CHANGE DETECTION IN MULTI-SPECTRAL, BI-TEMPORAL SPATIAL DATA USING ORTHOGONAL TRANSFORMATIONS

Allan Aasbjerg Nielsen
Institute of Mathematical Modelling
Technical University of Denmark, Building 321
DK–2800 Lyngby, Denmark
aa@imm.dtu.dk

ABSTRACT

This paper introduces the multivariate alteration detection (MAD) transformation which is based on the established canonical correlation analysis. The MAD transformation is invariant to linear scaling. It is therefore insensitive to for example differences in gain settings in a measuring device. Other multivariate change detection schemes described are principal component type analysis of simple difference images. A case study with SPOT HRV data using simple linear stretching and masking of the change images shows the usefulness of the new MAD change detection scheme.

INTRODUCTION

When analysing changes in panchromatic images taken at different points in time it is customary to analyse the difference between two images, possibly after some normalisation. The idea is of course that areas with no or little changes come out with zero or low absolute values and areas with large changes come out with large absolute values in the difference image. If we have two multivariate images with variables at a given pixel written as vectors

$$\boldsymbol{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_k \end{bmatrix} \quad \text{and} \quad \boldsymbol{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_k \end{bmatrix} \tag{1}$$

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

$$\boldsymbol{X} - \boldsymbol{Y} = \begin{bmatrix} X_1 - Y_1 \\ \vdots \\ X_k - Y_k \end{bmatrix}. \tag{2}$$

If our image data have more than three channels it is difficult to visualise change in all channels simultaneously. To overcome this problem and to concentrate information on change, linear transformations of the image data that optimise some design criterion can be considered. A linear transformation that will maximise a measure of change in the simple multispectral difference image is one that maximises deviations from no change for instance the variance

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

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 to request that $\boldsymbol{v}$ is a unit vector, $\boldsymbol{v}^T\boldsymbol{v} = 1$. This amounts to finding principal components of the simple difference images. Other change detection schemes based on simple difference images include factor analysis, min/max autocorrelation factor (MAF) analysis (Switzer and Green, 1984) and decorrelation methods.

A more parameter rich measure of change that allows different coefficients for $\boldsymbol{X}$ and $\boldsymbol{Y}$ and different number of spectral bands in the two sets, p and q respectively $(p \leq q)$, are linear combinations ($\boldsymbol{X}$ and $\boldsymbol{Y}$ here denote variables with mean value equal to zero)

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

and the difference between them $\boldsymbol{a}^T\boldsymbol{X} - \boldsymbol{b}^T\boldsymbol{Y}$. This measure also accounts for situations where the spectral bands are not the same but cover different spectral regions, for instance if one set of data comes from Landsat Thematic Mapper (TM) and the other set comes from SPOT High Resolution Visible (HRV). In this case one must be more cautious when interpreting the multivariate difference as multivariate change.

To find $\boldsymbol{a}$ and $\boldsymbol{b}$ Fung and LeDrew (1987) use principal components (PC) analysis on $\boldsymbol{X}$ and $\boldsymbol{Y}$ considered as one vector variable. This approach defines $\boldsymbol{a}$ and $\boldsymbol{b}$ simultaneously but the method does not have a clear design criterion. 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. 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 a different number 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 maximise the variance, this time $\text{Var}\{\boldsymbol{a}^T\boldsymbol{X} - \boldsymbol{b}^TY\}$. A multiplication of $\boldsymbol{a}$ and $\boldsymbol{b}$ with a constant c will multiply the variance with c^2. Therefore we 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: maximise $\text{Var}\{\boldsymbol{a}^T\boldsymbol{X} - \boldsymbol{b}^T\boldsymbol{Y}\}$ subject to the constraints $\text{Var}\{\boldsymbol{a}^T\boldsymbol{X}\} = \text{Var}\{\boldsymbol{b}^T\boldsymbol{Y}\} = 1$. Under these constraints we have

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

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 correlation (≥ 0). Determination of linear combinations with extreme correlations brings the theory of canonical correlation analysis to mind. This technique is described in most textbooks on multivariate statistics (good references are Cooley and Lohnes, 1971 and Anderson, 1984).

THE MULTIVARIATE ALTERATION DETECTION (MAD) TRANSFORMATION

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

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

where $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$ are the defining coefficients from a standard canonical correlation analysis. $\boldsymbol{X}$ and $\boldsymbol{Y}$ are vectors with mean zero. The dispersion matrix of the MAD variates is

$$\mathrm{D}\{\boldsymbol{a}^T\boldsymbol{X} - \boldsymbol{b}^T\boldsymbol{Y}\} = 2(\boldsymbol{I} - \boldsymbol{\Lambda}^{1/2}) \tag{7}$$

where $\boldsymbol{I}$ is the $p \times p$ unit matrix and $\boldsymbol{\Lambda}^{1/2}$ is a $p \times p$ 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 p'th 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 $\boldsymbol{Y}$ 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. As opposed to the principal components the MAD variates are invariant to linear scaling, which means that they are not sensitive to for example gain settings of a measuring device.

This type of multivariate change detection technique was first sketched in Conradsen and Nielsen (1991a, 1991b). Multivariate change detection techniques are also described in Hanaizumi and Fujimura (1992), and in Hanaizumi, Chino and Fujimura (1994) who work with multiple regression and canonical correlation methods applied to specific change detection. In a study methodologically related to the MAD transformation Shettigara and McGilchrist (1989) use a hybrid canonical correlation/principal components technique to enhance uncorrelated parts of Landsat TM equivalents of ATM data in gold exploration.

As a post-processing scheme for data with many spectral bands and/or situations where the first MADs are noisy the MAF transformation is useful.

For other descriptions of the transformation described here see Nielsen (1994) which describes the MAD transformation and introduces the MAF post-processing of the MADs, Conradsen and Nielsen (1995) in which the MAD transformation is described in more detail and from which the case study below is taken, and Simpson and Nielsen (1995) in which one case study very successfully applies the MAD transformation on pre-processed AVHRR data (correction for atmospheric water vapor attenuation, and cloud- and land-masking) in an oceanographic case study on ENSO-related mid-latitude warming events where principal components analysis of simple difference images fail. Other (land) cases in Simpson and Nielsen (1995) deal with urbanisation and tropical deforestation using Landsat MSS data.

CASE STUDY - SPOT HRV DATA

The applicability of the MAD transformation to multivariate and bi-temporal change detection studies is demonstrated in a case using SPOT-1 High Resolution Visible (HRV) data covering economically important coffee and pineapple plantations near Thika in the Kiambu District, Kenya. The data consist of three spectral bands with 512×512 20 meter pixels. The acquisition dates are 5 February 1987 and 12 February 1989. The intention of this case study is to illustrate the method and not to accurately assess the actual change on the ground.

The original HRV data are not shown. Figures 1 to 3 are stretched linearly from mean minus three standard deviations to mean plus three standard deviations. Areas that are very bright or very dark are maximum change areas. Figure 1 shows the simple difference images ("1989 minus 1987", HRV bands 1 to 3, also termed XS1 to XS3). Figure 2 shows the principle components of the simple difference images (PC1 to PC3). Figure 3 shows the MADs of the original HRV images (MAD1 to MAD3). Figure 4 shows RGB images of simple difference PCs one (red), two (green), and three (blue). In the left image the three components are stretched linearly from mean minus three standard deviations to mean plus three standard deviations. All areas in bright colours (including black and white) are areas of change. In the right image the absolute values of the three components are stretched linearly between mean (before taking absolute values) plus 1.5 standard deviations and mean plus 3.0 standard deviations. Values below mean plus 1.5 standard deviations are saturated in black and values above mean plus 3.0 standard deviations are saturated in the appropriate colour. In this case black means no change. Figure 5 shows RGB images of MAD components one (red), two (green), and three (blue) with the same stretching and masking.

For all change detection schemes an interpretation of the resulting change images can be based on correlations between the transformed variates and the original data. Table 1 shows the correlations between the MADs and the original SPOT HRV bands. MAD1 is positively correlated with 1987 XS3 and uncorrelated with 1987 XS2. MAD1 is negatively correlated with 1989 XS3 and positively correlated with 1989 XS2. Therefore MAD1 is a sort of vegetation index change detector. MAD2 is a change detector of the weighted means of XS1 and XS2, presumably a non-vegetation related change detector. MAD3 seems to be a non-vegetation related change detector also. MAD3 has its highest correlations (with opposite signs) with XS1 in both years and is therefore likely to represent change in atmospheric conditions. It seems that the MAD transformation has separated change due to atmospheric conditions and change related to vegetation in each end of the change "spectrum". However, the fact that MAD3 has a relatively high correlation with 1987 XS3 makes this conclusion a little uncertain.

	MAD1	MAD2	MAD3
XS1	0.0889	0.3869	0.2891
1987 XS2	–0.0849	0.4902	0.1758
XS3	0.5008	–0.0393	–0.2418
XS1	0.1260	–0.3291	–0.3226
1989 XS2	0.2750	–0.4348	–0.1714
XS3	–0.6047	–0.0583	0.0674

Table 1: Correlations between MADs and SPOT HRV bands

Simple No Change Simulation

As a simple simulation of a situation with no change in all bands we pad the two 512×512 scenes described above into the central part of 600×600 backgrounds with values 0 in all bands in both years. Change between the two 600×600 scenes is estimated by means of 1) simple differences, 2) principal components of simple differences (based on covariance matrix), 3) principal components of simple differences (based on correlation matrix), 4) varimax rotated

Figure 1: Simple difference images

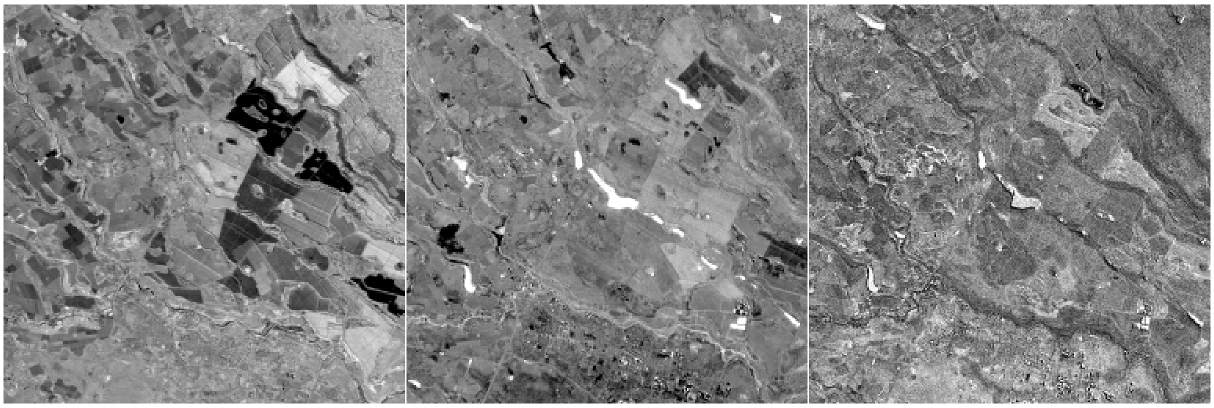

Figure 2: Principle components of simple difference images

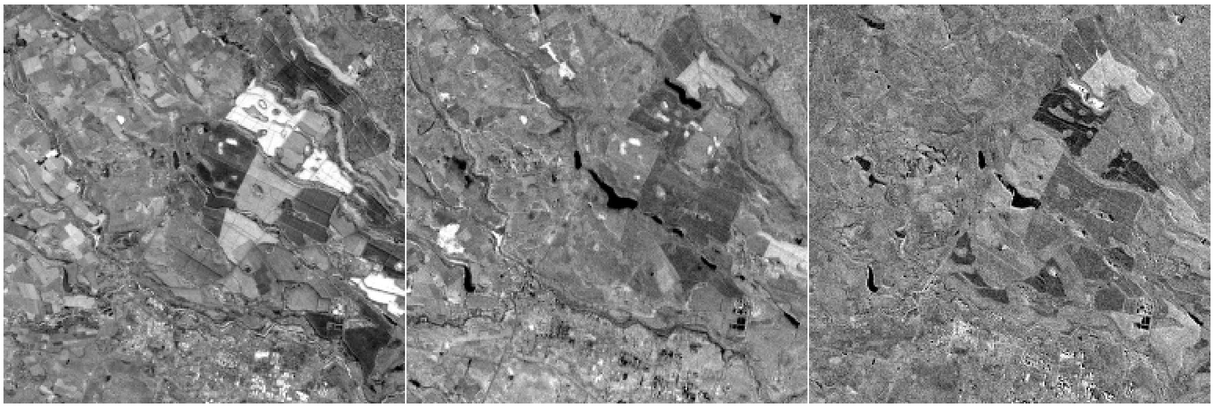

Figure 3: MAD images

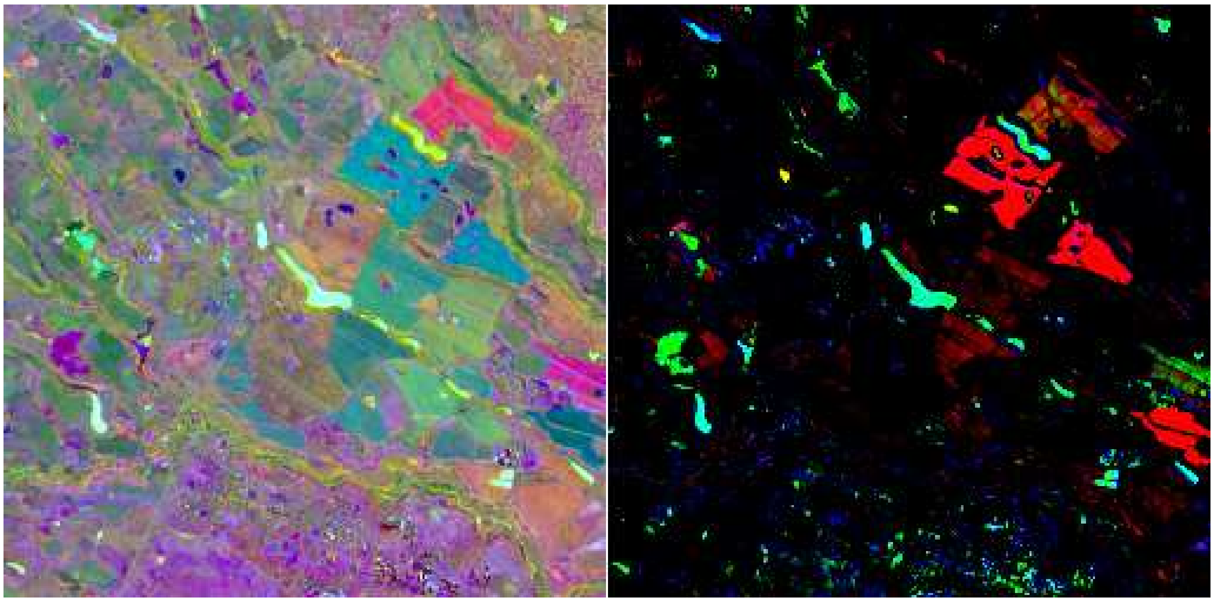

Figure 4: The three simple differences' PCs as RGB, left: areas in bright colours (including black and white) are areas of change; right: areas in bright colours are areas of change, black means no change

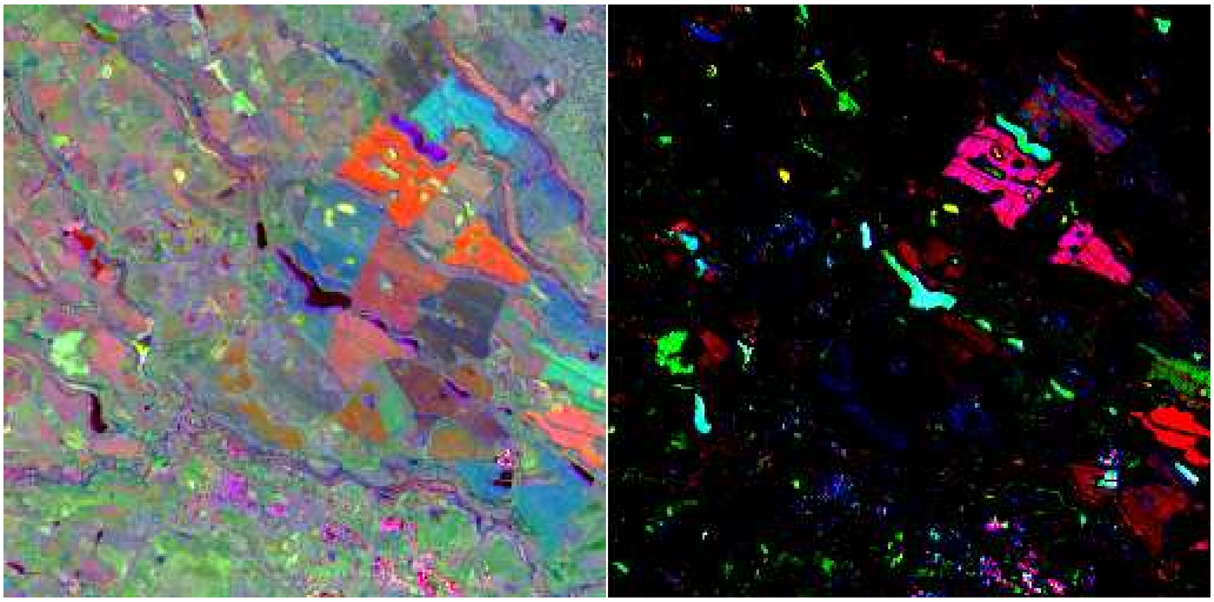

Figure 5: The three MAD components as RGB, left: areas in bright colours (including black and white) are areas of change; right: areas in bright colours are areas of change, black means no change

factors of simple differences, 5) MAFs of simple differences, and 6) MADs. Change detected in the region with no change (the 44 pixels broad edge around the actual image data) as given by standardised values of the results from the above change detection methods shows that MAD is the only multivariate technique tested here that perform well in this situation. All other methods give change values much higher than zero, see Table 2

Band/Component	1	2	3
1) Simple diff.	0	0	0
2) PC diff. (cov)	0.86	1.05	0.08
3) PC diff. (corr)	1.32	0.35	–0.04
4) Factors diff.	–1.23	0.58	0.07
5) MAF diff.	1.32	0.31	0.12
6) MAD	0.05	–0.03	–0.04

Table 2: Change detected (standardised scores) in no change region, value should be 0

The value in this region in the simple difference images is zero. Therefore the change detected here by most multivariate methods is due to the subtraction of the mean value of the entire image before calculating the relevant linear combinations of the original bands. This can easily be identified in this very simple situation, but had the no-change pixels been scattered as several fields inside the image this would not have been possible and the statistics for calculating the various transformations would have been exactly the same. Therefore the situation simulated though very simple is indeed a realistic one.

CONCLUSIONS

The proposed MAD transformation is based on the established canonical correlation analysis technique. It seems natural to base a change detection scheme on differences between transformed, orthogonal images that show maximum similarity as measured by correlation. The MAD transformation is invariant to affine transformations. It is therefore invariant to for instance differences in offset and gain in scanner electronics. This may sometimes be a desired property. If on the other hand, the change one wants to detect is one of simple shifts in level, the MAD transformation will not detect this. The MAD transformation detects more subtle change. A simple example with simulated no change regions shows that the MAD transformation is the only multivariate scheme among the ones tested that actually shows no change.

ACKNOWLEDGEMENTS

I wish to thank Professor Knut Conradsen for his cooperation on the development of the MAD concept. The calculations were performed by means of software written by Dr. Rasmus Larsen on my initiative and under my supervision. Also, I wish to thank Dr. James J. Simpson, Scripps Institution of Oceanography, for his immediate interest in the MAD scheme. Funding from Danida, the Danish International Development Agency, and the European Commission is highly appreciated.

REFERENCES

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

Cooley, W.W. and Lohnes, P.R. (1971): *Multivariate Data Analysis*. John Wiley & Sons.

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.

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. and Nielsen, A.A. (1995): Multivariate Alteration Detection (MAD) in Multispectral, Bi-temporal Image Data: A New Approach to Change Detection. Submitted.

Fung, T. and LeDrew, E. (1987): Application of Principal Components Analysis to Change Detection. *Photogrammetric Engineering and Remote Sensing*, **53**(12), pp. 1649–1658.

Gong, P. (1993): Change Detection Using Principal Component Analysis and Fuzzy Set Theory. *Canadian Journal of Remote Sensing*, **19**(1), pp. 22–29.

Hanaizumi, H., Chino, S. and Fujimura, S. (1994): A Method for Change Analysis with Weight of Significance Using Multi-temporal, Multi-spectral Images. In Desachy, J. (Ed.) *Proceedings SPIE 2315 European Symposium on Satellite Remote Sensing, Image and Signal Processing for Remote Sensing*, Rome, Italy, pp. 282–288.

Hanaizumi, H. and Fujimura, S. (1992): Change Detection from Remotely Sensed Multitemporal Images Using Multiple Regression. In *Proceedings from the 1992 International Geoscience and Remote Sensing Symposium, IGARSS*, Houston, Texas, USA, pp. 564–566.

Nielsen, A.A. (1994): *Analysis of Regularly and Irregularly Sampled Spatial, Multivariate, and Multi-temporal Data*. Ph.D. Thesis. Institute of Mathematical Modelling, Technical University of Denmark.

Shettigara, V.K. and McGilchrist, C.A. (1989): A Principal Component and Canonical Correlation Hybrid Technique for Change Detection in Two-Image Sets. In R.F. Barrett (Ed.) *ASSPA 89, signal processing, theories, implementations and applications*, Adelaide, Australia, pp. 47–52.

Simpson, J.J. and Nielsen, A.A. (1995): Accurate Detection of Patterns of Spatial and Temporal Change in Satellite Data Using Multivariate Statistical Transformations. Submitted.

Switzer, P. and Green, A.A. (1984): *Min/Max Autocorrelation Factors for Multivariate Spatial Imagery*. Technical Report. Department of Statistics, Stanford University.