FUNCTIONAL FACTOR ANALYSIS FOR PERIODIC REMOTE SENSING DATA[1]

BY CHONG LIU, SURAJIT RAY, GILES HOOKER AND MARK FRIEDL

Boston University, Boston University, Cornell University and Boston University

We present a new approach to factor rotation for functional data. This is achieved by rotating the functional principal components toward a predefined space of periodic functions designed to decompose the total variation into components that are nearly-periodic and nearly-aperiodic with a predefined period. We show that the factor rotation can be obtained by calculation of canonical correlations between appropriate spaces which make the methodology computationally efficient. Moreover, we demonstrate that our proposed rotations provide stable and interpretable results in the presence of highly complex covariance. This work is motivated by the goal of finding interpretable sources of variability in gridded time series of vegetation index measurements obtained from remote sensing, and we demonstrate our methodology through an application of factor rotation of this data.

1. Introduction. The goal of factor rotation is to find interpretable directions explaining the covariance of the variables. In the case of classical multivariate data interpretation of factors it is primarily carried out based on the grouping of factor loadings. However, these approaches are not always applicable to collections of random functions. Instead, we propose an interpretable factor rotation using a naturally predefined space of functions. The motivating data set for this paper consists of roughly weekly observations of vegetation acquired from remote sensing at regular intervals for multiple years. In this case, the dominant seasonal cycle provides a natural choice for dividing the variation into nearly-periodic and nearly-aperiodic sources of variation. More generally, our approach facilitates understanding highly complex forms of functional variation by dividing the total variation into two orthogonal parts, each of which may be explained by a smaller number of components with clear interpretation. Besides achieving the desired interpretability, these components are shown to be stable over the choice of the number of factors and can be obtained through computationally inexpensive steps.

While a large amount of methodological development in functional data analysis has been based on functional principal components analysis [Müller, Stadtmüller and Yao (2006)] and considerable theoretical attention devoted to its properties [Yao, Müller and Wang (2005), Hall, Müller and Wang (2006), Li and Hsing

---

(2010)], little attention has been given to finding rotations of the leading principal components to improve the interpretability of variance components in fPCA. In this context, Ramsay and Silverman (2005) propose a VARIMAX rotation, accomplished by evaluating derived principal components on a fine grid; VARIMAX rotations yield components that have either very high or very low values, effectively focusing variation on particular regions of the functional domain. In many contexts, this can be useful—their study of Canadian weather data neatly picks up the four seasons, for example—but there is considerable further scope for alternative notions of interpretability to be developed. In particular, existing rotation methods designed for multivariate data generally seek to emphasize particular variables or observations, but do not attempt to account for the ordering relations between variables that exist in functional data. We generally expect the loading at one time point to be close to the loading at a nearby time point. One way to achieve this is through smoothing penalties. Instead, we define a rotation toward an interpretable reference subspace of functions.

In the context of multi-year time series remote sensing data, the need for methods to extract interpretable sources of variation is particularly acute. The vegetation index considered in this paper consists of a 6-year time series of remote sensing images acquired at 8-day intervals for a site in central Massachusetts (see Section 2.1 and Figure 1). These data demonstrate a highly complex functional covariance structure. To illustrate, in Figure 2 we present a scree plot of eigenvalues for the data set. This scree plot shows exponential decay in explained variance, with no
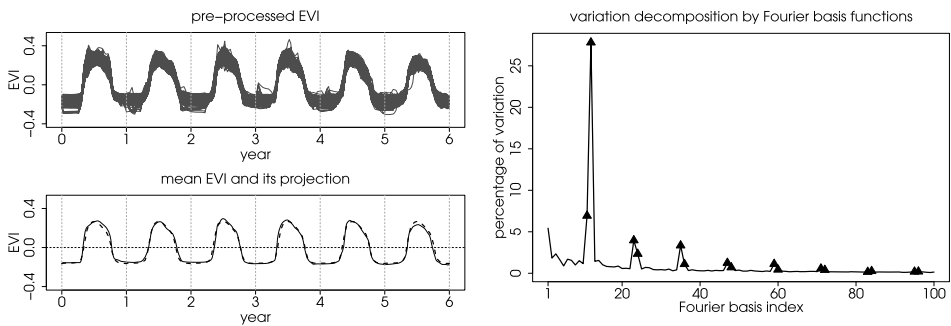

FIG. 1. *Upper left*: *Preprocessed EVI data is obtained by first smoothing raw EVI observations with saturated Fourier basis expansion and the penalty on the second derivative and then the raw EVI fit is time-series demeaned.* *Lower left*: *The solid line is the mean of preprocessed EVI curves. The dashed line is the projection of the mean onto the subspace spanned by all Fourier basis functions with annual period in the saturated basis system.* *Right*: *Percentage of variation explained by Fourier basis functions. Preprocessed EVI curves are projected onto each Fourier basis function. The variance of the projection scores and its percentage of the total variance are computed. The Fourier basis index starts from* $\sin(\omega t)$. *The function* $\sin(K\omega t)$ *has index* $2K - 1$ *and* $\cos(K\omega t)$ *has index* $2K$. *The solid triangles highlight the percentage score-variance associated with the annual Fourier basis which correspond to index* $11, 12, 23, 24, 35, 36, \ldots, 95, 96$. *The constant basis is not included in the calculation and index.*
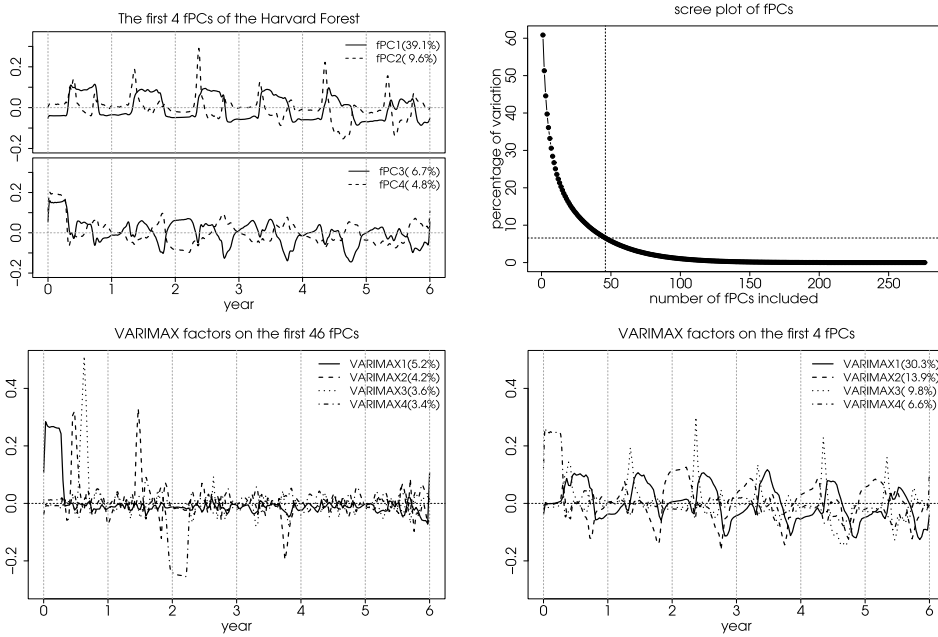
FIG. 2. *Upper left*: *the first* 4 *fPCs of the Harvard Forest data*; *Upper right*: *the scree plot of the fPC. The vertical dashed line stands at* 46 *and the horizontal dashed line shows the amount of variation not explained by the first* 46 *fPCs*; *Lower left*: *VARIMAX components derived by rotating the first* 46 *fPCs*; *Lower right*: *VARIMAX components derived by rotating the first* 4 *fPCs. Numbers in parentheses of the legend are percentage of variation explained by each component.*

evidence of the "elbow" that is frequently used to decide the number of eigenvalues to retain. Further, if we wished to explain 90% of variation—a frequently used criterion—over 30 components would need to be retained, and examination of the first few principal components suggests that interpretation of these components is problematic (see Figure 2), consisting of both strong periodic structure as well as trends and isolated features. Interpreting these sources of variation from this covariance structure is challenging and common techniques such as VARIMAX rotations (also shown in Figure 2) are clearly unhelpful in this case. There is, however, one clear and highly interpretable feature in the data: a strong periodic signal. This is naturally expected due to the strong seasonal forcing.

Basing an interpretation around seasonality is both visually satisfying and scientifically useful. Perhaps the most widely recognized feature of the global climate (e.g., temperature, precipitation) and ecosystem (e.g., vegetation) data is seasonality [Hartmann (1994)]. This can be illustrated by spectral decomposition of our data, shown in the right plot in Figure 1, where annual variation dominates. Meanwhile, because climate dynamics are produced by complex interactions among the Earth's oceans, atmosphere, cryosphere, and land masses, the Earth's weather and climate system, and hence indicators of ecosystem, does not behave in a strictly

periodic fashion [Holton (1992)]. Although sophisticated models have been developed for predicting climate-ecosystem dynamics, our understanding remains incomplete.

The contribution of this paper is to provide a new factor rotation technique that divides sources of variation into nearly-periodic and nearly-aperiodic components. While strictly periodic components could be obtained directly by projecting onto a basis of periodic functions, the year-to-year variation in season timing requires us to retain somewhat more flexibility so as not to overestimate the amount of nonseasonal variation. One approach to this would be to undertake a registration procedure [Gervini and Gasser (2004), Liu and Müller (2004), Ramsay and Silverman (2005), Kneip and Ramsay (2008)]. However, the registration is ill-posed and registration algorithms are computationally expensive, particularly for large and complex data sets. Instead, we keep within the framework of factor rotation and seek a rotation that rotates the largest sources of variation toward being periodic or a-periodic (see Figure 3). This is accomplished via a canonical correlations transform providing what we have labeled *principal periodic components* (PPCs).

In comparing VARIMAX and PPC, we perform both rotations on a sequence number of fPCs and compute the change in L^2 sense between the first rotated
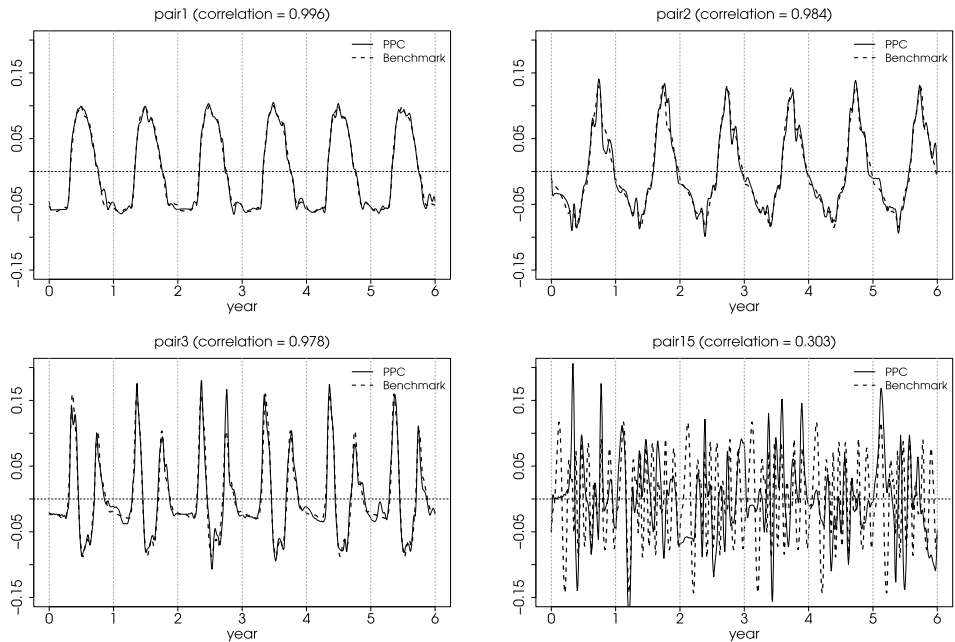

FIG. 3.   *PPC results on Harvard Forest data. PPCs are computed with 46 fPCs of preprocessed EVI curves. The solid curves are PPCs ξ_j and the dashed curves are benchmarks θ_j associated with ξ_j. The correlation is computed as the standardized inner product between θ_j and ξ_j. The pair index is ordered by the correlation.*

components derived from two consecutive numbers of fPCs. The L^2 change of PPC rotation is much smaller and more stable compared to VARIMAX rotation, suggesting PPC's robustness with respect to the number of fPCs used in rotation.

Simulation studies also show that PPCs perform very well in detecting periodic variation in the following two cases: (i) amount of periodic variation increases from 0 to a level only comparable to other source of variation where fPCs react slowly to the increasing periodic variation; (ii) total variation is dominated by increasing amount of high frequency disturbances where fPCs are quickly contaminated by disturbances and PPCs still capture the periodic source of variation.

To better understand the rotation and the relation between PPCs and the space of functions with strict annual cycle, we develop a heuristic test of whether the first PPC lies in that space. In the test, we create a set of curves under the null hypothesis as close to the original data as possible by either replacing the first PPC by its associated benchmark, or inflating the nearly-periodic component while controlling for Kullback–Leibler divergence of the sample functional covariance to the null functional covariance. The test on our motivating data set rejects the null hypothesis, suggesting that no strict annual variation is presented in the space spanned by PPCs.

A further aspect of the data is that it is gridded in a regular spatial distribution. This induces both spatial correlation as well as effects due to (unobserved) geographic and environmental factors. Our use of rotations will allow the effect of these structures to be empirically investigated in terms of both variation in cyclic ecological factors and in longer-term trends. Our functional data analysis approach differs from techniques using empirical orthogonal functions (EOFs) in spatio-temporal analysis [e.g., in Everson et al. (1996)] in considering observations as functions of time rather than of space. We believe that this approach is appropriate to the task of separating cyclical from other trends. We note that a similar rotation of EOFs toward a subspace of functions describing landscape features or other geographic gradients could be developed along similar lines to PPCs, but this is beyond the scope of the current paper.

The remainder of this paper is arranged as follows. In Section 2 we will show a motivating example in which we carry out smoothing and functional principal component analysis and demonstrate the motivation for PPCs. In Section 3 we introduce the framework of PPC and its results on our remote sensing data. Results of a simulation study are presented in Section 4 that illustrate the sensitivity and robustness of PPC in identifying periodic variation. Details of further simulation experiments concerning the power and size of the proposed tests are given in the supplemental article [Liu et al. (2012)]. We end with some concluding remarks and discussion of future research.

2. A motivating example.

2.1. *The data set*.   The data set used for this work consists of time series of remotely sensed images acquired over a site in central Massachusetts. Specifically,

we used surface spectral reflectance measurements from the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard NASA's Terra and Aqua satellites. Data from MODIS were extracted for a 25 by 25 pixel window (covering an area of $\approx 134 \text{ km}^2$) centered over the Harvard Forest Long Term Experimental Research site in Petersham, MA. This site is characteristic of mid-latitude temperate forests and is dominated by deciduous tree and understory species that exhibit strong seasonal variation in phenology. Data are provided at 8-day intervals (46 data points per year) for the period from January 1, 2001 to December 31, 2006. The spatial resolution of the data is 500-m on the ground.

Using MODIS surface spectral reflectances in the blue, red and near-infrared (NIR) wavelengths, we computed a quantity known as the "enhanced vegetation index" [EVI; Huete et al. (2002)]:

$$2.5 \times \frac{NIR - RED}{NIR + C_1 \times Red - C_2 \times Blue + L},$$

where NIR, Red and Blue are reflectances of the corresponding bands recorded by MODIS and C_1, C_2 and L are constant coefficients. The EVI exploits spectral reflectance properties of live vegetation, yielding an index that scales from -1 to 1 that is widely used for monitoring seasonal dynamics in vegetation. Because EVI data are sensitive to the presence of snow and include noise and missing values caused by clouds, the data were preprocessed prior to analysis to remove noise and fill gaps following the procedure described by Zhang, Friedl and Schaaf (2006). In the supplemental article [Liu et al. (2012)], we provide a detailed account of this preprocessing and criteria for excluding pixels with large blocks of missing observations.

The final data set consisted 276 EVI time series values for each of 423 pixels (excluding pixels with problematic observations), that is, 423 replicated curves, with each replication corresponding to a pixel in the area of interest. The regular spatial and temporal sampling of EVI data makes functional data analysis a useful framework for exploring variation among curves and facilitating the study of change in variation.

2.2. *Smoothing of EVI.* Denote the discrete observation at pixel i and time t_{ij} by Y_{ij}. We consider the following additive error model:

$$Y_{ij} = x_i(t_{ij}) + e_{ij}, \qquad 1 \le i \le N \text{ and } 1 \le j \le n_i,$$

where $x_i(t)$'s are the true realizations of the underlying random growing process $X(t)$ and e_{ij}'s are errors. To estimate $x_i(t)$, we choose a regularization approach based on basis expansion. Specifically, we fit our data with the saturated Fourier basis and explicitly penalize the total curvature. The Fourier basis is numerically convenient for our purposes; experiments with alternative B-spline bases indicated that our results are insensitive to this choice. The smoothing parameter is chosen to minimize the sum of generalized cross validation scores over all curves. This can

be implemented in R [R Development Core Team (2010)] using the FDA package [Ramsay et al. (2010)]. Let $\hat{x}_i(t)$ denote the fitted curve. Then, we further process this raw fitting by removing from each $\hat{x}_i(t)$ its time-series average. Then, we obtain the demeaned curve $z_i(t)$ as

$$z_i(t) = \hat{x}_i(t) - \frac{1}{T} \int_0^T \hat{x}_i(t)\, dt, \qquad 1 \le i \le N.$$

This demeaning process removes vertical variation and avoids defining it as either annual or nonannual. The centering removes heterogeneity in the overall growing level and allows us to focus on nonconstant modes of variation; see further discussion in Section 3. The pre-smoothed and demeaned EVI curves of the Harvard Forest data are shown in the upper left plot in Figure 1. While many methods have been developed to analyze the features and structure of the mean shape, in this paper we are interested in changes in vegetation dynamics manifested in terms of variance.

We decompose the total variation among the EVI curves by projecting EVI curves to the saturated Fourier basis system. This decomposition shows that variation explained by the annual Fourier basis function is a dominating source of variation in our example, as shown in the right plot in Figure 1. The data here are defined on a grid of observations taken every 8 days and thus could be considered a very high-dimensional multivariate data set. We have chosen to view these data as functional due to the underlying smooth greening process that they record, and because it facilitates the definition of periodicity which we employ to define a factor rotation below.

2.3. *Functional principal component analysis.* Functional principal component analysis (fPCA) is a well studied research area. It provides a way to extract the major mode of variation among curves and our proposed PPC is based on and motivated by fPCA. To introduce and fix notation for description of PPC in later sections, we give a brief review on fPCA. More references on fPCA can be found in Ramsay and Silverman (2002), Yao, Müller and Wang (2005) and Müller, Stadtmüller and Yao (2006). In particular, we look for a set of normalized and orthogonal functions $\gamma_j(t)$ such that the projection of all EVI curves onto each specific $\gamma_j(t)$ has the largest variability. These $\gamma_j(t)$'s are called the functional principal components (fPCs). Formally, suppose we have N smoothed and time-series demeaned EVI curves $z_i(t)$, $1 \le i \le N$. The sample cross-section mean process is $\hat{\mu}(t) = N^{-1} \sum_i z_i(t)$. Then the cross-section demeaned curve is obtained as $\tilde{z}_i(t) = z_i(t) - \hat{\mu}(t)$. $\gamma_j(t)$ is chosen to maximize $N^{-1} \sum_i (\int \gamma_j(t) \tilde{z}_i(t)\, dt)^2$ subject to the constraints that $\int \gamma_j(t) \gamma_k(t)\, dt = \delta_{jk}$ where δ_{jk} is the Kronecker delta. Given the estimated covariance kernel $\Omega(s, t) = N^{-1} \sum_{i=1}^N \tilde{z}_i(s) \tilde{z}_i(t)$, each fPC, $\gamma_j(t)$, satisfies the eigen-equation $\int \Omega(s, t) \gamma_j(t)\, dt = \lambda_j \gamma_j(s)$, where λ_j is the associated eigenvalue. By writing $\gamma_j(t)$ in expansion of basis functions, this problem can be reduced to the computation of matrix eigenvalues and eigenvectors. Here

we have pre-smoothed the data and applied a principle-components decomposition without additional penalty. fPCA can also be employed along with smoothing methods [Silverman (1996)] or by directly smoothing the covariance surface [Yao, Müller and Wang (2005)]. The methods developed below are applicable for an fPCA decomposition, irrespective of the method employed to derive it.

In order to explore the variation in EVI curves, we apply the standard fPCA techniques on Harvard Forest data. The first 4 fPCs of Harvard Forest are plotted in Figure 2 where each of the four fPCs contains some level of annual periodicity and pick up features of EVI variation at different times of year. For example, the first PC shows that the contrast of EVI between summer and winter is the most distinct feature that characterizes the vegetation growing in this area, however, with a decreasing trend suggesting the contrast between summer and winter has changed over the 6 years. The second fPC has a sharp peak roughly at the start of each growing season combined with noticeable dips during years 5 and 6. Due to the existence of the two negative bumps, it is hard to interpret the second fPC as the effect of growing season onset. A third fPC emphasizes the ending of growing season, characterizing variation in the timing of vegetation browning. However, as these fPCs are combined with nonannual signal, they are not designed to distinguish between annual and nonannual sources of variation. In Section 3 we discuss the appropriate rotation of fPCs to aid interpretation by separating annual and nonannual sources of variation. But first we discuss one widely used technique of rotation for functional data—the VARIMAX rotation.

2.4. *VARIMAX rotation.* VARIMAX is a widely used orthonormal transformation in multivariate analysis which can make multivariate principal components more interpretable. The functional VARIMAX rotation borrows readily the concept of multivariate VARIMAX rotation. Suppose we retain the first M fPCs and the subspace spanned by these M fPCs is denoted by $\mathbf{\Gamma}_M$. We use $\boldsymbol{\gamma}$ to refer to the vector valued function $(\gamma_1, \ldots, \gamma_M)'$. Let $\mathbf{B}$ be a $M \times n$ evaluation matrix of $\boldsymbol{\gamma}$ where $\mathbf{B}_{ij} = \gamma_i(t_j)$, $1 \leq j \leq n$. Given an orthonormal matrix $\mathbf{T}$, $\boldsymbol{v} = \mathbf{T}\boldsymbol{\gamma}$ gives us a new set of orthonormal functions. The evaluation matrix at the same t_j's of the rotated functions $\boldsymbol{v}$ is given by $\mathbf{A} = \mathbf{TB}$. Denote the ijth entry of matrix A by a_{ij}. Then the VARIMAX strategy for choosing the orthonormal rotation matrix $\mathbf{T}$ is to maximize the variation of a_{ij}^2 over all values of i and j.

The solution to the above maximization problem will encourage values a_{ij} to be either strongly positive, near zero, or strongly negative. This rotation tends to cluster information and make the components of variation easier to interpret. VARIMAX rotation on the first 46 fPCs and on the first 4 fPCs are shown in the two lower plots in Figure 2. If using only 4 fPCs, we do not have sufficient flexibility to provide improved interpretation. By contrast, using 46 fPCs provides so much concentration on individual time points that any natural interpretation is lost.

The rotation described here can be generalized to describe a rotation of principal components to find directions that lie close to an interpretable reference subspace.

In a multivariate context, this amounts to finding a rotation of factors Γ so that the leading components lie close to a subspace spanned by the columns of a matrix F_P, assumed to have interpretable relevance for the application at hand, and the mathematical development below can be read in an entirely multivariate context. It more generally applies to observations taking values on any Hilbert space. While the space of periodic functions is clearly relevant for our application, the choice of subspace is context-specific.

3. Principal periodic component (PPC). The VARIMAX rotation does not achieve our goal of separating annual and nonannual variation since its objective function is not designed to do so. We need to explicitly define an objective function which can extract annual variation. One natural way to do this is to order the rotated fPCs by their levels of annual periodicity. To measure annual periodicity, we will first define benchmarks which have strict annual periodicity. Then we compute the closeness between rotated fPCs and corresponding benchmarks and this computed closeness serves as the measure of annual periodicity of the rotated fPCs. Refer to Fourier basis functions with annual period as f_k, $1 \leq k \leq P$, the vector of them as **f** and the space spanned by them as $\mathbf{F}_P$. Hence, $\mathbf{F}_P$ is a space of functions with annual periodicity up to a certain frequency determined by P. P is limited to the set of periodic Fourier coefficients used to smooth the data. More generally, P can be set to N—allowing the interpolation of any N points that lie in a strictly periodic subspace. We construct our benchmarks as the linear combination of f_k's. Then benchmarks are in $\mathbf{F}_P$ and thus have exactly annual periodicity. Intuitively, we can consider $\boldsymbol{\gamma}$ and **f** as two frames of their own spaces $\boldsymbol{\Gamma}_M$ and $\mathbf{F}_P$. We can rotate the two frames and align them in the same direction as much as possible. If $\boldsymbol{\Gamma}_M$ contains direction which is exactly annual, then we will align the two spaces at least in one direction. The closeness between the rotated fPC and associated benchmark is computed as their standardized inner-product.

3.1. *Principal periodic component framework.* In this section we give a mathematical description of the PPC methodology. Recall that $\boldsymbol{\gamma}$ is a M dimensional vector of fPCs obtained from time-series demeaned curves and **f** is a P dimensional vector of Fourier basis functions with annual period. Define $\boldsymbol{\Sigma}_{\gamma f} = \langle \boldsymbol{\gamma}, \mathbf{f} \rangle$, where $\langle \cdot, \cdot \rangle$ is the inner-product in L^2 space and the ikth entry of $\boldsymbol{\Sigma}_{\gamma f}$ is given by $\langle \gamma_i, f_k \rangle$. We compute the singular value decomposition $\boldsymbol{\Sigma}_{\gamma f} = \hat{\mathbf{U}}' \mathbf{W} \hat{\mathbf{V}}$ and denote the jth row of $\hat{\mathbf{U}}$ by $\hat{\mathbf{u}}'_j$ and the jth row of $\hat{\mathbf{V}}$ by $\hat{\mathbf{v}}'_j$. The PPCs and associated benchmarks are then defined as follows,

$$(1) \qquad \xi_j = \hat{\mathbf{u}}'_j \boldsymbol{\gamma} \quad \text{and} \quad \theta_j = \hat{\mathbf{v}}'_j \mathbf{f}, \qquad j = 1, 2, \ldots, \min(M, P).$$

In the above definition, we call ξ_j the jth PPC and θ_j the associated benchmark of ξ_j.

In order to derive these estimates, denote any rotation on $\boldsymbol{\gamma}$ by **U** with $\mathbf{u}'_j$ being the jth row of **U**, and any rotation on **f** by **V** with $\mathbf{v}'_j$ being the jth row of **V**. Let

$\xi_j^0 = \mathbf{u}_j' \boldsymbol{\gamma}$ and $\theta_j^0 = \mathbf{v}_j' \mathbf{f}$. Then ξ_j^0 is the jth rotated fPC and θ_j^0 is a function with annual cycle. We define the closeness measure of the pair ξ_j^0 and θ_j^0 as the angle between them,

$$(2) \qquad \rho_j = \rho(\xi_j^0, \theta_j^0) = \frac{\langle \xi_j^0, \theta_j^0 \rangle}{\|\xi_j^0\| \|\theta_j^0\|} = \frac{\langle \mathbf{u}_j' \boldsymbol{\gamma}, \mathbf{v}_j' \mathbf{f} \rangle}{\|\mathbf{u}_j' \boldsymbol{\gamma}\| \|\mathbf{v}_j' \mathbf{f}\|}.$$

Given this closeness measure, we solve the following optimization problem for $j = 1, 2, \ldots, \min(M, P)$,

$$(3) \qquad (\hat{\mathbf{u}}_j, \hat{\mathbf{v}}_j) = \underset{\mathbf{u}_j, \mathbf{v}_j}{\arg\max}\, \rho(\xi_j^0, \theta_j^0) = \underset{\mathbf{u}_j, \mathbf{v}_j}{\arg\max}\, \frac{\mathbf{u}_j' \boldsymbol{\Sigma}_{\gamma f} \mathbf{v}_j}{\mathbf{u}_j' \boldsymbol{\Sigma}_{\gamma\gamma} \mathbf{u}_j \cdot \mathbf{v}_j' \boldsymbol{\Sigma}_{ff} \mathbf{v}_j},$$

subject to $\langle \xi_j^0, \xi_k^0 \rangle = \delta_{jk}$, $\langle \theta_j^0, \theta_k^0 \rangle = \delta_{jk}$, $\langle \xi_j^0, \theta_k^0 \rangle = 0$, where the ikth entry of $\boldsymbol{\Sigma}_{\gamma f}$, $\boldsymbol{\Sigma}_{\gamma\gamma}$ and $\boldsymbol{\Sigma}_{ff}$ are given by $\langle \gamma_i, f_k \rangle$, $\langle \gamma_i, \gamma_k \rangle$ and $\langle f_i, f_k \rangle$, respectively.

We observe that the objective in (3) has the same form as multivariate canonical correlation analysis (CCA) where random variables are replaced by functions. However, the sampling properties of the PPC rotation differ from CCA in that the frame of our reference subspace, $\mathbf{f}$, is deterministic while fPCs $\boldsymbol{\gamma}$ is random where randomness comes from sampling variation. See Mardia, Kent and Bibby (1979) for an overview of CCA; in the functional analysis context see Leurgans, Moyeed and Silverman (1993) and He, Müller and Wang (2003). According to the CCA results, we have the following solution:

(4) $\hat{\mathbf{u}}_j$ is proportional to the jth eigenvector of $\boldsymbol{\Sigma}_{\gamma\gamma}^{-1} \boldsymbol{\Sigma}_{\gamma f} \boldsymbol{\Sigma}_{ff}^{-1} \boldsymbol{\Sigma}_{\gamma f}'$ and $\hat{\mathbf{u}}_j' \boldsymbol{\Sigma}_{\gamma\gamma} \hat{\mathbf{u}}_j = 1$,

(5) $\hat{\mathbf{v}}_j$ is proportional to the jth eigenvector of $\boldsymbol{\Sigma}_{ff}^{-1} \boldsymbol{\Sigma}_{\gamma f}' \boldsymbol{\Sigma}_{\gamma\gamma}^{-1} \boldsymbol{\Sigma}_{\gamma f}$ and $\hat{\mathbf{v}}_j' \boldsymbol{\Sigma}_{ff} \hat{\mathbf{v}}_j = 1$.

Due to the orthogonality of fPCs and the Fourier basis system, we have $\boldsymbol{\Sigma}_{\gamma\gamma} = \mathbf{I}$ and $\boldsymbol{\Sigma}_{ff} = \mathbf{I}$. These two identities reduce (4) and (5) to the eigenanalysis of $\boldsymbol{\Sigma}_{\gamma f}$ and the results in (1) follows. In (1), $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ are two orthogonal rotation matrices on $\boldsymbol{\gamma}$ and $\mathbf{f}$, respectively. In a more general context, (4) and (5) can be employed if, for example, the space $\mathbf{F}_P$ is not parameterized by an orthogonal basis.

In this context, the motivation for removing the time series mean of the observations as described in Section 2 becomes apparent. We have not defined variation in terms of a constant vertical shift as being either periodic or aperiodic in nature. Demeaning the observations ensures that there is zero variation in this direction and, hence, all the computed fPCs will also integrate to zero. Had this step not been carried out, the constant shift would have been conflated with both periodic and aperiodic sources. If this constant source of variation were defined as periodic, a constant function could be added to the space $\mathbf{F}_P$.

3.2. *PPC results on Harvard forest data.* We now apply our PPC methodology on the Harvard Forest data. In the Harvard Forest data, periodicity is set to be annual and thus we have 46 Fourier basis functions with annual period. Thus, $P = 46$ and the space spanned by these functions is $\mathbf{F}_{46}$. We set $M = P = 46$ in our calculation in order to get pairwise match between the PPCs and benchmarks. $\mathbf{\Gamma}_{46}$ accounts for 93.4% of total variation. The robustness of PPC computation with respect to the choice of M is further discussed in Section 3.3. A selection of four pairs of PPCs and associated benchmarks with decreasing correlations are shown in Figure 3.

The first PPC suggests the most important annual variation is the contrast between summer and winter. The second PPC has the effect of shifting summer forward or backward in time, while the third PPC corresponds to combined effect of growing season length and summer maximum EVI. These leading PPCs demonstrate modes of variation which are most likely to repeat every year. From an ecological perspective, these sources of variance are of critical importance because they reflect signatures of climate variability in ecosystem processes. Thus, PPCs provide a tool for characterizing and understanding how subtle changes in climate, such as shifts in the timing of seasons, are affecting ecosystems [Parmesan and Yohe (2003), Piao et al. (2008)].

Note that benchmarks are always exactly annual and the correlation between PPCs and their benchmarks decreases as we extract more PPCs. We thus construct a set of orthonormal functions which are ordered by their level of annual periodicity. This shows that the amount of annual variation contained in PPCs decreases as the index increases. A trade-off in defining which components should be denoted "periodic" is detailed in Section 3.4

3.3. *Stability of PPC directions.* Choosing the number of fPC components, M is a statistically challenging task. This number depends on several factors, including strength of signals, the choice of smoothing parameter, and sampling error as well as the choice of fPCA methodology. An ideal factor rotation should be insensitive to the number of factors chosen. This is particularly important when there are many small components of variation and the number of components selected can be unstable. In the VARIMAX rotation the interpretation of rotated components is very sensitive to M. On the other hand, the PPCs provide a natural framework to achieve this goal when the principal sources of variation are periodic in nature. This is achieved due to the use of a well-defined reference subspace, thereby stabilizing the choice of "interesting" directions.

We explored the stability of the leading rotated component for a range of choices for M—the number of fPCs we rotate—from 5 to 50 in increments of 5. In these data, the first VARIMAX component was highly unstable, while the first PPC remained stable and retained most of its interpretation for the whole range of M (see Figure 4). Here we define the first important VARIMAX component in any of three ways: (i) the component which accounts for the most variation, (ii) the
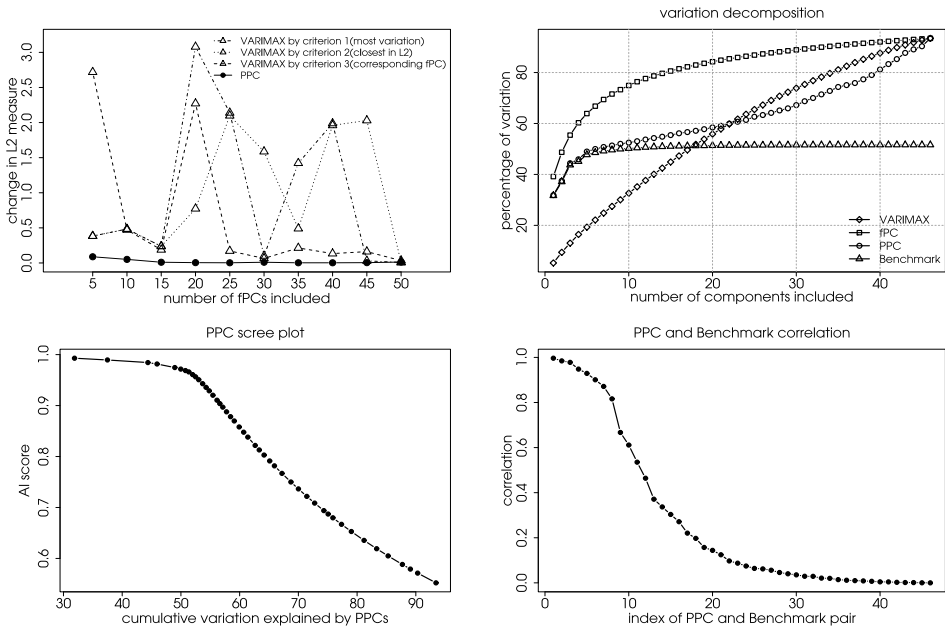
FIG. 4. *Upper left*: L^2 *difference between rotated components derived on consecutive values of* M. *The horizontal axis represents the value of* M, *the number of fPCs used in rotations. For a given* M, *the corresponding value on the vertical axis measures the* L^2 *difference between components obtained by rotating* M *and* $M - 5$ *fPCs. Three different dashed lines correspond to three definitions of the first VARIMAX component. The solid line corresponds to the PPC. Upper right: Percentage of variation explained. Diamonds are the cumulative variation explained by VARIMAX components. Squares are the cumulative variation explained by fPCs. Circles are the cumulative variation explained by PPCs. Triangles are the cumulative variation explained by benchmarks. Lower left: PPC scree plot, computed as the amount of cumulative variation explained by benchmarks as a proportion of cumulative variation explained by PPCs; Lower right: Correlation between PPCs and benchmarks.*

component of M fPCs that is closest to the first VARIMAX direction derived with $M - 5$ fPCs in the L^2 sense, and (iii) the component closest to the first fPC. To summarize the stability of these rotations, we explored the L^2 difference between components rotated using M and $M + 5$ fPCs under each of the three VARIMAX definitions above and using the first PPC component. The L^2 differences on PPC rotation are highly stable, whereas the measure for all of the VARIMAX rotations shows large change in both directions.

3.4. *Variation decomposition.* We demonstrate the variation decomposition using two sets of rotations, one being the standard VARIMAX rotation and the other being the PPC rotation described in this paper. For comparing the two techniques we define component scores as EVI curves projected on the set of orthogonal functions in which we are interested. Denote the VARIMAX components based

on 46 fPCs by ν_j. Then we have

$$s_{ij}^{\gamma} = \int_T \tilde{z}_i(t)\gamma_j(t)\,dt \quad \text{and} \quad \lambda_j^{\gamma} = \frac{1}{N-1}\sum_{i=1}^{N}(s_{ij}^{\gamma})^2,$$

$$s_{ij}^{\xi} = \int_T \tilde{z}_i(t)\xi_j(t)\,dt \quad \text{and} \quad \lambda_j^{\xi} = \frac{1}{N-1}\sum_{i=1}^{N}(s_{ij}^{\xi})^2,$$

$$s_{ij}^{\theta} = \int_T \tilde{z}_i(t)\theta_j(t)\,dt \quad \text{and} \quad \lambda_j^{\theta} = \frac{1}{N-1}\sum_{i=1}^{N}(s_{ij}^{\theta})^2,$$

$$s_{ij}^{\nu} = \int_T \tilde{z}_i(t)\nu_j(t)\,dt \quad \text{and} \quad \lambda_j^{\nu} = \frac{1}{N-1}\sum_{i=1}^{N}(s_{ij}^{\nu})^2.$$

The cumulative sum of $\lambda_j^{\gamma}, \lambda_j^{\xi}, \lambda_j^{\theta}$ and λ_j^{ν} are plotted in the upper right plot in Figure 4. The VARIMAX decomposition tends to produce equal decomposition, indicated by the low curvature of its cumulative sum. The fPCs decompose the total variation by their decreasing abilities to explain variation, producing the concave feature seen in its cumulative sum. Variation explained by the benchmarks goes flat, suggesting the annual variation represented by benchmarks with low correlation tends to be orthogonal to $\mathbf{\Gamma}_{46}$. The increasing gap from the left to the right between PPC decomposition and benchmark decomposition reflects the decreasing ability to line up the rotated frames of $\mathbf{\Gamma}_{46}$ and $\mathbf{F}_{46}$.

3.5. *Nearly-annual and nonannual trade-off.* In this subsection we develop an ad hoc methodology of choosing PPCs as nearly-annual, in order to separate annual variation from nonannual variation. Since the level of annual periodicity decreases, it suffices to find a cut-off position and include all PPCs before the cutoff as nearly-annual and all PPCs after the cutoff as nonannual. To this end, we measure the cumulative amount of variation explained by benchmarks as a proportion of cumulative variation explained by PPCs and call it annual information (AI). Specifically, we define

$$AI_j = \frac{\sum_{k=1}^{j}\lambda_k^{\theta}}{\sum_{k=1}^{j}\lambda_k^{\xi}}.$$

AI scores show an elbow around 8 PPCs (see the lower left plot in Figure 4). This elbow suggests a possible position to cutoff. This position is further supported by the plot of correlation between PPCs and benchmarks where a sudden drop is observed around 8 PPCs. In the supplemental article [Liu et al. (2012)] we detail a simulation study investigating the efficacy of AI as a visual diagnostic where we demonstrate that the appropriate number of PPC's is selected with high probability.

3.6. *Application of PPC.* PPCs are modes of variation which are ordered by their level of annual periodicity. Since PPCs are generated by orthogonally rotating the fPCs, PPCs form another empirical orthogonal basis which can be used to decompose EVI curves. Moreover, if we project EVI curves onto PPCs and fPCs, the approximation by PPCs is as good as the approximation by fPCs. However, we can further decompose EVI curves into nearly-annual and nonannual components. Suppose $P > M$ and thus we have M PPCs. If we have K fPCs in total, then we have the following decomposition:

$$z_i(t) = \hat{\mu}(t) + \sum_{j=1}^{J} s_{ij}^{\xi} \xi_j(t) + \sum_{j=J+1}^{M} s_{ij}^{\xi} \xi_j(t) + \sum_{j=M+1}^{K} s_{ij}^{\gamma} \gamma_j(t),$$

(6)
$$1 \le i \le N.$$

The first term on the right-hand side of (6) is the sample mean function. The second and the third terms are the nearly-annual component and nonannual component we determined in the last subsection. Note J in (6) is the number of PPCs we determined as nearly-annual. For the Harvard Forest data, J is taken as 8 based on the AI elbow and the correlation criterion described in Section 3.5. The last term in (6) is the contribution of fPCs associated with very small eigenvalues, which are removed when we truncate to a certain percentage of variation. These are retained in conducting the simulation studies below. The decomposition result is shown in Figure 5. This decomposition helps us reconstruct original EVI curves with restoring annual information as our priority.

Recall that our motivation of proposing PPCs is to separate annual and nonannual variation in the EVI curves. We expect that change in variability, if any, should be contained in the nonannual component. To uncover this information, we look into the fPCs of this nonannual component. There are distinct features in the first four nonannual fPCs; see Figure 5. In particular, a multiple of either the first original or nonannual fPC are added to and subtracted from the mean curve to facilitate interpretation. The plus signs represent the curves which receive positive fPC scores, while the negative signs represent the curves which receive negative fPC scores. It is observed that the first nonannual fPC is mostly positive in the first three years and mostly negative in the last three years. The real message of the first nonannual fPC is that the most dominant change of variation is the contrast of EVI relative to the cross-section mean between the first three years and the last three years. This contrast is also visualized by the gradual change of relative positions of plus and minus signs, shown in the upper panel of the lower right plot in Figure 5. There are also large peaks during the 5th and 6th years, indicating events specific to those years. The decreasing and last-two-year feature in the first nonannual fPC strongly correspond to and enhance the features observed previously in the first and second original fPCs. Further, the second to the fourth nonannual fPCs all capture information in particular years.
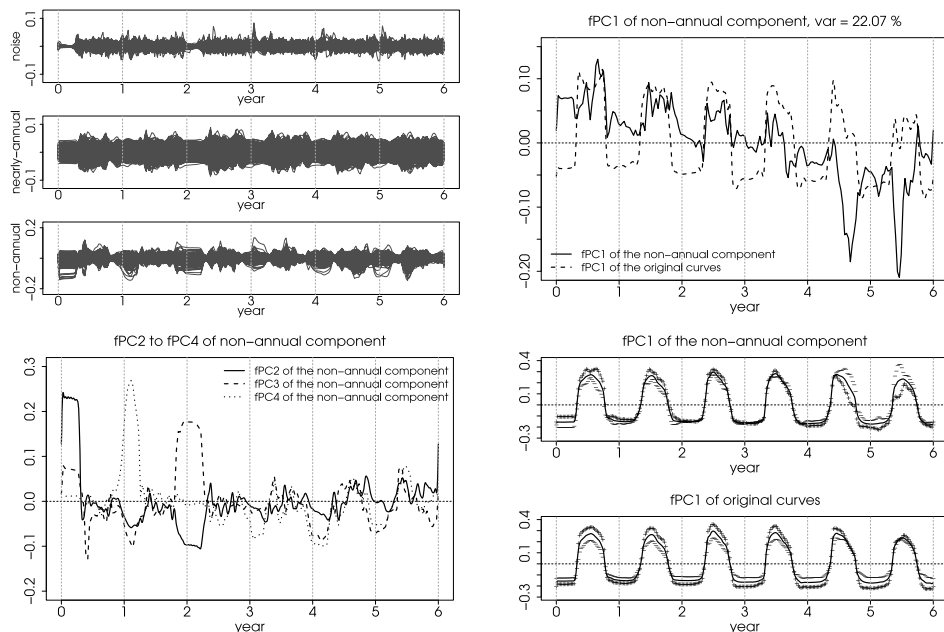
FIG. 5. *Upper left*: *Decomposition of signals. The top panel is noise which is removed when we retain the first* 46 *fPCs. The middle panel is the nearly-annual component and the bottom panel is the nonannual component. Cutoff between nearly-annual and nonannual is chosen at* 8 *first PPCs. Upper right*: *The first fPCs. The dashed curve is the first fPC of the original data. The solid curve is the first fPC of the nonannual component. Lower left*: *The second, the third and the fourth fPCs of the nonannual component. Lower right*: *Interpretation of the first fPCs of original data and nonannual component. Solid curves are the mean. In the upper panel, plus signs are mean curve plus multiple of the first nonannual fPC and minus signs are mean curve minus multiple of the first nonannual fPC. In the lower panel, plus and minus signs are the multiple of the first fPC of original data away from the mean curve.*

We can further investigate the spatial structure of the estimated aperiodic effects by plotting the scores of the first nonannual fPC on a map of pixels. The map on the right in Figure 6 show a noticeable south-west to north-east correlation structure that may be indicative of local geographic features. In preparing score maps in Figure 6, we imputed the pixels which had been excluded due to blocks of missing observations by using the functional covariance structure estimated from the retained pixels. The imputation procedure is discussed in detail in the supplemental article [Liu et al. (2012)]. The existence of evident spatial correlation may require new approaches to fPCA. Peng and Paul (2009) demonstrate that fPCA remains consistent under mild assumptions on spatial correlation. Alternatively, Allen, Grosenick and Taylor (2011) provide an approach to directly account for spatial correlation.
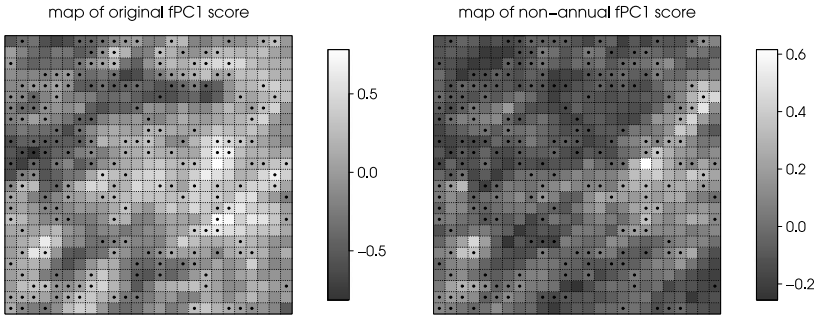
FIG. 6.   *Maps of component scores. The squares with dots in the middle represent the pixels we remove from the raw data set due to blocks of missing observations. These two score maps are naturally oriented with north at the top. Left: Projections of the first fPC onto EVI curves. Right: Projections of the first nonannual fPC onto EVI curves. A clear south-west to north-east pattern of correlation is evident in the nonannual fPC scores.*

3.7. *Tests of periodic variation.*   The high correlation between the first few PPCs and associated benchmarks gives rise to the question of whether there is exact annual variation contained in Γ_{46}, the space of leading fPCs, or PPCs (up to an orthogonal rotation). Note that the first PPC has the highest correlation with any linear combination of the annual basis. So the test of whether there is exactly annual variation contained in Γ_{46} is equivalent to testing the following hypothesis,

$$\mathbf{H}_0 : \rho_1 = 1,$$

$$\mathbf{H}_1 : \rho_1 < 1,$$

where ρ_1 is the correlation between the first PPC and its corresponding benchmark defined in (2). Note we have two ways to formulate this null hypothesis in terms of how we describe the leading fPC subspace Γ_{46}, either by the number of fPCs spanning it, or the percentage of variation it explains. We explore both formulations in the following analysis.

This null hypothesis does not follow the classical test of correlation coefficients in a multivariate setting [see, e.g., Mardia, Kent and Bibby (1979)]. Here we test that the leading principal components have a nontrivial intersection with a predefined subspace rather than the independence of pairs of linear combinations of two random vectors. To do so, we need to generate a null distribution for ρ_1 which is no longer invariant to the covariance under the null. We therefore seek an approximate least-favorable covariance by a minimal perturbation of the data so as to satisfy H_0 and then apply a bootstrap.

We first generate hypothesized curves to approximate the functional covariance under the null hypothesis based on curves $\tilde{z}_i(t)$'s. We rewrite (6) as

$$\tilde{z}_i(t) = s_{i1}^{\xi}\xi_1(t) + \sum_{j=2}^{M} s_{ij}^{\xi}\xi_j(t) + \sum_{j=M+1}^{K} s_{ij}^{\gamma}\gamma_j(t), \qquad 1 \le i \le N.$$

Under the null hypothesis, the first PPC and the first benchmark should be identical. Then we can replace the first PPC with its associated benchmark in the above equation and further write

$$(7) \qquad \bar{z}_i(t) = s_{i1}^{\xi}\theta_1(t) + \sum_{j=2}^{M} s_{ij}^{\xi}\xi_j(t) + \sum_{k=M+1}^{K} s_{ij}^{\gamma}\gamma_j(t), \qquad 1 \le i \le N.$$

$\bar{z}_i(t)$'s are called hypothesized curves under replacement. The eigenstructure of the covariance contained in $\bar{z}_i(t)$'s is an approximated least-favorable eigenstructure under the null. The correlation between the first PPC and its benchmark of $\bar{z}_i(t)$'s is 0.9999984 under both formulations of the null hypothesis, which we view as sufficiently close to 1. A distribution of the test statistic ρ_1 can now be generated based on this approximated null. One approach to obtaining a null distribution is to assume a distribution on component scores s_{ij}^{ξ} and s_{ij}^{γ}, and produce a Monte Carlo distribution of ρ_1. Here, we make no distributional assumptions and apply a bootstrap procedure instead.

We first sample with replacement from $\bar{z}_i(t)$'s to form bootstrap null curves. In order to accommodate the effect of pre-smoothing, we bootstrap residuals obtained from pre-smoothing and add them onto each bootstrapped null curve. Then we re-smooth these bootstrap observations and compute PPCs and first correlations. This testing procedure follows the same framework as that described in Li and Chiou (2011), where the authors tested the equality of functional means and covariances. Details of this procedure are provided in the supplemental article [Liu et al. (2012)].

The histogram of bootstrap correlations with fixed number of fPCs is shown in Figure 7. The correlation between the first PPC and its associated benchmark computed from the observed data is around 0.9965, which lies at the left tail of the bootstrap distribution, suggesting the major sources of variation do not cover
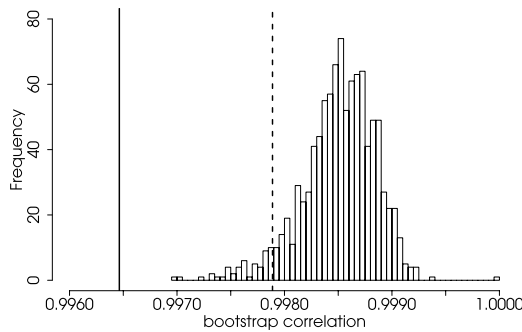

FIG. 7. *Histogram of the first correlation ρ_1 derived from bootstrap observations which are the sum of null curves bootstrapped from $\bar{z}_i(t)$'s and bootstrap residuals. The solid line is ρ_1 corresponding to the original curves of Harvard Forest data. The dashed line is the lower 0.05 critical value of the bootstrap distribution.*

the strictly periodic functions. We can also apply this test to examine a fixed percentage of variation explained instead of number of fPCs retained. The histogram of the null distribution from this test is very similar and the null hypothesis is also rejected. Readers are referred to the supplemental article [Liu et al. (2012)] for detailed results.

We have viewed the null hypothesis derived by replacing the first PPC with the first benchmark as being sufficiently close to the null hypothesis for our purposes. However, when this is not the case, the empirical first PPC correlation can be brought closer to 1 by inflating the first PPC scores. This method rescales the component score s_{ij}^{ξ} in (7) and keeps s_{ij}^{γ} fixed. This procedure allows the strength of annual signals be increased. Rescaled curves are expected to have ρ_1 enlarged toward 1. However, this rescaling should be done in a way that distorts $\bar{z}_i(t)$'s and the covariance kernel implied as little as possible. Hence, we put a penalty on the deviation of the hypothesized covariance kernel from the covariance kernel computed from $\bar{z}_i(t)$'s, then we solve an optimization problem which finds a balance between approximating the null hypothesis and controlling for divergence. Generally, let $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_M)$ be the rescaling vector. Then define hypothesized curves as

$$(8) \quad \check{z}_i(t, \boldsymbol{\tau}) = \tau_1 s_{i1}^{\xi} \theta_1(t) + \sum_{j=2}^{M} \tau_j s_{ij}^{\xi} \xi_j(t) + \sum_{k=M+1}^{K} s_{ij}^{\gamma} \gamma_j(t), \qquad 1 \leq i \leq N.$$

The covariance kernels under replacement and inflation are given by

$$\boldsymbol{\Omega}(s, t) = \lambda_1^{\theta} \theta_1(s) \theta_1(t) + \sum_{j=2}^{M} \lambda_j^{\xi} \xi(s) \xi(t) + \sum_{j=M+1}^{K} \lambda_j^{\gamma} \gamma(s) \gamma(t),$$

$$\boldsymbol{\Omega}_0(s, t, \boldsymbol{\tau}) = \tau_1^2 \lambda_1^{\theta} \theta_1(s) \theta_1(t) + \sum_{j=2}^{M} \tau_j^2 \lambda_j^{\xi} \xi(s) \xi(t) + \sum_{j=M+1}^{K} \lambda_j^{\gamma} \gamma(s) \gamma(t),$$

where $\boldsymbol{\Omega}(s, t)$ is the kernel based on curves under replacement and $\boldsymbol{\Omega}_0(s, t, \boldsymbol{\tau})$ is the hypothesized kernel based on rescaled curves $\check{z}_i(t, \boldsymbol{\tau})$'s. Under the null hypothesis, $\theta_1(t)$, $\{\xi_j(t)\}_{j=2}^{M}$ and $\{\gamma_j(t)\}_{j=M+1}^{K}$ are orthogonal to each other. It can be shown that the Kullback–Leibler divergence of $\boldsymbol{\Omega}_0(s, t, \boldsymbol{\tau})$ from $\boldsymbol{\Omega}(s, t)$ is given by

$$KL(\boldsymbol{\Omega}_0, \boldsymbol{\Omega}) = \frac{1}{2} \sum_{j=1}^{M} (\tau_j^2 - 1 - \log \tau_j^2).$$

Given $\check{z}_i(t, \boldsymbol{\tau})$'s which are functions of $\boldsymbol{\tau}$, we can compute PPCs and the first correlation $\check{\rho}_1(\boldsymbol{\tau})$. Ideally, we want to minimize $KL(\boldsymbol{\Omega}_0, \boldsymbol{\Omega})$ with the restriction that $\check{\rho}_1(\boldsymbol{\tau}) = 1$. This is achieved approximately by placing a large penalty on the difference between $\check{\rho}_1(\boldsymbol{\tau})$ and 1. Then we solve the following optimization:

$$(9) \qquad\qquad \min_{\boldsymbol{\tau}} KL(\boldsymbol{\Omega}_0, \boldsymbol{\Omega}) - \lambda \log \check{\rho}_1(\boldsymbol{\tau}),$$

where λ is a very large number. Denote the optimizer to (9) by $\hat{\boldsymbol{\tau}}$. Then, $\check{z}_i(t, \hat{\boldsymbol{\tau}})$'s are constructed according to (8). The eigenstructure implied by $\check{z}_i(t, \hat{\boldsymbol{\tau}})$'s is closer to the null hypothesis than that implied by $\bar{z}_i(t)$'s.

This procedure is investigated in detail in the supplemental article [Liu et al. (2012)] where (9) is solved with a sequence of λ values. While the first correlation obtained by this method increases, there is little effect on the test results.

4. Sampling properties of PPC. In this section we explore the stability and accuracy of PPC under random sampling. Two simulation schemes show the sensitivity and robustness of PPC in identifying annual variation.

4.1. *Sensitivity.* In this simulation scheme, we demonstrate how sensitive the PPC is in detecting annual variation. In the construction of the simulated curves, we take the linear combination of Fourier basis functions with different frequencies. We create 6 sets of simulated curves. Each set contains 200 curves and incorporates a different amount of annual variation by rescaling the coefficient of Fourier basis functions which are annual. In particular, denote the ith curve in the jth set by $a_i^j(t)$. These curves are generated as a linear combination of longer term components and annual components as follows:

(10)
$$a_i^j(t) = \sum_{k=1}^{3} \sigma_{kji1}\sin(k\omega t) + \sum_{k=1}^{3} \sigma_{kji2}\cos(k\omega t)$$
$$+ \sqrt{L_j}\bigl(\sigma_{4ji1}\sin(4\omega t) + \sigma_{4ji2}\cos(4\omega t)\bigr),$$

where $i = 1, 2, \ldots, 200$, $j = 1, 2, \ldots, 6$, $\omega = 2\pi/T$, $\sigma_{kjil} \sim \mathcal{N}(0, 1)$, i.i.d., $l = 1, 2$, $L_1 = 0$, $L_2 = 0.6$, $L_3 = 0.8$, $L_4 = 1$, $L_5 = 1.1$, and $L_6 = 1.3$.

T is the time span of the simulated curves. We take $T = 100$ and $a_i^j(t)$ spans over 4 years. Thus, $\sin(4\omega t)$ and $\cos(4\omega t)$ are sources of annual variation. The Fourier basis functions in the first two components of (10) are orthogonal to annual basis functions and thus do not contribute to the annual variation. The L_j's control the amount of annual variation. The larger the L_j, the greater the amount of annual variation. We compute PPCs with 80% of total variation cutoff in choosing how many fPCs we retain in all 6 sets. The result for $L_4 = 1$ is shown in the left 3 plots of Figure 8. The fPCs do not capture the underlying source of annual variation.

How much each sinusoidal function is reflected in retained fPCs depends on both the sample variance and covariance of σ_{k4il} and on their interaction with other sources of variation. However, $\sin(4\omega t)$ and $\cos(4\omega t)$ can be identified by PPCs even when their variation are on the same level ($L_4 = 1$) as other sources. The benchmarks exactly reproduce the annual signals, however, with phase shifting. The shifted phase is caused by the randomness in sampling σ_{44i1} and σ_{44i2}.
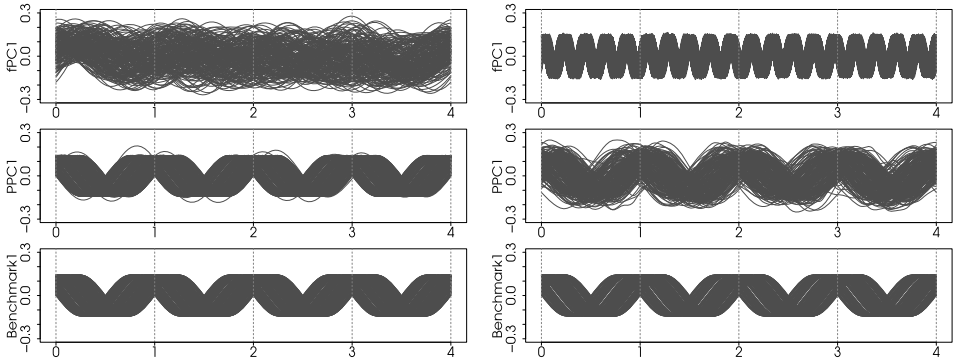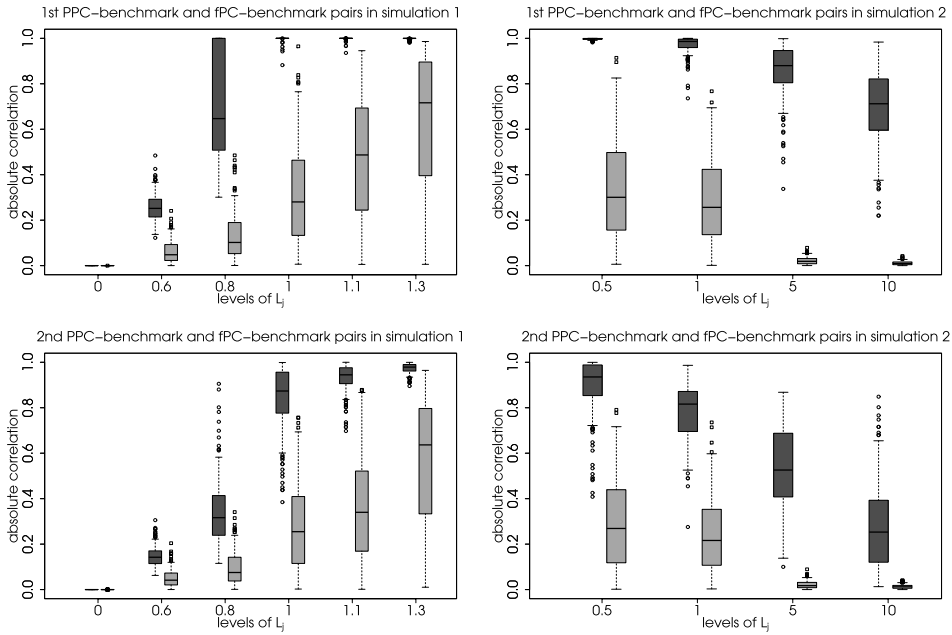
FIG. 9. *Simulation results*: *periodicity of estimated PPCs. Dark boxes are correlations between PPCs*, ξ_j *and associated benchmarks* θ_j. *Light boxes are correlations between fPCs*, γ_j *and associated benchmarks* θ_j. *Upper left*: *Simulation scheme* 1 *results on* ξ_1, γ_1 *and* θ_1; *Upper right*: *Simulation scheme* 2 *results on* ξ_1, γ_1 *and* θ_1; *Lower left*: *Simulation scheme* 1 *results on* ξ_2, γ_2 *and* θ_2; *Lower right*: *Simulation scheme* 2 *results on* ξ_2, γ_2 *and* θ_2.

the sources of annual variation which have the same amount of variation in the 4 sets of this simulation scheme. z is the frequency of HFD and is set to be 19 in our simulation. $\sin(z\omega t)$ and $\cos(z\omega t)$ are HFD whose amount of variation varies and are controlled by L_j's. Larger L_j value suggests greater amount of HFD and, hence, it is more difficult to extract annual signals for larger L_j's. In this scheme, we also use 80% as the cutoff to decide the number of fPCs we retain. The computed PPCs for $L_3 = 5$ is shown in Figure 8. With amount of HFD 5 times as great as annual variation, the fPCs are dominated by HFD and thus show a clear 19-periodic pattern. However, our first two PPCs still show a reasonably good annual pattern. To summarize results for all L_j's, we plot the fPC-benchmark and PPC-benchmark correlations of the first two pairs in the upper-right and lower-right plots in Figure 9. Again, for both pairs, the fPC-benchmark correlations are always lower than the PPC-benchmark correlations. Further, even for large HFD contamination ($L_j \geq 5$) when the fPC-benchmark correlations hover near zero, the PPC-benchmark correlations display much higher values, suggesting that the PPCs provide more robust directions compared to fPCs as the amount of HFD increases.

Based on these two simulations, we find PPCs are both sensitive and robust identifiers of the source of annual variation.

5. Conclusion. Despite the popularity of functional principal component analysis, little attention has been paid to the problem of factor rotation to improve the interpretability of modeled principal component directions. The smoothness, or ordering, properties of functional data analysis mean that factor rotation methods that are applicable for multivariate data are not always appropriate in a functional context. Conversely, new factor rotation methods may be applicable in functional data analysis that do not have analogues in multivariate statistics. As for all factor rotation methods, it is important to recall that the resulting directions are obtained as an interpretable means of representing the data, rather than independent mechanistic sources of variance.

In this paper, we have presented a factor rotation method motivated by remote sensing data and intended to improve our understanding of factors involved in ecological responses to climate change. In this data set we seek to differentiate seasonal sources of variation from both longer-term and localized effects. To do this, we present principal periodic components as a means of extracting nearly-periodic directions in the data. This factor rotation has the advantage of being efficiently implementable via canonical correlation analysis and effective at extracting periodic information. We have developed graphical tools to assess the level of periodicity in the data and to decide on thresholds between periodic and aperiodic signals. Further, a heuristic test of exact periodicity demonstrates that the addition of some further flexibility in our periodic signals is appropriate.

At its most general, our approach can be described as a rotation toward an interpretable subspace and applies to multivariate factor rotation as well as in functional data analysis. In our application, the set of periodic functions represents the most clearly relevant subspace for interpretation. However, alternative subspaces may be useful in other contexts; for example, in Koulis, Ramsay and Levitin (2008) a psychological experiment is described in which a stimulus is changed at prespecified times and a data-set of continuously-measured responses is recorded. In this case, a basis of step functions corresponding to change-times represents a relevant reference subspace with which to examine the functional response to the stimulus sequence. The choice of reference subspace depends strongly on the details of the application at hand. In our own application, we could have sought further rotations of aperiodic signals toward linear or exponential trends as a means of separating long-term effects from effects localized to individual years. Beyond this approach, we expect a more general exploration of sources of variation within the context of functional data analysis to be an important source of future research directions.

SUPPLEMENTARY MATERIAL

Description of data and details of simulation (DOI: 10.1214/11-AOAS518SUPP; .pdf). The supplementary material is divided into 3 sections. The first section provides a detailed description of the Harvard Forest data that is used in this article, including preprocessing steps. We also provide a detailed description of the imputation steps for pixels with missing observations. The second

section provides a description of *Annual Information* and its application is demonstrated through a simulation study. The last section provides results related to the bootstrap hypothesis testing procedure proposed in this article. In particular, we present the test results on the Harvard Forest data and simulation studies where we explore the empirical power curve and size on simulated data sets.

REFERENCES

ALLEN, G. I., GROSENICK, L. and TAYLOR, J. (2011). A generalized least squares matrix decomposition. Technical report, Rice Univ.

EVERSON, R., CORNILLON, P., SIROVICH, L. and WEBBER, A. (1996). Empirical eigenfunction analysis of sea surface temperatures in the Western North Atlantic. *AIP Conf. Proc.* **375** 563–590.

GERVINI, D. and GASSER, T. (2004). Self-modelling warping functions. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **66** 959–971. MR2102475

HALL, P., MÜLLER, H.-G. and WANG, J.-L. (2006). Properties of principal component methods for functional and longitudinal data analysis. *Ann. Statist.* **34** 1493–1517. MR2278365

HARTMANN, D. L. (1994). *Global Physical Climatology*. Academic Press, New York.

HE, G., MÜLLER, H.-G. and WANG, J.-L. (2003). Functional canonical analysis for square integrable stochastic processes. *J. Multivariate Anal.* **85** 54–77. MR1978177

HOLTON, J. (1992). *An Introduction to Dynamic Meteorology*. *International Geophysics Series*. Academic Press, San Diego, New York.

HUETE, A., DIDAN, K., MIURA, T., RODRIGUEZ, E. P., GAO, X. and FERREIRA, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* **83** 195–213.

KNEIP, A. and RAMSAY, J. O. (2008). Combining registration and fitting for functional models. *J. Amer. Statist. Assoc.* **20** 1266–1305.

KOULIS, T., RAMSAY, J. O. and LEVITIN, D. J. (2008). From zero to sixty: Calibrating real-time responses. *Psychometrika* **73** 321–339. MR2413184

LEURGANS, S. E., MOYEED, R. A. and SILVERMAN, B. W. (1993). Canonical correlation analysis when the data are curves. *J. Roy. Statist. Soc. Ser. B* **55** 725–740. MR1223939

LI, P.-L. and CHIOU, J.-M. (2011). Identifying cluster number for subspace projected functional data clustering. *Comput. Statist. Data Anal.* **55** 2090–2103. MR2785116

LI, Y. and HSING, T. (2010). Uniform convergence rates for nonparametric regression and principal component analysis in functional/longitudinal data. *Ann. Statist.* **38** 3321–3351. MR2766854

LIU, X. and MÜLLER, H.-G. (2004). Functional convex averaging and synchronization for time-warped random curves. *J. Amer. Statist. Assoc.* **99** 687–699. MR2090903

LIU, C., RAY, S., HOOKER, G. and FRIEDL, M. (2012). Supplement to "Functional factor analysis for periodic remote sensing data." DOI:10.1214/11-AOAS518SUPP.

MARDIA, K. V., KENT, J. T. and BIBBY, J. M. (1979). *Multivariate Analysis*. Academic Press, London. MR0560319

MÜLLER, H.-G., STADTMÜLLER, U. and YAO, F. (2006). Functional variance processes. *J. Amer. Statist. Assoc.* **101** 1007–1018. MR2324140

PARMESAN, C. and YOHE, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421** 37–42.

PENG, J. and PAUL, D. (2009). A geometric approach to maximum likelihood estimation of the functional principal components from sparse longitudinal data. *J. Comput. Graph. Statist.* **18** 995–1015. MR2598035

PIAO, S. L., CIAIS, P., FRIEDLINGSTEIN, P., PEYLIN, P., REICHSTEIN, M., LUYSSAERT, S., MARGOLIS, H., FANG, J. Y., BARR, A., CHEN, A. P., GRELLE, A., HOLLINGER, D. Y., LAURILA, T., LINDROTH, A., RICHARDSON, A. D. and VESALA, T. (2008). Net carbon dioxide losses of northern ecosystems in response to autumn warming. *Nature* **451** 49–52.

R DEVELOPMENT CORE TEAM. (2010). *R: A Language and Environment for Statistical Computing*. Vienna, Austria ISBN 3-900051-07-0.

RAMSAY, J. O. and SILVERMAN, B. W. (2002). *Applied Functional Data Analysis*. Springer, New York. MR1910407

RAMSAY, J. O. and SILVERMAN, B. W. (2005). *Functional Data Analysis*, 2nd ed. Springer, New York. MR2168993

RAMSAY, J. O., WICKHAM, H., GRAVES, S. and HOOKER, G. (2010). fda: Functional data analysis. R package version 2.2.2.

SILVERMAN, B. W. (1996). Smoothed functional principal components analysis by choice of norm. *Ann. Statist.* **24** 1–24. MR1389877

YAO, F., MÜLLER, H.-G. and WANG, J.-L. (2005). Functional linear regression analysis for longitudinal data. *Ann. Statist.* **33** 2873–2903. MR2253106

ZHANG, X., FRIEDL, M. A. and SCHAAF, C. B. (2006). Global vegetation phenology from moderate resolution imaging spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements. *Journal of Geophysical Research* **111** G04017.

C. LIU
S. RAY
DEPARTMENT OF MATHEMATICS AND STATISTICS
BOSTON UNIVERSITY
111 CUMMINGTON STREET
BOSTON, MASSACHUSETTS 02215
USA
E-MAIL: liuchong@math.bu.edu
          sray@bu.edu

G. HOOKER
DEPARTMENT OF STATISTICAL SCIENCE AND
   DEPARTMENT OF BIOLOGICAL STATISTICS
   AND COMPUTATIONAL BIOLOGY
CORNELL UNIVERSITY
1186 COMSTOCK HALL
ITHACA, NEW YORK 14853
USA
E-MAIL: gjh27@cornell.edu

M. FRIEDL
DEPARTMENT OF GEOGRAPHY
   AND ENVIRONMENT
BOSTON UNIVERSITY
675 COMMONWEALTH AVE
BOSTON, MASSACHUSETTS 02215
USA
E-MAIL: friedl@bu.edu