MULTI-CHANNEL REMOTE SENSING DATA AND ORTHOGONAL TRANSFORMATIONS FOR CHANGE DETECTION

Allan Aasbjerg Nielsen
IMM, Department of Mathematical Modelling
Technical University of Denmark, Building 321
DK–2800 Lyngby, Denmark
aa@imm.dtu.dk, http://www.imm.dtu.dk/∼aa

ABSTRACT

This paper describes the multivariate alteration detection (MAD) transformation which is based on the established canonical correlation analysis. It also proposes post-processing of the change detected by the MAD variates by means of maximum autocorrelation factor (MAF) analysis. As opposed to most other multivariate change detection schemes the MAD and the combined MAF/MAD transformations are invariant to affine transformations of the originally measured variables. Therefore, they are insensitive to, for example, differences in gain and off-set settings in a measuring device, and to the application of radiometric and atmospheric correction schemes that are linear or affine in the gray numbers of each image band. Other multivariate change detection schemes described are principal component type analysis of simple difference images. A case study with Landsat TM data using simple linear stretching and masking of the change images shows the usefulness of the new MAD and MAF/MAD change detection schemes. A simple simulation of a no-change situation shows the power of the MAD and MAF/MAD transformations.

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 change have zero or low absolute values and areas with large changes have large absolute values in the difference image. If we have two multivariate images with variables at a given location written as vectors (without loss of generality we assume that $\mathrm{E}\{\boldsymbol{X}\} = \mathrm{E}\{\boldsymbol{Y}\} = \boldsymbol{0}$)

$$\boldsymbol{X} = [X_1 \ \dots \ X_k]^T \quad \text{and} \quad \boldsymbol{Y} = [Y_1 \ \dots \ Y_k]^T \tag{1}$$

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

$$\boldsymbol{X} - \boldsymbol{Y} = [X_1 - Y_1 \ \dots \ X_k - Y_k]^T. \tag{2}$$

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

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

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 and maximum autocorrelation factor (MAF) analysis (Switzer and Green, 1984).

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

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

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 MultiSpectral Scanner (MSS) and the other set comes from Landsat Thematic Mapper (TM) or from SPOT High Resolution Visible (HRV) which may be valuable in historical change studies. 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 concatenated vector variable. Gong (1993) applies PC analysis to simple difference images as described above. The Fung and LeDrew (1987) 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. The Gong (1993) 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 maximise the variance, this time 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 Var$\{\boldsymbol{a}^T\boldsymbol{X}-\boldsymbol{b}^T\boldsymbol{Y}\}$ with Var$\{\boldsymbol{a}^T\boldsymbol{X}\}=$ Var$\{\boldsymbol{b}^T\boldsymbol{Y}\}=1$. With this choice 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 (non-negative) correlation. Determination of linear combinations with extreme correlations brings the theory of canonical correlations analysis to mind.

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. The technique was first described by Hotelling in 1936 and a treatment is given in most textbooks on multivariate statistics (good references are Cooley and Lohnes, 1971 and Anderson, 1984).

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. Wiemker et al. (1997) work with (iterated) PC analysis of the same variable at the two points in time and consider the second PC as a (marginal) change detector for that variable. They also introduce spatial measures such as inverse local variance weighting in statistics calculation and Markov random field modelling of the probability of change (vs. no change).

THE MULTIVARIATE ALTERATION DETECTION (MAD) TRANSFORMATION

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

$$[\boldsymbol{X}^T\ \boldsymbol{Y}^T]^T \rightarrow [\boldsymbol{a}_p^T\boldsymbol{X} - \boldsymbol{b}_p^T\boldsymbol{Y}\ \ldots\ \boldsymbol{a}_1^T\boldsymbol{X} - \boldsymbol{b}_1^T\boldsymbol{Y}]^T, \tag{6}$$

where $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$ are the defining coefficients from a standard canonical correlations analysis. $\boldsymbol{X}$ and $\boldsymbol{Y}$ are vectors with $\mathrm{E}\{\boldsymbol{X}\} = \mathrm{E}\{\boldsymbol{Y}\} = \boldsymbol{0}$. The dispersion matrix of the MAD variates is

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

where $\boldsymbol{I}$ is the $p\times p$ unit matrix and $\boldsymbol{R}$ 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, and to linear radiometric and atmospheric correction schemes. This type of multivariate change detection technique was first sketched in Conradsen and Nielsen (1991).

To find maximum change areas with high spatial autocorrelation a MAF post-processing of the MAD variates is suggested. The MAF transformation can be considered as a spatial extension of PC analysis in which the new variates maximise autocorrelation between neighbouring pixels rather than variance (as with PCs). Also the MAF transformation is invariant to linear scaling.

For other descriptions of the transformation described here see Nielsen (1994) who describes the MAD transformation and introduces the MAF post-processing of the MADs, Nielsen and Conradsen (1997) who describe the MAD transformation in more detail, and Nielsen et al. (1998)

where one case study very successfully applies the MAD transformation on pre-processed AVHRR data (correction for atmospheric water vapour attenuation, and cloud- and land-masking) in an oceanographic case study on ENSO-related (El Niño/Southern Oscillation) mid-latitude warming events where principal components analysis of simple difference images fail. Another case in Nielsen et al. (1998) deals with urbanisation in Australia using Landsat MSS data.

CASE STUDY - LANDSAT TM DATA

The applicability of the MAD transformation with the MAF post-processing to multivariate and bi-temporal change detection studies is demonstrated in a case using Landsat-5 Thematic Mapper (TM) data covering a small forested area approximately 20 kilometres north of Umeå in northern Sweden. The data consist of six spectral bands with 512×512 20 meter pixels rectified to the Swedish national grid by the Swedish Space Corporation who also provided the data. The acquisition dates are 6 June 1986 and 27 June 1988. The intention of this case study is to illustrate the method and not to accurately assess the actual change on the ground.

The images in Figures 1 and 2 are stretched linearly from mean minus three standard deviations to mean plus three standard deviations. Figure 1 shows the original Landsat TM data from 1986 in column one, the original Landsat TM data from 1988 in column two, the simple difference images (“1988 minus 1986”) in column three, and the principle components of the simple difference images in column four. The change information from all bands is concentrated in the first three or four PCs. Because of the stretching applied areas that are very bright or very dark are maximum change areas.

Figure 2 shows the canonical variates of the 1986 data in column one, the canonical variates of the 1988 data in column two, the MADs of the original TM images (the differences between the CVs in reverse order) in column three, and the MAF/MADs of the original TM images in column four. Again, areas that are very bright or very dark are maximum change areas. The low order MADs are quite noisy. It is to be expected that scanner noise and differences in atmospheric conditions are among the most different characteristics of the two scenes. In the MAF/MADs areas that are very bright or very dark are areas of change with high spatial autocorrelation. The change information from all bands is concentrated in the first three or four MAF/MADs.

Figure 3 shows negated absolute values stretched from mean (before taking absolute values) plus two standard deviations (before taking absolute values) to mean plus three standard deviations. Column one shows simple differences, column two principal components of simple differences, column three MADs, and column four MAF/MADs. This time bright regions are no-change areas and dark regions are maximum change areas.

Figure 4 shows the first three PCs of simple differences as RGB, stretched linearly from mean minus three standard deviations to mean plus three standard deviations (left), as (left) but with region from mean minus two standard deviations to mean plus two standard deviations collapsed to mean to enhance change regions (middle), and absolute values stretched linearly between mean (before taking absolute values) plus two standard deviations and mean plus six standard deviations, again to enhance change (right). Figure 5 shows the first three MAF/MAD components as RGB, stretching as in Figure 4.

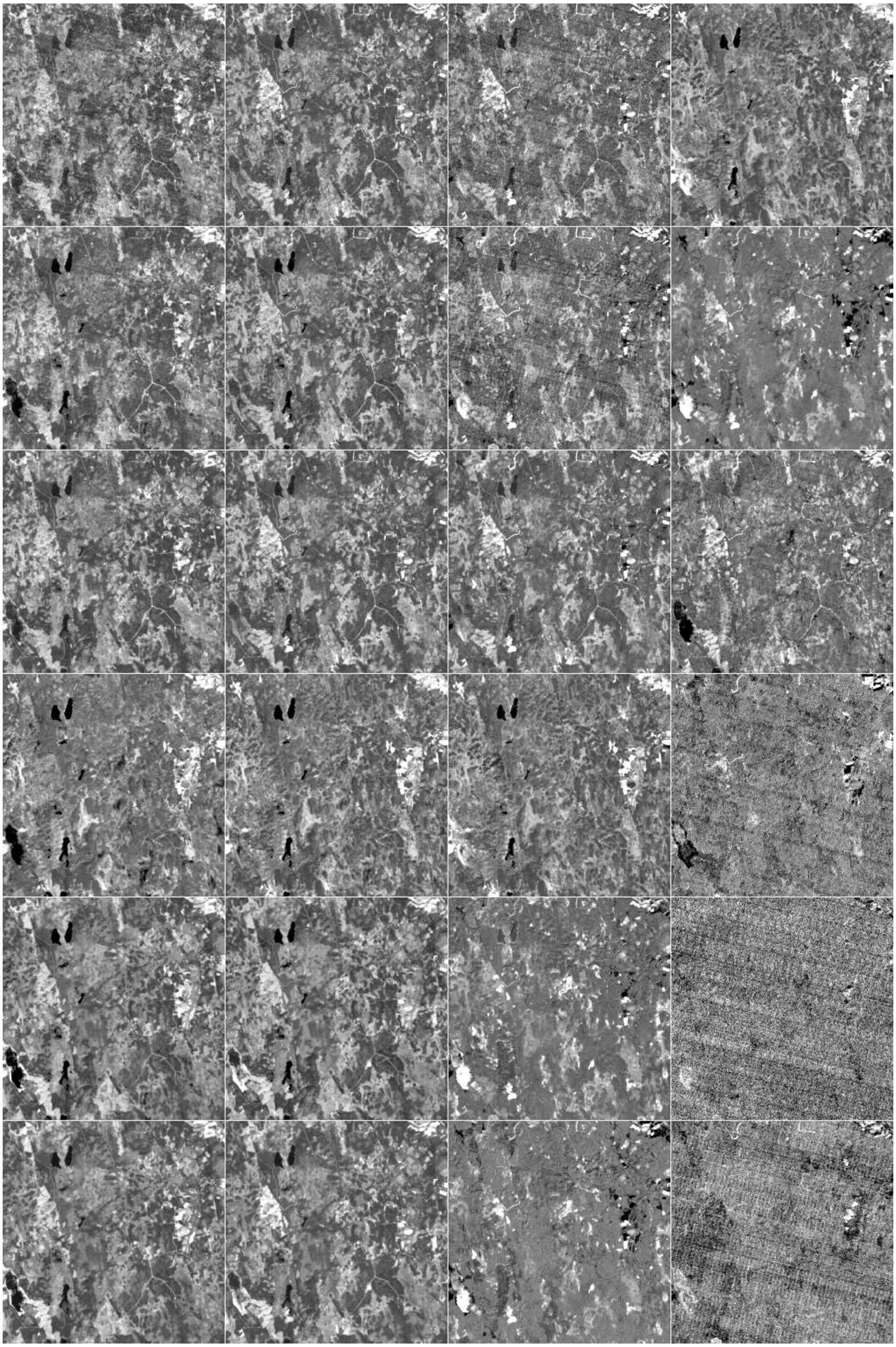

Figure 1: 1986 TM data (first column), 1988 TM data (second column), simple differences "1988–1986" (third column), and PCs of simple differences (fourth column)

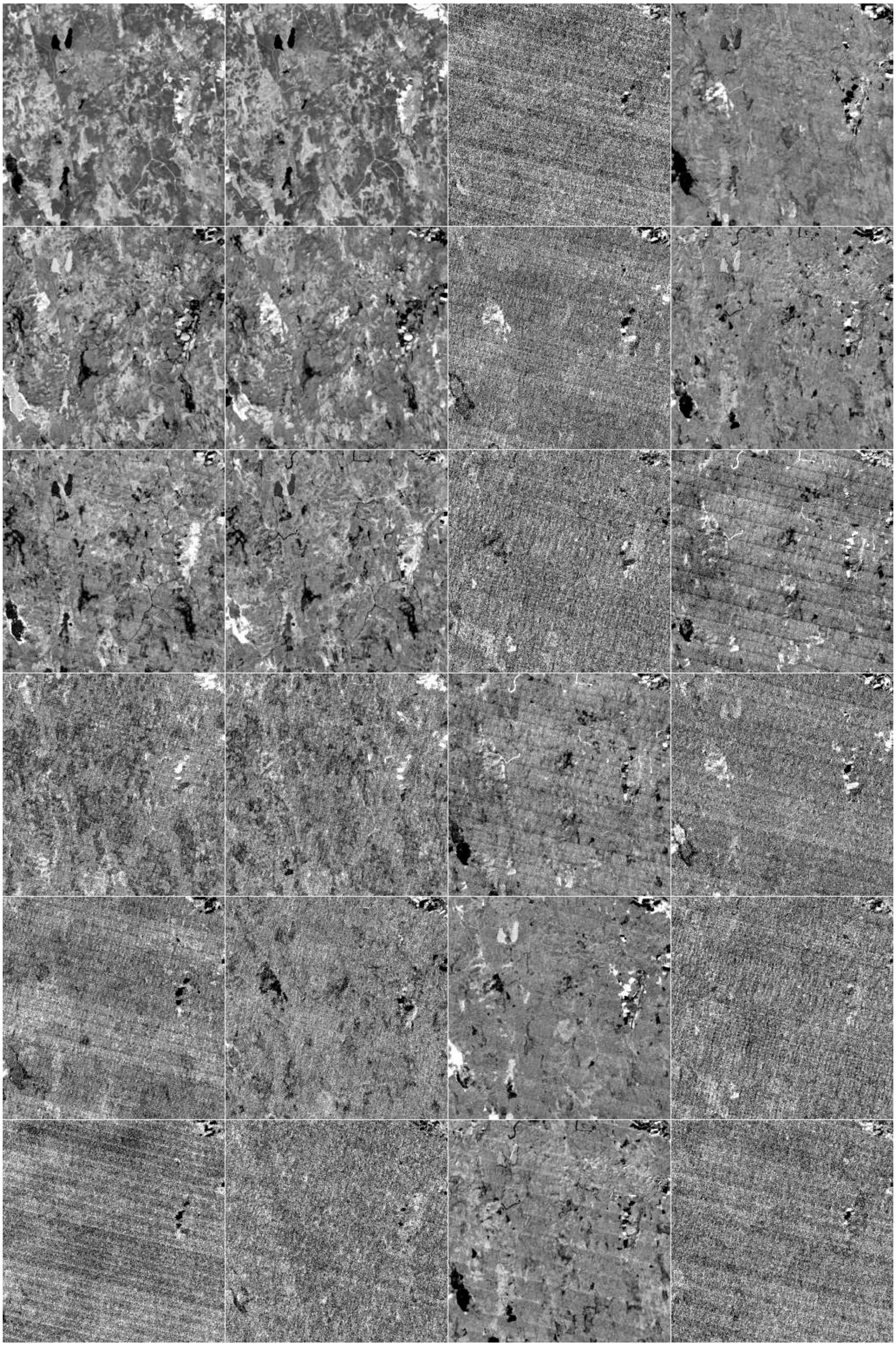

Figure 2: CVs of 1986 TM data (first column), CVs of 1988 TM data (second column), MADs (third column), and MAF/MADs (fourth column)

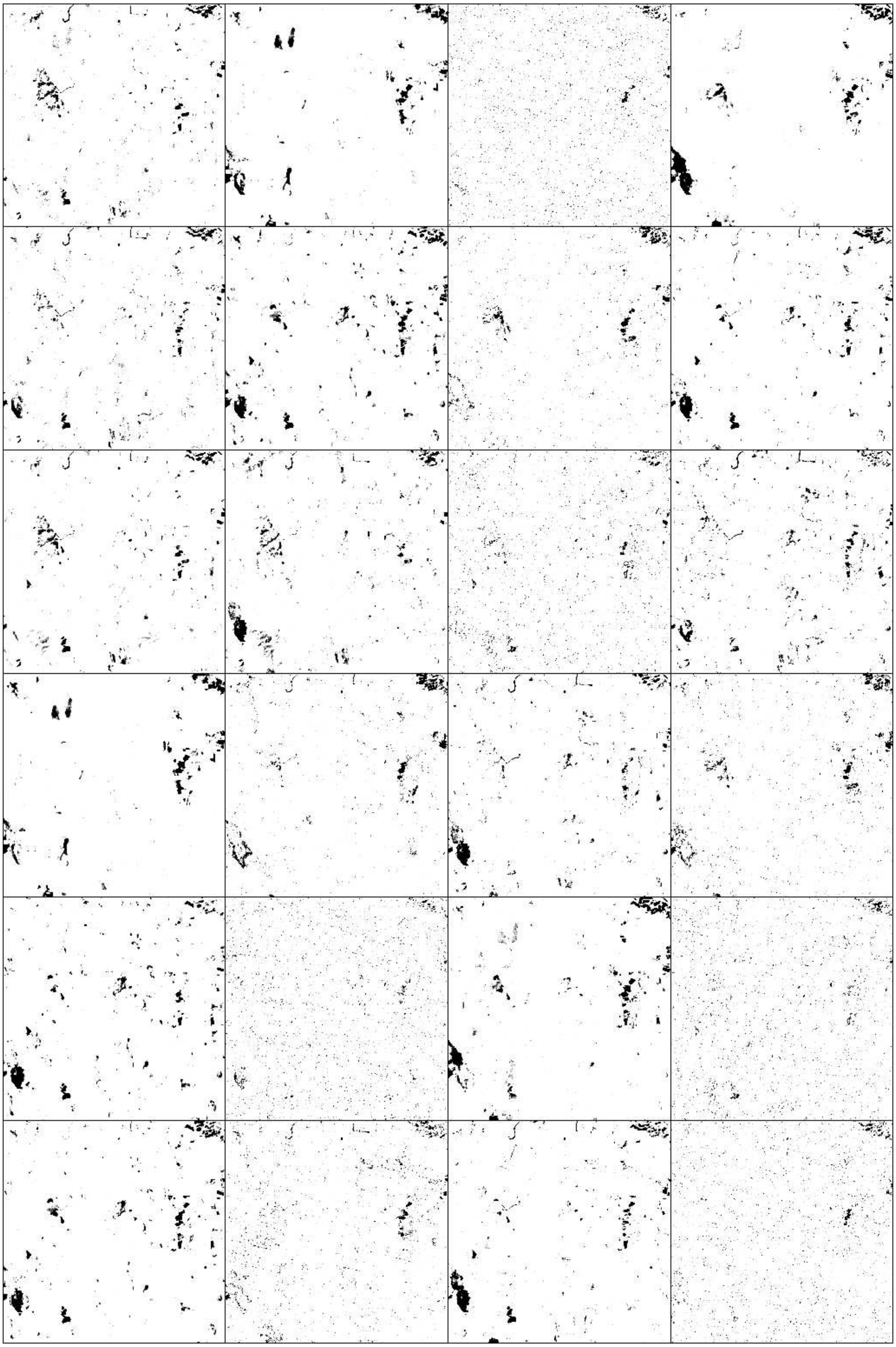

Figure 3: Negated absolute values of simple differences (first column), PCs of simple differences (second column), MADs (third column), and MAF/MADs (fourth column); white represents no-change areas

	MAF/ MAD1	MAF/ MAD2	MAF/ MAD3	MAF/ MAD4	MAF/ MAD5	MAF/ MAD6
1986 TM1	0.0089	0.5935	–0.1123	0.0142	–0.0399	0.1668
1986 TM2	0.1818	0.5589	–0.0567	0.1529	0.0260	0.0356
1986 TM3	0.0069	0.6548	–0.1052	0.0469	–0.1466	–0.0197
1986 TM4	0.4557	0.3039	0.0719	0.1630	–0.0217	0.1591
1986 TM5	0.1503	0.6145	0.1020	0.0729	–0.0636	0.0362
1986 TM7	–0.0320	0.6562	0.0816	–0.0990	–0.1077	–0.1077
1988 TM1	0.5343	–0.2053	0.2689	0.0991	–0.2147	–0.0293
1988 TM2	0.3632	–0.0964	0.3141	0.0098	–0.1771	0.1695
1988 TM3	0.6167	–0.2964	0.2591	0.0685	0.0021	0.0319
1988 TM4	–0.4385	0.3866	0.1113	0.0951	–0.0136	0.1285
1988 TM5	0.3976	–0.3634	–0.1730	–0.0676	–0.0137	–0.0148
1988 TM7	0.5884	–0.4006	–0.0838	0.1065	–0.0392	0.0488

Table 1: Correlations between MAF/MADs and Landsat TM bands

Band/Component	1	2	3	4	5	7/6
1) Simple diff.	0	0	0	0	0	0
2) PC diff. (cov)	–1.28	0.06	0.09	0.91	–0.03	0.26
3) PC diff. (corr)	–0.06	–1.53	0.19	0.40	0.03	–0.01
4) Factors diff.	–0.10	–0.81	1.22	0.22	–0.58	–0.05
5) MAF diff.	–1.52	0.05	0.46	0.03	0.01	0.01
6) MAD	0.01	0.01	0.01	0.01	0.01	–0.03
7) MAF/MAD	–0.00	0.01	0.02	0.00	–0.02	0.02

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

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 MAF/MADs and the original Landsat TM bands calculated only where the standardised score of any MAF/MAD variate is outside the plus/minus three standard deviations interval. MAF/MAD1 is positively correlated with 1986 TM4 and uncorrelated with 1986 TM3. MAF/MAD1 is negatively correlated with 1988 TM4 and positively correlated with 1988 TM3. Therefore MAF/MAD1 is a sort of vegetation index change detector. MAF/MAD2 is a change detector of the weighted mean of all bands except TM4, presumably a non-vegetation related change detector. MAF/MAD3 seems to be a non-vegetation related change detector also. MAF/MAD5 and 6 have their highest correlations with TM1 and TM2 and are therefore likely to represent change in atmospheric conditions. The MAF analysis of the MADs seems to have concentrated the change related to atmospheric conditions and change related to vegetation in each end of the autocorrelation “spectrum”. Also, scanner noise is concentrated in the high order MAF/MADs.

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 factors of simple differences, 5) MAFs of simple differences, 6) MADs, and 7) MAF/MADs. Change detected in the region with no change (the 44 pixels broad edge around the actual image data) as indicated by standardised values of the results from the above change detection methods is given in Table 2.

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 MADs would have been exactly the same. Therefore the situation simulated though very simple is indeed a realistic one.

It is obvious that MAD and MAF/MAD are the only multivariate techniques that perform well in this situation. All other methods give change values much higher than zero.

CONCLUSIONS

A new approach to change detection in multivariate, bi-temporal, spatial data based on canonical correlations analysis called the multivariate alteration detection (MAD) transformation and maximum autocorrelation factor (MAF) post-processing is described. The success of this scheme is demonstrated in a case with Landsat TM data from a forested region in Sweden. The difference between this scheme and the one based on principal components analysis of the simple difference images suggested by Gong (1993) is the invariance af the MAD and MAF transformations to linear and affine transformations which makes the use of the combined MAF/MAD transformation insensitive to for instance 1) differences in gain and off-set in a measuring device, and 2) linear radiometric and atmospheric calibration methods. A simple but realistic simulation of a no-change situation favours the MAD and the combined MAF/MAD methods.

ACKNOWLEDGEMENTS

I wish to thank Professor Knut Conradsen, IMM, for his cooperation on the development of the MAD concept. The calculations were performed by means of software written by Dr. Rasmus Larsen, IMM, on my initiative and under my supervision. Also, I wish to thank Dr. James J. Simpson, Scripps Institution of Oceanography, La Jolla, California, for his immediate interest in and comments on the MAD and the MAF/MAD schemes. The Landsat TM data were georectified and kindly provided by the Swedish Space Corporation. Funding from the European Commission and Danida, the Danish international development agency, 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. (1991): 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.

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 multi-temporal 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*. Department of Mathematical Modelling, Technical University of Denmark, Ph.D. Thesis No. 6. http://www.imm.dtu.dk/documents/users/aa/phd/.

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

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

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

Wiemker, R., Speck, A., Kulbach, D., Spitzer, H. and Beinlein, J. (1997): Unsupervised robust change detection on multispectral imagery using spectral and spatial features. In *Proceedings from the Third International Airborne Remote Sensing Conference and Exhibition*, Copenhagen, Denmark, vol. I, pp. 640-647.

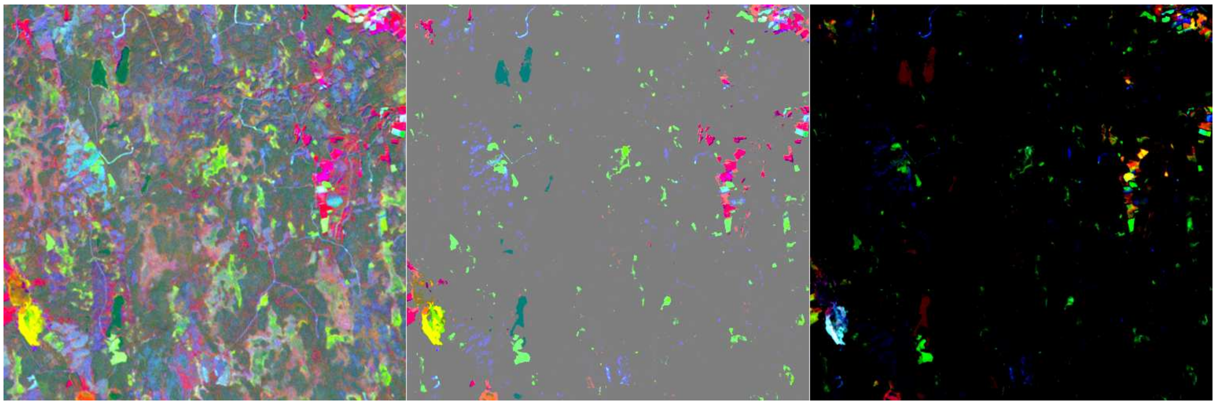

Figure 4: The first three PCs of simple differences as RGB, left: stretched linearly from mean minus three standard deviations to mean plus three standard deviations, middle: as left but with region from mean minus two standard deviations to mean plus two standard deviations collapsed to mean, right: absolute values stretched linearly between mean (before taking absolute values) plus two standard deviations and mean plus six standard deviations

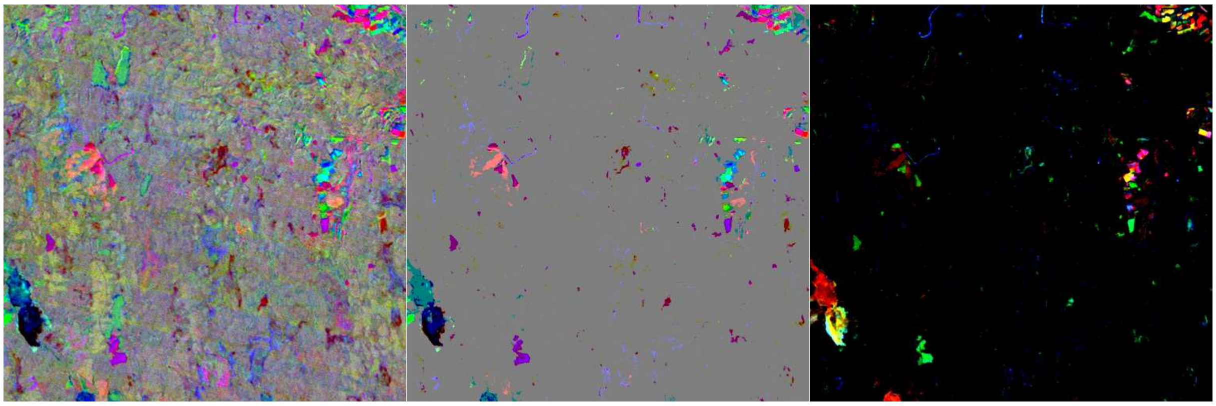

Figure 5: The first three MAF/MAD components as RGB, stretching as in Figure 4