Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

Bootstrap Confidence Region Estimation of the Motion of Rigid Bodies

Shyamal Das Peddada & Ted Chang

To cite this article: Shyamal Das Peddada & Ted Chang (1996) Bootstrap Confidence Region Estimation of the Motion of Rigid Bodies, Journal of the American Statistical Association, 91:433, 231-241

To link to this article: https://doi.org/10.1080/01621459.1996.10476681

Bootstrap Confidence Region Estimation of the Motion of Rigid Bodies

Shyamal Das PEDDADA and Ted CHANG

This article deals with statistical inference of the motion of rigid bodies using bootstrap methodology. We consider two types of motion: motion in p dimensional Euclidean space, and motion on a p-dimensional sphere. This article is motivated by problems of interest to polar scientists understanding the motion of ice pack at the North Pole and those to geoscientists studying the motion of tectonic plates. In addition to obtaining the point estimates for various parameters describing the motion of rigid bodies, we also perform confidence region estimation and testing of certain hypotheses. We consider two types of situations: when homologous data are available, and when the data consists of nonhomologous points. We illustrate the methodology by analyzing two data sets: a data set on the motion of arctic sea ice, and a data set from plate tectonics.

KEY WORDS: Asymptotic distribution; Multivariate nonlinear regression models; Orthogonal matrix; Procrustes analysis; Spherical regression.

1. INTRODUCTION

Consider a rigid body moving in a p-dimensional space with two sets of (unobserved true) prominent points, $\mathcal{S}_1 = \{P_1, P_2, \ldots, P_m\}$ and $\mathcal{S}_2 = \{Q_1, Q_2, \ldots, Q_n\}$, on the rigid body. Let the data recorded be measurements $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ of $\mathcal{S}_1$ at time T_0 and $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ of $\mathcal{S}_2$ at time T_1. Using such data sets, a scientist may be interested in estimating the motion of the rigid body. Typically, he or she may be interested in estimating the amount of rotation and translation. Such issues arise in many physical sciences. This article is motivated by two examples: (a) the estimation of motion of Arctic sea ice, which polar scientists believe to influence earth's atmosphere, and (b) the estimation of motion of tectonic plates, which geoscientists use to understand the theory of continental drift.

The rotation of a body can be characterized by a $p \times p$ orthogonal matrix (also referred to as a rotation matrix) $\boldsymbol{\beta}$ whose determinant value is $+1$, and the translation can be characterized by a vector $\boldsymbol{\alpha} \in \mathcal{R}^p$. Throughout this article we exploit a well-known fact from mathematics that an orthogonal matrix $\mathbf{O}$ whose determinant value is $+1$ can be expressed in terms of a skew symmetric matrix $\mathbf{H}$ as follows (see, e.g., Muirhead 1982, p. 593):

$$\mathbf{O} = \exp(\mathbf{H}) = \mathbf{I} + \mathbf{H} + \frac{\mathbf{H}^2}{2!} \ldots \tag{1}$$

Thus if an orthogonal matrix $\hat{\boldsymbol{\beta}}$ is an estimator of the rotation matrix $\boldsymbol{\beta}$, then we reparameterize in terms of a skew symmetric matrix $\hat{\mathbf{H}}$ as follows:

$$\boldsymbol{\beta}'\hat{\boldsymbol{\beta}} = \exp(\hat{\mathbf{H}}). \tag{2}$$

Such a reparameterization of the rotation matrix is very useful and rather natural. Heuristically, $\hat{\mathbf{H}}$ plays exactly the same role as $\hat{\eta} - \eta$ does in the confidence region estimation of an unknown parameter η using the point estimator $\hat{\eta}$. The statistical properties of $\hat{\mathbf{H}}$, and in particular its asymptotic covariance matrix, are much simpler than those of $\hat{\boldsymbol{\beta}}$.

When a body is in motion, it is generally unlikely that we can observe the same prominent point on the rigid body at two different times, T_0 and T_1. Thus it is possible that $\mathcal{S}_1 - \mathcal{S}_2 \neq \phi$ and, similarly, $\mathcal{S}_2 - \mathcal{S}_1 \neq \phi$. In fact it may even be possible that $\mathcal{S}_1 \cap \mathcal{S}_2 = \phi$. We say that a data set is *homologous* if $m = n$ and each $P_i = Q_i$ for all i, and it is *nonhomologous* otherwise. Because it is easier to develop statistical models for homologous data sets, we discuss the case of homologous data first. In Section 2 we review multivariate nonlinear models to describe the motion of a rigid body when we have homologous data sets. Suitable pivots are determined for constructing the confidence regions for various parameters. Bootstrap methodology is developed to construct confidence regions for translation and rotation parameters. The bootstrap methodology is an important alternative to the asymptotic theory described in Section 2.1 because of two practical reasons: (a) one rarely obtains homologous data sets when a rigid body is in motion, and (b) if one does observe homologous pairs of observations, they will not be too many in number. The bootstrap methodology developed in Section 2.2 is similar in spirit to the work of Fisher and Hall (1989), who described bootstrap confidence regions for the mean direction of a random vector on a $p-1$ dimensional sphere.

The methodology developed in Section 2.1 is illustrated by applying it to two data sets: Arctic ice data and plate tectonics data. Motivated by the theoretical justification pro-

Shyamal Das Peddada is Associate Professor of Statistics and Ted Chang is Professor of Statistics, Division of Statistics, University of Virginia, Charlottesville, VA 22903. The first author thanks the University of Virginia for providing support to visit the Department of Statistics at Rice University and Stanford University, whose hospitality was greatly appreciated. This work was partially supported by Office of Naval Research Grant N00014-92-J-1009, and additional support was received from National Science Foundation Grants DMS-91-01568 and MDA-904-91-H-0031. Some computations were performed on the supercomputers at the Pittsburgh Supercomputing Center through National Science Foundation Grant DMS 930004P. The authors are thankful to polar scientists Drew Rothrock and Roger Colony, Applied Physics Laboratory, University of Washington, Seattle; Jeff Banfield, Montana State University, Bozeman; Ramin Samadani, Department of Electrical Engineering, Stanford University; and Ron Kwok, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, for discussions on various problems involved in the tracking of Arctic sea ice and the importance of such a study. Also, the authors thank Jeff Banfield for providing relevant images of ice floes that were analyzed in this article.

vided for the pivot described in Section 2, we use the same pivot in Section 3 to construct confidence regions for rotation and translation parameters when we have nonhomologous points.

2. HOMOLOGOUS DATA

2.1 Preliminaries

Suppose that a rigid body is moving in a p-dimensional Euclidean space $\mathcal{R}^p$. Then the new position $\mathbf{y}_i$, at time T_1, of the ith prominent point on the rigid body can be described in terms of the past position $\mathbf{x}_i$ at time T_0 by the following multivariate regression model:

$$\mathbf{y}_i = \boldsymbol{\alpha} + \boldsymbol{\beta}(\mathbf{x}_i - \bar{\mathbf{x}}) + \boldsymbol{\varepsilon}_i, \qquad i = 1, 2, \ldots, n. \tag{3}$$

In this model $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are as described in the last section and $\bar{\mathbf{x}}$ denotes the sample mean of all the x's. The random errors $\boldsymbol{\varepsilon}_i$ are assumed to be independently distributed with mean zero and covariance matrix $\sigma^2\mathbf{I}$. Due to the orthogonality constraint on $\boldsymbol{\beta}$, (3) is a nonlinear regression model.

Chang and Ko (1995) considered M estimates obtained by minimizing an objective function of the form

$$S(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{i=1}^{n} \psi(\|\mathbf{y}_i - \boldsymbol{\alpha} - \boldsymbol{\beta}(\mathbf{x}_i - \bar{\mathbf{x}})\|). \tag{4}$$

Let $\hat{\mathbf{a}} = \hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}$ and let $\hat{\mathbf{H}}$ be defined as in (2). For any skew symmetric matrix $\mathbf{H}$, we set $\mathbf{h} = (h_{12}, h_{13}, \ldots, h_{1p}, h_{23}, h_{24}, \ldots, h_{(p-1)p})'$, the $p(p-1)/2$ column vector formed from its distinct elements. Let $\hat{\boldsymbol{\mu}} = (\hat{\mathbf{a}}, \hat{\mathbf{h}})$.

Although $\hat{\boldsymbol{\mu}}$ is unobservable, Chang and Ko (1995) derived an asymptotic pivot in terms of $\hat{\boldsymbol{\mu}}$, whose distribution together with the point estimates $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ can be used to construct confidence regions for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. We note that standard parametric confidence intervals can be viewed as being constructed from the parameter point estimate and the distribution of its deviation from the true parameter.

The asymptotic pivot they used is

$$\hat{\boldsymbol{\mu}}'\boldsymbol{\Omega}\hat{\boldsymbol{\mu}} \equiv \hat{\mathbf{a}}'\hat{\mathbf{a}} - \operatorname{Tr}(\hat{\mathbf{H}}\boldsymbol{\Sigma}\hat{\mathbf{H}}), \tag{5}$$

after dividing by a scale correction, where

$$\boldsymbol{\Sigma}_n = \frac{1}{n}\sum_{i=1}^{n} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})', \qquad \lim_{n\to\infty} \boldsymbol{\Sigma}_n = \boldsymbol{\Sigma}. \tag{6}$$

The same asymptotic pivot can be used for all spherically symmetric distributions and all the objective functions of the form (4).

Motivated by the problems in geophysical sciences where a geoscientist is interested in reconstructing the motion of tectonic plates, Chang (1986) considered the motion of a rigid body moving on a $p-1$ dimensional sphere in Euclidean p-dimensional space. Displacement of the body between two specific times is given by a rotation matrix $\boldsymbol{\beta}$ (alone without any translation). As defined earlier, let $\mathbf{x}_i$'s and $\mathbf{y}_j$'s be the observations recorded on the past and present positions. Because the points are on the surface of a sphere (without loss of generality a unit sphere), we note that $\mathbf{x}_i'\mathbf{x}_i = \mathbf{y}_i'\mathbf{y}_i = 1$.

Chang and Ko (1995) estimated the unknown rotation matrix $\boldsymbol{\beta}$ by minimizing an objective function of the form

$$S(\boldsymbol{\beta}) = \sum_{i=1}^{n} \psi(\mathbf{y}_i'\boldsymbol{\beta}\mathbf{x}_i). \tag{7}$$

For spherically symmetric error distributions, they obtained the asymptotic distribution of the skew symmetric matrix $\hat{\mathbf{H}}$ in the representation (2). It was proven that

$$\operatorname{Tr}(\hat{\mathbf{H}}^2\boldsymbol{\Sigma}) \tag{8}$$

after correction for scale is an asymptotic pivot, where for spherical data $\boldsymbol{\Sigma}_n$ is defined to be

$$\boldsymbol{\Sigma}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i\mathbf{x}_i' \quad \text{and} \quad \lim_{n\to\infty} \boldsymbol{\Sigma}_n = \boldsymbol{\Sigma}. \tag{9}$$

2.2 Bootstrap Methodology

Justification of the bootstrap methodology described in this section is provided in the Appendix. The arguments are essentially deduced from the theory developed by Bickel and Freedman (1981), Chang (1986), and Freedman (1981); hence we are very brief in our description.

In case of the Euclidean space data, let, for each $j, r_j = \|\hat{\boldsymbol{\varepsilon}}_j\|_2$, where $\hat{\boldsymbol{\varepsilon}}_j = \mathbf{y}_j - \hat{\boldsymbol{\alpha}}(n) - \hat{\boldsymbol{\beta}}(n)(\mathbf{x}_j - \bar{\mathbf{x}})$, and let $\hat{F}_n$ denote the empirical distribution of $\{r_1, r_2, \ldots, r_n\}$. Further, let $r_1^*(n), r_2^*(n) \ldots, r_m^*(n) \sim^{\text{iid}} \hat{F}_n$. Independently generate $\boldsymbol{\varepsilon}_i^*(n)$ uniformly on a sphere of radius $r_i^*(n)$. Using $\boldsymbol{\varepsilon}_i^*(n), i = 1, 2, \ldots, m$, we generate the "data" $\mathbf{y}_i^*(n)$ according to

$$\mathbf{y}_i^*(n) = \hat{\boldsymbol{\alpha}}(n) + \hat{\boldsymbol{\beta}}(n)(\mathbf{x}_i - \bar{\mathbf{x}}) + \boldsymbol{\varepsilon}_i^*(n), \qquad i = 1, 2, \ldots, m. \tag{10}$$

Usually we take $m = n$. For $m < n$, assumption (6) is reasonable for the x's in (10) if $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ do not occur in a sorted order but occur in a random order along the boundary of the rigid body.

Let $\hat{\boldsymbol{\alpha}}^*(n, m)$ and $\hat{\boldsymbol{\beta}}^*(n, m)$ be obtained by minimizing $S(\boldsymbol{\alpha}, \boldsymbol{\beta})$, for $\psi(t) = t^2$, and let

$$s^2 = \frac{S(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}})}{2n-3}, \qquad s^*(n,m)^2 = \frac{S(\hat{\boldsymbol{\alpha}}^*(n,m), \hat{\boldsymbol{\beta}}^*(n,m))}{2m-3}. \tag{11}$$

It can be verified that $\hat{\boldsymbol{\alpha}}^*(n, m), \hat{\boldsymbol{\beta}}^*(n, m)$, and $s^*(n, m)^2$ are consistent, as $(n, m) \to \infty$, for $\boldsymbol{\alpha}, \boldsymbol{\beta}$, and σ^2. Defining

$$\begin{aligned} \hat{\mathbf{a}}^*(n,m) &= \hat{\boldsymbol{\alpha}}^*(n,m) - \hat{\boldsymbol{\alpha}}(n), \\ \hat{\boldsymbol{\beta}}^*(n,m) &= \hat{\boldsymbol{\beta}}(n)\exp(\hat{\mathbf{H}}^*(n,m)), \end{aligned} \tag{12}$$

and dropping the symbols (n, m) for simplicity of notation, we obtain the bootstrap confidence region for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as follows:

Table 1. Bootstrap Critical Values for the Gulf of Aden Data

Nominal value	Bootstrap critical values	F critical values with (3, 30) df
.20	.33	.35
.50	.76	.81
.75	1.36	1.44
.90	2.16	2.28
.95	2.75	2.92
.99	4.57	4.51

$$C_{\text{boot}} = \left\{ (\boldsymbol{\alpha}, \boldsymbol{\beta}) | n \frac{\mathbf{a}'\mathbf{a} - \text{Tr}(\mathbf{H}\boldsymbol{\Sigma}_n\mathbf{H})}{s^2} \le c_\alpha, \quad \text{with} \right.$$

$$\left. \mathbf{a} = \hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha} \quad \text{and} \quad \exp(\mathbf{H}) = \boldsymbol{\beta}'\hat{\boldsymbol{\beta}} \right\}, \quad (13)$$

where c_α is obtained as

$$P\left\{ m \frac{\hat{\mathbf{a}}^{*\prime}\hat{\mathbf{a}}^* - \text{Tr}(\hat{\mathbf{H}}^*\boldsymbol{\Sigma}_n\hat{\mathbf{H}}^*)}{s^{*2}} \le c_\alpha | \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}} \right\} = 1 - \alpha. \quad (14)$$

For the spherical data, let $\hat{\boldsymbol{\beta}}(n)$ be the Ordinary Least Squares Estimator (OLSE); that is, the estimate obtained by using $\psi(t) = t$ in (7). Let $\hat{F}_n$ denote the empirical cdf of the angles between $\mathbf{y}_j$ and $\hat{\boldsymbol{\beta}}(n)\mathbf{x}_j, j = 1, 2, \ldots, n$. Generate bootstrap data $\mathbf{y}_i^*(n), i = 1, 2, \ldots, m$ as follows: $\mathbf{y}_i^*(n)$ has a colatitude; that is, an angle in radians measured from $\hat{\boldsymbol{\beta}}(n)\mathbf{x}_i$, which is distributed as $\hat{F}_n$ and is uniformly distributed on the "small circle" of dimension $p - 2$, defined as the points that have the chosen colatitude. Then $E(\mathbf{y}_i^*(n)|\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n) = \hat{c}_0(n)\hat{\boldsymbol{\beta}}(n)\mathbf{x}_i$ and $\text{cov}(\mathbf{y}_i^*(n)|\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n) = \hat{c}_1(n)\hat{\boldsymbol{\beta}}(n)\mathbf{x}_i\mathbf{x}_i'\hat{\boldsymbol{\beta}}(n)' + \hat{c}_2(n)\mathbf{I}$, where $\hat{c}_0(n), \hat{c}_1(n)$ and $\hat{c}_2(n)$ are as defined by Chang (1986). Let $\hat{\boldsymbol{\beta}}(n, m)^*$ be the OLSE of $\hat{\boldsymbol{\beta}}(n)$ under the bootstrap sample and similarly define $c_0^*(n, m)$ and $c_2^*(n, m)$. Consequently, the pivotal statistic is

$$- \frac{n\hat{c}_0^*(n, m)^2 \text{Tr}(\mathbf{H}(n, m)^*\boldsymbol{\Sigma}\mathbf{H}(n, m)^*)}{2\hat{c}_2^*(n, m)}. \quad (15)$$

As an example, consider the Gulf of Aden data set reported by Rivest (1989). This data set consists of 11 homologous pairs of points on the Somalian plate and the Arabian plate, which were coincident 5.37 million years ago. Geophysicists are interested in reconstructing the past position of the Somalian plate relative to the Arabian plate. This relative position can be described by a rotation matrix. The errors in tectonic data are extremely concentrated. Using this fact, and modeling the errors according to a Fisher distribution, Rivest (1989) calculated an asymptotic confidence region for the rotation using critical values from an F distribution with (3, 30) degrees of freedom. Nevertheless, Rivest noticed that 2 of the 11 pairs of points are outliers relative to the spread of the Fisher distribution. We are interested in comparing the F critical values to those obtained by bootstrap methods. We used 1,000 bootstrap resamples were used, and report the results in Table 1.

We simulated the coverage probabilities of these bootstrap confidence regions, using the $\mathbf{x}$ points of the Gulf of Aden data set and generating the $\mathbf{y}$ points according to the Fisher distribution with $\kappa = 1, 5, 100$. Distributions of $\mathbf{y}$ points with heavier tails than the Fisher distribution were also generated as follows. At each point $\boldsymbol{\beta}\mathbf{x}$, the square of colatitude between $\mathbf{y}$ and $\boldsymbol{\beta}\mathbf{x}$ was generated according to a $(2/c)F(2, df)$ distribution, for $c = 1, 5, 100$ and $df = 2, 5$. Conditionally on the generated colatitude, the position of $\mathbf{y}$ was uniformly chosen on the small circle with the specified colatitude, the position of $\mathbf{y}$ was uniformly chosen on the small circle with the specified colatitude. For large df and large c, this distribution of $\mathbf{y}$ is approximately Fisher with $\kappa = c$. Thus c and κ can be interpreted similarly.

All simulations were based on 1,000 Monte Carlo sets of 11 homologous pairs $(\mathbf{x}, \mathbf{y})$ and 1,000 bootstrap resamples for each Monte Carlo set. The results, reported in Table 2, have a 95% margin of error of $\pm.03$. It appears that $\kappa = 5$ for the Fisher-generated $\mathbf{y}$ or $c = 5$ for the heavy-tailed $\mathbf{y}$ is sufficient for the bootstrap approximation to the true coverage probabilities to be reasonable. We note that the errors in tectonic data are typically on the order of 10–20 kilometers, indicating values of c or κ approximately equal to 10^6.

In contrast to the standard linear model with normal errors, the speed of convergence to asymptotics for the spherical regression model with Fisher-distributed $\mathbf{y}$ points depends on the geometry of the $\mathbf{x}$ points (cf. Bingham, Chang, and Richards 1992). In general, convergence is slower for geometries in which the $\mathbf{x}$ points are concentrated in a small region of the sphere. We used simulations to explore the dependence of the bootstrap approximation for a variety of $\mathbf{x}$ point geometries. Surprisingly, for reasonably concentrated $\mathbf{y}$ point error distributions, the performance of the bootstrap confidence regions did not depend on the geometry of $\mathbf{x}$ points.

We considered 5 geometries of 11 $\mathbf{x}$ points: the original Gulf of Aden geometry, 3 geometries in which the $\mathbf{x}$ points were randomly generated according to Fisher distributions with $\kappa = 1, 5, 100$, and a geometry by polar projecting onto the sphere 11 bivariate normally distributed points with covariance matrix diag[.1, 10] (which we call *elliptic geometry*). The latter geometry concentrates the $\mathbf{x}$ points close to a great circle. Results are provided in Table 2.

2.3 Inference on the Motion of Several Rigid Bodies in the Euclidean Space

Suppose that there are k rigid bodies in motion and, corresponding to the ith rigid body, n_i homologous pairs of measurements $\mathbf{x}_{ij}, \mathbf{y}_{ij}, j = 1, 2, \ldots, n_i$ are recorded at times T_0 and T_1.

Let the motion of the ith rigid body be characterized by a translation vector $\boldsymbol{\alpha}_i$ and a rotation matrix $\boldsymbol{\beta}_i$. Then a multivariate nonlinear regression model, describing the motion of the rigid body, is given by

$$\mathbf{y}_{ij} = \boldsymbol{\alpha}_i + \boldsymbol{\beta}_i\mathbf{x}_{ij} + \boldsymbol{\varepsilon}_{ij}, \qquad i = 1, 2, \ldots, k,$$
$$j = 1, 2, \ldots, n_i. \quad (16)$$

Table 2. Coverage Probabilities of Bootstrap Confidence Regions

Geometry of x points	Nominal values	Geometry of y points: Fisher, κ = 1	Fisher, κ = 5	Fisher, κ = 100	F(2, 2), c = 1	F(2, 2), c = 5	F(2, 2), c = 100	F(2, 5), c = 1	F(2, 5), c = 5	F(2, 5), c = 100
Gulf of Aden data	.20	.16	.22	.22	.10	.21	.20	.09	.19	.20
	.50	.41	.52	.51	.27	.50	.42	.28	.51	.51
	.75	.62	.77	.76	.46	.74	.76	.47	.77	.75
	.90	.79	.90	.90	.65	.91	.90	.65	.90	.90
	.95	.88	.95	.94	.76	.94	.94	.76	.94	.95
	.99	.96	.99	.99	.90	.99	.98	.90	.99	.98
Elliptic data	.20	.14	.20	.20	.04	.21	.22	.04	.20	.20
	.50	.34	.48	.49	.14	.50	.50	.14	.50	.52
	.75	.57	.72	.72	.32	.73	.75	.37	.72	.76
	.90	.76	.88	.90	.60	.89	.89	.60	.87	.90
	.95	.84	.94	.94	.72	.92	.95	.72	.93	.95
	.99	.94	.99	.99	.89	.97	.99	.88	.98	.99
Fisher κ = 1	.20	.11	.21	.20	.02	.19	.20	.02	.19	.21
	.50	.31	.49	.50	.10	.50	.49	.10	.50	.51
	.75	.53	.72	.76	.28	.74	.75	.28	.71	.75
	.90	.74	.89	.91	.54	.89	.89	.54	.87	.89
	.95	.82	.94	.96	.68	.93	.94	.69	.93	.95
	.99	.92	.98	.99	.86	.98	.98	.87	.98	.98
Fisher κ = 5	.20	.11	.20	.20	.03	.17	.19	.03	.22	.19
	.50	.30	.48	.50	.12	.48	.50	.12	.50	.51
	.75	.51	.73	.75	.35	.73	.72	.35	.73	.74
	.90	.72	.89	.90	.61	.89	.88	.61	.87	.89
	.95	.82	.94	.95	.75	.93	.93	.74	.93	.95
	.99	.94	.98	.99	.90	.98	.99	.89	.98	.99
Fisher κ = 100	.20	.16	.20	.20	.11	.20	.21	.11	.21	.21
	.50	.40	.50	.50	.30	.50	.49	.26	.49	.49
	.75	.65	.74	.75	.43	.72	.72	.45	.72	.74
	.90	.81	.89	.89	.63	.87	.88	.65	.87	.87
	.95	.88	.95	.94	.75	.94	.94	.76	.93	.93
	.99	.96	.99	.99	.90	.99	.98	.90	.98	.99

When considering several bodies, it is inconvenient to center the $\mathbf{x}_{ij}$, because such a recentering interferes with the interpretation of translation parameter.

We shall assume, as before, that the random errors ε_{ij} are independent and identically spherically symmetrically distributed with mean zero and covariance matrix $\sigma^2\mathbf{I}$.

To understand the dynamics of the region, it is very reasonable for a scientist to ask whether all of the rigid bodies in a given region are rotating and translating by an equal amount. If they do, then what is an estimate of the common rotation and translation parameters? In other words, it is reasonable for a scientist to test the null hypothesis H_0: $(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i) = (\boldsymbol{\alpha}, \boldsymbol{\beta})$ for all i, where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the common unknown translation and rotation parameters, respectively. Because one can obtain a closed-form expression for the OLSE of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, in this section we work in terms of the OLSE.

To test H_0, we mimic the usual analysis of variance. Thus we compare the variability in motion between rigid bodies to that of the variability in motion within rigid bodies. Before constructing "between group sum of squares" and "within group sum of squares," we develop some useful notation.

Throughout this section, corresponding to any random vector $\mathbf{w}_{ij}$, we shall use the standard notation $\bar{\mathbf{w}}_{i\cdot} = \sum_{j=i}^{n_i} \mathbf{w}_{ij}/n_i$ and $\bar{\mathbf{w}}_{\cdot\cdot} = \sum_{i=1}^{k} n_i\bar{\mathbf{w}}_{i\cdot}/n, n = \sum_{i=1}^{k} n_i$. The OLSE's for $\boldsymbol{\beta}_i$ and $\boldsymbol{\alpha}_i$ are

$$\hat{\boldsymbol{\beta}}_i = \mathbf{O}_{2i}\mathbf{O}_{1i}', \quad \text{and} \quad \hat{\boldsymbol{\alpha}}_i = \bar{\mathbf{y}}_{i\cdot} - \hat{\boldsymbol{\beta}}_i\bar{\mathbf{x}}_{i\cdot}, \tag{17}$$

where $\mathbf{O}_{1i}\boldsymbol{\Lambda}_i\mathbf{O}_{2i}'$ is the modified singular value decomposition of $\sum_{i=1}^{n_i}(\mathbf{x}_{ij} - \bar{\mathbf{x}}_{i\cdot})(\mathbf{y}_{ij} - \bar{\mathbf{y}}_{i\cdot})'$, $\mathbf{O}_{1i}$ and $\mathbf{O}_{2i}$ are special orthogonal matrices, and $\boldsymbol{\Lambda}_i$ is a diagonal matrix with diagonal entries $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{p-1} \geq |\lambda_p|$. The sum of squares due to error (SSE; also called the within-group sum of squares) is given by

$$\text{SSE} = \sum_{i,j} (\mathbf{y}_{ij} - \bar{\mathbf{y}}_{i\cdot} - \hat{\boldsymbol{\beta}}_i(\mathbf{x}_{ij} - \bar{\mathbf{x}}_{i\cdot}))' \times (\mathbf{y}_{ij} - \bar{\mathbf{y}}_{i\cdot} - \hat{\boldsymbol{\beta}}_i(\mathbf{x}_{ij} - \bar{\mathbf{x}}_{i\cdot})). \tag{18}$$

Under the null hypothesis H_0, the model (16) reduces to

$$\mathbf{y}_{ij} = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{x}_{ij} + \varepsilon_{ij}. \tag{19}$$

The OLSE's of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ are

$$\hat{\boldsymbol{\beta}} = \mathbf{O}_2\mathbf{O}_1', \quad \text{and} \quad \hat{\boldsymbol{\alpha}} = \bar{\mathbf{y}}_{\cdot\cdot} - \hat{\boldsymbol{\beta}}\bar{\mathbf{x}}_{\cdot\cdot}, \tag{20}$$

where $\mathbf{O}_1\boldsymbol{\Lambda}\mathbf{O}_2'$ is the modified singular value decomposition of $\sum_{i,j}(\mathbf{x}_{ij} - \bar{\mathbf{x}}_{\cdot\cdot})(\mathbf{y}_{ij} - \bar{\mathbf{y}}_{\cdot\cdot})'$, $\mathbf{O}_1$ and $\mathbf{O}_2$ are orthog-

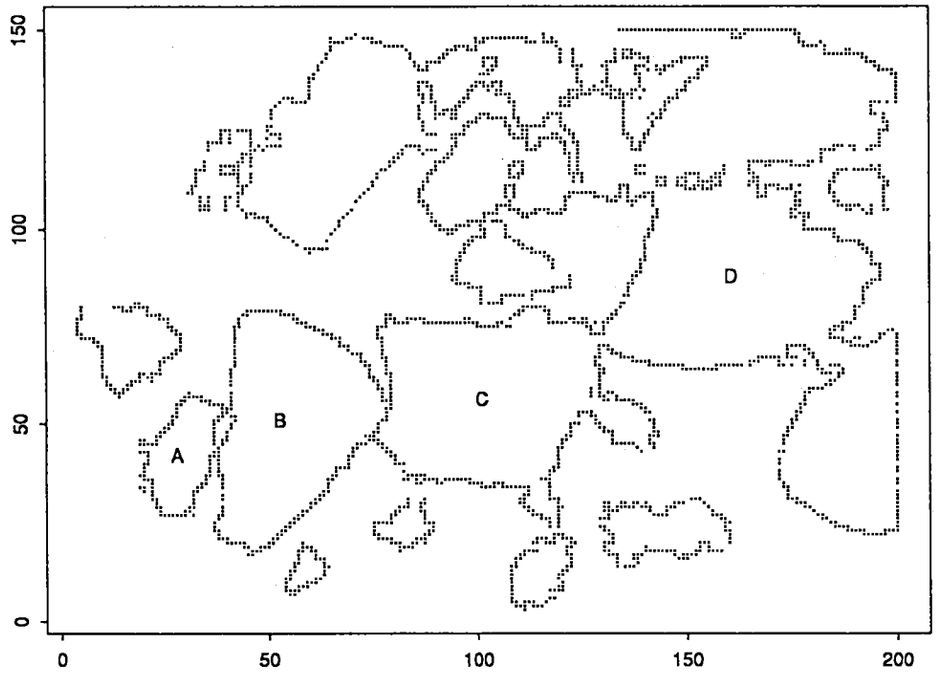

Figure 1. Image of Some Ice Floes in the Arctic Sea. Image processed by Banfield and Raftery (1992) (cleaned).

onal matrices, and Λ is a suitable diagonal matrix. The SSE under the null hypothesis is then given by

$$\mathrm{SSE}_{H_0} = \sum_{i,j} (\mathbf{y}_{ij} - \bar{\mathbf{y}}_{\cdot\cdot} - \hat{\boldsymbol{\beta}}(\mathbf{x}_{ij} - \bar{\mathbf{x}}_{\cdot\cdot}))' \times (\mathbf{y}_{ij} - \bar{\mathbf{y}}_{\cdot\cdot} - \hat{\boldsymbol{\beta}}(\mathbf{x}_{ij} - \bar{\mathbf{x}}_{\cdot\cdot})). \quad (21)$$

The between-group sum of squares is then given by $\mathrm{SSE}_{H_0} - \mathrm{SSE}$.

Under the model (16), the usual results of nonlinear regression imply that

$$F = \left\{ \frac{np - kp(p+1)/2}{(k-1)p(p+1)/2} \right\} \frac{\mathrm{SSE}_{H_0} - \mathrm{SSE}}{\mathrm{SSE}} \quad (22)$$

is approximately distributed as an F variable with $((k-1)p(p+1)/2, p(n-k(p+1)/2))$ degrees of freedom.

Typically, in the context of polar ice data, it is difficult to identify more than a few homologous points. Consequently,

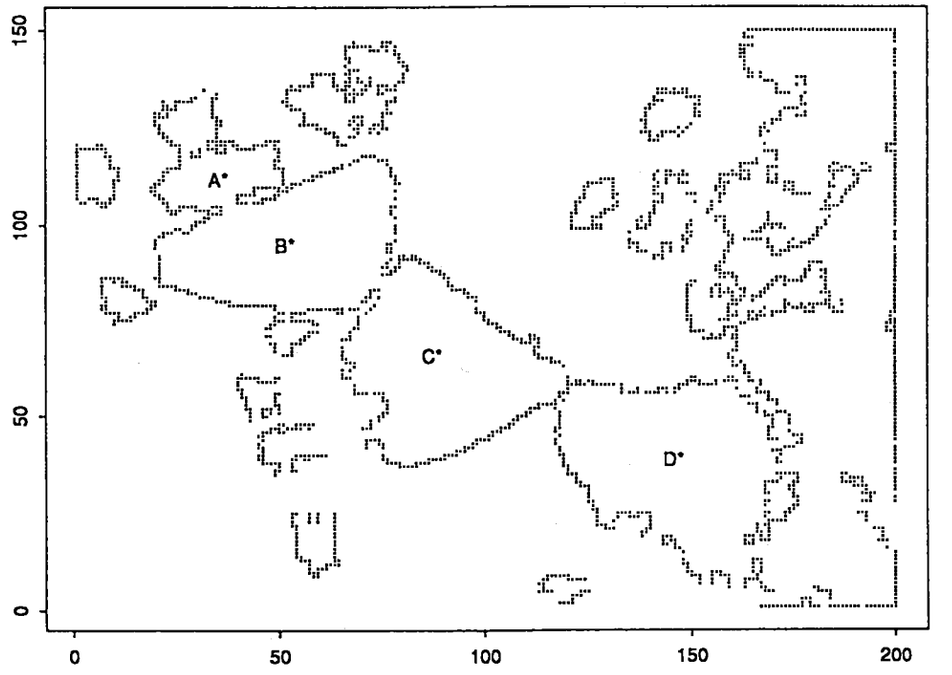

Figure 2. Figure 1 After a Few Days. Image processed by Banfield and Raftery (1992) (cleaned).

Table 3. Floes A, B in Figure 1 and A, B* in Figure 2*

A and A*		B and B*	
x	y	x	y
35, 37	30, 105	50, 19	23, 83
36, 37	30, 104	43, 75	73, 117
40, 53	50, 109	39, 27	26, 98
40, 54	50, 110	40, 20	22, 98
22, 31	23, 115		
23, 31	23, 114		

rather than using the aforementioned F approximation, we propose bootstrap methodology applied to (22).

Bootstrap methodology in the present situation can be developed along the same lines as in the previous section. Thus we do not prove any theorems but merely state the relevant results.

Let $\hat{\varepsilon}_{ij}, i = 1, 2, \ldots, k, j = 1, 2, \ldots, n_i$ denote the residuals $\mathbf{y}_{ij} - \hat{\boldsymbol{\alpha}}_i - \hat{\boldsymbol{\beta}}_i \mathbf{x}_{ij}$. Generate $\hat{\varepsilon}^*_{ij}$ as in the previous section from the joint empirical distribution of the ε_{ij} and let the bootstrap sample be obtained according to

$$\mathbf{y}^*_{ij} = \hat{\boldsymbol{\alpha}}_i + \hat{\boldsymbol{\beta}}_i \mathbf{x}_{ij} + \hat{\varepsilon}^*_{ij}. \tag{23}$$

Replacing $\mathbf{y}_{ij}, \bar{\mathbf{y}}_{i\cdot}, \bar{\mathbf{y}}_{\cdot\cdot}, \varepsilon_{ij}, \bar{\varepsilon}_{i\cdot}$, and $\bar{\varepsilon}_{\cdot\cdot}$ by $\mathbf{y}^*_{ij}, \bar{\mathbf{y}}^*_{i\cdot}, \bar{\mathbf{y}}^*_{\cdot\cdot}, \varepsilon^*_{ij}, \bar{\varepsilon}^*_{i\cdot}$, and $\bar{\varepsilon}^*_{\cdot\cdot}$, and repeating similar calculations as before, we obtain

$$\mathrm{SSE}^* = \sum_{i,j} (\mathbf{y}^*_{ij} - \bar{\mathbf{y}}^*_{i\cdot} - \hat{\boldsymbol{\beta}}^*_i(\mathbf{x}_{ij} - \bar{\mathbf{x}}_{i\cdot}))' \times (\mathbf{y}^*_{ij} - \bar{\mathbf{y}}^*_{i\cdot} - \hat{\boldsymbol{\beta}}^*_i(\mathbf{x}_{ij} - \bar{\mathbf{x}}_{i\cdot})) \tag{24}$$

and

$$\mathrm{SSE}^*_{H_0} = \sum_{i,j} (\mathbf{y}^*_{ij} - \bar{\mathbf{y}}^*_{\cdot\cdot} - \hat{\boldsymbol{\beta}}^*(\mathbf{x}_{ij} - \bar{\mathbf{x}}_{\cdot\cdot}))' \times (\mathbf{y}^*_{ij} - \bar{\mathbf{y}}^*_{\cdot\cdot} - \hat{\boldsymbol{\beta}}^*(\mathbf{x}_{ij} - \bar{\mathbf{x}}_{\cdot\cdot})). \tag{25}$$

Thus we have

$$F^* = \left\{ \frac{np - kp(p+1)/2}{(k-1)p(p+1)/2} \right\} \frac{\mathrm{SSE}^*_{H_0} - \mathrm{SSE}^*}{\mathrm{SSE}^*} \tag{26}$$

2.4 Illustration

In recent years there has been a considerable amount of interest among the polar scientists in studying the motion of the Arctic sea ice. We illustrate our methodology by analyzing two images, Figures 1 and 2, of a field of ice floes taken a few days apart by the SAR satellite. (These data were made available to us by Professor Jeff Banfield of Montana State University.) We estimate the motion (i.e., translation and rotation) of the two ice floes, which are labeled A and B in Figure 1 and whose new positions in Figure 2 are A* and B*. Ultimately, scientists are interested in estimating the field of motion of the Arctic icepack. An integral part of a tracking algorithm is identification of matching floes in two images, and we are using the error estimation developed here to determine the search space of a candidate ice floe. We will report on these results in a future work.

We chose as prominent points the approximate locations of the critical points of the curvature of the boundary. Thus, although a large number of data points are available along the boundary of each ice floe, it is possible to match only few points between the floes A and A* and B and B*. In our analysis we obtained six matched points between A and A* and four matched points between B and B*. The data are provided in Table 3.

To test the null hypothesis H_0: $(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i) = (\boldsymbol{\alpha}, \boldsymbol{\beta})$, we use the statistic

$$F = \left\{ \frac{2(n_1 + n_2 - 3)}{3} \right\} \frac{\mathrm{SSE}_{H_0} - \mathrm{SSE}}{\mathrm{SSE}}, \tag{27}$$

which is asymptotically distributed as an F variable with $(3, 2(n_1 + n_2 - 3))$ degrees of freedom. If the null hypothesis is not rejected, then we pool the data from the two ice floes and construct a simultaneous confidence region for the translation and the rotation parameters using the pivot given in (5).

Because the sample sizes are small, we use bootstrap methodology to perform the inference. The test of hypothesis H_0 is described in an analysis of variance (ANOVA) table, Table 4.

The null hypothesis is not rejected, and consequently, we pool the data from the two ice floes to estimate the joint rotation and translation. We now recenter the data as in (3). The OLSE of $\boldsymbol{\alpha}$ is obtained to be $\hat{\boldsymbol{\alpha}} = (35, 105.30)'$; the OLSE of the angle of rotation θ is $\hat{\theta} = -1.11$ radians. Note that in two-dimensional motion, the rotation matrix $\boldsymbol{\beta}$ is completely determined by the angle of rotation θ. Thus we write the motion parameters in terms of $(\boldsymbol{\alpha}', \theta)$.

The $(1 - \boldsymbol{\alpha})100\%$ bootstrap confidence region for $(\boldsymbol{\alpha}', \theta)$ is given by

$$C_{\mathrm{boot}} = \left\{ (\boldsymbol{\alpha}', \theta) \in \mathcal{R}^3 \left| \frac{(\hat{\boldsymbol{\alpha}}' - \boldsymbol{\alpha}', \hat{\theta} - \theta)\hat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\alpha}}' - \boldsymbol{\alpha}', \hat{\theta} - \theta)'}{s^2} \right. \right. \leq c_{\boldsymbol{\alpha}} \left. \right\}, \tag{28}$$

where

$$\frac{\hat{\boldsymbol{\Omega}}}{s^2} = \begin{pmatrix} 3.04 & 0 & 0 \\ 0 & 3.04 & 0 \\ 0 & 0 & 1{,}038.44 \end{pmatrix}$$

Table 4. Analysis of Variance Table

Source of variation	Sum of squares (SS)	F ratio	Bootstrap critical value, $\boldsymbol{\alpha} = .05$
Between floes	$\mathrm{SSE}(H_0) - \mathrm{SSE} = 11.24$	$\frac{14}{3} \frac{\mathrm{SSE}_{H_0} - \mathrm{SSE}}{\mathrm{SSE}} = 1.17$	3.32
Within floes	$\mathrm{SSE} = 44.68$		

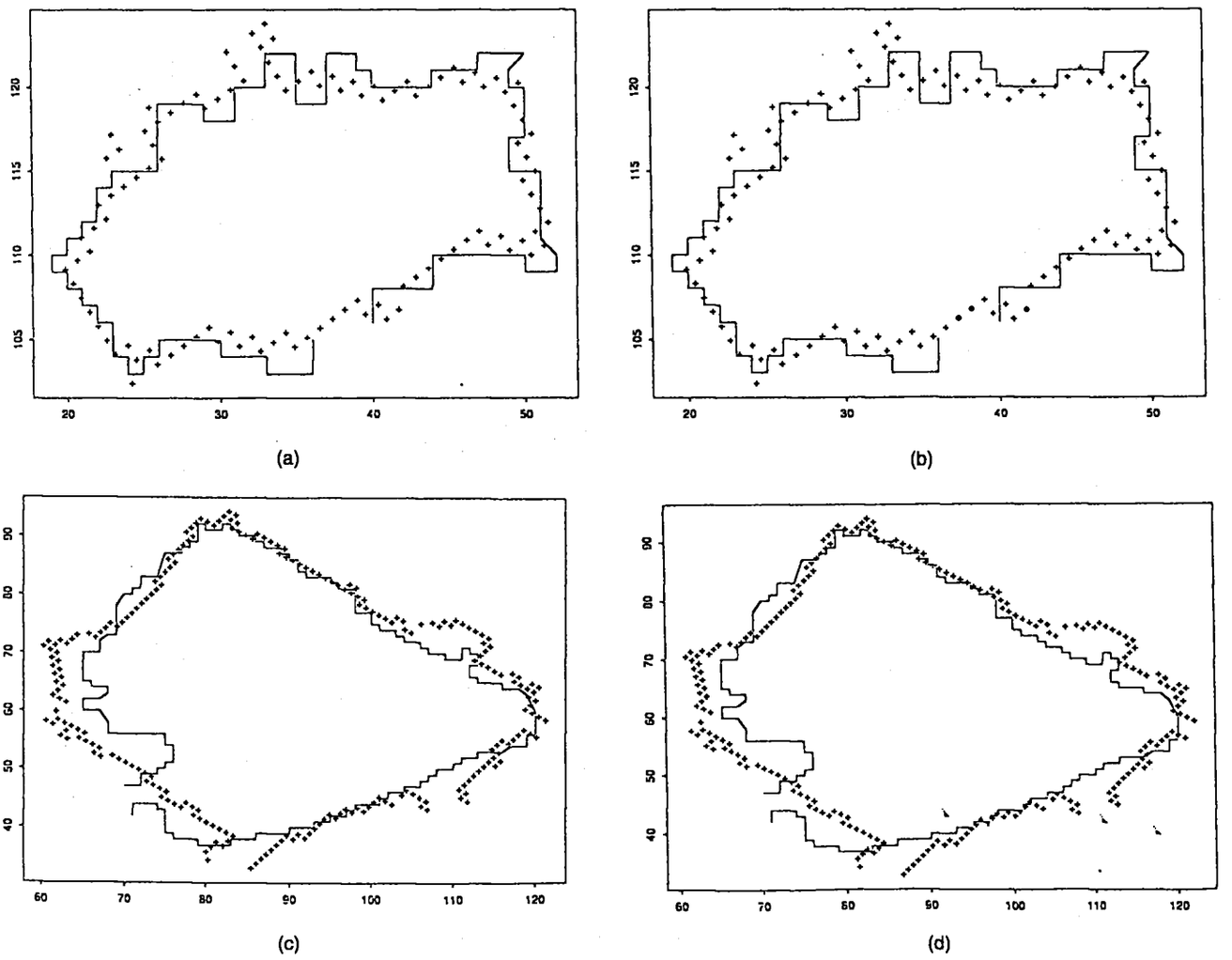

Figure 3. Nonhomologous Regression. Predicted position of floes A and C superimposed on their respective true positions A and C*. Solid line denotes the true position of an ice floe and + denotes the predicted position of an ice floe, and $\hat{\delta} = (\hat{\alpha}_1, \hat{\alpha}_2, \hat{\theta})'$ denotes estimated value of the motion parameters. (a) fit of A and A* under l_1 norm, $\hat{\delta} = (35.46, 113.37, -.99)'$; (b) fit of A and A* under l_2 norm, $\hat{\delta} = (35.50, 113.35, -.99)'$; (c) fit of C and C* under l_1 norm, $\hat{\delta} = (89.92, 63.20, -.65)'$; (d) fit of C and C* under l_2 norm, $\hat{\delta} = (89.81, 63.26, -.63)'$.*

and c_α is obtained from the bootstrap distribution of the pivot (5) such that

$$P\left\{\frac{(\hat{\boldsymbol{\alpha}}^{*\prime} - \hat{\boldsymbol{\alpha}}', \hat{\theta}^* - \hat{\theta})\hat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\alpha}}' - \boldsymbol{\alpha}', \hat{\theta} - \theta)'}{s^{*2}} \le c_\alpha | (\hat{\boldsymbol{\alpha}}', \hat{\theta})\right\} = 1 - \boldsymbol{\alpha}. \quad (29)$$

Table 5. Coverage Probability of Bootstrap Confidence Region

	Distribution of ϵ			
Nominal value	*Normal*	t_5	t_{10}	t_{20}
.20	.19	.20	.20	.20
.50	.50	.53	.52	.52
.70	.70	.69	.69	.69
.90	.90	.89	.90	.89
.95	.95	.94	.94	.95
.99	.99	.99	.99	.99

For the data in Table 3, we found that for a 95% confidence coefficient, $c_{.05} = 3.13$. These two bootstrap estimates were constructed using 1,000 resamples.

We conducted a Monte Carlo study of the coverage probability of the pivot (5) to gauge the accuracy of the bootstrap approximation. Corresponding to each of the four $\mathbf{x}$ values of the ice flow B given in Table 3, we constructed four $\mathbf{y}$ values by generating random error vectors $\boldsymbol{\epsilon}$ from bivariate t_n distribution with n degrees of freedom and from standard normal distribution. Using 1,000 Monte Carlo samples and 1,000 bootstrap resamples per Monte Carlo sample, we estimated the coverage probability of the confidence region (13). Results are reported in Table 5.

3. NONHOMOLOGOUS DATA

3.1 Point Estimation

In both the tectonic and sea ice contexts, the data as actually collected are not homologous. Models for nonhomologous data are necessarily complex, and rigorously proved theorems for them are much harder to establish. Thus we

use the better-established theory for homologous models to guide the approach to nonhomologous data. For example, the homologous theory yields theorems about the relationship of the shape of the boundary to asymptotics of least squares and other M estimators, the speed of convergence of the asymptotics, and the influence characteristics of the data points. In the tectonic contexts, some of these theorems have been heuristically observed for the actual nonhomologous data (Chang 1993).

An attractive theory would hypothesize an unknown boundary $\mathcal{C}$ and unknown points P_i and Q_j on $\mathcal{C}$ and $f(\mathcal{C})$, where f is the unknown rigid transformation. One would then model the distribution of $\mathbf{x}_i$ and $\mathbf{y}_j$ to be centered at P_i and Q_j, respectively. In a model of this type, P_i and Q_j become unknown parameters.

We believe that it is not profitable to model a distribution for the P_i and Q_j. Consider, for example, Figure 3. Although this image is a plot of the points $\mathbf{x}_i$ and $\mathbf{y}_j$, it would appear that the P_i and $f^{-1}(Q_j)$ are evenly distributed around $\mathcal{C}$. But if P_i and $f^{-1}(Q_j)$ were independently uniformly distributed in some probabilistic sense around $\mathcal{C}$, then one would observe that they would tend to coagulate in some parts of $\mathcal{C}$ and be relatively sparse in others. As noted by Feller (1968, p. 161), "to the untrained eye randomness appears as regularity or tendency to cluster." We believe that the even distribution of the points P_i and $f^{-1}(Q_j)$ is due to the mechanics of Banfield and Raftery algorithm (Banfield and Raftery 1992) that was used to identify the floe boundaries from gray-scale images.

In the tectonic context, the shape of the boundary $\mathcal{C}$ can be expressed using a finite number of parameters, and a rigorous analysis can be developed (Chang 1988) that estimates $\mathcal{C}$ and the unknown rotation simultaneously. Essentially, the parameters that describe $\mathcal{C}$, P_i, and Q_j are treated as nuisance parameters. In addition, only the distributions of the normal components of $\mathbf{x}_i$ and $\mathbf{y}_i$ to the boundary $\mathcal{C}$ become relevant, and for this reason, and due to the simple form of $\mathcal{C}$, it is not necessary to explicitly estimate the location of the points P_i and Q_j along $\mathcal{C}$.

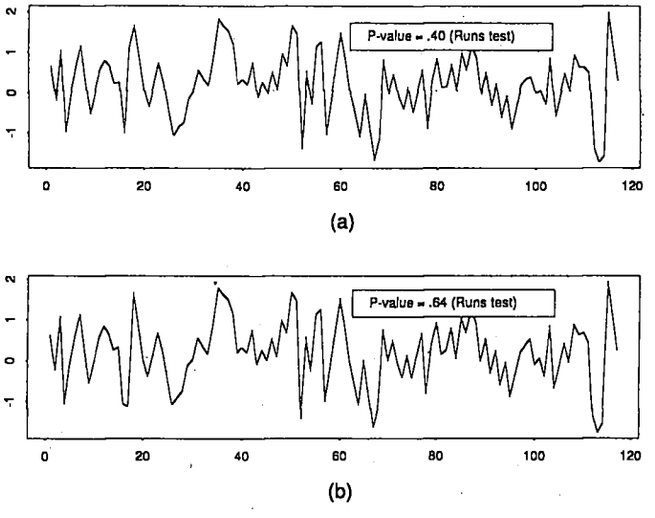

Figure 4. (a): Residual Plots for Floe Pair A and A. (a) L1 norm: p value = .40 (runs test); (b) L2 norm: p value = .64 (runs test).*

In the sea ice context, the shape of $\mathcal{C}$ is sufficiently complicated that parameterization in a reasonable number of parameters seems infeasible. For this reason, we replace the attractive model by a model that assumes that the $\mathbf{x}_i$ are not random and that defines $\mathcal{C}$ to the boundary curve through them. In addition, because the $\mathbf{x}_i$ are so numerous, we approximate $\mathcal{C}$ by the polygon through the $\mathbf{x}_i$. As in the context of tectonic plates, we model only the distribution of the normal components of the $\mathbf{y}_j$'s to the rotated and translated $\mathcal{C}$.

Let $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ denote the estimates of translation and rotation parameters, and let $\hat{\boldsymbol{\xi}}_j$ denote the projection of $\mathbf{y}_j$ onto the polygon formed by the vertices $\hat{\boldsymbol{\alpha}} + \hat{\boldsymbol{\beta}}(\mathbf{x}_1 - \bar{\mathbf{x}}), \ldots, \hat{\boldsymbol{\alpha}} + \hat{\boldsymbol{\beta}}(\mathbf{x}_m - \bar{\mathbf{x}})$. Let $\mathbf{e}_j = \mathbf{y}_j - \hat{\boldsymbol{\xi}}_j$ denote the jth residual vector, and let $r_j = \operatorname{sgn}(e_j)\|e_j\|_2$, where $\operatorname{sgn}(e_j) = +1$ if the normal vector from $\mathbf{y}_j$ onto the polygon is pointing outwardly toward the polygon and -1 otherwise.

During the course of motion of an ice floe, new boundaries are created due to melting or freezing of water around the ice floe. Consequently, it seems reasonable to believe that the residuals at neighboring points on an ice floe are correlated and that the correlation coefficient between the residuals at two points decreases as the distance between the points increases. In the case of homologous data, this was unnecessary, because generally the features that can be matched are so dispersed that any correlation is negligible.

Thus for $j = 1, 2, \ldots, n$, we propose the model

$$r_j = s_j + \sum_{i=1}^{(n+1)/2} \rho^i (s_{j+i} + s_{j-i})/2, \tag{30}$$

where $s_1, s_2, \ldots, s_n$ denote n univariate iid random variables with a common variance σ^2. Because the data are on a closed curve, we adopt the convention that $s_{-j} \equiv s_{n-j}$. In matrix notation, we have $\mathbf{r} = \mathbf{A}\mathbf{s}$, where $\mathbf{A}$ is a suitable circulant symmetric matrix.

We obtain $\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\xi}}_j$, and $\hat{\rho}$ by iteratively, solving two minimization problems using a Gauss–Seidel type of approach (Thisted 1988, p. 187). The first problem,

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \|\mathbf{A}^{-1}\hat{\mathbf{r}}\|, \tag{31}$$

fixes ρ. Here $\|\cdot\|$ denotes norm of a vector, and for a given $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}, \hat{\boldsymbol{\xi}}_j, j = 1, \ldots, n$ is the orthogonal projection of $\mathbf{y}_j$ onto the polygon formed by the vertices $\boldsymbol{\alpha} + \boldsymbol{\beta}(\mathbf{x}_1 - \bar{\mathbf{x}}), \ldots, \boldsymbol{\alpha} + \boldsymbol{\beta}(\mathbf{x}_m - \bar{\mathbf{x}})$; $\hat{\boldsymbol{\eta}}_j$ is the outward normal to the polygon; and $\hat{r}_j\hat{\boldsymbol{\eta}}_j = \mathbf{y}_j - \hat{\boldsymbol{\xi}}_j$. The second minimization,

$$\min_{\rho} \|\mathbf{A}^{-1}\hat{\mathbf{r}}\|, \tag{32}$$

fixes $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

The two norms considered in this article are the l_1 and l_2 norms. If a floe boundary undergoes substantial changes in shape due to erosion or melting or freezing, then it may be preferable to use the estimates based on l_1 norm rather than those based on the l_2 norm. Also, in such cases, one would expect ρ to be large.

Our estimation procedure is similar to iterated weighted least squares estimation. Starting with an initial value for ρ (e.g., taking $\rho = 0$), estimate $\boldsymbol{\alpha}, \boldsymbol{\beta}$, and $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_n$ by

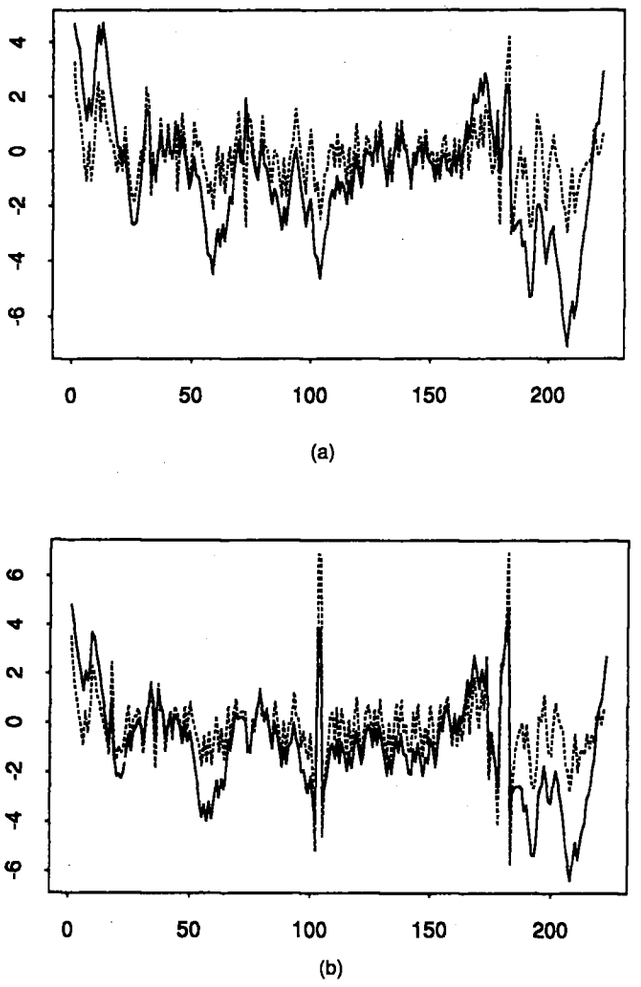

Figure 5. Residual Plots for Floe Pair C and C. Solid lines denote correlated residuals and dotted lines denote uncorrelated residuals. p-values are computed using the runs test. (a) under l_1 norm, p-value for the correlated residuals is < .0001 and the p-value for the uncorrelated residuals is .20; (b) under l_2 norm, p-value for the correlated residuals is < .00001 and the p-value for the uncorrelated residuals is .07).*

iteratively solving (31). Using these estimated values, we then estimate ρ (denoted by $\hat{\rho}$) by solving the minimization problem (32). Using this $\hat{\rho}$, we then reestimate $\boldsymbol{\alpha}, \beta$ and $\xi_1, \xi_2, \ldots, \xi_n$ by solving the problem (31) iteratively. This iterative process is continued until we converge to a solution.

Suitability of the correlation model (30) was examined using the ice floe pairs (A, A*) and (C, C*). Residual plot (Fig. 4), plot of $\hat{r}_i$ versus i, corresponding to floe pair (A, A*) suggests that the independence model (i.e., $\rho = 0$) describes the data very well. Note that to a very great extent,

Table 6. Estimates of $\delta' = (\alpha_1, \alpha_2, \theta)$

Floe pair	Estimates based on l_1 norm	Estimates based on l_2 norm
(A, A*)	(35.46, 113.37, −.99)	(35.50, 113.35, −.99)
(B, B*)	(53.35, 95.20, −1.03)	(51.72, 94.83, −1.08)
(C, C*)	(89.20, 63.20, −.65)	(89.81, 63.26, −.63)
(D, D*)	(147.38, 34.71, −1.01)	(147.72, 35.23, −.97)

Table 7. Bias (MSE)

Sample size	Parameter	Based on l_1 norm	Based on l_2 norm
21	$\boldsymbol{\alpha}$	.0009 (.5181)	.0039 (.6494)
21	θ	.0000 (.0002)	.0000 (.0002)
31	$\boldsymbol{\alpha}$	.0016 (.2032)	.0044 (.3817)
31	θ	.0000 (.0001)	.0000 (.0001)

the shape of ice floes A and A* are similar. On the other hand, the plot of $\hat{r}_i$ versus i corresponding to flow pair (C, C*) (solid lines in Fig. 5) suggests that $\rho > 0$. Taking $\rho = .80$ and plotting the corrected residuals $\hat{s}_i$ versus i (dotted lines in Fig. 5), we find the proposed correlation model (30) to be reasonable. In contrast to flow pair (A, A*), it seems that some amount of erosion has occurred in the southwest boundary of ice floe C as it moved from Figure 1 to Figure 2. Consequently, one would expect ρ to be large in this case.

Though the proposed correlation model (30) appears to be reasonable in a limited data set that we have considered, it is only ad hoc, and it will be worthwhile to consider other models. We invite the reader to explore other possible models.

To illustrate how well the proposed estimation procedure performs, we fitted the motion of the floes A, B, C, and D in Figure 1 against their counterparts A*, B*, C*, and D*∗ in Figure 2. The point estimates, under the two norms, of the angle of rotation (θ) and translational parameter vector $(\alpha_1, \alpha_2)'$ are given in Table 6. Using these point estimates, we plotted the predicted position of floes A and C on the top of the true positions A* and C* (Figure 3). We found equally good fits for other floes pairs as well.

Monte Carlo simulations were performed to estimate the squared bias, $(E(\hat{\boldsymbol{\alpha}}) - \boldsymbol{\alpha})'(E(\hat{\boldsymbol{\alpha}}) - \boldsymbol{\alpha})$ and $(E(\hat{\theta}) - \theta)^2$, and the mean squared error (MSE), $E(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})'(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})$ and $E(\hat{\theta} - \theta)^2$, of the proposed estimation procedure, under both l_1 and l_2 norms, where $\hat{\boldsymbol{\alpha}}$ is an estimator of the translation parameter $\boldsymbol{\alpha}$ and $\hat{\theta}$ is an estimator of the angle of rotation. To perform this simulation study, one needs to generate nonhomologous data $\mathbf{x}_1, \ldots, \mathbf{x}_m$ and $\mathbf{y}_1, \ldots, \mathbf{y}_n$, that describe the boundary of an ice floe at two different times. Because the proposed correlation model (30) seems appropriate, we use it in the generation of nonhomologous data sets as described later. For our simulation purposes, with out loss of generality, we generate nonhomologous data using ice floe A.

Let $\mathbf{x}_i, i = 1, 2, \ldots, m$, and $\mathbf{z}_j, j = 1, 2, \ldots, n$, be two sets of random samples of points from the boundary of the ice floe A, Let $\mathbf{u}_j$, denote the orthogonal projection of $\mathbf{z}_j$ onto the polygon formed by $\mathbf{x}_1 - \bar{\mathbf{x}}, \ldots, \mathbf{x}_m - \bar{\mathbf{x}}$, and let $\boldsymbol{\eta}_j$ denote the corresponding normal direction vector. Taking $\rho = .2$ and a random sample $s_j, j = 1, \ldots, n$, from $N(0, 1)$, we obtain correlated random variables $r_1, \ldots, r_n$ using the model (30). The $\mathbf{y}_j$'s are then obtained as

$$\mathbf{y}_j = \mathbf{u}_j + r_j \boldsymbol{\eta}_j. \tag{33}$$

The bias and the MSE's reported in Table 7 are based on 1,000 simulation runs, where in each run $\mathbf{y}_j$'s, $j = 1, 2, \ldots, n$, were generated according to the foregoing description with $m = n$.

Table 8. Bootstrap Percentiles

Nominal value	Percentiles based on l_1 norm	Percentiles based on l_2 norm
.50	.014	.858
.90	.017	1.091
.95	.019	1.167
.99	.022	1.286

3.2 Confidence Region Estimation

To construct the simultaneous confidence region for $\boldsymbol{\delta} = (\alpha_1, \alpha_2, \theta)'$, we use the pivotal statistic given in Section 2. In the case of nonhomologous data, the asymptotic distribution of this statistic is unknown. Thus we use bootstrap methodology.

Corresponding to each $\mathbf{y}_i$ and $\hat{\boldsymbol{\xi}}_i, i = 1, 2, \ldots, n$, let $\boldsymbol{\eta}_i$ denote a unit normal vector, pointing outwardly to the polygon formed by the points $\hat{\boldsymbol{\alpha}} + \hat{\boldsymbol{\beta}}(\hat{\mathbf{x}}_i - \bar{\mathbf{x}})$'s. Let $\hat{r}_1, \ldots, \hat{r}_n$ denote the estimated signed residuals obtained by solving the minimization problems (31) and (32), and let $\hat{s}_1, \ldots, \hat{s}_n$ be the "decorrelated" residuals, obtained by solving the system of equations (30) with ρ replaced by $\hat{\rho}$. Let $s_1^*, s_2^*, \ldots, s_n^*$ be a simple random sample from the empirical cdf of $\hat{s}_1, \ldots, \hat{s}_n$, and let $r_1^*, r_2^*, \ldots, r_n^*$ be obtained according to the correlation model (30), with s_i's replaced by s_i^*'s and ρ replaced by $\hat{\rho}$. The bootstrap sample $\mathbf{y}_j^*, j = 1, 2, \ldots, n$, is then given by

$$\mathbf{y}_j^* = \hat{\boldsymbol{\xi}}_j + r_j^* \boldsymbol{\eta}_j. \tag{34}$$

Corresponding to a given bootstrap sample, the bootstrap estimates $\hat{\boldsymbol{\alpha}}^*$ and $\hat{\boldsymbol{\beta}}^*$ are obtained by solving the minimization problems (31) and (32), with $\mathbf{y}_i$'s replaced by $\mathbf{y}_i^*$'s, $\hat{r}_i$'s replaced by $\hat{r}_i^*$'s, and $\hat{\rho}$ replaced by $\hat{\rho}^*$. The bootstrap pivotal statistic is obtained accordingly.

As an example, we obtained the bootstrap percentiles for the pivotal statistic for fitting the motion of ice floe A in Figure 1 against its counterpart, ice floe A*, in Figure 2. The percentiles, given in Table 8, were obtained using 300 bootstrap samples. Because the actual number of points on ice floe A is 112 and on ice floe A* is 117, we believe that 300 bootstrap samples are sufficiently large.

To study the performance of the proposed bootstrap confidence interval, a Monte Carlo simulation of the coverage probability was carried out using ice floe A. For simulation purposes, the nonhomologous data were generated from ice floe A according to the description given in the previous section with $m = 28$ and $n = 29$. Corresponding to the fixed $\mathbf{x}_i$'s, $i = 1, 2, \ldots, 28$, we generated 300 sets of $\mathbf{y}_j$'s, $j = 1, 2, \ldots, 29$, according to the model (33). For each simulation run (i.e., for each set of nonhomologous data $\mathbf{x}_1, \ldots, \mathbf{x}_{28}$ and $\mathbf{y}_1, \ldots, \mathbf{y}_{29}$), we generated 200 bootstrap samples according to the model (34). The estimated coverage probabilities are as given in Table 9. The coverage probabilities were estimated using both the l_1 and the l_2 norms. The results seem to indicate that the proposed bootstrap confidence intervals are indeed reasonable for upper tail probabilities.

Table 9. Estimated Coverage Probability (and Standard Error) of Bootstrap Confidence Regions

Nominal value	Point estimates based on l_1 norm	Point estimates based on l_2 norm
.20	.135 (.023)	.230 (.023)
.50	.445 (.029)	.505 (.029)
.70	.675 (.026)	.720 (.026)
.90	.885 (.017)	.915 (.017)
.95	.965 (.0126)	.970 (.0126)
.99	1.00 (.006)	.995 (.006)

4. CONCLUSIONS

Two types of rigid body motion were considered in this article: motion in a p-dimensional Euclidean space and motion in a p-dimensional sphere. A data set is said to be *homologous* if there exists a one-to-one correspondence between the observed points at two different times on the same rigid body. For when the data are homologous, we have developed a bootstrap methodology to construct confidence regions for the motion parameters. But in practice, one encounters *nonhomologous* data. To deal with such data sets, we have developed a new method of estimating motion parameters. This generalizes the principle of least squares, where along with the motion parameters, we also estimated the points on the boundary of a body at an earlier time that correspond to the observed points at a later time. Sometimes when a body is in motion, it undergoes some changes in shape. For instance, during the course of an ice floe's motion in the Arctic sea, it may melt or the water surrounding it may freeze. To deal with such changes in the shape, we modeled the data using a correlation model that seems to fit our data very well. For the nonhomologous data, under the proposed correlation model, we obtained the point estimates of motion along with the bootstrap confidence region for the motion parameters. We illustrated the methodology developed in this article with real data.

APPENDIX

A.1 Euclidean Space Data

Suppose that the errors $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, with mean zero and covariance matrix proportional to $\sigma^2\mathbf{I}$, are identically and independently spherically symmetrically distributed. Let $\hat{\varepsilon}_i$, denote the ith residual, $i = 1, 2, \ldots n$. Then we note from Seber and Wild (1989, p. 24), that

$$\hat{\varepsilon} \approx (\mathbf{I} - \mathbf{P})\varepsilon, \tag{A.1}$$

where $\varepsilon = [\varepsilon_1' : \varepsilon_2' : \ldots : \varepsilon_n']', \hat{\varepsilon} = [\hat{\varepsilon}_1' : \hat{\varepsilon}_2' : \ldots : \hat{\varepsilon}_n']', \mathbf{P} = \mathbf{D}(\mathbf{D}'\mathbf{D})\mathbf{D}'$, and $\mathbf{D}$ is the matrix $\mathbf{F}$ given by Seber and Wild (1989). Using (35) and following the arguments of Freedman (1981), we can prove that the empirical cdf of $\{\hat{\varepsilon}_1, \hat{\varepsilon}_2, \ldots, \hat{\varepsilon}_n\}$ converges weakly to the common cdf of $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$. Consequently, using lemmas 8.3 and 8.5 of Bickel and Freedman (1981), we deduce that $\hat{F}_n$ weakly converges to F, the common cdf of the iid random variables $\|\varepsilon_1\|_2, \|\varepsilon_2\|_2, \ldots, \|\varepsilon_n\|_2$. Now using the notations given in Section 2.2 and performing calculations similar to those of Seber (1989, p. 23), we can write

$$\hat{\boldsymbol{\mu}}^*(n, m) \approx m^{-1}\boldsymbol{\Omega}^{-1}\mathbf{D}'\boldsymbol{\varepsilon}^*(n, m). \tag{A.2}$$

Noting that $\varepsilon_i^*(n)$ are iid and weakly converge to the common cdf of $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, and repeating the arguments given in theorem 2.2 of Freedman (1981), we deduce that $\sqrt{m}\hat{\boldsymbol{\mu}}^*(n, m) \sim^{\text{asymptotically}}$

$N(\mathbf{0}, \sigma^2\mathbf{\Omega})$. Further,

$$\frac{m\hat{\boldsymbol{\mu}}^*(n,m)'\mathbf{\Omega}\hat{\boldsymbol{\mu}}^*(n,m)}{\sigma^2} \sim \chi^2 \quad \text{with } \nu_1 \text{ degrees of freedom.} \tag{A.3}$$

The result then follows.

A.2 Spherical Data

In the case of spherical regression, the validity of bootstrap methodology is verified along the lines of the proofs developed by Chang (1986). Denote $\mathbf{W}(n) = 1/n \sum_{i=1}^{n} \mathbf{x}_i\mathbf{y}_i'$ and $\mathbf{W}^*(n,m) = 1/m \sum_{j=1}^{m} \mathbf{x}_j\mathbf{y}_j^{*\prime}$. It follows from the Kolomogorov inequality that

$$P(\|\mathbf{W}^*(n,m) - \hat{c}_0(n)\mathbf{\Sigma}_m\hat{\boldsymbol{\beta}}(n)'\| > \epsilon|\hat{F}_n) \le \frac{k}{m\epsilon^2}, \tag{A.4}$$

where k is a constant not involving n. Thus, conditional on $\hat{F}_n$, as $m, n \to \infty$, $\mathbf{W}^*(n,m) - \hat{c}_0(n)\mathbf{\Sigma}_m\hat{\boldsymbol{\beta}}(n)' \to_p \mathbf{0}$. Consequently, conditional on $\hat{F}_n$, $\mathbf{W}^*(n,m) - \hat{c}_0(n)\mathbf{\Sigma}\hat{\boldsymbol{\beta}}(n)' \to_p \mathbf{0}$. Because the right side of (A.4) and $\|\mathbf{\Sigma}_m - \mathbf{\Sigma}\|$ do not depend on n, the foregoing convergence is true even unconditionally. Also, we know that $\hat{c}_0(n)\mathbf{\Sigma}\hat{\boldsymbol{\beta}}(n)' - c_0\mathbf{\Sigma}\boldsymbol{\beta}' \to_p 0$. Thus unconditionally, $\mathbf{W}^*(n,m) - c_0\mathbf{\Sigma}\boldsymbol{\beta}' \to_p 0$.

It follows from lemma 1 of Chang (1988) that as $(n,m) \to \infty, \hat{\boldsymbol{\beta}}(n,m)^* \to_p \boldsymbol{\beta}$ unconditionally and that conditional on $\hat{F}_n, \hat{\boldsymbol{\beta}}(n,m)^* \to_p \hat{\boldsymbol{\beta}}(n)$. Following Chang (1986), we can perform Taylor series expansion to get

$$-\mathrm{Tr}(\sqrt{m}\hat{\boldsymbol{\beta}}(n)\mathbf{B}\mathbf{W}^*(n,m)) = \hat{c}_0(n)\mathrm{Tr}(\sqrt{m}\mathbf{H}(n,m)^*\mathbf{\Sigma}\mathbf{B}) + R_m(n), \tag{A.5}$$

where $\mathbf{B}$ is a skew symmetric matrix and $R_m(n)$ is a reminder term satisfying

$$|R_m(n)| \le \sqrt{m}\|\mathbf{H}(n,m)^*\|\|\mathbf{B}\|\{\|\mathbf{H}(n,m)^*\| \exp(\|\mathbf{H}(n,m)^*\|) + \|\mathbf{W}^*(n,m) - \hat{c}_0(n)\mathbf{\Sigma}\hat{\boldsymbol{\beta}}(n)'\|\} \tag{A.6}$$

Noting from Chang (1986) that $\hat{\boldsymbol{\beta}}(n), \hat{c}_0(n)$, and $\hat{c}_2(n)$ are strongly consistent for $\boldsymbol{\beta}, c_0$, and c_2, and following calculations similar to those of Chang (1986), we conclude that asymptotically as $(n,m) \to \infty$, the distribution of $\mathbf{H}(n,m)^*$, conditional on $\hat{F}_n$, is normal with mean zero and density function proportional to $\exp(nc_0^2/2c_2)\ \mathrm{Tr}(\mathbf{H}^*\mathbf{\Sigma}\mathbf{H}^*)$. Thus the bootstrap methodology is valid.

[Received July 1993. Revised July 1995.]

REFERENCES

Banfield, J., and Raftery, A. E. (1992), "Ice Floe Identification in Satellite Images Using Mathematical Morphology and Clustering About Principal Curves," *Journal of the American Statistical Association*, 87, 7–16.

Bingham, C., Chang, T., and Richards, D. (1992), "Approximating the Matrix Fisher and Bingham Distributions: Applications to Spherical Regression and Procrustes Analysis," *Journal of Multivariate Analysis*, 41, 314–337.

Bickel, P. J., and Freedman, D. A. (1981), "Some Asymptotic Theory for the Bootstrap," *The Annals of Statistics*, 9, 1196–1217.

Chang, T. (1986), "Spherical Regression," *The Annals of Statistics*, 14, 907–924.

——— (1988), "Estimating the Relative Rotation of Two Tectonic Plates From Boundary Crossings," *Journal of the American Statistical Association*, 83, 1178–1183.

——— (1993), "Spherical Regression and the Statistics of Tectonic Plate Reconstructions," *International Statistical Reviews*, 61, 299–316.

Chang, T., and Ko, D. (1995), "*M* Estimates of Rigid Body Motion on the Sphere and in the Euclidean Space," *The Annals of Statistics*, to appear.

Feller, W. (1966), *An Introduction to Probability Theory and Its Applications*, Vol. 1, New York: John Wiley.

Fisher, N. I., and Hall, P. (1989), "Bootstrap Confidence Regions for Directional Data," *Journal of the American Statistical Association*, 84, 996–1002.

Freedman, D. A. (1981), "Bootstrapping Regression Models," *The Annals of Statistics*, 9, 1218–1228.

Goodall, C. (1991), "Procrustes Methods in the Statistical Analysis of Shape" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 53, 285–340.

Muirhead, R. J. (1982), *Aspects of Multivariate Statistical Theory*, New York: John Wiley.

Rivest, L. (1989), "Spherical Regression for Concentrated Fisher–von Mises Distributions," *The Annals of Statistics*, 17, 307–317.

Seber, G. A. F., and Wild, C. J. (1989), *Nonlinear Regression*, New York: John Wiley.

Thisted, R. A. (1988), *Elements of Statistical Computing*, New York: Chapman and Hall.