Model-Free Bootstrap for Markov processes

Li Pan
University of California–San Diego, La Jolla, USA - lipan@ucsd.edu

Dimitris N. Politis*
University of California–San Diego, La Jolla, USA - dpolitis@ucsd.edu

Abstract

Consider time series data $X_1, \ldots, X_n$ that arise as a (partial) sample path of a stationary Markov process of order $p \geq 1$. Under this general context, there are (at least) two resampling mechanisms available in the literature, namely: (a) bootstrap based on kernel estimates of the transition density of the Markov processes of Rajarshi (1990), and the Local Bootstrap for Markov processes of Paparoditis and Politis (2002). In this paper, we introduce a third option, the Model-Free Bootstrap for Markov Processes; this is a novel approach stemming from the Model-Free Prediction Principle of Politis (2013). The three approaches are compared as they apply to the problem of constructing confidence intervals for the conditional expectation function $\mu_y = E(X_t|Y_{t-1} = y)$ for some $y \in \mathbf{R}^p$ where we define $Y_t = (X_t, X_{t-1}, \cdots, X_{t-p+1})'$.

Keywords: Confidence intervals, Local Bootstrap, Model-Free Prediction.

1. Introduction

Bootstrap methods for time series have been the subject of active investigation for the last 25 years; see the recent review by Kreiss and Paparoditis (2011) for the state-of-the-art of the literature. In particular, when the data at hand are a sample from a Markov process, several different resampling schemes have been proposed; see e.g. Bertail and Clémencon (2006) and the references therein.

Consider time series data $X_1, \ldots, X_n$ that arise as a (partial) sample path of a stationary Markov process of order $p \geq 1$. Two popular methods for resampling Markov processes are: (a) the bootstrap based on a nonparametric–e.g. kernel smoothed–estimate of the transition density of Rajarshi (1990), and (b) the Local Bootstrap for Markov processes of Paparoditis and Politis (2002); the latter may be viewed as resampling from a nonparametric estimate of the transition cumulative distribution function as opposed to the transition density.

In this short paper, we introduce a third option, the *Model-Free Bootstrap* for Markov Processes; this is a novel approach stemming from the Model-Free Prediction Principle of Politis (2013). The key idea is to transform a given complex dataset into one that is i.i.d., and therefore easier to handle; this allows for conducting statistical inference in a new way. For instance, prediction intervals for the unobserved X_{n+1} were constructed in this new way by Pan and Politis (2014b).

Denote the conditional distribution and expectation of X_t given $Y_{t-1} = y \in \mathbf{R}^p$ respectively by

$$D_y(x) = P(X_t \leq x|Y_{t-1} = y) \;\text{ and }\; \mu_y = E(X_t|Y_{t-1} = y) \tag{1}$$

where $Y_t = (X_t, X_{t-1}, \cdots, X_{t-p+1})'$. In what follows, we will formulate the Model-Free Bootstrap (MFB) and show how can it be used to construct confidence intervals for parameters associated with the conditional distribution $D_y(\cdot)$. For concreteness, we will focus on the conditional expectation function μ_y as the parameter of interest but the algorithms remain true *verbatim* for other functionals of $D_y(\cdot)$, e.g., the conditional median, the conditional variance, etc.

2. Theoretical transformation

Here, and throughout the rest of the paper, we assume that $X = \{X_t, t = 1, 2, \cdots\}$ is a real-valued, strictly stationary process that is Markov of order p; our observations amount to the stretch $X_1, \ldots, X_n$.

For a natural number $i \leq p$, define the distributions with partial conditioning:

$$D_{y,i}(x) = P(X_t \leq x|Y_{t-1}^{(i)} = y) \tag{2}$$

where $Y_{t-1}^{(i)} = (X_{t-1}, \cdots, X_{t-i})'$ and $y \in \mathbf{R}^i$. In this notation, $D_{y,p}(x) = D_y(x)$, i.e, the distribution with full conditioning defined in eq. (1). We can also denote the unconditional distribution as $D_{y,0}(x) = P(X_t \le x)$ which does not depend on y. Throughout the paper, we will assume that, for any y and i, the function $D_{y,i}(\cdot)$ is continuous and invertible over its support.
A transformation from our Markov(p) dataset $X_1, \cdots, X_n$ to an i.i.d. dataset $\eta_1, \cdots, \eta_n$ can now be constructed as follows. Let

$$\eta_1 = D_{y,0}(X_1);\ \eta_2 = D_{Y_1^{(1)},1}(X_2);\ \eta_3 = D_{Y_2^{(2)},2}(X_3);\ \cdots;\ \eta_p = D_{Y_{p-1}^{(p-1)},p-1}(X_p) \tag{3}$$

$$\text{and } \eta_t = D_{Y_{t-1}}(X_t) \text{ for } t = p+1, p+2, \cdots, n. \tag{4}$$

Note that the transformation from the vector $(X_1, \cdots, X_m)'$ to the vector $(\eta_1, \cdots, \eta_m)'$ is one-to-one and invertible for any natural number m by construction. Now it is not difficult to see that the random variables $\eta_1, \cdots, \eta_n$ are i.i.d. Uniform(0,1). In fact, this is just an application of the Rosenblatt (1952) transformation in the case of Markov(p) sequences; see Pan and Politis (2014b) for details.

3. Estimating the Transformation from Data

To estimate the theoretical transformation from data, we would need to estimate the distributions $D_{y,i}(\cdot)$ for $i = 0, 1, \ldots, p-1$ and $D_y(\cdot)$. Note, however, that $D_{y,i}(\cdot)$ for $i < p$ can–in principle—be computed from $D_y(\cdot)$ since the latter uniquely specifies the whole distribution of the stationary Markov process. Hence, it should be sufficient to just estimate $D_y(\cdot)$ from our data. Another way of seeing this is to note that the p variables in eq. (3) can be considered as 'edge effects' or 'initial conditions'; the crucial part of the transformation is given by eq. (4), i.e., the one based on $D_y(\cdot)$. Assuming $D_y(\cdot)$ is smooth, e.g. continuous, in y, we can estimate $D_y(x)$ and μ_y by local averaging methods such as the kernel estimators

$$\hat{D}_y(x) = \frac{\sum_{i=p+1}^n 1_{\{X_i \le x\}} K(\frac{\|y - Y_{i-1}\|}{h})}{\sum_{k=p+1}^n K(\frac{\|y - Y_{k-1}\|}{h})} \text{ and } \hat{\mu}_y = \frac{\sum_{i=p+1}^n X_i K(\frac{\|y - Y_{i-1}\|}{h})}{\sum_{k=p+1}^n K(\frac{\|y - Y_{k-1}\|}{h})}; \tag{5}$$

here $K(\cdot)$ denotes the kernel, $h > 0$ the bandwidth, 1_A the indicator of set A, and $|| \cdot ||$ some norm on $\mathbf{R}^p$. Note that $\hat{D}_y$ in (5) is a step function; we can use linear interpolation on this step function to produce an estimate $\tilde{D}_y$ that is piecewise linear and strictly increasing (and therefore invertible); see Politis (2013). Estimator $\hat{D}_y$ is consistent for D_y under regularity conditions; see Paparoditis and Politis (2002). Furthermore, the consistency of $\tilde{D}_y$ follows from its close proximity to $\hat{D}_y$. Consequently, we may define

$$u_t = \tilde{D}_{Y_{t-1}}(X_t) \text{ for } t = p+1, \cdots, n. \tag{6}$$

The consistency of $\tilde{D}_y$ implies that $u_t \approx \eta_t$ where η_t was defined in eq. (4); thus, $\{u_t \text{ for } t = p+1, \cdots, n\}$ are approximately i.i.d. Uniform (0,1), Hence, the goal of transforming our data $X_1, \cdots, X_n$ to a sequence of (approximately) i.i.d. random variables u_t has been achieved. Note that the 'initial conditions' $u_1, \ldots, u_p$ were not explicitly generated in the above as they are not needed in the Model-Free Bootstrap algorithms.

4. Basic Algorithm for Confidence Intervals based on Model-Free Bootstrap (MFB)

Fix some $y \in \mathbf{R}^p$; the goal is to construct a $(1-\alpha)100\%$ confidence interval for μ_y. To do this, we will need to approximate the sampling distribution of the *root* $\mu_y - \hat{\mu}_y$ by a bootstrap distribution.

Basic MFB Algorithm

(1) Use eq. (6) to obtain the transformed data $u_{p+1}, \cdots, u_n$.

(2) (a) Resample randomly (with replacement) the transformed data $u_{p+1}, \cdots, u_n$ *to create the pseudo-data* $u^*_{-M}, u^*_{-M+1}, \cdots, u^*_0, u^*_1, \cdots, u^*_{n-1}, u^*_n$ *for some large positive integer* M.

(b) Draw $X^*_{-M}, \cdots, X^*_{-M+p-1}$ *randomly from any consecutive* p *values of the dataset* $(X_1, \cdots, X_n)$.

(c) Generate the bootstrap sample path $X^*_t = \tilde{D}^{-1}_{Y^*_{t-1}}(u^*_t)$ *for* $t = -M+p, \cdots, n$ *where* $Y^*_{t-1} = (X^*_{t-1}, \cdots, X^*_{t-p})'$.

(d) *Calculate the kernel estimator in the bootstrap world:* $\hat{\mu}_y^* = \frac{\sum_{i=p+1}^{n} X_i^* K(\frac{\|y-Y_{i-1}^*\|}{h})}{\sum_{k=p+1}^{n} K(\frac{\|y-Y_{k-1}^*\|}{h})}$.

(e) *Calculate the bootstrap root* $\hat{\mu}_y - \hat{\mu}_y^*$.

(3) *Repeat step (2) B times; the B bootstrap root replicates are collected in the form of an empirical distribution whose* α*-quantile is denoted* $q(\alpha)$.

(4) *The MFB* $(1-\alpha)100\%$ *equal-tailed confidence interval for* μ_y *is given by:* $[\hat{\mu}_y + q(\alpha/2), \hat{\mu}_y + q(1-\alpha/2)]$.

Remark 1. An alternative estimator for μ_y is $\tilde{\mu}_y = \frac{1}{n-p}\sum_{t=p+1}^{n} \tilde{D}_y^{-1}(u_t)$; as discussed by Politis (2013), $\tilde{\mu}_y$ and $\hat{\mu}_y$ are asymptotically equivalent and can be used interchangeably.

Remark 2. Recall that $\hat{D}_y(x)$ is a local average estimator, i.e., averaging the indicator $1_{\{X_i \le x\}}$ over data vectors Y_t that are close to y. If y is outside the range of the data vectors Y_t, then obviously estimator $\hat{D}_y(x)$ can not be constructed, and the same is true for $\tilde{D}_y(x)$. Similarly, if y is at the edges of the range of Y_t, e.g., within h of being outside the range, then $\hat{D}_y(x)$ and $\tilde{D}_y(x)$ will not be very accurate. Step 1 of Algorithm MFB can be modified to drop the u_is that are obtained from an x_i whose y_{i-1} is within h of the boundary.

Remark 3. Since the transformed data $u_{p+1}, \ldots, u_n$ are approximately i.i.d. Uniform(0,1), the resampling in step 2(a) of Algorithm MFB could alternatively be done using the Uniform distribution, i.e., generate $u_{-M}^*, u_{-M+1}^*, \cdots, u_0^*, u_1^*, \cdots, u_{n-1}^*, u_n^*$ as i.i.d. Uniform(0,1). Algorithm MFB still works fine with this choice but can not obviously be extended to include the use of "predictive residuals" as proposed in the next section.

5. Predictive Model-Free Bootstrap

Recall that the conditional distribution of interest is $D_y(x) = P(X_t \le x | Y_{t-1} = y)$ which is estimated by

$$\hat{D}_y(x) = \frac{\sum_{i=p+1}^{n} 1_{\{X_i \le x\}} K(\frac{\|y-Y_{i-1}\|}{h})}{\sum_{i=p+1}^{n} K(\frac{\|y-Y_{i-1}\|}{h})}.$$

Pan and Politis (2014a,b) introduced an alternative estimator. To define it, let the observations $X_1, \ldots, X_n$ take the values $x_1, \ldots, x_n$; similarly, let $Y_{t-1} = y_{t-1}$ for $t = p+1, \cdots, n$. Suppose we are interested in estimating $D_{y_{t-1}}(x_t)$. Obviously, the regression pair (y_{t-1}, x_t) was part of our data when $p < t \le n$. However, to avoid overfitting issues, we can estimate $D_{y_{t-1}}(x_t)$ based on a dataset that excludes the point (y_{t-1}, x_t) from the scatterplot. In other words, define the 'delete-one' estimator

$$\hat{D}_{y_{t-1}}^{(t)}(x_t) = \frac{\sum_{i=p+1, i\ne t}^{n} 1_{\{x_i \le x_t\}} K(\frac{\|y_{t-1}-y_{i-1}\|}{h})}{\sum_{k=p+1, k\ne t} K(\frac{\|y_{t-1}-y_{k-1}\|}{h})}, \text{ for } t = p+1, \cdots, n.$$

Linear interpolation on $\hat{D}_y^{(t)}(x)$ gives $\tilde{D}_y^{(t)}(x)$, and we can then define $u_t^{(t)} = \tilde{D}_{y_{t-1}}^{(t)}(x_t)$; here, the $u_t^{(t)}$ serve as analogs of the *predictive residuals* studied in Pan and Politis (2014a) in a nonparametric regression setup.

Predictive MFB (PMFB) Algorithm

The algorithm is identical to Algorithm MFB after substituting $u_{p+1}^{(p+1)}, \cdots, u_n^{(n)}$ *in place of* $u_{p+1}, \cdots, u_n$.

6. Smoothed Model-Free Bootstrap

In the above, we estimated the transition distribution $D_y(x) = P(X_t \le x | Y_{t-1} = y)$ by $\hat{D}_y(x)$ as defined in eq. (5). Noting that $D_y(x)$ is, by assumption, continuous in x while $\hat{D}_y(x)$ is not, the linearly interpolated, strictly increasing estimator $\tilde{D}_y(x)$ was used instead. However, $\tilde{D}_y(x)$ is piecewice linear, and therefore not differentiable in the argument x. In this section, we employ an alternative estimator of the conditional transition density that is smooth in x. To do this, we substitute the step function $1_{\{x_i \le x\}}$ in eq. (5) with $\Lambda(\frac{x-x_i}{h_0})$ where $\Lambda(\cdot)$ is an absolutely continuous, strictly increasing cumulative distribution function, and h_0 is a positive bandwidth parameter. The new estimator $\bar{D}_y(x)$ is defined by

$$\bar{D}_y(x) = \frac{\sum_{i=p+1}^{n} \Lambda(\frac{x-x_i}{h_0}) K(\frac{\|y-y_{i-1}\|}{h})}{\sum_{k=p+1}^{n} K(\frac{\|y-y_{k-1}\|}{h})}, \tag{7}$$

and the transformed data $\{v_t \text{ for } t = p+1, \cdots, n\}$ can be calculated by

$$v_t = \bar{D}_{y_{t-1}}(x_t). \tag{8}$$

Smoothed MFB (SMFB) Algorithm

(1) Use eq. (8) to obtain the transformed data $v_{p+1}, \cdots, v_n$.

(2) (a) Resample randomly (with replacement) the transformed data $v_{p+1}, \cdots, v_n$ to create the pseudo-data $v^*_{-M}, v^*_{-M+1}, \cdots, v^*_0, v^*_1, \cdots, v^*_{n-1}, v^*_n$ for some large positive integer M.

(b) Draw $X^*_{-M}, \cdots, X^*_{-M+p-1}$ randomly from any consecutive p values of the dataset $(X_1, \cdots, X_n)$.

(c) Generate the bootstrap sample path $X^*_t = \bar{D}^{-1}_{Y^*_{t-1}}(v^*_t)$ for $t = -M+p, \cdots, n$ where $Y^*_{t-1} = (X^*_{t-1}, \cdots, X^*_{t-p})'$.

(d) Calculate the kernel estimator in the bootstrap world: $\hat{\mu}^*_y = \frac{\sum_{i=p+1}^n X^*_i K(\frac{\|y-Y^*_{i-1}\|}{h})}{\sum_{k=p+1}^n K(\frac{\|y-Y^*_{k-1}\|}{h})}$.

(e) Calculate the bootstrap root $\hat{\mu}_y - \hat{\mu}^*_y$.

(3) Repeat step (2) B times; the B bootstrap root replicates are collected in the form of an empirical distribution whose α-quantile is denoted $q(\alpha)$.

(4) The SMFB $(1-\alpha)100\%$ equal-tailed confidence interval for μ_y is given by: $[\hat{\mu}_y + q(\alpha/2), \hat{\mu}_y + q(1-\alpha/2)]$.

Remark 4. As in Remark 2, Step 2 (a) of Algorithm SMFB can be modified to drop the v_is that are obtained from an X_is whose Y_{i-1} is within h of the boundary.

Remark 5. Pan and Politis (2014b) make the case that smoothing is advantageous; in order to take full advantage though, the two bandwidths must be chosen appropriately. Pan and Politis (2014b) recommend the choice $h_0 \sim h^2$ where h can be chosen through cross-validation.

As in Section 5, we can also use the delete-x_t estimator

$$\bar{D}^{(t)}_y(x_t) = \frac{\sum_{i=p+1, i\neq t}^n \Lambda(\frac{x_t - x_i}{h_0}) K(\frac{\|y_{t-1} - y_{i-1}\|}{h})}{\sum_{k=p+1, k\neq t} K(\frac{\|y_{t-1} - y_{k-1}\|}{h})} \text{ for } t = p+1, \cdots, n$$

in order to construct the transformed data: $v^{(t)}_t = \bar{D}^{(t)}_{y_{t-1}}(x_t)$ for $t = p+1, \cdots, n$; this leads to the Predictive Smoothed Model-Free Bootstrap.

Predictive Smoothed MFB (PSMFB) Algorithm

The algorithm is identical to Algorithm SMFB after substituting $v^{(p+1)}_{p+1}, \cdots, v^{(n)}_n$ in place of $v_{p+1}, \cdots, v_n$.

Remark 6. The above four Model-Free Bootstrap methods were discussed in terms of the concrete application of constructing a confidence interval for μ_y pointwise, i.e., for a given y. However, constructing simultaneous confidence intervals for $\{\mu_y \text{ with } y \in S\}$ for any finite set S is immediate as with any bootstrap method; see e.g. Wolf and Wunderli (2015). If the set S consists of points on a fine grid that span an interval, say $[a_1, a_2]$, then the assumed smoothness of μ_y can be used to turn the aforementioned simultaneous confidence intervals into a confidence band for $\{\mu_y \text{ with } y \in [a_1, a_2]\}$.

7. Finite-Sample Simulations

Table 1 shows empirical confidence interval coverages for different values of y using all the Model-Free Bootstrap methods, Local Bootstrap (LB) and Rajarshi's method (RAJ). 500 Markovian data sets with $n = 200$ were generated using each of the following five models with errors ϵ_t that were i.i.d. standard normal or Laplace.

- model 1: $X_{t+1} = \sin(X_t) + \epsilon_{t+1}$
- model 2: $X_{t+1} = 0.8\log(3X_t^2 + 1) + \epsilon_{t+1}$

Table 1: Empirical coverage level of (nominally) 95% equal-tailed confidence intervals.

	Normal Errors						Laplace errors					
model 1												
y	MFB	PMFB	SMFB	PSMFB	LB	RAJ	MFB	PMFB	SMFB	PSMFB	LB	RAJ
$-2\pi/3$	0.846	0.870	0.854	0.870	0.862	**0.880**	0.842	0.858	0.848	**0.870**	**0.870**	0.866
$-\pi/2$	0.898	0.910	0.898	0.910	0.910	**0.924**	0.870	**0.900**	0.872	0.898	0.894	0.892
$-\pi/6$	0.904	0.932	0.908	0.928	0.922	**0.936**	0.900	0.910	0.900	0.914	**0.924**	0.920
0	0.922	0.928	0.926	0.932	**0.938**	0.926	0.908	0.938	0.914	**0.934**	0.926	**0.934**
$\pi/6$	0.904	0.904	0.904	0.912	0.920	**0.924**	0.924	0.940	0.932	**0.952**	0.942	0.938
$\pi/2$	0.870	0.886	0.868	0.882	0.870	**0.892**	0.860	**0.904**	0.864	**0.904**	0.876	0.882
$2\pi/3$	0.838	0.850	0.844	0.858	0.846	**0.884**	0.862	0.884	0.870	0.888	0.874	**0.900**
model 2												
-0.5	0.842	0.866	0.844	0.858	0.860	**0.896**	0.822	0.842	0.822	0.842	0.850	**0.890**
0	0.880	0.896	0.868	0.896	0.892	**0.930**	0.868	0.888	0.864	0.890	0.886	**0.928**
0.5	0.906	0.922	0.904	0.922	0.910	**0.934**	0.884	0.916	0.880	0.908	0.916	**0.926**
1	0.922	0.934	0.922	0.936	0.928	**0.942**	0.924	0.948	0.916	**0.950**	0.940	0.942
1.5	0.940	0.958	0.942	0.954	**0.948**	**0.948**	0.910	0.936	0.906	**0.946**	0.928	0.944
2.5	0.860	0.880	0.868	0.892	0.878	**0.914**	0.908	0.930	0.906	0.930	0.922	**0.948**
3.5	0.864	0.894	0.872	0.890	**0.896**	0.908	0.850	0.896	0.848	0.894	0.878	**0.906**
model 3												
-2	0.818	0.834	0.832	**0.844**	0.814	0.842	0.842	0.846	0.846	**0.862**	0.820	0.830
-1.5	0.858	0.874	0.870	0.884	0.868	**0.886**	0.814	0.836	0.822	0.852	0.832	**0.866**
-1	0.860	0.880	0.884	**0.900**	0.878	0.894	0.834	0.860	0.862	0.888	0.870	**0.900**
-0.5	0.880	0.890	0.898	**0.906**	0.904	0.896	0.854	0.870	0.880	**0.912**	0.898	0.906
0	0.872	0.874	0.882	0.896	**0.900**	0.890	0.882	0.900	0.898	**0.922**	0.916	0.902
0.5	0.872	0.872	0.888	0.898	0.898	**0.908**	0.890	0.898	0.910	**0.924**	0.908	0.906
1	0.898	0.910	0.910	**0.916**	0.898	0.888	0.858	0.896	0.880	**0.918**	0.882	0.880
1.5	0.876	0.894	0.888	**0.904**	0.892	0.872	0.836	0.866	0.848	**0.882**	0.858	0.868
2	0.834	0.838	0.844	**0.846**	0.834	0.838	0.818	**0.852**	0.820	**0.852**	0.806	0.820
model 4												
$-2\pi/3$	0.786	0.806	0.794	0.806	0.798	**0.856**	0.806	0.826	0.804	0.826	0.822	**0.856**
$-\pi/2$	0.878	0.890	0.876	0.896	0.890	**0.906**	0.836	0.872	0.842	**0.880**	0.832	0.870
$-\pi/6$	0.894	0.906	0.896	0.906	0.902	**0.920**	0.882	0.916	0.878	0.906	0.886	**0.934**
0	0.944	**0.946**	0.940	0.942	0.954	0.944	0.930	0.948	0.918	**0.946**	0.944	0.936
$\pi/6$	0.904	0.918	0.896	0.922	0.910	**0.946**	0.926	0.940	0.944	**0.950**	0.916	0.944
$\pi/2$	0.864	0.880	0.872	0.884	0.876	**0.892**	0.840	0.880	0.842	0.888	0.844	**0.890**
$2\pi/3$	0.808	0.828	0.806	0.836	0.824	**0.856**	0.814	0.850	0.826	0.848	0.830	**0.872**
model 5												
-2	0.924	**0.938**	0.928	0.928	0.928	0.962	0.900	0.916	0.908	0.922	0.908	**0.930**
-1.5	0.936	0.942	0.940	**0.956**	0.958	0.966	0.890	0.924	0.896	0.930	0.916	**0.936**
-1	0.934	0.938	0.936	0.940	**0.948**	0.960	0.918	**0.952**	0.924	**0.952**	0.938	0.946
-0.5	0.930	**0.948**	0.934	0.954	0.946	0.954	0.904	0.938	0.908	0.938	0.932	**0.946**
0	0.900	0.932	0.900	**0.934**	0.924	0.932	0.910	0.938	0.914	0.940	0.924	**0.944**
0.5	0.896	0.916	0.896	0.916	**0.926**	0.912	0.900	0.922	0.902	0.926	0.914	**0.932**
1	0.888	0.912	0.886	0.902	0.902	**0.930**	0.908	0.936	0.906	**0.942**	0.914	0.932
1.5	0.888	0.910	0.890	**0.914**	0.904	**0.914**	0.872	0.918	0.868	0.922	0.908	**0.934**
2	0.888	0.912	0.890	0.908	0.886	**0.920**	0.884	0.914	0.874	**0.922**	0.904	0.912

- model 3: $X_{t+1} = -0.5\exp(-50X_t^2)Xt + \epsilon_{t+1}$
- model 4: $X_{t+1} = \sin(X_t) + \sqrt{0.5 + 0.25X_t^2}\epsilon_{t+1}$
- model 5: $X_{t+1} = 0.75X_t + 0.15X_t\epsilon_{t+1} + \epsilon_{t+1}$.

For the Local Bootstrap and the Model-Free methods, the bandwidth h was chosen via cross-validation; for SMFB and PSMFB, the extra bandwidth was taken to be $h_0 = h^2$ as in Remark 5. The kernels K and Λ had a normal shape. For Rajarshi's method the rule of thumb formula $h = 0.9\min(SD, IQ/1.34)n^{-1/4}$ was used where SD and IQ denote the sample standard deviation and IQ–range of the data; see Pan and Politis (2014b) for a discussion. The number of bootstrap repetitions was $B = 500$.
Some conclusions are as follows:
—Going from MFB (respectively SMFB) to PMFB (respectively PSMFB) one obtains better coverage level but larger variability of interval length (not shown).
—Smoothed model free methods have better coverage and less variability than their unsmoothed counterparts.
—The Local Bootstrap has better coverage than MFB but not as good as the one from PMFB.
—Among the four Model-Free methods, PSMFB has the best coverage levels.
—Rajarshi's method is the best with respect to coverage level for models with normal errors with the PSMFB coming a close second; this may be due in part to the different ways the bandwidth was chosen. In models with Laplace errors, PSMFB and Rajarshi's method have comparable coverage.
—In cases where PSMFB and Rajarshi's method led to similar coverage levels, it was found that PSMFB intervals had smaller (average) interval length which is highly desirable.

References

Bertail, P. and Clémencon, S. (2006). Regenerative block bootstrap for Markov chains. Bernoulli, vol. 12, no. 4, pp. 689-712.

Franke, J., Kreiss, J.-P. and Mammen, E. (2002). Bootstrap of kernel smoothing in nonlinear time series. Bernoulli, vol. 8, no. 1, pp. 1–37.

Kreiss, J.-P. and Paparoditis, E. (2011). Bootstrap methods for dependent data: a review. Journal of the Korean Statistical Society, vol. 40, no. 4, pp. 357-378.

Pan, Li and Politis, D.N. (2014a). Bootstrap prediction intervals for linear, nonlinear, and nonparametric autoregressions (with discussion). Journal of Statistical Planning and Inference.

Pan, Li and Politis, D.N. (2014b). Bootstrap prediction intervals for Markov processes, Working paper, Department of Economics, University of California.

Paparoditis, E. and Politis, D.N. (2002). The local bootstrap for Markov processes. Journal of Statistical Planning and Inference, vol. 108, no. 1, pp. 301–328.

Politis, D.N. (2013), Model-free model-fitting and predictive distributions (with discussion). Test, vol. 22, no. 2, pp. 183-250.

Rajarshi, M.B. (1990). Bootstrap in Markov-sequences based on estimates of transition density. Annals of the Institute of Statistical Mathematics, vol. 42, no. 2, pp. 253-268.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. Annals of Mathematical Statistics, pp. 470–472.

Wolf, M. and Wunderli, D. (2015). Bootstrap joint prediction regions. Journal of Time Series Analysis, (forthcoming).