

Center for Quality and Applied Statistics
Kate Gleason College of Engineering
Rochester Institute of Technology

Application of Nonnegative Principal Component Analysis in Hyperspectral Imaging

Peter Bajorski
Center for Quality and Applied Statistics
Rochester Institute of Technology
Peter.Bajorski@rit.edu

Technical Report 2006–3

July 2006



Application of Nonnegative Principal Component Analysis in Hyperspectral Imaging

Peter Bajorski

Graduate Statistics Department
Rochester Institute of Technology
98 Lomb Memorial Drive, Rochester, NY 14623-5604;
Email: Peter.Bajorski@rit.edu

ABSTRACT

The classic PCA (Principal Component Analysis) has been applied in hyperspectral imaging with varying success. One obstacle in its application is the potential physical interpretation of the principal components, which is questionable unless the principal component coefficients are nonnegative. In this paper, we show hyperspectral imaging applications of a recently developed methodology of nonnegative PCA, which overcomes this difficulty by constructing nonnegative principal components. We construct an approximation of a physics-derived target space, and suggest some interpretations of the resulting components.

Keywords: nonnegative PCA, hyperspectral image, physics-derived target space.

I. INTRODUCTION

The purpose of this paper is to investigate the methodology of nonnegative PCA in the context of hyperspectral images. The classic PCA (Principal Component Analysis) has been applied in hyperspectral imaging with varying success. One obstacle in its application is the potential physical interpretation of the principal components, which is questionable unless the principal component coefficients are nonnegative. In Section II, we briefly explain the methodology for Nonnegative Principal Component Analysis as recently presented in [1]. In Section III, we show some numerical results using a physics-derived target space developed in [2]. Conclusions are formulated in Section IV.

II. METHODOLOGY FOR NONNEGATIVE PCA

In this section, we explain the methodology for Nonnegative Principal Component Analysis as presented in [1]. Let Y be a p -dimensional random vector. In the hyperspectral imaging context, the realizations of Y are pixels of a hyperspectral image with p spectral bands. In classic PCA (Principal Component Analysis), we find linear combinations $X_i = e_i^T Y$, $i = 1, \dots, p$ that maximize variability, where e_i are eigenvectors of the variance-covariance matrix $\text{Var}(Y)$. Writing \mathbf{P} for a matrix of eigenvectors e_i as columns, we obtain

$$X = \mathbf{P}^T Y \quad \text{or} \quad Y = \mathbf{P} X ,$$

that is,

$$Y = e_1 X_1 + \dots + e_p X_p \tag{1}$$

The above can be regarded as a latent model, because all terms on the right-hand side are not directly observed. The representation (1) may assist in understanding variability in spectra described by Y . For example, the p -dimensional vectors e_i may represent features present in a hyperspectral image, while X_i could represent presence level of those features. For interpretation purposes, it would be desirable to have e_i vectors with nonnegative coordinates. For example, the vectors e_i could then be interpreted as spectral curves. Unfortunately, in classic PCA with e_i being orthogonal eigenvectors, we cannot expect the feature vectors (e_i) to be nonnegative. In order to overcome this difficulty, a method for nonnegative PCA was developed and investigated in [1]. We now briefly describe that methodology, since the paper [1] has not been published yet.

We assume that an observed p -dimensional random vector Y can be represented as

$$Y = \mathbf{A}X \quad (2)$$

where \mathbf{A} is a nonsingular matrix of constants, the components of X are uncorrelated and standardized so that $\text{Var}(X) = 1$, and the components of X are nonnegative with probability 1, that is, $P(X \geq 0) = 1$. We refer to X as a vector of nonnegative canonical components. For $1 \leq k \leq p$, define $\lambda_k = \sum_{j=1}^p a_{jk}^2$ to be the square of the length of the k -th column of \mathbf{A} , and put $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_p)$, that is, a diagonal matrix with λ_k values on the diagonal. Let $W = (W_1, \dots, W_p)^T = \mathbf{\Lambda}^{1/2} X$, that is, $W_j = \lambda_j^{1/2} X_j$. W_j is the nonnegative principal component (PC) corresponding to the nonnegative canonical component X_j . Clearly, $\text{Var}(W_j) = \lambda_j$.

If the components of X are ordered such that $\lambda_1 \geq \dots \geq \lambda_p$, then W_j is termed the j -th nonnegative PC. We can write $Y = \mathbf{A}\mathbf{\Lambda}^{-1/2}W = \mathbf{D}W$, where the columns of the matrix $\mathbf{D} = \mathbf{A}\mathbf{\Lambda}^{-1/2}$ have unit length. We have the following counterparts of the properties of the standard PCs:

$$\sum_{j=1}^p \text{Var}(Y_j) = \sum_{j=1}^p \text{Var}(W_j), \quad \text{corr}(W_j, Y_k) = \frac{d_{jk} \lambda_j^{1/2}}{\text{Var}(Y_k)} \quad (3)$$

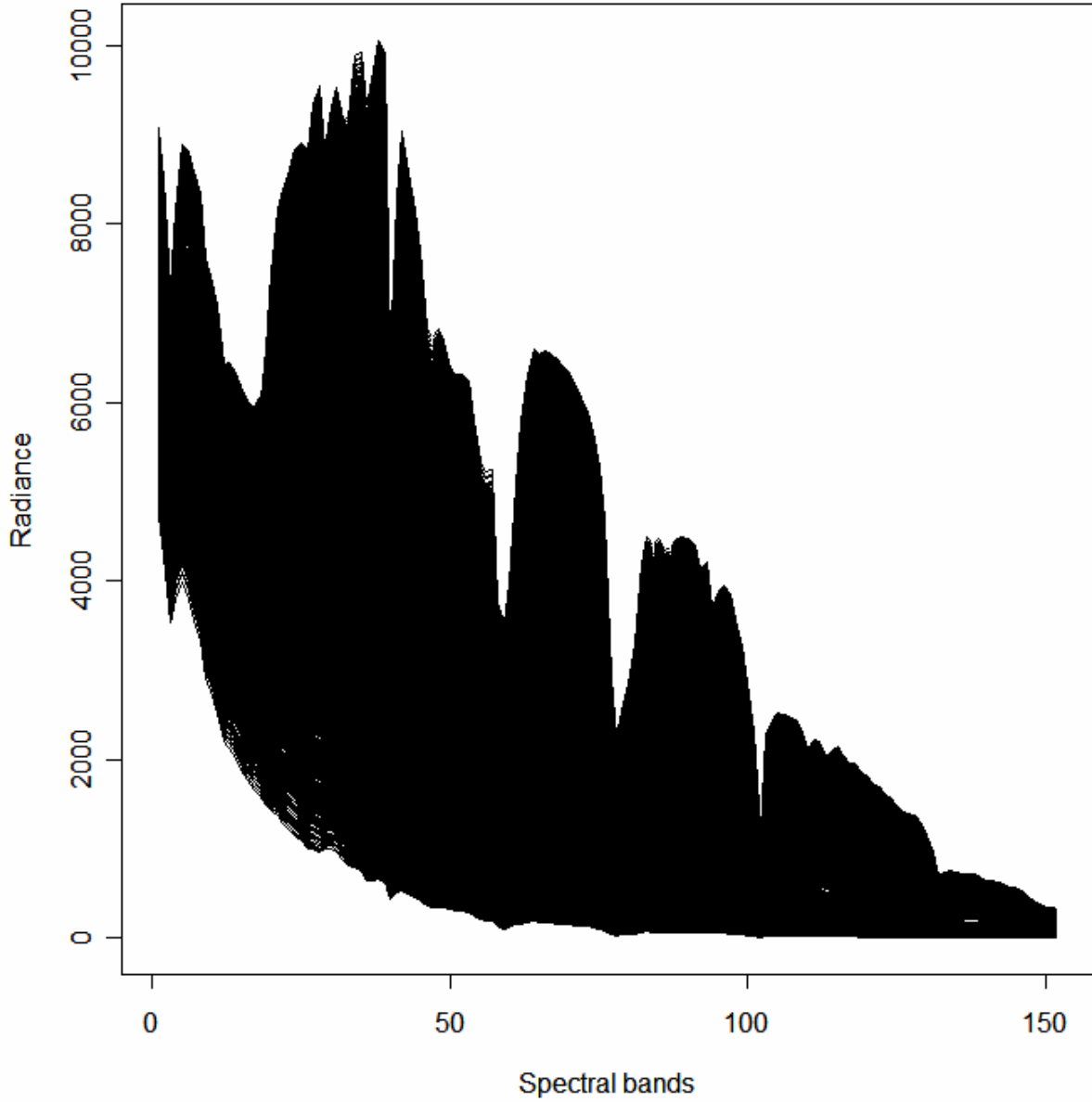
for $1 \leq j, k \leq p$, where d_{jk} denotes the (j, k) th component of \mathbf{D} . The first result in (3) asserts that the total variability in Y equals the total variability in W .

In classic PCA, we can always construct PCs. On the other hand, the nonnegative PCs may or may not exist for a given random vector Y . A necessary and sufficient condition for uniqueness of the nonnegative PCs is given in [1]. A method for estimation of nonnegative PCs and a related numerical algorithm are also described in [1].

III. NUMERICAL RESULTS

This section shows some numerical results on the estimation of nonnegative PCs using a physics-derived target space considered in [2]. The target space consists of 34,650 spectra, which are shown in Figure 1. More details about this target space can be found in [2].

Figure 1. 34,650 spectra from a physics-derived target space



Nonnegative PCs were estimated using the methods described in [1]. The resulting first five nonnegative PCs explain 99.98 percent of variability. Table 1 shows the percent of variability explained by those nonnegative PCs.

Table 1. Percent of variability explained by nonnegative PCs estimated from the Target Space data

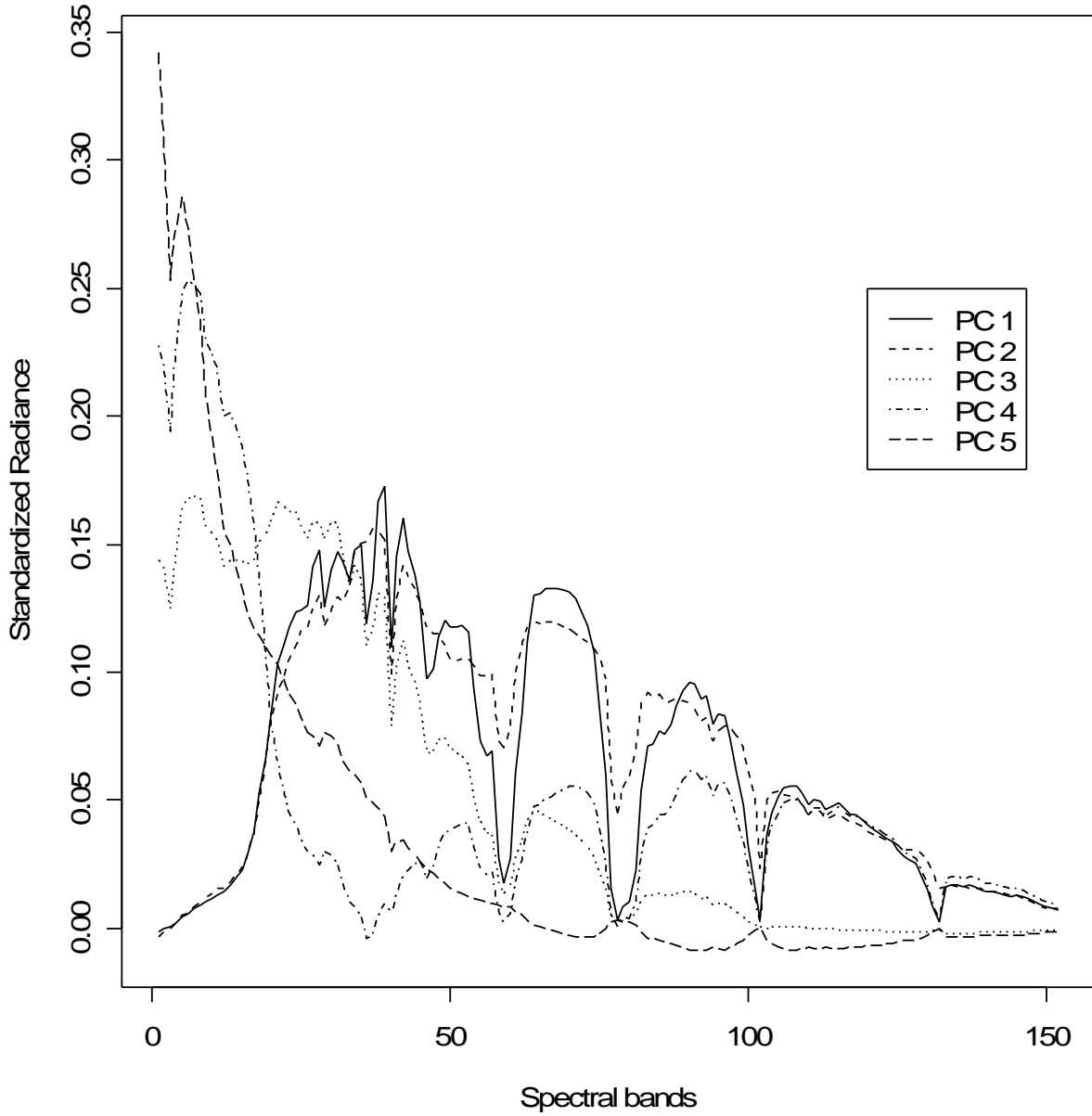
PC number	1	2	3	4	5
Percent of explained variability	51.9	38.1	5.03	4.67	0.28

The coefficients of nonnegative PCs (that is columns of the matrix \mathbf{D}) are shown in Figure 2. They can be interpreted as profiles of the nonnegative PCs, that is, they describe which ranges of spectra are mostly affected by a given nonnegative PC. The original random vector Y can be approximated as

$$Y \approx d_1 W_1 + \dots + d_5 W_5 \quad (4)$$

where d_i are columns of the matrix \mathbf{D} .

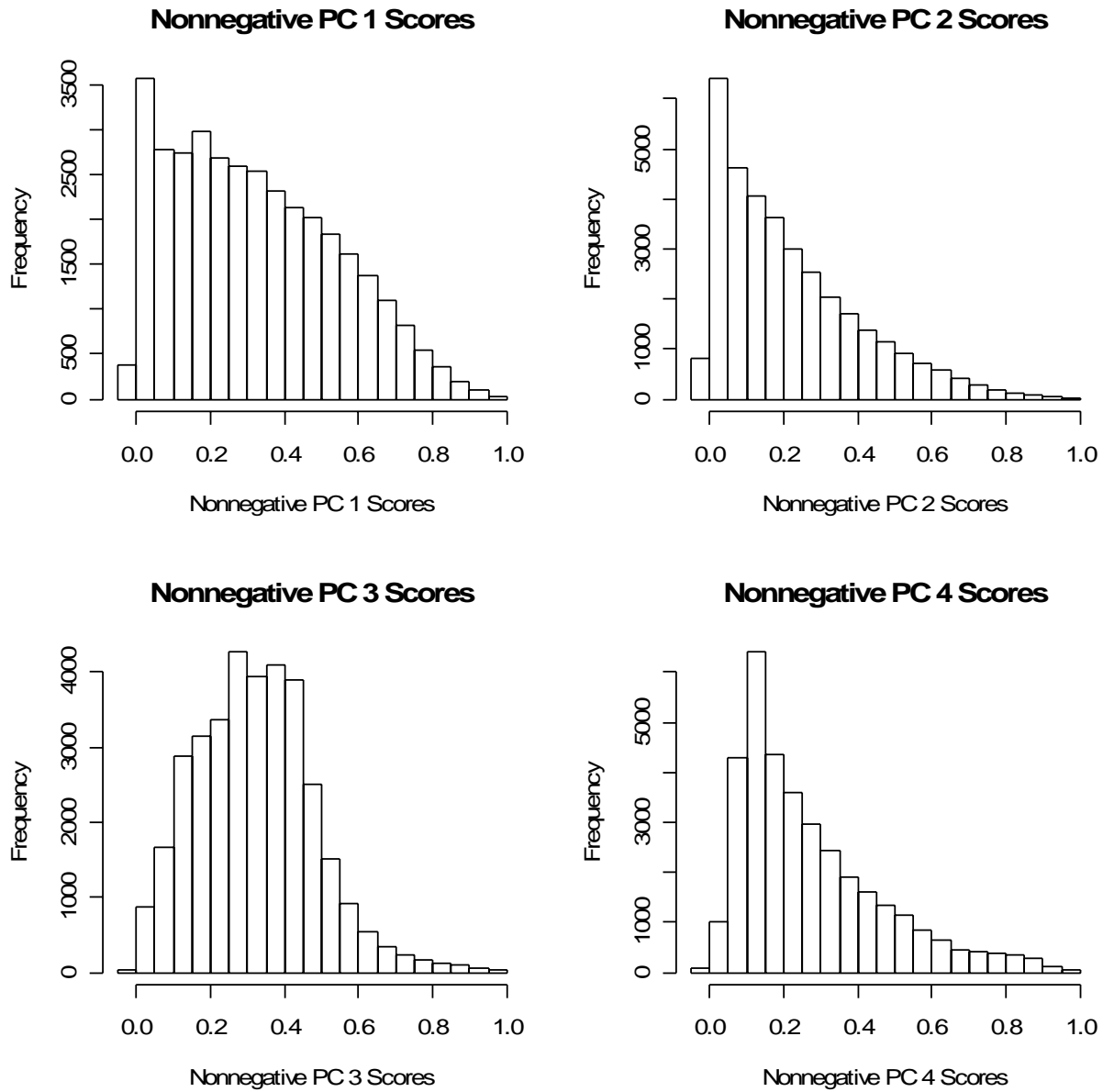
Figure 2. Profiles of the nonnegative PCs of a physics-derived target space



Since the fifth nonnegative PC explains less than 0.3 percent of variability, it was removed from further investigation. The way a nonnegative PC is removed from a set of remaining nonnegative PCs is somewhat different from the method used in classic PCA. In order to preserve nonnegativity, the data is not centered and the resulting nonnegative PCs have positive expected values. Consequently, when removing the fifth nonnegative PC, the approximation (4) becomes

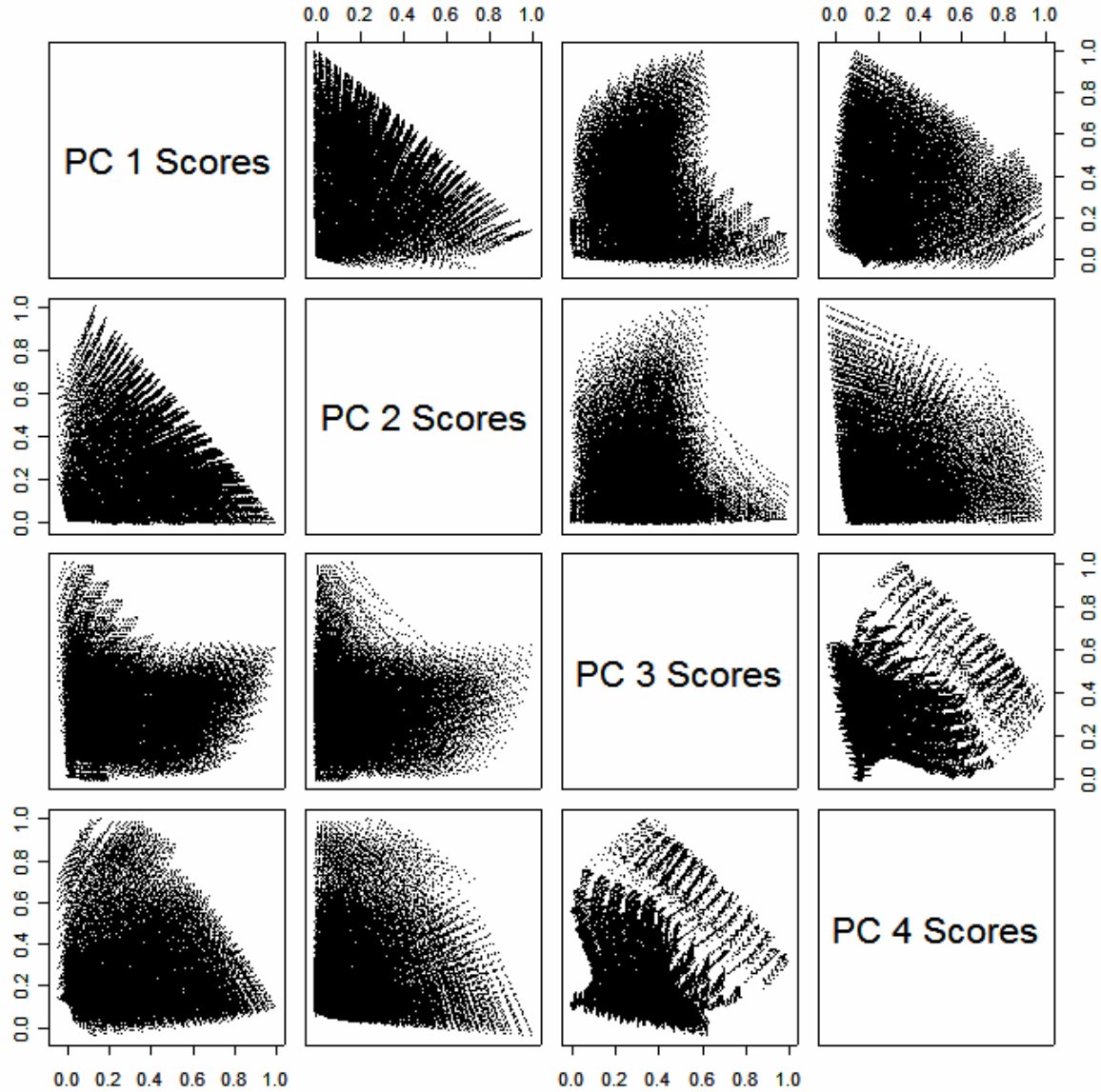
$$Y \approx d_1W_1 + \dots + d_4W_4 + d_5E(W_5) \tag{5}$$

Figure 3. Histograms of the empirical distributions of the scaled nonnegative PC scores



The last (nonrandom) term ($d_5 E(W_5)$) can be interpreted as a baseline spectrum, from which Y can be approximated using the incremental terms described by the first four nonnegative PCs. In order to facilitate interpretation of those incremental terms, we now scale the nonnegative PCs such that their maximum is equal to 1. The resulting empirical distributions of scaled nonnegative PC scores are shown in Figure 3 and their scatter plots are shown in Figure 4.

Figure 4. Scatter plots of the scaled nonnegative PC scores

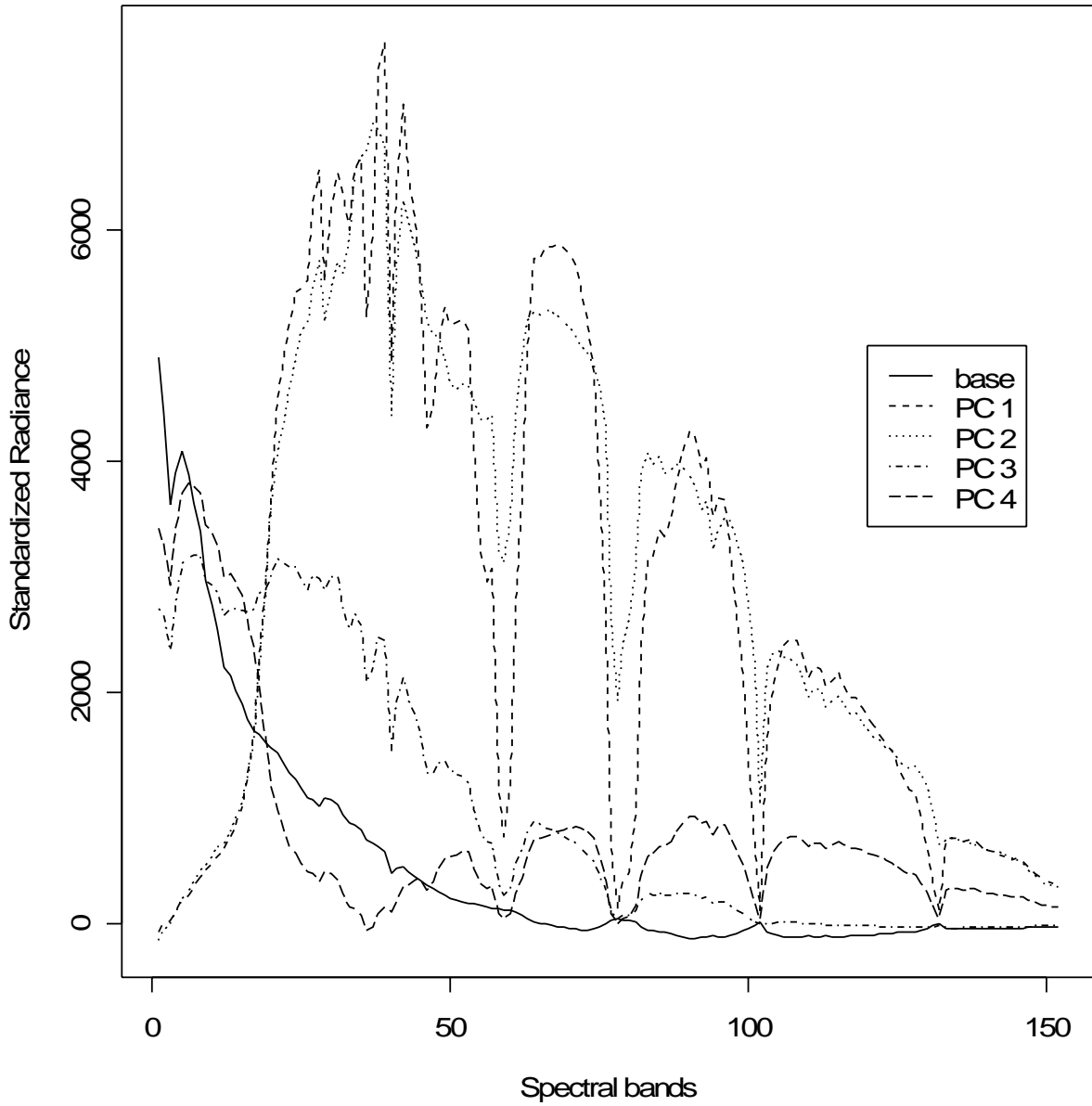


In order to preserve the approximation in (5), the scalars need to modify the profiles d_i , so that, the approximation becomes

$$Y \approx d_1^* W_1^* + \dots + d_4^* W_4^* + d_5^* E(W_5^*) \quad (6)$$

where stars denote the scaled quantities. The resulting profiles d_i^* , $i = 1, 2, 3, 4$ are plotted in Figure 5 together with the baseline spectrum $d_5^* E(W_5^*)$.

Figure 5. The baseline spectrum (base) together with the incremental terms d_i^* associated with the scaled nonnegative PCs



IV. CONCLUSIONS

In this paper, we briefly present a new statistical methodology for nonnegative Principal Component Analysis as described in [1]. We then show an application of that method to a physics-derived target space. The resulting approximation of the target space amounts to a baseline spectrum plus a set of incremental profiles with random coefficients that are uncorrelated. The baseline spectrum together with the incremental profiles and their coefficients are predominantly nonnegative, which should assist in practical implementation of these results. Further work is needed to utilize the representations shown in this paper. For example, one can investigate the relationship between the components defined in this paper and the factors that were used in the generation of the target space.

REFERENCES

1. P. Bajorski, P. Hall, "Methodology for Nonnegative Principal Component Analysis," [submitted for publication].
2. E. Ientilucci, P. Bajorski, "Statistical Models for Physically Derived Target Spaces," [submitted for publication]