ELSEVIER

Multivariate Alteration Detection (MAD) and MAF Postprocessing in Multispectral, Bitemporal Image Data: New Approaches to Change Detection Studies

Allan A. Nielsen, Knut Conradsen,* and James J. Simpson†*

This article introduces the multivariate alteration detection (MAD) transformation which is based on the established canonical correlations analysis. It also proposes using postprocessing of the change detected by the MAD variates using maximum autocorrelation factor (MAF) analysis. The MAD and the combined MAF/MAD transformations are invariant to linear scaling. Therefore, they are insensitive, for example, to differences in gain settings in a measuring device, or to linear radiometric and atmospheric correction schemes. Other multivariate change detection schemes described are principal component type analyses of simple difference images. Case studies with AHVRR and Landsat MSS data using simple linear stretching and masking of the change images show the usefulness of the new MAD and MAF/MAD change detection schemes. Ground truth observations confirm the detected changes. A simple simulation of a no-change situation shows the accuracy of the MAD and MAF/MAD transformations compared to principal components based methods. ©*Elsevier Science Inc., 1998*

INTRODUCTION

The biosphere–atmosphere–hydrosphere form a complex, highly interactive system in which change [e.g., urbanization (Graetz et al., 1992) and ENSO-related (El Niño–Southern Oscillation) change (Quiroz, 1983)] can occur over a broad spectrum of space and time scales. On the regional and global scale, multivariate satellite data provide the only practical way to monitor such changes. Often such large-scale environmental changes have significant social and economic impact. For example, there was a 40% reduction in fish catch within FAO Fishing Region 77 (5°S to 40°N; 80°W to 175°W) during the 1982–1983 ENSO and contemporaneous midlatitude oceanic coastal warming event (e.g., FAO, 1987; Simpson, 1992a), and simultaneously the Australian continent experienced severe drought. Thus, the scientific and socioeconomic needs for accurate change detection in sequences of satellite data are apparent.

When analyzing changes in panchromatic images taken at different points in time, it is customary to analyze the difference between two images, possibly after some normalization; areas with little or no change have zero or low absolute values, and areas with large changes have large absolute values in the difference image. Given two multivariate images with variables at a given location written as vectors (without loss of generality we assume that $E\{\mathbf{X}\}=E\{\mathbf{Y}\}=0$)

$$\mathbf{X}=[X_1 \cdots X_k]^T \quad \text{and} \quad \mathbf{Y}=[Y_1 \cdots Y_k]^T, \qquad (1)$$

where k is the number of spectral bands, then a simple change detection transformation is

$$\mathbf{X}-\mathbf{Y}=[X_1-Y_1 \cdots X_k-Y_k]^T. \qquad (2)$$

If our image data have more than three channels, it is difficult to visualize changes in all channels simultaneously. To overcome this problem and to concentrate information on change, linear transformations of the image data that optimize some design criterion can be considered. A linear transformation that will maximize a mea-

* Department of Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark

† Digital Image Analysis Laboratory, Scripps Institution of Oceanography, La Jolla, California

Address correspondence to A. A. Nielsen, Dept. of Mathematical Modelling, Technical Univ. of Denmark, Bldg. 321, DK-2800, Denmark.

Received 30 June 1995; revised 12 September 1997.

REMOTE SENS. ENVIRON. 64:1–19 (1998)

655 Avenue of the Americas, New York, NY 10010

0034-4257/98/$19.00
PII S0034-4257(97)00162-4

sure of change in the simple multispectral difference image is one that maximizes deviations from no change (e.g., the variance)

$$\mathrm{Var}\{v_1(X_1-Y_1)+\cdots+v_k(X_k-Y_k)\}=\mathrm{Var}\{\boldsymbol{v}^T(\boldsymbol{X}-\boldsymbol{Y})\}. \quad (3)$$

Areas in the image data with high absolute values of $\boldsymbol{v}^T(\boldsymbol{X}-\boldsymbol{Y})$ are maximum change areas. A multiplication of vector $\boldsymbol{v}$ with a constant c will multiply the variance with c^2. Therefore, we must make a choice concerning $\boldsymbol{v}$. A natural choice is unit normalization, $\boldsymbol{v}^T\boldsymbol{v}=1$. This amounts to finding principal components (Hotelling, 1933) of the simple difference images.

A more parameter rich measure of change that allows different coefficients for $\boldsymbol{X}$ and $\boldsymbol{Y}$ and different numbers of spectral bands in the two sets p and q, respectively ($p\leq q$), are linear combinations

$$\boldsymbol{a}^T\boldsymbol{X}=a_1X_1+\cdots+a_pX_p \quad \text{and} \quad \boldsymbol{b}^T\boldsymbol{Y}=b_1Y_1+\cdots+b_qY_q \quad (4)$$

and the difference between them $\boldsymbol{a}^T\boldsymbol{X}-\boldsymbol{b}^T\boldsymbol{Y}$. The measure of change we propose is: maximize $\mathrm{Var}\{\boldsymbol{a}^T\boldsymbol{X}-\boldsymbol{b}^T\boldsymbol{Y}\}$ with the chosen constraints $\mathrm{Var}\{\boldsymbol{a}^T\boldsymbol{X}\}=\mathrm{Var}\{\boldsymbol{b}^T\boldsymbol{Y}\}=1$. This formulation can be implemented using standard canonical correlations analysis (Hotelling, 1936). A treatment is given in most textbooks on multivariate statistics [good references are Cooley and Lohnes (1971) and Anderson (1984)]. This measure also accounts for situations where the spectral bands are not the same but cover different spectral regions; for instance, one set of data comes from Landsat Multi-Spectral Scanner (MSS) and the other set comes from Landsat Thematic Mapper (TM) or from SPOT High Resolution Visible (HRV) imager. This latter capability may be valuable in historical change detection studies using time series elements from different instruments. In this case one must be more cautious when interpreting the multivariate difference as multivariate change.

This article uses a new method of change detection (next section) based on canonical correlations analysis and the maximum autocorrelation factor (MAF) analysis (next section) developed by Switzer and Green (1984). Unlike traditional methods of change detection, it rank-orders the component differences in the multivariate data and simultaneously preserves the natural spatial structure of the data. Examples using AVHRR and Landsat MSS data show the usefulness of the new method, which is confirmed by ground truth data. Comparisons with traditional methods of change detection show the robustness of the new method. In fact, in one example standard methods fail.

STATISTICAL PROCEDURES

This section discusses: 1) the multivariate alteration detection (MAD) transformation; 2) the maximum autocorrelation factors (MAF) transformation; and 3) the physical motivation for the MAF/MAD analysis.

The Multivariate Alteration Detection (MAD) Transformation

Canonical correlations analysis investigates the relationship between two groups of variables. It finds two sets of linear combinations of the original variables, one for each group. The first two linear combinations are the ones with the largest correlation. This correlation is called the first canonical correlation, and the two linear combinations are called the first canonical variates. The second two linear combinations are the ones with the largest correlation subject to the condition that they are orthogonal to the first canonical variates. This correlation is called the second canonical correlation, and the two linear combinations are called the second canonical variates. Higher-order canonical correlations and canonical variates are defined similarly. Because corresponding pairs of canonical variates are linear combinations of the original variables ordered by correlation or similarity between pairs [Eq. (4)], it seems natural to base a change detection scheme on differences between these pairs of variates.

To find $\boldsymbol{a}$ and $\boldsymbol{b}$ [Eq. (4)], Fung and LeDrew (1987) use principal components (PC) analysis on $\boldsymbol{X}$ and $\boldsymbol{Y}$ concatenated to one vector variable. They define $\boldsymbol{a}$ and $\boldsymbol{b}$ simultaneously, but their method does not have a clear design criterion [e.g., an equation analogous to Eqs. (1)–(3) or (5)]. Also, bands are treated similarly whether or not they come from different points in time. Gong (1993) applies PC analysis to simple difference images as described above [Eq. (3)]. This approach depends on the scale at which the individual variables are measured (for instance, it depends on gain settings of a measuring device). Also, it forces the two sets of variables to have the same coefficients (with opposite sign), and it does not allow for the case where the two sets of images have different numbers of channels.

A potentially better approach is to define a set of $\boldsymbol{a}$ and $\boldsymbol{b}$ simultaneously in the fashion described below. Again, let us maximize the variance, this time $\mathrm{Var}\{\boldsymbol{a}^T\boldsymbol{X}-\boldsymbol{b}^T\boldsymbol{Y}\}$. We also must make choices concerning $\boldsymbol{a}$ and $\boldsymbol{b}$, and natural choices in this case are requesting unit variance of $\boldsymbol{a}^T\boldsymbol{X}$ and $\boldsymbol{b}^T\boldsymbol{Y}$. The criterion then is: maximize $\mathrm{Var}\{\boldsymbol{a}^T\boldsymbol{X}-\boldsymbol{b}^T\boldsymbol{Y}\}$ subject to the constraints $\mathrm{Var}\{\boldsymbol{a}^T\boldsymbol{X}\}=\mathrm{Var}\{\boldsymbol{b}^T\boldsymbol{Y}\}=1$. Then we have

$$\begin{aligned}\mathrm{Var}\{\boldsymbol{a}^T\boldsymbol{X}-\boldsymbol{b}^T\boldsymbol{Y}\}&=\mathrm{Var}\{\boldsymbol{a}^T\boldsymbol{X}\}+\mathrm{Var}\{\boldsymbol{b}^T\boldsymbol{Y}\}-2\mathrm{Cov}\{\boldsymbol{a}^T\boldsymbol{X},\boldsymbol{b}^T\boldsymbol{Y}\}\\&=2(1-\mathrm{Corr}\{\boldsymbol{a}^T\boldsymbol{X},\boldsymbol{b}^T\boldsymbol{Y}\}).\end{aligned} \quad (5)$$

We shall request that $\boldsymbol{a}^T\boldsymbol{X}$ and $\boldsymbol{b}^T\boldsymbol{Y}$ are positively correlated. Therefore, determining the difference between linear combinations with maximum variance corresponds to determining linear combinations with minimum (nonnegative) correlation; this complies with canonical correlations analysis.

In accordance with the above, we define the multivariate alteration detection (MAD) transformation as

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \rightarrow \begin{bmatrix} \boldsymbol{a}_p^T\mathbf{X} - \boldsymbol{b}_p^T\mathbf{Y} \\ \cdot \\ \cdot \\ \cdot \\ \boldsymbol{a}_1^T\mathbf{X} - \boldsymbol{b}_1^T\mathbf{Y} \end{bmatrix}, \tag{6}$$

where $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$ are the defining coefficients from a standard canonical correlations analysis (see the Appendix). $\mathbf{X}$ and $\mathbf{Y}$ are vectors with expectation values $E\{\mathbf{X}\}=E\{\mathbf{Y}\}=0$. We see that the MAD transform consists of the variates we get when we subtract corresponding canonical variates in reverse order. The dispersion matrix of the MAD variates is

$$D\{\boldsymbol{a}^T\mathbf{X} - \boldsymbol{b}^T\mathbf{Y}\} = 2(\boldsymbol{I} - \boldsymbol{R}), \tag{7}$$

where $\boldsymbol{I}$ is the $p \times p$ unit matrix and $\boldsymbol{R}$ is the $p \times p$ matrix containing the sorted canonical correlations on the diagonal and zeros off the diagonal.

The MAD transformation has the very important property that if we consider linear combinations of two sets of p respectively q ($p \leq q$) variables that are positively correlated then the pth difference shows maximum variance among such variables. The $(p-j)$th difference shows maximum variance subject to the constraint that this difference is uncorrelated with the previous j ones. In this way we may sequentially extract uncorrelated difference images where each new image shows maximum difference (change) under the constraint of being uncorrelated with the previous ones. If $p<q$, then the projection of $\mathbf{Y}$ on the eigenvectors corresponding to the eigenvalues 0 will be independent of $\mathbf{X}$. That part may of course be considered the extreme case of multivariate change detection. As opposed to the principal components, the MAD variates are invariant to linear and affine scaling (see Nielsen and Conradsen, 1997). This means that they are not sensitive, for example, to the offsets or gain settings of a measuring device, or to radiometric and atmospheric correction schemes that show a linear relationship with brightness counts. This type of multivariate change detection technique was first sketched in Conradsen and Nielsen (1991a,b). Multivariate change detection techniques are also described in Hanaizumi and Fujimura (1992), and in Hanaizumi et al. (1994) who work with multiple regression and canonical correlations methods applied to specific change detection. Mathematical details of the MAD transformation are given by Nielsen and Conradsen (1997).

The Maximum Autocorrelation Factor (MAF) Transformation

To find maximum change areas with high spatial autocorrelation a MAF postprocessing of the MAD variates is carried out. The MAF transformation can be considered as a spatial extension of principal components (PC) analysis in which the new variates maximize autocorrelation between neighboring pixels rather than variance (as with PCs). MAF1 is the linear combination of the original variables that maximizes autocorrelation, MAF2 is the linear combination of the original variables that maximizes autocorrelation subject to the condition that it is orthogonal to MAF1. Higher order MAFs are defined similarly. Expressed mathematically, the MAFs are constructed by finding new variables defined by a vector of coefficients $\boldsymbol{c}$, where $\mathrm{Corr}\{\boldsymbol{c}^T\mathbf{Z}(\boldsymbol{x}), \boldsymbol{c}^T\mathbf{Z}(\boldsymbol{x}+\boldsymbol{\Delta})\}$ is maximized, thereby preserving information contained in the spatial structure of the data itself. $\mathbf{Z}(\boldsymbol{x})$ is the variable in question (here $\boldsymbol{a}^T\mathbf{X} - \boldsymbol{b}^T\mathbf{Y}$) at location $\boldsymbol{x}$ and $\boldsymbol{\Delta}=[\Delta_x, \Delta_y]^T$ is a spatial shift. Like the MAD transformation, the MAF transformation is invariant to linear scaling. For more detailed mathematical descriptions of the MAF and related transformations, see Switzer and Green (1984), Conradsen et al. (1985), Green et al. (1988), Ersbøll (1989), Conradsen and Ersbøll (1991), and Nielsen (1994).

Physical Motivation for a MAF/MAD Analysis

Studies of change detection using satellite data are concerned with accurately estimating rates of change of a given quantity of interest. Simple differencing of two coregistered satellite scenes separated in time provides a point by point estimate of the desired rate. A PC transformation of the simple difference image is an improvement over simple differencing because it produces concentrated change information in uncorrelated variables which contain decreasing amounts of variance (and thus change) with decreasing eigenvalues.

Physical and biological phenomena observed in satellite images (e.g., lakes, agricultural fields), however, generally occupy many pixels in a scene. Hence, it is very desirable to analyze rates of change within the spatial context of the image data itself (i.e., there is information in the spatial structure of data that is not obtained simply by analyzing the data with pixel-based operations). The combination of MAD and MAF analyses used in this article provides a statistically rigorous way to retain the spatial context of the data in the analyzed results.

MAD analysis takes the difference between linear combinations of the original data that have maximal correlation. All such differences are orthogonal. MADs, like PCs, however, are performed as point operations. Because the MAFs retain spatial structure in the data (through maximizing the correlation between the original and the shifted data), taking the MAFs of the MADs provides a way to retain the spatial context of the neighborhood pixels in the image in the final change detection analysis. Regions of little or no change in the composite image data will have MADs and MAFs of MADs (herein after referred to as MAF/MADs) with near zero or low absolute value whereas regions of large change will have large absolute value.

Often, multitemporal imagery are concatenated and transformed using the PC transformation to yield a se-

quence of patterns contained in the original image sequence [see Richards (1993) for details]. Empirical orthogonal function (EOF) analysis (Preisendorfer, 1988) extends this concept further by evaluating the time series of temporal amplitudes associated with the PCs. Such methods, however, are not optimized for change detection because: 1) They incorporate no explicit design criterion for detecting change in the state of a natural system [e.g., equations analogous to Eqs. (1)–(3) and (5) are not included in the formulations]; 2) the sequence of PCs can be dominated by a single element in the image sequence which has a disproportionate high variance compared to other elements in the image sequence [see detailed example in Gallaudet and Simpson (1994)]; and 3) such PC-based approaches are sensitive to errors in instrument calibration. The MAD and MAF/MAD analyses have explicit design criteria to detect spatially coherent change [Eq. (5) and the previous subsection] and are not affected by issues 2) and 3) cited above.

SATELLITE DATA AND PREPROCESSING

Data for the terrestrial case study were taken with the Multispectral Scanner System (MSS) aboard the Landsat satellites. The MSS is a four-channel instrument with Channels 1–4 having bandwidths of 0.5–0.6 μm (green), 0.6–0.7 μm (red), 0.7–0.8 μm (red to near-infrared) and 0.8–1.1 μm (near-infrared), respectively. Historically, channel 1 data have been used to map coastal features in sediment laden water, channel 2 data have been used to map roads and urban areas, and data from channels 3 and 4 have been used for vegetation studies and for mapping land/water boundaries (Matra Marconi Space, 1994). The swath width of the MSS is 185 km and the pixel resolution is 80 m. A more detailed description of the MSS sensor is given by Markham (1985) and the references contained therein. Data were chosen for their cloud-free qualities and an interchannel relative calibration was performed by Australia's Commonwealth Scientific and Industrial Research Organization (CSIRO).

AVHRR data consist of five channels having bandwidths 0.55–0.68 μm (red), 0.73–1.10 μm (red to near-infrared), 3.55–3.93 μm (mid-infrared), 10.3–11.3 μm (thermal infrared), and (if present) 11.5–12.5 μm (thermal infrared), respectively. These data were downlinked by the Scripps Satellite Oceanography Center, calibrated to geophysical units (Lauriston et al., 1979 and updates), and coregistered to a common geophysical grid (e.g., Legeckis and Pritchard, 1976). Clouds were removed from the data using the procedure of Simpson and Humphrey (1990). A common cloud/land mask for the AVHRR images was created using the results from the individual image cloud masks, coastline grids, morphological transformations, and polygon fills (Simpson, 1992b). Water vapor atmospheric corrections and multichannel sea surface temperatures (MCSST) were calculated using

$$\mathrm{MCSST}=1.0095T_4+\frac{1.6692T_4}{R_{12}}-\frac{1.6682T_5}{R_{12}}-2.64, \quad (8)$$

where R_{12} is computed on a pixel by pixel basis as $R_{12}=\Delta T_5/\Delta T_4$ and ΔT_i is the temperature difference between adjacent pixels in Channels 4 and 5, respectively (Harris and Mason, 1992). This assumes that the atmospheric variation in water vapor changes slowly over the spatial scales of the local SST gradient which allows an additional nonlinear atmospheric transmittance correction to be computed from the local brightness temperature. The effectiveness of this approach was demonstrated by Harris and Mason (1992) and independently confirmed for the California Current using a large number of images and a neural network classifier (Yhann and Simpson, 1995).

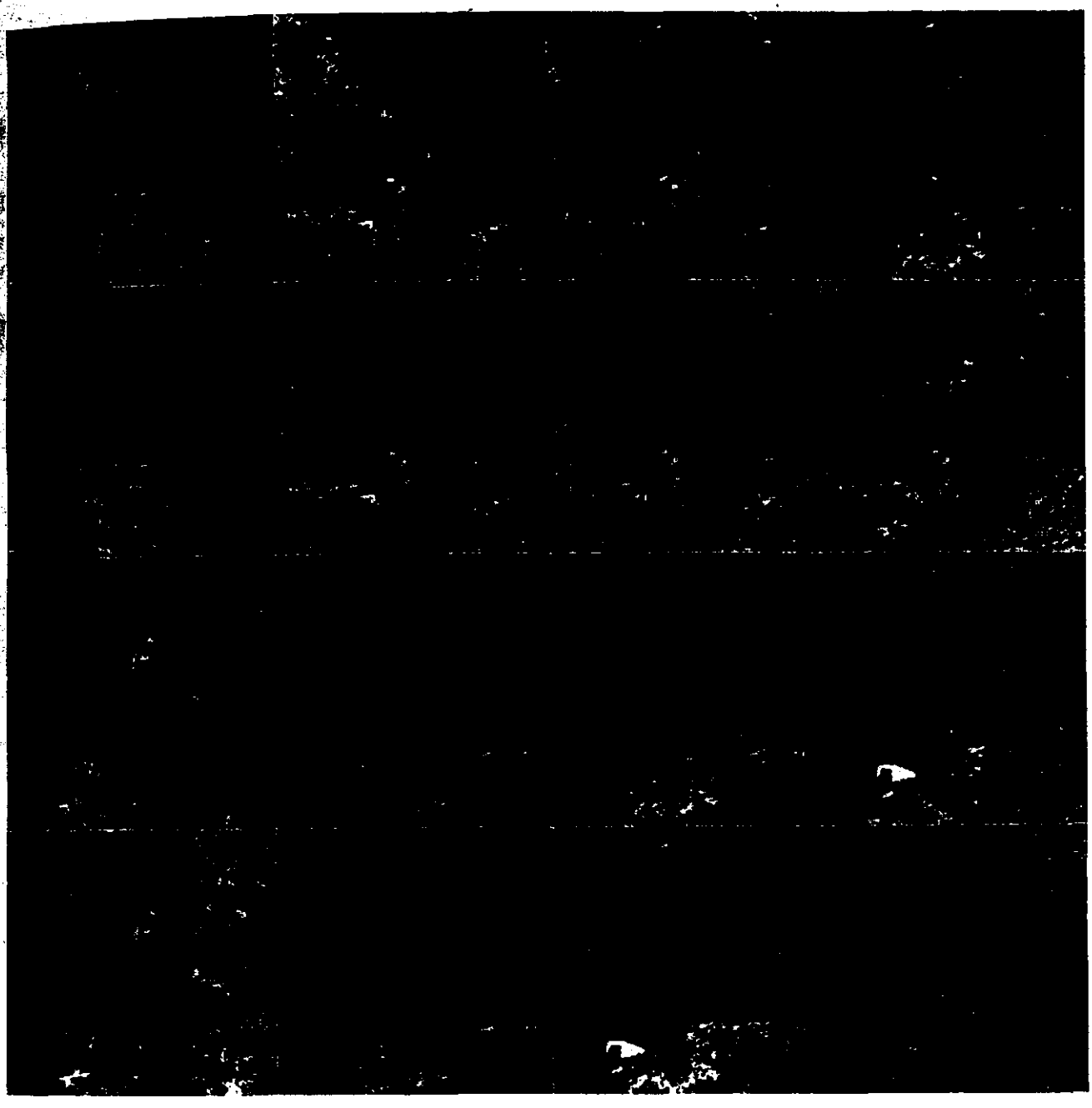
RESULTS

PC, MAD and MAF/MAD transformations are used to detect the spatial patterns of change found in a sequence of Landsat MSS data (terrestrial case) and a sequence of AVHRR data (oceanic case).

Terrestrial Case: Urbanization of Queensland, Australia

Graetz et al. (1992) used two Landsat MSS images taken on 11 November 1972 and 16 November 1988 to examine urbanization in Queensland, Australia, over the 16-year period. In 1991 the Brisbane Statistical Division registered a population of 2,978,617 people with an estimated growth rate of about 2% per year. The most striking man-made feature pointed out by Graetz et al. (1992) was the establishment and filling of three large water reservoirs to accommodate the population growth. The largest of these was developed to the west of the cities of Brisbane and Ipswitch. PC and MAF/MAD analyses are applied on the relevant subsections of the two Landsat MSS scenes to detect the pattern of land cover change produced by the development and filling of the largest reservoir.

Figure 1 shows the western region of the MSS image (Bands 1–4, panels a–d, respectively) taken on 11 November 1972. Each image is 500×500 pixels. For the present analysis, the feature of most importance in these panels is a section of the river clearly visible in all four bands of the MSS image. Figure 1 (panels e–h) shows the same area 16 years later. Figure 1 (panels i–l) shows the simple difference images, and Figure 1 (panels m–p) shows the principal components of these differences. The formation of the reservoir is clearly seen in the second, third and fourth columns. Spectral means (Table 1) and covariance/correlation structure of the original data (Table 2) are consistent with the changes observed in Figure 1. Figure 2 (panels a–d, e–h) shows the canonical variates of the MSS data in the two years. Figure 2 (pan-



Key:

a	e	i	m
b	f	j	n
c	g	k	o
d	h	l	p

Figure 1. Landsat MSS Bands 1–4, 11 November 1972 (panels a–d), 16 November 1988 (panels e–h), simple differences (panels i–l), and PCs of simple differences (panels m–p).

Table 1. Means for the Terrestrial Case Study

	MSS1	*MSS2*	*MSS3*	*MSS4*
1972	67.38	53.21	99.18	78.65
1988	67.61	57.52	96.48	23.41
Difference	0.23	4.31	−2.70	−55.24

els i–l) shows the corresponding MAD variates. Figure 2 (panels m–p) shows the MAF/MADs. PC1 (Figure 1, panel m) contains maximum variance (see Table 3) and shows the pattern of reservoir formation as well as a region of urbanization and hobby farms in the lower part of the PC1 [see Graetz et al. (1992) for additional supporting information]. PC4 (Figure 1, panel p) contains minimum variance (Table 3) and primarily shows a noise pattern in the MSS data. The general patterns found in the PCs also are found in the MADs (Fig. 2, panels i–l) except that neither the reservoir and hobby farm formation nor the MSS noise are confined to a single MAD.

Figure 3 (panels a–c) shows the composite of the three first PCs, PC1 in red, PC2 in green, and PC3 in blue. Figure 3 (panels d–f) shows the composite of the three most spatially coherent MAF/MADs, MAF/MAD1 in red, MAF/MAD2 in green, and MAF/MAD3 in blue. In panels b and e, no-change regions in the images are masked to gray. In panels c and f, no-change regions in the images of the absolute values are masked to black. Note that the MAF transformation maximizes the autocorrelation in the data rather than maximizing the variance. This is clear in Figure 2, panels m–p. Here, reservoir formation is shown in MAF/MAD1 (panel m), while the MAF/MAD4 (panel p) largely is noise in the composite dataset (i.e., the least spatially coherent component of the composite dataset, Table 3). In order, the three most spatially coherent patterns of change in Figure 3 are the reservoir formation (red), urbanization and hobby farm development (green) in the lower part of the image, and changes in the river structure (blue). The RGB composite MAF/MAD visually preserves this rank order of change.

Simple No-Change Simulation.

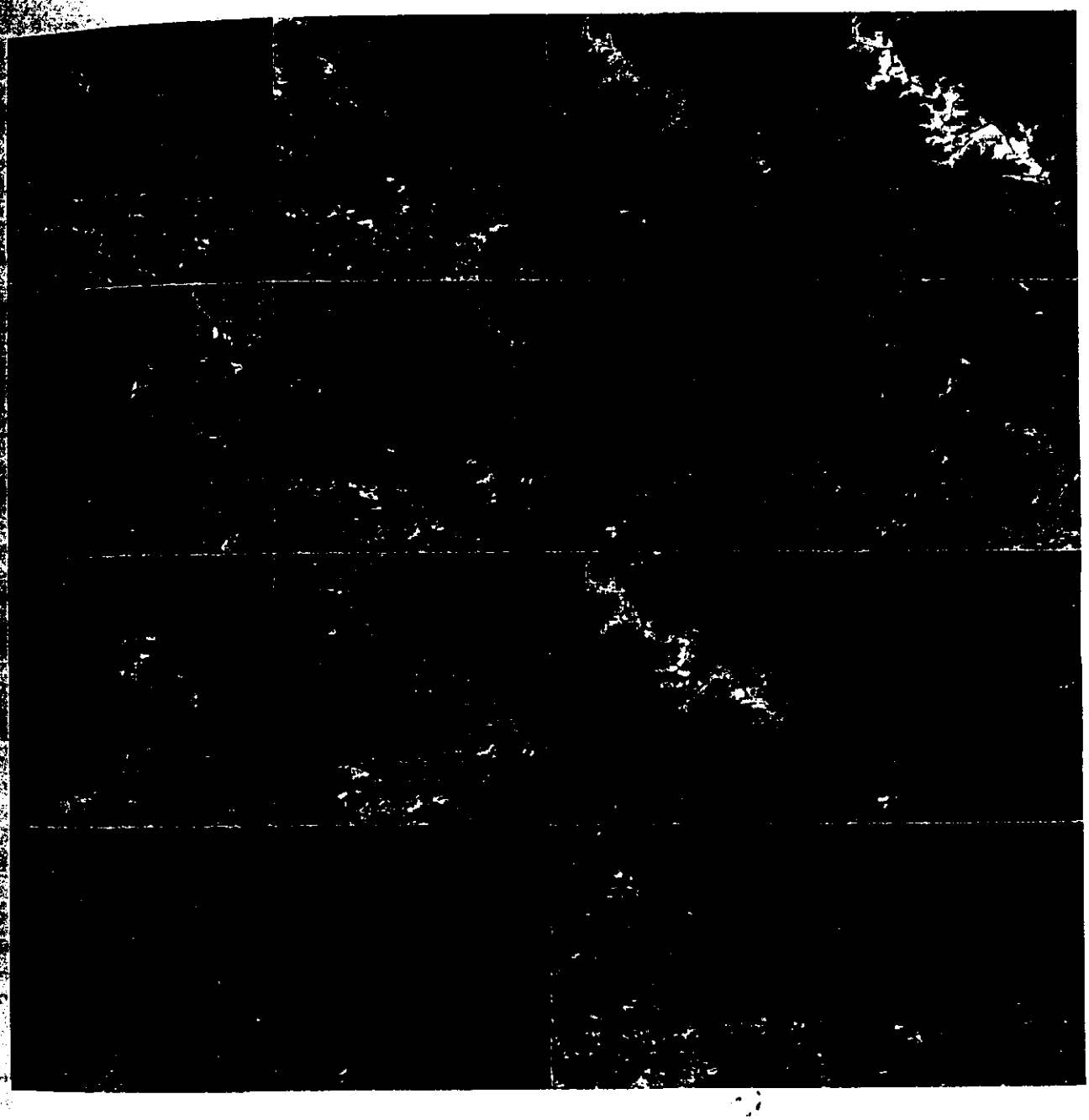
As a simple simulation of a situation with no change in all bands we pad the two 500×500 Landsat MSS scenes described above into the central part of 550×550 backgrounds with values 0 in all bands in both years. Change between the two 550×550 scenes is estimated by means of: 1) simple differences; 2) principal components of simple differences (based on the covariance matrix); 3) principal components of simple differences (based on the correlation matrix); 4) varimax rotated factors of simple differences; 5) MAFs of simple differences; 6) MADs; and 7) MAF/MADs. Change detected in the region with no change (the 25 pixels broad edge around the actual image data) as indicated by standardized values of the results from the different change detection methods is given in Table 4. It is obvious that MAD and MAF/MAD are the only multivariate techniques that perform well in this situation. All other methods give values much higher than 0.

Oceanic Case: ENSO-Related Mid-Latitude Coastal Warming

El Niño–Southern Oscillation (ENSO) events are characterized by large-scale changes in the structure and dynamics of the tropical atmosphere–ocean system (e.g., Bjerknes, 1966; Horel and Wallace, 1981). ENSO events, however, are not confined to tropical regions; generally they are accompanied by contemporaneous midlatitude warming events and changes in polar circulation. The 1982–1983 ENSO event (e.g., Rasmusson and Wallace, 1983) was accompanied by a major expansion and intensification of the Aleutian Low Atmospheric Circulation (e.g., Simpson, 1983) and by major changes in ocean circulation and property distributions from the equatorial eastern tropical Pacific to Alaska (e.g., Lynn, 1983; Simpson, 1984a,b). In particular, pronounced downwelling occurred along the west coast of North America which contributed to the formation of anomalous warmer surface and subsurface water in the inshore domain of the California Current System (CCS). Plankton and nekton distributions also were modified in response to these changes in the physical structure and dynamics of the CCS (e.g., Fiedler, 1984; McGowan, 1985).

Table 2. Covariance/*Correlation* Matrix for the Terrestrial Case Study

		1972				*1988*			
		MSS1	*MSS2*	*MSS3*	*MSS4*	*MSS1*	*MSS2*	*MSS3*	*MSS4*
1972	MSS1	77.30	96.36	76.24	39.39	54.87	92.09	11.11	−4.95
1972	MSS2	*0.91*	144.60	61.84	12.13	74.92	128.68	18.71	−5.04
1972	MSS3	*0.49*	*0.29*	310.95	269.25	44.31	66.40	85.42	16.43
1972	MSS4	*0.27*	*0.06*	*0.95*	260.83	16.52	19.15	81.36	19.41
1988	MSS1	*0.57*	*0.57*	*0.23*	*0.09*	119.87	192.62	198.53	43.96
1988	MSS2	*0.58*	*0.59*	*0.21*	*0.07*	*0.97*	328.51	286.42	61.62
1988	MSS3	*0.04*	*0.05*	*0.16*	*0.17*	*0.61*	*0.53*	880.78	235.44
1988	MSS4	*−0.07*	*−0.05*	*0.12*	*0.15*	*0.50*	*0.42*	*0.98*	65.48



Key:

a	e	i	m
b	f	j	n
c	g	k	o
d	h	l	p

Figure 2. Canonical variates (CVs) of 1972 MSS data (panels a–d), CVs of 1988 MSS data (panels e–h), MADs (panels i–l), and MAF/MADs (panels m–p).

Table 3. Key Parameters for PC, MAD and MAF/MAD Transformations for the Terrestrial Case Study

Eigen number	*PC Component Variance*	*MAD Canonical Correlation*	*MAD Component Variance*	*MAF/MAD Correlation*
1	1320.56	0.03	1.95	0.78
2	205.77	0.18	1.64	0.74
3	76.14	0.23	1.54	0.46
4	9.10	0.73	0.54	0.22

AVHRR images, representative of conditions in the CCS near the beginning (29 January 1982) and during the advanced stages (10 January 1983) of the 1982–1983 warming event were selected for analysis. Each image is 1024×512 pixels, centered at 31° 30′N latitude and 119° 30′W longitude and covers part of the southern California and Baja California sectors of FAO Fishing Region 77.

In this case we used the AVHRR channel 3 and the MCSST data only. Figures 4a and 4b show the 1982 data. Figures 4c and 4d show the 1983 data. The 1982 image contains relatively little noise in AVHRR channel 3 data (Fig. 4a), but there is considerable noise in the 1983 channel 3 image (Fig. 4c). There is little or no noise in the MCSST images. Mean values and covariance/correlation structure are given in Tables 5 and 6. The PC analysis of the simple difference data is shown in Figures 5a and 5b. Parameters for the PC analysis are given in Table 7. Because of the extremely high levels of noise in the 1983 channel 3 data, the PC analysis failed (i.e., a significant noise fraction of the uncorrected channel 3 data contaminates each principal component). PC analysis only produces valid results (computed but not shown) if the channel 3 data are first Wiener filtered to remove noise (Simpson and Yhann, 1994). The MAF/MAD analysis (Figs. 5c and 5d) isolated the highly spatially coherent pattern of change in MAF/MAD1 (Fig. 5c, Table 7) and confines the noise to MAF/MAD2 (Fig. 5d).

False color representations of MAF/MAD1 are shown in Figure 6. The pattern of positive change (yellow to red in Fig. 6a) and negative change (blue to magenta in Fig. 6a) is consistent with downwelling near the coast and onshore transport of cooler Pacific Subarctic water from the offshore California Current observed during the 1982–1983 event (e.g., Lynn, 1983; Simpson, 1984a,b; McGowan, 1985). In Figure 6b, the no-change regions are masked to black. Both the MADs alone and the MAF/MADs (Figs. 5c and 5d) were highly effective in isolating the AVHRR channel 3 noise. A more detailed treatment of noise isolation in satellite data using the MAF/MAD transformations is given in the discussion.

In situ observations of near-surface temperature (1-m depth) were made with a Neil Brown CTD from *R/V David Starr Jordan* and *R/V Townsend Cromwell* as part of the 1982 and 1983 California Cooperative Oceanic Fisheries Investigations (CalCOFI) surveys, respectively. These data were extrapolated to the surface and interpolated to a 1-km grid to produce *in situ* maps of sea surface temperature. The dates of the two CalCOFI cruises are approximately 2 weeks later than the respective 1982 and 1983 satellite images used in this study due to cloud cover. Each cruise lasted approximately 3 weeks. Figures 7a and 7c show satellite derived MCSST distributions for both years. Figures 7b and 7d show image representations of the CalCOFI shipboard data interpolated to the satellite grid. The CalCOFI cruise tracks did not cover the full area viewed by the satellite. Therefore, to avoid errors in interpolation associated with regions of no ship data, such areas were masked from the data. Thus, the ship data mask (Figs. 7b and 7d) is different from the satellite data mask (Figs. 7a and 7c). In general, there is a good large-scale agreement between the shipboard SST and the satellite MCSST for both 1982 and 1983. The correlations between the satellite and *in situ* data (shipboard observation and nearest satellite pixel) are 0.95 and 0.82 for the 1982 and 1983 data sets, respectively. The lower correlation for the 1983 data is consistent with anomalous high near-surface warming rates and near-surface static stability observed in the region for this period (see Simpson, 1992a, Figs. 6 and 11). These processes tend to distort the large-scale coherent patterns associated with long-term mean conditions in the regions (Lynn and Simpson, 1987).

The MCSST difference (Fig. 8a) and the corresponding *in situ* data difference (Fig. 8b) show similar large-scale structures; consistent with the correlations cited above. The *in situ* data, however, do not capture the fine scale structure shown in the satellite data because of their coarser spatial resolution and subsequent interpolation. The mean and standard deviations for the MCSST *in situ* temperature differences are $-0.45°C \pm 0.4°C$ for 1982 and $-0.34°C \pm 0.41°C$ for 1983. As expected, the MCSSTs are colder than the *in situ* data, and these differences (both magnitude and sign) are consistent with ocean molecular boundary layer effects at the air–sea interface (Paulson and Simpson, 1981). Coastal regions in the difference image contain even larger, positive temperature differences. The patterns of change detected by the MAF/MAD analysis (Figs. 5 and 6) are consistent with those found in the shipboard data.

A vertical section of temperature (Fig. 9) also was

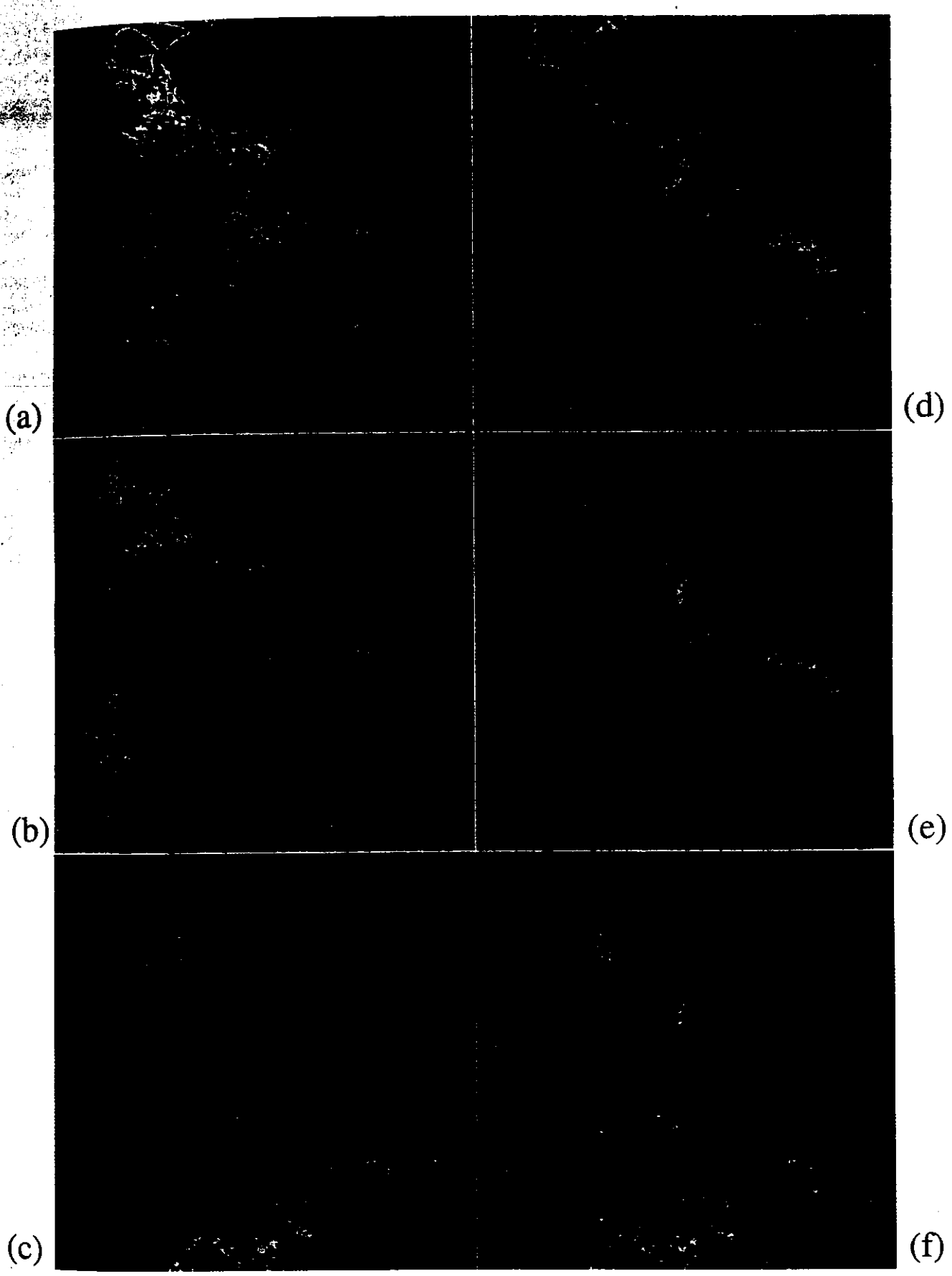


Figure 3. Landsat MSS color composites of PCs (panels a–c) and MAF/MADs (panels d–f).

Table 4. Change Detected in No-Change Regions for Landsat MSS Data, Value Should Be 0

Channel/Component	*1*	*2*	*3*	*4*
1) Simple difference	0	0	0	0
2) PC difference (cov)	0.77	−1.72	0.56	−0.06
3) PC difference (corr)	0.23	1.59	1.13	0.07
4) Factors difference	−0.11	1.90	−0.48	−0.05
5) MAF difference	1.94	0.32	0.06	−0.04
6) MAD	−0.06	−0.01	0.11	0.07
7) MAF/MAD	0.05	0.06	−0.03	0.12

constructed from these data along a line of CTD stations marked with white crosses in Figure 4d. These data were used to compute subsurface thermal anomaly structure relative to the long-term, large-scale CalCOFI mean structure determined from 28 years of shipboard observations [see Simpson (1992a) for details of the computation]. They confirm that the surface signature viewed by the satellite was indicative of a deeper ocean process. A more detailed discussion of near-surface and subsurface oceanic response in the region to ENSO-related forcing is given by Simpson (1992a) and the references contained therein.

The above analysis was repeated using three-channel AVHRR images (channels 3, 4, and 5) in place of the two channel images (channel 3 and MCSST) shown in Figure 4. Again, the MAD and the MAF/MAD transformation (computed but not shown) proved superior to PC-based methods for separation of the channel 3 noise found in the January 1983 data and subsequent detection of the change in California Current sea surface temperature observed during the 1982–1983 ENSO midlatitude coastal warming event.

DISCUSSION

Advantages of the MAF/MAD Transformation for Change Detection

Simple differencing of two coregistered satellite scenes separated in time provides a point-by-point estimate of the rate of change of the quantity under study. Simple differencing has the advantage that it is physically intuitive. Simple differencing, however, also has some inherent disadvantages: 1) the simple differences are affected significantly by the absolute accuracy and temporal stability of the calibration of the instrument on the satellite (e.g., Frouin and Simpson, 1995); 2) generally, the bands in the differenced image are themselves correlated because the original data usually show some degree of correlation [e.g., data from AVHRR Bands 4 and 5 typically have a correlation of 0.99 at 0 spatial lag (Simpson and Yhann, 1994)]; and 3) time-varying sensor noise may render a given channel difference difficult to use [e.g., AVHRR channel 3 sensor noise (Simpson and Yhann, 1994)].

The principal component (PC) transformation of the simple differenced image partially mitigates some of the problems cited above because the principal components of a multivariate variable are produced by a linear transformation which produces uncorrelated variables of decreasing variance (Hotelling, 1933). The PC analysis, however, also has some inherent disadvantages: 1) the PC transformation is not invariant to a linear scaling of the input data (e.g., Preisendorfer, 1988); 2) the PC transformation, because it maximizes the variance in the first few PCs, can be artificially weighted by a single image in the sequence which contains a disproportionate amount of variance (e.g., Gallaudet and Simpson, 1994); 3) the traditional PC analysis is performed as a set of pixelwise operations, and thus does not take into account the inherent spatial structure of the image data (e.g., Banet and Lebart, 1984); and 4) certain forms of noise in the data (e.g., Fig. 4c) can render PC analysis unfruitful (e.g., Figs. 5a and 5b).

The multivariate alteration detection (MAD) transformation of multispectral, bitemporal satellite images is an extension of traditional canonical correlations analysis. The MAD transformation simply is the difference of the canonical variates of the two sets of multivariate data in reverse order. MADs, like PCs, produce uncorrelated differences (i.e., the MADs are orthogonal). The MAD transformation has several advantages over the PC transformation: 1) MADs are invariant to a linear scaling of the input data (Nielsen, 1994); 2) because the MADs are based on canonical variates, all data are weighted equally, and therefore MADs are not sensitive to disproportionate amounts of variance being isolated in one or a few images (as are PCs); 3) the MAD analysis is less subject to noise contamination than the PC analysis (Fig. 5); and 4) the MADs form a more generalized difference than does the PC operating on the simple differences of images. Thus, the MAD transformation provides a way of combining different types of data (e.g., Landsat TM and Landsat MSS) which may be useful in historical change detection studies (Nielsen, 1994). MADs, however, like PCs are performed as local point operations. Thus, MADs, like PCs, fail to retain the spatial context of the data in the change detection results.

Unlike the pixel-based PC and MAD transformations, the MAF transformation incorporates the spatial structure of the data in the image into the analysis by using the dispersion matrix of the difference between the data and the spatially shifted data (Switzer and Green, 1984). The MAF transformation differs from the PC transformation in that the MAF maximizes the autocorrelation between neighboring observations whereas the PC transformation maximizes the variance. MAFs, like MADs, are invariant to linear scaling of the input data. MAFs also are uncorrelated, that is, the MAF transformation also is an orthogonal transformation (Switzer and Green, 1984).

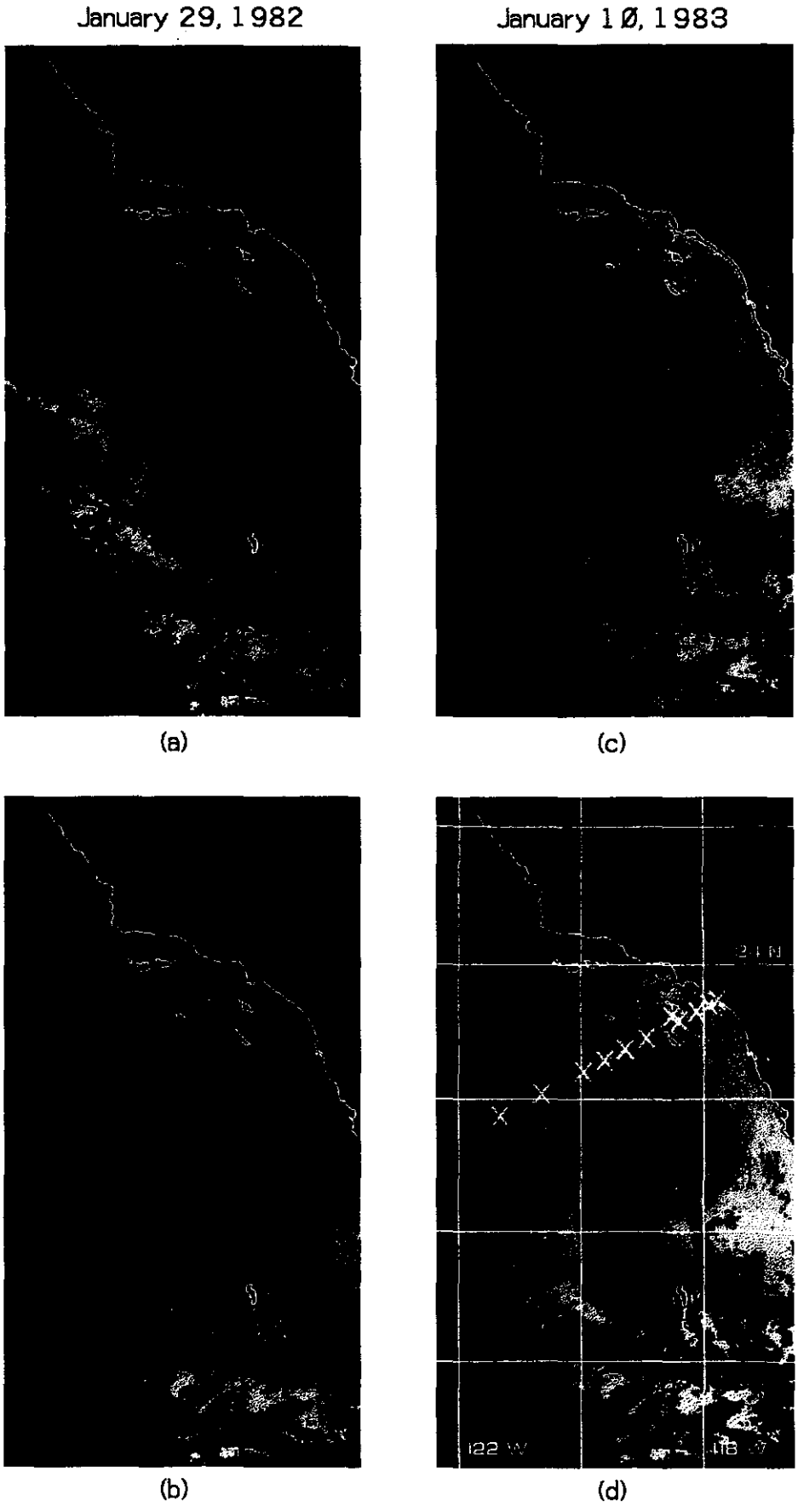


Figure 4. AVHRR channel 3 and MCSST, 29 January 1982 (panels a and b), 10 January 1983 (panels c and d). White X's (panel d) show station locations of CalCOFI line 90 vertical profiles.

Table 5. Means for the Oceanic Case Study

	AVHRR3	*MCSST*
1982	13.48	13.92
1983	14.47	15.51
Difference	0.99	1.59

Central to most global change studies are data and analysis needed to 1) accurately determine the effective magnitude and spatial distribution of the change (e.g., deforestation) and 2) determine the rates of change that effect such biomes in a given region. A combined MAF/MAD analysis provides a statistically rigorous way to accurately determine the spatially coherent patterns of major change in an image sequence by retaining the spatial context of the neighborhood pixels in the analysis. Linear transformation invariance, however, precludes computing an estimate of the effective magnitude of the change. Simple band by band differencing of properly calibrated data provides an estimate of the rate of change, but this estimate lacks a spatial context. MAF/MAD analysis (to statistically determine patterns of coherent change in the data), coupled with the simple differencing of properly intercalibrated images, is an effective combination to 1) compute accurate rates of change and 2) place these rates in a statistically meaningful spatial context. This can be achieved by assigning significance only to the simple differenced values in regions where the correlation between the simple differences and the MAF/MADs of the simple differences is high.

Interpretation of MADs and MAF/MADs

A few simple statistical measures provide valuable information for the interpretation of the PCs, MADs, and MAF/MADs. Assuming a stable, accurate instrument calibration, the means of the individual spectral channels provide a gross estimate of the observed spectral change over the time between images. The diagonal elements of the covariance matrix of the original composite dataset provide estimates of the variances in the composite dataset as a function of wavelength over time. The off-diagonal elements provide information on the interband correlation in the composite dataset. Large changes in the means and the variances over time, especially if these changes occur in some but not all of the spectral bands, provide insight into the type of process that may have produced the observed change. Changes in MSS channels 3 and 4, for example, are useful indicators of either vegetation-based or land/water boundary-based change. The amount of spatial correlation in each MAF/MAD can provide a basis for separating signal from some types of noise in the composite dataset. Finally, the correlation matrices between the original composite dataset and the MADs and the MAF/MADs provide an indication of the processes that have contributed to the change detected by a given component of the MADs and MAF/MADs.

Consider, for example, the terrestrial case study. The means of the composite dataset (Table 1) show very little change in most of the bands except band 4 where the mean drops from 78.65 to 23.41. The diagonal elements of the covariance matrix (Table 2) show that, in general, the spectral variances are higher in 1988 than in 1972 except for the variance in MSS channel 4 which shows a fourfold drop in value (from 260.83 in 1972 to 65.48 in 1988). These changes are consistent with two known facts for the region in the times measured: 1) a major fraction of the scene changed from natural vegetation along a riverbank (1972) to a large reservoir (1988). The reflectance properties of water are more uniform and lower (e.g., water is a near blackbody with an emissivity of about 0.98) than those of vegetation; 2) the 1972 data were taken under very dry, nongreen vegetation conditions whereas the 1988 data were taken under wetter conditions with much greener vegetation (Graetz et al., 1992). Table 8 shows that MAF/MAD1 is positively correlated with all bands in the 1972 data and negatively correlated with all the bands in the 1988 data. These correlations are consistent with conversion of natural vegetation to reservoir and to greening up of the vegetation in 1988 compared to 1972. MAF/MAD2 shows maximum positive correlation with 1972 MSS4 data (another indication of vegetation change) and maximum negative correlation with 1972 MSS2 data. Moreover, MAF/MAD2 has the form of a vegetation index in 1972 and a negative vegetation index in 1988, a result consistent with the vegetation changes reported by Graetz et al. (1992). MAF/MAD3 shows maximum correlations with MSS3 data in both years and MSS4 data in 1988. These correlations are consistent with reservoir formation. MAF/MAD4 shows little correlation with any of the data, consistent with the residual noise pattern found in MAF/MAD4. Analogous statistical information for the oceanic case is given in Tables 5, 6, 7, and 9.

Table 6. Covariance/*Correlation* Matrix for the Oceanic Case Study

		1982		*1983*	
		AVHRR3	*MCSST*	*AVHRR3*	*MCSST*
1982	AVHRR3	2.65	1.49	0.70	0.74
1982	MCSST	*0.76*	1.46	1.00	1.07
1983	AVHRR3	*0.17*	*0.33*	6.32	1.08
1983	MCSST	*0.42*	*0.82*	*0.40*	1.17

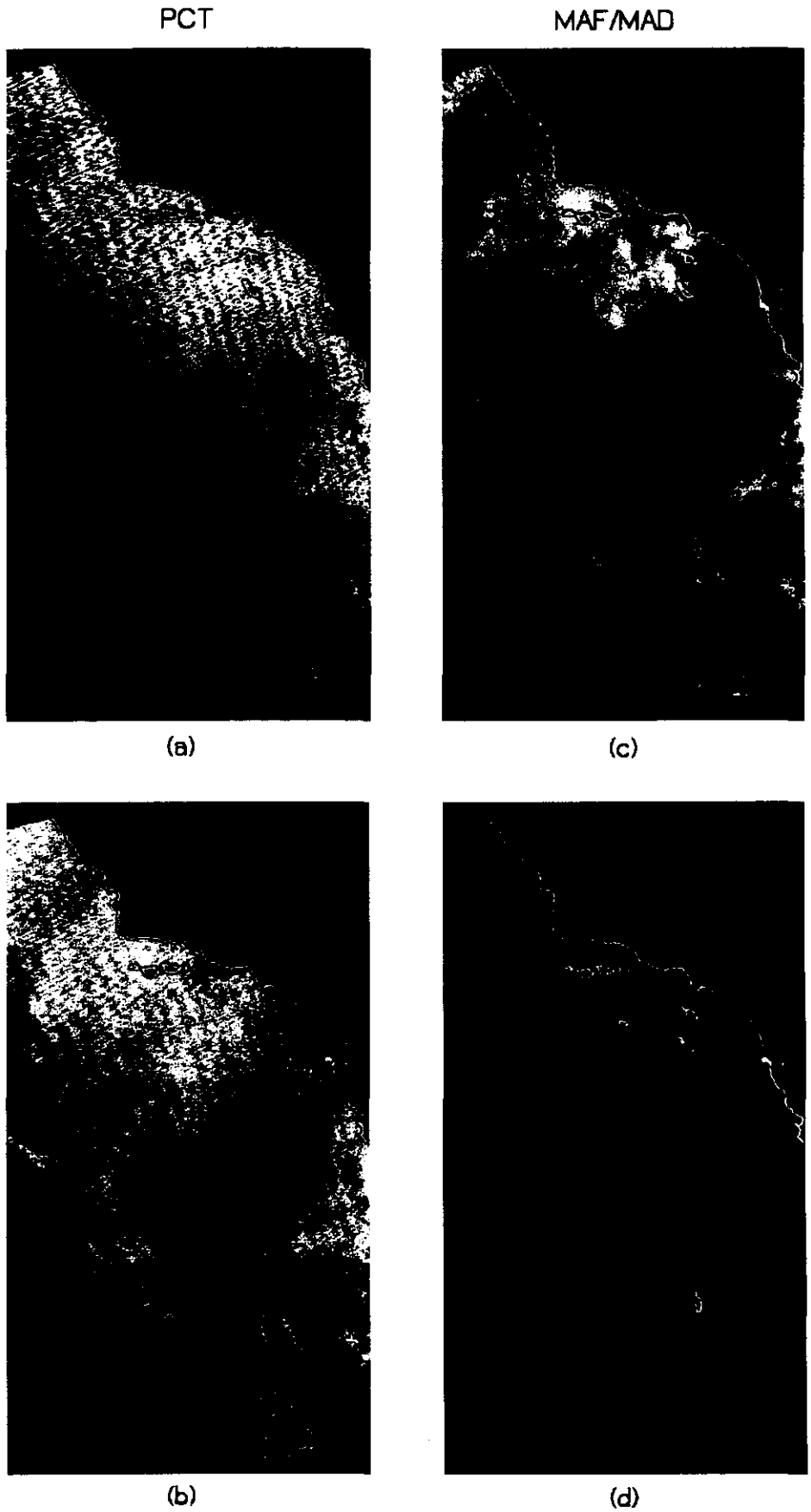


Figure 5. PCs of simple AVHRR differences (panels a and b), MAF/MADs (panels c and d).

Table 7. Key Parameters for PC, MAD, and MAF/MAD Transformations for the Oceanic Case Study

Eigen number	*PC Component Variance*	*MAD Canonical Correlation*	*MAD Component Variance*	*MAF/MAD Correlation*
1	7.67	0.00	1.99	0.96
2	0.40	0.88	0.25	0.68

Noise Isolation Using MAF/MADs

Noise in image data can be either spatially coherent or incoherent. Because the MAF transformation maximizes the autocorrelation rather than maximizing the variance (as does the PC) of the data, one expects that spatially incoherent noise would be isolated in higher-order MAFs. The oceanic case study shows that both the MAD (not shown) and MAF/MAD analyses correctly isolated AVHRR channel 3 noise whereas the PC analysis was distorted by the noise (Figs. 5 and 6). PC1 isolated only part of the channel 3 noise. PC2 shows the actual pattern of oceanic change, but it is heavily compromised by residual channel 3 noise not removed by PC1. The MAF/MAD analysis isolates the signal variance to MAF/MAD1 (Fig. 5c) because the ENSO-induced change in SST was spatially more coherent than the noise in the data. Noise is correctly isolated in MAF/MAD2 (Fig. 5d).

Because the striping in the Landsat MSS data used in the terrestrial case is periodic (i.e., it is a form of spatially coherent noise) and occurs in all the bands of the composite dataset, none of the transformations will completely remove it from the data. In cases such as this, the periodic noise can be removed by using finite impulse response filtering techniques, either in the spatial or the Fourier domain, to minimize the effects of the stripes prior to change detection analysis. Such proce-

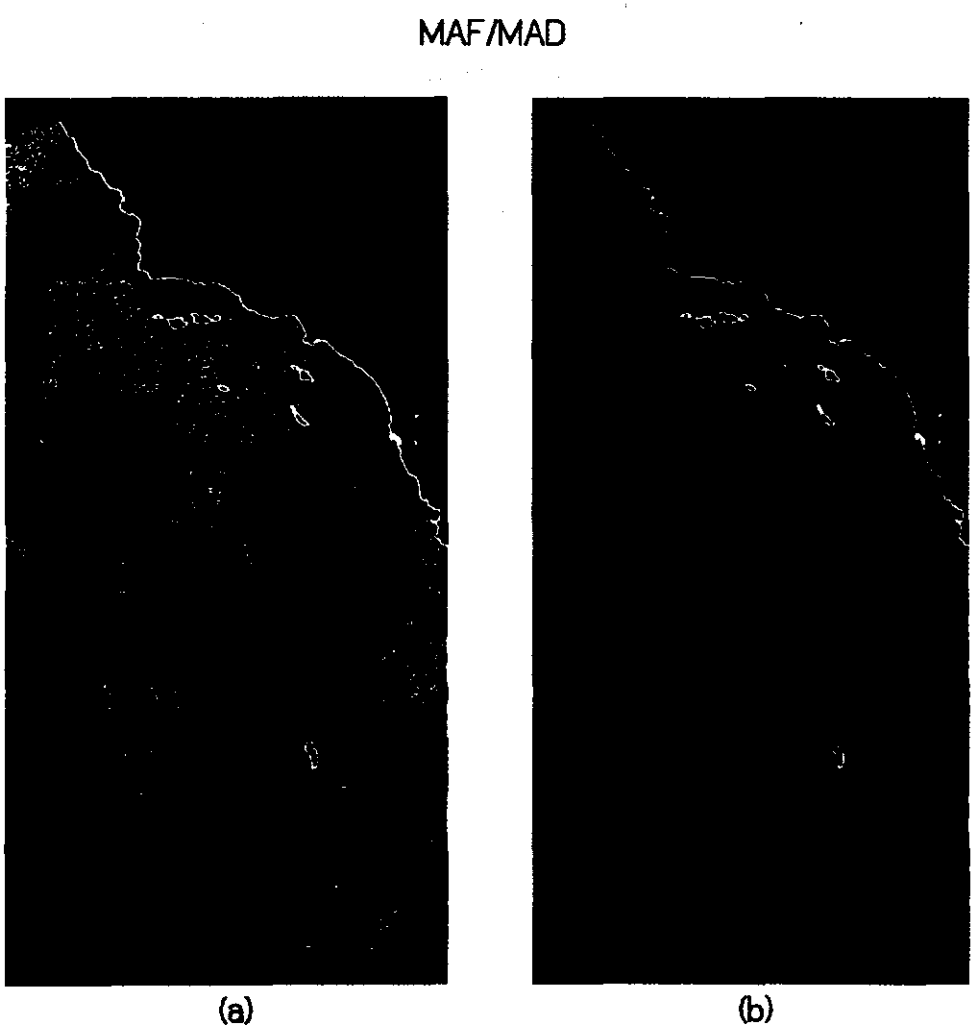


Figure 6. a) AVHRR false color plot of MAF/MAD1; b) regions of maximum change (positive= red; negative=blue/magenta) of MAF/MAD1.

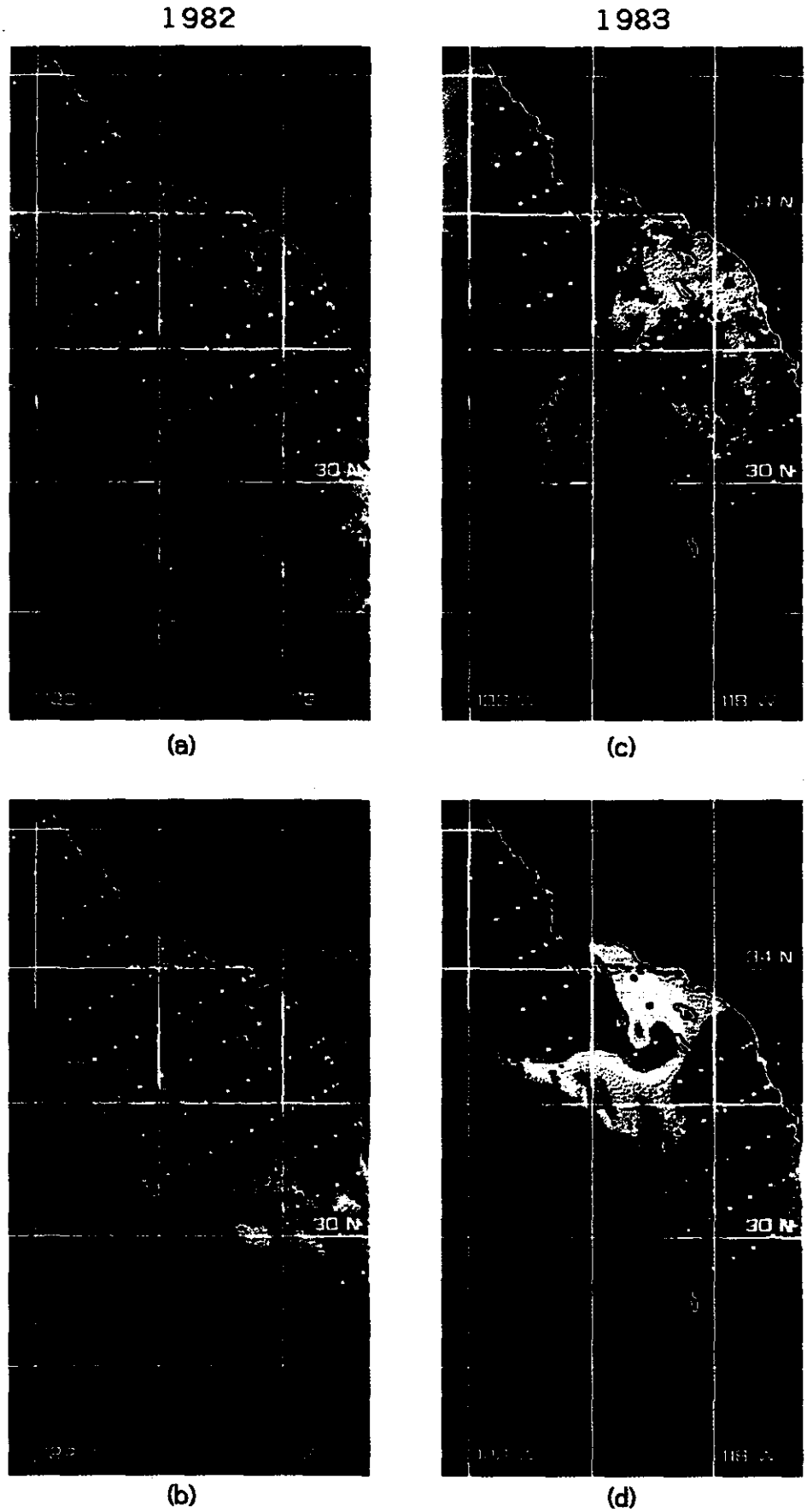


Figure 7. Satellite based MCSST and interpolated CalCOFI in situ surface temperature for 1982 (panels a and b). Analogous data for 1983 are shown in panels c and d.

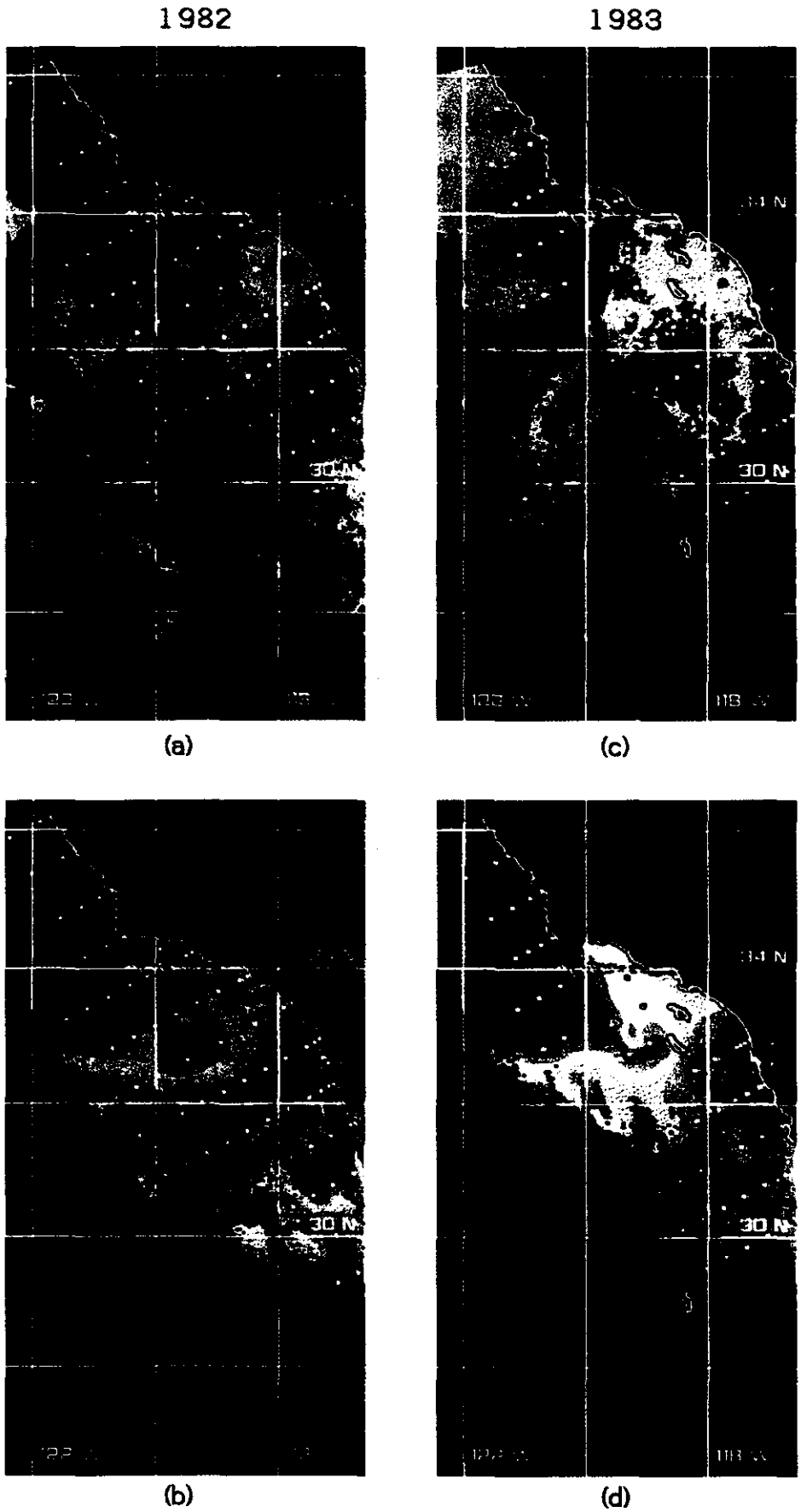


Figure 7. Satellite based MCSST and interpolated CalCOFI in situ surface temperature for 1982 (panels a and b). Analogous data for 1983 are shown in panels c and d.

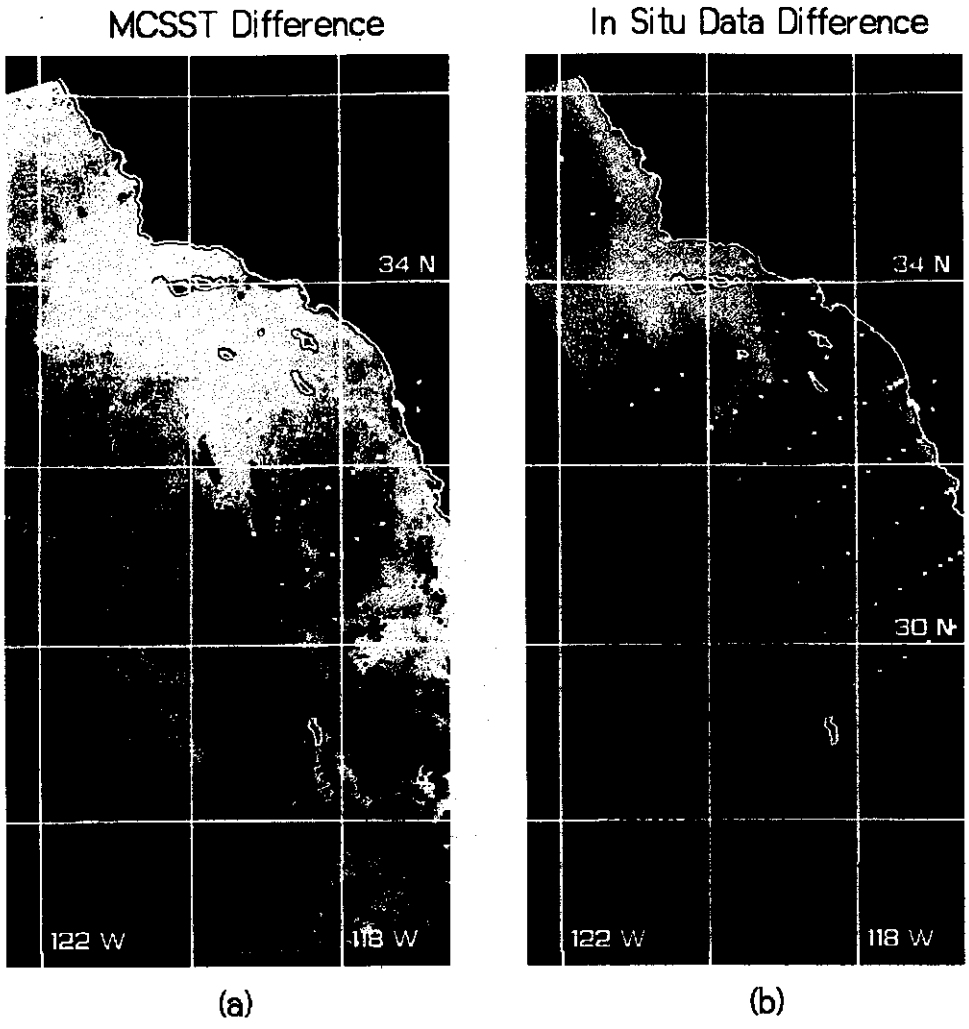


Figure 8. a) Satellite MCSST difference (1983–1982); b) *in situ* temperature difference for the same period.

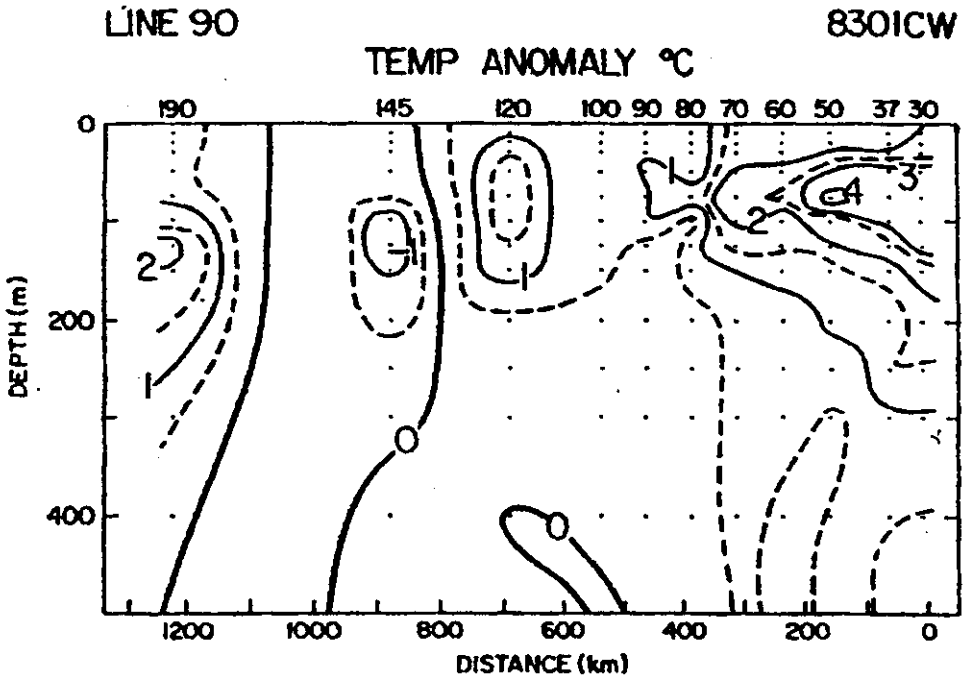


Figure 9. Vertical section of ocean temperature. Data were taken along CalCOFI line 90 (see Fig. 4d).

Table 8. Correlations between Input Data and MADs and MAF/MADS for the Terrestrial Case Study

		MAD1	MAD2	MAD3	MAD4	MAF/ MAD1	MAF/ MAD2	MAF/ MAD3	MAF/ MAD4
1972	MSS1	−0.02	0.11	0.26	0.57	0.46	−0.27	0.32	0.12
1972	MSS2	0.02	−0.25	0.14	0.56	0.37	−0.47	0.05	−0.03
1972	MSS3	−0.09	0.41	0.65	0.11	0.59	0.35	0.34	−0.08
1972	MSS4	0.05	0.41	0.67	−0.02	0.58	0.45	0.21	0.38
1988	MSS1	−0.41	0.28	−0.61	−0.43	−0.78	0.18	0.23	−0.04
1988	MSS2	−0.25	0.36	−0.43	−0.41	−0.59	0.28	0.21	0.08
1988	MSS3	−0.42	0.13	−0.69	0.01	−0.66	−0.23	0.33	0.04
1988	MSS4	−0.47	0.18	−0.65	0.07	−0.62	−0.24	0.41	0.02

times to remove periodic stripes 1) in hyperspectral airborne scanner data (GERIS) are given by Nielsen (1994) and Nielsen and Larsen (1994); and 2) in GOES data by Simpson et al. (1995). With minor frequency tuning, both techniques could be successfully applied to Landsat MSS data.

Detection of Outliers in the Data

Three factors often can produce erroneous change detection independent of the method used for the change detection analysis: 1) errors in the coregistration of the two scenes; 2) inadequate cloud detection; and 3) onboard satellite amplifier hysteresis at sharp boundaries (e.g., cloud–land, ocean–land). Errors in coregistration can, for example, incorrectly align a land pixel in the one image with an ocean pixel in the other. Accurate cloud detection is essential because clouds typically are the most transient features in satellite scenes. Hysteresis, which causes incorrect sensor integration of counts on the subpixel level, can produce a pixel value near a sharp boundary whose brightness count is not representative of either of the classes on the opposite sides of the boundary. This often leads to misclassification and to large apparent changes at the affected pixel locations. All these effects can produce regions of apparent (but false) large change. If undetected, these areas of false change have the potential to hide regions of actual change.

In principle, image masking would minimize these effects. In practice, however, a few bad pixels simply may not be detected. Because such pixels almost always appear as pixels with maximum change (but spatially isolated and few in numbers) in the MAD analysis, the MAD analysis can be used iteratively to locate and isolate them from the data. Thus, MAD analysis provides an additional and efficient method for data quality control in geophysical data.

Table 9. Correlations between Input Data and MADS and MAF/MADs for the Oceanic Case Study

		MAD1	MAD2	MAF/ MAD1	MAF/ MAD2
1982	AVHRR3	−0.67	−0.17	−0.20	−0.64
1982	MCSST	−0.31	−0.32	−0.33	−0.26
1983	AVHRR3	0.64	0.34	0.37	0.59
1983	MCSST	0.05	0.62	0.62	−0.05

CONCLUSIONS

Two multivariate statistical transformations, the multivariate alteration detection (MAD) transformation of Nielsen and Conradsen (1997) and the maximum autocorrelation factor (MAF) transformation of Switzer and Green (1984) are used to accurately detect coherent patterns of spatial change in sequences of satellite data. Urbanization during 1972–1988 in Queensland, Australia and ENSO-related midlatitude coastal warming off the west coast of North America during 1982–1983 were examined. The methods proved superior to the more traditional principal component (PC) transformation of simple differenced data for change detection studies. MAD and MAF transformations also are effective in removing incoherent noise from image data and for detecting outliers (e.g., coregistration errors). Both MADs and MAFs are invariant to linear and affine transformations such as: 1) unit variance normalization of the data; 2) changes in sensor gain and offset; and 3) radiometric and atmospheric corrections that are linear in the gray level numbers (desirable characteristics in a statistical analysis of geophysical data). Linear transformation invariance, however, precludes computing an estimate of the effective magnitude of the change. MAF/MAD analysis (to statistically determine patterns of coherent change in the data), coupled with simple differencing of properly intercalibrated data, is an effective combination to 1) compute accurate rates of change and 2) place these rates in a statistically meaningful spatial context. This combination can be achieved by only assigning significance to the simple differenced values in regions where the correlation between the simple differences and the MAF/MAD components is high.

APPENDIX: CANONICAL CORRELATIONS ANALYSIS

We consider a $(p+q)$-dimensional random variable $(p \leq q)$ ideally following a Gaussian distribution split into two groups of dimensions p and q, respectively

$$\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}\in N(\boldsymbol{\mu},\boldsymbol{\Sigma})=N\left(\begin{bmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{bmatrix},\begin{bmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{bmatrix}\right), \tag{A.1}$$

and we assume that $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\Sigma}_{22}$ (and $\boldsymbol{\Sigma}$) are nonsingular. Also, without loss of generality, we assume that $\boldsymbol{\mu}_1=\boldsymbol{\mu}_2=\mathbf{0}$. We are searching for linear combinations of $\mathbf{X}$ and $\mathbf{Y}$

$$\boldsymbol{U}=\boldsymbol{a}^T\mathbf{X}, \qquad \mathrm{Var}\{\boldsymbol{U}\}=\boldsymbol{a}^T\boldsymbol{\Sigma}_{11}\boldsymbol{a}, \tag{A.2}$$

$$\boldsymbol{V}=\boldsymbol{b}^T\mathbf{Y}, \qquad \mathrm{Var}\{\boldsymbol{V}\}=\boldsymbol{b}^T\boldsymbol{\Sigma}_{22}\boldsymbol{b}, \tag{A.3}$$

with maximum correlation

$$\rho=\mathrm{Corr}\{\boldsymbol{U},\boldsymbol{V}\}=\frac{\mathrm{Cov}\{\boldsymbol{U},\boldsymbol{V}\}}{\sqrt{\mathrm{Var}\{\boldsymbol{U}\}\ \mathrm{Var}\{\boldsymbol{V}\}}}=\frac{\boldsymbol{a}^T\boldsymbol{\Sigma}_{12}\boldsymbol{b}}{\sqrt{\boldsymbol{a}^T\boldsymbol{\Sigma}_{11}\boldsymbol{a}\boldsymbol{b}^T\boldsymbol{\Sigma}_{22}\boldsymbol{b}}} \tag{A.4}$$

If $(\boldsymbol{ab})$ is a solution, so is $(c_1\boldsymbol{a}, c_2\boldsymbol{b})$ where c_i is any scalar. We choose $(\boldsymbol{ab})$ so that $\mathrm{Var}\{\boldsymbol{U}\}=\mathrm{Var}\{\boldsymbol{V}\}=1$, introduce Lagrange multipliers $\lambda/2$ and $\nu/2$ and maximize

$$F=\boldsymbol{a}^T\boldsymbol{\Sigma}_{12}\boldsymbol{b}-\frac{\lambda}{2}(\boldsymbol{a}^T\boldsymbol{\Sigma}_{11}\boldsymbol{a}-1)-\frac{\nu}{2}(\boldsymbol{b}^T\boldsymbol{\Sigma}_{22}\boldsymbol{b}-1). \tag{A.5}$$

By setting $\partial F/\partial\boldsymbol{b}=\mathbf{0}$ and $\partial F/\partial\boldsymbol{a}=\mathbf{0}$, and inserting the results into Eq. (A.4), we get

$$\rho^2=\frac{\boldsymbol{a}^T\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{a}}{\boldsymbol{a}^T\boldsymbol{\Sigma}_{11}\boldsymbol{a}}=\frac{\boldsymbol{b}^T\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{b}}{\boldsymbol{b}^T\boldsymbol{\Sigma}_{22}\boldsymbol{b}} \tag{A.6}$$

$$\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{a}=\rho^2\boldsymbol{\Sigma}_{11}\boldsymbol{a}, \tag{A.7}$$

$$\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{b}=\rho^2\boldsymbol{\Sigma}_{22}\boldsymbol{b}, \tag{A.8}$$

that is, we find the desired projections for $\mathbf{X}$ by considering the conjugate eigenvectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p$ corresponding to the eigenvalues $\rho_1^2\geqslant\cdots\geqslant\rho_p^2$ of $\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$ with respect to $\boldsymbol{\Sigma}_{11}$. Similarly, we find the desired projections of $\mathbf{Y}$ by considering the conjugate eigenvectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_p$ of $\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$ with respect to $\boldsymbol{\Sigma}_{22}$ corresponding to the *same* eigenvalues ρ_i^2. If $p=q$, this will be all the eigenvalues and eigenvectors of $\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$. If $q>p$, the last eigenvalue will be 0 with multiplicity $q-p$.

A. A. Nielsen was funded in part by Danida, the Danish International Development Agency, under grant No. 104.Dan.8/410, the Commission of the European Union under Contracts Nos. MA2M-CT90-0010 and BRE2-CT92-0201, and the Danish National Technical Research Council under Grant No. 5.26.09.07. J. J. Simpson was funded in part by the Marine Life Research Group (MLRG) of the Scripps Institution of Oceanography and the NOAA Office of Global Program's Climate Change Data and Detection subprogram. Special thanks to T. Karl and W. Murray for their encouragement and support. Dr. M. Mullin (MLRG) also provided financial assistance for publication expenses. At the U.S. end, R. Keller and A. Schmidt assisted with computer programming. Helpful comments from two Danish colleagues, Dr. B. K. Ersbøll and Dr. R. Larsen, are gratefully acknowledged. Dr. R. Larsen also did most of the computer programming at the Danish end. V. Batchelder assisted with final manuscript preparation. Thanks also to the anonymous reviewers for comments and suggestions that have made the overall presentation clearer and more focused.

REFERENCES

Anderson, T. W. (1984), *An Introduction to Multivariate Statistical Analysis*, 2nd ed., Wiley, New York, 675 pp.

Banet, T. A., and Lebart, L. (1984), Local and partial principal components analysis (PAC) and correspondence analysis. COMPSTAT 1984. In *Proceedings in Computational Statistics*, Physica-Verlag, Vienna for IASC (International Association for Statistical Computing), pp. 113–118.

Bjerknes, J. (1966), A possible response of the atmospheric Hadley circulation to equatorial anomalies in ocean temperature. *Tellus* 18:820–829.

Conradsen, K., and Ersbøll, B. K. (1991), Data dependent orthogonal transformations of multichannel image data, Technical Report, IMSOR, Technical University of Denmark, Lyngby, 35 pp.

Conradsen, K., and Nielsen, A. A. (1991a), Remote sensing in forecasting agricultural statistics in Kenya, Danida, the Danish International Development Agency, Technical Report Contract No. 104.Dan.8/410, Institute of Mathematical Statistics and Operations Research, Technical University of Denmark, Lyngby.

Conradsen, K., and Nielsen, A. A. (1991b), Multivariate change detection in multispectral, multitemporal images. In *Abstracts and Notes from Seminar on Near Real-Time Remote Sensing for Land and Ocean Applications*, Eurimage and ESA/Earthnet, Rome, Italy.

Conradsen, K., Nielsen, B. K., and Thyrsted, T. (1985), A comparison of min/max autocorrelation factor analysis and ordinary factor analysis. In *Proceedings from Symposium in Applied Statistics*, Lyngby, Denmark, pp. 47–56.

Cooley, W. W., and Lohnes, P. R. (1971), *Multivariate Data Analysis*, Wiley, New York.

Ersbøll, B. K. (1989), Transformations and classification of remotely sensed data, Ph.D. thesis, Institute of Mathematical Statistics and Operations Research, Technical University of Denmark, Lyngby, 297 pp.

FAO (1987), *Yearbook of Fisheries Statistics and Commodities*, Rome: Food and Agricultural Organization (FAO) of the United Nations, Rome, Vol. 65, 369 pp.

Fiedler, P. C. (1984), Satellite observations of the 1982-83 El Niño along the U.S. Pacific Coast. *Science* 224:1251–1254.

Frouin, R., and Simpson, J. J. (1995), Radiometric calibration of GOES-7 VISSR solar channels. *Remote Sens. Environ.* 52:95–115.

Fung, T., and LeDrew, E. (1987), Application of principal components analysis to change detection. *Photogramm. Eng. Remote Sens.* 53:1649–1658.

Gallaudet, T. C., and Simpson, J. J. (1994), An empirical orthogonal function analysis of remotely sensed sea surface temperature variability and its relation to interior oceanic processes off Baja California. *Remote Sens. Environ.* 47:375–389.

Gong, P. (1993), Change detection using principal component analysis and fuzzy set theory. *Can. J. Remote Sens.* 19:22–29.

Graetz, D., Fisher, R., and Wilson, M. (1992), Looking back: the changing face of the Australian continent, 1972–1992, CSIRO/COSSA Publication No. 029, Canberra, 159 pp.

Green, A. A., Berman, M., Switzer, P., and Craig, M. D. (1988), A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. Geosci. Remote Sens.* 26:65–74.

...zumi, H., and Fujimura, S. (1992), Change detection ...om remotely sensed multi-temporal images using multiple ...gression. In *Proceedings from the 1992 International Geo...ce and Remote Sensing Symposium, IGARSS*, Houston, ...X, pp. 564–566.

...izumi, H., Chino, S., and Fujimura, S. (1994), A method ...change analysis with weight of significance using multi-...temporal, multi-spectral images. In *Proceedings SPIE 2315 European Symposium on Satellite Remote Sensing, Image and Signal Processing for Remote Sensing* (J. Desachy, Ed.), ...Rome, Italy, pp. 282–288.

...is, A. R., and Mason, I. M. (1992), An extension to the ...plit-window technique giving improved atmospheric correc...tion and total water vapour. *Int. J. Remote Sens.* 3:881–892.

...el, J. D., and Wallace, J. M. (1981), Planetary-scale atmo...spheric phenomena associated with southern oscillation. *...Mon. Weath. Rev.* 109:813–829.

...ling, H. (1933), Analysis of a complex of statistical vari...ables into principal components. *J. Ed. Psychol.* 24:417–441.

...ling, H. (1936), Relations between two sets of variates. *...ometrika* XXVIII:321–377.

...iston, L., Nelson, G. G., and Porto, R. W. (1979 and up...dates), Data extraction and calibration of TIROSN/NOAA A-G ...Radiometers, NOAA Technical Memorandum NESS 107, ...U.S. Department of Commerce, Washington, DC, 58 pp.

...okis, R., and Pritchard, J. (1976), Algorithms for correcting ...VHRR imagery for geometric distortion due to Earth curva...ture, Earth rotation, and space craft roll attitude, NOAA ...Technical Memorandum NESS 77, U.S. Department of ...Commerce, Washington, DC.

...Lynn, R. J. (1983), The 1982–83 warm episode in the Califor...nia Current. *Geophys. Res. Lett.* 10:1093–1095.

...Lynn, R. J., and Simpson, J. J. (1987), The California Current ...System—the seasonal variability of its physical characteris...tics. *J. Geophys. Res.* 92:12,947–12,966.

...Markham, B. L. (1985), The Landsat sensors spatial response. *...IEEE Trans. Geosci. Remote Sens.* 23:864–875.

...Matra Marconi Space (1994), *Earth Observation Spacecraft Di...rectory*, 2nd ed., Toulouse, France, 65 pp.

...McGowan, J. A. (1985), El Niño 1983 in the Southern Califor...nia Bight. In *El Niño Effects in the Eastern Subarctic Pacific Ocean* (W. S. Wooster and D. L. Fluharty, eds.), University ...of Washington Press, Seattle, pp. 166–184.

...Nielsen, A. A. (1994), Analysis of regularly and irregularly sam...pled spatial, multivariate and multi-temporal data, Ph.D. dis...sertation, IMM, Technical University of Denmark. Lyngby, ...189 pp. http://www.imm.dtu.dk/documents/users/aa/phd.

...Nielsen, A. A., and Conradsen, K. (1997), Multivariate alteration ...detection (MAD) in multispectral, bitemporal image data: a ...new approach to change detection studies, Technical Report 1997-11, IMM, Technical University of Denmark, Lyngby. http://www.imm.dtu.dk/documents/users/aa/tech-rep-1997-11/.

Nielsen, A. A., and Larsen, R. (1994), Restoration of GERIS data using the maximum noise fractions transform. In *Proceedings from the First International Airborne Remote Sensing Conference and Exhibition*, ERIM, Strasbourg, France, pp. 557–568.

Paulson, C. A., and Simpson, J. J. (1981), The temperature difference across the cool skin of the ocean. *J. Geophys. Res.* 86:11,044–11,053.

Preisendorfer, R. W. (1988), *Principal Component Analysis in Meteorology and Oceanography*, Elsevier, Amsterdam, 397 pp.

Quiroz, R. S. (1983), The climate of the "El Niño" winter of 1982–83: a season of extraordinary climatic anomalies. *Mon. Weath. Rev.* 111:1695–1706.

Rasmusson, E. M., and Wallace, J. M. (1983), Meteorological aspects of the El Niño/southern oscillation. *Science* 222: 1195–1202.

Richards, J. A. (1993), *Remote Sensing Digital Image Analysis: An Introduction*, 2nd ed., Springer-Verlag, Berlin, New York, 340 pp.

Simpson, J. J. (1983), Large-scale thermal anomalies in the California current during the 1982–83 El Niño. *Geophys. Res. Lett.* 10:937–940.

Simpson, J. J. (1984a), El Niño-induced onshore transport in the California current during the 1982–83 El Niño. *Geophys. Res. Lett.* 11:233–236.

Simpson, J. J. (1984b), A simple model of the 1982–83 Californian "El Niño." *Geophys. Res. Lett.* 11:237–240.

Simpson, J. J. (1992a), Response of the southern California current system to the mid-latitude North Pacific coastal warming events of 1982–83 and 1940–41. *Fish. Oceanogr.* 1:57–79.

Simpson, J. J. (1992b), Image masking using polygon fill and morphological operations. *Remote Sens. Environ.* 40:161–183.

Simpson, J. J., and Humphrey, C. (1990), An automated cloud screening algorithm for daytime AVHRR data. *J. Geophys. Res.* 95:13,459–13,481.

Simpson, J. J., and Yhann, S. (1994), Reduction of noise in AVHRR channel 3 data with minimum distortion. *IEEE Trans. Geosci. Remote Sens.* 32:315–328.

Simpson, J. J., Gobat, J. I., and Frouin, R. (1995), Improved destriping of GOES images using finite impulse response filters. *Remote Sens. Environ.* 52:15–35.

Switzer, P., and Green, A. A. (1984), Min/max autocorrelation factors for multivariate spatial imagery, Technical Report No. 6, Department of Statistics, Stanford University, Stanford, CA, 14 pp.

Yhann, S., and Simpson, J. J. (1995), Application of neural networks to AVHRR cloud segmentation. *IEEE Trans. Geosci. Remote Sens.* 33:590–604.