Proceedings PSFVIP-3

March 18 – 21 2001, Maui, Hawaii, U.S.A.

F3399

Wavelet-Based Functional Data Analysis: Theory, Applications and Ramifications[1]

BRANI VIDAKOVIC
Georgia Institute of Technology
Atlanta, GA 30332-0205, USA

Abstract. A majority of applications of wavelets in statistical data analysis are in the area of nonlinear regression and function estimation. Given a single time series, or image, it might not be possible to utilize various statistical techniques that for their implementation require more than a single observation at a fixed point in time/space. If the researcher has repeated measurements in form of functions/images, than wavelets can be successfully employed in a functional-type data analysis. In this paper we focus on wavelet based functional ANOVA procedure in which the noise is separated from the signal for different treatments, and at the same time the treatment responses are additionally split on the mean response and the treatment effect, in the spirit of traditional ANOVA. The key properties utilized are abilities of wavelets to decorrelate and regularize the inputs. Different strategies for (multivariate) shrinkage separation of treatment effects and significance testing in the wavelet domain are discussed. The methods are illustrated and evaluated on test-functions and images with controlled signal-to-noise ratios and on real data from satellite monitoring.

Keywords: ANOVA, Functional Data Analysis, Wavelet Domain.

1 Introduction

In conventional statistical practice, an observation is usually a number or a vector. But in many real-life situations, observed values are continuous curves, vectors of curves, images, or vectors of images. Prototypical examples are growth curves (e.g., measurements of height and

[1]Parts of this paper were communicated at PSFVIP-3 meeting, March 18 –21 2001, Maui, Hawaii. The research was sponsored by the NSF-DMS 0079923 award at Georgia Institute of Technology and Duke University.

weight in children at particular age, brain potentials, and a variety of responses in biological, chemical, and geophysical experiments and measurements. A vibrant statistical research in this area dealing with univariate measurements is summarized in the monograph by Ramsay and Silverman (1997).

Two characteristics are common to any functional data analysis (FDA): *a strong link with the multivariate statistical paradigm* and *the need for regularization.* The strong link with multivariate statistics arises from the fact that methods, such as principal component analysis, multivariate linear modeling, canonical correlation analysis, etc. can be applied within the functional data analysis framework. Function values $f(\mathbf{t})$, where $\mathbf{t}$ is continuous s-dimensional variable, are observed at discrete grid points.

Regularization in FDA consists of assuming a particular class of smooth estimators. The implementation is carried out in various ways. For example, one can penalize roughness as part of the fitting criteria or use particular representations that possess inherent smoothness (e.g., splines, wavelets, neural networks). Often there are links between magnitudes of coefficients in a particular representation and the regularity features of the represented objects (wavelets, pursuit methods).

In making an inference in longitudinal and spatial functional data models, two problems are of particular concern for the inference maker. The most forbidding problem is dimensionality. Explanatory variables are sampled curves or surfaces, and depending on the sampling rate, the number of variables can exceed the number of responses by factor of hundreds or thousands. Connected to the problem of dimensionality is the dependence problem. The responses from design points close in time/space are necessarily dependent and should be modeled as such. Even if the measurement errors are assumed independent, other components of the model are highly correlated in time. A standard answer to these two problems is to make inference in the domain of principal components (Besse and Ramsay, 1986; Rice and Silverman, 1993; Ramsay and Silverman, 1997) or in the wavelet domain (Fan 1996; Raz and Turetsky, 1999; Rosner and Vidakovic, 2000). The benefit of wavelet-based methods is that they solve both of the problems simultaneously and automatically.

In the next chapter we describe the statistical model and discuss its wavelet counterpart.

2 The Model, Methodology, and Simulational Example

The functional ANOVA (FANOVA) model has been utilized by several authors. For example, Ramsay and his team use the FANOVA to model lip motion from the acoustical data (Ramsay *et al.*,1996) and Fan and Lin apply it to test longitudinal effects of business advertisment.

We give a definition that is in fact "functionalized" version of a standard one way ANOVA formulation, and remark that variety of more complex models can be functionalized and dealt with in a semi-parametric or nonparametric fashion.

Suppose that for any fixed $\mathbf{t} \in \mathbf{T} \subset \mathbb{R}^s$, the observations $\mathbf{y}$ are modeled by a fixed effect

ANOVA model,

$$y_{il}(\mathbf{t}) = \mu(\mathbf{t}) + \alpha_i(\mathbf{t}) + \epsilon_{il}(\mathbf{t}),\ i = 1, \ldots, p,\ l = 1, \ldots, n_i;\ \sum_{i=1}^{p} n_i = n, \tag{1}$$

where $\epsilon_{il}(\mathbf{t})$ are independent $\mathcal{N}(0, \sigma^2)$ errors. On the other hand, when i and l are fixed we assume that functions $\mu(\mathbf{t})$ and $\alpha_i(\mathbf{t})$ are in $\mathbb{L}_2(\mathbf{T})$, and $\epsilon_{il}(\mathbf{t})$ is a normal random field. To ensure identifiability of treatment functions α_i, it is standardly imposed:

$$\int |\sum_i n_i \alpha_i(\mathbf{t})| dt = 0. \tag{2}$$

In the rest of the paper we will assume that measurements $\mathbf{y}$ are taken at dyadic grid in s-dimensional space $\mathbf{t}_m = (t_{1,m}, \ldots, t_{s,m})$,

$$t_{i,m} = m2^{-N},\ 1 \le i \le s, 1 \le m \le 2^N,$$

and that N is chosen as a power of 2.

The standard least square estimators for $\mu(\mathbf{t})$ and $\alpha_i(\mathbf{t})$

$$\begin{aligned} \hat{\mu}(\mathbf{t}) &= \bar{y}_{..}(\mathbf{t}) = \frac{1}{n} \sum_{i,l} y_{il}(\mathbf{t}), \\ \hat{\alpha_i}(\mathbf{t}) &= \bar{y}_{i.}(\mathbf{t}) - \bar{y}_{..}(\mathbf{t}). \end{aligned}$$

where $\bar{y}_{i.}(\mathbf{t}) = \frac{1}{n_i} \sum_l y_{il}(\mathbf{t})$, are obtained by minimizing the discrete version of LMSSE (Ramsay and Silverman, 1997, p. 141),

$$LMSSE = \sum_{\mathbf{t}} \sum_{i,l} [y_{il}(\mathbf{t}) - (\mu(\mathbf{t}) + \alpha_i(\mathbf{t}))]^2, \tag{3}$$

subject to the discrete version of constraint (2), $(\forall \mathbf{t})\ \sum_i n_i \alpha_i(\mathbf{t}) = 0$.

The fundamental ANOVA identity becomes functional identity,

$$\text{SST}(\mathbf{t}) = \text{SSTr}(\mathbf{t}) + \text{SSE}(\mathbf{t}), \tag{4}$$

with $\text{SST}(\mathbf{t}) = \sum_{i,l} [y_{il}(\mathbf{t}) - \bar{y}_{..}(\mathbf{t})]^2$, $\text{SSTr}(\mathbf{t}) = \sum_i n_i [y_{i.}(\mathbf{t}) - \bar{y}_{..}(\mathbf{t})]^2$, and $\text{SSE}(\mathbf{t}) = \sum_{i,l} [y_{il}(\mathbf{t}) - \bar{y}_{i.}(\mathbf{t})]^2$. If $\text{MSE}(\mathbf{t}) = \text{SSE}(\mathbf{t})/(n-p)$ and $\text{MSTr}(\mathbf{t}) = \text{SSTr}(\mathbf{t})/(p-1)$, then for each $\mathbf{t}$, the function

$$F(\mathbf{t}) = \frac{\text{MSTr}(\mathbf{t})}{\text{MSE}(\mathbf{t})} \tag{5}$$

is distributed as non-central $F_{p-1,n-p}\left(\frac{\sum_i n_i \alpha_i^2(\mathbf{t})}{\sigma^2}\right)$.

As hinted in the introduction, an automatic pointwise application of the standard ANOVA in the above method is hindered by dependence and dimensionality. Wavelet transformations are known to be whitening transformation. Their decorrelation properties are now well explored and quantified (Wornell, 1996). Thus, it is natural to transform the discretized FANOVA model to the wavelet domain and proceed with statistical inference there. In addition to preserving the structure of the original linear model, observations in the wavelet domain will be almost uncorrelated and suitably "prepared" for dimension reduction.

Let $\mathbf{d}$ be a wavelet transformation of $\mathbf{y}$, $\mathbf{d} = \mathbb{W}\mathbf{y}$. The discrete "times" $\mathbf{t}_m$ are replaced by a standard multiresolution indexing $(j, \mathbf{k})$, in which j corresponds to a scale level and $\mathbf{k}$ corresponds to location shifts. Denote by $\mathcal{I}$ the set of all indices $(j, \mathbf{k})$ in the transformed data. Due to linearity and orthogonality of $\mathbb{W}$,

$$\begin{aligned} d_{il}(j,\mathbf{k}) &= \theta_i(j,\mathbf{k}) + \epsilon'_{il}(j,\mathbf{k}) \\ &= \theta(j,\mathbf{k}) + \tau_i(j,\mathbf{k}) + \epsilon'_{il}(j,\mathbf{k}), \end{aligned} \tag{6}$$

is the wavelet transform of model (1). Of course, $\sum_i \tau_i(j,\mathbf{k}) = 0$ for any fixed $(j,\mathbf{k}) \in \mathcal{I}$. Moreover, the ANOVA estimators in the time domain, and inverse transformations of the estimators in the wavelet domain coincide. The following result is an easy consequence of orthogonality of wavelet transformations.

Result: Let $\hat{\mu}$ and $\hat{\alpha}_i$ be least-square estimators of μ and α_i. If $\hat{\theta}$ and $\hat{\tau}_i$ are least-square estimators of θ and τ_i, and

$$\begin{aligned} \tilde{\mu} &= \mathbb{W}^{-1}\hat{\theta}, \\ \tilde{\alpha}_i &= \mathbb{W}^{-1}\hat{\tau}_i,\ i = 1, \ldots, p, \end{aligned}$$

then $\hat{\mu} \equiv \tilde{\mu}$ and $\hat{\alpha}_i \equiv \tilde{\alpha}_i$. The *energy preservation* of orthogonal wavelet transformations also implies,

$$\sum_{\mathbf{t}\in\mathbf{T}} \mathrm{MSE}(\mathbf{t}) = \sum_{j,\mathbf{k}\in\mathcal{I}} \mathit{WMSE}(j,\mathbf{k}) \tag{7}$$

$$\sum_{\mathbf{t}\in\mathbf{T}} \mathrm{MSTr}(\mathbf{t}) = \sum_{j,\mathbf{k}\in\mathcal{I}} \mathit{WMSTr}(j,\mathbf{k}) \tag{8}$$

where *WMSE* and *WMSTr* are the wavelet-domain counterparts of MSE and MSTr.

2.1 Inference in the Wavelet Domain

Benefits of carrying out an analysis in the wavelet domain, rather than in the domain of original observations, comes from decorrelation and regularization considerations for the estimators, and from dimension reduction (energy compression) for the tests.

Ramsay and Silverman (1997) discuss several approaches for incorporating regularization in functional linear models. Standard methods involve criteria that penalize roughness. Also

choosing relatively few basis functions in some basis representation of a model can lead to regular estimators. This second method, sometimes intertwined with particular roughness penalty methods, is especially convenient when observations are wavelet-transformed. This comes from the fact that wavelets are unconditional bases for a variety of smoothness spaces, and appropriate regularizations can easily be implemented by selecting and shrinking wavelet coefficients. Research on statistical estimation using wavelets is vibrant, an overview can be found in Vidakovic (1999) and Walden and Percival (2000).

Problem of testing involving curves and surfaces is considered in time series literature and in spatial statistics (see Shumway 1998 and Cressie 1111). Hall and Hart (1990) propose bootstrap testing procedures, Eubank and Hart (1992) and Eubank and LaRiccia (1992) propose cross-validation based tests. The use of Fourier and wavelet domains is suggested by Fan (1996) and Fan and Lin (1998).

We illustrate the need for dimension reduction in the testing problem by an example.

Example. Many tests involving sampled curves and surfaces can be simplified to a paradigmatic problem of testing the multivariate normal mean. Suppose we wish to test

$$H_0 : \theta = 0 \ \ vs. \ \ H_1 : \theta \neq 0, \tag{9}$$

if $\mathbf{X} \ \sim \ \mathcal{MVN}_n(\theta, I_n)$ is observed. Sufficient statistics is $||\mathbf{X}||^2$. If the alternative is $\theta = \theta_1$ then the α-level ML ratio test has approximate power

$$1 - \Phi\left(\frac{z_{1-\alpha} - ||\theta_1||^2/\sqrt{2n}}{\sqrt{1 + 2||\theta_1||^2/n}}\right) \approx 1 - \Phi\left(z_{1-\alpha} - ||\theta_1||^2/\sqrt{2n}\right). \tag{10}$$

When $||\theta_1||^2$ goes to infinity we expect the power to tend to 1, but if $||\theta_1||^2 = o(\sqrt{n})$, the power tends to α when n increases. Thus, in testing too many dimensions, the discriminatory quantity $||\theta_1||^2$ is weakened to $||\theta_1||^2/\sqrt{2n}$

Fan (1996) proposed testing only m out of total n dimensions via an adaptive Neyman truncation test. For the m-dimensional subproblem, Fan's adaptive Neyman test statistics is

$$T^*_{AN} = \max_{1 \leq m \leq n} \left\{ \frac{1}{\sqrt{m}} \sum_{j=1}^{m} (X_j^2 - 1) \right\}. \tag{11}$$

Too large values of T^*_{AN} reject H_0. Fan and Lin (1998) tabulate critical values of

$$T_{AN} = a_n T^*_{AN} - b_n, \tag{12}$$

where $a_n = \sqrt{2 \log \log n}$ and $b_n = 2 \log log n + 0.5 \log \log \log n - 0.5 \log(4\pi)$. for selected values of n and level α. They also show that T_{AN} converges in distribution to extremal III distribution, but the convergence is slow.

2.2 Testing in the WANOVA setup

Let in the model (1) we are interested in testing

$$H_0 : \alpha_i(\mathbf{t}) = 0,\ i = 1 \ldots, p, \tag{13}$$

versus the general alternative. This is equivalent to test

$$H_0 : \theta_i(j, \mathbf{k}) = \theta(j, \mathbf{k}),\ i = 1, \ldots, p, \tag{14}$$

for $(j, \mathbf{k}) \in \mathcal{I}$.

There are two challenges of this approach. First, what are representative indices and what is the optimal number of them. In the wavelet context it is natural that the optimal indices possess the most of the energy, i.e., select m wavelet coefficients with largest magnitude.

As regards the optimal m, several approaches are possible. Wavelet thresholding techniques provide a variety of selections for m.

Suppose that m is fixed and that $\mathcal{I}_m$ is a set of indices $(j, \mathbf{k})$ with m largest magnitudes. Then,

$$T_m^* = \frac{1}{\sqrt{2m(p-1)}} \left[\sum_{(j,\mathbf{k}) \in \mathcal{I}_m} \sum_{i=1}^{p} \left(\frac{\bar{d}_i(j, \mathbf{k}) - \bar{d}(j, \mathbf{k})}{\hat{\sigma}^2(j, \mathbf{k})/\sqrt{n_i}} \right)^2 - m(p-1) \right], \tag{15}$$

where $\bar{d}_i(j, \mathbf{k}) = 1/n_i \sum_{l=1}^{n_i} d_{il}(j, \mathbf{k}), \bar{d}_(j, \mathbf{k}) = 1/n \sum_{i=1}^{p} n_i \bar{d}_i(j, \mathbf{k})$, and $\hat{\sigma}^2(j, \mathbf{k}) = WMSE(j, \mathbf{k})$.

The level α test rejects H_0 if

$$T_m^* \geq \frac{1}{\sqrt{2m(p-1)}} \left(\chi^2_{m(p-1)}(1-\alpha) - m(p-1) \right). \tag{16}$$

or, alternatively, one can normalize T_m^* as in (12) and use the tables provided by Fan and Lin (1998).

2.3 A Simulational Study

To illustrate the proposed methodology we provide a simulational study with known components of the model of a 2-D signal. For examples involving 1-D signals see Rosner and Vidakovic (2000). The components were circle image $\mu(\mathbf{t})$ as the main effect or mean image and 4 treatment effect images, $\alpha_i(\mathbf{t}),\ i = 1, \ldots, 4$. The images $\alpha_1(\mathbf{t}), \alpha_2(\mathbf{t})$ and $\alpha_3(\mathbf{t})$ are test images `lena, cameraman,` and `barbara`, respectively, (for sources see `www.isye.gatech.edu/~brani/datapro.html`), while $\alpha_4(\mathbf{t}) = -\frac{1}{n_4}(n_1\alpha_1(\mathbf{t}) + n_2\alpha_2(\mathbf{t}) + n_3\alpha_3(\mathbf{t}))$ is fixed because of the constraint (2).

The number of observations was $n = 20$ and the treatment sample sizes were $n_1 = 5, n_2 = 5, n_3 = 5$, and $n_4 = 5$. Unbalanced designs are possible, as well.

Each observation from the treatment i is an 256×256 image, $\mu_i(\mathbf{t}) + \epsilon(\mathbf{t}) = \mu(\mathbf{t}) + \alpha_i(\mathbf{t}) + \epsilon(\mathbf{t})$ taken at equally spaced grid points $\mathbf{t}_m$ The wavelet used was Coiflet with 3 vanishing

moments for the scaling function. All functional components of the model are rescaled so that the signal-to-noise ratio [SNR] of $\mu + \alpha$ (signal) to ϵ (noise) is 0.4 and the size of noise is 200 (i.e., each pixel is distributed as $N(0, 200^2)$).

To get estimators of $\mu(\mathbf{t})$ and $\alpha_i(\mathbf{t})$, we applied the Bayesian FDR procedure with level $\alpha = 0.1$, and posterior probabilities given by (**??**) with $\mu = (\hat{\sigma^2})^{-1}$ where $\hat{\sigma^2}$ is a robust estimator of the variance (exhibited using the level of finest detail in the corresponding wavelet decomposition).

function	time domain LS estimate	wavelet domail LS estimate after shrinkage [from 2.7% coeffs]
μ	1997	1170
α_1	6012	3509
α_2	5963	3560
α_3	6006	3369
α_4	5977	3179
total	25955	14787

Table 1: Mean square errors in the domains of original data and wavelets.

In Table 1, the first column of numbers are squared errors of the ANOVA estimators from the data domain, while the second column contains the errors of wavelet processed estimators. For example $1/256^2 \sum_{\mathbf{t}}[\hat{\mu}(\mathbf{t}) - \mu(\mathbf{t})]^2 = 1997, 1/(256)^2 \sum_{\mathbf{t}}[\tilde{\mu}(\mathbf{t}) - \mu(\mathbf{t})]^2 = 1170$. The errors are exactly calculated since the true value of images are known.

3 An Application in Cloud/Temperature Mapping

EUMETSAT is an intergovernmental organization created through an international convention agreed by 17 European Member States. EUMETSAT's Meteosat system is intended primarily to support the National Meteorological Services (NMS) of Member States. The NMS in turn distribute the image data to other end users, notably through the provision of forecasts on television several times a day. The Meteosat system provides continuous meteorological observations from space to a large user community. In addition to the provision of images of the Earth and its atmosphere every half an hour in three spectral channels (Visible, Infrared and Water Vapour), a range of processed meteorological parameters is produced.

The satellite receives that part of the sun radiation which is reflected by the earth surface or by cloudiness. It is a so-called *window channel* which means that radiation is not significantly absorbed by the gases in the troposphere. The satellite receives radiation which is emitted by the earth and the clouds because of their temperature. Infra Red (IR) images via is window channel (Wavelength 10-13 μ) are useful for day and night cloud-mapping and determination of surface temperature. Different grey shades in the IR channel represent different temperatures of the radiating surface which can be the earth surface or the cloud tops. In practice,

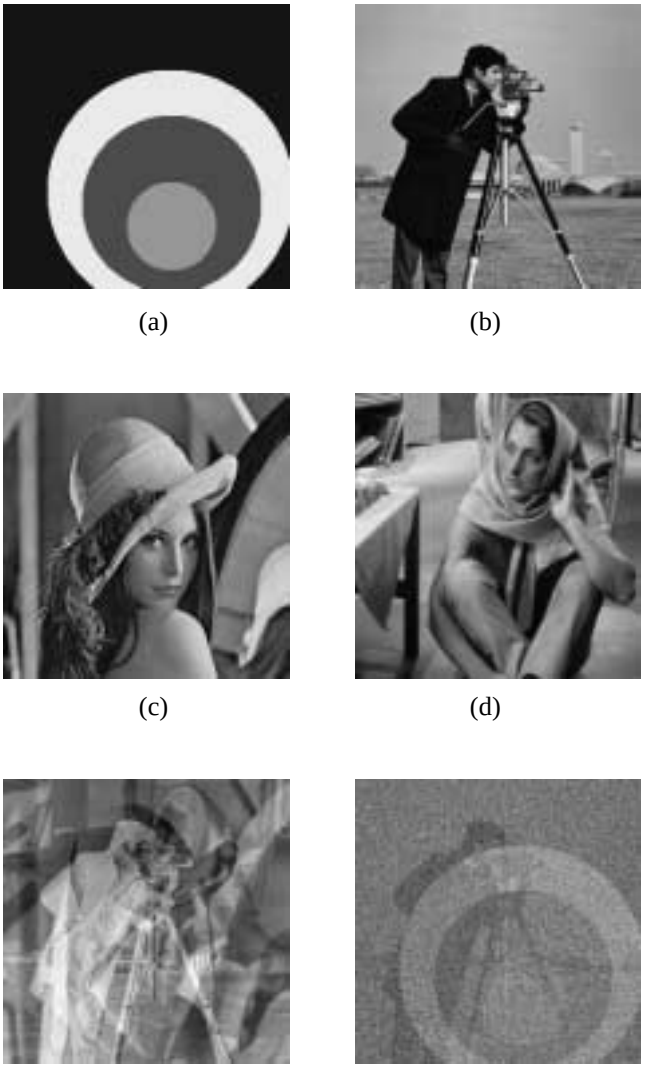


Figure 1: (a) The true mean image $\mu(\mathbf{t})$ (circles); (b) The treatment image $\alpha_1(\mathbf{t})$ (cameraman); (c) The treatment image $\alpha_2(\mathbf{t})$ (lena); (d) The treatment image $\alpha_3(\mathbf{t})$ (barbara); (e) Mish-mash treatment image. It is a negative superposition of other treatment images: cameraman, lena, and barbara. The sum of all four treatment images is set to zero because of identifiability reasons. (f) A typical observation, one of 20 used in the simulational study. This is a superposition of the mean image (circles), treatment image (cameraman), and the noise. Signal-to-noise ratio (measured as ratio of standard deviations) is 0.4;

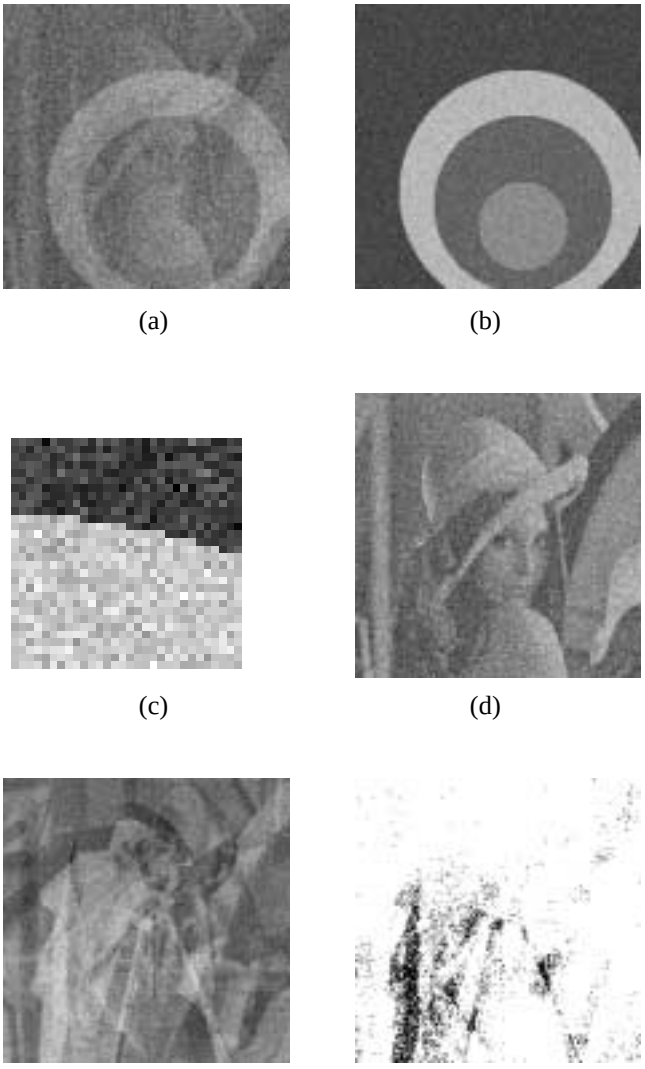


Figure 2: (a) The observation $\mu(\mathbf{t})+\alpha_2(\mathbf{t})+\epsilon(\mathbf{t})$ (circles, lena, and noise); (b) The least-square estimator of $\hat{\mu}(\mathbf{t})$; (c) Enlarged portion of the estimator form the panel (b); (d) Estimator of the treatment image $\hat{\alpha}_2(\mathbf{t})$ (lena); (e) Estimator of the mish-mash treatment, $\hat{\alpha}_4(\mathbf{t})$; (f) F-test for equality of treatments was performed. Significant F statistics (> 3.24 white) and nonsignificant (< 3.24 black) are depicted. There is about 85% significant F statistics in the image. Value 3.24 is 95% percentile of $F_{3,16}$ distribution.

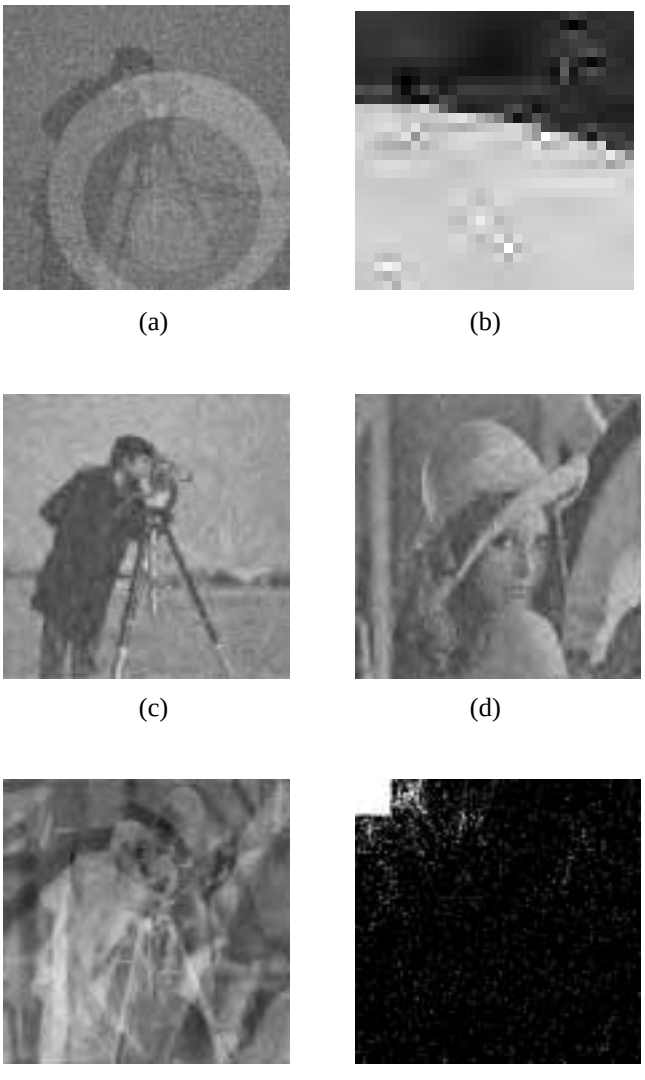


Figure 3: Estimators from the wavelet domain: (a) Estimator of $\tilde{\mu}(\mathbf{t})$; (b) Enlarged portion of the estimator form panel (b); (c-d) Estimators of treatment images $\tilde{\alpha}_1(\mathbf{t})$ (cameraman), $\tilde{\alpha}_2(\mathbf{t})$ (lena), and $\tilde{\alpha}_4(\mathbf{t})$ (mish-mash); (f) Significant F statistics in the wavelet domain (white, $F > 3.24$). Note that only 2.70% of the positions have a significant F statistic.

grey shades presented in the IR images are reversed (white becomes black) to get appearance similar to that of a visual image (Wavelength 0.5 μ).

Our data set contains 36 IR satelite images of the Gulf of Guinea (West cost of Africa and South Atlantic Ocean) taken at 12 consequtive days (1/4/2001 - 1/15/2001). The images are divided into 3 groups according to the hour of their acquisition: (i) morning (6:00am) group, (ii) noon (12:00pm) group, and (iii) evening (18:00pm) group. A typical observation (6:00 am, 1/4/01) is depicted in Figure 4(a).

WANOVA estimator of the grand mean ($\mu(\mathbf{t})$) is the image shown in Figure 4(b). We focus on one treatment (6:00am) only. Panels (c) and (d) in Figure 4 depict the estimators of the treatment effect ($\alpha_1(\mathbf{t})$) and treatment mean ($\mu_1(\mathbf{t})$).

Panels (e) and (f) emphasize differences of WANOVA estimators and pointwise ANOVA estimators in the time domain. Shown are differences of the mean effect and the treatment effect (6:00 am), respectively.

Figure (3) schematically depicts the decompositions $\mu_i(\mathbf{t}) = \mu(\mathbf{t}) + \alpha_i(\mathbf{t})$ for the three treatments, $i = 1, 2$ and 3.

4 Discussion

In this article we addressed some benefits of making a functional ANOVA inference in the wavelet domain. Various generalizations are possible: (i) to extend class of functional models amenable to wavelet transformations, and (ii) to consider different shrinkage techniques.

In this paper it was assumed that the errors were homoscedastic in both – time and treatments. Straightforward generalizations are possible to relax this equality of variances.

It is also possible to use different wavelet bases for data coming from different treatments, or to select bases adaptively, from the wavelet packet library, for instance. These issues are subject of ongoing research. Readers interested in mimicking the calculations from this paper are welcome to request matlab m-files from the author. Donoho's `wavelab` package was used.

References

[1] ABRAMOVICH, F. and BENJAMINI, Y. (1995). Thresholding of wavelet coefficients as a multiple hypotheses testing procedure, In: *Wavelets and Statistics*, Editors A. Antoniadis and G. Oppenheim. Lecture Notes in Statistics, **103**, 5–14. Springer-Verlag, New York.

[2] ABRAMOVICH, F. and BENJAMINI, Y. (1996). Adaptive thresholding of wavelet coefficients. *Comput. Statist. Data Anal.* **22**, 351–361.

[3] BENJAMINI, Y. and HOCHBERG, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, *J. Roy. Statist. Soc. Ser. B.* **57,** 289–300.

[4] BERGER, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*, Springer Verlag.

[5] CLYDE, M. A., PARMIGIANI, G., and VIDAKOVIC, B. (1998). Multiple Shrinkage and Subset Selection in Wavelets. *Biometrika*, **85**, 391–402.

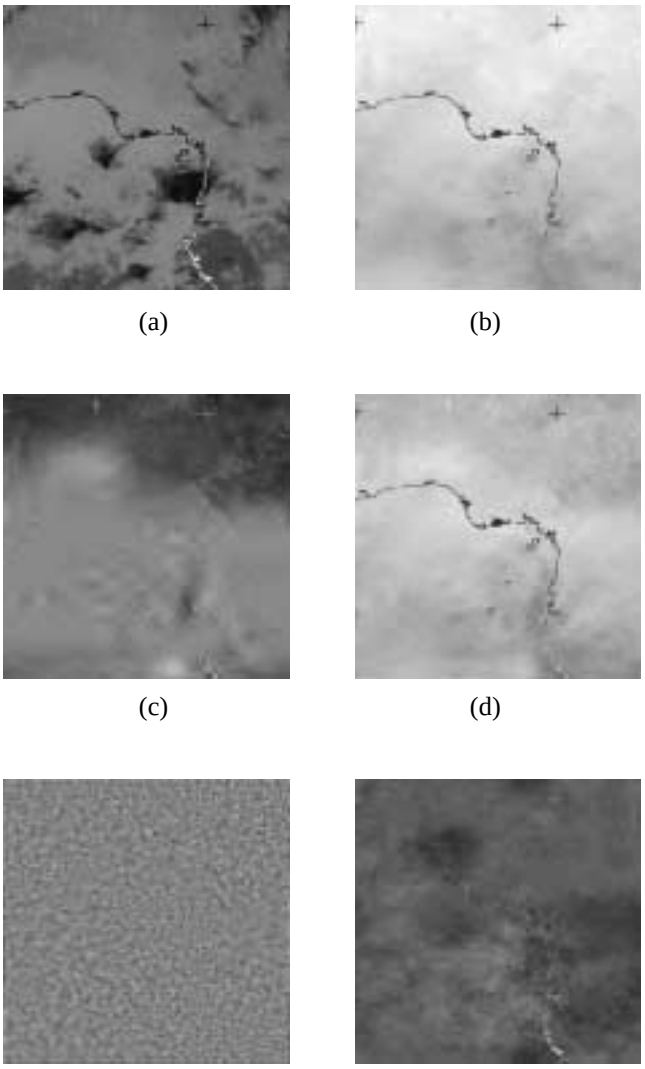
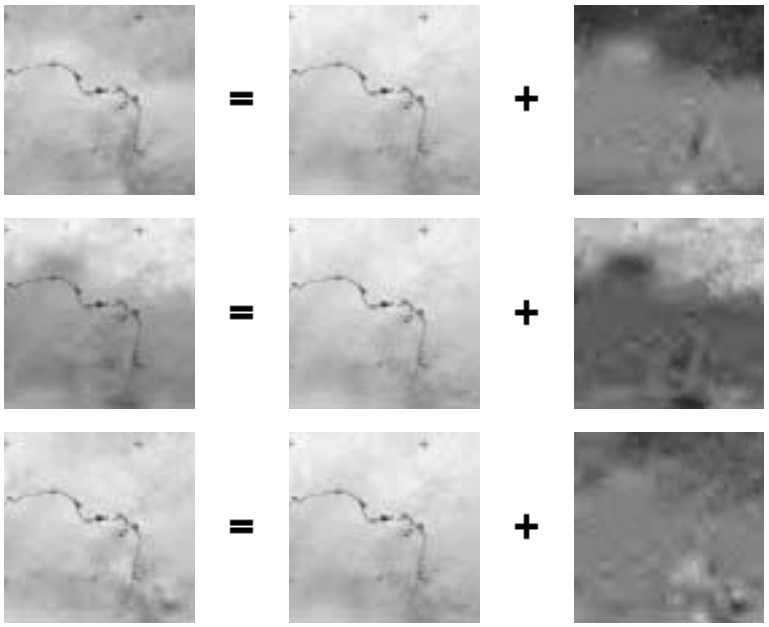


Figure 4: (a) A typical observation, 4/1/2001 at 6:00am; WANOVA estimators of (b)$\mu(\mathbf{t})$, (c)$\alpha_1(\mathbf{t})$, and (d) $\mu_1(\mathbf{t})$; (e) difference of two estimators for $\mu(\mathbf{t})$ WANOVA and direct point-wise ANOVA; (f) The same as (e) but for $\alpha_1(\mathbf{t})$.



[6] FAN, J. (1996). Test of significance based on wavelet thresholding and Neyman's truncation. *J. Amer. Statist. Assoc.* **91**, 674 – 688.

[7] FAN, J. and LIN, S-K. (1998). Test of significance when data are corves, *J. Amer. Statist. Assoc.* **93**, 1007–1021.

[8] KNEIP, A. and GASSER, T. (1992). Statistical tools to analyze data representing a sample of curves. *Annals of Statistics,* **20,** 1236–1265.

[9] MÜLLER, P. and VIDAKOVIC, B. (1999). MCMC Methods in Wavelet Shrinkage: Non-Equally Spaced Regression, Density and Spectral Density Estimation, In: *Bayesian Inference In Wavelet Based Models*, Editors P. Müller and B. Vidakovic, Lecture Notes in Statistics,**141**, 187–202, Springer Verlag, New York.

[10] RAMSAY, J. O. and DALZELL, C. J. (1991). Some tools for functional data analysis (with discussion.) *Journal of the Royal Statistical Society, Series B,* **53,** 539–572.

[11] RAMSAY, J.O. and SILVERMAN, B. W. (1997). *Functional Data Analysis,* Springer-Verlag, New York.

[12] RAMSAY, J. O., BOCK, R. D., and GASSER, T. (1995) Comparison of height acceleration curves in the Fels, Zurich and Berkeley growth data. *Annals of Human Biology,* **22,** 413–426.

[13] RAMSAY, J. O. (1995). Principal differential analysis: Data reduction by differential operators. *Journal of the Royal Statistical Society,* **58,** 495–508.

[14] RAMSAY, J. O., MUNHALL, K. G., GRACCO, V. L. and OSTRY, D. J. (1996) Functional data analysis of lip motion. *Journal of the Acoustical Society of America,* **99,** 3718–3727.

[15] RAZ J. and TURETSKY B. (1999). Wavelet ANOVA and fMRI. Wavelet Applications in Signal and Image Processing VII, Proceedings of the SPIE, 3813, 561–570.

[16] RICE, J. A. and SILVERMAN, B. W. (1991). Estimating the mean and covariance structure nonparametrically when the data are curves. *Journal of the Royal Statistical Society, Series B,* **53,** 233–243.

[17] ROSNER, G. AND VIDAKOVIC, B. (2000). Wavelet Functional ANOVA, Bayesian False Discovery Rate, and Longitudinal Measurements of Oxygen Pressure in Rats, Technical Report 1/2000, ISyE, Georgia Institute of Technology.

[18] SILVERMAN, B. W. (1995). Incorporating parametric effects into functional principal components analysis. *Journal of the Royal Statistical Society, Series B.*

[19] VIDAKOVIC, B. (1998). Nonlinear Wavelet Shrinkage With Bayes Rules and Bayes Factors, *Journal of the American Statistical Association*, **93**, 441, 173–179.

[20] WORNELL, G. W. (1996). *Signal Processing with Fractals: A Wavelet Based Approach*, Prentice Hall, Englewood Cliffs, NJ.

BRANI VIDAKOVIC
INDUSTRIAL AND SYSTEMS
ENGINEERING
GEORGIA INSTITUTE OF TECHNOLOGY
765 FERST DR., GROSECLOSE BLDG 437
ATLANTA, GA 30332
brani@isye.gatech.edu

