Functional Data Analysis: An Initiative Approach for Hyperspectral Data

S Sugianto\textsuperscript{1*}, M Rusdi\textsuperscript{1}, Syakur\textsuperscript{2}

\textsuperscript{1} Remote Sensing and Cartography Laboratory, Faculty of Agriculture, Syiah Kuala University, Darussalam 2311, Indonesia.
\textsuperscript{2} Soil Science, Faculty of Agriculture, Syiah Kuala University, Darussalam 2311

\textsuperscript{*}sugianto@unsyiah.ac.id

Abstract. This paper summarizes initiative approach to analyze spectral variation and distribution for given spectral range of hyperspectral data using Functional Data Analysis (FDA). The FDA can be used to analyse hyperspectral data which restrain information as curves or functions. Vertical shift and horizontal shift are two examples of directions of variation, modes of variation in hyperspectral data. Shifts of curves in certain wavelength obtained from remote sensing are often used as a key indicator to obtain information from a target. As hyperspectral data use radiance or reflectance defined as the ratio of intensity of light received or reflected from the target to the intensity of the light incident on it, the radiance or reflectance values are then used to derive information about the composition of samples. In this paper, FDA method is introduced to present an alternative analysis in hyperspectral data as a function of spectral reflectance, distribution and variance of hyperspectral radiance or reflectance data.

1. Background

As hyperspectral data consists of contiguous narrow spectral bands that can form a curve or function of wavelength, can be considered as functional data. To analyze curve or image data we rely on characteristic shapes. Vertical shift and horizontal shift are two examples of directions of variation, modes of variation, in functional data, and they are of interest in many applications. These shifts are considered as key factors for data interpretation either as pixel based or as object-based hyperspectral image classification. Similarly, shifts of curves in certain wavelength obtained from remote sensing are often used as a key indicator to obtain information from a target. Compared to conventional analysis of hyperspectral remote sensing data is dominated by multivariate analysis, or data fusion methods to spectral information analysis based on spectral response [1] or combination of hyperspectral data and LIDAR [2], hyperspectral data is more complex to extract the earth surface phenomena. To accomplish this complexity of hyperspectral data, statistical inference is often necessary, ranging from simple descriptive statistics to multivariate geostatistics. Multivariate analysis permits the study of data, based on a finite set of observations, but not as a functional representation.

Hyperspectral remote sensing, a measured of extents of spectral reflectance is known as well as functions. The reflectance can be assumed as a function of wavelength and it could represent of the target characteristics. FDA is therefore a suitable approach for the analysis of this data. Functional data is originally assumed by a number of scientists associated with continuous-
time monitoring processes, whose final outputs are samples of functions [3]. The increase in functional data collection follow the development of very detailed real-time measurement instrument, like hyperspectral sensors. The theoretical development of hyperspectral data in this field is clearly gaps behind the potential for its applications. Consequently, some effort is required to adjust a standard statistical technique for functional data, for use in hyperspectral remote sensing.

A simple functional relationship is typically written as: \( y = f(x) \), where \( y \) is the dependent variable and \( x \) is an independent variable. To understand the characteristics of functional data, the functional relationship between dependent and independent variables can be developed and considered using some subtle but powerful data analysis tools, such as multivariate data analysis, using parametric or non-parametric analysis [4]–[6].

2. Functional data in remote sensing

The functional association between spectral reflectance and the factors affecting it are more complex. For example, the radiance value recorded by a sensor, after noise correction, can be considered a function of position (\( x \)), time (\( t \)), wavelength (\( \lambda \)) and observing geometry (\( \theta \)) of the earth component [7], and stated as

\[
R = f(x, t, \lambda, \theta).
\] (1)

An adequate change in one of \( x, t, \lambda \) or \( \theta \) can cause a change in reflectance (\( R \)). Viewing geometry (\( \theta \)) is one of the factors that can be controlled (minimised or maximised) by analysts.

As a first step in building up more complex predictive models of functional relationships, simple models are used to analyse the functional relationship between the factors affecting it, such as the directional spectral reflectance function, which involves the dependent variable (\( R \)) and the independent variables (\( x, t, \lambda \) or \( \theta \)). The spectral reflectance (\( R \)) can be reflected as a function, for example, of vegetation cover type for a multi-temporal variable and measured as a vegetation index [8].

2.1. Functional scheme in remote sensing data

There are two central informational paradigms in environmental sciences: advancement and shape [9]. In remote sensing data analysis, there is resultant to function and shape and land-use and land cover, and classification [10]. Spectral classes can be readily identified based on land cover (as structure classes) or different classes due to sensor scanning mode that can produce difference spectral reflectance ‘classes’ even though at a single land cover. The variety in these classes can be thought of as a function. A ‘function’ is theoretically much more complex than a ‘shape’ since the latter is reorganized in terms of measures and it interacts between land-cover components. Spectral reflectance data obtained from remote sensing data are considered as single entities rather than as continuous. This assumption remains because the way multi-spectral data is measured and presented, as an abrupt connection of spectral measurements between bands (Fig.1). In contrast, hyperspectral data can generate continuous curves with contiguous channels.

The functional data analysis approach retains the structure of the data. It treats each function as a series of data points that are likely to relate, rather than as series of independent data points. By assuming the data points, the structural features of the data in any analysis can be retained, without violating statistical assumptions of independence between the samples (Figure 2).
Figure 1. Three common spectral curves representing soil, water and vegetation derived from multi-spectral and hyperspectral at Iffley site, New South Wales Australia. Both curve plots represent the concept of ‘functional data’ in remote sensing.

Figure 2. Illustration of hyperspectral dataset, its spectral and functional space representation showing the functional relationships $y_1$-$y_n$ as dependent variable for each dependent variable $t_1$-$t_n$.

3. The FDA Method
FDA consists of two principal roots, data and bases, followed by the different analyses available (Figure 3). In FDA, discrete data can be considered as a functional form of its origin, even though experimental data are usually recorded separately. However, distinct raw data may contain an error, so altering raw data into essentially into a smooth practical form requires proficient smoothing procedures. The distinctive of functional data analysis approach is to proper of individual curve independently, with the expansion in bases functions.

A fascinating idea is that the FDA approaches can work on the statistical figures of the bases extension, a prominent less computational problem. A supplementary benefit of dealing with functions is its ability to accommodate functional pre-processing such as derivation, integration, regularisation, smoothing, prediction and the modelling of data in a functional manner.
Spectral data collected from field and space-borne remote sensing data can be considered as discrete points of reflectance intensity at different scanning angle or bands \( (t_j) \) and the relationship between spectral reflectance and scanning angle or number of bands can be formulated as:

\[
y_j = x(t_j) + \epsilon_j,
\]

where \( \epsilon_j \) is the unobserved error component and the sequence of \( y_j \) are the observed outcomes of the process. A major underlying idea is that the function, \( x(t) \), is differentiable to some or other order. A key aspect for the success of functional statistics consideration in this experiment is the use of the implicit sampling rate, determined by the sequence of angle differences \( (t_j - t_{j-1}) \), to model the linear function of spectral reflectance under different angle positions.

### 3.1. Bases functions

The alteration of raw data to functional procedure needs two-steps: selecting and defining a set of bases functions, and calculating the finest linear combination for each fixed of reflectance spectra. Selecting comparatively few bases functions in certain bases representations of a prototypical, leading to asymmetrical estimation. The bases functions are formulated as:

\[
x_j(t) = \sum_{k=1}^{K} c_{ik} \phi_k(t),
\]

where \( \phi_k(t) \) is the fixed-bases functions and \( c_{ik} \) is coefficients \( \phi_k(t) \). The extent of the data \( y_j = x(t_j) \) are curved, rather than precisely replicated or interpolated, is determined by total of \( K \) bases functions. Precise representation is usually possible if \( K = n \), one can select the coefficients \( c_i \) to produce \( x(t_j) = y_j \) for each \( j \). The transformation of data into a functional data object quantities to computing and storing the coefficients \( c_{ik} \) of the bases representation in an \( I \) by \( K \) matrix. A bases is specified by four elements, the bases type, the range of argument values \( [T_0, \ldots \]
To perform a functional data analysis, an appropriate set of bases functions need to be determined.

### 3.2 Smoothing bases function

Since spectral data frequently contains noise, some smoothing is needed. A common smoothing technique used in remote sensing data is the Savitzky–Golay filter [12], which is based on a least-squares approximation of short segments of the spectrum. But most of the smoothing used on hyperspectral data mentioned do not consider spline functions. The use of the cubic spline technique in FDA overcomes some of the problems encountered previously when smoothing hyperspectral data. Figure 4 shows how to construct a smooth curve from B-splines with given coefficients.

![Figure 4](image)

**Figure 4.** Typical bases smoothing of cotton spectra with different bases function at 10°. The plot on the left uses 10-bases functions, while the plot on the right uses 20 bases functions. RMS residual is given to show variation between the modeled and the original data value [13]

By implementing the smoothing principle, the smoothing function can be applied for hyperspectral data. If $x_{ij}$ represents the matrix of observations of spectral reflectance data in which the $i$'th reflectance function of wavelength $t_j$ at an angle observation, it may obtain the bases representation of $x(t)$ by a least squares criterion, that is a minimization for each $i$ with respect to the $c_{ik}$'s (here $j \leq k$).

$$\sum_{j=1}^{f} \left( x_{ij} \sum_{k=j}^{c} c_{ik} \phi_k(t_j) \right)^2$$

where $k$ defines the smoothness of fit. A minor $k$ results in a smooth fit, while fitting using a large $k$ is less smooth. This can be done by a least squares criterion with an added roughness penalty by minimization for each factor of $k$ where $g(x)$ is a measure of roughness and $k$ is the smoothing weight.

### 3.3 Functional principal component analysis (fPCA)

The fPCA in this approach is described in terms of spectral data. It can be expected that the spectral reflectance function $y(t,x)$ can be constructed as a bases analysis in fPCA at point $x$ of wavelength at band or angle $t$. The spectral measurement is taken over the target $x=1,\ldots,p$ at bands or angle $t=1,\ldots,n$. Therefore, the concept mentioned above is collections [$z(t,x)$: $x=1,\ldots,p$] of readings $z(t,x)$ reserved at each $n$ time of spectral bands or view angles scanning mode $t$, and which have been centred on their bands or angle averages. These collections can be assumed of as a $p$ by $l$ vector $\tilde{z}(t)=[z(t,1),\ldots,z(t,p)]^T$ creating a group of points about the
source of a $p$-dimensional Euclidian space $E_p$. The symbol $T$ represents the reorder process. From this point in $E_p$ a $p$ by $p$ symmetric matrix $S$ and the value of $S$ is determined by the sum of vector matrix $z$ at band or sensor view angles $t$:

$$S = \sum_{i=1}^{n} z(t) z^T(t).$$  \hspace{1cm} (5)

This matrix has a fixed of orthonormal eigenvectors $\{e_j(1), \ldots, e_j(p)\}^T$, $j=1, \ldots, p$. These are the observed orthogonal functions. They are observed because they arise from data, and orthogonal because they are uncorrelated.

As with the multivariate PCA, the first functional PCA band comprises the largest proportion of data adjustment and the second functional PCA contains the second largest data adjustment. The last functional PCA functions perform noisy because they contain very little inconsistency, much of which is due to noise in the original spectral data. Functional PCA can be implemented as follows. Deliberate an integral transform of all $x_i(t)$ values $y$ as:

$$y = \int \xi(t)x_i(t) dt,$$  \hspace{1cm} (6)

where the mutual weight function $\xi(t)$ is determined by lessening, $\sum_{i=1}^{I} y_i^2$, within a piece of functions subject to the limitation $\int \xi^2(t) dt = 1$. The reducing function is denoted $\xi_1(t)$ and known as the first harmonic function. The conforming reducing $y_i$'s are the principal component scores. The minimizing problem is considered again to obtain another weight function $\xi_2(t)$ subject to the added orthogonal restriction:

$$\int \xi_1(t)\xi_2(t) dt = 0.$$  \hspace{1cm} (7)

This is the second principal component function. This process can be continued to derive a sequence of principal component functions (or harmonics) $\{\xi_k(t); k=1,2,\ldots,K\}$ orthonormal to one another. Here $K$ must be less than $I$. Each harmonic is calculated a definite portion of the variability in the original functional data substances, decreasing for increasing $K$.

Another way to consider functional principal components is to find least squares approximating functions in terms of an orthonormal bases $\{\xi_k(t); k=1,2,\ldots,K\}$, i.e.

$$\hat{x}_i(t) = \sum_{k=1}^{K} y_k \xi_k(t)$$  \hspace{1cm} (8)

therefore:

$$\sum_{i=1}^{I} \int (x_i(t) - \hat{x}(t))^2 dt$$  \hspace{1cm} (9)

is minimized. An optimal choice of these bases is identical to the solution of the problem above (Equation 8 and 9). Since these functions are determined by the data, they are often named empirical orthonormal functions [14].

A principal curve determination is not unique and solutions that are just as good, but perhaps more easily interpreted, may be obtained by rotation. Mathematically the principal component problem above leads to a so-called functional eigenvalue problem for the covariance function:

$$\nu(s,t) = \frac{1}{I} \sum_{i=1}^{I} \int x_i(s)x_i(t)$$  \hspace{1cm} (10)
The numerical implementation of a covariance function is based on an approximation that brings the problem into the realm of standard linear algebra. The simplest approach is discretization of the functional objects into a fine grid spanning the interval in question, leading to a standard matrix eigenvalue problem [15].

3.4. Functional linear model and functional analysis of variance

For this context, describing and comparing different estimation procedure such as incomplete least squares and principal component regression, should be taken into consideration together the number of descriptive variables (which may overdo the model size) and the high correspondence between these variables [16]. Several authors have developed models which allow the description of the “functional” nature of explanatory variables of chemometrics data and temporal Landsat images [17], [18].

Typical the correlation between response of spectral reflectance, $Y(t)$, and forecaster, $X(t)$, together measured completed period by the calculation,

$$Y(t) = a_0(t) + \beta_0^T X(t) + \epsilon(t)$$  \hspace{1cm} (11)

where $a_0(t)$ is a smoothed function of $t$, $\beta_0$ is a fixed but unidentified vector of regression coefficients and $\epsilon(t)$ is a nil mean fixed Gaussian process.

To describe the estimation procedure using a functional linear model, the normal equations described above cannot be resolved directly. There will constantly be more indefinite parameters than subjects in the approximation of $\beta(t)$, succeeding in an infinite number of solutions with a perfect but pointless fit. Therefore, the bases functions should be selected to reflect the appearances of the data. An unsuitable choice of bases functions would need a large number of functions in order to sufficiently model the data see Eq. 4. Having chosen the bases function, the functional linear regression can be generated based on vector bases coefficients and formulated as follows: $Y = \Theta(\{X(t), t, \nu\} + \gamma$, where $\Theta$ is a constant linear model and $\gamma$ is a centred actual arbitrary variable with variance $\Phi^2$ and self-determining from $X$. This model may also be formulated as:

$$Y = \int_0^1 \psi(t) X(t) dt + \epsilon$$ \hspace{1cm} (12).

This functional model may be adopted to solve problems in multi-angular hyperspectral data where different sensor view angles may not provide the same spectral reflectance in all directions. In this context, the principles of functional linear models mentioned above can be used to assess the effects. The simple hypothesis of the model is that the random variable of spectral reflectance, for different azimuth or zenith angles, may be predicted through functional regression models. The factor $(t)$ can be considered as sensor angles or bands (channels). Thus, functional linear models can accommodate the functional regression model of multi-angular spectral reflectance data either based on the bands or the view angles. To examine the significance of the functional linear model, functional analysis of variance (FANOVA) can be adopted as is now discussed.

3.5. Functional analysis of variance (FANOVA)

The FANOVA [3], [14] is a generalisation of the standard ANOVA. In this example, the individual soil and cotton spectrum is considered as a categorical variable in wavelength or azimuth/zenith bases analysis. The spectrum, $ref$, is substantial at a specific wavelength or angle, $t$, for each fixed $t$ indicated by
\[ \text{ref}_{(ig)}(t) = \mu(t) + \alpha_{(ig)} + \varepsilon_{(ig)}(t); \]
\[ i = 1, ..., I, g = 1, ..., G \]

where \( \mu \) is the outstanding mean function, \( \alpha_{(g)} \) signifies the specific effect on the reflectance attribute (angle or bands), \( \varepsilon_{(ig)}(t) \) is the error (residual). In order to uniquely identify the reflectance, it requires the constraint \( \sum_{g=1}^{G} \alpha_g(t) = 0 \) for entirely \( t \).

The remaining function \( \varepsilon_{ig} \) is the unsolved dissimilarity detailed for the \( i \)th band or angle within the functional observations. To test the standard null hypothesis \( H_0 = \forall t(t) = . . . = \alpha_G(t) = 0 \) for all \( t \) of spectral reflectance as result of multi-view angles, and \( H_1 = \forall t(t) \neq . . . \neq \forall \alpha_G(t) \neq 0 \).

By assuming that the function reflectance \( x(t) \), \( \alpha_{g(t)} \) and \( \varepsilon_{ig(t)} \) are spanned by B-Spline bases \( B_k(t) \). Thus, \( x(t) = E_k m_B B_k(t) \), \( \forall t(t) = I_{\forall g} B_k(t) \) and \( \varepsilon_{ig} = E_{\forall ig(t)} \). So, the equation becomes:

\[ \sum_{k} C_{igk} B_k(t) = \sum_{k} m_k B_k(t) + \sum_{k} \alpha_{gk} B_k(t) + \sum_{k} \varepsilon_{igk} B_k(t). \]

The orthogonality of the B-Spline bases results, \( C_{igk} = m_k + \forall g \alpha_{gk} \) for entirely \( K \).

The constraint the equation becomes;

\[ \sum_{g} \alpha_{gk} = 0 \text{ for all } k. \]

Reducing each \( k \) using the normal least squares fitting standard \( \sum_{g} (c_{ig} - m_k - \alpha_{gk})^2 \) focuses to the constraint \( \varepsilon_{gk} = 0 \) provides the least square estimates \( \hat{m}_k \) and \( \hat{\alpha}_{gk} \). This then gives estimate \( \hat{\mu} \) and \( \hat{\alpha}_g \) of functional parameter and \( \alpha_g \) correspondingly:

\[ \hat{\mu}(t) = \sum_k \hat{m}_k B_k(t), \quad \hat{\alpha}_g(t) = \sum_k \hat{\alpha}_{gk} B_k(t). \]

To assess the goodness of fit for the model we use the \( R^2 \):

\[ R^2(t) = \frac{(SSY(t) - (SSE(t))}{SSY(t)}, \text{ where } SSY(t) = \sum_{g=1}^{G} \sum_{i=1}^{ng} (y_{ig}(t) - \hat{\mu}(t))^2, \]

\[ SSE(t) = \sum_{g=1}^{G} \sum_{i=1}^{ng} (y_{ig}(t) - \hat{\mu}(t) - \hat{\alpha}_g(t))^2, \]

and, finally, the F ratio can be calculated as

\[ F - \text{ratio} = \frac{(SSY - SSE)}{SSE}. \]

By knowing the F-ratio, the significance of the difference among functions of spectral reflectance of hyperspectral data can be determined and compared with standard statistical procedures and significance tests.

4. Conclusion

Remote sensing uses spectral radiance or reflectance to define the intensity of light as a diagnostic measure. This spectral radiance or reflectance can be considered as functional data analysed with FDA. Functional data analysis approaches such as functional bases smoothing, functional principal component analysis, and functional linear representations can be utilised to analyse the dissemination and variance of multi-angular spectral radiance or reflectance data.

The steps involved in the functional data analysis process are the conversion of data to a functional data form, the selection of the bases function, and the analysis of the data in
functional models such as functional PCA, functional linear models, and the analysis of variance.

Two features are common to any functional data analysis, a robust relationship with the multivariate statistical model and the requirement for smoothing. The robust relationship with multivariate statistics come with the fact that approaches such as principal component analysis, multivariate linear modelling, and canonical correlation analysis can be adopted within the FDA model.

5. References

[1] Yusuf Y, Sumantyo J T S and Kuze H 2013 “Spectral information analysis of image fusion data for remote sensing applications,” Geocarto Int., vol. 28, no. 4, pp. 291–310
[2] Liu X and Bo Y 2015 “Object-Based Crop Species Classification Based on the Combination of Airborne Hyperspectral Images and LiDAR Data,” Remote Sens., vol. 7, no. 1, pp. 922–950.
[3] Laukaitis A and Račkauskas A 2005, “Functional data analysis for clients segmentation tasks,” Eur. J. Oper. Res., vol. 163, pp. 210–216.
[4] Besse P C, Cardot H and Ferraty F 1997 “Simultaneous non-parametric regressions of unbalanced longitudinal data,” Comput. Stat. Data Anal., vol. 24, pp. 255–270.
[5] Bosq D 1991 “Modelization, Nonparametric Estimation and Prediction for Continuous Time Processes,” in Nonparametric Functional Estimation and Related Topics, Roussas G Editor Dordrecht: Springer Netherlands, 509–529.
[6] Fernandes R and Leblanc S G 2005 “Parametric (modified least squares) and non-parametric (Theil–Sen) linear regressions for predicting biophysical parameters in the presence of measurement errors,” Remote Sens. Environ., vol. 95, pp. 303–316.
[7] Kumar K, Schmidt K, Dury S and Skidmore A 2001 “Imaging spectrometry and vegetation science,” in Imaging Spectrometry, S. M. De Meer, Freek D. van der; Jong, Ed. Dordrecht: Springer Netherlands, pp. 111–155.
[8] Cardot H, Faivre R and Goulard M 2003 “Functional approaches for predicting land use with the temporal evolution of coarse resolution remote sensing data,” J. Appl. Stat., vol. 30, no. 10, pp. 1185–1199.
[9] Kellman K, Kellman M and Tackaberry R 1997 Tropical Environments: The Functioning and Management of Tropical Ecosystems. London: Routledge
[10] Gong P 1992, “Frequency-based contextual classification and gray-level vector reduction for land-use identification,” Photogramm. Eng. Remote, vol. 58, pp. 423–437
[11] Abdou W A 2001 “Ground measurements of surface BRF and HDRF using PARABOLA III,” J. Geophys. Res. Atmos., vol. 106, no. D11, pp. 11967–11976
[12] Savitzky A and Golay M J E 1964 “Smoothing and Differentiation of Data by Simplified Least Squares Procedures.,” Anal. Chem., vol. 36, no. 8, pp. 1627–1639.
[13] Sugianto S and Laffan S 2012 “Functional data analysis of multi-angular hyperspectral data on vegetation,” Aceh Int. J. Sci. Technol., vol. 1, no. 1, pp. 30–39
[14] Ramsay J O and Silverman B 2007 Applied functional data analysis: methods and case studies. New York: Springer-Verlag, 2007.
[15] Ramsay J O and Dalzell C J 1991 “Some tools for functional data analysis,” J. R. Stat. Soc. Ser. B, vol. 53, no. 3, pp. 539–572
[16] Frank I E and Friedman J H 1993 “A Statistical of Some Chemometrics View Regression Tools,” Technometrics, vol. 35, no. 2, pp. 109–135
[17] Cardot H, Ferraty F, Mas A and Sarda P 2003 “Testing Hypotheses in the Functional Linear Model,” Scand. J. Stat., vol. 30, no. 1, pp. 241–255
[18] Cardot H, Ferraty F, Mas A and Sarda P 2003 “Spline estimators for the functional linear model,” Stat. Sin., vol. 13, pp. 571–591