\boldsymbol{\epsilon}\| \leq \chi$:

$$U^*(\chi, L, \mathbf{G}, \mathsf{C}) \equiv \sup_{\substack{\boldsymbol{\delta} = \mathbf{G}z + \boldsymbol{\epsilon} \\ z \in \mathsf{C} \\ \|\boldsymbol{\epsilon}\| \leq \chi}} U(\chi, L, \mathbf{G}, \mathsf{C}, \boldsymbol{\delta}). \quad (13)$$

Define the "modulus of continuity of L" [*Donoho*, 1989]

$$\omega(\chi, L, \mathbf{G}, \mathsf{C}) \equiv \sup\{|Lx - Ly| : \|\mathbf{G}x - \mathbf{G}y\| \leq \chi \text{ and } x, y \in \mathsf{C}\}. \quad (14)$$

(This is uncertainty measure U2.) The modulus measures how much L can differ between two models in C that produce almost the same data. We claim that

$$U^*(\chi, L, \mathbf{G}, \mathsf{C}) = \omega(2\chi, L, \mathbf{G}, \mathsf{C}) \quad (15)$$

whenever the modulus is finite. We usually suppress the arguments $L, \mathbf{G}$, and C.

Proof. By the triangle inequality,

$$\{(x,y) : \|\mathbf{G}x - \boldsymbol{\delta}\| \le \chi,\ \|\mathbf{G}y - \boldsymbol{\delta}\| \le \chi\} \subset \{(x,y) : \|\mathbf{G}x - \mathbf{G}y\| \le 2\chi\}. \tag{16}$$

Thus $U^*(\chi) \le \omega(2\chi)$. For $\gamma > 0$, let x^γ and $y^\gamma \in \mathsf{C}$ satisfy $\|\mathbf{G}x^\gamma - \mathbf{G}y^\gamma\| \le 2\chi$ and $|Lx^\gamma - Ly^\gamma| \ge \omega(2\chi) - \gamma$. Define $z^\gamma \equiv (x^\gamma + y^\gamma)/2$. By the convexity of C, $z^\gamma \in \mathsf{C}$. We have

$$\|\mathbf{G}x^\gamma - \mathbf{G}z^\gamma\| = \left\|\frac{1}{2}(\mathbf{G}x^\gamma - \mathbf{G}y^\gamma)\right\| \tag{17}$$

$$\le \chi, \tag{18}$$

using the linearity of $\mathbf{G}$ and the positive scalar homogeneity of the norm. The same inequality holds for $\|\mathbf{G}y^\gamma - \mathbf{G}z^\gamma\|$, so for $\boldsymbol{\delta}^\gamma \equiv \mathbf{G}z^\gamma$,

$$\omega(2\chi, L, \mathbf{G}, \mathsf{C}) \ge U^*(\chi, L, \mathbf{G}, \mathsf{C}) \ge U(\chi, L, \mathbf{G}, \mathsf{C}, \boldsymbol{\delta}^\gamma) \ge \omega(2\chi, L, \mathbf{G}, \mathsf{C}) - \gamma. \tag{19}$$

But γ was arbitrary, so (15) holds. □

Thus the optimization problem (14) gives the uncertainty of estimating Lx_0 corresponding to measures U2 and U3. We specialize now to $Lx = x(\hat{s})$, the point-evaluator.

Uncertainties for Local Values When C *is a Finite-Dimensional Subspace.*

Suppose C is a finite-dimensional subspace of X: the span of the spherical harmonics of degree $l \le l_{max}$. On C introduce the seminorm

$$\|x\|_G \equiv \sqrt{\mathbf{G}x \cdot \mathbf{G}x}, \tag{20}$$

which may fail to be a norm if it is positive-semidefinite rather than positive-definite. This can be checked as follows:

Define the bilinear functional

$$< x|y >_G \equiv \mathbf{G}x \cdot \mathbf{G}y. \tag{21}$$

Let $\{x_j\}_{j=1}^N$, $N = (l_{max}+1)^2$, be an enumeration of the Schmidt seminormalized spherical harmonics Y_l^m, $0 \leq l \leq l_{max}$, $|m| \leq l$ [*Backus*, 1986]. If the N by N matrix $\boldsymbol{\Gamma}$ with elements

$$\Gamma_{ij} \equiv < x_i|x_j >_G \tag{22}$$

is positive definite, then $< \cdot|\cdot >_G$ is an inner product and $\| \cdot \|_G$ is a norm. This can be verified numerically. Henceforth we assume $\| \cdot \|_G$ is a norm.

The space C spanned by $\{x_j\}_{j=1}^N$ with inner product $< \cdot|\cdot >_G$ and norm $\| \cdot \|_G$ is a finite-dimensional Hilbert space, and thus possesses a reproducing kernel $K_G(\hat{r}, \hat{s})$ [*Aronszajn*, 1950]. The properties of $K_G(\hat{r}, \hat{s})$ we use are:

$$< x(\hat{r})|K_G(\hat{r}, \hat{s}) >_G = x(\hat{s}) \tag{23}$$

$$|x(\hat{s})| \leq \|x\|_G \sqrt{K_G(\hat{s}, \hat{s})}, \tag{24}$$

and that $K_G(\hat{r}, \hat{s})$ has the expression

$$K_G(\hat{r}, \hat{s}) = \sum_{j,k=1}^{N} \lambda_{jk} x_j(\hat{r}) x_k(\hat{s}), \tag{25}$$

where $\lambda_{jk} = (\boldsymbol{\Gamma}^{-1})_{jk}$. Equation (23) is the defining property of a reproducing kernel: for fixed $\hat{s}$, $K(\hat{r}, \hat{s}) \in \mathsf{C}$, and the inner product of any element $x \in \mathsf{C}$ with $K(\hat{r}, \hat{s})$ evaluates x at the point $\hat{s}$. Inequality (24) follows from (23) by the Cauchy-Schwarz inequality. It says that we can bound x at a point $\hat{s}$ in terms of the norm of x and the reproducing kernel. Equation (25) shows that once we construct $\boldsymbol{\Gamma}$ we essentially have $K_G(\hat{r}, \hat{s})$ as well. See *Aronszajn* [1950] for proofs.

The uncertainty in $x(\hat{s}) =< K_G(\hat{r},\hat{s})|x(\hat{r}) >_G$ is

$$\begin{aligned}
\omega(2\chi) &= \sup\{|x(\hat{s}) - y(\hat{s})| : \|\mathbf{G}x - \mathbf{G}y\| \leq 2\chi,\ x, y \in \mathsf{C}\} && (26)\\
&= \sup\{|x(\hat{s})| : \|\mathbf{G}x\| \leq 2\chi,\ x \in \mathsf{C}\} && (27)\\
&= \sup\{|x(\hat{s})| : \|x\|_G \leq 2\chi,\ x \in \mathsf{C}\} && (28)\\
&= 2\chi\sqrt{K_G(\hat{s},\hat{s})}, && (29)
\end{aligned}$$

from (24) and the fact that C is a subspace. (This does not hold for general convex C.) Thus the reproducing kernel gives the uncertainty of the CMB. This development does not depend on the fact that the finite-dimensional subspace consists of spherical harmonics—the same procedure works for a pixel basis or any other finite-dimensional space of CMB models, with different entries in $\boldsymbol{\Gamma}$.

Subject to assumptions about the distribution of sources, we can show that the modulus converges to zero. Suppose sources are randomly distributed in a way that the resulting CMB hit points are independently distributed with a distribution $\mathcal{F}$. Then each entry in the Gram matrix $\boldsymbol{\Gamma}$ is the sum of n iid random variables with finite variance (the distributions will generally be different for different elements of the Gram matrix; the variances are finite since the spherical harmonics are bounded). Each sum is n times the average of the random variables. By the weak law of large numbers, each average converges (in probability) to a fixed mean value. Thus as $n \to \infty$, the Gram matrix converges to n times a constant matrix $\boldsymbol{\Gamma}_0$. Provided $\mathcal{F}$ satisfies mild conditions (having a density that is nonzero on an open subset of the CMB will do), $\boldsymbol{\Gamma}_0$ is invertable. The inverse of the finite n Gram matrix will then tend to $1/n$ times the inverse of $\boldsymbol{\Gamma}_0$; i.e. each element will converge to zero at rate n^{-1}. The modulus at any point is the square-root of a linear combination of elements of the Gram

matrix (substitute equation (25) into equation (29)), so it converges to zero at rate $n^{-1/2}$.

Extremal Models.

The reproducing kernel also gives the model in C with the largest amplitude at the point $\hat{s}$, subject to perturbing the data by at most χ (uncertainty measure U1): suppose we seek

$$x_{\hat{s}}^*(\hat{r}) = \arg\max_{x \in \mathsf{C}} \{x(\hat{s}) : \|\mathbf{G}x\| \leq \chi\}. \tag{30}$$

Let $x = \sum_j \beta_j x_j$, where $\{x_j\}_{j=1}^N$ is a basis for C; let $\boldsymbol{\beta}$ be the N-vector with components β_j; and let $\boldsymbol{\gamma}$ be the N-vector with components

$$\gamma_j = x_j(\hat{s}). \tag{31}$$

We need

$$\begin{aligned} \max\{x(\hat{s}) : \|\mathbf{G}x\| \leq \chi\} &= \max\{\boldsymbol{\gamma} \cdot \boldsymbol{\beta} : \boldsymbol{\beta} \cdot \boldsymbol{\Gamma} \cdot \boldsymbol{\beta} \leq \chi^2\} & (32) \\ &= -\max_{\lambda \geq 0} \min_{\boldsymbol{\beta}} \left\{-\boldsymbol{\gamma} \cdot \boldsymbol{\beta} + \lambda(\boldsymbol{\beta} \cdot \boldsymbol{\Gamma} \cdot \boldsymbol{\beta} - \chi^2)\right\}. & (33) \end{aligned}$$

The last equality follows from Lagrangian duality [*Luenberger*, 1969] if the modulus is finite. Let $\psi(\lambda, \boldsymbol{\beta})$ denote the term in brackets in the last equation. The functional ψ is differentiable with respect to both $\boldsymbol{\beta}$ and λ. Stationarity in $\boldsymbol{\beta}$ yields

$$\boldsymbol{\beta} = \frac{1}{2\lambda} \boldsymbol{\gamma} \cdot \boldsymbol{\Gamma}^{-1}. \tag{34}$$

Substituting (34) into (33), we find

$$\max\{x(\hat{s}) : \|\mathbf{G}x\| \leq \chi\} = -\max_{\lambda \geq 0} \left\{-\frac{1}{4\lambda} \boldsymbol{\gamma} \cdot \boldsymbol{\Gamma}^{-1} \cdot \boldsymbol{\gamma} - \lambda \chi^2\right\}. \tag{35}$$

Differentiating with respect to λ gives the stationary point

$$\lambda = \frac{1}{2\chi} \sqrt{\boldsymbol{\gamma} \cdot \boldsymbol{\Gamma}^{-1} \cdot \boldsymbol{\gamma}}\,. \tag{36}$$

Now $\gamma \cdot \Gamma^{-1} \cdot \gamma = K_G(\hat{s}, \hat{s})$, so

$$\beta = \frac{\chi \gamma \cdot \Gamma^{-1}}{\sqrt{K_G(\hat{s}, \hat{s})}}, \tag{37}$$

$$x_{\hat{s}}^*(\hat{r}) = \frac{\chi K_G(\hat{r}, \hat{s})}{\sqrt{K_G(\hat{s}, \hat{s})}}. \tag{38}$$

Thus the model $x_{\hat{s}}^*(\hat{r})$ maximizing the perturbation at the point $\hat{s}$ is proportional to the reproducing kernel. Setting $\hat{s} = \hat{r}$, we find

$$x_{\hat{s}}^*(\hat{s}) = \chi \sqrt{K_G(\hat{s}, \hat{s})}, \tag{39}$$

half the width of the bounds. Since $-x_{\hat{s}}^*$ also lies in C and produces data with norm less than χ, the total range of values of $x(\hat{s})$ that perturb the data by at most χ and lie in C is $2\chi\sqrt{K_G(\hat{s}, \hat{s})} = \omega(2\chi)$, as claimed.

Extensions: Quantitative Smoothness and Infinite-Dimensions.

There are a number of ways to incorporate a quantitative bound on the roughness of the CMB. Suppose

$$|x_0|_S \equiv \sqrt{\int_{\|\hat{r}\|=1} |\nabla_1^2 x_0|^2} \leq M, \tag{40}$$

where ∇_1^2 is the dimensionless surface Laplacian [*Backus*, 1986]. We present three ways to use the constraint (40). Two lead to optimistic uncertainty estimates; the third (conservative) approach may be useful in other problems.

Let

$$\mathcal{N} \equiv \text{span}\{x_j\}_{j=1}^N. \tag{41}$$

Redefine C to be the convex set

$$\mathsf{C} \equiv \{x(\hat{r}) : |x|_S \leq M\}. \tag{42}$$

We seek the modulus (14) with this new definition of C:

$$\omega(\chi) = \sup\{|x(\hat{s}) - y(\hat{s})| : \|\mathbf{G}x - \mathbf{G}y\| \leq \chi,\ x, y \in \mathsf{C}\} \tag{43}$$

$$= \sup\{x(\hat{s}) : \|\mathbf{G}x\| \leq \chi, |x|_S \leq 2M\}, \tag{44}$$

by the symmetry of C and the set $\{x : \|\mathbf{G}x\| \leq \chi\}$.

Upper Bounds. On the orthogonal complement $\mathcal{N}^\perp$ of $\mathcal{N}$, $|\cdot|_S$ is a norm (since the null-space of ∇_1^2 is in the span of the data kernels G_j) strong enough that the resulting space has a reproducing kernel $K_{SN}(\hat{r}, \hat{s})$. Since $\mathcal{N}$ is finite-dimensional and therefore closed, X can be decomposed into the direct sum

$$\mathsf{X} = \mathcal{N} + \mathcal{N}^\perp. \tag{45}$$

I.e., any $x \in \mathsf{X}$ can be written uniquely

$$x = x_{\mathcal{N}} + x_{\mathcal{N}^\perp}, \tag{46}$$

with $x_{\mathcal{N}} \in \mathcal{N}$ and $x_{\mathcal{N}^\perp} \in \mathcal{N}^\perp$. On X, define the inner product

$$< x|y >_{GSN} \equiv < x_{\mathcal{N}}|y_{\mathcal{N}} >_G + \int_{\|\hat{r}\|=1} \nabla_1^2 x_{\mathcal{N}^\perp} \nabla_1^2 y_{\mathcal{N}^\perp}. \tag{47}$$

If two Hilbert spaces $\mathcal{Y}$ and $\mathcal{Z}$ have reproducing kernels K_Y and K_Z, and $\mathcal{W} = \mathcal{Y} \oplus \mathcal{Z}$, then $\mathcal{W}$ has the reproducing kernel $K_W = K_Y + K_Z$ [*Aronszajn*, 1950]. Thus the reproducing kernel for X with inner product $< x|y >_{GSN}$ is

$$K_{GSN}(\hat{r}, \hat{s}) = K_G(\hat{r}, \hat{s}) + K_{SN}(\hat{r}, \hat{s}), \tag{48}$$

and so

$$\omega(\chi) = \sup\{x(\hat{s}) : \|\mathbf{G}x\| \leq \chi \text{ and } |x|_S \leq 2M\} \tag{49}$$

$$\leq \sup\{x(\hat{s}) : \|\mathbf{G}x\| \leq \chi \text{ and } x \in \mathcal{N}\} +$$

$$+\sup\{x(\hat{s}) : |x|_S \leq 2M \text{ and } x \in \mathcal{N}^{\perp}\} \tag{50}$$

$$= \chi\sqrt{K_G(\hat{s},\hat{s})} + 2M\sqrt{K_{SN}(\hat{s},\hat{s})}. \tag{51}$$

The reproducing kernel $K_{SN}(\hat{r},\hat{s})$ can be bounded numerically by estimating series involving Legendre polynomials to get upper bounds on $\omega(\chi)$.

Lower Bounds. We present two ways to get lower bounds on the modulus $\omega(\chi)$. One uses a reproducing kernel, the other simply solves the optimization problem for ω.

First Technique. On $\mathcal{N}$ define the bilinear functional

$$< x|y >_S \equiv \int_{\|\hat{r}\|=1} \nabla_1^2 x \nabla_1^2 y, \tag{52}$$

and let $\boldsymbol{\Lambda}$ be the $N \times N$ matrix with elements

$$\Lambda_{jk} =< x_j|x_k >_S . \tag{53}$$

The spherical harmonics $x_j = Y_l^m$ are eigenfunctions of the surface Laplacian ∇_1^2 with eigenvalues $\lambda_{lm} = \lambda_l = -l(l+1)$ and are orthogonal with norm $(2l+1)^{-1}$ [*Backus*, 1986]. Therefore $\boldsymbol{\Lambda}$ is diagonal with entries $\Lambda_{jj} = l^2(l+1)^2(2l+1)^{-1}$ for the appropriate value of l. Recall the definition of $\boldsymbol{\gamma}$ (31).

$$\omega(\chi) = \sup\{x(\hat{s}) : \|\mathbf{G}x\| \leq \chi, |x|_S \leq 2M\} \tag{54}$$

$$\geq \sup\{x(\hat{s}) : \|\mathbf{G}x\| \leq \chi, |x|_S \leq 2M,\ x \in \mathcal{N}\} \tag{55}$$

$$= \max_{\boldsymbol{\beta}\in\mathbf{R}^N}\{\boldsymbol{\gamma}\cdot\boldsymbol{\beta} : \boldsymbol{\beta}\cdot\boldsymbol{\Gamma}\cdot\boldsymbol{\beta} \leq \chi^2,\ \boldsymbol{\beta}\cdot\boldsymbol{\Lambda}\cdot\boldsymbol{\beta} \leq 4M^2\} \tag{56}$$

$$= -\max_{\lambda_1,\lambda_2\geq 0}\min_{\boldsymbol{\beta}\in\mathbf{R}^n}\{\boldsymbol{\gamma}\cdot\boldsymbol{\beta} + \lambda_1(\boldsymbol{\beta}\cdot\boldsymbol{\Gamma}\cdot\boldsymbol{\beta} - \chi^2) +$$

$$+\lambda_2(\boldsymbol{\beta}\cdot\boldsymbol{\Lambda}\cdot\boldsymbol{\beta} - 4M^2)\}, \tag{57}$$

which can be found numerically. The solution can be written *const* $\times \sqrt{K_\lambda(\hat{s},\hat{s})}$, where $K_\lambda(\hat{s},\hat{s})$ is the reproducing kernel

$$K_\lambda(\hat{r},\hat{s}) = \sum_{jk} \left((\boldsymbol{\Gamma} + \lambda \boldsymbol{\Lambda})^{-1}\right)_{jk} x_j(\hat{r}) \bar{x}_k(\hat{s}). \tag{58}$$

Second Technique. Alternatively, consider the inner product on $\mathcal{N}$

$$< x|y >_{GS} \equiv < x|y >_G + < x|y >_S, \tag{59}$$

and the induced norm

$$\|x\|_{GS} \equiv \sqrt{< x|x >_{GS}}. \tag{60}$$

There is a reproducing kernel K_{GS} corresponding to this norm; the relevant matrix is $\boldsymbol{\Gamma} + \boldsymbol{\Lambda}$. The constraints $\|\mathbf{G}x\| \leq \chi$ and $|x|_S \leq 2M$ are less restrictive than the constraints $x \in \mathcal{N}$ and $\|x\|_{GS} \leq \min\{\chi, 2M\}$, so

$$\omega(\chi) \geq \sup\{x(\hat{s}) : x \in \mathcal{N},\ \|x\|_{GS} \leq \min\{\chi, 2M\}\} \tag{61}$$

$$= \min\{\chi, 2M\}\sqrt{K_{GS}}. \tag{62}$$

Thus we can accomodate a quantitative bound $|x|_S \leq M$ and still get optimistic uncertainty estimates for $x_0(\hat{s})$. We have not pursued this numerically since we are unaware of a persuasive argument for a particular value of M, or even the existence of a finite M.

Results.

Data and Computations.

R.J. Pulliam supplied us with processed ISC data, consisting of arrival times from over 46,000 events between 1964-1987. The events were relocated in the IASP91 model [*Kennett and Engdahl*, 1991] using all direct P arrival times to constrain the locations. Only events

with at least 50 reported arrivals were retained. In all 84,947 PcP times were identified in the epicentral distance range 30° to 38° and 47° to 75°. There were 33,602 PKPab arrivals in the range 146° to 180°; 308,846 PKPbc arrivals in the range 146° to 154°; and 434,510 PKPdf arrivals in the range 114° to 180°. In each case, the search for arrivals was in a $\pm 30s$ window around the expected arrival. These rays were traced in the IASP91 model to find their crossing or bounce points on the CMB. The total number of rays was 861,905 and the total number of sample points on the CMB was 1,638,863.

Figures 1a(i)-1c(i) plot the relative density of hits of rays on the CMB in cells one degree on a side, normalized by area and scaled to a maximum of 100. The plots are truncated at 10 to show some structure. Figures 1a(ii)-1c(ii) plot 1° by 1° cells on the surface of the CMB that are hit at least once. All maps are in the Aitoff equal-area projection, with longitude zero at the left boundary and latitude zero in the middle. All plots (except figure 5) were produced by programs written by R.L. Parker.

For this set of rays we computed the matrix $\boldsymbol{\Gamma}$ of inner products $< x_j | x_k >_G$ by evaluating, for each source-receiver pair, the linearized change to the traveltimes induced by spherical harmonic perturbations x_j and x_k, and summing over source-receiver pairs. In the linearization the perturbation does not change the parameter of the ray joining the source and receiver. We computed the change in path lengths through material just above and below the CMB and divided by the corresponding velocities to find the traveltime perturbations. The spherical harmonics were evaluated at the CMB hit points using code written by Robert L. Parker. This was done once, for the maximum dimension $l_{max} = 20$ ($N = 21^2$). The dimension of $\boldsymbol{\Gamma}$ is $N \times N$, regardless of the number n of data. We used NAG routines [*Numerical Algorithms Group, Inc.* Mark 12] to invert submatrices of $\boldsymbol{\Gamma}$ corresponding to

$l_{max} \in \{4, 8, 12, 16, 20\}$. We computed the "diagonal" of the reproducing kernel, $K_G(\hat{s}, \hat{s})$, using equation (25) and evaluated it numerically on a grid of 360 longitude by 180 latitude points using R.L. Parker's spherical harmonic routines. The unnormalized uncertainties $\omega(1) = \sqrt{K_G(\hat{s}, \hat{s})}$ for the five values of l_{max} are plotted in figure 2 for the combined PcP-PKP data set. Figure 3 shows degree 4 and 20 models that perturb the data by only $1sec$ in the norm and maximize the structure southwest of Australia and beneath the mid-Pacific, respectively; see the section on extremal models above. We show below that the degree 20 model could penetrate $92.7km$ into the mantle but perturb the data by just $0.1sec$ rms.

Interpretation.

Figures 2 and 3 must be multiplied by 2χ to be meaningful. The choice of χ affects the interpretation of the figures. We present a deterministic interpretation and several probabilistic interpretations.

Deterministic and Crude Statistical Interpretations. The first is the crudest. Suppose we can find a finite value of χ such that

$$Prob\{\|\epsilon\| \leq \chi\} \geq 1 - \alpha, \tag{63}$$

for some known $\alpha \in [0, 1)$. For $\alpha = 0$, we get deterministic uncertainty bounds for $x(\hat{s})$: figure 2 times 2χ gives the widths of bounds guaranteed to contain $x(\hat{r})$ for all $\|\hat{r}\| = 1$. For $\alpha \neq 0$, figure 2 times 2χ bounds the widths of simultaneous $1 - \alpha$ confidence envelopes for $x(\hat{r})$ for all $\|\hat{r}\| = 1$: the length of a "strict bounds" $1 - \alpha$ confidence envelope [*Stark*, 1992] for $x(\hat{r})$ with the worst possible data is given by the figures times 2χ. These interpretations use minimal assumptions about ϵ, but yield extremely wide bounds.

Consider the deterministic choice of χ. If the rms error in traveltimes is less than $0.25sec$

(it is closer to $1.0sec$ for most epicentral distances), then $\chi = 0.25\sqrt{n}$, and we must normalize figure 2 by $0.5\sqrt{n}$ to get uncertainty bounds (see table 1). The minimum uncertainty using all the PcP and PKP data is $0.5 \times 928.4 \times 0.017 = 7.89km$ ($= 0.25 \times$ entry in table $1 \times$ entry in table 6), even if the CMB has a degree 4 spherical harmonic expansion. The maximum uncertainty for a degree 20 core would be $0.5 \times 928.4 \times 0.998 = 463.3km$.

For the probabilistic choice, if the errors ϵ_j are iid Gaussian random variables with mean zero and variance σ^2, the value of χ needed to guarantee $1 - \alpha$ coverage probability is

$$\chi = \sigma\sqrt{\chi^2_{\alpha,n}}, \tag{64}$$

where $\chi^2_{\alpha,n}$ is the $100(1 - \alpha)$ percentile of the chi-square distribution with n degrees of freedom. For large n (here n is nearly a million),

$$\chi^2_{\alpha,n} \approx n + \sqrt{2n}\mathcal{Z}_\alpha, \tag{65}$$

where $\mathcal{Z}_\alpha$ is the $100(1 - \alpha)$ percentile of the standard Normal distribution. The value of χ grows very quickly with n for this approach. With stronger assumptions, χ can be chosen much smaller and still yield $1 - \alpha$ coverage probability.

Least-Squares Regression. *Morelli and Dziewonski* [1987] use least-squares to estimate a truncated spherical harmonic expansion of the CMB. In this case, C is an N-dimensional subspace of X, and linear regression theory provides both simultaneous confidence envelopes (measure U6) and individual confidence intervals (measure U4), assuming the errors ϵ_j are iid with mean zero and variance $\sigma^2 < \infty$. The resulting choice of χ is much smaller than in the preceeding section. It grows with the number of model coefficients, not the number of data.

Recall the linearized problem (8) $\boldsymbol{\delta} = \mathbf{G}x_0 + \boldsymbol{\epsilon}$ with $\mathsf{C} = \text{span}\{x_j\}_{j=1}^n$. Let $\boldsymbol{\beta}$ be the N-vector $(\beta_j)_{j=1}^N$. Define

$$\boldsymbol{\beta} \cdot \mathbf{x}(\hat{r}) \equiv \sum_{j=1}^{N} \beta_j x_j(\hat{r}),$$

and let $\boldsymbol{\beta}_0$ denote the coefficients of x_0. Then

$$\boldsymbol{\delta} = \mathbf{G}(\boldsymbol{\beta}_0 \cdot \mathbf{x}(\hat{r})) + \boldsymbol{\epsilon} \tag{66}$$

$$= (\mathbf{G}\mathbf{x}(\hat{r})) \cdot \boldsymbol{\beta}_0 + \boldsymbol{\epsilon} \tag{67}$$

$$= \mathbf{A} \cdot \boldsymbol{\beta}_0 + \boldsymbol{\epsilon}, \tag{68}$$

where $\mathbf{A} \equiv \mathbf{G}\mathbf{x}(\hat{r})$ is the $n \times N$ matrix with entries $A_{ij} = G_i x_j$.

The least-squares estimate of $\boldsymbol{\beta}_0$ is

$$\tilde{\boldsymbol{\beta}} = (A^T A)^{-1} A^T \boldsymbol{\delta} \tag{69}$$

$$= \boldsymbol{\Gamma}^{-1} A^T \boldsymbol{\delta}, \tag{70}$$

with $\boldsymbol{\Gamma}$ as in equation (22). When $\{\epsilon_j\}$ are iid and $n >> N$ (as here), the distribution of the coefficient error vector $\boldsymbol{\eta} \equiv \tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0$ is approximately joint normal, with covariance matrix

$$\Sigma \equiv \sigma^2 \boldsymbol{\Gamma}^{-1}; \tag{71}$$

i.e.

$$\boldsymbol{\eta} \equiv \tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0 \sim N(0, \sigma^2 \boldsymbol{\Gamma}^{-1}). \tag{72}$$

See, for example, *Hoel et al.* [1971], *Menke* [1989], or *Tarantola* [1987]. The corresponding least-squares estimate $\tilde{x}(\hat{r})$ of $x_0(\hat{r})$ is

$$\tilde{x}(\hat{r}) = \tilde{\boldsymbol{\beta}} \cdot \mathbf{x}(\hat{r}), \tag{73}$$

and the error in the estimate at $\hat{s}$ is

$$\tilde{x}(\hat{s}) - x_0(\hat{s}) = \boldsymbol{\eta} \cdot \boldsymbol{\gamma}, \tag{74}$$

where $\boldsymbol{\gamma}$ is given by (31). The error at $\hat{s}$ has zero mean and variance

$$\boldsymbol{\gamma} \cdot \sigma^2 \boldsymbol{\Gamma}^{-1} \cdot \boldsymbol{\gamma} = \sigma^2 K_G(\hat{s}, \hat{s}). \tag{75}$$

Thus a $1 - \alpha$ confidence interval for $x_0(\hat{s})$ has width

$$2\sigma \mathcal{Z}_{\alpha/2} \sqrt{K_G(\hat{s}, \hat{s})}. \tag{76}$$

This is measure U4; table 3 compares these values with optimal results for non-simultaneous inference (measure U5) computed below.

We can also use the regression formulation to find a simultaneous confidence envelope for all $\hat{r}$. For each $\hat{r}$, we have

$$\begin{aligned} |\tilde{x}(\hat{r}) - x_0(\hat{r})| &= |\boldsymbol{\eta} \cdot \mathbf{x}(\hat{r})| \\ &= |\boldsymbol{\eta} \cdot \boldsymbol{\Gamma}^{1/2} \cdot \boldsymbol{\Gamma}^{-1/2} \cdot \mathbf{x}(\hat{r})| \\ &\leq \|\boldsymbol{\eta} \cdot \boldsymbol{\Gamma}^{1/2}\| \, \|\boldsymbol{\Gamma}^{-1/2} \cdot \mathbf{x}(\hat{r})\|. \end{aligned} \tag{77}$$

The covariance matrix of $\boldsymbol{\eta}$ is $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{\Gamma}^{-1}$, so $\boldsymbol{\eta} \cdot \boldsymbol{\Gamma}^{1/2}$ has covariance $\sigma^2 I$, where I is the $N \times N$ identity matrix. Thus the first term in (77) is the square-root of σ^2 times a chi-square random variable with $N = \dim(\mathsf{C})$ degrees of freedom. The second term is $\sqrt{K_G(\hat{r}, \hat{r})}$; see (25). A simultaneous $1 - \alpha$ confidence envelope for $x_0(\hat{r})$ for all $\hat{r}$ (measure U6) therefore has width

$$2\sqrt{\chi^2_{\alpha,N}} \sqrt{K_G(\hat{r}, \hat{r})}. \tag{78}$$

This is considerably smaller than (64), where the number of degrees of freedom in the chi-square distribution was n (the number of data), which is much larger than $N = \dim \mathsf{C}$. These are the narrowest simultaneous confidence envelopes we have found. Table 2 gives the resulting normalizing constants for figure 2 as a function of the maximum degree of the model and the confidence level.

If the truncation level l_{max} is different from the maximum degree in the least-squares parametrization, the confidence intervals and envelopes must be lengthened (e.g. if CMB structure extends to degree 16 but we only solve for degree 4 and less, the confidence intervals and envelopes need to be wider than for degree 16). The regression formulation can also be combined with inequality (51) to get upper bounds on the width of a simultaneous $1 - \alpha$ confidence envelope using a quantitative smoothness constraint.

Non-Simultaneous Inference. The uncertainty measure that gives the smallest constant χ is U5, a confidence interval for a single point on the CMB. This applies when we are interested in the height of the CMB at a particular point, rather than the shape of the CMB. Inferences about variations in the CMB require simultaneous coverage probability. Nonetheless, it is interesting to know how well the CMB can be located at an individual point, and in some problems this is the primary question. Since U5 gives much shorter confidence intervals, we urge extreme caution in interpreting the results.

If the errors ϵ_j are iid $N(0, \sigma^2)$, the elegant theory of Donoho [*Donoho*, 1989; *Donoho and Low*, 1992; *Stark*, 1992a] shows that $\omega(2\sigma \mathcal{Z}_\alpha)$ is a lower bound on the length of a valid $1 - \alpha$ confidence interval for $x_0(\hat{s})$. The appropriate values of 2χ for $\sigma = 1$ are given in table 3. Values for $\sigma \neq 1$ are obtained by multiplying the tabulated values by σ. These intervals are shorter than the least-squares confidence intervals by the factor $\mathcal{Z}_\alpha / \mathcal{Z}_{\alpha/2}$.

Discussion.

The maximum uncertainties grow quickly with the maximum spherical harmonic degree l_{max}. Within ray theory the uncertainty is unbounded as $l_{max} \to \infty$, except where rays hit the CMB. For small l_{max}, the spatial correlation imposed by the parametrization causes the largest uncertainties to appear where there is some ray coverage, rather than exactly in the largest gaps. As l_{max} increases, the maximum uncertainties migrate to the biggest gaps, which are under the mid-Pacific. Under the western coast of Africa and to the west and southwest of Australia the uncertainties are large too. Although there are relatively few PcP data, and the uncertainty using the PcP data alone is huge, they reduce the maximum uncertainty substantially when used with PKP data. They do not change the average uncertainty much.

Distribution Function of the Uncertainty.

We can summarize the results by the fraction of the core that can be estimated with specified accuracy. Figure 4 plots the distribution of the unnormalized uncertainties $\omega(1)$ for the combined PKP and PcP data and the various maximum model degrees. The abscissa x of the plots is the unnormalized uncertainty; the ordinate y is the fraction of the unit sphere that has unnormalized uncertainty less than x:

$$y = \int_{\|\hat{r}\|=1} 1_{\omega(1)\leq x},$$

where $1_{\omega(1)\leq x}$ is the indicator function of the set where $\omega(1)$ is at most x. (Here x and y are real numbers, not models.) Tables 4-6 give the minimum, quartiles, average, and maximum unnormalized uncertainties, summarizing figure 4.

By normalizing Figure 4 with constants from tables 1-2, we can find the fraction of

the core's surface on which the topography can be estimated with specified accuracy. For example, suppose we want a simultaneous 95% confidence envelope for the entire CMB, based on least-squares estimates of the topography from the combined PKP and PcP data sets. Assume the traveltime errors are iid with mean zero and standard deviation $0.25sec$. If CMB topography is of degree 4 and less, then 50% of the CMB can be estimated to better than $\pm 0.29km$, and the entire CMB to about $\pm 0.53km$. If CMB structure involves up to degree 8, the uncertainty over 50% of the CMB is more than $\pm 1.62km$, with a maximum of $\pm 4.20km$. For $l_{max} = 12$, the uncertainty is larger than $\pm 4.81km$ over half the CMB and larger than $\pm 7.56km$ over 25% of the core, with a maximum of $\pm 18.4km$. For $l_{max} = 16$, the median uncertainty is $\pm 10.7km$, the upper quartile is $\pm 17.5km$ and the maximum is $\pm 54.5km$. For $l_{max} = 20$, the median, upper quartile and maximum are $\pm 20.3km$, $\pm 33.6km$ and $\pm 122.5km$, respectively. See table 7 for a more comprehensive list for 99.9% confidence. Many seismologists believe the traveltime standard deviation is closer to $1.0sec$ than to $0.25sec$. If the standard deviation is $0.5sec$, the numbers above must be multiplied by 2.

Experimental Design.

The reproducing kernel approach provides a framework to optimize the location of new stations to best improve the reliability of CMB images. Experimental design requires a precise goal; we believe it is especially desirable to decrease the maximum uncertainty.

Given smoothness constraints, our results confirm quantitatively the intuition that the local uncertainty depends on the density of "hits" on the CMB, which suggests a strategy: Target a patch on the CMB where the uncertainty is largest and trace rays from all source locations that intersect that patch. Concentrations of surface hits of those rays are reasonable places for a new station. From this naive starting point, trace rays from all sources to the

receiver point to build a reduced Gram matrix for just that receiver. Gram matrices for different source-receiver pairs add, so the reduced Gram matrix can be added to the fixed Gram matrix for all remaining source-receiver pairs to find the Gram matrix for the combined actual/hypothetical data set. Invert the Gram matrix to find the new maximum uncertainty, perturb the location of the new receiver, and iterate. The additivity of Gram matrices for different data sets allows most of the computations to be done only once.

Implications for Previous CMB Images.

There have been several studies of the shape of the CMB from traveltime data over the last six years, for example *Creager and Jordan* [1986], *Doornbos and Hilton* [1989], *Gudmundsson* [1989], and *Morelli and Dziewonski* [1987]. Data processing before inversion may include:

1. relocating events
2. correcting for elevation and ellipticity
3. truncating outliers
4. applying station corrections (possibly with azimuthal dependence)
5. applying event corrections
6. forming summary rays
7. correcting for the crust and mantle
8. correcting for core anisotropy

The resulting summary traveltime anomalies are inverted by least-squares or damped least-squares for a finite-dimensional CMB model. Each step has uncertainties, and it has not been

established that together they succeed in reducing "signal" not due to the CMB. In fact, since the effect of perturbing these parameters is not orthogonal to the effect of perturbing the CMB, small changes in the corrections can induce large changes in inferred CMB topography.

The first CMB studies from ISC traveltimes, the standards to which others compare themselves, are those of *Creager and Jordan* [1986] (CJ), who use a pixel basis and stochastic inversion, and *Morelli and Dziewonski* [1987] (MD) who use a degree 4 spherical harmonic expansion and least-squares. CJ, MD and the studies that have followed constrain the CMB to be smooth through low-degree spherical harmonic parametrizations, restricted pixel sizes, regularization, or a combination of these. CJ do not solve for topography; instead they map traveltime anomalies to the CMB. They propose the existence of "continents on the core," finding a chemical boundary layer more plausible than the $\pm 10km$ topography needed to explain the traveltime anomalies.

MD argue from the correlation of their results using PcP and PKP traveltimes separately that the CMB topography of about $\pm 5km$ they find is likely to be real. They quote a significance level of less than 0.001 for the correlation, but do not state the null hypothesis or the test used. We find correlation arguments unpersuasive. (See, e.g. [*Pulliam and Stark*, 1992].) The high correlation between PcP and PKP models is misleading because MD's models are restricted to degree 4 spherical harmonics. MD structure is large where the uncertainty in figure 2 is large. The correlation coefficient is high due to the similarity of the largest structure in MD PcP and PKP models, which occurs where there are no data.

This often happens in fitting smooth global functions to noisy data with spatial gaps using least-squares. Figure 5 shows two sets of 10 data on the discrete circle with 20 points, along with least-squares fits of 5 low frequency sinusoids to each data set. The correlation

of the data is 0.064. The correlation of the models is 0.887. Both models have their largest structure where there are no data. This large, unconstrained structure is similar in the two models, increasing the correlation. Fitting low-frequency sinusoids also smooths the data in the range where there are observations, which increases the correlation: the correlation of the models restricted to the interval where there are data is still 0.871, an order of magnitude larger than the correlation of the original data. We believe this explains the large correlation observed by MD.

Later studies are more skeptical. *Doornbos and Hilton* [1989], who use both spherical harmonic and pixel parametrizations, tabulate the correlation of their models with CJ and MD. *Gudmundsson* [1989], who uses a pixels, also compares his results with CJ and MD. *Doornbos and Hilton* [1989] find structure of about $\pm 4km$, and *Gudmundsson* [1989] finds structure of about $\pm 2km$, with and without corrections for inner core anisotropy. Both confirm that different parametrizations or regularization schemes yield different results, and both argue for boundary layer structure in addition to CMB topography.

It is difficult to assess the significance of the (generally low) correlations among the published CMB models for several reasons, including lack of information about how spherical harmonic coefficients were computed from pixel models and how pixels with no hits were treated in the transformation to a spherical harmonic representation. *Eckhardt* [1984] studies the statistical significance of correlation between spherical harmonic models.

Table 7 gives the spatial distribution function of the width of a simultaneous 99.9% confidence envelope for the CMB for least-squares estimates with various maximum degrees using the combined PKP and PcP data sets, assuming the errors in traveltimes are iid zero-mean random variables with standard deviation $\sigma = 0.25s$, that source locations are

perfect, and that linearization error can be ignored. Under these optimistic assumptions, the confidence envelope is wider than $2.90km$ over 75% of the core, $3.97km$ over half the core, and $5.93km$ over 25% of the core, even if core structure is limited a priori to degree 8 or less. If the CMB has structure up to degree 20, those values increase to $28.3km$, $44.4km$ and $73.8km$, respectively. The uncertainty in the regions where MD infer the largest structure ($\pm 5km$) is about $10.3km$ for maximum degree 8 and $269.km$ for maximum degree 20.

The largest structure in CJ, *Doornbos and Hilton* [1989], *Gudmundsson* [1989], and MD all correlate visually with regions of high uncertainty. We conclude that within ray theory the CMB structure inferred previously in the literature is below the noise level: large variations in the location of the CMB may be consistent with ISC PcP and PKP traveltimes, but the data do not require such variations. A more powerful hypothesis test (with lower probability of a type II error for the alternative hypothesis of asymmetry) could be constructed by looking at just two well-sampled points on the CMB and using confidence intervals for just those points to see if the data require different radii. The nonsimultaneous confidence intervals described above can be used for this test; the significance level of the resulting test is less than 2α.

Jault and LeMouël [1990] show that topography of $\pm 5km$ or so would impose severe constraints on fluid flows at the CMB to avoid torques grossly in excess of those manifest in length-of-day fluctuations. *Forte and Peltier* [1991] used the MD model to test a number of convection models through their induced CMB topography, arguing that the seismically inferred topography is strong evidence for whole-mantle convection. Given the large uncertainties in the seismically inferred CMB topography, we believe it is premature to it as a constraint.

Is There A Way Out?

Our conclusion is that the package—traveltimes, ray theory, noise of $0.25s$, CMB structure of degree 20 or less—leads to wide simultaneous confidence envelopes for the CMB. An hypothesis test based on least-squares estimates of the whole CMB would fail to reject the null-model that the CMB is spherical, since the confidence envelope is larger than the structure. Traveltime data may contain useful information about the CMB, and published models may be basically correct, but the uncertainty under the current set of assumptions is large. There are a number of directions in which progress is possible.

To estimate the location of the CMB point-by-point requires assuming the CMB is smooth, which seems unlikely and has not been established (perhaps mineral physics could provide a substantiating argument). As a result, we advocate abandoning the attempt to locate the CMB at every point, and looking instead at functionals that average the CMB location. Such functionals are less sensitive to smoothness assumptions, and are better constrained by the data. Examples include the average height of the CMB on a patch, the average squared height of the CMB, the "inner product" of the CMB with a particular spherical harmonic, etc. We must keep in mind that even if we know the $l = 2$ component with very high accuracy, that does not mean the CMB looks like an $l = 2$ spherical harmonic: for example, the cosine transform (at any fixed frequency) of a delta function is 1, but a delta function does not look like a cosine.

Another way to circumvent the pessimistic conclusions is to approximate the physics better. As mentioned above, ray theory requires assumptions about the smoothness of the CMB. Doing the physics more precisely shows that traveltimes average CMB structure in a well-defined way. The fact that the traveltimes are smoothing functionals of CMB position

could enable us to recover averages of CMB structure without assuming the CMB is smooth.

Stark and Nikolayev [1992] study the sense in which traveltimes average Earth structure in a simple model. *Stark* [1992b] provides theory that can be adapted to make conservative inferences about other functionals of the CMB using those averages. Procedures for constructing minimax confidence intervals for linear functionals are given by *Donoho* [1989] (*Stark* [1992a] presents a geomagnetic application). Although the authors find it implausible, it is possible that underparametrized least-squares estimates of the CMB using ray theory give reliable estimates of averages of CMB position. *Pulliam and Stark* [1992] give evidence to the contrary, that underparametrized least-squares inversions are likely to confound source-receiver geometry with CMB structure.

It has been suggested (Mark Richards, personal communication, 1992; Thorne Lay, personal communication, 1992) that inverting for structure only where the ray coverage is good using a local basis rather than spherical harmonics might yield lower uncertainties. For comparable amounts of smoothness, the uncertainties are the same regardless of the basis used. That is, if the CMB is divided into pixels comparable in size to the oscillations of a degree 20 spherical harmonic, the modulus in the pixel basis will be comparable to the modulus for the spherical harmonic basis. In our analysis the uncertainties depend on the spherical harmonic basis primarily through the smoothness it affords, and through the number of parameters in the model. The uncertainties are huge where there are no data, and Table 7 shows they are also quite large ($> 5.5km$) even where the core is densely sampled, if the CMB has degree 16 structure. The uncertainty everywhere depends critically on smoothness assumptions when we use ray theory; using a different basis will not alleviate the problem. However, it takes fewer parameters to model part of the CMB at a given level of smoothness

than it does to model the entire CMB. The number of parameters enters the least-squares confidence envelopes through $\chi^2_{\alpha,N}$, which is smaller (since N is smaller) if we parametrize only part of the CMB using a local basis.

Conclusions.

By means of reproducing kernels one may compute many measures of spatial uncertainty in linear inverse problems with smoothness constraints. In certain linear problems, several intuitive measures of uncertainty are equivalent up to constant scaling factors: the widths of confidence intervals and simultaneous confidence envelopes around least-squares estimates, the width of a confidence interval around the best possible (minimax) estimate, the size of a perturbation that produces a negligable change in the data, and the largest difference between models that produce nearly the same data. Within ray theory, the uncertainty in the shape of the CMB inferred by least-squares collocation of low-degree spherical harmonic models depends critically on CMB smoothness. If the CMB has structure up to degree 16, the uncertainty measured by the width of a 95% simultaneous confidence envelope around the least-squares estimate exceeds the CMB structure inferred in previous studies. To reduce the uncertainty significantly will require abandoning the attempt to locate the CMB point-by-point, and using a better approximation than ray theory.

Acknowledgements. We are grateful to D.L. Donoho for suggesting the reproducing kernel approach, to P.S. Earle for code for ISC event extraction, to L.R. Johnson for ray-tracing code, to R.L. Parker for spherical harmonic subroutines and contouring and plotting packages, to R.J. Pulliam for relocated ISC events and rays traced through the IASP91 model, and to T. Lay and an anonymous referee for helpful comments on an earlier draft. This work was supported by the National Science Foundation through PYI award DMS-8957573 and

grant DMS-8810192.

References

[1] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337–404, 1950.

[2] G.E. Backus. Poloidal and toroidal fields in geomagnetic field modeling. *Rev. Geophys.*, 24:75–109, 1986.

[3] Bulletin of the International Seismological Centre: Catalogue of Events and Associated Observations (years 1964-1987). International Seismological Centre, Newbury UK, 1967-1989.

[4] R.W. Clayton and R.P. Comer. A tomographic analysis of mantle heterogeneities from body wave travel time data. *EOS Trans. Am. geophys. Un.*, 64:776, 1983.

[5] K.C. Creager and T.H. Jordan. Aspherical structure of the core-mantle boundary from PKP travel times. *Geophys. Res. Lett.*, 13:1497–1500, 1986.

[6] D.L. Donoho. Statistical estimation and optimal recovery. Technical Report 214, Department of Statistics, UC Berkeley, 1989.

[7] D.L. Donoho and M. Low. Renormalization exponents and optimal pointwise rates of convergence. *Ann. Stat.*, to appear, 1992.

[8] D.J. Doornbos and T. Hilton. Models of the core-mantle boundary and the travel times of internally reflected core phases. *J. Geophys. Res.*, 94:15,741–15,751, 1989.

[9] A.M. Dziewonski. Mapping the lower mantle: Determination of lateral heterogeneity in P velocity up to degree and order 6. *J. Geophys. Res.*, 89:5929–5952, 1984.

[10] A.M. Dziewonski, B.H. Hager, and R.J. O'Connell. Large-scale heterogeneities in the lower mantle. *J. Geophys. Res.*, 82:239–255, 1977.

[11] A.M. Dziewonski and J.H. Woodhouse. Global images of the earth's interior. *Science*, 236:37–48, 1987.

[12] D.H. Eckhardt. Correlations between global features of terrestrial fields. *Math. Geol.*, 16:155–171, 1984.

[13] A.M. Forte and W.R. Peltier. Mantle convection and core-mantle boundary topography: explanations and implications. *Tectonophysics*, 187:91–116, 1991.

[14] O. Gudmundsson. *Some Problems in Global Tomography: Modeling the core-mantle boundary and statistical analysis of travel-time data.* PhD thesis, California Inst. of Technology, 1989.

[15] P.G. Hoel, S.C. Port, and C.J. Stone. *Introduction to Statistical Theory.* Houghton Mifflin Co., 1971.

[16] H. Inoue, Y. Fukao, K. Tanabe, and Y. Ogata. Whole mantle P-wave travel time tomography. *Phys. Earth Planet. Inter.*, 59:294–328, 1990.

[17] D. Jault and J.-L. Le Mouël. Core-mantle boundary shape: constraints inferred from the pressure torque acting between the core and the mantle. *Geophys. J. Intl.*, 101:233–241, 1991.

[18] T.H. Jordan and K.C. Creager. Chemical boundary layers and large-scale flow in the mantle and core. *EOS Trans. Am. geophys. Un.*, 67:311, 1986.

[19] B.L.N. Kennett and E.R. Engdahl. Traveltimes for global earthquake location and phase identification. *Geophys. J. Intl.*, 105:429–465, 1991.

[20] D.G. Luenberger. *Optimization by Vector Space Methods.* John Wiley and Sons, Inc., New York, 1969.

[21] William Menke. *Geophysical Data Analysis: Discrete Inverse Theory.* Academic Press, San Diego, 1989.

[22] A. Morelli and A.M. Dziewonski. Topography of the core-mantle boundary and lateral homogeneity of the liquid core. *Nature*, 325:678–683, 1987.

[23] The NAG FORTRAN Library, Mark 12. Downer's Grove IL.

[24] R.J. Pulliam and P.B. Stark. Bumps on the core-mantle boundary: Are they facts or artifacts? Technical Report 342, Dept. Stat. Univ. California, Berkeley, 1992.

[25] R.J. Pulliam, D. Vasco, and L.R. Johnson. Tomographic inversions for mantle P-wave velocity structure based on the minimization of l^2 and l^1 norms of ISC travel time residuals. *J. Geophys. Res.*, in press, 1992.

[26] M.K. Sengupta and M.N. Toksöz. Three dimensional model of seismic velocity variation in the earth's mantle. *Geophys. Res. Lett.*, 3:84–86, 1976.

[27] P.B. Stark. Inference in infinite-dimensional inverse problems: Discretization and duality. *J. Geophys. Res.*, to appear, 1992.

[28] P.B. Stark. Minimax confidence intervals in geomagnetism. *Geophys. J. Intl.*, 108:329–338, 1992.

[29] P.B. Stark and D.I. Nikolayev. Toward tubular tomography. Technical Report 353, Dept. Stat. Univ. California, Berkeley, 1992.

[30] Albert Tarantola. *Inverse Problem Theory: methods for data fitting and model parameter estimation.* Elsevier Science Publishing Co., 1987.

Figure Captions.

Fig. 1: The distribution of PcP and PKP hits on the CMB that comprise our data set. Figure 1a(i) shows a shaded plot of PcP hits per unit area on the CMB. The data are normalized so that the maximum number of hits per unit area is 100. The hit density is extremely inhomogeneous; to show some structure we truncated the scale at 10 hits per unit area. Figure 1a(ii) shows 1° by 1° cells hit by at least 1 PcP ray. Figure 1b(i) is the same as 1a(i), but for PKP data. Figure 1b(ii) is as 1a(ii), but for PKP data. Figures 1c(i) and 1c(ii) are the same as 1b(i) and 1b(ii) but for PcP and PKP data combined. Note the gap in data coverage in the Pacific.

Fig. 2: Unnormalized uncertainty $\omega(1)$ in the location of the CMB as a function of angular position, for various spherical harmonic truncation levels (l_{max}) for the combined PcP and PKP data set. The function $\omega(1)$ is is defined by an optimization problem solved by the reproducing kernel in equation (29). The values of ω must be multiplied by a scaling factor 2χ to be interpreted as uncertainties in the location of the CMB. The constant 2χ depends on the noise level, and on the interpretation one wishes to make of the figures.

Fig. 3: Extremal models. Figure 3a shows a CMB model with maximum spherical harmonic

degree 4 that produces a perturbation to the combined PcP and PKP data sets of $1sec$ in the two-norm, and maximizes the height of the CMB southwest of Australia. 3b shows a CMB model of degree 20 that produces PcP and PKP data with norm $1sec$ and maximizes the height of the CMB beneath the mid-Pacific.

Fig. 4: Distribution function of the unnormalized uncertainty; i.e. the fraction of the CMB for which $\omega(1) \leq x$ (see text for the definition and meaning of $\omega(\cdot)$). Panels a, b, and c respectively give the uncertainties using PcP only, PKP only, and PcP and PKP data together.

Fig. 5: Illustration of illusory correlation of least-squares fits of smooth periodic functions to noisy, gappy data. There are two sets of 10 data (stars and plusses) given at the first 10 points of a discrete circle with 20 points. The correlation of the two sets of data is 0.064. The curves are fits of 5 low frequency sinusoids to each of the sets of data. The largest structure in the fitted models occurs where there are no data. The correlation of the fitted models is 0.887. Even restricted to the range where there are data, the correlation of the fitted models is 0.871. As a general rule, smoothing data, for example by fitting smooth functions, increases apparent correlation. Often when smooth functions are fit to rough data with gaps, the largest structure in the model occurs where there are no data.

Data Type	No.Rays	CMB points	Constant
PcP	84,947	84,947	582.9
PKP_{ab}	33,602	67,204	366.6
PKP_{bc}	308,846	617,692	1,111.5
PKP_{df}	434,510	869,020	1,318.3
PKP	776,958	1,553,916	1,762.9
Both	861,905	1,638,863	1,856.8

Table 1: Normalization constants for deterministic traveltime errors, for 1.0 sec rms misfit. For other values of the rms misfit, multiply by the rms. The constants do not depend on the degree of the model.

l_{max}	N	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.001$
4	25	75.4	88.6	105.2
8	81	206.0	227.0	252.3
12	169	400.6	429.4	463.2
16	289	659.3	695.7	738.1
20	441	981.9	1026.0	1077.1

Table 2: Normalization constants for simultaneous confidence envelopes around parametrized least-squares estimates of the CMB, assuming Gaussian traveltime errors with standard deviation σ = 1.0sec. The constants do not depend on the number of data. Column 1: maximum degree of the perturbation. Column 2: number of spherical harmonic coefficients of that degree or less. Columns 3-5: normalization constants for coverage probabilities 0.95, 0.99, and 0.999 respectively.

Procedure	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.001$
LS	3.920	5.152	6.582
Minimax	3.290	4.652	6.180

Table 3: Normalization constants for individual (not simultaneous) confidence intervals, assuming Gaussian traveltime errors with σ = 1.0sec. The constants do not depend on the number of data or the degree of the model. Columns 1-3: coverage probabilities 0.95, 0.99 and 0.999 respectively. Row 1: Constants for least-squares confidence intervals. Row 2: Constants for lower bounds on minimax confidence intervals using Donoho's theory.

l_{max}	min	LQ	Med	UQ	max	mean
4	0.052	0.120	0.182	0.300	0.475	0.214
8	0.065	0.231	0.391	0.659	1.672	0.476
12	0.076	0.346	0.612	1.099	5.107	0.803
16	0.085	0.460	0.837	1.598	14.75	1.232
20	0.093	0.578	1.075	2.186	40.46	1.850

Table 4: Distribution of unnormalized uncertainty $\omega(1)$ for PcP data only. Column 1: maximum model degree. Col. 2-7: minimum, lower quartile (by area), median, upper quartile, and average of $\omega(1)$.

l_{max}	min	LQ	Med	UQ	max	mean
4	0.017	0.027	0.032	0.042	0.057	0.034
8	0.024	0.048	0.065	0.099	0.242	0.076
12	0.028	0.069	0.101	0.161	0.609	0.125
16	0.031	0.089	0.137	0.227	1.213	0.180
20	0.034	0.110	0.174	0.294	2.251	0.240

Table 5: Same as table 4, but for PKP data only.

l_{max}	min	LQ	Med	UQ	max	mean
4	0.017	0.025	0.031	0.041	0.056	0.033
8	0.022	0.046	0.063	0.094	0.163	0.071
12	0.027	0.065	0.096	0.151	0.367	0.113
16	0.030	0.085	0.130	0.212	0.661	0.158
20	0.033	0.105	0.165	0.274	0.998	0.206

Table 6: Same as table 4 but for PcP and PKP data combined.

l_{max}	min	LQ	Med	UQ	max	mean
4	0.45	0.66	0.82	1.08	1.47	0.87
8	1.42	2.90	3.97	5.93	10.3	4.45
12	3.13	7.53	11.1	17.5	42.5	13.1
16	5.54	15.7	24.0	39.1	122.	29.2
20	8.94	28.3	44.4	73.8	269.	55.5

Table 7: Distribution of the width of a simultaneous 99.9% confidence region for the CMB, assuming the data errors are independent and the model is found by least-squares using all PcP and PKP data. This table combines results from tables 2 and 6. Column 1: maximum model degree. Col. 2-7: minimum, lower quartile (by area), median, upper quartile, and mean width of the envelope in *km*.

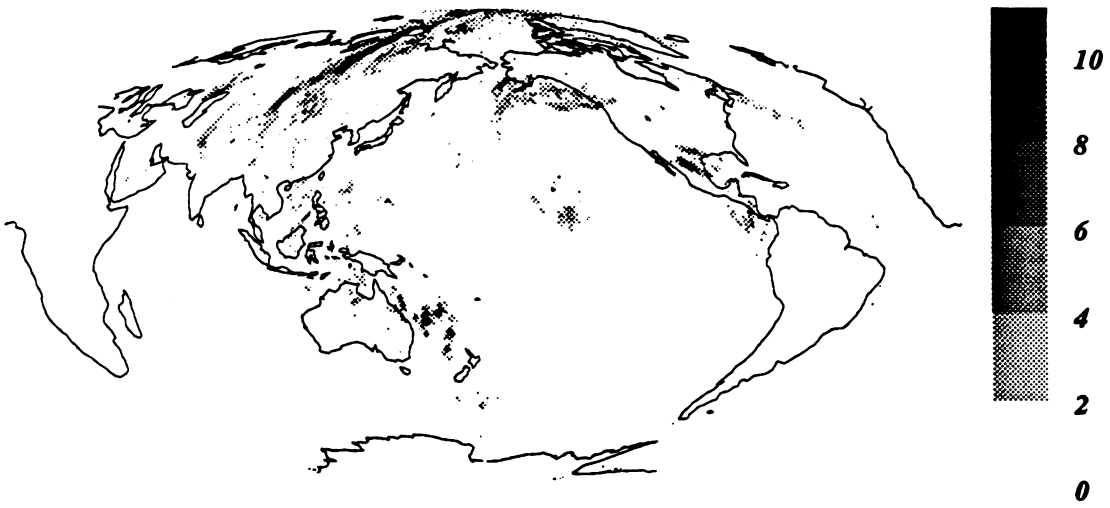


Fig. 1a(i): normalized PcP hits/unit area on CMB. max=100

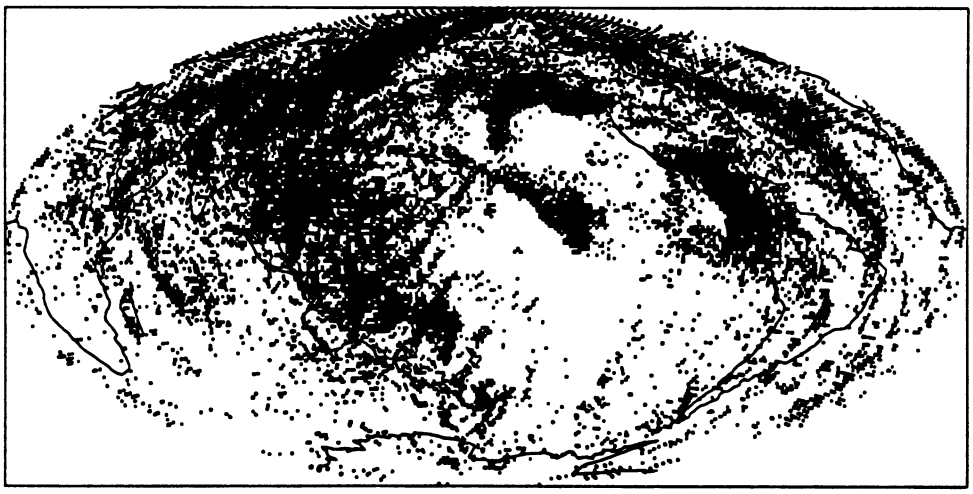


Fig. 1a(ii): Cells on the CMB hit by PcP

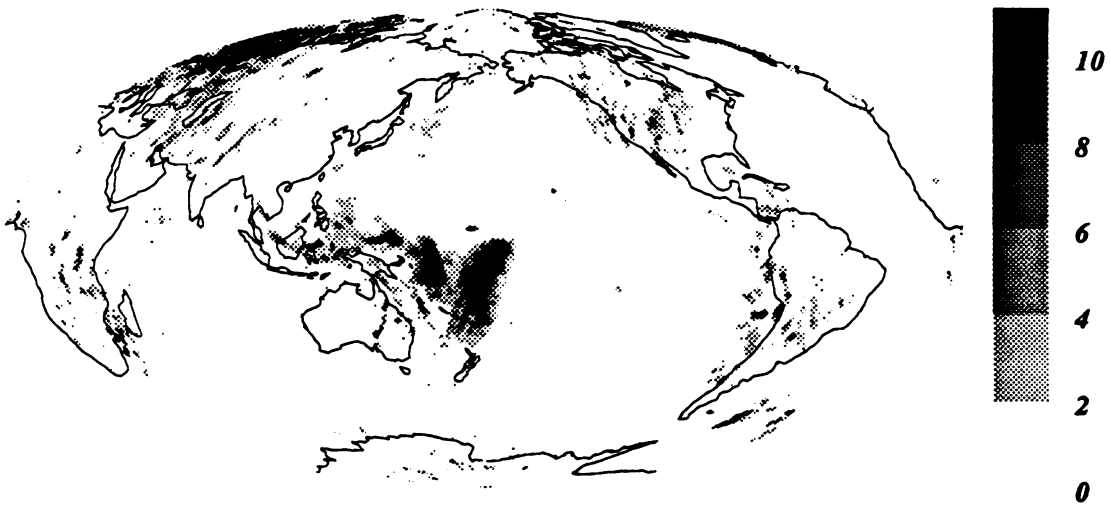


Fig. 1b(i): normalized PKP hits/unit area on CMB. max=100

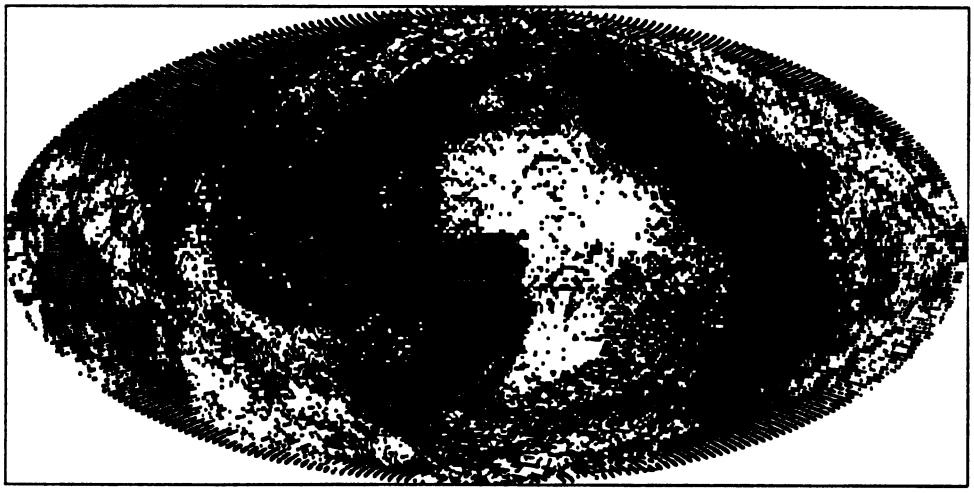


Fig. 1b(ii): Cells on the CMB hit by PKP

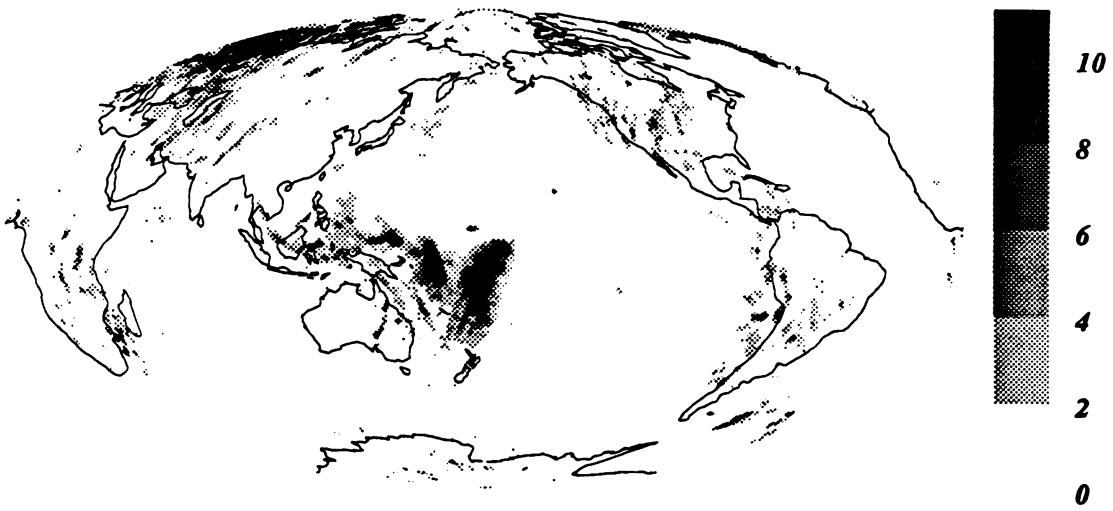


Fig. 1c(i): normalized PcP and PKP hits/unit area on CMB. max=100

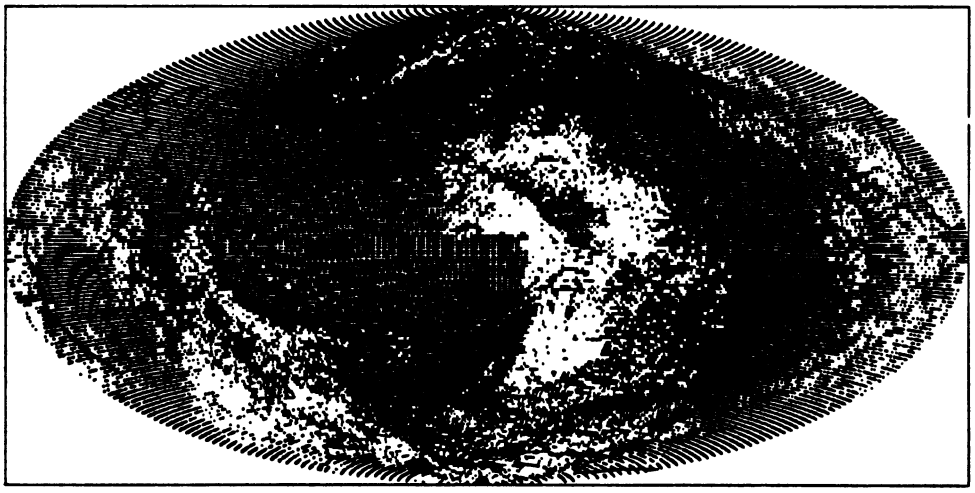


Fig. 1c(ii): Cells on the CMB hit by PcP or PKP

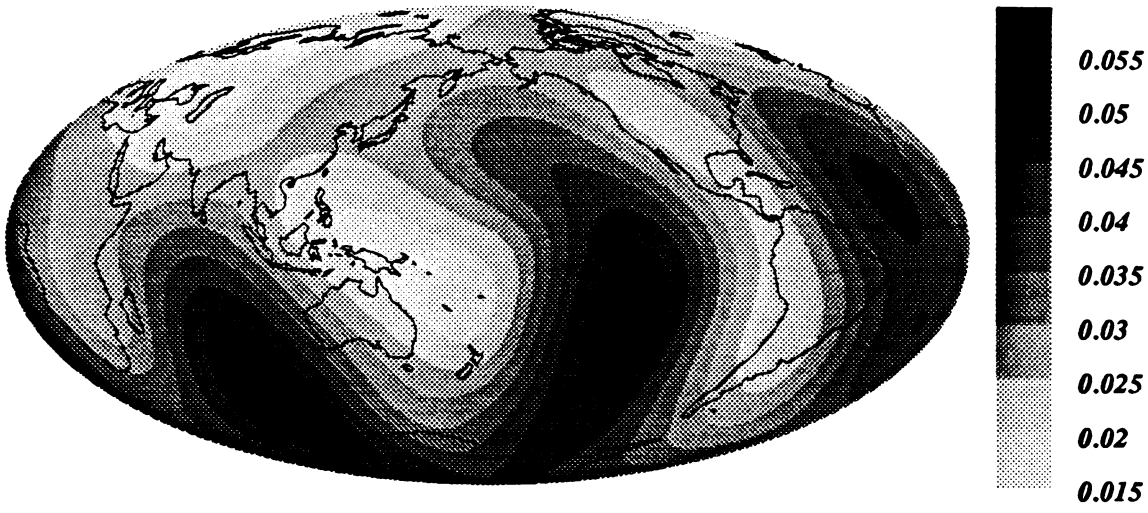


Fig. 2a: omega(1), Max l=4. PcP and PKP. min=0.0168, max=0.0562

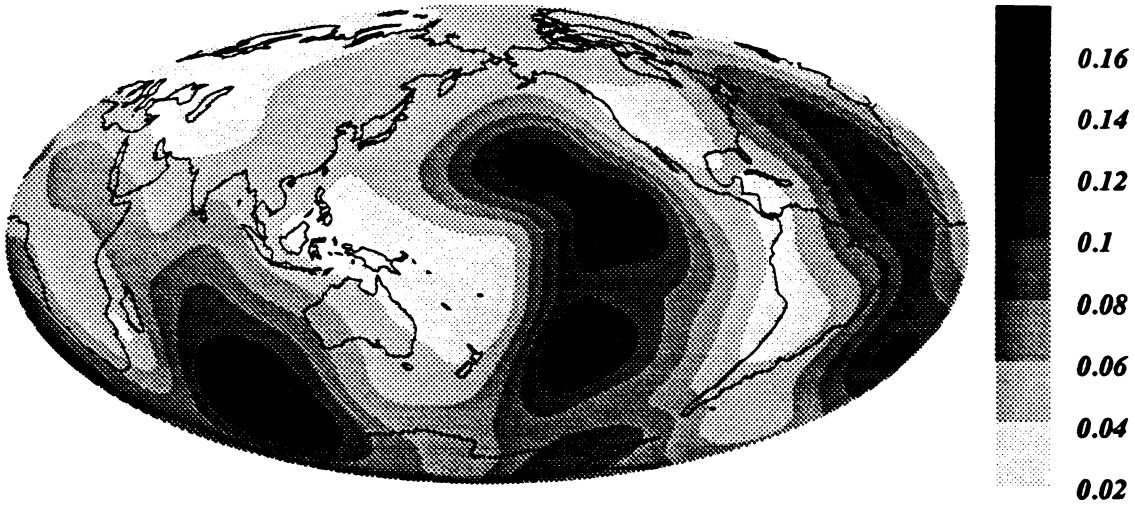


Fig. 2b: omega(1), Max l=8. PcP and PKP. min=0.0225, max=0.1632

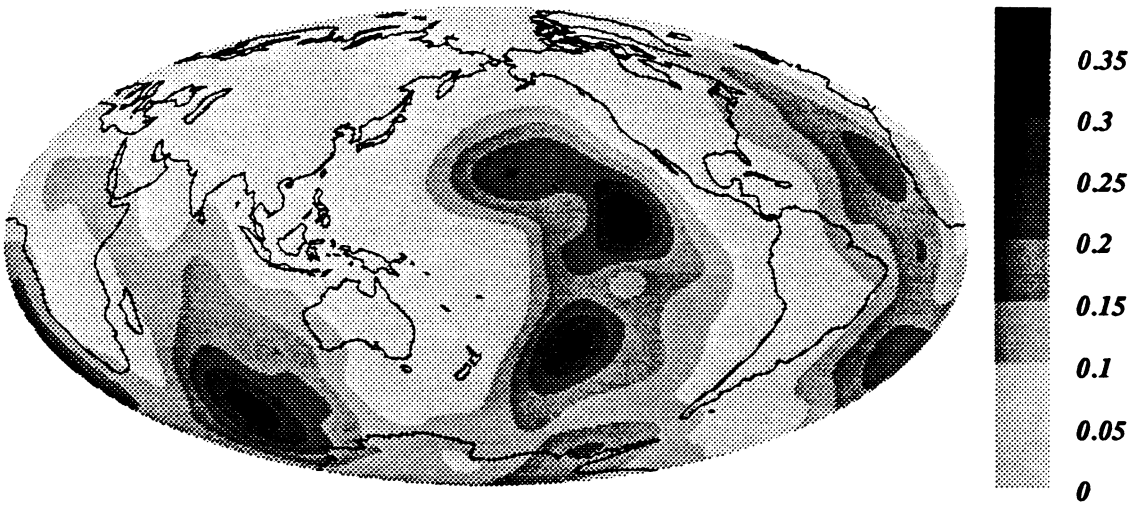


Fig. 2c: omega(1), Max l=12. PcP and PKP. min=0.0267, max=0.3672

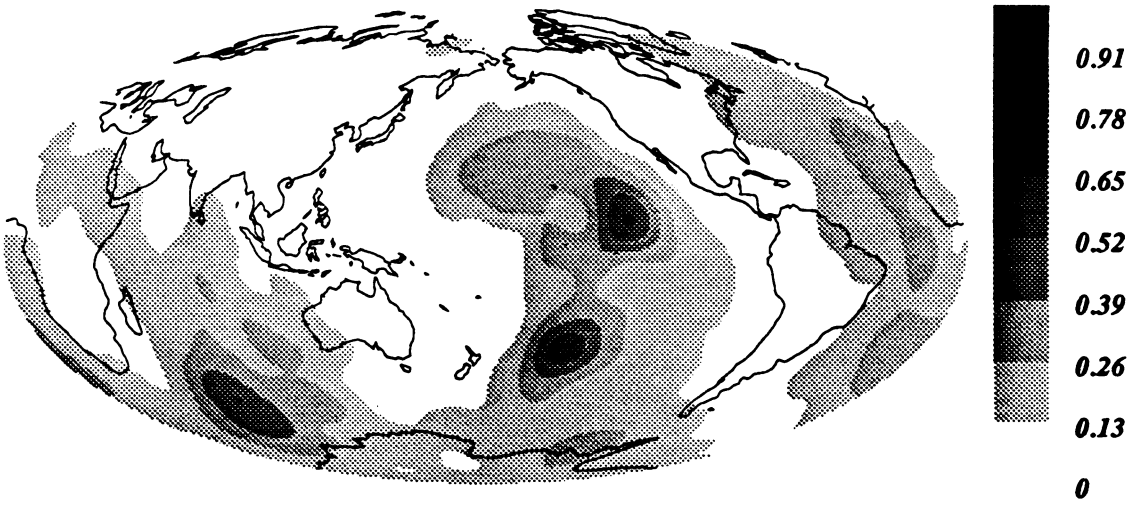


Fig. 2d: omega(1), Max l=16. PcP and PKP. min=0.0301, max=0.6608

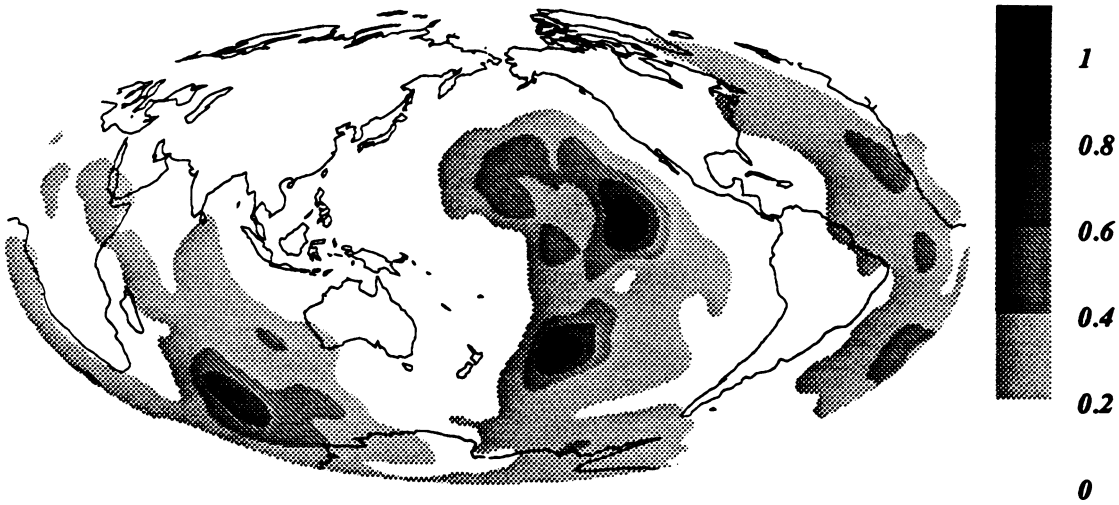


Fig. 2e: omega(1), Max l=20. PcP and PKP. min=0.033, max=0.998

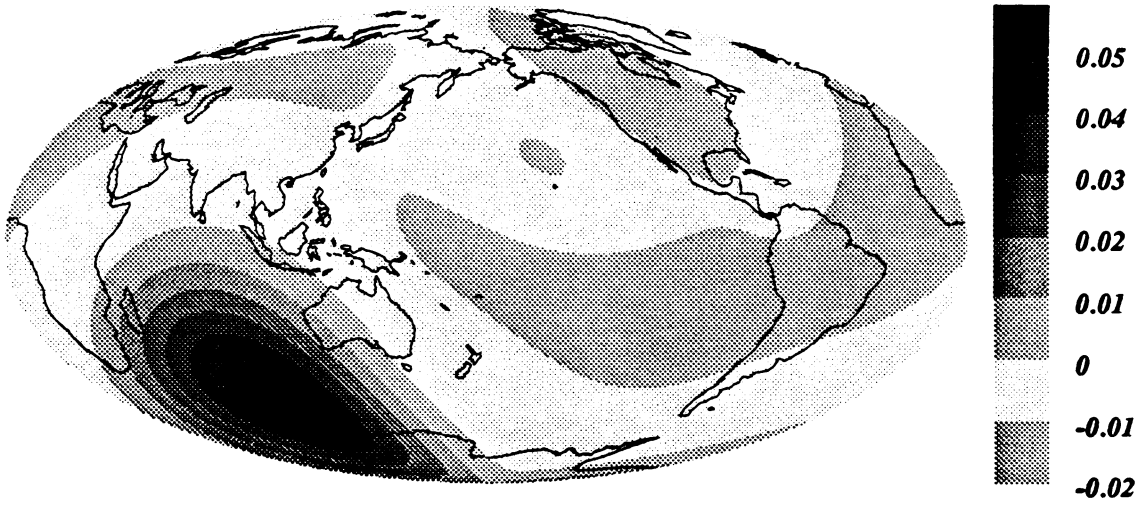


Fig. 3a: Extremal model, lmax=4

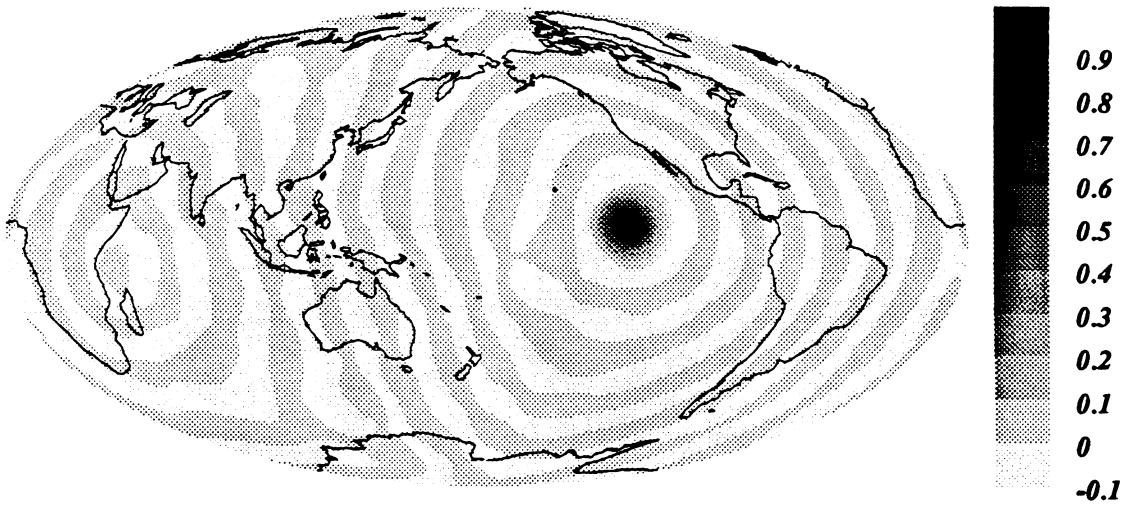


Fig. 3b: Extremal model, lmax=20

Fig. 4a: Distribution of ω(1), PcP Data

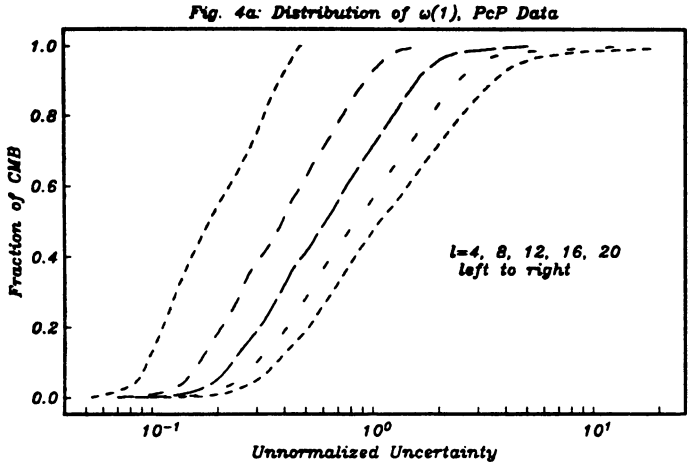


Fig. 4b: Distribution of ω(1), PKP Data

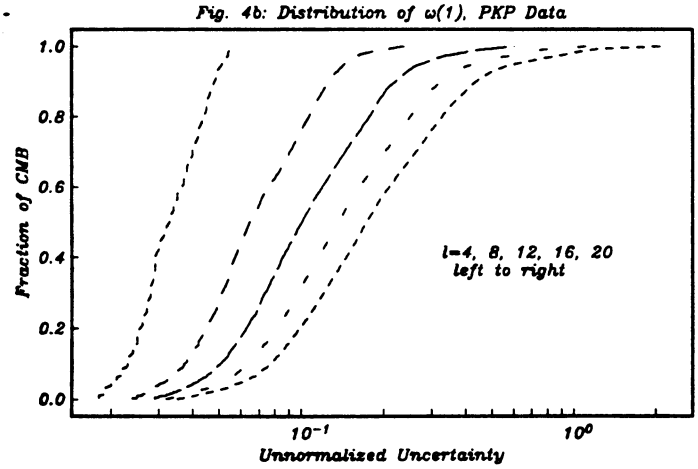
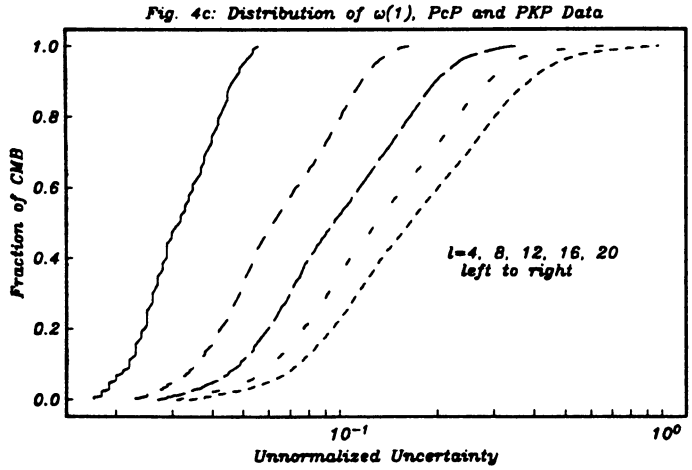


Fig. 4c: Distribution of ω(1), PcP and PKP Data



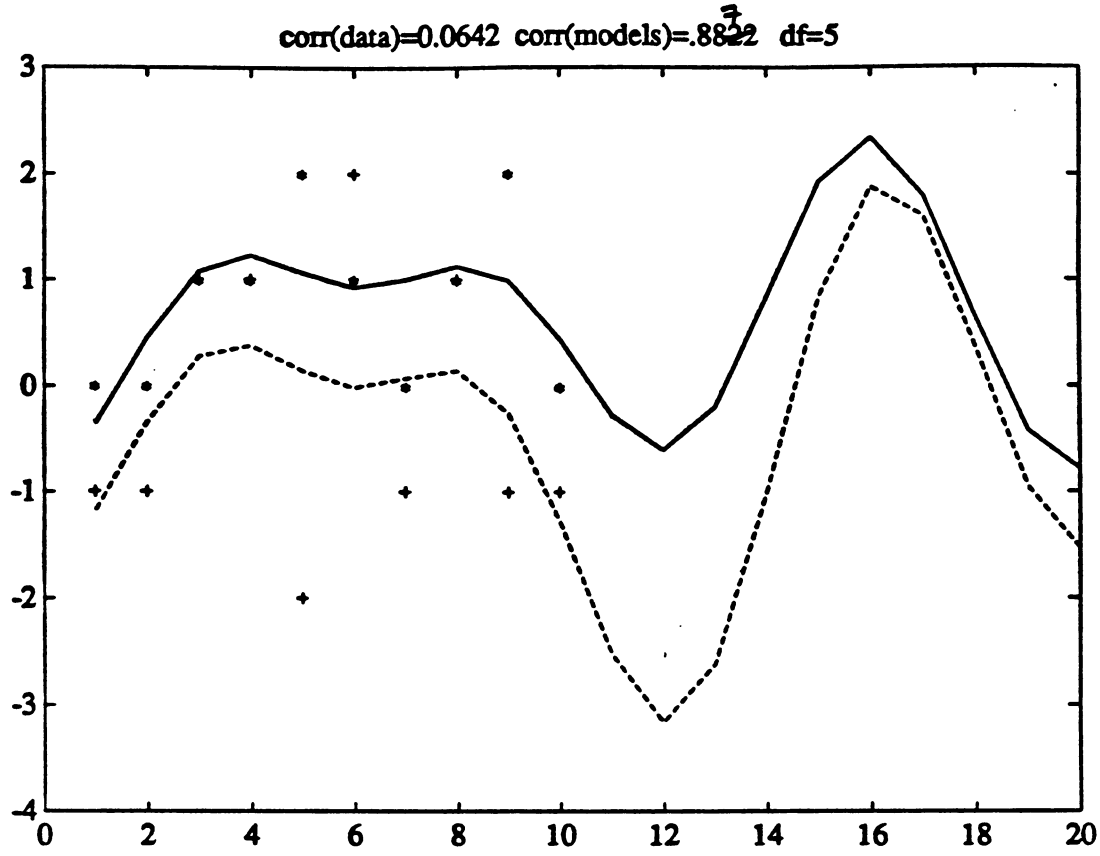

corr(data)=0.0642 corr(models)=.8822 df=5
3
2
1
0
-1
-2
-3
-4
0
2
4
6
8
10
12
14
16
18
20

Figure 5.