Multivariate Change Detection in Multispectral, Multitemporal Images

Knut Conradsen
Allan Aasbjerg Nielsen

IMSOR
Institute of Mathematical Statistics and Operations Research
Technical University of Denmark, Building 321
DK–2800 Lyngby, Denmark

Abstract

This paper introduces a new orthogonal transform the **multivariate change detection** (MCD) transform based on an established multivariate statistical technique **canonical correlation analysis**. The theory for canonical correlation analysis is sketched and modified to be more directly applicable in our context. As opposed to traditional univariate change detection schemes our scheme transforms two sets of multivariate observations (e.g. two multispectral satellite images aquired at different points in time) into a difference between two linear combinations of the original variables explaining maximal change (i.e. the difference explaining maximal variance) in all variables simultaneously. A case study using multispectral SPOT data from 1987 and 1989 covering coffee and pineapple plantations near Thika, Kiambu District, Kenya, shows the usefulness of this new concept.

1 Introduction

When analyzing changes in panchromatic imagery taken at different points in time it is customary to analyze the difference between two images.

If we have multivariate images with outcomes at a given pixel

$$\boldsymbol{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_k \end{bmatrix} \quad \text{resp.} \quad \boldsymbol{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_k \end{bmatrix}$$

then a naïve change detection transform would be

$$\boldsymbol{X} - \boldsymbol{Y} = \begin{bmatrix} X_1 - Y_1 \\ \vdots \\ X_k - Y_k \end{bmatrix}.$$

If we have seasonal changes in say vegetation cover this would cause shifts of "energy" from one wavelength to others and the changes may therefore be "smeared" out. This would make it difficult to actually see the changes in all the channels. The problem is obviously that the channels are correlated. We shall therefore try to overcome this problem by looking at linear combinations

$$\begin{aligned} \boldsymbol{a}'\boldsymbol{X} &= a_1X_1 + \cdots + a_kX_k \\ \boldsymbol{b}'\boldsymbol{Y} &= b_1Y_1 + \cdots + b_kY_k \end{aligned}$$

and then consider the difference between these, i.e.

$$\boldsymbol{a}'\boldsymbol{X} - \boldsymbol{b}'\boldsymbol{Y}.$$

For any choice of $\boldsymbol{a}$ and $\boldsymbol{b}$ this will give a measure of change. One could use pricipal components analysis on $\boldsymbol{X}$ to find an optimal $\boldsymbol{a}$ and on $\boldsymbol{Y}$ to find an optimal $\boldsymbol{b}$ (independent of $\boldsymbol{a}$). An improvement of this technique is to use principal components analysis on $\boldsymbol{X}$ and $\boldsymbol{Y}$ considered as *one* variable, cf. Fung and LeDrew (1987). This approach does not, however, guarantee a separation of $\boldsymbol{X}$ and $\boldsymbol{Y}$. It might as well give, say, differences between shortwave and longwave image bands. A better approach is to define an optimal set of $\boldsymbol{a}$ and $\boldsymbol{b}$ simultaneously. Emphasizing changes is the same as saying we want the difference to show a great variation or in other words that we want to maximize the variance $V(\boldsymbol{a}'\boldsymbol{X} - \boldsymbol{b}'\boldsymbol{Y})$. Now, trivially a multiplication of $\boldsymbol{a}$ and $\boldsymbol{b}$ with a constant c will multiply the variance with c^2. Therefore we must put some restrictions on $\boldsymbol{a}$ and $\boldsymbol{b}$, and natural restrictions are requesting unit variance of $\boldsymbol{a}'\boldsymbol{X}$ and $\boldsymbol{b}'\boldsymbol{Y}$.

The criterion would thus be

$$\text{maximize } V(\boldsymbol{a}'\boldsymbol{X} - \boldsymbol{b}'\boldsymbol{Y})$$

subject to the constraints

$$V(\boldsymbol{a}'\boldsymbol{X}) = 1 \quad \text{and} \quad V(\boldsymbol{b}'\boldsymbol{Y}) = 1.$$

Under these constraints we have

$$\begin{aligned} V(\boldsymbol{a}'\boldsymbol{X} - \boldsymbol{b}'\boldsymbol{Y}) &= V(\boldsymbol{a}'\boldsymbol{X}) + V(\boldsymbol{b}'\boldsymbol{Y}) - 2\,Cov(\boldsymbol{a}'\boldsymbol{X}, \boldsymbol{b}'\boldsymbol{Y}) \\ &= 2(1 - Corre(\boldsymbol{a}'\boldsymbol{X}, \boldsymbol{b}'\boldsymbol{Y})). \end{aligned}$$

Therefore, determining the combinations with maximum variance corresponds to having minimum correlation.

In the sequel we shall assume that $\boldsymbol{a}$ and $\boldsymbol{b}$ are chosen so that the correlation between $\boldsymbol{a}'\boldsymbol{X}$ and $\boldsymbol{b}'\boldsymbol{Y}$ is positive. Positive correlation may simply be obtained by a change of sign if necessary. Determining linear combinations with extreme correlations brings the theory of **canonical correlation analysis** into mind. Canonical correlation 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 the two linear combinations are called the first canonical variates. The second two linear combinations are the ones with the second 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 the two linear combinations are called the second canonical variates. Higher order canonical correlations and canonical variates are defined similarly. The technique is described in most standard textbooks on multivariate statistics, cf. e.g. Anderson (1984).

The main idea presented in this paper is now to try to modify the theory used in defining **canonical variates**. This could be viewed upon as a time analogue to the introduction of minimum/maximum autocorrelation factors in the spatial domain. Minimum/Maximum autocorrelation factor (MAF) analysis was first described by Switzer and Green, 1984. We shall therefore summarize the theory of canonical correlations and then modify the theorems so that they are more directly applicable in our context.

2 Canonical Correlation Analysis

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

$$\begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \end{bmatrix} \in N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = N(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix})$$

and we assume that $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\Sigma}_{22}$ (and $\boldsymbol{\Sigma}$) are non-singular.

We consider the conjugate eigenvectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p$ corresponding to the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_p$ of $\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$ with respect to $\boldsymbol{\Sigma}_{11}$, i.e.

$$(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} - \lambda_i \boldsymbol{\Sigma}_{11})\boldsymbol{a}_i = \mathbf{o}.$$

If we put

$$\boldsymbol{b}_i = \frac{1}{\sqrt{\lambda_i}} \boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{a}_i$$

we have

$$\begin{aligned} \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} &= \frac{1}{\sqrt{\lambda_i}} \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{a}_i \\ &= \sqrt{\lambda_i}\boldsymbol{\Sigma}_{21}\boldsymbol{a}_i \\ &= \lambda_i \boldsymbol{\Sigma}_{22}\boldsymbol{b}_i \end{aligned}$$

i.e. $\boldsymbol{b}_i$ is an eigenvector of $\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$ with respect to $\boldsymbol{\Sigma}_{22}$ corresponding to eigenvalue λ_i. If $p = q$ this will be all the eigenvalues and -vectors of $\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$. If $q > p$ then then last eigenvalue will be 0 with multiplicity $q - p$.

<u>Theorem</u>. Letting δ_{ij} be the Kronecker delta ($\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ otherwise) we have

$$\boldsymbol{a}_i'\boldsymbol{\Sigma}_{11}\boldsymbol{a}_j = \boldsymbol{b}_i'\boldsymbol{\Sigma}_{22}\boldsymbol{b}_j = \delta_{ij}$$

$$\boldsymbol{a}_i'\boldsymbol{\Sigma}_{12}\boldsymbol{b}_j = \sqrt{\lambda_i}\delta_{ij}.$$

<u>Proof</u>. The result for $\boldsymbol{a}_i$ follows by definition. We then obtain

$$\begin{aligned}
\boldsymbol{b}_i'\boldsymbol{\Sigma}_{22}\boldsymbol{b}_j &= \frac{1}{\sqrt{\lambda_i\lambda_j}}\boldsymbol{a}_i'\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{22}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{a}_j \\
&= \sqrt{\frac{\lambda_i}{\lambda_j}}\boldsymbol{a}_i'\boldsymbol{\Sigma}_{11}\boldsymbol{a}_j \\
&= \delta_{ij}.
\end{aligned}$$

Similarly we obtain

$$\begin{aligned}
\boldsymbol{a}_i'\boldsymbol{\Sigma}_{12}\boldsymbol{b}_j &= \frac{1}{\sqrt{\lambda_j}}\boldsymbol{a}_i'\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{a}_j \\
&= \sqrt{\lambda_j}\delta_{ij}.
\end{aligned}$$

□

We are now able to introduce the **canonical variates**

$$U_i = \boldsymbol{a}_i'\boldsymbol{X},\ i = 1, \ldots, p$$

$$V_i = \boldsymbol{b}_i'\boldsymbol{Y},\ i = 1, \ldots, p$$

and with an obvious choice of notation

$$\boldsymbol{U} = \boldsymbol{A}'\boldsymbol{X} \quad \text{and} \quad \boldsymbol{V} = \boldsymbol{B}'\boldsymbol{Y},$$

where

$$\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p) \text{ is } p \times p$$

$$\boldsymbol{B} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_p) \text{ is } q \times p.$$

We then have

<u>Theorem</u>. We consider the random variable $\boldsymbol{Z} = \boldsymbol{U} - \boldsymbol{V}$ and have that the dispersion matrix is

$$D(\boldsymbol{Z}) = D(\boldsymbol{U} - \boldsymbol{V}) = 2\begin{bmatrix} 1 - \sqrt{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 - \sqrt{\lambda_p} \end{bmatrix} = 2(\boldsymbol{I} - \boldsymbol{\Lambda}^{\frac{1}{2}}).$$

Proof. Straightforward

$$\begin{aligned} D(U-V) &= D(U)+D(V)-2Cov(U,V) \\ &= A'\Sigma_{11}A + B'\Sigma_{22}B - 2A'\Sigma_{12}B \\ &= I + I - 2\Lambda^{\frac{1}{2}}. \end{aligned} \qquad \square$$

3 A Minimizing Property of Canonical Variates

Normally the stepwise definition of canonical variates starts at the set with maximal correlation as mentioned in the introduction. From our point of view it will be more natural to start with the component of Z yielding the largest variance i.e. the canonical variates with the smallest correlation.

We assume that $c'X$ is independent of $U_{j+1}, \ldots, U_p$. We have

$$Cov(c'X, a_k'X) = c'\Sigma_{11}a_k = 0, \; k = j+1, \ldots, p.$$

Now, c may be written as $\gamma_1 a_1 + \cdots + \gamma_p a_p$ and this implies

$$c'\Sigma_{11}a_k = \gamma_k = 0, \; k = j+1, \ldots, p$$

i.e. $c'X$ may be written as

$$c'X = (\gamma_1 a_1' + \cdots + \gamma_j a_j')X = \gamma_1 U_1 + \cdots + \gamma_j U_j.$$

Similarly, if $d'Y$ is independent of $V_{j+1}, \ldots, V_p$ we may write

$$d'Y = \nu_1 V_1 + \cdots + \nu_j V_j.$$

We now want to minimize the absolute value of the correlation between $c'X$ and $d'Y$ i.e. minimize

$$\frac{c'\Sigma_{12}d}{c'\Sigma_{11}d\Sigma_{22}d} = \frac{\gamma_*'\Lambda_*^{\frac{1}{2}}\nu_*}{\sqrt{\gamma_*'\gamma_*\nu_*'\nu_*}} \tag{1}$$

or equivalently minimize

$$(\gamma_1, \ldots, \gamma_j)\begin{bmatrix} \sqrt{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{\lambda_j} \end{bmatrix}\begin{bmatrix} \nu_1 \\ \vdots \\ \nu_j \end{bmatrix} = \gamma_*'\Lambda_*^{\frac{1}{2}}\nu_*$$

subject to the constraints

$$\gamma_*'\gamma_* = 1 \quad \text{and} \quad \nu_*'\nu_* = 1.$$

We introduce the Lagrange expression

$$F = \gamma_*' \Lambda_*^{\frac{1}{2}} \nu_* - \frac{\alpha}{2}(\gamma_*'\gamma_* - 1) - \frac{\beta}{2}(\nu_*'\nu_* - 1),$$

and have at optimum

$$\frac{\partial F}{\partial \gamma_*} = \Lambda_*^{\frac{1}{2}} \nu_* - \alpha\gamma_* = 0 \Leftrightarrow \alpha\gamma_* = \Lambda_*^{\frac{1}{2}} \nu_*$$

$$\frac{\partial F}{\partial \nu_*} = \Lambda_*^{\frac{1}{2}} \gamma_* - \beta\nu_* = 0 \Leftrightarrow \beta\nu_* = \Lambda_*^{\frac{1}{2}} \gamma_*.$$

We insert this in Equation 1 and obtain the expression

$$\frac{|\beta|}{\beta} \frac{\gamma_*'\Lambda_*\gamma_*}{\sqrt{\gamma_*'\gamma_*\gamma_*'\Lambda_*\gamma_*}} = \operatorname{sign}(\beta)\sqrt{\frac{\gamma_*'\Lambda_*\gamma_*}{\gamma_*'\gamma_*}}.$$

The square of this expression is

$$\frac{\lambda_1\gamma_1^2 + \cdots + \lambda_j\gamma_j^2}{\gamma_1^2 + \cdots + \gamma_j^2}$$

and this has a minimum for $\gamma_1 = \cdots = \gamma_{j-1} = 0$ and $\gamma_j = 1$, which corresponds to choosing $\boldsymbol{c}'\boldsymbol{X}$ as U_j. We have now proven the following

Theorem. The canonical variates have the property that the j'th canonical variate shows minimal correlation amongst linear combinations independent of the previous $p-j$ least correlated canonical variates. In the case $q > p$ correlations between any of the U's and the projection on the eigenvectors of $\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$ with respect to $\boldsymbol{\Sigma}_{22}$ corresponding to the eigenvalue 0 are exactly equal to $\mathbf{o}$.

4 The Multivariate Change Detection (MCD) Transform

Having established this result we are now ready to define the multivariate change detection (MCD) transform as

$$\begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \end{bmatrix} \rightarrow \boldsymbol{U} - \boldsymbol{V} = \boldsymbol{A}'\boldsymbol{X} - \boldsymbol{B}'\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{a}_1'\boldsymbol{X} - \boldsymbol{b}_1'\boldsymbol{Y} \\ \vdots \\ \boldsymbol{a}_p'\boldsymbol{X} - \boldsymbol{b}_p'\boldsymbol{Y} \end{bmatrix},$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ etc. are defined as above i.e. $\boldsymbol{A}$ and $\boldsymbol{B}$ are the defining coefficients from a standard canonical correlation analysis. The MCD transform has the very important property that if we consider linear combinations of the two sets of variables that are positively correlated then the p'th difference will show maximum variance among such variables. The $(p-j)$'th difference will show maximum variance subject to the constraint that it should be uncorrelated with the previous j ones etc. In this way we may sequentially extract uncorrelated difference images where each new image shows maximum change under the constraint of being uncorrelated with the previous ones.

If $p < q$ then the projection of $\boldsymbol{Y}$ on the eigenvectors corresponding to the eigenvalues 0 will be independent of $\boldsymbol{X}$. That part may of course be considered the extreme case of multivariate change detection.

5 Case Study – SPOT XS Data, Kenya

A 512×512 sub-scene from multispectral 1987 and 1989 SPOT XS data is used as a pilot area. This area contains economically important coffee and pineapple fields near Thika, Kiambu District, Kenya. In Plate 1 we show false colour composites of multispectral SPOT scenes from 1987 and 1989. The area is dominated by large pineapple fields to the northeast and coffee fields to the northwest. To the south is Thika town. We also show the naïve change detection image. The major differences are due to the changes primarily in the pineapple fields. Pineapple is a triennial crop and therefore we observe changes from one year to another. Since the changes are connected to change in vegetation it seems natural to study the change in the normalized difference vegetation index

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{R}}{\mathrm{NIR} + \mathrm{R} + 1}$$

where NIR is the near-infrared channel (XS3) and R is the red channel (XS2). The philosophy behind the NDVI is that healthy green matter reflects the near-infrared light strongly and absorbs the red light. Therefore the NDVI will be large in vegetated areas and small in non-vegetated areas. An interesting study on NDVI change detection based on NOAA AVHRR decade (10 day) GAC data from Sudan covering a period of nearly 7 years the result of which was presented as a video was made by Stern (1989). In Plate 1 we finally show the 1989 NDVI as red and 1987 NDVI as cyan (causing no change to be represented by a grey scale). This image enhances the differences between fields in a much clearer way than the naïve change detection image. However, this enhancement of differences may also be due to differences between fields in the *same* season. In Plate 2 we show the canonical variates for the 1987 and the 1989 data. Enhancing many interesting features it readily follows that these images are useful by themselves. We also show the MCD_3 variable, i.e. the difference between linear combinations from the two years showing maximal difference. In Plate 2 we finally show all three MCDs. An inspection of these images and a comparison with the naïve change detection image show that there is a much better distinction between different types of changes. In the naïve change detection image cyan is dominating but in the MCD image we see that a much better discrimination has been achieved. We therefore conclude that the MCD transformation is a useful extension of more naïve multivariate change detection schemes.

6 Annex

In this annex we show the numerical results from the canonical correlation analysis of the two scenes. We first report the basic statistics

Means and Standard Deviations

	1987		1989	
	Mean	Std Dev	Mean	Std Dev
XS1	45.00	5.40	32.27	4.79
XS2	36.86	7.12	22.88	4.87
XS3	74.15	12.55	62.33	10.66

Correlations Among the Original Variables

	1987			1989		
	XS1	XS2	XS3	XS1	XS2	XS3
XS1	1.0000	0.9057	−0.3336	0.5116	0.3955	−0.0082
1987 XS2	0.9057	1.0000	−0.4196	0.4352	0.4140	−0.0381
XS3	−0.3336	−0.4196	1.0000	−0.3477	−0.2644	0.2492
XS1	0.5116	0.4352	−0.3477	1.0000	0.8866	−0.2609
1989 XS2	0.3955	0.4140	−0.2644	0.8866	1.0000	−0.4191
XS3	−0.0082	−0.0381	0.2492	−0.2609	−0.4191	1.0000

It is noted that the correlation structure is basically the same in the two years considered.

Canonical Correlation Analysis

	Canonical Correlation	Approx Standard Error	Squared Canonical Correlation
1	0.6505	0.0011	0.4232
2	0.4024	0.0016	0.1619
3	0.2403	0.0018	0.0577

Test of H_0: The canonical correlations in the current row and all that follow are zero

	Likelihood Ratio	Approx F	Numerator DF	Denominator DF	Pr > F
1	0.4555	27039	9	637975	0.0
2	0.7897	16425	4	524278	0.0
3	0.9423	16058	1	262140	0.0

Standardized canonical coefficients

	1987 CAN1	1987 CAN2	1987 CAN3	1989 CAN1	1989 CAN2	1989 CAN3
XS1	-1.8816	-0.6862	1.2787	-2.0441	-0.8151	0.4247
XS2	1.5328	1.6894	-0.9417	1.5120	1.7877	-0.4430
XS3	0.5938	0.4081	0.8441	0.2616	0.6431	0.9063

Canonical Structure

Correlations between original variables and canonical variables

	1987 CAN1	1987 CAN2	1987 CAN3	1989 CAN1	1989 CAN2	1989 CAN3
1987 XS1	-0.6915	0.7078	0.1442	-0.4499	0.2848	0.0347
1987 XS2	-0.4206	0.8967	-0.1377	-0.2736	0.3609	-0.0331
1987 XS3	0.5784	-0.0719	0.8126	0.3763	-0.0289	0.1952
1989 XS1	-0.5021	0.2423	-0.0491	-0.7718	0.6021	-0.2045
1989 XS2	-0.2667	0.3201	-0.1072	-0.4099	0.7955	-0.4462
1989 XS3	0.1050	0.0429	0.2357	0.1613	0.1067	0.9811

Canonical Redundancy Analysis

Standardized variance of 1987 XS

	Explained by Their Own Canonical Variables: Proportion	Their Own Canonical Variables: Cumulative Proportion	The Opposite Canonical Variables: Canonical R^2	The Opposite Canonical Variables: Proportion	The Opposite Canonical Variables: Cumulative Proportion
1	0.3299	0.3299	0.4232	0.1396	0.1396
2	0.4368	0.7667	0.1619	0.0707	0.2103
3	0.2333	1.0000	0.0577	0.0135	0.2238

Standardized variance of 1989 XS

	Explained by Their Own Canonical Variables: Proportion	Their Own Canonical Variables: Cumulative Proportion	The Opposite Canonical Variables: Canonical R^2	The Opposite Canonical Variables: Proportion	The Opposite Canonical Variables: Cumulative Proportion
1	0.2632	0.2632	0.4232	0.1114	0.1114
2	0.3356	0.5988	0.1619	0.0543	0.1658
3	0.4012	1.0000	0.0577	0.0232	0.1889

Squared multiple correlations (R^2) between 1987 XS and the first M canonical variates of 1989 XS, and squared multiple correlations (R^2) between 1989 XS and the first M canonical variates of 1987 XS

	R^2(1987 XS, 1989 CAN)			R^2(1989 XS, 1987 CAN)		
M	1	2	3	1	2	3
XS1	0.2024	0.2835	0.2847	0.2521	0.3108	0.3132
XS2	0.0749	0.2051	0.2062	0.0711	0.1736	0.1851
XS3	0.1416	0.1424	0.1805	0.0110	0.0129	0.0684

Thus the canonical variates for the 1987 XS data are

$$\begin{bmatrix} \text{CAN1} \\ \text{CAN2} \\ \text{CAN3} \end{bmatrix} = \begin{bmatrix} -1.8816 & 1.5238 & 0.5938 \\ -0.6862 & 1.6894 & 0.4081 \\ 1.2787 & -0.9417 & 0.8441 \end{bmatrix} \begin{bmatrix} (\text{XS1} - 45.00)/5.40 \\ (\text{XS2} - 36.86)/7.12 \\ (\text{XS3} - 74.15)/12.6 \end{bmatrix}$$

and the canonical variates for the 1989 XS data are

$$\begin{bmatrix} \text{CAN1} \\ \text{CAN2} \\ \text{CAN3} \end{bmatrix} = \begin{bmatrix} -2.0441 & 1.5120 & 0.2616 \\ -0.8151 & 1.7877 & 0.6431 \\ 0.4247 & -0.4430 & 0.9063 \end{bmatrix} \begin{bmatrix} (\text{XS1} - 32.27)/4.79 \\ (\text{XS2} - 22.88)/4.87 \\ (\text{XS3} - 62.33)/10.7 \end{bmatrix}.$$

Acknowledgements

The authors wish to thank Bjarne Kjær Ersbøll, IMSOR, for many good discussions on multivariate statistics and analysis of spatial data.

The work reported was funded by Danida, the Danish International Development Agency, under grant No. 104.Dan.8/410.

References

T.W. Anderson (1984): *An Introduction to Multivariate Statistical Analysis.* 2nd Edition. John Wiley and Sons, 675 pp.

T. Fung and E. LeDrew (1987): Application of Principal Components Analysis to Change Detection. *Photogrammetric Engineering and Remote Sensing*, Vol. 53, No. 12, pp. 1649–1658.

M. Stern (1989): *pers. comm.*, Space Media Networks/IMSOR.

P. Switzer and A.A. Green (1984): *Min/Max Autocorrelation Factors for Multivariate Spatial Imagery.* Technical Report No. 6. Department of Statistics, Stanford University, 10 pp.

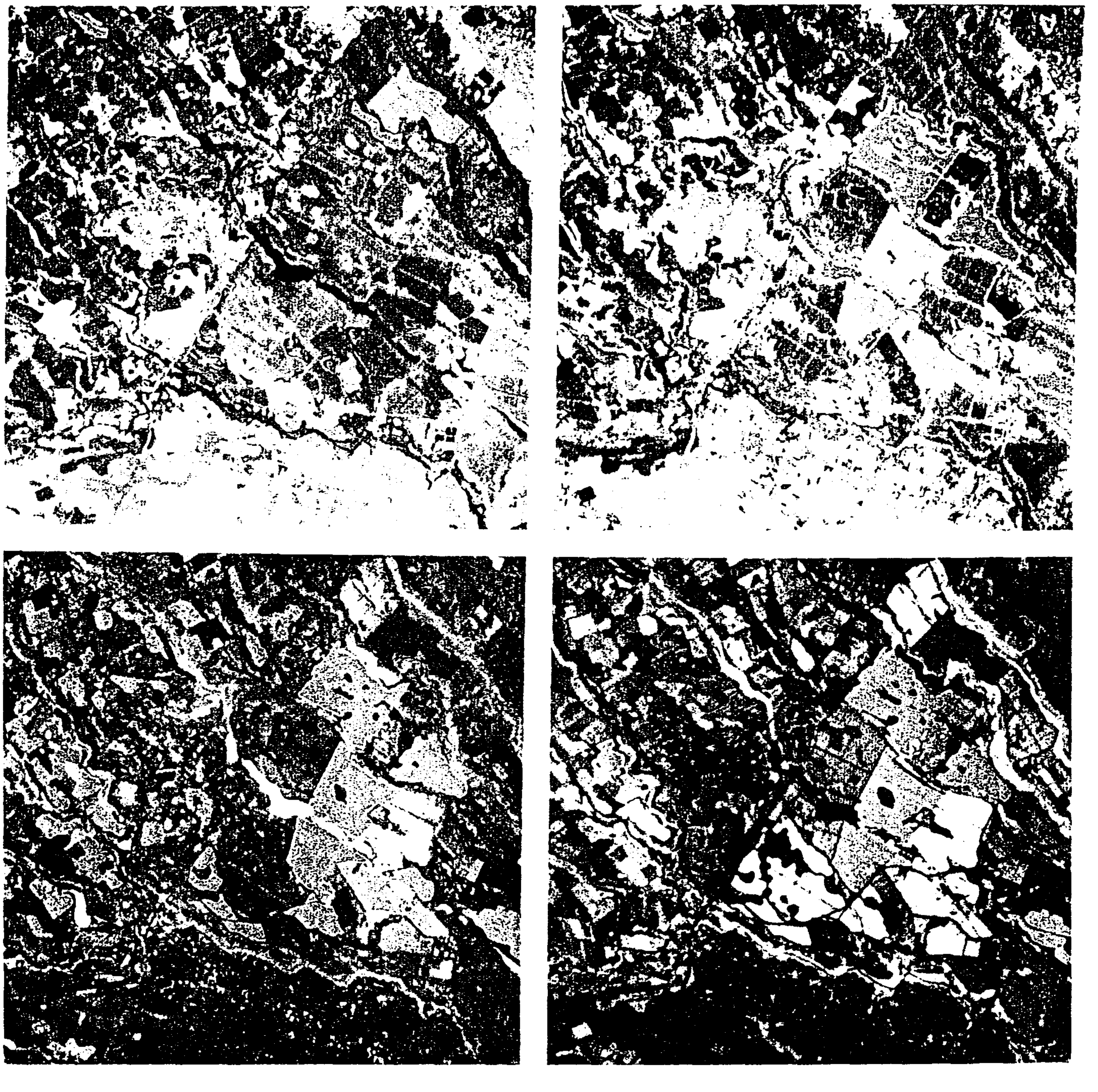

Plate 1 (left-to-right top-to-bottom): (a) 1987 SPOT XS, (b) 1989 SPOT XS, (c) Naïve Change SPOT XS, (d) Change SPOT NDVI.

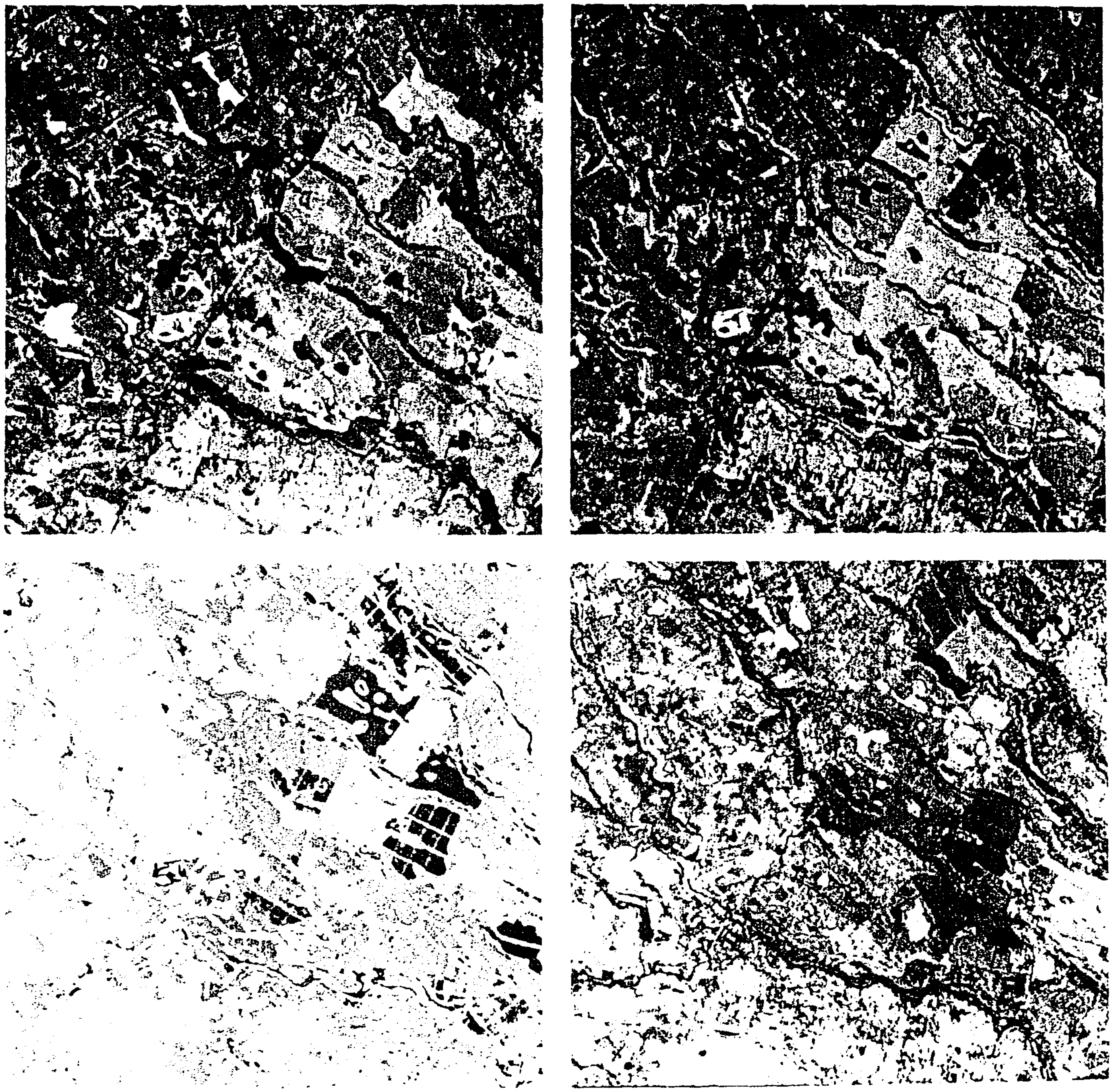

Plate 2 (right-to-left top-to-bottom): (a) 1987 SPOT Can. Var., (b) 1989 SPOT Can. Var., (c) SPOT MCD_3, (d) SPOT MCD.