

An Iterative Extension to the MAD Transformation for Change Detection in Multi- and Hyperspectral Remote Sensing Data

Allan Aasbjerg Nielsen

Technical University of Denmark
Informatics and Mathematical Modelling
DK-2800 Lyngby, Denmark
www.imm.dtu.dk/~aa
aa@imm.dtu.dk

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 a regularized, 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 gives a better no-change background against which to detect change than the original MAD method, and that the IR-MAD method and its regularized version depict the change detected in less noisy components.

Keywords: canonical correlation analysis, CCA; iteratively reweighted multivariate alteration detection, IR-MAD; orthogonal transformations; regularization or penalization.

1 INTRODUCTION

This paper deals with construction of more general difference images than simple band-wise differences in multi- and hypervariate change detection. This is done via a regularized or penalized, iterated version [12] of the canonical correlation analysis (CCA) [1], [6], [10] based multivariate alteration detection (MAD) method [15] 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 [2], [3], [7]. This can be used to establish a single change/no-change image based on the generalized multivariate difference images. 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 [4].

Results (not shown here) from partly simulated multivariate data indicate an improved performance of the iterated scheme over the original MAD method [12]. Also, a few comparisons with established methods for calculation of robust statistics for the CCA indicate that the scheme suggested here performs better [12], [22].

Regularization or penalization issues typically important in connection with the analysis of hyperspectral data are dealt with in [13], [16]–[18], [20], [21] and described here also. Among many other things [18] describes CCA in a functional setting.

This paper focuses on mathematical/statistical methodology development in multi- and hyperspectral remote sensing change detection and not on thorough scrutiny of actual change on the ground including analysis of *in-situ* data.

2 THE MAD TRANSFORMATION

Simple band-wise 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 pairs of canonical variates from a canonical correlation analysis, CCA. Based on N observations (pixels) of geometrically co-registered $p \times 1$ zero mean data X from one point in time and $q \times 1$ zero mean data Y from another point in time ($p \geq q$), CCA establishes new mutually orthogonal variables, $a^T X$ and $b^T Y$ termed canonical variates, CVs, with unit variance and with maximum correlation termed the canonical correlation, $\rho = \text{Corr}\{a^T X, b^T Y\} \geq 0$. The canonical variates can be found by solving the generalized eigenvalue problem

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

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 $p \times 1$ eigenvector containing the weights with which to multiply X from the one point in time and b is the $q \times 1$ eigenvector containing the weights with which to multiply Y from the other point in time. A (perhaps) more well-known formulation of the CCA problem is given by the coupled eigenvalue problems

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

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

To do change detection we form the canonical variates $U_i = a_i^T X$ and $V_i = b_i^T Y$, $i = 1, \dots, p$, 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 with low canonical correlations ρ .

2.1 Iteratively reweighted MAD, IR-MAD

Ideally, the sum of squared standardized MAD variates will follow a χ^2 distribution with p degrees of freedom [12], i.e., we have approximately

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

where σ_{Z_i} ideally is the standard deviation of the no-change observations. These can be found by methods proposed in [2], [3], [7].

In the iteratively reweighted (IR) MAD method [12] 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

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

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.

2.2 Regularization, CCA and (IR-)MAD

In ordinary least squares (OLS) regression [19]

$$y = X\theta + e \quad (7)$$

where y is the $N \times 1$ response variable, X is $N \times p$ with one column for each of p regressors or input variables, and θ is the $p \times 1$ parameter vector, N is the number of observations, we solve for the parameter vector θ by minimizing $\frac{1}{2}e^T e$, i.e., we minimize (half) the sum of the squared differences between the data and the model. This leads to the so-called normal equations

$$(X^T X) \hat{\theta} = X^T y \quad (8)$$

or (formally) $\hat{\theta} = (X^T X)^{-1} X^T y$. To avoid possible (near) singularity problems in $X^T X$ we may minimize $\frac{1}{2}(e^T e + k \theta^T \theta)$ instead, i.e., we minimize (half) the sum of the squared differences between the data and the model, and the size or the length of the parameter vector θ ; the scalar $k \geq 0$ controls how much weight we want to put on penalizing the size of θ . This leads to new normal equations

$$(X^T X + k I) \hat{\theta} = X^T y \quad (9)$$

where I is the $p \times p$ unit matrix. Note that $X^T X$ is proportional to the variance-covariance matrix of the regressors. To ensure the same influence of the regularization on all regressors it is customary to normalize the input variables to zero mean and unit variance. We see that by solving Equation 9 instead of Equation 8 we punish or penalize high values of the elements of θ . In other words with increasing k the elements of θ tend to become closer to zero. k can be chosen subjectively or estimated from the data by cross-validation. This type of regression is termed ridge regression [9].

More generally, we may penalize other characteristics of θ than size by minimizing $\frac{1}{2}[e^T e + k (L\theta)^T (L\theta)]$ where L is some matrix. This leads to

$$(X^T X + k \Omega) \hat{\theta} = X^T y \quad (10)$$

with $\Omega = L^T L$. In the above simple case we have $\Omega = I$ ($= L = L_0 = L_0^T L_0$).

Say instead we wanted to force all elements of θ to be equal. This can be done by setting $L_1 \theta = 0$ where L_1 is $(p-1) \times p$ with

$$L_1 = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix} \quad (11)$$

leading to the desired

$$L_1 \theta = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_p \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (12)$$

i.e., $\theta_1 = \theta_2, \theta_2 = \theta_3, \dots, \theta_{p-1} = \theta_p$ or $\theta_1 = \theta_2 = \dots = \theta_p$.

Rather than forcing the elements of θ to be equal we may want to penalize curvature in the elements of θ . For this to make sense some ordering of the elements of θ is assumed; in remote sensing this ordering could be by wavelength of the spectral bands. The desired minimum curvature can be achieved by using the usual discrete approximation to the second order derivative and setting $L_2 \theta = 0$ where L_2 is $(p-2) \times p$ with

$$L_2 = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -2 & 1 \end{bmatrix} \quad (13)$$

leading to $p \times p$

$$\Omega = L_2^T L_2 = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2 & 5 & -4 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 6 & -4 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -4 & 6 & -4 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -4 & 6 & -4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & -4 & 6 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & -4 & 5 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & -2 & 1 \end{bmatrix} \quad (14)$$

which is penta-diagonal.

Of course we can combine these different ways of penalizing the elements of θ to obtain a desired structure or desired characteristic of the solution by setting

$$\Omega = w_0 L_0^T L_0 + w_1 L_1^T L_1 + w_2 L_2^T L_2 + \cdots \quad (15)$$

where w_i are weights.

If we wish to apply regularization to CCA we could solve

$$\begin{bmatrix} 0 & \Sigma_{12} \\ \Sigma_{21} & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \rho \begin{bmatrix} \Sigma_{11} + \lambda_1 \Omega & 0 \\ 0 & \Sigma_{22} + \lambda_2 \Omega \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \quad (16)$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ determine the amount of regularization and Ω is designed to penalize or minimize, e.g., size, slope or curvature in a and b considered as functions of wavelength [13], [14], [16], [18]. Again, to ensure the same influence of the regularization on all the variables it is customary to normalize them to zero mean and unit variance. If X and Y are the same type of data as in our case we often set $\lambda_1 = \lambda_2 = \lambda$. We can either choose λ subjectively or determine it by cross-validation on the leading (i.e., the first) canonical variates only [18].

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 (for example by means of principal component [1], maximum autocorrelation factor [8] or independent component analysis [11]) before change detection by both the original and the iterated MAD methods [12], [17].

Both types of regularization can be applied to multispectral as well as hyperspectral data. For hyperspectral data regularization might be crucial due to problems with (near) singular or ill-conditioned variance-covariance matrices.

3. CASE STUDY

To illustrate the techniques we use 44 channels of 400 rows by 270 columns 5 m pixels HyMap [5] data covering a small agricultural area near Lake Waging-Taching in Bavaria, Germany. We have chosen bands 19 to 62 covering the wavelength region from 0.707 to 1.340 μm with approximately 15 nm spacing. Figure 1 shows HyMap bands 62, 40 and 19 at 30 June 2003 8:43 UTC and 4 August 2003 10:23 UTC as RGB images. In [3] Landsat TM data are subjected to this type of analysis followed by unsupervised classification into no-change and several change clusters.

Figure 2 shows the canonical correlations for the IR-MAD method over the seven iterations needed to stabilize the correlations to within 0.01 for no regularization and regularization with $\lambda_1 = \lambda_2 = \lambda = 0.1$. The regularization applied penalizes curvature, i.e., the second order derivative, of the weights in the CCA considered as functions of wavelength. Figure 3 shows the weights obtained by the IR-MAD method for the leading (i.e., the first pair of) canonical variates for no regularization and regularization with $\lambda = 0.1$. We see that with this heavy regularization we obtain many canonical correlations close to zero. As expected we get much smoother weights (eigenvectors) which facilitates interpretation. Also, we see that for the low order CVs the first iteration is most important and that most correlations especially of lower order increase, 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 standard deviations of all observations and not by means of the ideal no-change observations standard deviations only. Less heavy regularization with $\lambda = 0.01$ or $\lambda = 0.001$ gives slightly more wiggly weight functions and very similar canonical correlations and IR-MAD variates.

Figure 4 shows variates 43, 42 and 41 for the MAD method, for the IR-MAD method, and for the regularized IR-MAD method as RGB images. All variates are stretched linearly between the mean value (which is zero) minus four standard deviations and the mean value plus four standard deviations.

Figure 5 shows MAD variates 44, 43, 42 and 41, Figure 6 shows IR-MAD variates 44, 43, 42 and 41, and Figure 7 shows regularized IR-MAD variates 44, 43, 42 and 41. Again, all stretches are \mp four standard deviations. For all methods, MAD variate 44, i.e., the difference between the leading CVs with maximum correlation, seem dominated by changes associated with edges probably due to possible inaccuracies in geometric registration and different sun angles (there are five weeks between acquisitions, the first is taken at 8:43 UTC and the second is taken at 10:23 UTC). Therefore they are left out of the RGB composites in Figure 4. The average of the autocorrelations in the four main directions in these images are shown in Table 1. The strong tendency towards higher autocorrelations underline the better spatial togetherness of the change images of both the IR-MAD method and its regularized version over the original MAD method. This is also quite conspicuous visually.

Table 1. Mean autocorrelations between neighbouring pixels for the high order original MADs, the IR-MADs and the regularized IR-MADs.

	Orig. MAD	IR-MAD	Reg. IR-MAD
MAD 44	0.43	0.40	0.50
MAD 43	0.50	0.49	0.66
MAD 42	0.32	0.49	0.61
MAD 41	0.38	0.50	0.69

Often maximum autocorrelation factors (MAF) analysis [8] is used to post-process the MAD variates. The MAF transformation enhances signal-to-noise in the resulting low order components by maximizing spatial autocorrelation in the change observations. For space limitation reasons this is not shown here.

In these change detection transformations, scores close to zero indicate no-change and scores far from zero, i.e., very high (positive) or very low (negative) scores indicate strong change. The same stretching is applied to all score images. Therefore, the more extreme graytones in Figures 7 and 6 compared to Figure 5 and the more saturated colours in Figures 4(c) and 4(b) compared to Figure 4(a) are results of more extreme scores. This shows that the IR-MAD method and especially its regularized version are much better at differentiating between change and no-change observations than the original MAD method. Also, as mentioned above, the IR-MAD method and its regularized version produce less noisy change components, see Table 1.

4 CONCLUSIONS

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

Compared to the MAD method the IR-MAD method and its regularized version provide better backgrounds of no-change against which to detect change resulting in a much greater difference between scores for change observations and no-change observations for the IR-MAD methods. Also, (especially the higher order) IR-MAD variates are much less noisy than the MAD variates, and the canonical correlations for the regularized IR-MAD method fall off to zero much more rapidly than for both the MAD and the IR-MAD methods.

Regularization is useful and may even be necessary if the number of observations (pixels) is small relative to the number of variables (spectral bands) or if we work with hyperspectral data.

ACKNOWLEDGMENTS

The author wishes to thank Andreas Müller, DLR, Oberpfaffenhofen, Germany, for cooperation on hyperspectral data work over many years and for the geometrically and atmospherically corrected HyMap data. Thanks also to Dr. Mort Canty, FZ-Jülich, Germany, for cooperation on change detection methods.

This work is done partly within the EU funded Network of Excellence Global Monitoring for Security and Stability, GMOSS, <http://gmooss.jrc.cec.eu.int>.

REFERENCES

- [1] T. W. Anderson, "An Introduction to Multivariate Statistics," Wiley, third edition, 2003.
- [2] 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.
- [3] 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>.
- [4] 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. Internet <http://www.imm.dtu.dk/pubdb/p.php?2815>.
- [5] T. Cocks, R. Jenssen, A. Stewart, I. Wilson, and T. Shields, "The HyMap airborne hyperspectral sensor: the system, calibration, and performance," in *Proceedings of the 1st EARSeL Workshop on Imaging Spectroscopy*, pp. 37–42, Zürich, Switzerland, 6–8 October 1998, Internet <http://www.hyvista.com> and <http://www.intspec.com>.
- [6] W. W. Cooley and P. R. Lohnes, "Multivariate Data Analysis," Wiley, 1971.
- [7] I. Gath and A. B. Geva, "Unsupervised optimal fuzzy clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 3, no. 3, pp. 773–781, 1988.
- [8] 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.
- [9] A. E. Hoerl and R. W. Kennard, "Ridge regression. Biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.
- [10] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. XXVIII, pp. 321–377, 1936.
- [11] A. Hyvärinen and J. Karhunen and E. Oja, "Independent Component Analysis," Wiley, 2001.
- [12] A. A. Nielsen, "Iteratively reweighted multivariate alteration detection in multi- and hyperspectral data," submitted, 2005.
- [13] 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>.
- [14] A. A. Nielsen, "Multi- and hyperspectral remote sensing change detection with generalized difference images by the IR-MAD method," accepted for *MultiTemp 2005*, Biloxi, Mississippi, USA, 16–18 May 2005, Internet <http://www.imm.dtu.dk/pubdb/p.php?3630>.
- [15] 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. Internet <http://www.imm.dtu.dk/pubdb/p.php?1220>.
- [16] A. A. Nielsen and A. Müller, "Change detection by the MAD method in hyperspectral image data," in *Proceedings of the 3rd EARSeL Workshop on Imaging Spectroscopy*, 2003, pp. 115–116, Internet <http://www.imm.dtu.dk/pubdb/p.php?2420>.
- [17] 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>.
- [18] J. Ramsay and B. Silverman, *Functional Data Analysis*, Springer, 1997.
- [19] J. A. Rice, *Mathematical Statistics and Data Analysis*, Duxbury Press/Wadsworth, second edition, 1995.
- [20] H. D. Vinod, "Canonical ridge and econometrics of joint production," *Journal of Econometrics*, vol. 4, pp. 147–162, 1976.
- [21] B. Yu, I. M. Ostland, P. Gong, and R. Pu, "Penalized discriminant analysis of *in situ* hyperspectral data for conifer species recognition," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 5, pp. 2569–2577, 1999.
- [22] 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.



(a) 30 June 2003 8:43 UTC.

(b) 4 August 2003 10:23 UTC.

Figure 1. HyMap bands 62, 40 and 19 as RGB.

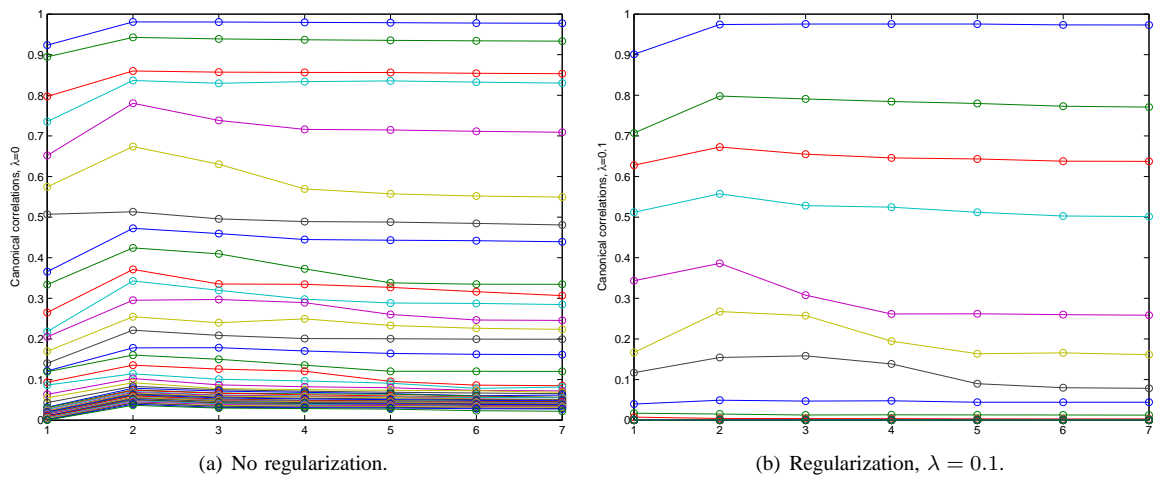


Figure 2. Canonical correlations over seven iterations.

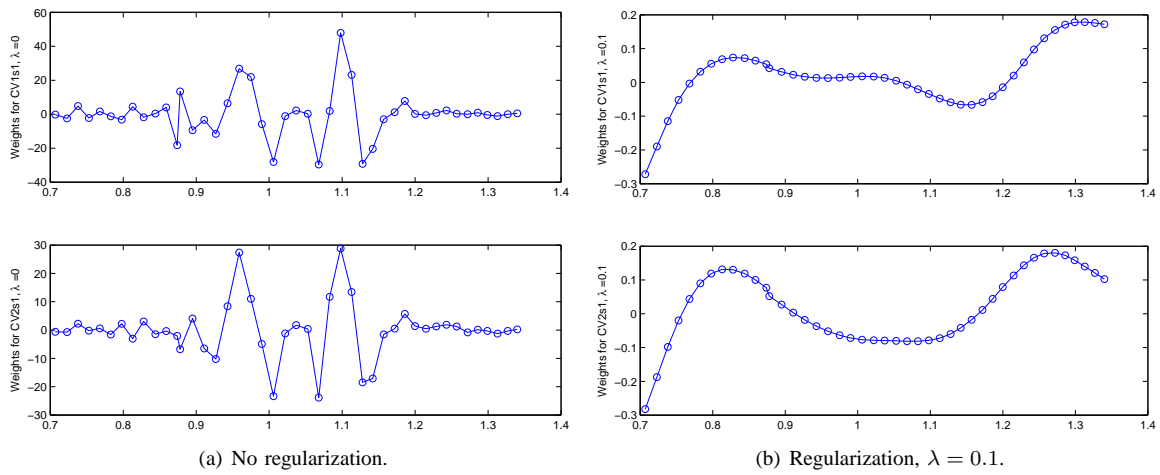


Figure 3. Weights for leading canonical variates.

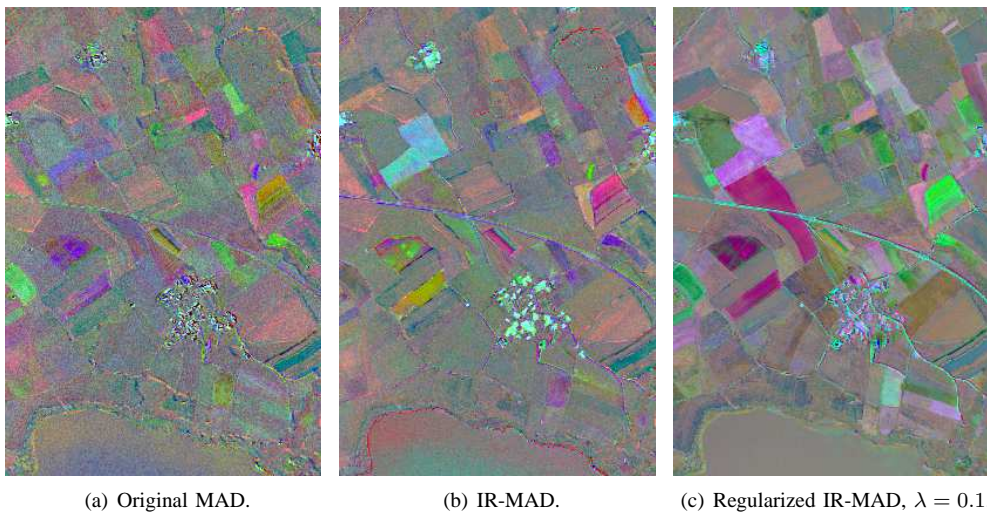


Figure 4. MADs 43, 42 and 41 as RGB.



Figure 5. MADs 44, 43, 42 and 41.



Figure 6. IR-MADs 44, 43, 42 and 41.



Figure 7. Regularized IR-MADs 44, 43, 42 and 41, $\lambda = 0.1$.