TWO-STAGE IDENTIFICATION OF LOCALLY STATIONARY AUTOREGRESSIVE PROCESSES AND ITS APPLICATION TO THE PARAMETRIC SPECTRUM ESTIMATION

Maciej Niedźwiecki and Marcin Ciołek

Faculty of Electronics, Telecommunications and Computer Science, Department of Automatic Control
Gdańsk University of Technology, Narutowicza 11/12, 80-233 Gdańsk , Poland
maciekn@eti.pg.gda.pl, marcin.ciolek@pg.gda.pl

ABSTRACT

The problem of identification of a nonstationary autoregressive process with unknown, and possibly time-varying, rate of parameter changes, is considered and solved using the parallel estimation approach. The proposed two-stage estimation scheme, which combines the local estimation approach with the basis function one, offers both quantitative and qualitative improvements compared with the currently used single-stage methods.

Index Terms— Identification of nonstationary processes, selection of estimation bandwidth, parametric spectrum estimation.

1. INTRODUCTION

Due to mathematical simplicity and good approximation properties time-varying autoregressive (AR) models have found their way to a large number of practical applications, such as signal prediction [1], [2], elimination of impulsive disturbances [3], [4], analysis of biomedical signals [5], [6], [7], [8], [9], [10], equalization of telecommunication channels [11], [12], and spectrum estimation [13], [14], [15], among many others. To take full advantage of the AR model based technique, one needs good identification algorithms, capable of estimating, with sufficient accuracy, time-varying model parameters.

The existing approaches to identification of nonstationary processes can be broadly divided into local estimation methods (based on the assumption that the analyzed process can be regarded as stationary in sufficiently short time intervals), Kalman filter methods (based on explicit stochastic models of parameter variation), and basis function methods (based on deterministic models of parameter changes). A comparative analysis of different estimation techniques mentioned above can be found e.g. in [16] and [17].

In this paper, extending results presented in [18], we will describe a new approach to identification of nonstationary AR processes, which combines the local estimation method, incorporated at the first stage of identification, with the basis function method, used at the second stage to smooth local parameter estimates in a way that takes into account their accuracy. We will show that such a combined method yields better results than the component methods.

This work was partially supported by the National Science Center under the agreement UMO-2015/17/B/ST7/03772. Calculations were carried out at the Academic Computer Centre in Gdańsk.

2. LOCALLY STATIONARY AUTOREGRESSIVE PROCESS

Consider a nonstationary autoregressive process $\{y(t)\}$ governed by

$$y(t) = \sum_{l=1}^{n} a_l(t)y(t-l) + e(t) = \boldsymbol{\alpha}^{\mathrm{T}}(t)\boldsymbol{\varphi}(t) + e(t) \qquad (1)$$

where $t = \ldots, -1, 0, 1, \ldots$ denotes normalized (dimensionless) discrete time, $\boldsymbol{\alpha}(t) = [a_1(t), \ldots, a_n(t)]^{\mathrm{T}}$ denotes the vector of time-varying autoregressive coefficients, $\boldsymbol{\varphi}(t) = [y(t-1), \ldots, y(t-n)]^{\mathrm{T}}$ denotes regression vector, and $\{e(t)\}$ is the sequence of independent Gaussian variables with zero mean and time-varying variance $\rho(t)$.

When AR coefficients in (1) vary smoothly with time, and the forming filter $1/A[z^{-1}, \boldsymbol{\alpha}(t)]$, where

$$A[z^{-1}, \boldsymbol{\alpha}(t)] = 1 - \sum_{l=1}^{n} a_l(t)z^{-l}$$

is uniformly stable in the analyzed time interval, the signal $\{y(t)\}$ belongs to the class of locally stationary processes with well-defined evolutionary spectral representation [19], [20], [21]. Under such conditions (see [20] for more technical details) the time-varying spectral density function of $\{y(t)\}$ can be uniquely defined as

$$S(\omega, t) = \frac{\rho(t)}{|A[e^{j\omega}, \boldsymbol{\alpha}(t)]|^2} \qquad (2)$$

where $\omega \in (-\pi, \pi]$ denotes the normalized angular frequency.

3. LOCAL ESTIMATION APPROACH

3.1. Weighted Yule-Walker estimates

Locally stationary processes can be identified using local estimation techniques, namely the estimates of $\boldsymbol{\alpha}(t)$ and $\rho(t)$ can be obtained by fitting a time-invariant model to a fixed-length data segment $\{y(t-k), \ldots, y(t+k)\}$ of width $2k+1$, centered at t. The integer number k, which determines the size of local analysis window $T_k(t) = [t-k, t+k]$, will be further called an estimation bandwidth coefficient. The name stems from the fact that k controls the estimation bandwidth of the local estimation algorithm, i.e., the frequency range in which signal parameters can be tracked "successfully" (the larger k the smaller the bandwidth) [17].

Following [19] and [18], we will use for local estimation purposes the weighted Yule-Walker (WYW) scheme. The corresponding estimates can be expressed in the form

$$\widehat{\boldsymbol{\alpha}}_k(t) = [\widehat{a}_{1|k}(t), \ldots, \widehat{a}_{n|k}(t)]^{\mathrm{T}} = \widehat{\mathbf{R}}_k^{-1}(t)\widehat{\mathbf{r}}_k(t) \qquad (3)$$

$$\widehat{\rho}_k(t) = \widehat{r}_{0|k}(t) - \widehat{\mathbf{r}}_k^{\mathrm{T}}(t)\widehat{\boldsymbol{\alpha}}_k(t) \qquad (4)$$

where

$$\widehat{\mathbf{R}}_k(t) = \begin{bmatrix} \widehat{r}_{0|k}(t) & \dots & \widehat{r}_{n-1|k}(t) \\ \vdots & \ddots & \vdots \\ \widehat{r}_{n-1|k}(t) & \dots & \widehat{r}_{0|k}(t) \end{bmatrix}$$

$$\widehat{\mathbf{r}}_k(t) = \begin{bmatrix} \widehat{r}_{1|k}(t) & \dots & \widehat{r}_{n|k}(t) \end{bmatrix}^{\mathrm{T}},$$

$\widehat{r}_{l|k}(t), l = 0, \dots, n$ denote the local estimates of the autocorrelation coefficients

$$\widehat{r}_{l|k}(t) = \frac{p_{l|k}(t)}{L_k}, \quad L_k = \sum_{i=-k}^{k} w_k^2(i)$$

$$p_{l|k}(t) = \sum_{i=-k+l}^{k} y_k(t+i|t) y_k(t+i-l|t),$$

and $y_k(t+i|t) = w_k(i)y(t+i)$, $i = -k, \dots, k$, denotes the weighted (tapered) data sequence. We will assume that data taper $w_k(i)$, $w_k(0) = 1$, is a bell-shaped, nonnegative sequence defined on the interval $[-k, k]$. According to [18], there is a strong empirical evidence in favor of the cosinusoidal taper $w_k(i) = \cos\{\pi i/[2(k+1)]\}$, which not only offers a good tradeoff between the bias and variance components of the mean-squared parameter tracking errors, but also allows for time-recursive computation of $p_{l|k}(t)$. The estimates $\widehat{\boldsymbol{\alpha}}_k(t)$ and $\widehat{\rho}_k(t)$ can be computed using the Levinson-Durbin algorithm [22].

Based on the local identification results, the parametric spectrum estimator can be expressed in the form

$$\widehat{S}_k(\omega, t) = \frac{\widehat{\rho}_k(t)}{|A[e^{j\omega}, \widehat{\boldsymbol{\alpha}}_k(t)]|^2}. \tag{5}$$

Parametric spectrum estimators, such as (5), have some well-known advantages over the nonparametric (Fourier transform based) solutions – they offer higher estimation accuracy (especially for small data sets) and yield results which, due to elimination of spurious spectral peaks, are usually easier to interpret [23], [24].

Remark

The choice of the particular estimation method summarized above was motivated by the fact that the WYW estimators minimize the so-called Whittle likelihood, i.e., the Kullback-Leibler divergence between the parametric AR spectrum and the nonparametric weighted periodogram one [19]. It is therefore very well suited to the task of spectrum estimation.

3.2. Bandwidth-adaptive solution

If the rate of signal nonstationarity is itself time-varying, it may be difficult to find a single value of bandwidth parameter that would guarantee good parameter tracking at all instants of time. To overcome this limitation, one can simultaneously run several local estimation algorithms equipped with different bandwidth parameters $k_i, i = 1, \dots, K$. As shown in [18], the reasonable estimate of the most appropriate value of k at the instant t can be obtained from

$$\widehat{k}(t) = \arg\min_{k \in \mathcal{K}} \mathrm{FPE}_k(t) \tag{6}$$

where $\mathcal{K} = \{k_1, \dots, k_K\}$ and $\mathrm{FPE}_k(t)$ denotes the suitably modified Akaike's FPE statistic [25], [26]

$$\mathrm{FPE}_k(t) = \frac{1 + \frac{n}{N_k}}{1 - \frac{n}{N_k}} \widehat{\rho}_k(t), \quad N_k = \frac{[\sum_{i=-k}^{k} w_k^2(i)]^2}{\sum_{i=-k}^{k} w_k^4(i)}. \tag{7}$$

The corresponding estimate of the instantaneous spectrum can be obtained in the form

$$\widehat{S}(\omega, t) = \frac{\widehat{\rho}_{\widehat{k}(t)}(t)}{|A[e^{j\omega}, \widehat{\boldsymbol{\alpha}}_{\widehat{k}(t)}(t)]|^2}. \tag{8}$$

The parallel estimation scheme described above is robust to unknown and/or changing rate of parameter variation and often outperforms, in terms of spectral estimation errors, all component fixed-bandwidth algorithms.

4. TWO-STAGE IDENTIFICATION PROCEDURE

4.1. Local smoothing of weighted Yule-Walker estimates

When the local estimation technique is used, the estimated parameter trajectories and consequently also the obtained evolutionary spectra, are not smooth functions of time – even if true process parameters change in a smooth manner (as assumed here). In parallel estimation schemes this ruggedness effect may be further strengthened due to bandwidth "switching". In applications that require qualitative (e.g. visual) evaluation of spectral estimation results such a lack of smoothness may be considered a drawback. The way out of difficulty, proposed in this paper, is local smoothing of the results obtained at the first stage of identification, described in the previous section.

Denote by $T_0 = [t_0 - M, t_0 + M]$ the local approximation interval of length $2M + 1$, centered at the instant t_0. To smooth parameter estimates $\widehat{\boldsymbol{\alpha}}_{\widehat{k}(t)}(t)$, $\widehat{\rho}_{\widehat{k}(t)}(t)$, $t \in T_0$, we will apply the following approximations

$$\widehat{a}_{l|\widehat{k}(t)}(t) \cong \mathbf{f}^{\mathrm{T}}(t - t_0)\boldsymbol{\eta}_l(t_0), \;\; t \in T_0, \;\; l = 1, \dots, n \tag{9}$$

$$\widehat{\rho}_{\widehat{k}(t)}(t) \cong \mathbf{f}^{\mathrm{T}}(t - t_0)\boldsymbol{\gamma}(t_0), \;\; t \in T_0 \tag{10}$$

where $\mathbf{f}(t) = [f_1(t), \dots, f_m(t)]^{\mathrm{T}}$ is the vector made up of m known, linearly independent discrete-time functions defined on T_0, further referred to as basis functions (BF), and $\boldsymbol{\eta}_l(t_0), l = 1, \dots, n$, $\boldsymbol{\gamma}(t_0)$ denote m-dimensional vectors of approximating coefficients. The frequently chosen set of basis functions consists of powers of time $[f_i(t) = t^{i-1}, i = 1, \dots, m]$, which corresponds to the Taylor series (polynomial) approximation of the smoothed sequence.

Note that (9) can be rewritten in a more compact form as

$$\widehat{\boldsymbol{\alpha}}_{\widehat{k}(t)}(t) \cong \mathbf{F}^{\mathrm{T}}(t - t_0)\boldsymbol{\beta}(t_0)$$

where

$$\mathbf{F}(t) = \begin{bmatrix} \mathbf{f}(t) & & 0 \\ & \ddots & \\ 0 & & \mathbf{f}(t) \end{bmatrix}_{nm \times n} \quad \boldsymbol{\beta}(t_0) = \begin{bmatrix} \boldsymbol{\eta}_1(t_0) \\ \vdots \\ \boldsymbol{\eta}_n(t_0) \end{bmatrix}_{nm \times 1}$$

In the stationary case, where $\boldsymbol{\alpha}(t) = \boldsymbol{\alpha}$, $\rho(t) = \rho$, $\forall t$, it holds that [27]

$$\mathrm{cov}[\widehat{\boldsymbol{\alpha}}_k(t)] \cong \frac{\rho \mathbf{R}^{-1}}{N_k} = \mathbf{Q}_k, \quad \mathrm{var}[\widehat{\rho}_k(t)] \cong \frac{2\rho^2}{N_k} = q_k$$

where $\mathbf{R} = \mathrm{E}[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^T(t)]$ denotes covariance matrix of the regression vector. Therefore, to account for accuracy of the local estimators that are combined, we will minimize the weighted least squares measure of fit

$$\widehat{\boldsymbol{\beta}}(t_0) = \arg\min_{\boldsymbol{\beta}} \sum_{t \in T_0} \| \widehat{\boldsymbol{\alpha}}_{\widehat{k}(t)}(t) - \mathbf{F}^{\mathrm{T}}(t - t_0)\boldsymbol{\beta} \|^2_{\widehat{\mathbf{Q}}^{-1}_{\widehat{k}(t)}(t)} \tag{11}$$

$$\widehat{\boldsymbol{\gamma}}(t_0) = \arg\min_{\boldsymbol{\gamma}} \sum_{t\in T_0} \frac{[\widehat{\rho}_{\widehat{k}(t)}(t) - \mathbf{f}^{\mathrm{T}}(t-t_0)\boldsymbol{\gamma}]^2}{\widehat{q}_{\widehat{k}(t)}(t)} \tag{12}$$

where

$$\widehat{\mathbf{Q}}_k(t) = \frac{\widehat{\rho}_k(t)\widehat{\mathbf{R}}_k^{-1}(t)}{N_k}\ , \quad \widehat{q}_k(t) = \frac{2\widehat{\rho}_k^2(t)}{N_k}\ .$$

Since $\widehat{\mathbf{R}}_k(t)\widehat{\boldsymbol{\alpha}}_k(t) = \widehat{\mathbf{r}}_k(t), \forall k, t$ [cf. (3)], one arrives at

$$\begin{aligned}\widehat{\boldsymbol{\beta}}(t_0) &= \left[\sum_{t\in T_0} \mathbf{F}(t-t_0)\ \frac{N_{\widehat{k}(t)}\widehat{\mathbf{R}}_{\widehat{k}(t)}(t)}{\widehat{\rho}_{\widehat{k}(t)}(t)}\ \mathbf{F}^{\mathrm{T}}(t-t_0)\right]^{-1} \\ &\quad \times \left[\sum_{t\in T_0} \mathbf{F}(t-t_0)\ \frac{N_{\widehat{k}(t)}\widehat{\mathbf{r}}_{\widehat{k}(t)}(t)}{\widehat{\rho}_{\widehat{k}(t)}(t)}\right] \\ &= \left\{\sum_{t\in T_0} \frac{N_{\widehat{k}(t)}\widehat{\mathbf{R}}_{\widehat{k}(t)}(t)}{\widehat{\rho}_{\widehat{k}(t)}(t)} \otimes [\mathbf{f}(t-t_0)\mathbf{f}^{\mathrm{T}}(t-t_0)]\right\}^{-1} \\ &\quad \times \left[\sum_{t\in T_0} \frac{N_{\widehat{k}(t)}\widehat{\mathbf{r}}_{\widehat{k}(t)}(t)}{\widehat{\rho}_{\widehat{k}(t)}(t)} \otimes \mathbf{f}(t-t_0)\right]\end{aligned} \tag{13}$$

where $\otimes$ denotes Kronecker product of the respective vectors/matrices. The corresponding expression for $\widehat{\boldsymbol{\gamma}}(t_0)$ takes the form

$$\begin{aligned}\widehat{\boldsymbol{\gamma}}(t_0) &= \left[\sum_{t\in T_0} \frac{N_{\widehat{k}(t)}\mathbf{f}(t-t_0)\mathbf{f}^{\mathrm{T}}(t-t_0)}{\widehat{\rho}^2_{\widehat{k}(t)}(t)}\right]^{-1} \\ &\quad \times \left[\sum_{t\in T_0} \frac{N_{\widehat{k}(t)}\mathbf{f}(t-t_0)}{\widehat{\rho}_{\widehat{k}(t)}(t)}\right]\ .\end{aligned} \tag{14}$$

Once the approximation coefficients are determined using (13) and (14), the smoothed parameter trajectories can be obtained from

$$\widetilde{\boldsymbol{\alpha}}(t) = \mathbf{F}^{\mathrm{T}}(t-t_0)\widehat{\boldsymbol{\beta}}(t_0),\ \widetilde{\rho}(t) = \mathbf{f}^{\mathrm{T}}(t-t_0)\widehat{\boldsymbol{\gamma}}(t_0) \quad t\in T_0 \tag{15}$$

which results in the following estimate of $S(\omega, t)$

$$\widetilde{S}(\omega, t) = \frac{\widetilde{\rho}(t)}{|A[e^{j\omega}, \widetilde{\boldsymbol{\alpha}}(t)]|^2}\ . \tag{16}$$

Remark 1

We note that in the special case where ρ is unknown but constant, and k is fixed, one can adopt $\widehat{\mathbf{Q}}_k(t) = \widehat{\rho}_k(t_0)\widehat{\mathbf{R}}_k^{-1}(t)/N_k$, leading to

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_k(t_0) &= \left\{\sum_{t\in T_0} \widehat{\mathbf{R}}_k(t) \otimes [\mathbf{f}(t-t_0)\mathbf{f}^{\mathrm{T}}(t-t_0)]\right\}^{-1} \\ &\quad \times \left[\sum_{t\in T_0} \widehat{\mathbf{r}}_k(t) \otimes \mathbf{f}(t-t_0)\right]\end{aligned} \tag{17}$$

which is identical with the approximation formula derived by Dahlhaus [19]. The corresponding estimates of $\widehat{\boldsymbol{\alpha}}_k(t)$ and $\widehat{\rho}_k(t_0)$ can be obtained from

$$\begin{aligned}\widehat{\boldsymbol{\alpha}}_k(t) &= \mathbf{F}^{\mathrm{T}}(t-t_0)\widehat{\boldsymbol{\beta}}_k(t_0),\ \ t\in T_0 \\ \widehat{\rho}_k(t_0) &= \frac{1}{2M+1}\sum_{t\in T_0}[y(t) - \widehat{\boldsymbol{\alpha}}_k^{\mathrm{T}}(t)\boldsymbol{\varphi}(t)]^2.\end{aligned} \tag{18}$$

Remark 2

The proposed postfiltering scheme differs from the result of direct application of the method of basis functions [28] – [35]. In this approach process parameters are modeled as linear combinations of the functions $f_1(t), \ldots, f_m(t)$, i.e., it is assumed that

$$\boldsymbol{\alpha}(t) = \mathbf{F}^{\mathrm{T}}(t-t_0)\boldsymbol{\beta}(t_0), \quad t\in T_0$$

which leads to

$$y(t) = \boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\beta}(t_0) + e(t) \tag{19}$$

where $\boldsymbol{\psi}(t) = \boldsymbol{\varphi}(t) \otimes \mathbf{f}(t-t_0)$.
Based on (19), the estimate of $\boldsymbol{\beta}(t_0)$ can be obtained in the form

$$\begin{aligned}\widehat{\boldsymbol{\beta}}(t_0) &= \arg\min_{\boldsymbol{\beta}} \sum_{t\in T_0}[y(t) - \boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\beta}]^2 \\ &= \left\{\sum_{t\in T_0}[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)] \otimes [\mathbf{f}(t-t_0)\mathbf{f}^{\mathrm{T}}(t-t_0)]\right\}^{-1} \\ &\quad \times \left\{\sum_{t\in T_0}[y(t)\boldsymbol{\varphi}(t)] \otimes \mathbf{f}(t-t_0)\right\}.\end{aligned} \tag{20}$$

Since the quantities $\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)$ and $y(t)\boldsymbol{\varphi}(t)$ can be viewed as pointwise "pre-estimates" of $\mathbf{R}(t)$ and $\mathbf{r}(t)$, respectively, the formula (20) bears some resemblance to (17).

4.2. Clipping

It is known that accuracy of the basis function approximations generally decreases towards both ends of the approximation interval $T_0 = [t_0 - M, t_0 + M]$ [33]. This effect is quite understandable, since while in the middle of T_0 approximation is based on both "past" and "future" data samples, at the instants close to $t_0 + M$ and $t_0 - M$ only the "past" or "future" data are available, respectively. When identification is carried out for a finite impulse response (FIR) system subject to a stationary excitation, such a behavior can be explained analytically using the concept of a frequency response associated with the BF algorithm [17].

Excessive estimation errors, observed at both ends of T_0, can be avoided if the BF estimates are computed for the entire approximation interval, but used in a smaller sub-interval $T_0^* = [t_0 - M + l, t_0 + M - l]$, where $1 \le l \le M$. Such an approach was advocated e.g. in [19] and [36]. To apply "clipping" technique described above, one should work with partially overlapping approximation intervals, i.e., instead of setting $t_0 = M + 1 + 2Mi,\ i = 0, 1, \ldots$, one should use the following rule $t_0 = M + 1 + 2(M-l)i,\ i = 0, 1, \ldots$.

Finally, we note that whenever a decision delay of $k_{\max} + M$ sampling intervals is tolerable, where $k_{\max} = \max\{k \in \mathcal{K}\}$, the proposed two-stage identification method can be operated in the so-called nearly real-time mode.

5. SIMULATION RESULTS

The time-varying AR model of order $n = 8$, used for simulation purposes, had 4 pairs of complex-conjugate poles, evenly spread in terms of their radial location and with the same magnitude equal to 0.995 – see Fig. 1. As symbolically depicted in Fig. 2, this fixed pole constellation was moved (rotated), with a constant speed, from the low-frequency position A, to the high-frequency position B (A-B) and *vice versa* (B-A). Additionally, there were 3 periods of time invariance (A-A, B-B, A-A).

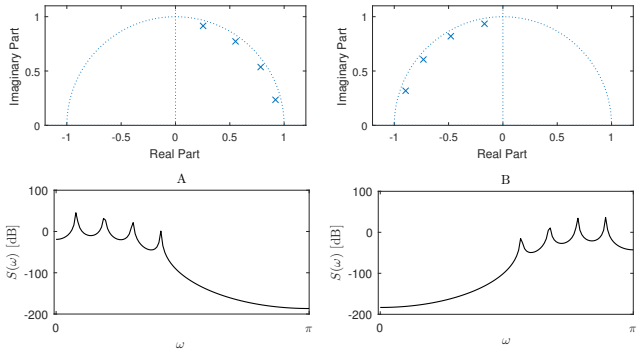

Fig. 1. Two terminal constellations of AR model poles and the corresponding time-invariant spectra.

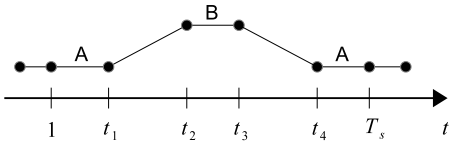

Fig. 2. Simulation scenario.

To check performance of the compared methods under different rates of signal nonstationarity, 3 different values of simulation time T_s were considered (56000, 28000 and 14000), resulting in 3 different speeds of parameter variation (SoV): S_1 (slow), S_2 (medium) and S_3 (fast), respectively. To guarantee that estimation can be always started at the instant 1 and ended at the instant T_s, data generation was started 1000 instants prior to $t = 1$ and continued for 1000 instants after $t = T_s$.

Table 1 shows comparison – for different orders of the polynomial basis – of the results yielded by the proposed two-stage variable-bandwidth algorithm (15)-(16), combining estimates provided by 3 local WYW algorithms ($k_1 = 225$, $k_2 = 337$, $k_3 = 505$, cosinusoidal taper), with the analogous results obtained for 3 constant-bandwidth Dahlhaus algorithms (17)-(18) [D(k_1), D(k_2), D(k_3)], and the classical BF algorithm (20). The squared parameter estimation errors and Itakura-Saito spectral distortion measures [37], shown in Table 1, were averaged over $t \in [1, T_s]$ and 100 independent realizations of $\{y(t)\}$. In all cases M (the half-width of the local approximation interval) was set to 50 and l was set to 10 (20% overlap).

Note that in almost all cases the proposed algorithm outperformed the best-fitted Dahlhaus algorithm, and that its performance was much better than performance of the BF algorithm. The obtained results were also better than the results yielded by the bandwidth-adaptive algorithm (3)-(8) based on local estimation only – see Table 2. Finally, note that in the case considered the best results were obtained for a small number of basis functions ($m = 2$).

6. CONCLUSION

A new two-stage procedure for identification of locally stationary autoregressive processes was described and compared with the method proposed by Dahlhaus, and with the classical method of basis functions. The proposed algorithm adapts to the unknown and possibly time-varying rate of process nonstationarity, yields better results than the local estimation approach, used at the first stage of identification, and provides better results than the direct basis function approach, used at the second stage.

Table 1. Results of comparison of 3 identification algorithms, described in the text, obtained for different quality measures (mean-squared parameter estimation error, Itakura-Saito spectral distortion measure), different number of basis functions ($m = 1, \ldots, 5$), and different peeds of parameter variation (S_1, S_2, S_3).

parameter estimation errors

SoV	m	D(k_1)	D(k_2)	D(k_3)	BF	Proposed
S_1	1	3.178	0.579	**0.208**	1.365	**0.208**
	2	3.326	0.585	**0.200**	2.833	0.203
	3	3.362	0.590	**0.201**	4.208	0.204
	4	3.376	0.591	**0.201**	6.072	0.204
	5	3.384	0.592	**0.201**	9.255	0.204
S_2	1	3.269	0.713	0.449	1.406	**0.345**
	2	3.385	0.688	0.409	2.884	**0.313**
	3	3.420	0.691	0.410	4.274	**0.315**
	4	3.433	0.693	0.410	6.232	**0.315**
	5	3.441	0.693	0.410	9.498	**0.315**
S_3	1	3.484	1.416	2.180	1.566	**0.826**
	2	3.417	1.242	1.980	2.885	**0.674**
	3	3.447	1.246	1.982	4.279	**0.677**
	4	3.460	1.247	1.982	6.278	**0.678**
	5	3.468	1.248	1.982	9.502	**0.678**

spectrum estimation errors

SoV	m	D(k_1)	D(k_2)	D(k_3)	BF	Proposed
S_1	1	0.140	0.062	0.050	0.246	**0.042**
	2	0.144	0.060	0.048	0.546	**0.040**
	3	0.146	0.061	0.048	0.942	**0.041**
	4	0.147	0.061	0.048	1.604	**0.041**
	5	0.147	0.061	0.048	3.048	**0.041**
S_2	1	0.166	0.104	0.131	0.254	**0.074**
	2	0.162	0.095	0.122	0.549	**0.066**
	3	0.165	0.096	0.123	0.950	**0.067**
	4	0.165	0.096	0.123	1.630	**0.067**
	5	0.166	0.096	0.123	3.107	**0.067**
S_3	1	0.268	0.296	0.496	0.284	**0.204**
	2	0.235	0.263	0.471	0.547	**0.179**
	3	0.240	0.266	0.474	0.944	**0.181**
	4	0.241	0.266	0.475	1.622	**0.181**
	5	0.241	0.266	0.475	3.080	**0.181**

Table 2. Identification results obtained using 3 fixed-bandwidth WYW algorithms (k_1, k_2, k_3) and the variable-bandwidth algorithm based on minimization of the FPE statistic, described in [18].

parameter estimation errors

SoV	k_1	k_2	k_3	FPE
S_1	4.187	0.686	**0.209**	0.234
S_2	4.242	0.785	0.418	**0.346**
S_3	4.258	1.337	1.991	**0.713**

spectrum estimation errors

SoV	k_1	k_2	k_3	FPE
S_1	0.165	0.064	0.046	**0.043**
S_2	0.185	0.111	0.160	**0.070**
S_3	0.289	0.339	0.582	**0.184**

7. REFERENCES

[1] A. Elnagar and K. Gupta, "Motion prediction of moving objects based on autoregressive model," *IEEE Trans. Syst. Man Cyb.*, Part A, vol. 28, pp. 803-810, 1998.

[2] N. Channouf, P. LEcuyer, A. Ingolfsson and A.N. Avramidis, "The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta," Health Care Manag. Sci., vol. 10, pp. 25–45, 2007.

[3] M. Niedźwiecki, M. Ciołek, and K. Cisowski, "Elimination of impulsive disturbances from stereo audio recordings using vector autoregressive modeling and variable-order Kalman filtering," *IEEE Trans. Audio, Speech, and Lang. Process.*, vol. 23, pp. 970-981, 2015.

[4] M. Niedźwiecki and M. Ciołek, "New semi-causal and non-causal techniques for detection of impulsive disturbances in multivariate signals with audio applications," *IEEE Trans. Sign. Process.*, vol. 65, pp. 3881-3892, 2017.

[5] T. Wada, M. Jinnouchi and Y. Matsumura, "Application of autoregressive modelling for the analysis of clinical and other biological data," *Ann. Inst. Statist. Math.*, vol. 40, pp. 211-227, 1998.

[6] *Handbook of Brain Connectivity*, V.K. Jirsa and A.R. McIntosh, Eds., Springer-Verlag, 2007.

[7] A. Schlögl, *The Electroencephalogram and the Adaptive Autoregressive Model: Theory and Applications*. Aachen, Germany: Shaker Verlag, 2000.

[8] Z.G. Zhang, H.T. Liu, S.C. Chan, K.D.K. Luk, and Y. Hu, "Time-dependent power spectral density estimation of surface electromyography during isometric muscle contraction: Methods and comparisons," *J. Electromyogr. Kinesiol.*, vol. 20, pp. 89-101, 2010.

[9] A. A. Zaman, M. Ferdjallah, and A. Khamayseh, "Muscle fatigue analysis for healthy adults using TVAR model with instantaneous frequency estimation, in *Proc. IEEE 38th Southeast. Symp. Syst. Theory (SSST)*, Cookeville, TN, USA, pp. 244-247, 2006.

[10] Z.G. Zhang, Y.S. Hung, and S.C. Chan, "Autoregressive models with application to time-frequency analysis of event-related EEG," *IEEE Trans. Biomed. Eng.*, vol. 58, pp. 553-566, 2011.

[11] K.E. Baddour, N.C. Beaulieu, "Autoregressive models for fading channel simulation," IEEE Trans. Wireless Comm., vol. 4, pp. 1650–1662, 2005.

[12] J.F. Hayes and T.V.J. Ganesh Babu, *Modeling and Analysis of Telecommunication Networks,* Wiley, 2004.

[13] P. Lesage, F. Glangeaud and J. Mars, "Applications of autoregressive models and time-frequency analysis to the study of volcanic tremor and long-period events," *J. Volc. Geotherm. Res.*, vol. 114, pp. 391417, 2002.

[14] C. Li and R.L. Nowack, "Application of autoregressive extrapolation to seismic tomography," Bull. Seism. Soc. Amer., vol., pp. 1456-1466, 2004.

[15] D. Brillinger, E.A. Robinson and F.P. Schoenberg, Eds., *Time Series Analysis and Applications to Geophysical Systems*, Springer, 2012.

[16] L. Ljung and S. Gunnarsson, "Adaptative tracking in system identification A survey, *Automatica*, vol. 26, pp. 7-22, 1990.

[17] M. Niedźwiecki, *Identification of Time-varying Processes*, Wiley, 2000.

[18] M. Niedźwiecki, M. Ciołek, and Y. Kajikawa, "On adaptive covariance and spectrum estimation of locally stationary multivariate processes," *Automatica*, vol. 82, pp. 1–12, 2017.

[19] R. Dahlhaus, "Fitting time series models to nonstationary processes," *Ann. Statist.*, vol. 25, pp. 1–37, 1997.

[20] R. Dahlhaus, "Local inference for locally stationary time series based on the empirical spectral measure," *J. Econometrics*, vol. 151, pp. 101–112, 2009.

[21] R. Dahlhaus, "Locally stationary processes," *Handbook Statist.*, vol. 25, pp. 1–37, 2012.

[22] A. H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley, 2003.

[23] S. Kay, *Modern Spectrum Estimation: Theory and Application*, New Jersey: Prentice-Hall, 1999.

[24] P.M.T. Broersen, "Facts and fiction in spectral analysis,", *IEEE Trans. Instrument. Meas.*, vol. 49, pp. 766–772, 2000.

[25] H. Akaike, "Fitting autoregressive models for prediction," *Ann. Inst. Statist. Math.*, vol. 21, pp. 243–247, 1969.

[26] H. Akaike, "Statistical predictor identification," *Ann. Inst. Statist. Math.*, vol. 22, pp. 203–217, 1970.

[27] M. Niedźwiecki, "On the localized estimators and generalized Akaike's criteria," *IEEE Trans. Automat. Contr.*, vol. 29, pp. 970–983, 1984.

[28] T. Subba Rao, "The fitting of nonstationary time-series models with time-dependent parameters," *J. R. Statist. Soc. B*, vol. 32, pp. 312-322, 1970.

[29] J.M. Mendel, *Discrete Techniques of Parameter Estimation: The Equation Error Formulation*, New York: Marcel Dekker, 1973.

[30] J.M. Liporace, "Linear estimation of nonstationary signals," *J. Acoust. Soc. Amer.*, vol. 58, pp. 1288-1295, 1975.

[31] M. Hall, A.V. Oppenheim, and A. Willsky, "Time-varying parametric modeling of speech," *Signal Processing*, vol. 5, pp. 267–285, 1983.

[32] R. Charbonnier, M. Barlaud, G. Alengrin and J. Menez, "Results on AR-modelling of non-stationary signals," *Signal Processing*, vol. 12, pp. 143–151, 1987.

[33] M. Niedźwiecki, "Functional series modeling approach to identification of nonstationary stochastic systems," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 955-961, 1988.

[34] R.B. Mrad, S.D. Fassois, and J.A. Levitt, "A polynomial-algebraic method for non-stationary TARMA signal analysis-Part I: The method," *Signal Processing*, vol. 65, pp. 1–19, 1998.

[35] A.G. Poulimenos and S.D. Fassois, "Parametric time-domain methods for non-stationary random vibration modelling and analysisa critical survey and comparison," *Mechanical Systems and Signal Processing*, vol. 20, pp. 763–816, 2006.

[36] M. Niedźwiecki, "Functional series modeling identification of nonstationary stochastic systems – the clipping technique," *Proc. 1st IAESTED Symp. Signal Process. App.*, Brisbane, Australia, pp. 321–326, 1987.

[37] F. Itakura and S. Saito, "A statistical method for estimation of speech spectral density and formant frequencies," *Electron. Commun. Jap.*, vol. 53-A, pp. 36–43, 1970.