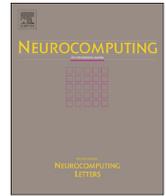




ELSEVIER

Contents lists available at ScienceDirect

## Neurocomputing

journal homepage: [www.elsevier.com/locate/neucom](http://www.elsevier.com/locate/neucom)

# Algebraic segmentation of short nonstationary time series based on evolutionary prediction algorithms



Rita Palivonaite, Kristina Lukoseviciute, Minvydas Ragulskis\*

Research Group for Mathematical and Numerical Analysis of Dynamical Systems, Kaunas University of Technology, Studentu 50-222, Kaunas LT-51368, Lithuania

## ARTICLE INFO

### Article history:

Received 27 July 2012

Received in revised form

16 February 2013

Accepted 2 May 2013

Communicated by P. Zhang

Available online 16 July 2013

### Keywords:

Segmentation

Time series prediction

Rank of a sequence

## ABSTRACT

Algebraic segmentation of short nonstationary time series is presented in this paper. The proposed algorithm is based on the algebraic one step-forward predictor which is used to identify a temporal near-optimal algebraic model of the real-world time series. A combinatorial algorithm is used to identify intervals where prediction errors are lower than a predefined level of acceptable accuracy. Special deterministic strategy is developed for the selection of this acceptable level of prediction accuracy and is individually determined for every time series. The nonparametric identification of quasistationary segments is performed without the employment of any statistical estimator. Several standard real-world time series are used to demonstrate the efficiency of the proposed technique.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

The evolution of complex systems in many cases can be considered being composed of bouts of quasistationarity in which time-varying pseudoparameters remain more or less unchanged. In general, a proper segmentation of a time series provides a useful portrait of the local properties for investigating and modeling nonstationary systems [1]. Such segmentation serves as a valuable tool in different areas including physics [2–4], biology [5–7], image and signal processing [8,9] and other disciplines. A number of different segmentation techniques have been introduced and successfully exploited in various computational set-ups. The extraction of Kramers–Moyal coefficients of representative stochastic equations from time series data is proposed in [10]. A fixed characteristic scale (window length) is used to assess the scenario of mixed statistics and to produce a statistically meaningful segmentation in [11]. Statistical methods for the identification of the discrepancy between the segments on both sides of a moving pointer are used in [12]. The algorithm based on Student's *t*-statistics is used to test the significance of the null hypothesis of equal means in appropriate segments in [13,14]. The Jensen–Shannon divergence is exploited for the segmentation of symbolic sequences in [15]. A segmentation approach to generate surrogate signals by randomly removing data segments from stationary signals with

different types of long-range correlations is considered in [16]. An exploratory analysis algorithm for functional data clustering and segmentation is proposed in [17]. A new kind of similarity measure specifically tailored to subspace representation and piecewise segmentation of artificial and real data time series is proposed in [18]. A variational Bayesian formulation of a manifold-constrained hidden Markov model is applied in [19] to the segmentation of a set of multivariate times series of electromyographic recordings.

It is well known that long data sets are one of the prime requirements of time series analysis techniques to unravel the dynamics of an underlying system, though acquiring long data sets is often not possible. The question of whether it is still possible to understand the complete dynamics of a system if only short (but many) time series are observed and if a single long time series can be generated from these short segments using the concept of recurrences in phase space is addressed in [20]. The main idea of our approach is based on the identification of skeleton algebraic sequences representing local models of short time series. In that sense our methodology is somewhat similar to the switching state-space model introduced in [21] and the adaptive segmentation technique proposed in [22]. AR predictor is used in the pioneering work of Bodenstein and Praetorius [22] to monitor the error rate of a one-step predictor in order to detect abrupt changes and then to use that information for the segmentation of the time series. Since then many AR based techniques have been implemented for time series segmentation. Our approach also belongs to the class of methods originated by [22]. But the main advantage of our methodology is based on the concept of skeleton algebraic sequences. Moreover, we do not only detect the moment

\* Corresponding author. Tel.: +370 69 822456.

E-mail addresses: [rita.palivonaite@ktu.lt](mailto:rita.palivonaite@ktu.lt) (R. Palivonaite), [kristina.lukoseviciute@ktu.lt](mailto:kristina.lukoseviciute@ktu.lt) (K. Lukoseviciute), [minvydas.ragulskis@ktu.lt](mailto:minvydas.ragulskis@ktu.lt) (M. Ragulskis).

URL: <http://www.personalas.ktu.lt/~mragul> (M. Ragulskis).

of a potential change in the evolution of the process. Our technique classifies skeleton sequences into separate classes (what enables an effective application of a novel combinatorial algorithm).

The concept of skeleton algebraic sequences has been introduced in [23] and has been successfully exploited for the prediction of short real-world time series. We will exploit this algebraic one step-forward prediction technique for the nonparametric segmentation of short nonstationary real-world time series.

## 2. Preliminaries

Time series prediction, especially short time series prediction, is a challenging problem in many fields of science and engineering. The objective of all forecasting techniques is to build a model of the process and then use this model on the last values of the time series to extrapolate past behavior into future. Forecasting procedures include different techniques and models, and although the search for a best time series forecasting method continues, it is agreeable that no single method will outperform all others in all situations.

It is common that the forecasting horizon of short time series predictors is one step-forward only [24–27]. As mentioned previously, we will also use a one point forward prediction technique. But the main difference of the algebraic prediction technique introduced in [23] from other alternative short time series predictors is in the fact that the algebraic predictor identifies the algebraic complexity of the time series by means of the H-rank (or the Hankel rank) of a sequence. The identification of the H-rank will serve as key computational tool in the segmentation procedure of the analyzed time series.

Though the definition of the H-rank is explicitly introduced in [23], we will give a concise overview of the computational techniques used for the identification of H-ranks, what will help to interpret the segmentation algorithm presented in the following section.

The Hankel matrix, named after Hermann Hankel, is widely used for system identification when given a sequence of output data and a realization of an underlying state-space model is desired. A first solution to this challenging system-theoretic problem that became known as the state-space realization problem was provided in 1965 in [28]. The key tool for solving this problem is the Hankel matrix, whose factorization into the product of an observability matrix and controllability matrix is known as the Ho–Kalman realization method [28]. The Hankel matrix-based models are appropriate to describe linear input/output mappings by infinitely many parameters, in general, since they might be obtained directly from available input/output data on the system. It took years of research to go from the theoretical results described in [28] to a numerically reliable realization algorithm [29]. The combination of deterministic realization theory based on the factorization of the Hankel matrix, with the theory of Markovian and innovations representations, gave rise to the stochastic theory of minimal realizations. The stochastic realization problem was studied intensively during the early 1970s in connection with innovations theory and spectral factorization theory [30,31]. Many new innovative applications based on the Hankel matrix have been developed in diverse areas of science and engineering. Gathering outputs from an impulse-response simulation into a generalized Hankel matrix and its singular value decomposition (SVD) helps to obtain reduced order models for high dimensional linear dynamical systems [32]. An infinite polynomial block Hankel matrix, as well as its associate  $\tau$ -finite polynomial block Hankel matrices, is used in [33–37] in order to relate the spectral controllability and observability properties of

minimal realizations with the minimum feasible finite rank of such a Hankel matrix.

Let  $S$  is a sequence of real or complex numbers:

$$S := (x_0, x_1, x_2, \dots) \tag{1}$$

The Hankel transform of the sequence  $S$  yields a sequence of determinants of Hankel catalecticant matrices  $\{d_n\}_{n \geq 0}$  [38]:

$$d_n := (x_{i+j-2})_{1 \leq i, j \leq n+1}$$

$$= \begin{vmatrix} x_0 & x_1 & \dots & x_n \\ x_1 & x_2 & \dots & x_{n+1} \\ \vdots & \vdots & \dots & \vdots \\ x_n & x_{n+1} & \dots & x_{2n} \end{vmatrix} \tag{2}$$

**Definition 1.** The H-rank of the sequence  $S$  is equal to  $m$ ;  $m \in \mathbb{N}$  if

$$d_{m+k-1} = 0 \tag{3}$$

for all  $k \in 1, 2, \dots$ , but  $d_{m-1} \neq 0$ .

**Example 1.** Though the class of considered algebraic models is much wider than periodic sequences, let us consider a periodic sequence  $S = (a, b, c, a, b, c, \dots)$ .

Then

$$d_0 = a; \quad d_1 = \begin{vmatrix} a & b \\ b & c \end{vmatrix}; \quad d_2 = \begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix};$$

$$d_3 = \begin{vmatrix} a & b & c & a \\ b & c & a & b \\ c & a & b & c \\ a & b & c & a \end{vmatrix} = 0; \quad d_4 = d_5 = \dots = 0; \quad Hr = 3.$$

**Theorem 1.**  $HrS = m$  if and only if elements of the deterministic algebraic sequence  $S$  are expressed in the form:

$$x_n = \sum_{k=1}^r \sum_{l=0}^{n_k-1} \mu_{kl} \binom{n}{l} \rho_k^{n-l}, \quad n = 0, 1, 2, \dots \tag{4}$$

where the characteristic roots  $\rho_k \in \mathbb{C}$ ;  $k = 1, 2, \dots, r$  can be determined from the Hankel characteristic equation

$$\begin{vmatrix} x_0 & x_1 & \dots & x_m \\ x_1 & x_2 & \dots & x_{m+1} \\ \vdots & \vdots & \dots & \vdots \\ x_{m-1} & x_m & \dots & x_{2m-1} \\ 1 & \rho & \dots & \rho^m \end{vmatrix} = 0; \tag{5}$$

the recurrence indexes of these roots  $n_k$  ( $n_k \in \mathbb{N}$ ) satisfy the equality  $n_1 + n_2 + \dots + n_r = m$ ; coefficients  $\mu_{kl} \in \mathbb{C}$ ;  $k = 1, 2, \dots, r$ ;  $l = 0, 1, \dots, n_k - 1$  can be determined from a system of linear algebraic equations which can be formed from the systems of equalities in Eq. (4), and this system has a unique solution. Rigorous proof of Theorem1 is given in [39].

In general, deterministic algebraic sequences belong to the large class of auto-regressive models (which may become nonlinear if the roots of the Hankel characteristic equation are recurrent).

The algebraic prediction technique in [23] exploits the concept of the H-rank and performs the extrapolation of the reconstructed algebraic model into the future. But a random sequence does not have a rank (otherwise an explicit algebraic law would govern the evolution of that sequence) – therefore a real-world time series does not have an H-rank simply due to an inevitable contamination by noise. Thus special evolutionary computational strategies are developed in [23] for the identification of a closest skeleton algebraic sequence to the real-world time series. As

mentioned previously, we will exploit this algebraic predictor for the segmentation of the time series.

The idea is based on the assumption that the sequence  $S$  is produced by adding noise to some unknown algebraic sequence. The H-rank of that unknown algebraic sequence is assumed to be equal to  $m$ ;  $2m + 1$  observations are available for building the algebraic model of the process;  $x_{2m}$  is the value of the observation at the present moment. In other words, it is assumed that the sequence  $\tilde{S} := (x_0 - \varepsilon_0, x_1 - \varepsilon_1, x_2 - \varepsilon_2, \dots)$  is some sort of a skeleton sequence determining the global dynamics of the time series ( $Hr\tilde{S} = m$ ). Corrections  $\varepsilon_k$ ;  $k = 0, 1, 2, \dots, 2m$  must be identified before any prediction could be made. Since the goal is to minimize distortions of the original time series, the fitness function for the set of corrections  $\{\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{2m}\}$  is introduced in [23]

$$F(\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{2m}) = \frac{1}{a|\tilde{d}_m| + \sum_{k=0}^{2m} \lambda_k |\varepsilon_k|}; \quad a > 0; \quad (6)$$

where  $\tilde{d}_m$  is the  $m+1$ -th Hankel determinant of  $\tilde{S}$  ( $\tilde{d}_m$  should be equal to 0 if all corrections would be exact);  $a$  is the penalty proportion between the magnitude of the determinant and the sum of weighted corrections; coefficients  $\lambda_k$  determine the tolerance corridor for the corrections (all corrections would be the same if  $\lambda_k = 1/(2m + 1)$ ;  $k = 0, 1, \dots, 2m + 1$ ) [23]. The fitness function would reach its maximum at  $\varepsilon_0 = \varepsilon_1 = \dots = \varepsilon_{2m} = 0$  if the original time series  $S$  is an algebraic sequence and  $d_m = 0$  ( $F(0, 0, \dots, 0) = +\infty$  then). Otherwise evolutionary algorithms are used to identify the near-optimal set of corrections  $\{\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{2m}\}$  and to compute one step-forward prediction  $x_{2m+1}$  using Eq. (4).

### 3. The construction of the segmentation algorithm

We will describe the proposed segmentation algorithm in detail in this section. But before depicting the details of this algorithm we construct an artificial time series which will be used for the validation purposes.

#### 3.1. The artificial time series

The artificial scalar time series comprises 191 elements (Fig. 1). The first segment of the series represents a periodic sequence. The period length is 8 and elements in the period are: 0.5, 0.7, 0.1, 0.9, 0.3, 0.2, 0.8 and 0.4. The period is repeated 6 times; the length of the first segment is 48 elements.

The second segment is constructed as a periodic sequence with the trend. The periodic part comprises 5 elements: 0.6, 0.2, 0.7, 0.1 and 0.4. A step 0.05 is added consecutively to every element in this segment. Thus, elements in the first part of the segment read: 0.6, 0.25, 0.8, 0.25, 0.6; elements in the second part read: 0.85, 0.5, 1.05, 0.5, 0.85; the process is repeated 7 times (the second segment comprises 35 elements).

The third segment comprises 4 periods of 11 elements (2.5, 1.9, 2.7, 1.7, 2.1, 2.0, 1.5, 2.6, 1.8, 2.3, 1.5). The fourth segment contains 28 elements – the periodic sequence 1.5, 0.6, 1.1, 0.8 is repeated 7 times. Finally, 9 elements (0.2, 0.7, 0.4, 0.9, 0.1, 0.8, 0.5, 0.3, 0.6) are repeated 4 times in the fifth segment. Algebraic H-ranks are shown at appropriate segments in Fig. 1(b).

It is clear that the generated artificial time series cannot be considered as a good representation of a real-world process simply due to explicit algebraic relationships between elements of the sequence (at appropriate segments). In order to test the functionality of the segmentation algorithm on realistic signals we add the evenly distributed noise in the interval  $[-0.15, 0.15]$  to all elements of the generated sequence – the graph of the sequence with the additive noise is shown as a solid line in Fig. 1(a).

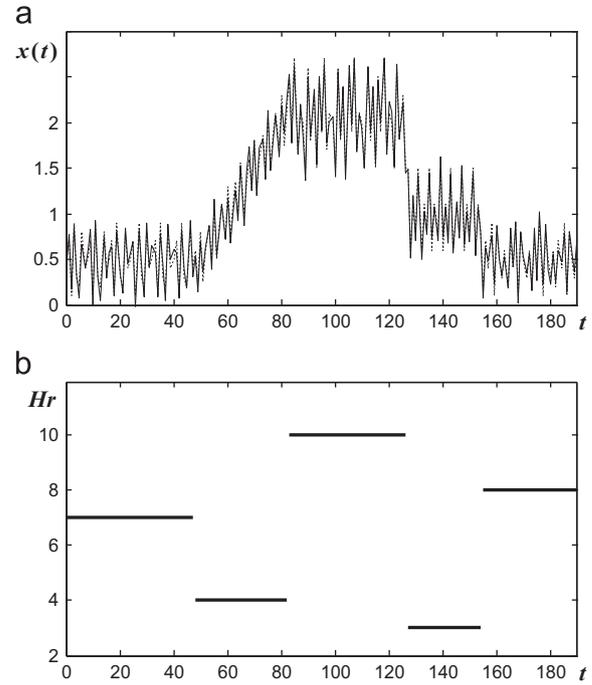


Fig. 1. The artificial time series constructed from 5 stationary segments. The dashed line represents the noiseless time series and the solid line stands for the time series with the additive noise evenly distributed in the interval  $[-0.15, 0.15]$  (part (a)). Numerical values of H-ranks at appropriate segments of the noiseless time series are illustrated in part (b).

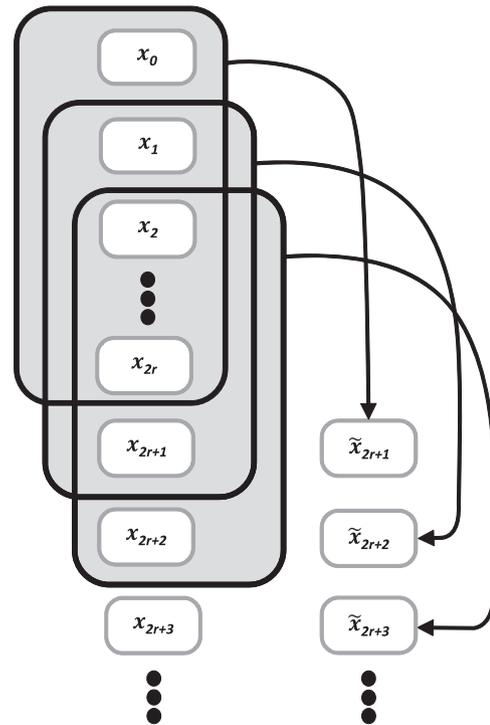


Fig. 2. The schematic diagram illustrating the one step-forward prediction technique exploited in the segmentation algorithm; it is assumed that  $Hr\tilde{S} = r$  for the whole time series. Gray-shaded blocks of  $2r+1$  elements represent observation windows necessary to produce algebraic estimates of the next elements of the sequence.

#### 3.2. One step-forward algebraic prediction of time series

As mentioned previously, we will use the time series prediction algorithm based on the identification of skeleton algebraic

sequences [23] for the segmentation of the time series. But instead of trying to identify the most appropriate H-rank of the time series at the beginning of the prediction process, we will perform the prediction at different preset values of the H-rank.

The selection of the effective range of H-ranks is the first step of the segmentation algorithm. In general, this selection can be free, though too wide range of H-ranks would raise the computational costs required by the proposed technique. We preselect  $3 \leq Hr \leq 12$  for the artificial time series with additive noise.

The schematic diagram of the prediction process is illustrated in Fig. 2. Given a sequence  $(x_0, x_1, x_2, \dots)$  let us assume that the H-rank is set to  $r$ . Then, according to Eq. (5),  $2r+1$  elements are required to form the Hankel characteristic equation (the first block of  $2r+1$  elements is illustrated as the top gray-shaded block in Fig. 2). Note that we do not check if the determinant of the Hankel matrix  $d_r$  is equal to zero. We simply use the prediction algorithm presented in [23] and extrapolate the skeleton sequence by one element into the future:  $\tilde{x}_{2r+1}$  is the algebraic prediction of the sequence  $(x_0, x_1, \dots, x_{2r})$  (Fig. 2). Next, we shift the observation window by one element and predict  $\tilde{x}_{2r+2}$  (Fig. 2). The process is continued until the last element of the original data sequence is predicted. It must be noted that the first element we can predict is  $\tilde{x}_{2r+1}$ . The higher the preselected H-rank  $r$ , the larger the amount of data we have to accumulate in order to perform the first prediction.

The next step is the selection of the tolerable error level  $\delta$  for the algebraic prediction of the analyzed time series. The basic idea of the proposed technique is straightforward: the preselected

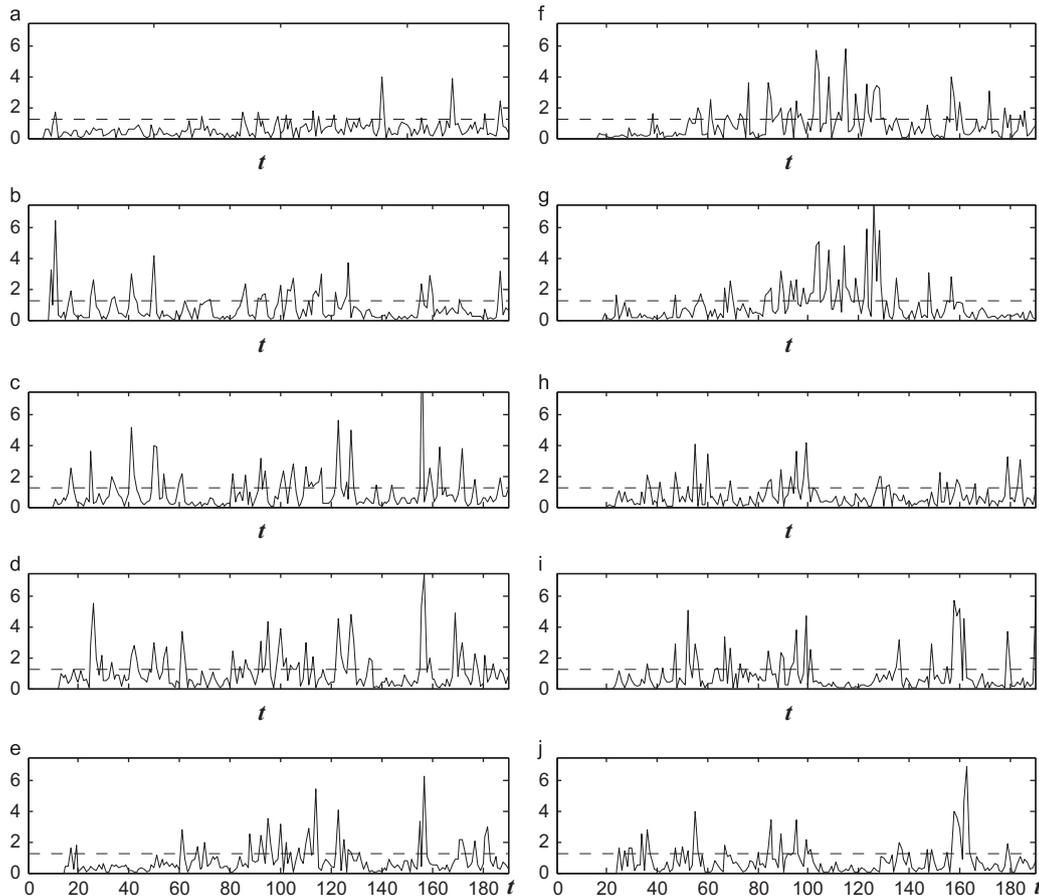
algebraic model is sufficiently accurate if extrapolation errors of the prediction are lower than  $\delta$ .

Initially we select  $\delta = 1.2015$  for the artificial time series with the additive noise and perform the prediction of this time series for  $Hr = 3, 4, \dots, 12$  (Fig. 3). Note that the H-rank is not changed during the prediction process – we start with a preselected H-rank and continue the prediction until the end of the time series.

The lower bound of the effective range of H-ranks is predetermined by the fact that the condition  $Hr < 3$  results into primitive time sequences ( $Hr=2$  for an elementary arithmetic progression;  $Hr=1$  for an elementary geometric progression). On the other hand, the length of the vector of corrections  $\{\epsilon_k\}$  is equal to 27 already at  $Hr=12$  (what raises computational costs of the prediction algorithm). At least 27 elements of the original time series are required in order to produce a single one step-forward forecast at  $Hr=12$ . But we need to have more elements in order to be able to make any comparison between the original and the predicted time series. We entail that the number of elements required to make at least one step-forward prediction (at the highest H-rank) should be not larger than 20% of the whole time series. Then, the upper bound of the effective range of H-ranks ( $Hr=12$ ) is a decent selection for a short time series comprising about 150 elements.

### 3.3. Combinatorial aspects of the segmentation algorithm

Before defining an explicit rule for the selection of  $\delta$ , we will describe combinatorial aspects of the segmentation algorithm. Let us return to the previous step where predictions are made at



**Fig. 3.** Absolute prediction errors for the artificial time series with the additive noise at  $HrS=3$  (part (a));  $HrS=4$  (part (b)); ...;  $HrS=12$  (part (j)). Horizontal dashed lines in all parts represent the acceptable level of prediction errors  $\delta = 1.2015$ . The percentage of successful predictions (when the absolute prediction error is lower than  $\delta$ ) is shown for every H-rank. (a)  $p=0.89$   $HrS=3$ , (b)  $p=0.80$   $HrS=4$ , (c)  $p=0.76$   $HrS=5$ , (d)  $p=0.71$   $HrS=6$ , (e)  $p=0.82$   $HrS=7$ , (f)  $p=0.77$   $HrS=8$ , (g)  $p=0.75$   $HrS=9$ , (h)  $p=0.85$   $HrS=10$ , (i)  $p=0.86$   $HrS=11$  and (j)  $p=0.81$   $HrS=12$ .

a preselected H-rank equal to  $r$  (Fig. 2) and assume that a certain level  $\delta$  is preset at the beginning of the computational experiment (Fig. 3). Now, if the absolute prediction error  $|\tilde{x}_{2r+k} - x_{2r+k}|$  is lower than  $\delta$  ( $k = 1, 2, \dots$ ) we plot a black dot; otherwise a white dot is left unmarked at appropriate value of  $k$  (see horizontal dotted lines in Fig. 4). Note that adjacent black dots merge forming black line intervals. Of course, such computational experiments are performed for all values of  $r$  in the effective range of H-ranks (Fig. 3). The percentage of successful predictions (when the absolute prediction error is lower than  $\delta$ ) is computed for every H-rank ( $p = 0.89$  in Fig. 3(a) denotes that 89% of predictions were successful at  $Hr = 3$ ).

Next, we propose a combinatorial segmentation algorithm for the identification of longest continuous intervals of successful predictions in the effective range of H-ranks, at a predefined level of prediction errors  $\delta$ . A schematic diagram of this algorithm is illustrated in Fig. 5. Note that the diagram in Fig. 5 is constructed for illustrative purposes only and does not represent any particular time series.

Step A: Set the level  $\delta$  ( $\delta > 0$ ) and perform the algebraic forecasting of the given data series at different preselected ranks. Mark intervals of the time series where the forecasting errors were lower than  $\delta$ . Such marking is schematically illustrated in Fig. 5(a) (the vertical axis stands for the H-rank  $r$ ). For example, the interval  $(t_0; t_5)$  is associated to the lowest H-rank in the effective range of H-ranks; the interval  $(t_1; t_2)$  is associated to the highest H-rank in the schematic diagram in Fig. 5(a). Note that these marked intervals may overlap for different H-ranks. Also, some intervals can be left unassociated to any particular H-rank. For example, forecasting errors in the interval  $(t_7; t_8)$  are higher than  $\delta$  for all  $r$  in the effective range of H-ranks (Fig. 5(a)).

Step B: Identify the longest continuous interval (in the whole range of effective H-ranks). The longest interval is  $(t_4; t_7)$ ; it is marked by a gray shaded area in Fig. 5(a).

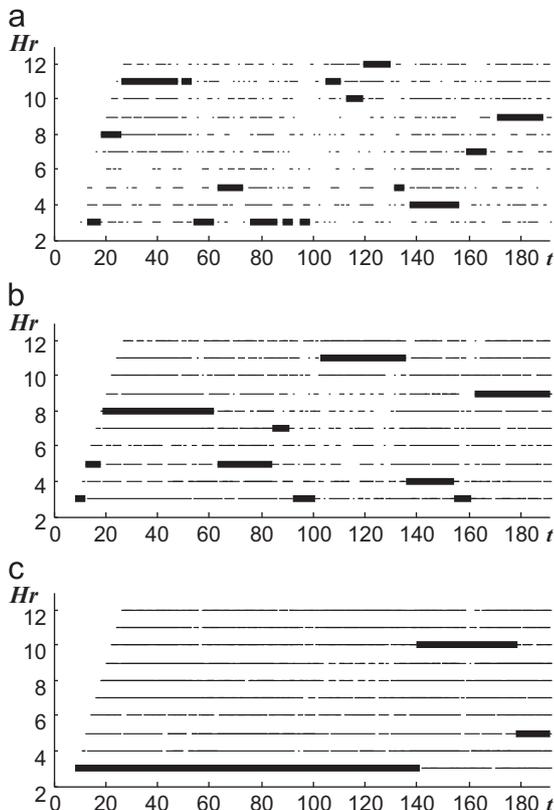


Fig. 4. The result of the segmentation algorithm for the artificial time series with the additive noise for  $\delta = 0.4923$  (part (a));  $\delta = 1.2015$  (part (b)) and  $\delta = 2.5712$  (part (c)).

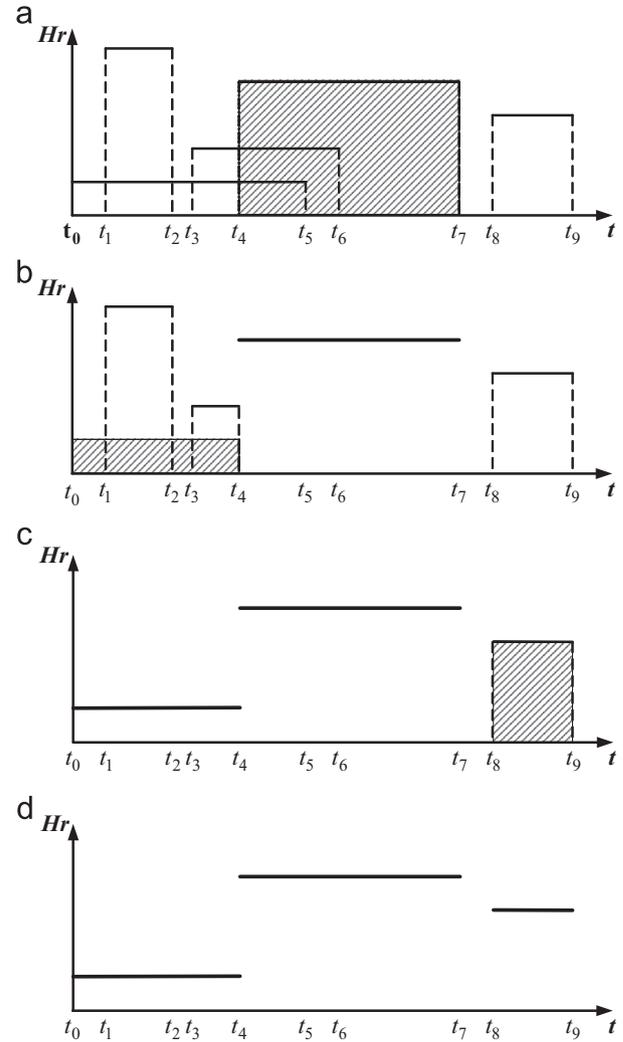


Fig. 5. The illustration of the combinatorial segmentation algorithm. Horizontal lines in part (a) show intervals where algebraic prediction errors are lower than the preset level  $\delta$  (the height of a line stands for the appropriate H-rank). The gray-shaded area in part (a) illustrates the longest continuous line interval which is associated to a separate segment in part (b). The process is continued through parts (b–d) until the whole sequence is split into separate segments. Thin lines show intervals where prediction errors are smaller than  $\delta$ ; thick solid lines represent the result of the segmentation algorithm.

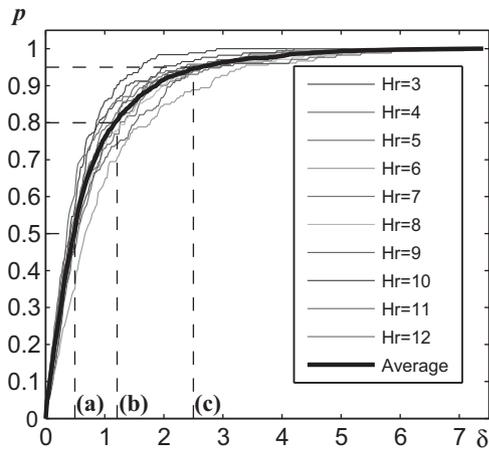
Step C: Denote the marked interval as the segment associated to the according H-rank; erase all information about predictions at other H-ranks in the marked segment. The marked segment is illustrated by a thick solid horizontal line in Fig. 5(b).

Step D: Identify the longest continuous interval in the zones not occupied by the marked segments (return back to step B). The longest interval  $(t_0; t_4)$  is marked by a gray shaded area in Fig. 5(b).

Step E: Continue until all possible intervals are marked as segments. The interval  $(t_0; t_4)$  is associated to the according H-rank in Fig. 5(c). The last marked interval in Fig. 5(c) is  $(t_8; t_9)$ . Finally, the segmentation algorithm identifies four distinct segments:  $(t_0; t_4)$ ;  $(t_4; t_7)$ ;  $(t_7; t_8)$  and  $(t_8; t_9)$  in Fig. 5. Note that the unmarked interval  $(t_7; t_8)$  is not associated to any particular H-rank.

### 3.4. The strategy for the selection of $\delta$

The results of the segmentation (for the artificial time series contaminated with noise) are illustrated in Fig. 4 by thick solid lines. It is clear that the plotted dichotomous lines of prediction errors in Fig. 4 would be sparse if  $\delta$  is considerably lower than the



**Fig. 6.** A diagram illustrating the selection of the acceptable level of prediction errors  $\delta$  – thin solid lines represent percentages of successful predictions of the artificial time series with the additive noise for different H-ranks. The thick solid line represents the average of all percentages for a fixed  $\delta$ . The average percentage  $p=0.5$  corresponds to  $\delta = 0.4923$  (marked as (a));  $p=0.8$  corresponds to  $\delta = 1.2015$  (marked as (b));  $p=0.95$  corresponds to  $\delta = 2.5712$  (marked as (c)).

average level of absolute prediction errors (Fig. 4(a)). On the contrary, continuous intervals of acceptable predictions would be long if  $\delta$  is much higher than the average level of absolute prediction errors (Fig. 4(c)). One has to identify the appropriate level of  $\delta$  which would result into a realistic segmentation (Fig. 4(b)). Note that the darker lines in Fig. 4 represent the H-rank for that segment as identified by the proposed segmentation algorithm.

The selection of the acceptable level of prediction errors  $\delta$  is illustrated by the diagram in Fig. 6. As mentioned previously, we perform the algebraic prediction of the time series and plot absolute prediction errors for every single H-rank in the effective range of H-ranks. We fix a particular level of  $\delta_{kl}$  and compute the percentage of satisfactory predictions in terms of  $\delta$  (note that  $\delta$  is the same for all H-ranks in Fig. 3). We repeat such computations for different values of  $\delta$  and calculate the percentage of average satisfactory predictions  $p$  as the arithmetic mean for all H-ranks in the effective range of H-ranks (Fig. 6). It is clear that  $\lim_{\delta \rightarrow 0} p = 0$  for real-world time series because the inevitable noise does not allow the exact reconstruction of the algebraic model of the time series. On the other hand,  $p$  saturates to 1 (corresponding to 100%) when  $\delta$  reaches the level of the highest absolute prediction error (Fig. 6). The values  $p=0.5, 0.8$  and  $0.95$  result into  $\delta = 0.4923, 1.2015$  and  $2.5712$ , respectively (Fig. 6).

As mentioned previously, different levels of absolute prediction errors  $\delta$  result into different segmentations of the original time series. Fig. 4 illustrates the segmentation of the artificial time series with the additive noise at  $\delta = 0.4923; 1.2015$  and  $2.5712$ . The nearest segmentation corresponding to the original formation of the artificial time series is observed at  $p=0.8$ . Thus we fix this value of  $p$  and continue with the segmentation experiments with real-world time series.

So far, such a selection of the parameter  $p$  is based only on computational experiments with the artificial time series contaminated with noise. Nevertheless, segmentation experiments with other artificial time series (different algebraic sequences, different levels of additive noise) also suggest that  $p=0.8$  is an optimal choice for the segmentation of a time series with an embedded deterministic law. The variation of H-ranks and the evolutionary strategy for the identification of nearest algebraic skeleton sequences help to construct an effective deterministic algorithm for unsupervised segmentation. The parameter  $p$  plays a role of a criterion which is used to declare the fact that a previously

assumed algebraic model cannot be longer extrapolated outside the identified segment. Moreover, this dimensionless parameter does not directly depend on such factors as the signal range, the signal-noise ratio, absolute prediction errors. The parameter  $p$  depends on  $\delta$  – but the range of  $\delta$  is not predetermined at the beginning of the segmentation experiment. The value  $p=0.8$  serves as a good conciliation among two extremities – the situation when prediction errors are unacceptable almost everywhere (at any H-rank) and the situation when prediction errors are acceptable everywhere for all possible H-ranks. The proposed segmentation algorithm is not only robust to noise – it does not breakdown even if it is used to segment the noise itself, though such segmentation may not have a direct physical meaning because the prediction errors are several times higher than the signal itself.

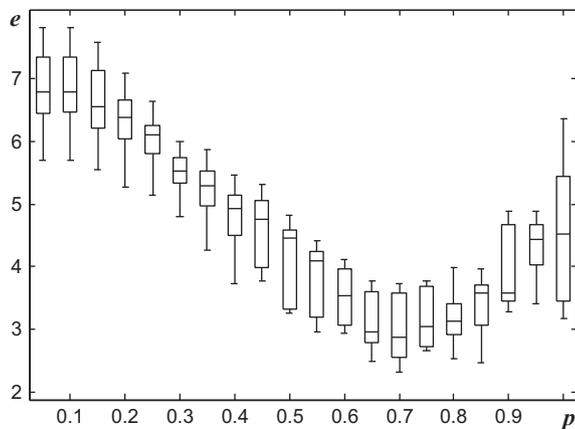
### 3.5. The strategy for the selection of $p$

Different levels of absolute prediction errors  $\delta$  result into different segmentations of the original time series. Fig. 4 illustrates the segmentation of the artificial time series with the additive noise at  $\delta = 0.4923; 1.2015$  and  $2.5712$  (the nearest segmentation corresponding to the original distribution of H-ranks for the artificial time series is observed at  $p=0.8$ ).

The primary objective of the proposed segmentation algorithm is to identify segments where H-ranks are stationary. In other words, we do not only identify intervals of quasistationary in a time series. We do reconstruct a deterministic algebraic model governing the evolution of the time series in each of those segments. The H-rank could be considered as a parameter of the prediction method – but it also serves as a powerful tool for the algebraic identification of the model of the system and provides a deep insight into the underlying structure of the time series (in the current segment). The algebraic prediction method uses the current setting of the H-rank and tries to identify characteristic roots  $\rho_k$  and coefficients  $\mu_{kl}$  (Eq. (4)) for the data in the analyzed observation window (Fig. 2). Such identification of the algebraic model would be exact for a noise-free time series. But since all real-world time series are inevitably contaminated with noise, we are forced to exploit evolutionary algorithms (Eq. (6)) in order to identify underlying algebraic models. The variation of H-ranks and the evolutionary strategy for the identification of nearest algebraic skeleton sequences help to construct an effective deterministic algorithm for unsupervised segmentation.

As mentioned previously, the main objective of all segmentation algorithms is to identify intervals of quasistationarity in a time series. In our case we request even more – the proposed segmentation algorithm should identify not only intervals of quasistationarity, but also hidden algebraic models governing the evolution of the system at those intervals. In other words, the goal of the proposed algorithm is to produce the distribution of H-ranks which is as close as possible to the underlying actual distribution of H-ranks. Let us denote  $h_n$  as the exact value of the H-rank at the location of the  $n$ -th element of the time series and  $\tilde{h}_n$  as the value of the H-rank produced by the proposed segmentation algorithm. The optimal value of  $p$  should minimize the RMSE of the reconstructed distribution of H-ranks  $e(p) = \sqrt{(1/N) \sum_n (h_n - \tilde{h}_n)^2}$ ; ( $N$  is the total length of the time series).

So far, all computational experiments were produced using the artificial time series contaminated with noise (Fig. 1). Now we construct 25 different artificial time series – all contaminated with different types and levels of noise; the distribution of H-ranks is different for every artificial time series. Note that we do not use real-world time series in this computational experiment (because we do not know the actual distribution of H-ranks for a real-world time series). All 25 artificial time series are segmented using

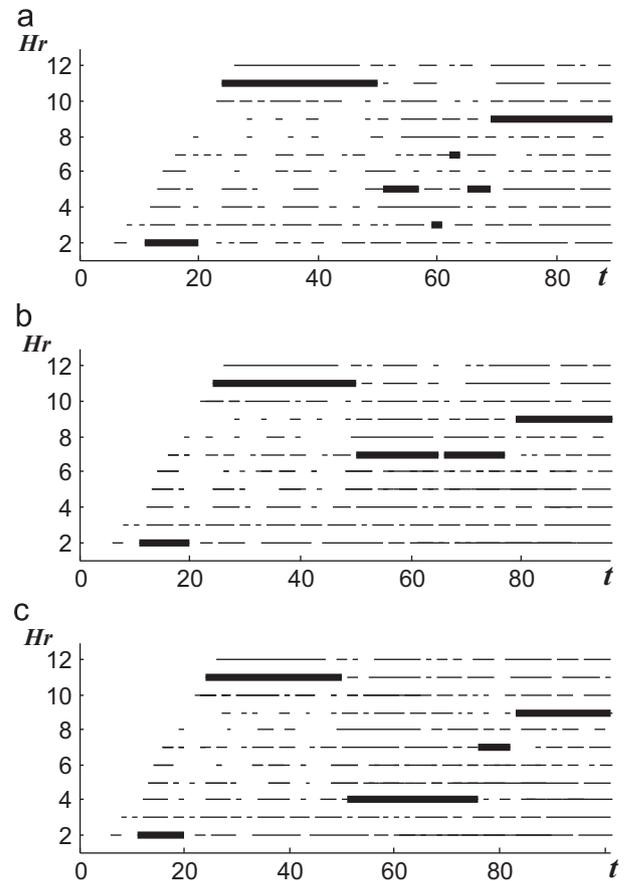


**Fig. 7.** The distribution of errors of reconstruction of H-ranks for 25 different artificial time series. The central thick line in the middle of each box marker is the median of the reconstruction error; the edges of the box are the 25th and 75th percentiles; the whiskers extend to the most extreme data points.

different levels of  $p$ ; the distribution of  $e(p)$  is plotted in Fig. 7. The central thick line in the middle of each box marker is the median; the edges of the box are the 25th and 75th percentiles; the whiskers extend to the most extreme data points. It can be observed that the value  $p=0.8$  serves as a good conciliation among two extremities – the situation when prediction errors are unacceptable almost everywhere (at any H-rank) and the situation when prediction errors are acceptable everywhere for all possible H-ranks; moreover this value of  $p$  corresponds to the best reconstruction of the underlying algebraic model of the time series. Thus we fix  $p=0.8$  and continue with the segmentation experiments with real-world time series.

### 3.6. The sensitivity of the method to the length of a segment

It is clear that RMSE is not enough to elucidate what is the length of the identified segment. RMSE is tightly linked with the concept of the H-rank of a sequence in our approach. RMSE of the prediction errors would have no meaning in the context of the segmentation without the H-rank. Dynamic observation of these two intertwined parameters (RMSE and H-rank) lies at the foundation of the presented methodology. The minimum length of a segment in order to be detected depends on the H-rank of the algebraic model. At least  $(2r+1)$  elements are required for the identification of a segment if the H-rank of the algebraic model is equal to  $r$ . That can be seen in Fig. 8 – empty gaps at the left sides of parts a–c in Fig. 8 are proportional to H-ranks of algebraic models. Experiments with variable length segments are carried out with the aim to know to what extent the results are confident in terms of the length of the estimated segments. We exploit the same artificially generated time series for that purpose. But instead of using five different segments (as shown in Fig. 1) we use only three. The first segment corresponds to the third segment of the original artificial time series (4 periods of 11 elements); the third segment corresponds to the fifth segment of the original artificial time series (4 periods of 9 elements). But the middle segment (corresponding to the fourth segment of the artificial time series) is constructed as a variable length segment: 8 elements in part (a); 15 elements in part (b) and 20 elements in part (c) of Fig. 8. Note that a randomly distributed noise is added to all elements of the sequence (analogously as to the original artificial time series). The H-rank of the first segment would be equal to 11 (if the elements of that segment would not be perturbed by noise); 23 elements are required to detect the algebraic law in the



**Fig. 8.** The sensitivity of the segmentation algorithm to the length of the segment. The length of the second segment is 8 elements in part (a), 15 elements in part (b) and 20 elements in part (c).

segment. The proposed method detects the hidden algebraic law in the first segment as soon as the number of analyzed points is sufficient to construct the appropriate algebraic relationship (Fig. 8 (a)–(c)). The H-rank of the second segment is not detected in Fig. 8 (a) but the segmentation results of the middle interval are rather good in Fig. 8(b) and (c). One must keep in mind that we do test the functionality of the proposed methodology on a signal contaminated by noise. It is well known in the statistical literature that estimating the rank of noise contaminated data is very difficult [40,41]. A straightforward identification of the H-rank in our model becomes an ill-posed problem because a random sequence does not possess a finite H-rank (otherwise a deterministic algebraic law generating the random sequence could be reconstructed). It is well known that the summation of two sequences results into the H-rank not lower than the maximum H-rank of one of the sequences [42]. Therefore, the H-rank of noise contaminated data is infinite. Nevertheless, near-optimal identification of skeleton sequences in noise contaminated data enables an efficient reconstruction of the underlying algebraic model [23]. The results of the segmentation of the noise contaminated data presented in Fig. 8 demonstrate the robustness of the proposed technique. Moreover, the proposed method is capable to identify the H-rank as soon as the number of elements is sufficient to reconstruct the underlying algebraic model.

## 4. Computational experiments with real-world time series

We test the functionality of the proposed segmentation algorithm using real-world time series. We select a standard odonovan7.

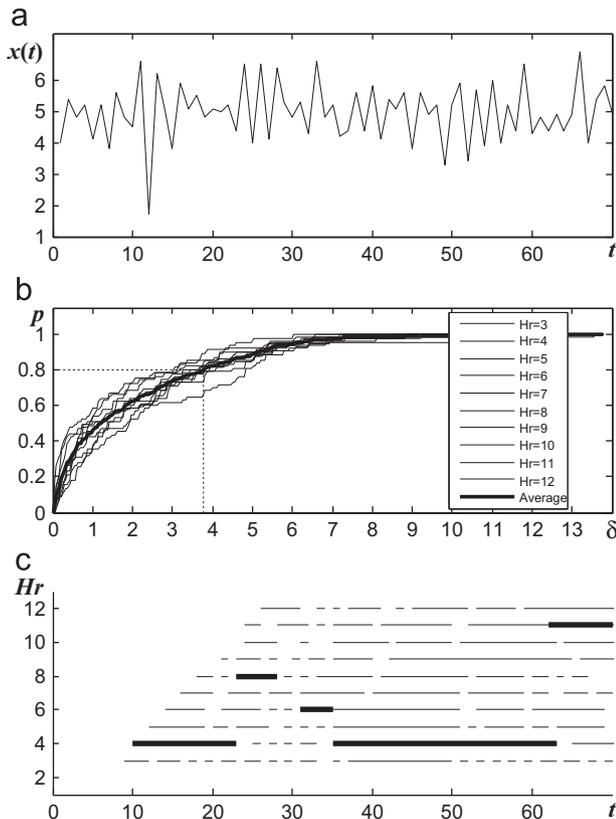
dat time series describing 70 consecutive readings of batch chemical process [43] (Fig. 9(a)). This time series is short because the available number of elements is too small for training any classifier or network. Nevertheless, the proposed algorithm copes well with the segmentation task. Percentages of satisfactory predictions in the effective range of H-ranks ( $3 \leq Hr \leq 12$ ) are illustrated in Fig. 9(b). The average percentage of satisfactory predictions  $p=0.8$  results into the absolute prediction error level  $\delta=3.76$  (Fig. 9(b)). The combinatorial segmentation algorithm (at  $\delta=3.76$ ) produces the segmentation illustrated in Fig. 9(c).

We continue computational experiments with a standard barison.dat time series describing monthly basic iron production in Australia in thousand tons in the time period between January 1956 and August 1995 [43] (210 available discrete data points are plotted in Fig. 10(a)). Percentages of satisfactory predictions in the effective range of H-ranks ( $3 \leq Hr \leq 12$ ) are illustrated in Fig. 10(b). The average percentage of satisfactory predictions  $p=0.8$  results into the absolute prediction error level  $\delta=0.9591$  (Fig. 10(b)). The combinatorial segmentation algorithm (at  $\delta=0.9591$ ) produces the segmentation illustrated in Fig. 10(c).

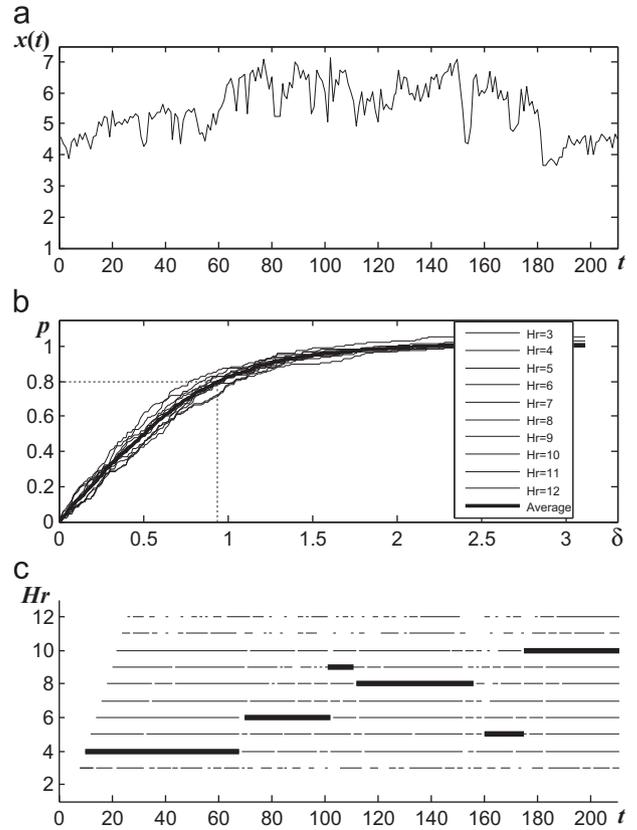
The produced segmentation results provide a deep physical insight into evolution of the real-world time series. We are able to observe intervals where algebraic laws governing the evolution of the process are stationary. Also we are able to identify potential changes in the evolution of the process but there is no way for the current methodology to estimate how abrupt the changes are.

### 5. Comparisons with other segmentation techniques

A comparative assessment of the functionality of the proposed technique with other typical segmentation methods is required

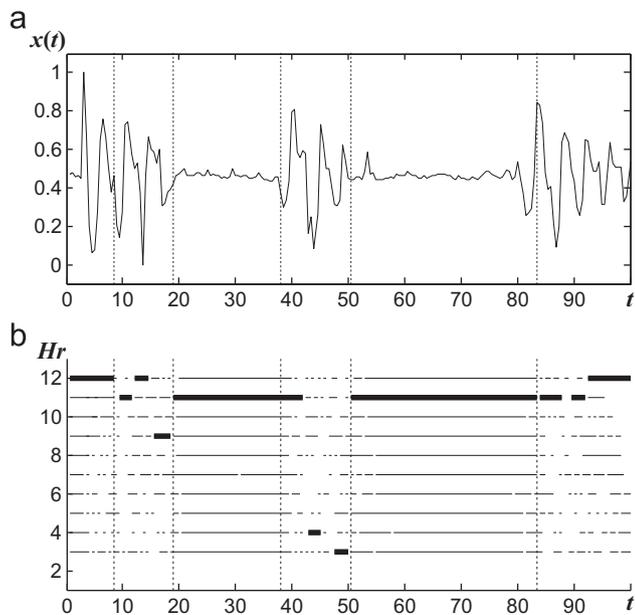


**Fig. 9.** The segmentation of odonovan7.dat time series. The time series is illustrated in part (a); average percentages of successful predictions are shown as a thick solid line in part (b);  $p=0.8$  corresponds to  $\delta=3.76$ . The result of the segmentation is presented in part (c).



**Fig. 10.** The segmentation of barinson.dat time series. The time series is illustrated in part (a); average percentages of successful predictions are shown as a thick solid line in part (b);  $p=0.8$  corresponds to  $\delta=0.9591$ . The result of the segmentation is presented in part (c).

in order to understand if our methodology does outperform other methods or not, and under which conditions does it happen. The first comparison is performed with the segmentation method based on switching state-space models [21]. This model combines and generalizes two of the most widely used stochastic time series models – the hidden Markov model and the linear dynamical system. It is demonstrated in [21] that switching state-space models are useful in modeling time series which have nonlinear dynamics characterized by several different regimes. To illustrate this point, Ghahramani and Hinton examined a psychological data set from a patient tentatively diagnosed with sleep apnea, which is a medical condition in which patients intermittently stop breathing during sleep, which results in a reflex arousal and gasps of breath. The data was obtained from the repository of time series data sets associated with Santa Fe Time Series Analysis and Prediction Competition and is described in detail in Rigney et al. [44]. The task is simple – it is necessary to highlight the fact that the respiration pattern in sleep apnea is characterized by at least two regimes – no breathing and gasping breathing. Note that Ghahramani and Hinton used samples 6201–7200 for training and 5201–6200 for testing [21]. Our methodology does not require any training at all – we can start the segmentation right from the first samples. Thus we apply our segmentation algorithm to samples 5201–5401; segmentation results are presented in Fig. 11. It can be noted that our methodology does not produce only two different types of segments. Therefore, though it is easy to locate no breathing regimes in our segmentation results, the location of gasping breathing regimes is more difficult compared to the results produced by [21]. That can be considered as a definite drawback of our methodology. On the other hand, the original time series is rather simple (from the point of the segmentation process). No breathing

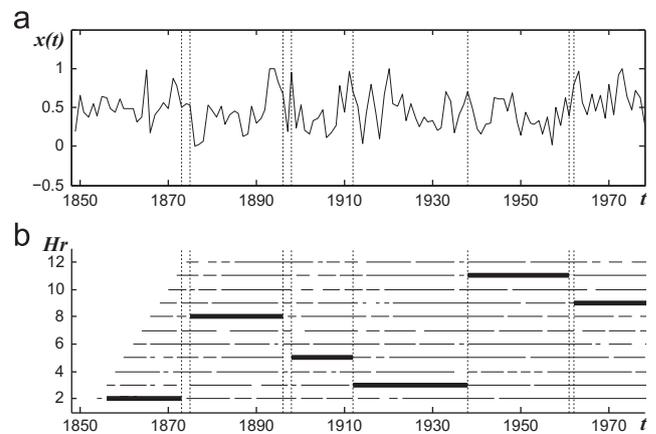


**Fig. 11.** Segmentation results for the patient breathing data during sleep: the time series is shown in part (a); segmentation results are illustrated in part (b).

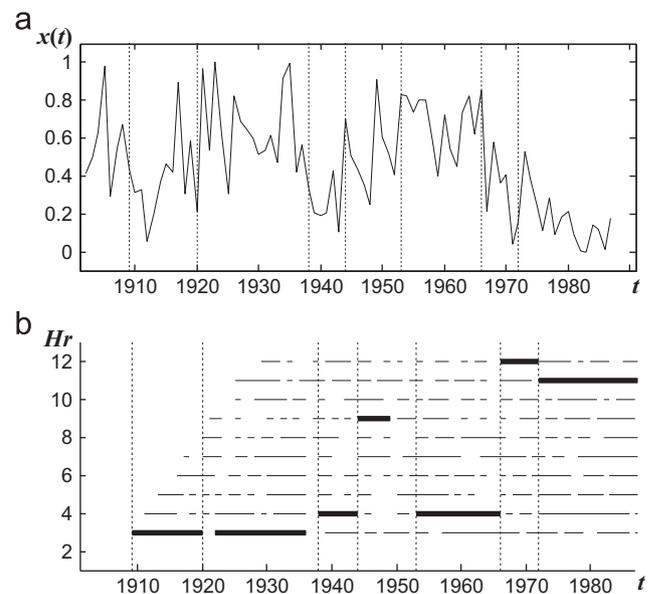
and gasping breathing regimes can be easily identified by a naked eye; a sophisticated segmentation method is not necessary for the interpretation of data. Our methodology outperforms the switching state-space model from that point of view – we are able to locate algebraic relationships also in the gasping breathing regimes. It is well known that time series representing human physiological data are chaotic [45]. No algebraic relationship (linear or nonlinear) can describe long-term evolution of a chaotic signal [42]. In that sense our segmentation results provide a deeper insight into the evolution of the process than a simple classification into two different states.

The second comparison is performed with the segmentation method proposed by Aksoy et al. [46]. The time series has a length of 131 years and consists of the annual total precipitation data (in mm) at Fortaleza, Brazil, for period 1849–1979 [47]. The segmentation results produced by [46] single out the following intervals: 1848–1893, 1894–1897, 1898–1962 and 1963–1979, which is the highest order segmentation accepted by the Scheffe test. With the exception of the 4 year segment during the period 1893–1896, the annual precipitation at Fortaleza can be considered stable for more than a century until 1962, after which an increase is observed up to the end of the period, 1979. Our segmentation methodology has singled out the following intervals: 1848–1873, 1875–1896, 1898–1912, 1912–1938, 1939–1962 and 1963–1979 (Fig. 12). Thus, our methodology was able to detect the major change points located in [47] but still managed to find additional change points. Therefore, we may conclude that our methodology is more sensitive to changes in process evolution compared to [46].

The last comparison is performed with the time series segmentation method with shifting means hidden Markov models [48]. The experiment is performed with the Senegal River annual discharge data, measured at the Bakel station for the years 1903–1988 [49]. The results of our segmentation are shown in Fig. 13. The algorithm [48] produces breaks at years 1921, 1938, 1949 and 1967. Our methodology is able to detect the break at 1920 (1921), 1938, 1967 but is not able to locate the break at 1949. It is well known that no single time series prediction method will outperform all others in all situations. Our prediction methodology is tightly related to the one step forward algebraic predictor introduced in [48]. Thus, poor prediction may result into poor segmentation. And though our segmentation methodology does show rather promising



**Fig. 12.** Segmentation results for the annual total precipitation data (in mm) at Fortaleza, Brazil, for period 1849–1979: the time series is shown in part (a); segmentation results are illustrated in part (b).



**Fig. 13.** Segmentation results for the Senegal River annual discharge data, measured at the Bakel station for the years 1903–1988: the time series is shown in part (a); segmentation results are illustrated in part (b).

results, it is natural to expect that there exist other segmentation methods which do outperform our results.

## 6. Concluding remarks

It is important to note that the proposed segmentation algorithm is based on an efficient computational strategy. Algebraic predictions are made for different H-ranks only once. All further computations are performed with absolute prediction errors, but the predictions do not need to be repeated. The acceptable level of prediction errors  $\delta$  is varied from 0 to the maximum absolute prediction error and the percentages of successful predictions are computed for the already available data. Our computations show that the average percentage of successful predictions  $p=0.8$  yields an optimal value of  $\delta$  – the optimality is considered as the closest segmentation to the underlying intervals of quasistationarity.

The strategy for the selection of the percentage of successful predictions  $p$  is thoroughly discussed in Section 3.5. We did perform computation experiments with 25 different time series and best segmentation results have been observed at  $p=0.8$ . This

particular value of  $p$  serves as a good conciliation among two extremities – the situation when prediction errors are unacceptable almost everywhere (at any H-rank) and the situation when prediction errors are acceptable everywhere for all possible H-ranks. Nevertheless, a question remains open if  $p=0.8$  is the best value for any time series. In general, one could perform additional tuning of the parameter  $p$  – but a large database of time series with known segmentation results should be available for that purpose. Unfortunately, almost all available segmentation results of real world time series are more or less empirical. Different authors compare the functionality of their segmentation algorithms, but a “standard” segmentation cannot be found (known) beforehand expect for an artificial time series (possibly contaminated by noise). That is the main reason why we fix the value  $p=0.8$  for all real world time series.

The proposed segmentation algorithm uses the one step-forward algebraic predictor. There exist numerous alternative predictors which could be used for short-term and long-term forecasting of noisy real-world time series. But the algebraic predictor (which we use in our algorithm) is based on the concept of H-ranks. In other words, the predictor identifies a near-optimal algebraic model of the time series and extrapolates that model into the future. We propose a segmentation algorithm based on the identification of changes in the mimicking algebraic model of the time series.

The proposed algorithm belongs to the class of level-set computational algorithms. The proposed algorithm belongs to the class of level-set computational algorithms. We do not compute statistical estimators of the prediction quality. Instead we do classify the time series into dichotomous intervals according to one step forward predictions. But instead of simply detecting the moment when absolute prediction errors exceed a predefined level, we develop a strategy applicable for nonparametric identification of quasistationary segments. We could use the same algebraic predictor and move with one step-forward forecasts until the prediction error at some point becomes higher than a preset level. Then, we should have to identify a new best fitting H-rank for the next interval and continue until the prediction error exceeds the preset level again. Unfortunately, such an approach possesses two serious drawbacks. The first one is related to the accumulation of data before the algebraic prediction can be commenced ( $2r+1$  data points are required for the algebraic prediction at  $HrS=r$ ). Thus, relatively long intervals between adjacent segments would be left without an association to any segment. The second drawback is related to a rather complex identification of the best-fitting H-rank. The proposed strategy liberates the user from the necessity of searching a best-fitting H-rank. Predictions are performed for all different H-ranks (in a pre-selected range) and a combinatorial level-set based algorithm is used for the identification of appropriate segments.

Such segmentation has a deep physical meaning. We do identify bouts of quasistationarity; the evolution of the process is governed by a fixed algebraic law in each reconstructed segment. The proposed segmentation algorithm does not apply formal algebraic relationships for the observed data. It reveals that the hidden structure of the time series is able to identify potential changes in the evolution of the process and exploits predictability as a tool for the characterization of complexity [50].

## Acknowledgments

Financial support from the Lithuanian Science Council under Project no. MIP-100/2012 is acknowledged.

## References

- [1] S. Camargo, S.M. Duarte Queirós, C. Anteneodo, Nonparametric segmentation of nonstationary time series, *Phys. Rev. E* 84 (2011) 46702.
- [2] I. Neri, N. Kern, A. Parmeggiani, Totally asymmetric simple exclusion process on networks, *Phys. Rev. Lett.* 107 (2011) 68702.
- [3] A. Kothe, V.S. Zykov, H. Engel, Second universal limit of wave segment propagation in excitable media, *Phys. Rev. Lett.* 103 (2009) 154102.
- [4] A. Diaz, P. Tirtik, M. Guizar-Sicairos, A. Menzel, P. Thibault, O. Bunk, Quantitative x-ray phase nanotomography, *Phys. Rev. B* 85 (2012) 020104.
- [5] A. Szabo, E.D. Perryn, A. Cziorok, Network formation of tissue cells via preferential attraction to elongated structures, *Phys. Rev. Lett.* 98 (2007) 038102.
- [6] E.P. Petrov, T. Ohrt, R.G. Winkler, P. Schuille, Diffusion and segmental dynamics of double-stranded DNA, *Phys. Rev. Lett.* 97 (2006) 258101.
- [7] Y. Zhang, Relations between Shannon entropy and genome order index in segmenting DNA sequences, *Phys. Rev. E* 79 (2009) 041918.
- [8] M. Brand, V. Kettner, Discovery and segmentation of activities in video, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (8) (2000) 844–851.
- [9] M. Weigel, Connected-component identification and cluster update on graphics processing units, *Phys. Rev. E* 84 (2011) 036709.
- [10] C. Anteneodo, S.M. Duarte Queirós, Low-sampling-rate Kramers–Moyal coefficients, *Phys. Rev. E* 82 (2010) 041122.
- [11] C. Beck, E.G.D. Cohen, Superstatistics, *Physica A* 322 (2003) 267–275.
- [12] I. Berkes, R. Gebrys, L. Horváth, P. Kokoszka, Detecting changes in the mean of functional observations, *J. R. Stat. Soc. B* 71 (2009) 927–946.
- [13] P. Bernaola-Galván, P.Ch. Ivanov, L.A.N. Amaral, H.E. Stanley, Scale invariance in the nonstationarity of human heart rate, *Phys. Rev. Lett.* 87 (2001) 168105.
- [14] K. Fukuda, H.E. Stanley, L.A.N. Amaral, Heuristic segmentation of a nonstationary time series, *Phys. Rev. E* 69 (2004) 021108.
- [15] P. Carpena, J.L. Oliver, M. Hackenberg, A.V. Coronado, G. Barturen, P. Bernaola-Galván, High-level organization of isochores into gigantic superstructures in the human genome, *Phys. Rev. E* 81 (2010) 031908.
- [16] Q.D.Y. Ma, R.P. Bartsch, P. Bernaola-Galván, M. Yoneyama, P.Ch. Ivanov, Effect of extreme data loss on long-range correlated and anti-correlated signals quantified by detrended fluctuation analysis, *Phys. Rev. E* 81 (2010) 031101.
- [17] G. Hébraïl, B. Huguency, Y. Lechevallier, F. Rossi, Exploratory analysis of functional data via clustering and optimal segmentation, *Neurocomputing* 73 (2010) 1125–1141.
- [18] E. Fuchs, T. Gruber, H. Pree, B. Sick, Temporal data mining using shape space representations of time series, *Neurocomputing* 74 (2010) 379–393.
- [19] I. Olier, J. Amengual, A. Vellido, A variational Bayesian approach for the robust analysis of the cortical silent period from EMG recordings of brain stroke patients, *Neurocomputing* 74 (2011) 1301–1314.
- [20] C. Komalapiya, M. Thiel, M.C. Romano, N. Marwan, U. Schwarz, J. Kurths, Reconstruction of a system's dynamics from short trajectories, *Phys. Rev. E* 78 (2008) 066217.
- [21] Z. Ghahramani, G.E. Hinton, Switching State-space Models, Technical Report CRG-TR-96-3, Department of Computer Science, University of Toronto, Toronto, 1996.
- [22] G. Bodenstern, H.M. Praetorius, Feature extraction from the electroencephalogram by adaptive segmentation, *Proc. IEEE* 65 (5) (1977) 642–652.
- [23] M. Ragulskis, K. Lukoseviciute, Z. Navickas, R. Palivonaite, Short-term time series forecasting based on the identification of skeleton algebraic sequences, *Neurocomputing* 74 (2011) 1735–1747.
- [24] M.C. Lee, C.N. Ko, Timeseries prediction using RBF neural networks with a nonlinear time-varying evolution PSO algorithm, *Neurocomputing* 73 (2009) 449–460.
- [25] Z.A. Bashir, M.E. El-Hawary, Applying wavelets to short-term load forecasting using PSO-based neural networks, *IEEE Trans. Power Syst.* 24 (2009) 20–27.
- [26] D. Niua, D. Liua, D.D. Wu, A softcomputing system for day-ahead electricity price forecasting, *Appl. Soft Comput.* 10 (2010) 868–875.
- [27] S.T.S. Bukkapatnam, C. Cheng, Forecasting the evolution of nonlinear and nonstationary systems using recurrence-based local Gaussian process models, *Phys. Rev. E* 82 (2010) 056206.
- [28] B.L. Ho, R.E. Kalman, Effective construction of linear state-variable models from input–output functions, *Regelungstechnik* 12 (1965) 545–548.
- [29] L.S. de Jong, Numerical aspects of recursive realization algorithms, *SIAM J. Control Optim.* 16 (1978) 646–659.
- [30] H. Akaike, Stochastic theory of minimal realization, *IEEE Trans. Autom. Control* 26 (1974) 667–673.
- [31] M. Gevers, T. Kailath, An innovations approach to least-squares estimation, part vi: discrete-time innovations representations and recursive estimation, *IEEE Trans. Autom. Control* 18 (1973) 588–600.
- [32] J. Juang, R. Pappa, An eigensystem realization algorithm for modal parameter identification and model reduction, *J. Guid. Control Dyn.* 8 (1985) 620–627.
- [33] I. Gohberg, M.A. Kaashoek, L. Lerer, On minimality in the partial realization problem, *Syst. Control Lett.* 9 (1987) 97–104.
- [34] D. Hinrichsen, W. Manthey, U. Helmke, Minimal partial realization by descriptor systems, *Linear Algebra Appl.* 326 (2001) 45–84.
- [35] L. Bakule, J. Rodellar, J.M. Rossell, Controllability–observability of expanded composite systems, *Linear Algebra Appl.* 332 (2001) 381–400.
- [36] S. Chaib, D. Boutat, A. Benali, J.P. Barbot, Observability of the discrete state for dynamical piecewise hybrid systems, *Nonlinear Anal.: Theory Methods Appl.* 63 (2005) 423–438.
- [37] C. Coll, M.J. Fullana, E. Sanchez, Reachability and observability indices of a discrete-time periodic descriptor system, *Appl. Math. Comput.* 153 (2004) 485–496.
- [38] F. Partington, An Introduction to Hankel Operators. LMS Student Texts, Cambridge University Press, Cambridge, 1988.

- [39] Z. Navickas, L. Bikulciene, Expressions of solutions of ordinary differential equations by standard functions, *Math. Model. Anal.* 11 (2006) 399–412.
- [40] R.K. Lim, M.Q. Phan, R.W. Longman, *State-space System Identification with Identified Hankel Matrix*, Mechanical and Aerospace Engineering Technical Report 3045, Princeton University, 1998.
- [41] M. Fazel, T.K. Pong, D. Sun, P. Tseng, Hankel matrix rank minimization with applications in system identification and realization, *SIAM. J. Matrix Anal. & Appl.* 34, (2013) 946–977.
- [42] M. Ragulskis, Z. Navickas, R. Palivonaite, M. Landauskas, Algebraic approach for the exploration of the onset of chaos in discrete nonlinear dynamical systems, *Commun. Nonlinear Sci. Numer. Simul.* 17 (2012) 4304–4315.
- [43] R.J. Hyndman, *Time Series Data Library* (<http://robjhyndman.com/TSDL/>) (accessed 13-February-2012).
- [44] D.R. Rigney, A.L. Goldberger, W.C. Ocasio, Y. Ichimaru, G.B. Moody, R.G. Mark, Multi-channel physiological data: description and analysis, in: A.S. Weigend, N.A. Gershenfeld (Eds.), *Time Series Prediction: Forecasting the Future and Understanding the Past*, Addison-Wesley, Reading, MA, 1993, pp. 105–129.
- [45] L. Glass, Introduction to controversial topics in nonlinear science: is the normal heart rate chaotic? *Chaos* 19 (2009) 028501.
- [46] H. Aksoy, A. Gedikli, N.E. Unal, A. Kehagias, Fast segmentation algorithms for long hydrometeorological time series, *Hydrol. Process.* 22 (2008) 4600–4608.
- [47] P.A. Morettin, A.R. Mesquita, J.G.C. Rocha, Rainfall at Fortaleza in Brazil revisited, *Time Ser. Anal. Theory Pract.* 6 (1987) 67–85.
- [48] A. Kehagias, V. Fortin, Time series segmentation with shifting means hidden Markov models, *Nonlinear Processes Geophys.* 13 (2006) 339–352.
- [49] P. Hubert, Change points in meteorological analysis, in: T.S. Rao, M.B. Priestley, O. Lessi (Eds.), *Applications of Time Series Analysis in Astronomy and Meteorology*, Chapman and Hall, 1997.
- [50] G. Boffetta, M. Cencini, F. Falcioni, A. Vulpiani, Predictability: a way to characterize complexity, *Phys. Rep.* 356 (2002) 367–474.



**Rita Palivonaite** received the M.Sc. degree in mathematics in 2008 from Kaunas University of Technology, Lithuania. She is currently an assistant lecturer and a Ph.D. degree student within the Research Group for Mathematical and Numerical Analysis of Dynamical Systems, Kaunas University of Technology. Her current areas of research interests are combinatorial optimization and scientific visualization.



**Kristina Lukoseviciute** received the Ph.D. degree in informatics in 2012 from Kaunas University of Technology, Lithuania. Since 2012 she is a lecturer at the Department of Mathematical Research in Systems, Kaunas University of Technology. Her current research interests are evolutionary computations and time series analysis.



**Minvydas Ragulskis** received the Ph.D. degree in 1992 from Kaunas University of Technology, Lithuania. Since 2002 he is a professor at the Department of Mathematical Research in Systems, Kaunas University of Technology. His research interests include nonlinear dynamical systems and numerical analysis. He is the founder and head of the Research Group for Mathematical and Numerical Analysis of Dynamical Systems.