Multi- and Hyperspectral Remote Sensing Change Detection with Generalized Difference Images by the IR-MAD Method

Allan A. Nielsen
Technical University of Denmark
Informatics and Mathematical Modelling
DK-2800 Lyngby, Denmark
Email: aa@imm.dtu.dk

Morton J. Canty
Jülich Research Center
Systems Analysis and Technology Evaluation
D-52425 Jülich, Germany
Email: m.canty@fz-juelich.de

Abstract— **Change detection methods for multi- and hypervariate data aim at identifying differences in data acquired over the same area at different points in time. In this contribution an iterative extension to the multivariate alteration detection (MAD) transformation for change detection is sketched and applied. The MAD transformation is based on canonical correlation analysis (CCA), which is an established technique in multivariate statistics. The extension in an iterative scheme seeks to establish an increasingly better background of no-change against which to detect change. This is done by putting higher weights on observations of no-change in the calculation of the statistics for the CCA. The differences found may be due to noise or differences in (atmospheric etc.) conditions at the two acquisition time points. To prevent a change detection method from detecting uninteresting change due to noise or arbitrary spurious differences the application of regularization, also known as penalization, and other types of robustification of the change detection method may be important especially when applied to hyperspectral data. Among other things results show that the new iterated scheme does give a better no-change background against which to detect change than the original, non-iterative MAD method and that the IR-MAD method depicts the change detected in less noisy components.**

I. INTRODUCTION

This contribution focuses on construction of more general difference images than simple differences in multivariate change detection. This is done via an iterated version [1] of the canonical correlation analysis (CCA) [2] based multivariate alteration detection (MAD) method [3] that could, moreover, be combined with an expectation-maximization (EM) based method for determining thresholds for differentiating between change and no-change in the difference images, and for estimating the variance-covariance structure of the no-change observations [4], [5]. The variances can be used to establish a single change/no-change image based on the general multivariate difference image. The resulting imagery from MAD based change detection is invariant to linear and affine transformations of the input including, e.g., affine corrections to normalize data between the two acquisition time points. This is an enormous advantage over other multivariate change detection methods. The resulting single change/no-change image can be used to establish both change regions and to extract observations with which a fully automated orthogonal regression analysis based normalization of the multivariate data between the two points in time can be developed [6].

Results (not shown here) from partly simulated multivariate data indicate an improved performance of the iterated scheme over the original MAD method [1]. Also, a few comparisons with established methods for calculation of robust statistics for the CCA indicate that the scheme suggested here performs better, see also [7].

Regularization issues typically important in connection with the analysis of hyperspectral data are dealt with in [8]–[10] and briefly mentioned here.

II. THE MAD TRANSFORMATION

Band-wise simple differences for change detection make sense only when the data are calibrated or at least when the data at the two points in time are normalized to a common zero and scale.

The so called MAD variates consist of differences (in reverse order) between canonical variates from CCA. The canonical variates can be found by solving this generalized eigenvalue problem for geometrically co-registered $p \times 1$ data X from one point in time and $q \times 1$ data Y from another point in time $(p \geq q)$

$$\begin{bmatrix} 0 & \Sigma_{12} \\ \Sigma_{21} & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \rho \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \quad (1)$$

or

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = (\rho + 1) \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}. \quad (2)$$

Σ_{11} is the $p \times p$ variance-covariance matrix of X, Σ_{22} is the $q \times q$ variance-covariance matrix of Y and Σ_{12} is the $p \times q$ covariance matrix between the two, $\Sigma_{21} = \Sigma_{12}^T$. The quantity a is the eigenvector containing the weights with which to multiply X from the one point in time and b is the eigenvector containing the weights with which to multiply Y from the other point in time. A more well-known formulation of the
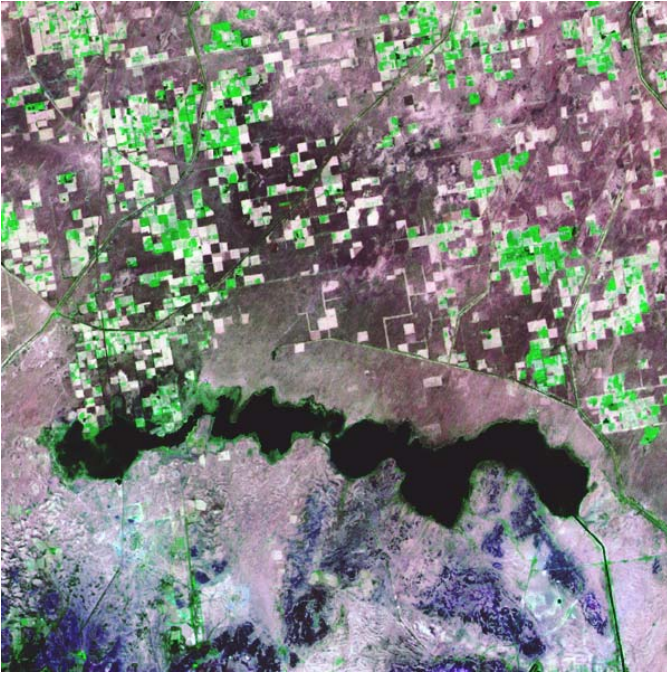
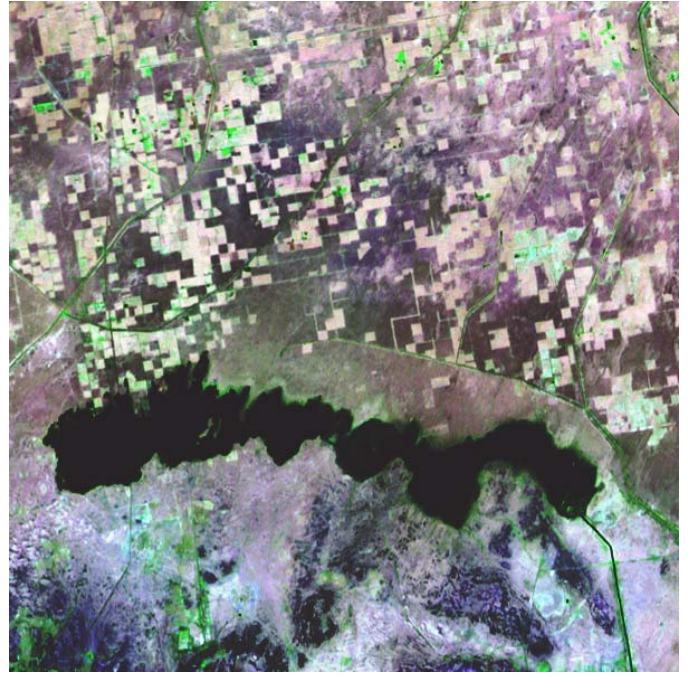Fig. 1.   Landsat TM bands 7, 4 and 5 at 29 March 1998 as RGB.

Fig. 2.   Landsat TM bands 7, 4 and 5 at 16 May 1998 as RGB.

CCA problem is given by the coupled eigenvalue problems

$$\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\ a\ =\ \rho^2 \Sigma_{11}\ a \qquad (3)$$

$$\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\ b\ =\ \rho^2 \Sigma_{22}\ b. \qquad (4)$$

To do change detection we form the canonical variates $U_i = a_i^T X$ and $V_i = b_i^T Y$ and the MAD change detector as the difference $Z_i = U_i - V_i$ between them ($V_i = 0$ for $i > q$). The MAD variates Z_i are orthogonal and have variances $2(1-\rho_i)$, hence the reverse ordering which maximizes variance in the low order MAD variates which are the differences between the high order canonical variates.

A. IR-MAD

Ideally, the sum of squared standardized MAD variates will follow a χ^2 distribution with p degrees of freedom, i.e., we have approximately

$$\sum_{i=1}^{p}\left(\frac{Z_i}{\sigma_{Z_i}}\right)^2 \in \chi^2(p) \qquad (5)$$

where σ_{Z_i} ideally is the standard deviation of the no-change observations. These can be found by the methods proposed in [4], [5].

In the iteratively reweighted (IR) MAD method [1] we put increasing weight on observations that exhibit little change over time. For the statistics calculations we weight observation j in the next iteration by w_j which is a measure of no change, namely the probability of finding a greater value of the χ^2 value in Equation 5

$$w_j = P\left\{ > \sum_{i=1}^{p}\left(\frac{Z_i}{\sigma_{Z_i}}\right)^2 \right\}_j \simeq P\{> \chi^2(p)\}_j. \qquad (6)$$

This establishes a better background of no-change against which to detect change. Iterations stop when ρ stops changing (substantially).

Since the (IR-)MAD transformation is based on CCA the (IR-)MAD variates, like the canonical variates, are invariant to affine (and linear) transformations to the original data including linear and affine radiometric normalization or calibration. This makes them good generalized multivariate differences between all variables at the two time points of acquisition.

B. Regularized CCA and (IR-)MAD

If we wish to apply regularization in CCA we could solve

$$\begin{bmatrix} 0 & \Sigma_{12} \\ \Sigma_{21} & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \qquad (7)$$

$$= \rho \begin{bmatrix} \Sigma_{11} + k_1\Omega & 0 \\ 0 & \Sigma_{22} + k_2\Omega \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

where k_1 and k_2 determine the amount of regularization and Ω is designed to minimize, e.g., size, slope or curvature in a and b considered as functions of wavelength [8], [9].

Alternatively, regularization can be based on exploitation of the affine transformation invariance of the MAD method. The data at the two points in time can be orthogonally transformed separately to reduce redundancy and dimensionality before change detection by both the original and the iterated MAD methods [1], [10]. Both types of regularization can be applied to multispectral as well as hyperspectral data. In the latter case regularization might be crucial due to problems with (near) singular or ill-conditioned variance-covariance matrices.
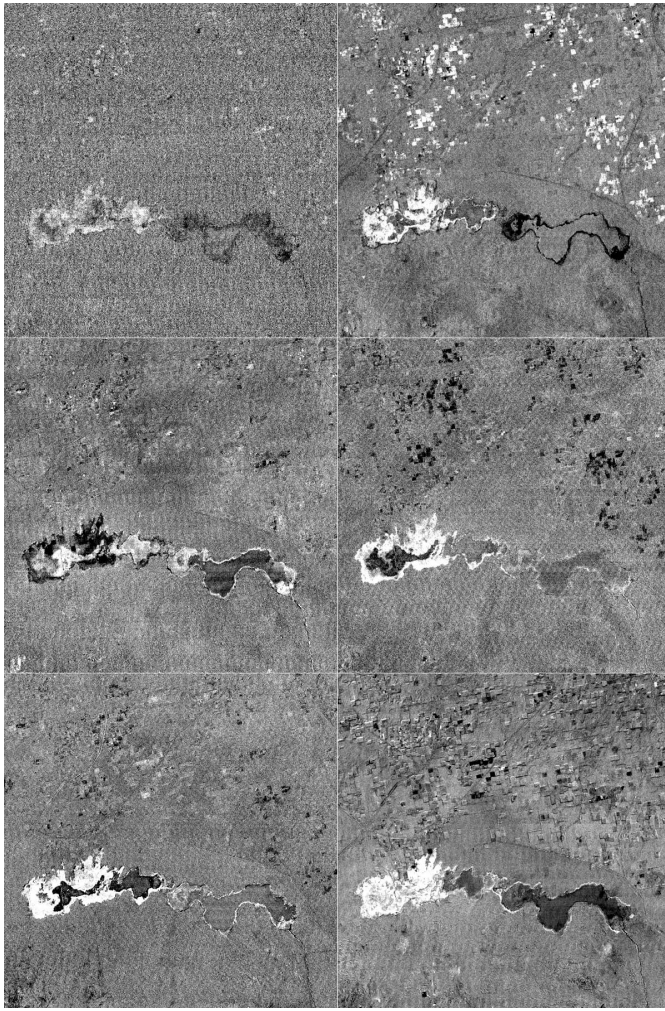
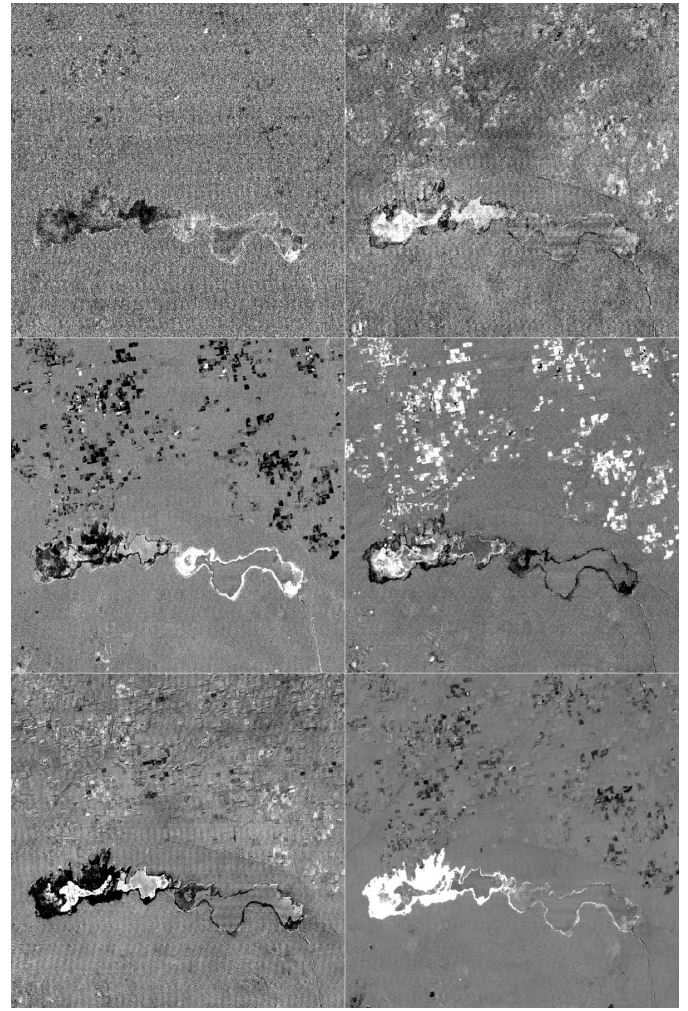Fig. 3.   Generalized differences: MAD variates 1 to 6, row wise.

Fig. 4.   Generalized differences: IR-MAD variates 1 to 6, row wise.

III. CASE STUDY

Figures 1 and 2 show 600×600 30 m pixels Landsat TM scenes from two points in time covering an area over a semi-arid agricultural region in Hindustan, India, see also [11] where the same data are subjected to this type of analysis followed by unsupervised classification into no-change and several different change clusters. The scenes were co-registered by applying an automatic contour matching algorithm [12] using a first-order polynomial and nearest-neighbor re-sampling. The RMS errors were less than or equal to 0.5 pixel. Because of the good atmospheric conditions and the invariance of the MAD variates under affine transformations, no radiometric corrections were applied to the two scenes. The most evident changes that took place between the acquisitions is some shallow flooding at the western end of the lake or reservoir and the vegetation changes in some of the bright rectangular (square) areas scattered over the northern part of the scenes.

Figures 3 and 4 show the original MAD variates and the IR-MAD variates after seven iterations, respectively. Note,

that since these change images are based on a solution to an eigenproblem their signs are arbitrary. We see that the reservoir formation is associated with most of the MAD variates and particularly so for the high order MAD variates corresponding to the low order also known as the leading canonical variates. Vegetation changes are associated with MAD variates number two and four, and with IR-MAD variates three and four (and to some extent variate six).

The higher order IR-MAD variates clearly appear much less noisy than the higher order MAD variates. For example, the average autocorrelations in IR-MAD and MAD variates six in the four main directions are 0.94 and 0.70, respectively. This underlines the visually conspicuous better spatial togetherness of both change and no-change observations for the IR-MAD method.

Figure 5 shows the canonical correlations over the seven iterations needed to stabilize the correlations to within 0.001. We see that the first iteration is most important and that the correlations increase steadily, corresponding to the gradual exclusion of the change observations from the canonical correlation analysis. Most of the changes in ρ take place within

4-5 iterations.

Here the standardization to unit variance in Equation 5 is done by means of the ideal no-change standard deviations only and not by means of the standard deviations of all observations. These no-change standard deviations were determined with the fuzzy maximum likelihood estimation (FMLE) clustering scheme suggested in [4] which is a form of EM algorithm. Two cluster means and variance-covariance matrices, representing no-change and change observations, were fit to all the IR-MAD variates after each iteration using the EM algorithm. For comparison, the variance-covariance matrix of the IR-MAD variates after the 7 iterations is given in Table I, while Table II shows the variance-covariance matrix of the no-change cluster only. The latter is seen to be approximately diagonal with monotonically decreasing diagonal elements (variances).

Figure 6 shows χ^2 for the original MAD method (left) and for the IR-MAD method (right). We see that since we normalize by smaller variances for IR-MAD χ^2 becomes larger thus depicting change more conspicuously.

Figure 7 shows MAD variates 6, 5 and 4 as RGB. Figure 8 shows IR-MAD variates 6, 5 and 4 as RGB. We see a huge difference in the ability to differentiate between change and no-change observations (note, that Figure 7 is stretched linearly between $\mp$ five standard deviations whereas Figure 8 is stretched linearly between $\mp$ ten standard deviations. Figure 9 shows maximum autocorrelation factors (MAF) [13] 1, 2 and 3 of the six IR-MAD variates. This transformation enhances signal-to-noise in the resulting low order components by maximizing spatial autocorrelation in the change observations. In both Figures 8 and 9 we see a very clear depiction of the change in the lake both to the west and more subtle change
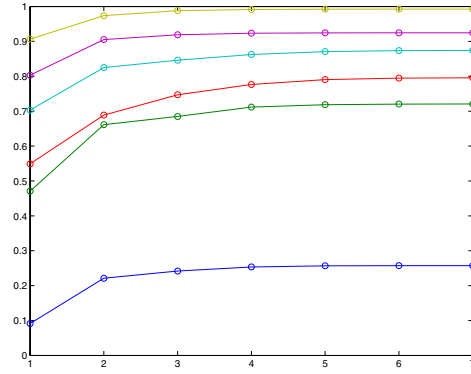

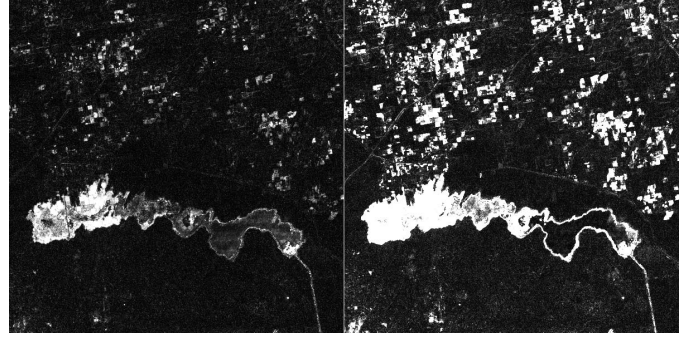Fig. 5.   Canonical correlations over seven iterations.

Fig. 6. χ^2 for the original MAD method (left) and for the IR-MAD method (right) stretched linearly between 0 and 50.

along the shore line to the east, as well as vegetation changes to the north of the lake.

IV. CONCLUSIONS

Simple differencing (provided it makes sense to do the calculation at all) gives correlated difference images ordered by wavelength. The MAD and IR-MAD methods provide generalized orthogonal difference images ordered by similarity as measured by linear correlation. These generalized differences are invariant to linear and affine transformations of the original variables, a enormous advantage over the simple differences.

Compared to the MAD method the IR-MAD method provides a better background of no-change against which to detect change resulting in a much greater difference between scores for change observations and no-change obervations for the IR-MAD method. Also, (especially the higher order) IR-MAD variates are much less noisy than the MAD variates.

Regularization may be useful or even necessary if the number of observations (pixels) is small relative to the number of variables (spectral bands). This typically applies to hyperspectral data.

TABLE I

VARIANCE-COVARIANCE AND CORRELATION MATRICES (ABOVE DIAGONAL) OF IR-MAD VARIATES, ALL OBSERVATIONS.

2.5246	−0.1037	0.1803	−0.1372	−0.0343	−0.0626
−0.1805	**1.1996**	−0.1897	0.3168	0.1794	−0.0599
0.5784	−0.4193	**4.0748**	−0.7539	−0.0945	0.0020
−0.3658	0.5821	−2.5533	**2.8153**	0.2968	−0.2812
−0.0464	0.1671	−0.1622	0.4235	**0.7232**	−0.5836
−0.0851	−0.0561	0.0034	−0.4037	−0.4247	**0.7322**

TABLE II

VARIANCE-COVARIANCE AND CORRELATION MATRICES (ABOVE DIAGONAL) OF IR-MAD VARIATES, NO-CHANGE OBSERVATIONS ONLY.

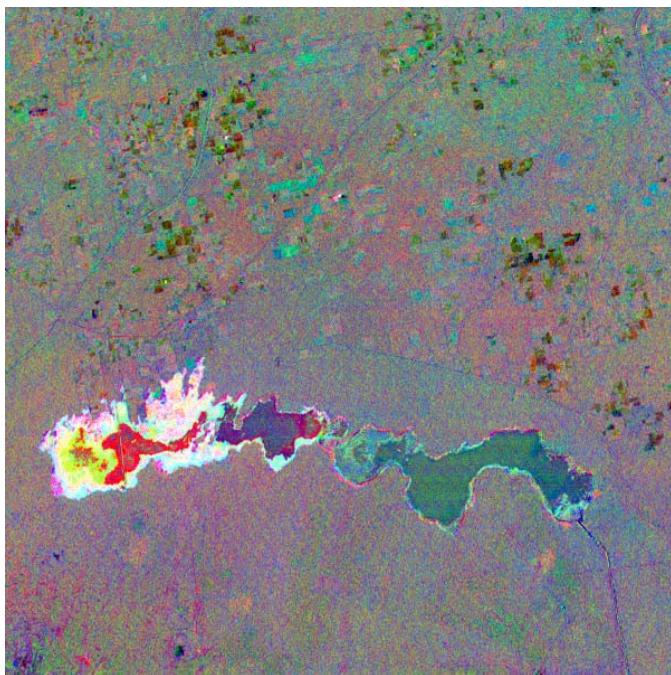|					
2.1617	−0.0093	0.0035	−0.0036	−0.0112	−0.0013
−0.0125	**0.8329**	−0.0070	0.0336	0.0167	−0.0246
0.0041	−0.0051	**0.6414**	−0.0917	0.0218	0.0536
−0.0034	0.0196	−0.0469	**0.4077**	0.0305	−0.0371
−0.0081	0.0075	0.0086	0.0096	**0.2425**	0.0569
−0.0003	−0.0036	0.0069	−0.0038	0.0045	**0.0258**

Fig. 7. MAD variates 6, 5 and 4 as RGB stretched linearly between ∓ five standard deviations for the no-change observations as calculated by the FMLE clustering algorithm.
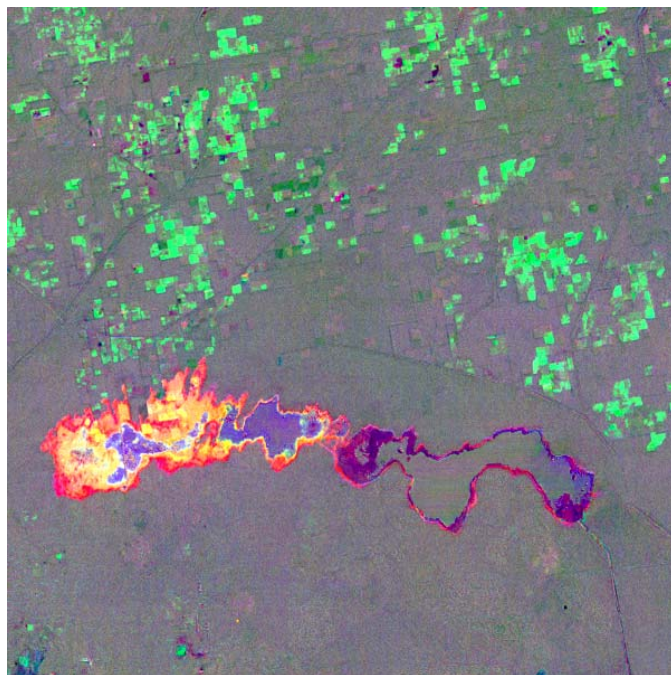

Fig. 9. MAF variates 1, 2 and 3 of all six IR-MAD variates as RGB stretched linearly between ∓ ten standard deviations.
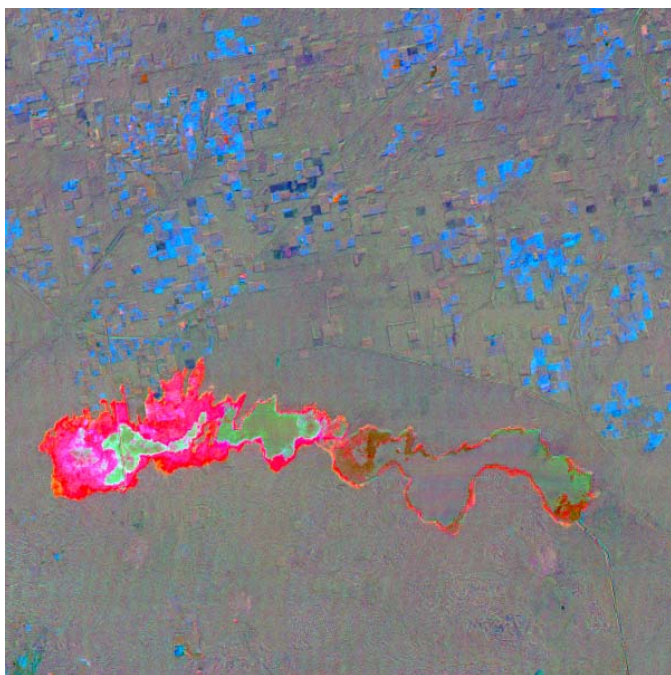

Fig. 8. IR-MAD variates 6, 5 and 4 as RGB stretched linearly between ∓ ten standard deviations for the no-change observations as calculated by the FMLE clustering algorithm.

REFERENCES

[1] A. A. Nielsen, "Iteratively reweighted multivariate alteration detection in multi- and hyperspectral data," submitted, 2005.

[2] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. XXVIII, pp. 321–377, 1936.

[3] A. A. Nielsen, K. Conradsen, and J. J. Simpson, "Multivariate alteration detection (MAD) and MAF post-processing in multispectral, bi-temporal image data: New approaches to change detection studies," *Remote Sensing of Environment*, vol. 64, pp. 1–19, 1998.

[4] I. Gath and A. B. Geva, "Unsupervised optimal fuzzy clustering," *IEEE Transactions on Pattern Analysis and Machine Intellegence*, vol. 3, no. 3, pp. 773–781, 1988.

[5] L. Bruzzone and D. F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 3, pp. 1171–1182, 2000.

[6] M. J. Canty, A. A. Nielsen, and M. Schmidt, "Automatic radiometric normalization of multitemporal satellite imagery," *Remote Sensing of Environment*, vol. 91, pp. 441–451, 2004.

[7] W. Zhang and M. Liao and Y. Wang and L. Lu and Y. Wang, "Robust approach to the MAD change detection method," in *Proceedings of the 11th SPIE International Symposium on Remote Sensing X*, vol. 5574, pp. 184–193, Maspalomas, Gran Canaria, Spain, 13-16 September 2004.

[8] J. Ramsay and B. Silverman, *Functional Data Analysis*, Springer, 1997.

[9] A. A. Nielsen, "Regularisation in multi- and hyperspectral remote sensing change detection," on CD-ROM *Proceedings of 6th Geomatic Week Conference*, Barcelona, Spain, 8-10 February 2005, Internet http://www.imm.dtu.dk/pubdb/p.php?3387.

[10] A. A. Nielsen, A. Müller, and W. Dorigo, "Hyperspectral data, change detection and the MAD transformation," on CD-ROM *Proceedings of the 12th Australasian Remote Sensing and Photogrammetry Association Conference*, 18-22 October 2004, Internet http://www.imm.dtu.dk/pubdb/p.php?3176.

[11] M. J. Canty and A. A. Nielsen, "Unsupervised classification of changes in multispectral satellite imagery," in *Proceedings of the 11th SPIE International Symposium on Remote Sensing X*, vol. 5573, pp. 356–363, Maspalomas, Gran Canaria, Spain, 13-16 September 2004, Internet http://www.imm.dtu.dk/pubdb/p.php?3425.

[12] H. Li, B. S. Manjunath, and S. K. Mitra, "A contour-based approach to multisensor image registration," *IEEE Transactions on Image Processing*, vol. 4, no. 3, pp. 320–334, 1995.

[13] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 65–74, 1988.