UC3M Working Papers
Statistics and Econometrics
15-06
ISSN 2387-0303
April 2015

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-48

A Random Walk Test for Functional Time Series

Nicola Mingotti[a], Rosa Lillo[b], Juan Romo[c]

Abstract

In this paper we introduce a Random Walk test for Functional Autoregressive Processes of Order One. The test is non parametric, based on Bootstrap and Functional Principal Components. The power of the test is shown through an extensive Montecarlo simulation. We apply the test to two real dataset, Bitcoin prices and electrical energy consumption in France.

Keywords: Autoregressive Process, FAR(1), unit root, Bootstrap, Computational Statistics, hypothesis test, Principal Components.

[a] Corresponding author. Phd student in the Department of Statistics, Universidad Carlos III de Madrid, Getafe, Madrid, Spain (E-mail: nicola.mingotti@uc3m.es).

[b] Professor in the Department of Statistics, Universidad Carlos III de Madrid, Getafe, Madrid, Spain (E-mail: rosaelvira.lillo@uc3m.es).

[c] Professor in the Department of Statistics, Universidad Carlos III de Madrid, Getafe, Madrid, Spain (E-mail: juan.romo@uc3m.es).

Acknowledgements: the authors acknowledge financial support from the Spanish Ministry of Economy and Competition, research project ECO2012-38442.

A Random Walk Test for Functional Time Series

Nicola Mingotti, Rosa E. Lillo, Juan Romo

April 30, 2015

Abstract

In this paper we introduce a Random Walk test for Functional Autoregressive Processes of Order One. The test is non parametric, based on Bootstrap and Functional Principal Components. The power of the test is shown through an extensive Montecarlo simulation. We apply the test to two real dataset, Bitcoin prices and electrical energy consumption in France.

Key Words. Autoregressive Process, FAR(1), unit root, Bootstrap, Computational Statistics, hypothesis test, Principal Components.

1 Introduction

Testing if an Autoregressive Process of order one is a Random Walk is a well known subject of classical Time Series analysis that was introduced firstly by [Box and Jenkins, 1970] and then improved in posterior works as for example [Dickey and Fuller, 1979], [Phillips and Perron, 1988] and [Ferretti and Romo, 1996]. Random Walks are of especial interest in Economics, they were used since [Bachelier, 1900] to model stock prices. The works of [Nelson and Plosser, 1982] and [Meese and Rogoff, 1983] have shown Random Walks can be found not only stock prices but even in unsuspected time series, as unemployment rate, gross domestic product or currency exchange. In this article we export this classic test to a functional context and introduce an algorithm to check if Functional Autoregressive Process of order one can be considered a Random Walk.

The transposition of Time Series techniques to the Functional Data context is motivated by the same arguments as Functional Data Analysis in general: continuous nature of the data, dimension reduction and an expected

improvement in the predictive power of the model. Anyway, in Time Series, FDA can bring some notable simplification in the models, we will show it with an example. Suppose you are going to model the daily electricity consumption in a country, the classic way to do it is to join a deterministic trend with a more or less complex stochastic model, for example an ARIMA. One of the main difficulties is that there are many important periodic factors to take into account occurring in a year, for example all weekends but also special occasions as Christmas, Easter and summer holidays. Every one of these special events has to be separately taken into account and this make the model complex and loaded with exceptions. One of the *desiderata* in using a Functional Model is that it could be able to cope automatically with many of the recurring periodic events. Indeed, if we model electricity consumption as a function that cover one full year, then year-2 will take into account at least what happened in the previous year, year-1, where all of the holidays and weekends were already present so, in principle, we would not need to adjust for many of the seasonalities. One exception to this scheme are moving holidays, like Easter, that would need still manual correction.

The last ten years have seen a lot of advances in Functional Time Series and Functional Autoregressive processes, both in theory and in applications. For example, [Horváth et al., 2010] proposed a method to check if a model can be considered constant in time, [Battaglia, 2005] proposed a method to identify outliers, [Kokoszka and Reimherr, 2013] a method to establish the order of an autoregressive process. In applications, [Damon and Guillas, 2002] used functional autoregressive processes for Ozone forecasting, [Besse et al., 2000] to forecast ocean temperatures, [Guillas et al., 2011] to forecast the seabed evolution and maintain navigability channels. As of today, [Bosq, 2000] is the *de facto* reference for the theoretical aspects of Functional Autoregressive Processes while the recent [Horváth and Kokoszka, 2012] collects many recent results and is directed to the researcher as well as to the practitioner in FDA.

In this article we are going to focus on a very special topic of Functional Time Series. We present a Random Walk test for Functional Autoregressive Processes of order one. As always in Functional Data, the first impulse is to start by looking at what was done in the past to solve the unfunctional problem. Some of the classical methods to test the Random Walk hypothesis in an AR(1) process were cited at the beginning of the article, all of them, in one way or another, try to estimate the value ρ in $X_{i+1} = \rho\ X_i + \varepsilon_{i+1}$ and compare it against the condition H_0: $\rho = 1$. We choose not to follow this direct approach in the functional context because the ruling equation for a FAR(1) process is $X_{i+1}(t) = \Psi\ X_i(t) + \varepsilon_{i+1}(t)$ and the available estimator for Ψ converges very slowly and in an unexpected[1] way, see [Horváth and Kokoszka, 2012, pg.240], [Kokoszka and Zhang, 2010].

1. The estimator depends on principal components but increasing the number of principal components reduces the performance of the estimator.

More precisely, in the functional context, the model corresponding to the Autoregressive Process of order one, AR(1), defined by

$$X_{i+1} = \rho\, X_i + \varepsilon_{i+1}, \tag{1}$$

is the Functional Autoregressive Process of order one, denoted for short as FAR(1) and defined by

$$X_{i+1}(t) = \Psi(X_i)(t) + \epsilon_{i+1}(t), \tag{2}$$

where X_i have mean zero. In the general setting of [Bosq, 2000], $X_i(t)$ are functions in an Hilbert space, Ψ is a bounded linear operator from H to H and $\epsilon_i(t)$ are H-white noise. In this work we will restrict our attention to the framework most used in applications, as in [Horváth and Kokoszka, 2012]. H will be $L^2[0,1]$, the space of functions in $[0,1]$ which are square integrable according to Lebesgue. The scalar product is the common $<f,g> := \int_0^1 f\,g$, $\epsilon_i(t)$ are i.i.d. with $E(\epsilon_i(t)) = 0$ and $E\,(||\,\epsilon_i(t)||^2) < \infty$ for all i. As Ψ, we consider only integral operators of type

$$\Psi(f)(t) = \int_0^1 \psi(t,s)\, f(s)\, d\, s, \tag{3}$$

which are always linear and bounded if the kernel function $\psi(t,s)$ is continuous. It might be useful to consider this operator as an integral average of f respect to a set of weight functions ψ_t. These operators are not too restrictive on what we can express, on the contrary, they are quite general. Indeed, quoting from [Gohberg et al., 1990, ch.7], they are like an *universal model* for Hilbert-Schmidt[2] operators because for each H-S operator A: $H \to H$ there exists a unitary operator U such that $U\,A\,U^{-1}$ is an integral operator as Eq.3.

We want to check if the dataset $X_1, ..., X_m$, which we suppose was generated by a process as Eq.2, can be considered a Random Walk. The idea is to compare the covariance of the original data set with the covariance of the same dataset resampled under the null hypothesis that Ψ is the Identity operator, using the Empirical Functional Principal Principal Component (EFPCA).

Further in this article, Section 2 contains a detailed description of the test and of how to apply it in practice. In Section 3 is explained the Montecarlo simulation scheme and presented and analysis of its results. Section 4 shows

2. An operator A on an Hilbert space, is an Hilbert-Schmidt operators if it is linear, continuous and such that $\sum_{i=1}^{\infty} ||\,A\,\phi_i||^2 < \infty$ for an orthonormal basis $\phi_1, \phi_2, \ldots$. There are also other equivalent definitions, see [Gohberg et al., 1990, pg.140].

two applications to real datasets, we check for Random Walk the series of yearly electricity consumption in France and diary prices of Bitcoin. It also clarifies the procedure given in section 2 and can be red just after it by the reader most interested in the test deployment. In Section 5 there are the conclusions and finally in the Appendix are collected the simulation results.

2 Test Procedure

We start with a functional data set $X_1(t)$, ..., $X_m(t)$ and compute its first $p \leq m-1$ *Empirical Functional Principal Components*. The EFPC are eigenfunctions and eigenvalues of the empirical covariance operator and they will be denoted with with $(\hat{\xi}_i, \hat{\lambda}_i)_{i=1..p}$, see [Ramsay and Silverman, 2005, ch.8]. It is known that, under mild conditions, when m goes to infinity, $\hat{\lambda}_i$ converges to λ_i, the eigenvalues of the populational covariance operator, see [Horváth and Kokoszka, 2012, pg.31], [Bosq, 2000, sec.4.2], [Dauxois et al., 1982].

The Schmidt Norm of an Hilbert-Schmidt operator A can be computed as

$$|| A||_S^2 = \sum_{j=1}^{\infty} || A\,\phi_j||^2 \tag{4}$$

where $\phi_1, \phi_2, ...$ is an orthonormal basis, see [Gohberg et al., 1990, pg.141-143]. In our case A is the covariance operator K_Ψ and choosing as orthonormal basis its eigenfunctions given by PCA: $\xi_1, \xi_2, ...$ we get

$$|| K_\Psi||_S^2 = \sum_{j=1}^{\infty} || K_\Psi\,\xi_j||^2 = \sum_{j=1}^{\infty} || \lambda_i\,\xi_j||^2 = \sum_{i=1}^{\infty} \lambda_i^2. \tag{5}$$

In applications we will use $\hat{\lambda}_i$ as estimator for λ_i so we will actually compute $\widehat{|| K_\Psi||}_S^2$ but, using the aforementioned result of convergence and the continuity of the norm, for large values of m, $\widehat{|| K_\Psi||}_S^2$ converges to $|| K_\Psi||_S^2$. It must be stressed that there are not problems of convergence for the series in Eq.5 since we are summing only a finite number of terms $\hat{\lambda}_i$ for $i = 1...m-1$.

Under the null hypothesis that the FAR(1) process is a Random Walk we can estimate its innovations as

$$\hat{\epsilon}_{i+1}(t) := X_{i+1}(t) - X_i(t). \tag{6}$$

Resampling $\hat{\epsilon}_i(t)$ and applying again the null hypothesis H_0 we can compute B Bootstrap copies of the observations set $\{X_i(t)\}_i$ as

$$X^{b,*}_{i+1}(t) := X^{b,*}_i(t) + \epsilon^{b,*}_{i+1}(t) \qquad \text{for } b \text{ in } 1 \dots B. \tag{7}$$

For each set of resampled observations $\{X^{b,*}_i(t)\}_i$ we compute the estimated Schmidt norm of its covariance $(\widehat{||K_{Id}||}^{b,*}_S)^2$ through principal components as done before. Then, build its empirical distribution function $\mathcal{E}$. Finally, we compute $\widehat{||K_\Psi||}^2_S$ for the original data set and reject the null hypothesis that Ψ is Id if $\mathcal{E}(\widehat{||K_\Psi||}^2_S)$ is smaller than some predefined threshold α.

From the Bootstrap structure and the statistic involved it is apparent the nature of the test is

$$\begin{cases} H_0 \colon \Psi = I\,d \\ H_1 \colon ||\,\Psi|| < 1. \end{cases} \tag{8}$$

Indeed, if Ψ has norm one then there will be no reduction on the impact of all old innovations (ϵ_j with $j \leqslant i$) on X_i, so X_i could grow very fast and consequently also the covariance matrix and its eigenvalues, just as in the case of $\Psi = I\,d$. This condition make Ψ and Id indistinguishable to our test.

Suppose now our dataset comes as a matrix $M_{n,m}$ where data to be considered functions are the columns $c_1, ..., c_m$. We take for granted here that data are already aligned and equispaced, techniques to make a dataset in this form are discussed in [Ramsay and Silverman, 2005].

Before starting the Random Walk test, by our understanding of the real phenomenon underlying the data, we check if there is a structure that can be considered deterministic and remove it applying a proper transformation. Bosq suggests to remove trend but not seasonality, if the seasonality component can be modeled through a FAR(1), see [Bosq, 2000, pg.152, 240].

Perhaps, to appreciate the importance of the previous step the following imaginary experiment can be of help. Suppose, you are studying the daily phone calls functions in the Italian mobile telephone network. It is well known to all Italian people, that on New Year's Eve night the mobile network collapses, and it becomes difficult to send the important "Happy new year!" messages. This will happen every year, it is a deterministic structure and you need to remove it because the FAR(1) model has no way to get it straight since it looks only at what happened on the night before, that is on the not so special night of 30 December. On the contrary, suppose now you are studying the yearly phone calls functions. In this case you should not, in principle, adjust for New Year's Eve because the previous year also

contains it.

To apply the Random Walk test the following steps must be performed.

1. Smooth the dataset fitting each column $c_1 \dots c_m$ to a base of your preference. We used BSpline and Fourier. The number of basis functions is largely to be selected depending on the nature of the problem but a good starting point could be $\sqrt{n}$. At the end of this step we will have the set of functional observations $X_1(t) \dots X_m(t)$.

2. Choose a number of principal components you want to use for your analysis. We suggest to start with $p=3$ and adjust it *ex post* if necessary. There are classical ways of selecting the number of principal components based for example on the *scree plot* or on the *explained variance* but, according to our simulations, the power of the test is concentrated only on the first eigenvalue. Using $p>1$ is instrumental in showing if the sequence of eigenvalues displays a continuous decay, which is characteristic of ordinary FAR(1) processes. Or a sudden extreme drop after the first eigenvalue, which is a characteristic of a Random Walk. A graphical representation of this phenomenon is presented in the Applications section.

3. Find the Empirical Functional Principal Components of your dataset and their associated eigenvalues $\hat{\lambda}_1 \dots \hat{\lambda}_p$, then compute the estimated (squared) Schmidt Norm $\widehat{||K_\Psi||}_S^2 \leftarrow \sum_{i=1}^{p} \hat{\lambda}_i^2$. It is important to observe that we compute the EFPC always on the centered[3] dataset, in this case it is $\{\widetilde{X}_i | \widetilde{X}_i := X_i - \bar{X} \text{ for } i=1 \dots m\}$.

4. Resample the dataset under the null hypothesis H_0 that $\Psi = \mathrm{Id}$. First compute the estimated innovations,

$$\hat{\epsilon}_{i+1}(t) \leftarrow X_{i+1}(t) - X_i(t) \qquad \text{for } i=2 \dots m$$

   Center the $\hat{\epsilon}_{i+1}(t)$ subtracting their common mean and find the resampled observations. From here on all the $\hat{\epsilon}_i$ will be considered centered.

$$\begin{cases} X_1^{b,*}(t) \leftarrow \mathrm{Mean}(X_1(t), ..., X_m(t)) \\ X_{i+1}^{b,*}(t) \leftarrow X_i^{b,*}(t) + \hat{\epsilon}_{i+1}^{b,*}(t) \qquad \text{for } i=2 \dots m \end{cases}$$

3. If you use *R* with *fda* package for your computations, this can be achieved automatically setting the parameter `centerfns=TRUE` in function `pca.fd` .

Compute the associated Schmidt norm $(\widehat{||K_\Psi||}_s^{b,*})^2$ as in step (3). Some observations about the choice of the mean as first bootstrapped observations will be given at the end of the procedure.

5. Repeat step (4) for the necessary amount of iterations until you get B estimations of $(\widehat{||K_\Psi||}_s^{b,*})^2$. We usually start start with $B = 200$. Find the empirical distributions $\mathcal{E}$ of all the $(\widehat{||K_\Psi||}_s^{b,*})^2$. Let P-value be $\mathcal{E}(\widehat{||K_\Psi||}_s^2)$. Traditionally reject H_0 if P-value < 0.05, or some other threshold of your choice.

6. Toggle the number of basis functions, principal components and Bootstrap replications to make sure results are stable.

The initial bootstrapped observation $X_1^{*,b}$ is a free variable in this problem because we can estimate only $m-1$ innovations. We choose to set it to $\bar{X}$ following this heuristic. If we want the Bootstrap to replicate something we have to give it a chance to do it right, the best chance to start seems to be the middle of the original dataset. Other possibilities are reasonable but were not tested, for example one could start the simulation picking randomly one X_i, or the median. A graphical comparison of the initial dataset and some Bootstrap replications can give some hints about the suitability of the selected starting point criterion.

3 Simulation Study

In this section we will describe in detail the simulation procedure. It has been implemented in the programming language *R* and it uses the external package *fda* which is presented in [Ramsay et al., 2009]. Steps in which *fda* function library is central will be denoted with *(fda)*.

There are a number of global numerical parameters that control the simulation, they are summarized here in the next table for ease of reference.

n.obs	Number of functional observations. $X_1(t) \dots X_{\boldsymbol{n.obs}}(t)$.
n.pt	Number of raw (x, y) coordinates describing each function $X_i(t)$.
n.basis	Number of basis functions used to represent each $X_i(t)$.
n.pc	Number of functional principal components to consider.
n.boot	Number of bootstrap replications.
sd	Raw indicator of error function magnitude. It appears in the simulation of *Brownian Motion* and *Brownian Bridge*.

The following list will describe all steps to perform in order to simulate a FAR(1) process and test it against the Random Walk null hypothesis.

1. Define an operator Ψ of type (3) providing a kernel function $\psi(s,t)$.

2. Multiply the kernel function by a constant C to be able to see how the test performs with kernels that have same algebraic structure but different size.

3. Compute constant $\hat{C}_{Id}$ such that the following relation is satisfied

$$|| \Psi||_S^2 = \int\int_U \hat{C}_{Id}^2 \, \psi^2(t,s) \, \mathrm{dt} \, \mathrm{ds} = 1. \tag{9}$$

4. Set the random number generator to a fixed value to provide reproducible results.

5. Define a tuple of constants that will multiply C_{Id} and vary the kernel size.

$$\mathrm{kset} \leftarrow (0.5, 0.7, 0.8, 0.9, 0.95, 0.99, 1.0) \tag{10}$$

6. For each value K in kset repeat what follows 100 times and store the final result.

    a. Create an initial FAR(1) data set with $X_0(t) = f_0(t)$ and $X_{i+1}(t) = \Psi(X_i(t)) + \epsilon_{i+1}(t)$. f_0 is set initially to $f_0(t) = (x^2+1) + Sin(8\pi x)$. There are no special reasons to choose this function, but error functions parameters were selected have a reasonable order of magnitude compared to it, the precise *signal to noise ratio* in general depends on applications. The error functions $\epsilon_i(t)$ can be independent trajectories of *Brownian Bridge* (BB) or *Brownian Motion* (BM). They are generated from a cumulative sum of 100 Normal independent random variables with $\mu = 0$ and σ to be set to **sd**. A detailed discussion can be found in [Iacus, 2009, sec.1.6, 1.8]. We use ten different kernel functions $\psi(s,\, t)$, some of them appear in literature like the the *Gaussian*, *Wiener* and *Parabolic kernels* in [Gabrys et al., 2010], the others were introduced by the authors.

    b. (fda) Following a common rule of thumb we create a BSpline or Fourier functions basis with number of elements $\boldsymbol{n}$**.basis** $\leftarrow$

$\lfloor \sqrt{\mathbf{n.pt}} \rfloor$ and fit all the $X_i(t)$ to the new functions space.

c. (fda) Perform a principal component analysis on the functional data set, extract the eigenvalues $\hat{\lambda}_1 ... \hat{\lambda}_p$ corresponding to the first **n.pc** empirical principal components and set $\widehat{||K_\Psi||}_S^2 \leftarrow \sum_{i=1}^{\mathbf{n.pc}} \hat{\lambda}_i^2$.

d. Compute the estimated residuals $\hat{\epsilon}_i(t)$ under H_0 and center them subtracting their common mean.

$$\hat{\epsilon}_{i+1}(t) \leftarrow X_{i+1}(t) - X_i(t) \tag{11}$$

From here on all occurences of $\hat{\varepsilon}_i$ will denote centered residuals.

e. Create **n.boot** copies of the data data set $\{X_i(t)\}_i$, each time resampling the residuals with the following rule

$$X_{i+1}^{b,*}(t) \leftarrow X_i^{b,*}(t) + \hat{\epsilon}_{i+1}^{b,*}(t). \tag{12}$$

The parameter b is the bootstrap index and varies in $1 ... \mathbf{n.boot}$.

f. (fda) For each family of bootstrapped observations compute the first **n.pc** empirical principal components, with their respective eigenvalues and set $\left(\widehat{\|K_{\mathrm{Id}}\|}_S^{b,*}\right)^2 \leftarrow \sum_{i=1}^{p} (\hat{\lambda}_i^{b,*})^2$.

g. Build the empirical distribution function of the $\left(\widehat{\|K_{\mathrm{Id}}\|}_S^{b,*}\right)^2$ values and name it $\mathcal{E}$.

h. The P-value associated to the current experiment will be $\mathcal{E}(\widehat{||K_\Psi||}_S^2)$. We reject H_0 that is, we reject that Ψ is the Identity if P-value < 0.05.

7. In the results tables we read the rejection rate for each K, the number of experiments in which H_0 was rejected divided by 100.

We applied the previous procedure to several possible setups. All experiments results are in tabular format and all tables can be found in the Appendix. The table following this paragraph (Table 1) is a reference to

all of the simulation results. In the first column it tells the *code* of the experiment, a unique identifier by which it is possible to find the appropriate table in the Appendix. In the second, it tells if Ψ was a constant, if not, it reports the formula for its kernel $\psi(s,t)$. "N" tells the number of Montecarlo replications used to estimate the power of the test and *base* specifies if we are using the BSpline of Fourier basis. Columns **n.boot**, **n.pc** and **sd** have the same meaning previously defined.

code	Ψ	N	**n.boot**	**n.pc**	base	**sd**
b1	constant	100	200	1	S	0.5
b2	constant	100	200	1	S	0.05
b4	constant	100	1000	1	S	0.5
b5	constant	100	1000	1	S	0.05
b7	constant	100	200	3	S	0.5
b8	constant	100	200	3	S	0.05
b10	constant	100	200	1	F	0.5
b11	constant	100	200	1	F	0.05
k1b1	$e^{-(s^2+t^2)}$	100	200	1	S	0.5
k1b2	$e^{-(s^2+t^2)}$	100	200	1	S	0.05
k1b7	$e^{-(s^2+t^2)}$	100	200	3	S	0.5
k2b1	$e^{(s^2+t^2)}$	100	200	1	S	0.5
k2b2	$e^{(s^2+t^2)}$	100	200	1	S	0.05
k3b1	$min(s,t)$	100	200	1	S	0.5
k3b2	$min(s,t)$	100	200	1	S	0.05
k4b1	$(t-\frac{1}{2})^2+(s-\frac{1}{2})^2$	100	200	1	S	0.5
k4b2	$(t-\frac{1}{2})^2+(s-\frac{1}{2})^2$	100	200	1	S	0.05
k5b1	$(t+\frac{1}{2})^2+(s+\frac{1}{2})^2$	100	200	1	S	0.5
k5b2	$(t+\frac{1}{2})^2+(s+\frac{1}{2})^2$	100	200	1	S	0.05
k6b1	$Sin(2\pi t+s)$	100	200	1	S	0.5
k7b1	$Sin(2\pi s+t)$	100	200	1	S	0.5
k8b1	$Sin(2\pi s)Sin(2\pi t)$	100	200	1	S	0.5
k9b1	$\lvert Sin(2\pi s)Sin(2\pi t)\rvert$	100	200	1	S	0.5
k10b1	$Sin(8\pi s)Sin(8\pi t)$	100	200	1	S	0.5

Table 1. Experiments reference.

Suppose we are interested in simulations with a *Gaussian Kernel* so choosing for example the experiment *k1b2*, here below is a copy of what we will find in the Appendix. In the first column it is stated if error is of type *Brownian Motion* or *Brownian Bridge*. In the second the sample size, that is the number of simulated functional observations $X_1 \ldots X_m$ entering the test. The third column tells C_{id} was multiplied by 0.5, and the remaining tells C_{id} was multiplied by $0.7, 0.8$, etc.

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	1	0.98	0.86	0.75	0.64
	100	1	1	1	1	0.9	0.64
	200	1	1	1	1	1	0.75
	500	1	1	1	1	1	0.87
bb	50	1	0.96	0.9	0.64	0.36	0.2
	100	1	1	1	0.92	0.68	0.33
	200	1	1	1	1	0.95	0.36
	500	1	1	1	1	1	0.67

Table 2. Tabular output corresponding to experiment **k1b2**.

It must be noticed that in this table there is not $K = 1.00$ multiplying C_{Id}, the last value is 0.99. This was made on purpose to stress that when $||\Psi|| = 1$ ($||\Psi||_S$ is the best estimator of $||\Psi||$ we have) our test has no chances to determine if $\Psi = I\,d$.

3.1 Analysis

There are very different kind of operators in the simulation from which different results are expected. Experiments with code **b1**, **b2**, ..., **b11** all have a constant operator and are a special case to see how the algorithm perform when H_0 is true, when $\Psi = I\,d$. Operators kernels in experiments **k1X**, **k2X**, **k3X**, **k4X**, **k9X** are all positive and symmetric. The remaining kernels in experiments **k6X**, **k7X**, **k8X**, **k10X** are not positive on all the domain $[0, 1] \times [0, 1]$.

We observe that the power of the test is much higher when there are non positive kernels. In all corresponding tables the power of the test is very high, even with a small sample size. This fact is easy to explain intuitively, a fixed sign kernel is, generally speaking, more similar to the identity operator than a non fixed sign operator.

All experiments with constant operator and positive definite kernels display a similar triangular structure in the *power* distribution. Power of the test increases increasing the sample size, and decreases when the kernel Schmidt norm goes near to one. The performance increasing with the sample size is what we expect from every statistical test, in general, more information is available, more the decision task is simplified. The identity operator has uniform[4] norm one. The Schmidt norm is an estimator from the top of the uniform norm. When the Schmidt norm of the kernel is forced by multiplicative constant to go near to one then the operator is forced to become,

4. The uniform norm for an operator A is defined as $||A||_L := \underset{||x|| \leq 1}{sup} ||A\,x||$, it is proved that $||A||_L \leq ||A||_S$, that is, the Schmidt norm dominates the uniform norm, see [Bosq, 2000] pg.34-35

from the point of view of the norm, similar to the identity. Consequently, becoming the operator more similar to Id, it becomes more difficult for the test to discriminate between them.

There is an operator that is positive definite but gets *power* much higher then the other ones. It is the *Wiener* kernel, **k3X**. Its power is as high as the one associated to non-positive kernels. Probably the reason for this anomaly is that the kernel $\psi(s,t)$ of the Wiener process is just the covariance of Brownian Motion, which is the model for our innovations.

Changing the error term from Brownian Bridge to Brownian motion had little influence on the power. Also changing the size of the error, from $\mathrm{sd}=0.5$ to $\mathrm{sd}=0.05$ had not visible impact.

Comparing **k1b7-k1b1** and **b1-b8**, we do not see an important change in performance changing the number of principal components. Comparing **b1-b10** and **b2-b11** we do not observe important changes in power passing to Fourier basis. Comparing **b1-b4** and **b2-b5** again display not important changes, so also increasing the number of bootstrap resamples does not seem to increase much the power of the test.

4 Applications to Real Dataset

4.1 Electrical energy consumption in France

We analyze France electrical energy consumption from the beginning of 1996 to the end of 2012. The dataset is available from *RTE France* one the Web[5]. As stated by the provider, data covers power consumption in metropolitan France area, except Corsica. It includes losses on the network but it does not take into account power withdrawn by hydroelectric installations. The paper by [Cho et al., 2013] uses our same dataset and provides supplementary informations about energy consumption and factors influencing it.

Electricity consumption is observed every 30 minutes. Following [Cho et al., 2013] we study the series of weekly average consumption. We apply a logarithm transformation to cope with the apparent increase in variance and remove the trend which was estimated with a LOESS using *R* default parameters. The original dataset and its appearance after each transformation is shown in Fig.1

After the transformations we are working with a 52x16 matrix, on columns we have years, on rows weeks. Each column of the matrix is transformed (smoothed) into a functional object in $[0,1]$ using a BSpline basis with $\lfloor\sqrt{52}\rfloor$

5. `http://clients.rte-france.com`

basis functions, knots are distributed equidistantly between zero and one, included. Applying our test with **n.pc** $\leftarrow$3, and **n.boot** $\leftarrow$200 we get P-value zero. The estimated $\widehat{||K_\Psi||}_S^2$ is $8.0 \cdot 10^{-7}$ and the quartiles for $\left(\widehat{\|K_{\mathrm{Id}}\|}_S^{b,*}\right)^2$ are $1.7 \cdot 10^{-6}$, $1.6 \cdot 10^{-5}$, $2.7 \cdot 10^{-5}$, $5.3 \cdot 10^{-5}$, $1.09 \cdot 10^{-3}$. Changing the number of principal components, doubling the number of basis functions and doubling the bootstrap replications did not affect the result significantly. The two plots in Fig.2 show that, besides the Schmidt norm, there is a significant difference in the structure of the estimated eigenvalues of K_Ψ and K_{Id}. The first ones decrease smoothly in magnitudes. For the Random Walks instead the first eigenvalues dominates all the others.

In conclusion, we can reject the null hypothesis that Electrical Energy consumption in France is a Random walk because the P-value is very small and the decay of the eigenvalues of the original dataset is quite continuous whilst, under the Random Walk hypothesis, the first eigenvalue dominates all the others.

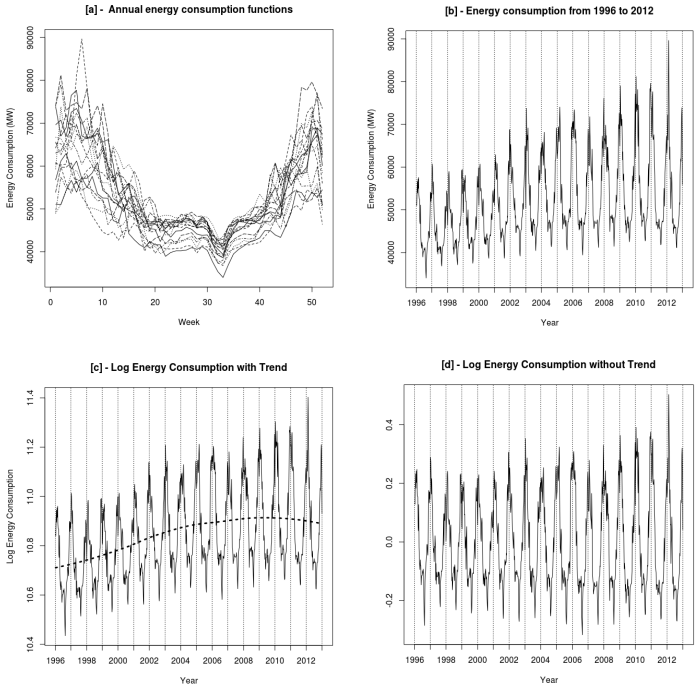

Figure 1. These plots represent electrical power consumption in France. [a] All yearly consumption's are plotted together. [b] Historic plot of energy consumption, from 1996 to 2012. [c] Log is applied to previous plot with LOESS trend function superimposed. [d] The trend is removed from the previous.

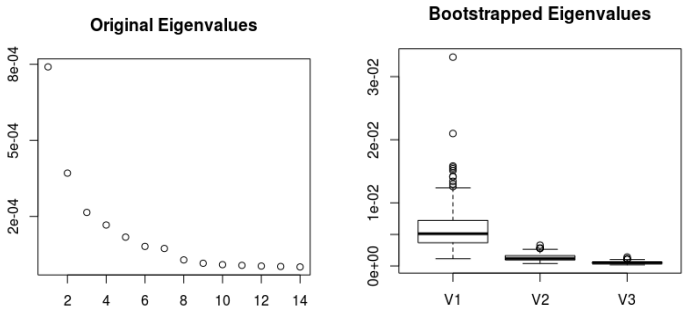

Figure 2. The two plots show the estimated eigenvalues of K_Ψ on the right, and the bootstrapped estimated eigenvalues of K_{Id} on the left.

4.2 Bitcoin daily prices

Bitcoin is a virtual currency introduced by [Nakamoto, 2008]. Bitcoins are traded twenty-four hours per day, all days of the year, prices are known to have large variability and suffered a burst on beginning of 2014 after a period of explosive growth. In the recent paper by [Kristoufek, 2013], it was shown that the series of average daily prices between 1-May-2011 and 30-June-2013 is non stationary. The paper provides some introductory information about the Bitcoin currency to which the interested reader may refer.

We follow [Kristoufek, 2013] using the same dataset[6] and the same temporal window for our investigations. But, instead of using daily average prices, we will consider the much more detailed series of daily prices, where each day is seen as a functional observation $X_i(t)$.

The Bitcoin prices we have are the ones processed at *Mt.Gox*, once the largest currency trading center. They are available for free on the Internet. The number of trades in *Mt.Gox* during the considered time period is extremely variable. For six days there were no transactions at all, these days were removed from our analysis. Excluding zeros, the minimum number of transactions per day was 373, the maximum 66293, the other deciles: 2042, 2722, 3448, 4255, 5000, 5994, 7322, 9483, 13736.

The series of trade prices was passed to logarithm and detrended. Fig.3 illustrates the dataset at each step. Then, the dataset was divided in day blocks. For each daily data was build a piecewise linear interpolating function with domain in $[0, 1]$. Each function was sampled in 2000 equidistant points from zero to one creating a matrix of 2000x786 elements where every column corresponds to a day. The data matrix was converted to a set of func-

6. File `mtgoxUSD.csv` at `http://api.bitcoincharts.com/v1/csv/`

tional observations respect to a BSpline basis with $\lfloor\sqrt{2000}\rfloor$ basis function, knots were equispaced between zero and one. We applied our test for unit root setting **n.pc** $\leftarrow$3 and **n.boot** $\leftarrow$200. The resulting P-value was 0.63, the estimated $\widehat{||K_\Psi||}_S^2$ is 0.43 and the quartiles for $\left(\widehat{\|K_{\mathrm{Id}}\|}_S^{b,*}\right)^2$ are 0.011, 0.14, 0.30, 0.62, 14.29. The null hypothesis can not be rejected. Moreover, comparing the structure of eigenvalues in Fig.4 we see they are similar, the first eigenvalue dominates all the others and also, the magnitude of the first eigenvalue for K_Ψ is approximately the median of the bootstrapped eigenvalues of K_{Id}. Increasing the number of principal components and doubling the number of basis functions did not affect the P-value significantly.

Finally, we can not reject the null hypothesis that Bitcoin daily prices are ruled by a Random Walk. Indeed, the P-value is substantially larger than 0.05 and also, the decay of eigenvalues of the original dataset is similar to the one observed on the bootstrapped datasets under the null hypothesis that the series is a Random Walk.

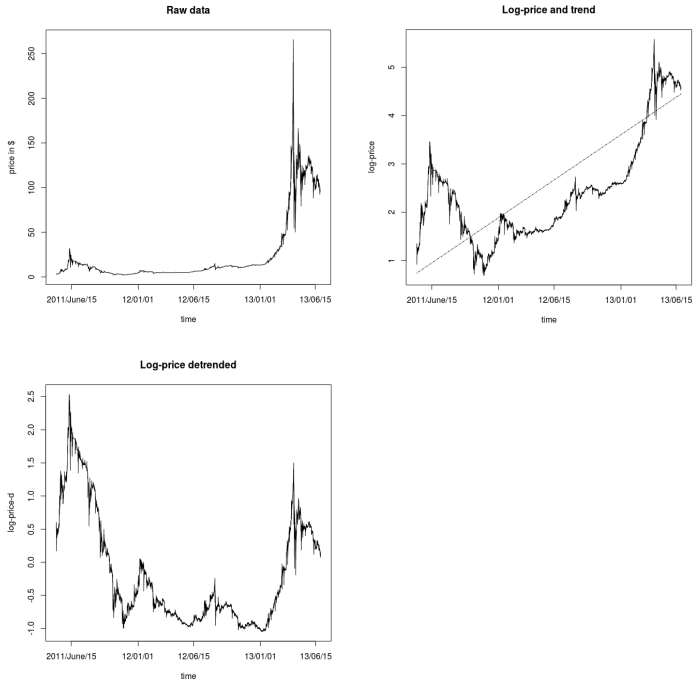

Figure 3. In the top left pane is represented the original dataset, Bitcoin trading price in U.S. dollars. On the top-right, the series after applying the logarithm with a dashed line superimposed for the trend. On the bottom, the the data after logarithm transform and removal of the trend.

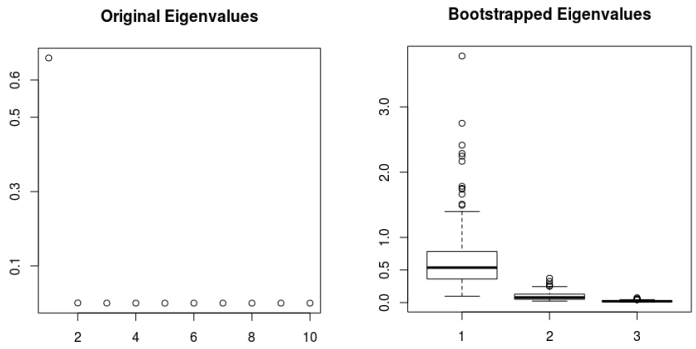

Figure 4. Plot of K_Ψ eigenvalues and the boxplots of bootstrapped K_{Id} eigenvalues.

5 Conclusions

Using $\widehat{||K_\Psi||}_S^2$ as test statistic associated with our Bootstrap scheme has given very good results in the numerical simulations. Comparing our power results to the ones previously found for AR(1) in [Ferretti and Romo, 1996] we see they are surprisingly high. The comparison is not completely correct because we are working on a different framework but there is no other Random Walk test on FAR(1) against which we could compare.

It has emerged from our investigations that a Random Walk tends to have a first large eigenvalue that dominates all the following ones. Indeed, this feature can be of help in deciding if to reject the null hypothesis in cases in which the P-value would be near to the rejection region. Or also, in cases in which the P-value alone would give a result in sharp contrast with our intuition about the problem.

Applying the test to the two real data sets has given the results we were expecting from visual inspection. Yearly innovations in France electrical energy consumption can not be considered a Random Walk. On the other side, it is not rejected that Bitcoin daily prices could be a random walk. It was necessary to apply the common tools of time series analysis before entering the test: remove the trend and adjust for variance. Indeed, a FAR(1) process defined according to equations (2) and (3) has not enough analytical

freedom to cope with this conditions which are not of local nature, but global external factors.

As all simulations in functional data, there are many parameters that could be tuned: the choice of the basis, the number of basis functions, the number points in the real data set and their relative distance, the errors, their dependence structure, the data smoothing, the empirical distribution smoothing and so on. We tried to stick to the most widespread choices, the most common errors types and operators, and the most popular initial analysis setups. There are large possibilities for further experimentation and, of course, for a desirable theoretical development in support of the numerical evidence.

Bibliography

[Bachelier, 1900] Bachelier, L. (1900). *Théorie de la spéculation*. Gauthier-Villars.

[Battaglia, 2005] Battaglia, F. (2005). Outliers in functional autoregressive time series. *Statistics & probability letters*, 72(4):323–332.

[Besse et al., 2000] Besse, P. C., Cardot, H., and Stephenson, D. B. (2000). Autoregressive forecasting of some functional climatic variations. *Scandinavian Journal of Statistics*, 27(4):673–687.

[Bosq, 2000] Bosq, D. (2000). *Linear processes in function spaces: theory and applications*, volume 149. Springer.

[Box and Jenkins, 1970] Box, G. E. and Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. Holden-Day.

[Cho et al., 2013] Cho, H., Goude, Y., Brossat, X., and Yao, Q. (2013). Modeling and forecasting daily electricity load curves: a hybrid approach. *Journal of the American Statistical Association*, 108(501):7–21.

[Damon and Guillas, 2002] Damon, J. and Guillas, S. (2002). The inclusion of exogenous variables in functional autoregressive ozone forecasting. *Environmetrics*, 13(7):759–774.

[Dauxois et al., 1982] Dauxois, J., Pousse, A., and Romain, Y. (1982). Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference. *Journal of multivariate analysis*, 12(1):136–154.

[Dickey and Fuller, 1979] Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431.

[Ferretti and Romo, 1996] Ferretti, N. and Romo, J. (1996). Unit root bootstrap tests for ar (1) models. *Biometrika*, 83(4):849–860.

[Gabrys et al., 2010] Gabrys, R., Horváth, L., and Kokoszka, P. (2010). Tests for error correlation in the functional linear model. *Journal of the American Statistical Association*, 105(491):1113–1125.

[Gohberg et al., 1990] Gohberg, I., Goldberg, S., and Kaashoek, M. (1990). Classes of linear operators, 1990.

[Guillas et al., 2011] Guillas, S., Bakare, A., Morley, J., and Simons, R. (2011). Functional autoregressive forecasting of long-term seabed evolution. *Journal of Coastal Conservation*, 15(3):337–351.

[Horváth et al., 2010] Horváth, L., Hušková, M., and Kokoszka, P. (2010). Testing the stability of the functional autoregressive process. *Journal of Multivariate Analysis*, 101(2):352–367.

[Horváth and Kokoszka, 2012] Horváth, L. and Kokoszka, P. (2012). *Inference for functional data with applications*, volume 200. Springer Science & Business Media.

[Iacus, 2009] Iacus, S. M. (2009). *Simulation and inference for stochastic differential equations: with R examples*. Springer.

[Kokoszka and Reimherr, 2013] Kokoszka, P. and Reimherr, M. (2013). Determining the order of the functional autoregressive model. *Journal of Time Series Analysis*, 34(1):116–129.

[Kokoszka and Zhang, 2010] Kokoszka, P. and Zhang, X. (2010). Improved estimation of the kernel of the functional autoregressive process. Technical report, Technical Report. Utah State University.

[Kristoufek, 2013] Kristoufek, L. (2013). Bitcoin meets google trends and wikipedia: Quantifying the relationship between phenomena of the internet era. *Scientific reports*, 3.

[Meese and Rogoff, 1983] Meese, R. A. and Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of international economics*, 14(1):3–24.

[Nakamoto, 2008] Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Consulted*, 1(2012):28.

[Nelson and Plosser, 1982] Nelson, C. R. and Plosser, C. R. (1982). Trends and random walks in macroeconmic time series: some evidence and implications. *Journal of monetary economics*, 10(2):139–162.

[Phillips and Perron, 1988] Phillips, P. C. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2):335–346.

[Ramsay et al., 2009] Ramsay, J., Hooker, G., and Graves, S. (2009). *Functional data analysis with R and MATLAB*. Springer Verlag.

[Ramsay and Silverman, 2005] Ramsay, J. and Silverman, B. (2005). *Functional data analysis*. Springer, New York.

Appendix

b1

Error	m	0.5	0.7	0.8	0.9	0.95	0.99	1.00
bm	50	1	0.91	0.62	0.27	0.18	0.06	0.03
	100	1	1	0.98	0.68	0.3	0.16	0.05
	200	1	1	1	1	0.69	0.14	0.04
	500	1	1	1	1	1	0.35	0.05
bb	50	1	0.97	0.78	0.37	0.17	0.06	0.05
	100	1	1	0.99	0.84	0.51	0.15	0.06
	200	1	1	1	0.99	0.89	0.23	0.06
	500	1	1	1	1	1	0.46	0.03

b2

Error	m	0.5	0.7	0.8	0.9	0.95	0.99	1.00
bm	50	1	0.93	0.61	0.3	0.26	0.16	0.11
	100	1	1	0.98	0.71	0.31	0.18	0.08
	200	1	1	1	1	0.63	0.16	0.06
	500	1	1	1	1	1	0.33	0.08
bb	50	1	0.92	0.58	0.31	0.2	0.49	0.34
	100	1	1	1	0.76	0.33	0.24	0.13
	200	1	1	1	1	0.87	0.23	0.13
	500	1	1	1	1	1	0.43	0.05

b4

Error	m	0.5	0.7	0.8	0.9	0.95	0.99	1.00
bm	50	1	0.93	0.59	0.22	0.07	0.02	0.05
	100	1	1	0.99	0.68	0.3	0.04	0.09
	200	1	1	1	0.99	0.63	0.15	0.06
	500	1	1	1	1	1	0.29	0.05
bb	50	1	0.98	0.85	0.37	0.28	0.04	0.05
	100	1	1	1	0.86	0.44	0.13	0.06
	200	1	1	1	1	0.88	0.11	0.09
	500	1	1	1	1	1	0.46	0.02

b5

Error	m	0.5	0.7	0.8	0.9	0.95	0.99	1.00
bm	50	1	0.87	0.57	0.24	0.12	0.13	0.13
	100	1	1	0.99	0.64	0.3	0.11	0.15
	200	1	1	1	0.99	0.63	0.14	0.08
	500	1	1	1	1	1	0.28	0.06
bb	50	1	0.96	0.62	0.17	0.23	0.34	0.31
	100	1	1	1	0.78	0.37	0.23	0.16
	200	1	1	1	1	0.84	0.25	0.11
	500	1	1	1	1	1	0.44	0.03

b7

Error	m	0.5	0.7	0.8	0.9	0.95	0.99	1.00
bm	50	1	0.96	0.73	0.35	0.2	0.06	0.03
	100	1	1	0.98	0.72	0.35	0.17	0.05
	200	1	1	1	1	0.76	0.14	0.04
	500	1	1	1	1	1	0.42	0.06
bb	50	1	0.99	0.83	0.46	0.21	0.07	0.05
	100	1	1	1	0.95	0.56	0.17	0.04
	200	1	1	1	1	0.92	0.25	0.06
	500	1	1	1	1	1	0.51	0.03

b8

Error	m	0.5	0.7	0.8	0.9	0.95	0.99	1.00
bm	50	1	0.96	0.69	0.35	0.27	0.16	0.1
	100	1	1	0.99	0.76	0.39	0.2	0.06
	200	1	1	1	1	0.72	0.18	0.07
	500	1	1	1	1	1	0.37	0.08
bb	50	1	0.97	0.75	0.39	0.25	0.53	0.35
	100	1	1	1	0.8	0.41	0.25	0.13
	200	1	1	1	1	0.93	0.23	0.13
	500	1	1	1	1	1	0.52	0.05

b10

Error	m	0.5	0.7	0.8	0.9	0.95	0.99	1.00
bm	50	1	0.91	0.64	0.27	0.17	0.06	0.03
	100	1	1	0.98	0.66	0.3	0.16	0.05
	200	1	1	1	1	0.69	0.14	0.04
	500	1	1	1	1	1	0.35	0.05
bb	50	1	0.98	0.78	0.38	0.18	0.09	0.06
	100	1	1	0.99	0.84	0.52	0.15	0.06
	200	1	1	1	0.99	0.89	0.22	0.06
	500	1	1	1	1	1	0.47	0.03

b11

Error	m	0.5	0.7	0.8	0.9	0.95	0.99	1.00
bm	50	1	0.93	0.63	0.31	0.26	0.16	0.11
	100	1	1	0.98	0.71	0.31	0.18	0.09
	200	1	1	1	1	0.62	0.16	0.07
	500	1	1	1	1	1	0.34	0.08
bb	50	1	0.92	0.59	0.31	0.19	0.51	0.36
	100	1	1	1	0.76	0.34	0.24	0.13
	200	1	1	1	1	0.89	0.24	0.13
	500	1	1	1	1	1	0.46	0.04

k1b1

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	1	1	0.88	0.72	0.5
	100	1	1	1	0.99	0.94	0.63
	200	1	1	1	1	1	0.75
	500	1	1	1	1	1	0.89
bb	50	1	0.95	0.9	0.65	0.36	0.19
	100	1	1	1	0.91	0.69	0.33
	200	1	1	1	1	0.95	0.35
	500	1	1	1	1	1	0.66

k1b2

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	1	0.98	0.86	0.75	0.64
	100	1	1	1	1	0.9	0.64
	200	1	1	1	1	1	0.75
	500	1	1	1	1	1	0.87
bb	50	1	0.96	0.9	0.64	0.36	0.2
	100	1	1	1	0.92	0.68	0.33
	200	1	1	1	1	0.95	0.36
	500	1	1	1	1	1	0.67

k1b7

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	1	1	0.92	0.78	0.54
	100	1	1	1	1	0.97	0.7
	200	1	1	1	1	1	0.83
	500	1	1	1	1	1	0.94
bb	50	1	0.96	0.94	0.74	0.43	0.2
	100	1	1	1	0.97	0.74	0.4
	200	1	1	1	1	0.97	0.43
	500	1	1	1	1	1	0.72

k2b1

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	0.93	0.88	0.59	0.4	0.22
	100	1	1	1	0.88	0.6	0.28
	200	1	1	1	1	0.93	0.42
	500	1	1	1	1	1	0.6
bb	50	1	0.97	0.98	0.83	0.53	0.37
	100	1	1	1	0.99	0.84	0.46
	200	1	1	1	1	0.97	0.55
	500	1	1	1	1	1	0.8

k2b2

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	0.97	0.81	0.57	0.44	0.31
	100	1	1	1	0.86	0.6	0.35
	200	1	1	1	0.99	0.92	0.41
	500	1	1	1	1	1	0.58
bb	50	1	0.98	0.98	0.82	0.53	0.37
	100	1	1	1	0.99	0.83	0.47
	200	1	1	1	1	0.97	0.54
	500	1	1	1	1	1	0.8

k3b1

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	200	1	1	1	1	1	1
	500	1	1	1	1	1	1
bb	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	200	1	1	1	1	1	1
	500	1	1	1	1	1	1

k3b2

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	200	1	1	1	1	1	1
	500	1	1	1	1	1	1
bb	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	200	1	1	1	1	1	1
	500	1	1	1	1	1	1

k4b1

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	1	1	0.9	0.72	0.48
	100	1	1	1	1	0.97	0.66
	200	1	1	1	1	1	0.87
	500	1	1	1	1	1	0.98
bb	50	1	1	1	0.95	0.83	0.69
	100	1	1	1	1	1	0.81
	200	1	1	1	1	1	0.96
	500	1	1	1	1	1	0.99

k4b2

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	1	0.98	0.87	0.78	0.67
	100	1	1	1	1	0.91	0.63
	200	1	1	1	1	1	0.83
	500	1	1	1	1	1	0.96
bb	50	1	1	1	0.96	0.84	0.68
	100	1	1	1	1	1	0.79
	200	1	1	1	1	1	0.96
	500	1	1	1	1	1	1

k5b1

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	0.93	0.88	0.6	0.42	0.21
	100	1	1	1	0.87	0.6	0.27
	200	1	1	1	1	0.92	0.41
	500	1	1	1	1	1	0.6
bb	50	1	0.97	0.93	0.73	0.44	0.26
	100	1	1	1	0.97	0.75	0.39
	200	1	1	1	1	0.96	0.47
	500	1	1	1	1	1	0.74

k5b2

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	0.96	0.82	0.56	0.45	0.32
	100	1	1	1	0.84	0.61	0.34
	200	1	1	1	0.99	0.91	0.39
	500	1	1	1	1	1	0.6
bb	50	1	0.97	0.91	0.72	0.44	0.27
	100	1	1	1	0.97	0.78	0.41
	200	1	1	1	1	0.96	0.46
	500	1	1	1	1	1	0.74

k6b1

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	200	1	1	1	1	1	1
	500	1	1	1	1	1	1
bb	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	200	1	1	1	1	1	1
	500	1	1	1	1	1	1

k7b1

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	200	1	1	1	1	1	1
	500	1	1	1	1	1	1
bb	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	200	1	1	1	1	1	1
	500	1	1	1	1	1	1

k8b1

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	1	1	1	1	0.98
	100	1	1	1	1	1	0.99
	200	1	1	1	1	1	1
	500	1	1	1	1	1	1
bb	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	200	1	1	1	1	1	1
	500	1	1	1	1	1	1

k9b1

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	1	0.99	0.82	0.67	0.42
	100	1	1	1	0.99	0.91	0.49
	200	1	1	1	1	1	0.66
	500	1	1	1	1	1	0.81
bb	50	1	0.96	0.95	0.75	0.45	0.24
	100	1	1	1	0.97	0.77	0.42
	200	1	1	1	1	0.97	0.46
	500	1	1	1	1	1	0.71

k10b1

Error	m	0.5	0.7	0.8	0.9	0.95	0.99
bm	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	200	1	1	1	1	1	1
	500	1	1	1	1	1	1
bb	50	1	1	1	1	1	1
	100	1	1	1	1	1	1
	200	1	1	1	1	1	1
	500	1	1	1	1	1	1