

# KERNEL MAD ALGORITHM FOR RELATIVE RADIOMETRIC NORMALIZATION

Yang Bai <sup>a,b</sup>, Ping Tang <sup>b</sup>, Changmiao Hu <sup>b,\*</sup>

<sup>a</sup> University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing, China

<sup>b</sup> Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, No.20 Datun Road, Beijing, China – (tangping, huchangmiao, baiyang) @radi.ac.cn

Commission I, WG I/4

**KEY WORDS:** Relative radiometric normalization, multivariate alteration detection (MAD), canonical correlation analysis (CCA), kernel version of canonical correlation analysis (KCCA)

## ABSTRACT:

The multivariate alteration detection (MAD) algorithm is commonly used in relative radiometric normalization. This algorithm is based on linear canonical correlation analysis (CCA) which can analyze only linear relationships among bands. Therefore, we first introduce a new version of MAD in this study based on the established method known as kernel canonical correlation analysis (KCCA). The proposed method effectively extracts the non-linear and complex relationships among variables. We then conduct relative radiometric normalization experiments on both the linear CCA and KCCA version of the MAD algorithm with the use of Landsat-8 data of Beijing, China, and Gaofen-1(GF-1) data derived from South China. Finally, we analyze the difference between the two methods. Results show that the KCCA-based MAD can be satisfactorily applied to relative radiometric normalization, this algorithm can well describe the nonlinear relationship between multi-temporal images. This work is the first attempt to apply a KCCA-based MAD algorithm to relative radiometric normalization.

## 1. INTRODUCTION

As the growing need of multi-sensor, multi-temporal remote sensing data together used for land-use and land-cover change (LUCC) monitoring, global resource environment analysis and climate change monitoring, relative radiometric normalization has practical significance for engineering applications.

Relative radiometric normalization can establish the correction equation for each corresponding multi-spectral band in multi-temporal remote sensing data directly applying the pixel value of the image, without requiring any other parameters, such as atmospheric conditions on the day that remote sensing data obtained. Relative radiometric normalization can not only correct the differences caused by atmospheric conditions, but can also reduce the radiation difference induced by other factors, such as integration of multi-temporal remote sensing data.

The pseudo-invariant feature (PIF) method is commonly used in relative radiometric normalization. The core task of this approach involves is the selection of pseudo-invariant feature points (PIFs) (Schott et al. (1988)). Canty (2004) presented a new idea for the automatic extraction of PIFs in linear relative radiometric normalization. This method is on the basis of the multivariate alteration detection (MAD) method proposed by Nielson (2002,1998). The transformation of MAD is applied to a pair of multi-spectral scenes, acquired at times  $t_1$  and  $t_2$ , from which PIFs are selected automatically and independently of surface properties, using the statistical properties of the data. Unlike the manual selection of PIFs, this method is simple, fast, fully automatic, and does not require the determination of the complex threshold. Canty (2008) further proposed the iteratively reweighted multivariate alteration detection (IR-MAD) method which improves on accuracy and robustness of the MAD algorithm with the iterative updating of

weights. At present, this method is commonly used in the relative radiometric normalization process.

MAD transformation is based on linear canonical correlation analysis (CCA). CCA method aims at finding a linear relationship of two sets of variables, so the spectral curves of the PIFs detected by MAD are all linearly changed. In relative radiometric normalization the whole scene is corrected by the linear relationship on the PIFs. The PIFs generally distribute only on asphalt roof, gravel surface, concrete apron, clean water, concrete and sand etc. However, the spectral curves of some other objects change nonlinearly over an interval  $[t_1, t_2]$ . Therefore, it is unsuitable to apply the linear correlation information on the PIFs to those nonlinearly changed objects in relative radiometric normalization.

In this work, we introduce a new method based on the kernel version of canonical correlation analysis (KCCA) that help us select PIFs, and we discuss the difference in the MAD versions based on CCA and KCCA for relative radiometric normalization. The results of this study indicate that the KCCA-based MAD can also be employed in relative radiometric normalization, and the PIFs it detected can well describe the nonlinear relationship between multi-temporal images.

This paper is organized as follows. Section 2 introduces the principle of CCA-based (standard) MAD and KCCA-based (kernel-based) MAD algorithms, and present the obtained data and the experiment subsequently. Finally, Section 3 concludes the paper.

## 2. METHOD

### 2.1 The CCA Transformation

Hotelling (1936) proposed that CCA is a multivariate feature extraction method that aims at finding the rotation of two sets of variables that maximizes their joint correlation. The essence of CCA involves the selection of a number of representative indicators (linear combination of variables) from the two groups of random variables. These indicators can express the correlation between both sets of variables. As Cauty (2004) suggested, we first form linear combinations of the intensities for all  $n$  channels in two images, acquired in times  $t_1$  and  $t_2$ . The random vectors namely  $\mathbf{x}$  and  $\mathbf{y}$  represent the images. It can be defined as:

$$\mathbf{u} = \mathbf{a}^T \mathbf{x} ; \mathbf{v} = \mathbf{b}^T \mathbf{y} \quad (1)$$

Where  $\mathbf{x}$ =image1  
 $\mathbf{y}$ = image2  
 $\mathbf{a}, \mathbf{b}$  =random vectors  
 $\mathbf{u}, \mathbf{v}$  = canonical variable

The pearson correlation coefficient is used to measure the relationship between image1 and image2. If we can determine a set of optimal solutions ( $\mathbf{a}, \mathbf{b}$ ) that can maximize the pearson correlation coefficient, then  $\mathbf{u}$  and  $\mathbf{v}$  reach the maximum correlation. The pearson correlation coefficient can be defined as:

$$\rho_{\mathbf{u},\mathbf{v}} = \text{corr}(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{a}^T \Sigma_{12} \mathbf{b}}{\sqrt{\mathbf{a}^T \Sigma_{11} \mathbf{a}} \sqrt{\mathbf{b}^T \Sigma_{22} \mathbf{b}}} \quad (2)$$

Where  $\Sigma_{11}$ = covariance matrix of  $\mathbf{x}$   
 $\Sigma_{12}$ = covariance matrix of  $(\mathbf{x}, \mathbf{y})$   
 $\Sigma_{22}$ = covariance matrix of  $\mathbf{y}$

As first described by Hotelling (1936), the condition of this optimization problem involves maximizing  $\mathbf{a}^T \Sigma_{12} \mathbf{b}$ , which is in turn subject to  $\mathbf{a}^T \Sigma_{11} \mathbf{a}=1$  and  $\mathbf{b}^T \Sigma_{22} \mathbf{b}=1$ . The solution is to construct the Lagrange equation. Thus we can derive the following:

$$\begin{bmatrix} 0 & \Sigma_{12} \\ \Sigma_{21} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \lambda \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \quad (3)$$

Where  $\lambda = \mathbf{a}^T \Sigma_{12} \mathbf{b}$

We assume that  $\Sigma_{11}$  and  $\Sigma_{22}$  are invertible matrices. Eq(3) is equivalent to:

$$\begin{aligned} B &= \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{bmatrix} \\ A &= \begin{bmatrix} 0 & \Sigma_{12} \\ \Sigma_{21} & 0 \end{bmatrix} \\ w &= \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \\ B^{-1}Aw &= \lambda w \end{aligned} \quad (4)$$

Where  $\lambda$  = the eigenvalue of  $B^{-1}A$

The formula above generates the solution to the eigenvalue problem. In accordance with the process described above, we can determine the eigenvalues  $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_n$  (where  $n$  is matrix dimension). Then the first group of canonical variable coefficients namely  $\mathbf{a}_1$  and  $\mathbf{b}_1$  can be calculated when  $\lambda$  is the maximum eigenvalue  $\lambda_1$ . Then the second group of canonical variable coefficients namely  $\mathbf{a}_2$  and  $\mathbf{b}_2$  can be computed when  $\lambda = \lambda_2$ , and so on.

### 2.2 The Kernel CCA Transformation

Prior to introducing the kernel version of CCA, we must first understand the kernel function. A kernel function can convert an  $n$  dimensional inner product in low-dimensional space into an  $m$ -dimensional inner product in high-dimensional space ( $m>n$ ). The kernel function is the theoretical basis for solving a complex classification or regression problem in high-dimensional feature space. The commonly used kernel functions are expressed as follows:

(1)Linear Kernel:

$$k(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y}$$

(2)Polynomial Kernel:

$$k(\mathbf{x}, \mathbf{y}) = (\gamma \mathbf{x}^T \mathbf{y} + c)^n$$

(3)Hyperbolic Tangent (Sigmoid) Kernel:

$$k(\mathbf{x}, \mathbf{y}) = \tanh(\gamma \mathbf{x}^T \mathbf{y} + c)$$

(4) Radial Basis Function Kernel:

$$k(\mathbf{x}, \mathbf{y}) = \exp(-\gamma \|\mathbf{x} - \mathbf{y}\|^2)$$

In the KCCA transformation the two sets of variables can be defined as follows:

$$\mathbf{u} = \mathbf{c}^T \phi_x(\mathbf{x}); \mathbf{v} = \mathbf{d}^T \phi_y(\mathbf{y}) \quad (5)$$

Where  $\phi_x = \mathbf{x}_n \rightarrow \mathbf{x}_m$ , low dimension to high dimension  
 $\mathbf{c}, \mathbf{d}$  =  $m$  dimension random vector  
 $\mathbf{x}$  =image1  
 $\mathbf{y}$  =image2

In the standard CCA transformation, the pearson correlation coefficient between  $\mathbf{u}$  and  $\mathbf{v}$  can be calculated with Eq.(4) However, we must first introduce a kernel function that help us analyze the nonlinear relationship between two images in the KCCA transformation.

$$k(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle \quad (6)$$

We can obtain the kernel matrix using the kernel function written in Eq.(5):

$$k_{i,j} = k(\mathbf{x}_i, \mathbf{y}_i) \quad (7)$$

Eq(4) can satisfy Eq(5) through the construction of a Lagrangian function.

$$\begin{aligned} L &= E[(\mathbf{u} - E[\mathbf{u}])(\mathbf{v} - E[\mathbf{v}])] \\ &\quad - \frac{\lambda_1}{2} E[(\mathbf{u} - E[\mathbf{u}])^2] \\ &\quad - \frac{\lambda_2}{2} E[(\mathbf{v} - E[\mathbf{v}])^2] \end{aligned} \quad (8)$$

In this expression, the derivative to be assumed is zero, and the canonical variable coefficients namely  $\mathbf{c}$  and  $\mathbf{d}$  are updated to the following:

$$\mathbf{c} = \sum_i \alpha_i \phi_x(\mathbf{x}_i) \quad (9)$$

$$\mathbf{d} = \sum_i \beta_i \phi_y(\mathbf{y}_i) \quad (10)$$

The following can be derived based on Eq (4,8,9):

$$\mathbf{u} = \sum_i \alpha_i \langle \phi_x(\mathbf{x}_i), \phi_x(\mathbf{x}) \rangle \quad (11)$$

$$\mathbf{v} = \sum_i \beta_i \langle \phi_y(\mathbf{y}_i), \phi_y(\mathbf{y}) \rangle \quad (12)$$

$$\text{var}(\mathbf{u}) = \mathbf{c}^T \text{var}(\phi_x(\mathbf{x})) \mathbf{c} = \alpha^T k_x k_x \alpha \quad (13)$$

$$\text{var}(\mathbf{v}) = \mathbf{d}^T \text{var}(\phi_y(\mathbf{y})) \mathbf{d} = \beta^T k_y k_y \beta \quad (14)$$

$$\text{cov}(\mathbf{u}, \mathbf{v}) = \mathbf{c}^T \text{cov}(\phi_x(x), \phi_y(y)) \mathbf{d} = \alpha^T k_x k_y \beta \quad (15)$$

$$\text{corr}(\mathbf{u}, \mathbf{v}) = \frac{\alpha^T k_x k_y \beta}{\sqrt{\alpha^T k_x k_x \alpha} \sqrt{\beta^T k_y k_y \beta}} \quad (16)$$

We can see that the look of the equations are basically similar as the standard CCA transformation. The correlation coefficient just replace covariance matrix with  $k_x k_y$ . So the solution process is similar too. We can solve the generalized eigenvalue problem of KCCA as below:

$$\begin{bmatrix} 0 & k_x k_y \\ k_y k_x & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \lambda \begin{bmatrix} k_x k_x & 0 \\ 0 & k_y k_y \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (17)$$

As in the standard CCA transformation, we easily obtain the eigenvalues  $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_n$  (where  $n$  is matrix dimension) in the KCCA transformation. A series of canonical variables can subsequently be derived from these eigenvalues. The correlation coefficient corresponding to the canonical variable can then be calculated.

### 2.3 Relative radiometric normalization

According to the iteratively reweighted multivariate alteration detection (IR-MAD) method proposed by Canty (2008), the PIFs must meet the following condition:

$$\text{MAD} = \mathbf{u} - \mathbf{v} \quad (18)$$

$$\omega = \sum_{i=1}^n \frac{(\text{MAD}_{ij})^2}{\sigma_{\text{mad}_i}} < \tau$$

Where  $\text{MAD} = |\text{MAD}_{ij}|_{\substack{i=1,2,\dots,N \\ j=1,2,\dots,M}}$

$$\sigma_{\text{mad}}^2 = 2(1 - \text{corr}(\mathbf{u}, \mathbf{v}))$$

$$\tau = \text{threshold}$$

Under both CCA and KCCA transformation methods, the points that satisfy the aforementioned formula are the PIFs. The radiation values of these points are linearly correlated in the images obtained at different time phases. The regression model can be established according to the linear relationship, by applying the radiation values of the PIFs.

$$y = p * x + q$$

$$p = \frac{\sum_{i=1}^M (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^M (x_i - \bar{x})^2} \quad (19)$$

$$q = \bar{y} - p * \bar{x}$$

where  $x, y$  = radiation values of pixels in image1, image2;  
 $x_i, y_i$  = the radiation value of PIFs  
 $\bar{x}, \bar{y}$  = the mean radiation value of PIFs  
 $p$  = gradient  
 $q$  = interception

Both  $p$  and  $q$  can easily be counted according to the formula above. Subsequently, the radiation correction of the entire image can be realized.

### 2.4 Examples and Results

In this work, two Landsat-8 satellite images of the region of Beijing, China, and two GF-1 satellite images of South China are obtained as test data at two different points in time, as shown in Figure 1. The object categories in the region are varied, which contains lake, vegetation, and artificial construction. Figure 1(a) was acquired on Oct. 3, 2013(t1), and Figure 1(b) was obtained on Nov. 4, 2013(t2). Although time

interval is very short, vegetation in North China turns yellow and begins to fall down in November. So the surface properties of vegetation significantly change between both images over the interval [t1, t2]. Figure 1(c) was acquired on Dec. 5, 2013(t3), and Figure 1(d) was obtained on Jan. 31, 2013(t4). The vegetation is always green.

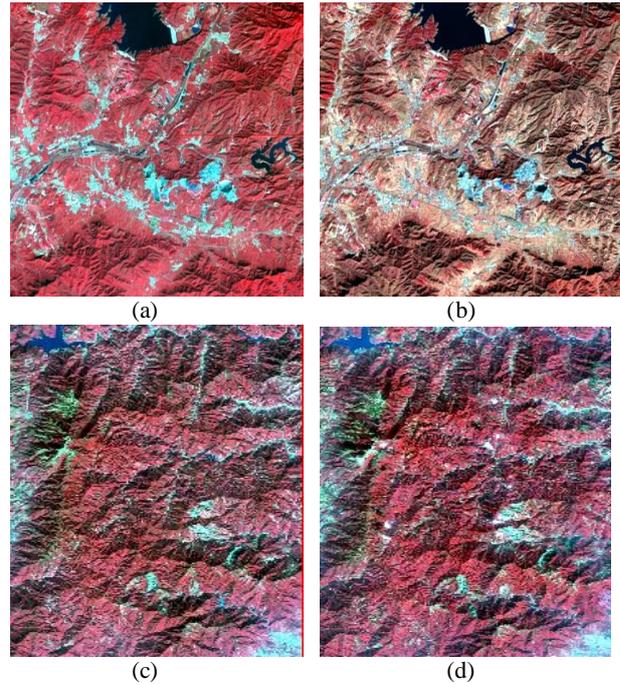
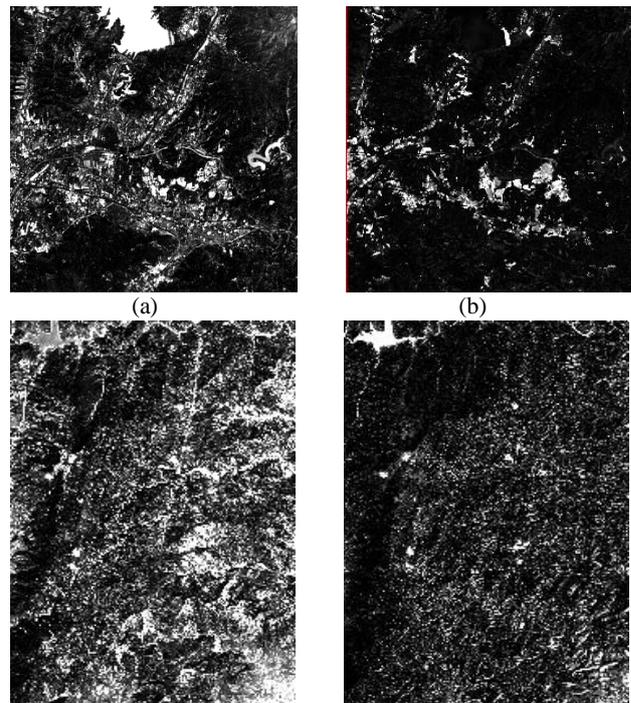


Figure 1. Experimental data. (a) is Landsat-8 data acquired in 20131003 ,(b) is Landsat-8 data acquired in 20131104, (c) is GF-1 data acquired in 20131205, (d) is GF-1 data acquired in 20130131

Subsequently, we choose PIFs according to the MAD methods. The weight of each pixel  $\omega$  is calculated, and the results are presented as follows Figure 2:



(c) (d)  
 Figure 2. MAD results. (a) is  $\omega$  map of Landsat-8 extracted by CCA. (b) is  $\omega$  map of Landsat-8 extracted by KCCA, (c) is  $\omega$  map of GF-1 extracted by CCA, (d) is  $\omega$  map of GF-1 extracted by KCCA

The standard MAD and the KCCA-based MAD algorithm detected 8,358 and 36,230 PIFs respectively for subsequent calculations in Landsat-8 image pairs, and detected 25,507 and 238,972 PIFs respectively in GF-1 image pairs. These PIFs are shown in Figure 3:

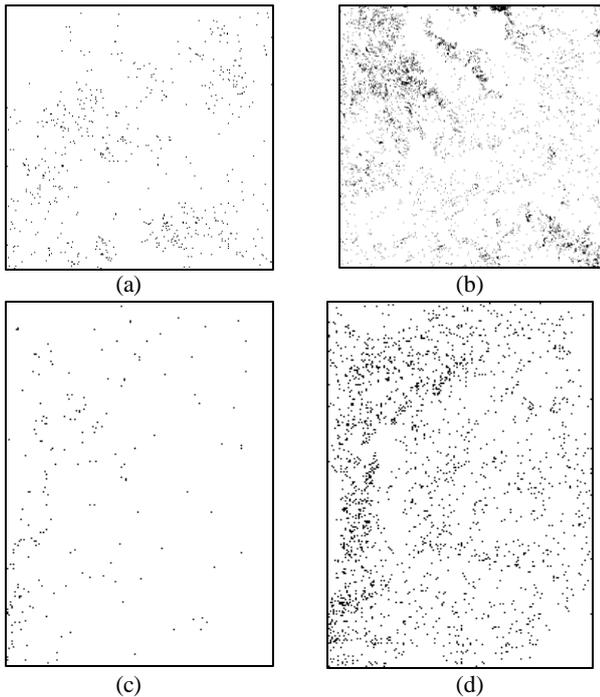


Figure 3. The distribution map of PIFs. (a) is results of CCA(Landsat-8); (b) is results of KCCA(Landsat-8); (c) is results of CCA(GF-1); (d) is results of KCCA(GF-1)

Relative radiometric normalization is performed with the regression model. The outcomes suggest that under the condition set in this study, the KCCA-based MAD algorithm can be satisfactorily applied to relative radiation normalization. The results of relative radiometric normalization are as follows Figure 4:

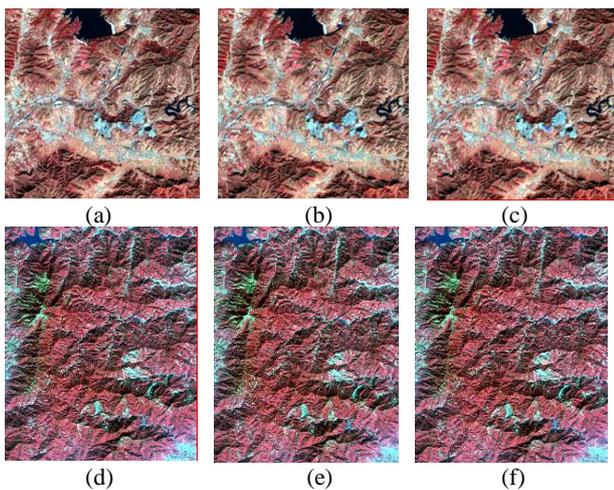


Figure 4. Relative radiometric correction results

(a) is the original Landsat-8 image (b) is relative radiometric correction result of CCA; (c) is relative radiometric correction result of KCCA; (d) is the original GF-1 image, (e) is relative radiometric correction result of CCA; (f) is relative radiometric correction result of KCCA

The standard MAD method is a commonly used relative radiometric correction algorithm, and its outcomes can be considered correct. We evaluate accuracy of the radiation correction results obtained with the KCCA-based MAD method based on the findings of the standard method. The difference between both images is calculated, and the D-values of the various bands are displayed in Figure 5. This leaves no-change area in black, and change areas in white. The discrepancies are mainly concentrated in the vegetation area.

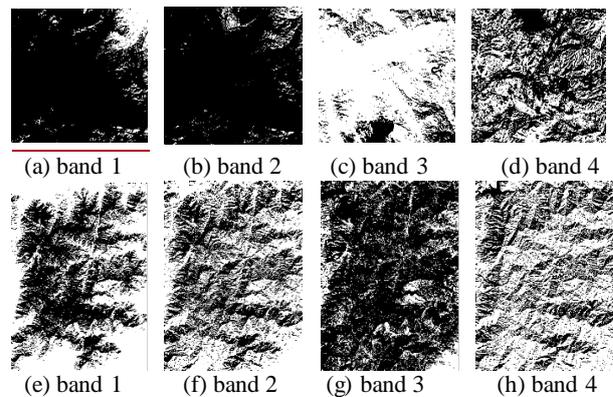


Figure 5. The difference map of two methods ( $result_{KMAD} - result_{MAD}$ ). The first row is the difference of Landsat-8 data, the second row is the difference map of GF-1 data

### 3. CONCLUSIONS

The study results indicate that both CCA-based and KCCA-based MAD methods can be applied to relative radiometric normalization. The KCCA-based MAD algorithm detected more PIFs than the former, and the PIFs it detected can well describe the nonlinear relationship between multi-temporal images. These PIFs can not only be used in relative radiometric normalization, but also can be further applied to multivariate change detection. In future studies, we can focus on how to filter these points to generate enhanced results.

### ACKNOWLEDGEMENTS (OPTIONAL)

This work was supported by the National High Technology Research and Development Program (863) of China under Grant 2013AA12A301.

### REFERENCES

- Nielsen, A. A., Conradsen, K., & Andersen, O. B. (2002). A change oriented extension of EOF analysis applied to the 1996–1997 AVHRR sea surface temperature data. *Physics and Chemistry of the Earth*, 27(32–34), pp.1379–1386.
- Nielsen, A. A., Conradsen, K., & Simpson, J. J. (1998). *Multivariate alteration detection (MAD) and MAF post-processing in multispectral, bitemporal image data: New approaches to change detection studies. Remote Sensing of Environment*, 64, pp.1–19.

Canty, M. J., Nielsen, A. A., & Schmidt, M. (2004). *Automatic radiometric normalization of multitemporal satellite imagery*. *Remote Sensing of Environment*, 91(3 - 4), pp.441–451. Internet <http://www.imm.dtu.dk/pubdb/p.php?2815>

M. J. Canty and A. A. Nielsen. (2008). *Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation*. *Remote Sensing of Environment*, vol. 112, no. 3, pp. 1025–1036. Internet <http://www2.imm.dtu.dk/pubdb/p.php?5362>

Nielsen, A. A., Vestergaard, J.S. (2013) *A kernel version of multivariate alteration detection*. *IEEE Int. Geosci. Remote Sens. Symp. IGARSS, Melbourne(AUS)*, pp.3451-3454.

Schott, J. R., Salvaggio, C., & Volchok, W. J. (1988). *Radiometric scene normalization using pseudo-invariant features*. *Remote Sensing of Environment*, vol.26, pp.1– 16.

Hotelling, H. (1936). *Relations between two sets of variates*. *Biometrika*, vol.28, pp.321–377.

P. L. Lai and C. Fyfe. (2000). *Kernel and nonlinear canonical correlation analysis*. *International Journal of Neural Systems*, vol. 10, no. 5, pp. 365–377

M. J. Canty and A. A. Nielsen. (2012). *Linear and kernel methods for multivariate change detection*. *Computers & Geosciences*, vol.38, pp. 107-114