Intensive Care Med (1998) 24: 1305–1314

ORIGINAL

M. Imhoff
M. Bauer
U. Gather
D. Löhlein

Statistical pattern detection in univariate time series of intensive care on-line monitoring data

Received: 29 April 1998
Accepted: 5 October 1998

This work was supported by the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 475. Preliminary results of this study were presented at the 10th European Congress of Intensive Care Medicine, Paris, 1997

M. Imhoff (✉) · D. Löhlein
Chirurgische Klinik, Städtische Kliniken, Beurhausstrasse 40, D-44137 Dortmund, Germany
email: mike@imhoff.de
Tel.: + 49 (231) 5020021
Fax: + 49 (231) 5020081

M. Bauer · U. Gather
Fachbereich Statistik, Universität Dortmund, D-44221 Dortmund, Germany

Abstract *Objectives:* To determine how different mathematical time series approaches can be implemented for the detection of qualitative patterns in physiologic monitoring data, and which of these approaches could be suitable as a basis for future bedside time series analysis.
Design: Off-line time series analysis.
Setting: Surgical intensive care unit of a teaching hospital.
Patients: 19 patients requiring hemodynamic monitoring with a pulmonary artery catheter.
Interventions: None.
Measurements and results: Hemodynamic data were acquired in 1-min intervals from a clinical information system and exported into statistical software for further analysis. Altogether, 134 time series for heart rate, mean arterial pressure, and mean pulmonary artery pressure were visually classified by a senior intensivist into five patterns: no change, outlier, temporary level change, permanent level change, and trend. The same series were analyzed with low-order autoregressive (AR) models and with phase space (PS) models. The resulting classifications from both models were compared to the initial classification. Outliers and level changes were detected in most instances with both methods. Trend detection could only be done indirectly. Both methods were more sensitive to pattern changes than they were clinically relevant. Especially with outlier detection, 95 % confidence intervals were too close. AR models require direct user interaction, whereas PS models offer opportunities for fully automated time series analysis in this context.
Conclusion: Statistical patterns in univariate intensive care time series can reliably be detected with AR models and with PS models. For most bedside problems both methods are too sensitive. AR models are highly interactive, and both methods require that users have an explicit knowledge of statistics. While AR models and PS models can be extremely useful in the scientific off-line analysis, routine bedside clinical use cannot yet be recommended.

Key words Time series analysis · Autoregressive model · Phase space model · Decision support · Critical care · Patient monitoring

Introduction

Today most of our bedside decisions are based on subjective judgment and experience, rather than on hard data analysis. Most of the changes of a variable over time are more important than one pathological value at the time of observation. Over the past three decades mathematical methods have been developed that allow the assessment of single or multiple variables over time.

There are various approaches to describe time-dependent data generated from dynamical systems reflecting the different natures of the underlying processes. Time-dependent data are either generated in controlled scientific and engineering experiments or observed in medical, biological, environmental, and econometric studies. Although in neighboring sciences that deal with dynamical systems some of the same phenomenona are investigated, different terminologies and interpretations of the data-generating mechanism are applied. This also leads to the use of a different calculus for data analysis.

Methods used for describing these systems assume that the process under consideration is deterministic or stochastic or a combination of the two. In mathematics and theoretical physics, methods from the theory of dynamical systems [1] are often used to describe deterministic processes, whereas in physical and engineering applications approaches from the fields of systems theory [2] and digital signal processing are frequently applied [3]. Most of this work deals with fully deterministic processes as well as deterministic processes and additive noise. In that context, noise is seen as measurement errors or additive random effects, which are superimposed on the deterministic signal. By contrast, in statistical time series analysis the models are constructed in a way that random effects also drive the processes themselves. This approach is suitable to model medical, biological, environmental, and econometric variables like blood pressure or stock prices, because many of these processes can be regarded as stochastic processes [4–10]. We, therefore, concentrate here on time series analysis techniques for analyzing intensive care data.

Only a few investigations have employed this methodology in the field of intensive care medicine. In general it has been shown that time series analysis techniques are suitable for retrospective analysis of physiologic variables [8–10]. A statistically similar, but methodologically more challenging task is the on-line analysis of intensive care monitoring data. Statistically, approaches to this problem are rare and not readily available for application in clinical practice [11–13].

For this study, two entirely different statistical methods were used to describe critical care time series: autoregressive models, which have been used in numerous applications since their introduction in the 1970s, and phase space models, which represent a new approach to time series data. In autoregressive models the current value of a process is expressed as a finite, linear aggregate of previous values of the process and a stochastic term [14, 15]. In phase space models, after a transformation, the time series data are regarded as a multivariate sample with dependent observations [16].

Several authors have applied autoregressive models in the field of critical care [8], in longitudinal physiologic experiments [10], as well as in studies of laboratory data of the chronically ill [9]. Autoregressive models have also been successfully used to describe on-line time series from intensive care bedside monitors [8]. Clinical investigations with autoregressive models into therapeutic effects have recently been reported from the fields of cardiology [10] rheumatology [17] neurology [18], psychiatry [19], and nursing research [20].

The phase space approach originally came from the theory of nonlinear systems and is used for analysis of complex, deterministic, and especially chaotic systems. Several authors have applied measures which are based on phase space embeddings to judge the complexity of heart rate dynamics [21–23]. The application of these methods to heart rate dynamics is only feasible when data are recorded with high sampling frequency (e.g., 50 ms [21]). In clinical information systems data are often recorded in longer time intervals (e.g., 1 min) and, therefore, a different approach for analysis has to be selected. A new application of phase space models was recently introduced in the context of intensive care data, where phase space models are formulated in the statistical framework of linear stochastic systems [16]. This allows a meaningful application of these models to different physiologic variables, even with small sample sizes.

The detection of qualitative patterns in physiologic monitoring data (e.g., outliers, level changes, trends) is one of the basic applications of medical time series analysis. Traditional statistical methods for pattern identification like cluster or discriminant analysis are appropriate for time-independent data [24] but cannot be used for identifying and discriminating time series patterns. One possible statistical approach to the identification of patterns in time series are state space models, first used in the engineering sciences [25]. Several applications of Kalman Filter techniques to intensive care data exist [12, 26]. But these procedures are not very reliable in the identification of patterns and they require significant computational power [26, 27].

Recently, neural networks have been used for describing and controlling dynamical systems [28, 29]. The transfer of neural networks to pattern recognition in intensive care data is difficult: every patient has to be controlled individually, and it is not clear how to construct training phases. It is also difficult to decide with the help of neural networks whether or not the state of a critically ill patient is improving, because too little information about the health of the patient is available for the training phase [30, 31].

In the following investigation we chose autoregressive and phase space models, because they are suitable to model the underlying dynamic processes of intensive care variables and seem to be promising for the identification of patterns [8, 9, 16]. This study evaluates two aspects of the application of time series analysis to on-line monitoring data: (a) can all patterns be correctly identified with the applied statistical methods? (b) are the applied methods adequate for clinical use?

Table 1 Physiological time series included in the study: variables and patterns, as identified a priori by an intensivist. Length of estimation and prediction periods for autoregressive models

	Variables and patterns					
	No change	Outlier	Level change		Trend	Total
			Temporary	Permanent		
Variable						
Heart rate	8	6	7	11	8	40
Arterial pressure (mean)	5	24	5	7	0	41
Pulmonary artery pressure (mean)	10	5	12	24	2	53
Total	23	35	24	42	10	134
Estimation period for AR models						
Mean	208	163	144	177	184	173
Maximum	501	481	334	451	410	501
Minimum	81	51	41	50	80	41
Prediction period for AR models						
Mean	98	94	152	150	94	123
Maximum	299	270	470	337	150	470
Minimum	20	20	40	50	50	20

Patients and methods

On a 16-bed surgical intensive care unit in a 2000-bed teaching hospital, on-line monitoring data were acquired from 19 consecutive critically ill patients (8 females, 11 males, mean age 65 years), who had pulmonary artery catheters for extended hemodynamic monitoring, in 1-min intervals from a standard clinical information system. These data were transferred into a secondary SQL database and exported into standard statistical software for further analysis. The system configuration was comprised of the following: (a) Clinical Information System (CIS): Emtek Continuum 2000, Version 4.1M3, (b) Decision Support System (DSS): Sybase SQL server 4.9.2, (c) Statistical Software: SPSS version 6.1, SAS version 6.12 with additional programs in SAS/IML, and (d) System platform: Sun Sparc, Sun Solaris 1.1.2 (CIS, DSS) and 2.5 (statistical software).

From a total of 550000 single observations of seven variables (heart rate and invasive pressures), segments of 150 to 500 observations for each variable were visually classified by a senior intensivist into five medically relevant patterns: no change, outlier, temporary level change, permanent level change, and trend. The intensivist did not have to define objective criteria to explain why he chose a specific classification. A total of 134 time series were included in the study. The classifications are listed in Table 1. All these a priori classifications were done by one senior intensivist. After all time series were classified, they were presented again in a different order for reclassification by the same intensivist. There were no differences in the classification, which was attributable to the obviousness of the patterns as shown in the examples in Figs. 2 to 6.

The same segments were analyzed with second order autoregressive [AR(2)] models and with phase space (PS) models.

Each series of autoregressive models was split into a model estimation period, and a prediction period, where the pattern should be diagnosed. The average length of the estimation period was 173 min, that of the prediction period 123 min (Table 1). The actual measurements were compared to the confidence intervals (CI) for the prediction period. According to the percentage of values outside the CI, the variation was classified into the different patterns. Values outside the CI were classified as an outlier, if less than five consecutive observations (= minutes) were outside the CI. A level change was identified by five or more consecutive observations outside the CI, and was called a temporary level change with less than 50 % of the prediction period outside the CI, and a permanent level change with 50 % or more of the prediction period outside the CI.

Trend patterns were indirectly identified by deviations of the autocorrelation function (ACF) of the residuals and the Durbin-Watson statistics. In this case, the ACF of the original series was analyzed for trend typical patterns. Typically, the ACF plot of a time series without a trend declines exponentially (Fig. 1). In the case of a trend (i. e., a nonstationary series) the ACF plot fades only slowly over a larger number of lags (Fig. 1). In these cases, an AR(2) model of the first derivative of the time series was calculated. If this model showed sufficient goodness of fit, a significant trend of the time series was assumed.

Because the correct CI width was unknown before the study, a posteriori adjustment of the CI was done and the analysis rerun with this CI. This was done in cases where an inexplicable difference between the visual classification and the percentage of outliers occurred. The margin of the CI determines the "sensitivity" of the model. In this study an autoregressive or phase space model was considered too sensitive if at least one outlier was detected by the model that was not described in the initial visual classification.

In order to compare the CIs between different time series as one measure of the sensitivity of the predicted model, the relative confidence interval (CI_{rel}) was calculated as the difference between the upper *(UCL)* and lower confidence limits *(LCL)* in relation to the fitted model *(FIT)* for the entire prediction period expressed in percent:

$$CI_{rel}[\%] = 100 \cdot \frac{UCL - LCL}{FIT}$$

For PS models, the first 60 observations were taken and retrospectively analyzed (i. e., outlying regions were estimated and patterns in this time interval identified). After this, a time window of length 60 was moved through the data. That means, that at time point 61 it was determined if the phase space vector y_{61} was in a distant region. If not, then no pattern was present, and the estimated outlying region was replaced by a new one that was estimated from the last 60 observations $y_2, \ldots, y_{61}$. This was repeated for every new observation as long as for the time point t the phase space vector $\vec{y}_t$ was in a distant region. Then the system was not in a steady state, and after analyzing the consecutive observations $y_{t+1}, y_{t+2}, \ldots$, it

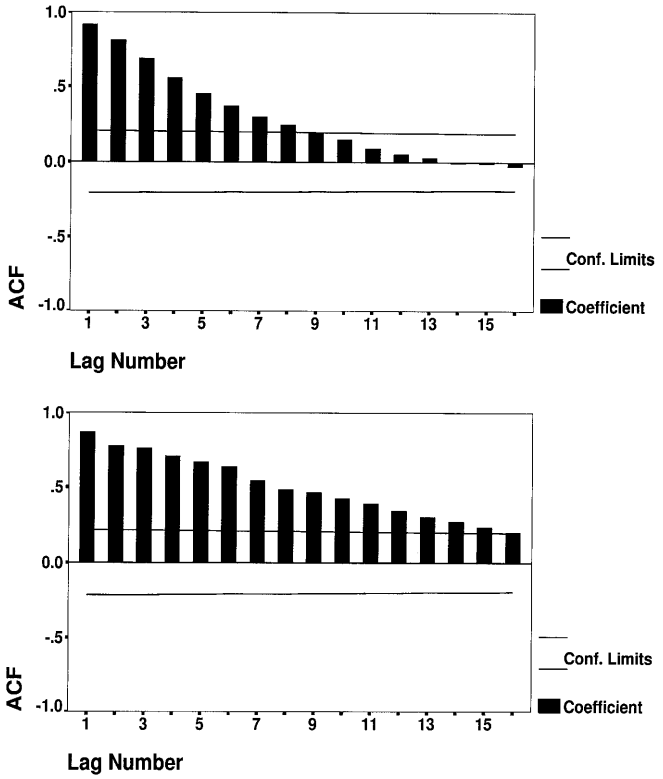

Fig. 1 **Top** Plot of the autocorrelation function for the first 16 lags of a series without a trend (prediction period of the series from Fig. 4). Rapid decline of the coefficients for the ACF. **Bottom** Plot of the autocorrelation function for the first 16 lags of a series with a trend (prediction period of the series from Fig. 7). Slow decline of the coefficients for the ACF. *Solid bars* coefficients of the ACF or PACF, *solid lines* 95% confidence limit for the ACF or PACF

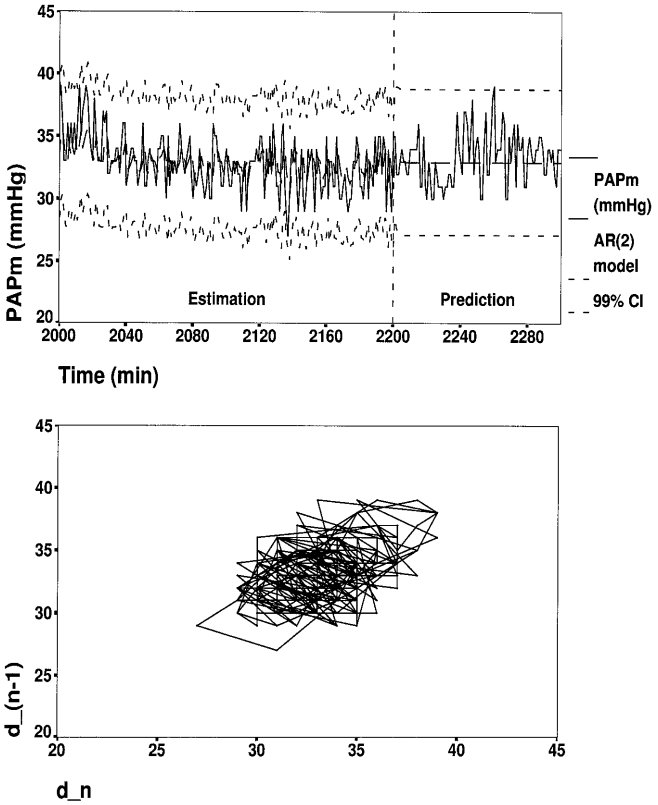

Fig. 2 **Top** Autoregressive model for a time series without significant change: AR model for pulmonary artery pressure *PAPm.* All data points for the prediction period are within the 99% CI. *Solid line* Measured series of PAPm; *dashed line* fitted AR(2) model of PAPm; *dotted lines* 99% confidence interval for AR(2) model of PAPm; *time* time after start of first measure in minutes. **Bottom** Phase space model for the same time series: PS model for PAPm. All vectors of the PS model are within an imaginary elippsoid. *d_n* value for PAPm at time "n"; d_(n – 1) value for PAPm at time "n – 1" (i.e., one observation prior to d_n)

was decided if a pattern was present similar to the retrospective analysis.

A detailed description of the concepts of autoregressive and phase space models and the underlying statistics is given in appendix 1.

Results

With autoregressive models, both outliers and level changes could always be identified. All series that were classified as "without change" were also correctly identified by autoregressive models. Phase space models allowed the identification of series without any change and with outliers in every instance, too. Temporary level changes were correctly identified in 20 of 24 series, and permanent level changes in 37 of 42. In the instances where identification failed, the level changes were marked by a very slow decrease or increase of the observed values.

Figures 2 to 6 display typical examples of each pattern analyzed by the two different methods. Figure 2 shows the same series for mean pulmonary artery pressure, which was clinically classified as "no change", while in the AR model the confidence interval was adjusted to 99% to allow the correct classification. The graphic representation of the series without a change is obvious in the PS model and is confirmed by the phase space procedure.

Typical graphs for outliers are shown in a time series for heart rate in Fig. 3. As before, the 95% CI for the AR model was too close for the clinical situation. With a 99.9% CI, two outliers could be identified that extrude from the CI. The same outliers are graphically represented by two vectors in the PS model that protrude from the imaginary ellipsoid of the vector cloud.

The analysis of a time series for heart rate with a temporary level change is shown in Fig. 4. In the PS model the temporary level change is represented by a small secondary ellipsoid of vectors that extrude from and fall back into the main ellipsoid, when the series returns

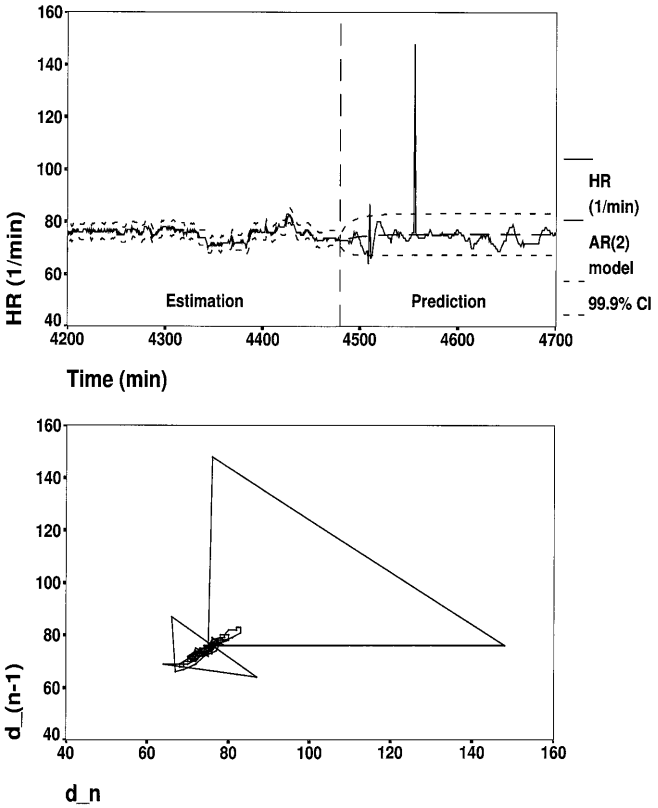

Fig. 3 Top Autoregressive model for a time series with an outlier: AR model for mean AP. Three data points (outliers) are outside the 99.9% CI. *Solid line* Measured series of heart rate *HR; dashed line* fitted AR(2) model of HR; *dotted lines* 99.9% confidence interval for AR(2) model of HR; *time* time after start of first measure in minutes. **Bottom** Phase space model for the same time series with an outlier: PS model for mean AP. Few vectors extrude from the imaginary ellipsoid. Each vector representing an outlier. *d_n* value for HR at time "n"; *d_(n – 1)* value for HR at time "n – 1" (i.e., one observation prior to d_n)

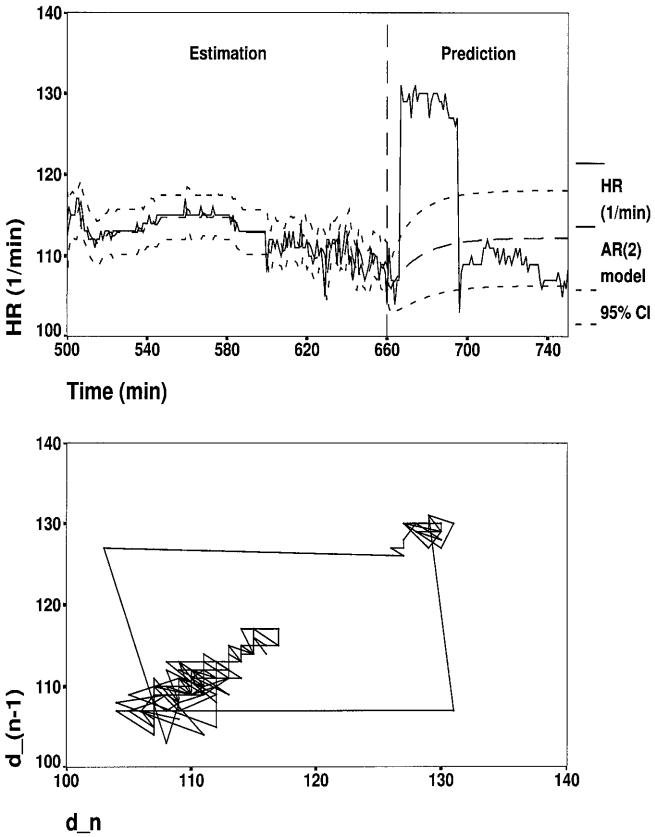

Fig. 4 Top Autoregressive model for a time series with a temporary level change: AR model for heart rate *HR*. A series of values is outside the 95% CI. Quantification with additional regressor. *Solid line* Measured series of HR; *dashed line* fitted AR(2) model of HR; *dotted lines* 99.9% confidence interval for AR(2) model of HR; *time* time after start of first measure in minutes. **Bottom** Phase space model for the same time series: PS model for HR. Several vectors extrude from the imaginary ellipsoid and form an additional ellipsoid, which falls back to the main ellipsoid. Changes cannot be quantified. *d_n* value for HR at time "n"; *d_(n – 1)* value for HR at time "n – 1" (i.e., one observation prior to d_n)

to baseline values. Similarly, the time series lies outside the 95% CI of the AR for the time of the temporary level change. In the PS model of a permanent level change the vectors from the secondary ellipsoid do not fall back into the main ellipsoid, as the time series in the AR model will not return to the baseline within the prediction period (Fig. 5).

The detection of trend cannot be done directly by either method. Trend detection required complete model diagnosis with AR models, as described in the methods section. After first order differentiation, the AR model was fitted to the time series in the estimation period. As shown in Fig. 6, due to the differentiation of the series the 95% CI widens rapidly after the start of the prediction period. This phenomenon precludes sensitive detection of changes during the prediction period (Table 2). With PS models a trend can only be detected by the oblong shape of the vector ellipsoid, which is a relatively insensitive method to distinguish between series without a trend and with a slight trend (Fig. 6).

Both methods were more sensitive to pattern changes than clinically relevant. Especially with outlier detection, 95% confidence intervals for autoregressive models were too close. In a second run, the confidence intervals were adjusted until clinically relevant results were found. This problem was most pronounced when the series had a very small variance during the estimation period. For those series where the CI was adjusted after initial analysis, the 95% CI was as close as an average of 9.7% on both sides of the fitted model with minimum values as low as 1.2% (Table 2). Although these small values are statistically significant, clinically they are not meaningful, as the small confidence intervals do not reflect therapeutically important changes and

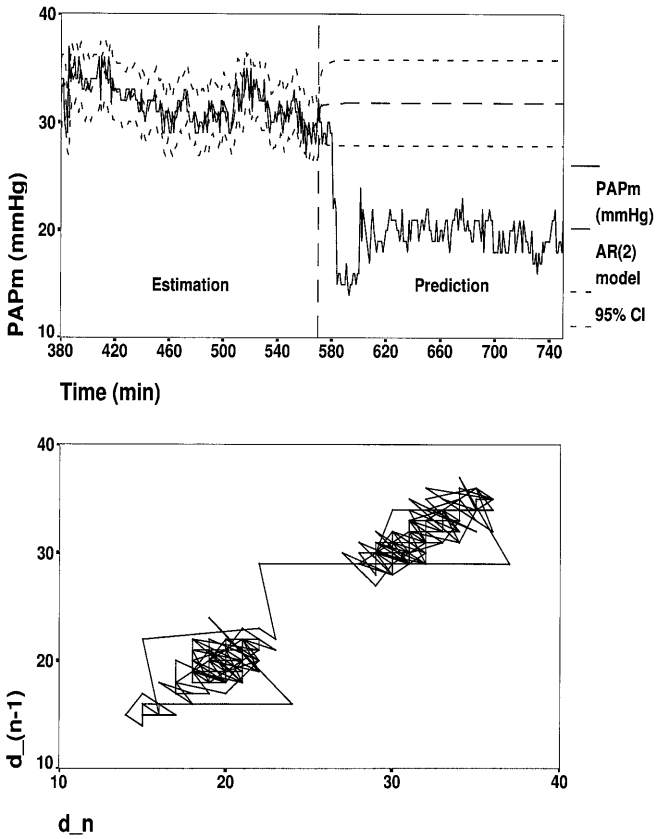

Fig. 5 Top Autoregressive model for a time series with a permanent level change: AR model for mean pulmonary artery pressure *PAPm.* All observations for the prediction period are outside the 95% CI. Quantification with additional regressor. *Solid line* Measured series of PAPm; *dashed line* fitted AR(2) model of PAPm; *dotted lines* 99% confidence interval for AR(2) model of PAPm; *time* time after start of first measure in minutes. **Bottom** Phase space model for the same time series: PS model for mean PAP. Numerous vectors extrude from the imaginary ellipsoid and form an additional ellipsoid, which does not fall back. Changes cannot be quantified. *d_n* value for PAPm at time "n"; *d_(n – 1)* value for PAPm at time "n – 1" (i.e., one observation prior to d_n)

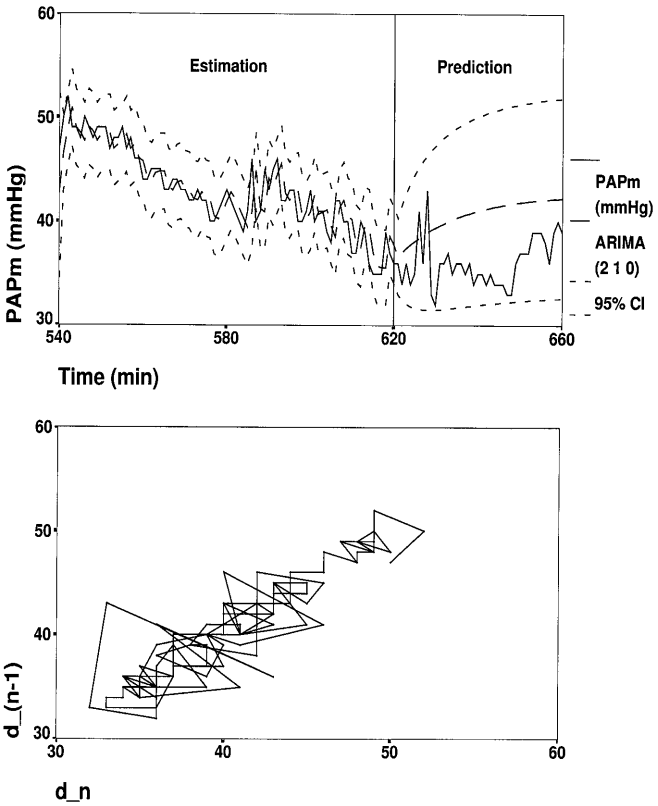

Fig. 6 Top Autoregressive model for a time series with a trend: AR model for mean pulmonary artery pressure *PAPm.* The trend can only be described indirectly by first order differentiation. The resulting model is an ARIMA(2 1 0) model. 95% CI very wide for prediction period due to differentiation. *Solid line* Measured series of PAPm; *dashed line* fitted AR(2) model of PAPm; *dotted lines* 99% confidence interval for AR(2) model of PAPm; *time* time after start of first measure in minutes. **Bottom** Phase space model for the same time series: PS model for PAPm. Vectors form a distorted ellipsoid without significant outliers. Trend can be diagnosed indirectly from the shape of the ellipsoid. *d_n* value for PAPm at time "n"; *d_(n – 1)* value for PAPm at time "n – 1" (i.e., one observation prior to d_n)

may even be smaller than the overall error of the method of measurement.

In five cases of outlier detection, the confidence intervals were adjusted to 99.99%. For a very sensitive detection of outliers, in some instances the CI was reduced to 90% (Table 3). In PS models, a level of 99.99% was chosen as well for all series.

Comparison between precisely diagnosed AR models and overdetermined models (AR order higher than necessary) showed that overdetermined models allow a "semi"automatic pattern detection without any tradeoff in clinical sensitivity. AR models require direct user interaction, while PS models offer opportunities for fully automated time series analysis in this context. Moreover, AR models require a priori assumptions about the approximate location of the disturbance of interest within the time series.

Discussion

In our study patterns of univariate physiologic time series could reliably be identified both with low order autoregressive models and phase space models. The only exception was trend patterns where both approaches have methodological shortcomings.

With on-line monitoring data that are sampled at very short time intervals, the number of available observations should not pose a problem for the application of time series analysis. But with on-line monitoring data,

Table 2 Relative confidence intervals *CI* for series with readjusted confidence intervals: initial CI 95 %. All figures for prediction periods of AR(2) models. Comparison of final CI after readjustment for different patterns. Values are % of final model values

Pattern of change	Relative CI (CI_{rel}) from fitted model (%)			No.
	Mean	Maximum	Minimum	
Final CI value: 90 %				
No change	22.33	26.14	18.51	2
Spike/artifact	5.29	5.29	5.29	1
Temporary level change	3.79	3.79	3.79	1
Permanent level change	13.12	24.63	5.80	6
Final CI value: 95 % (= initial CI value)				
No change	12.82	38.18	2.78	16
Spike/artifact	13.47	34.46	3.36	15
Temporary level change	9.56	23.91	2.74	12
Permanent level change	12.54	53.66	2.87	32
Trend	30.42	92.83	6.11	10
Final CI value: 99 %				
No change	7.49	14.37	3.47	4
Spike/artifact	9.80	23.59	3.14	4
Temporary level change	12.38	33.76	1.55	10
Permanent level change	10.95	13.08	7.86	4
Final CI value: 99.9 %				
No change	5.61	5.61	5.61	1
Spike/artifact	15.18	38.65	2.33	20
Temporary level change	4.74	4.74	4.74	1
Final CI value: 99.99 %				
Spike/artifact	17.78	23.90	13.58	5

reliable algorithms for artifact rejection have to be employed before the series can be subjected to statistical analysis. Robustness against artifacts and outliers is still a major problem with most time series methods [16, 32, 33]. For most bedside problems both methods are too sensitive. AR models seem to be better in this regard than PS models (Table 3). But a direct comparison is difficult because, in the estimation period of AR models, no pattern detection is performed. Thus, there is no possibility of misclassifying patterns in this period, whereas PS models look for patterns from the onset.

A possible task for further research is to replace the assumption of the normal distribution by probability distributions with more weight in the tails for the observed variables. This would reduce the sensitivity of the procedures. A disadvantage would be that estimation procedures are more complicated and would demand more computational effort. Another way to reduce the sensitivity is to use an automatically adjusted level. A low level should be chosen (i.e., less sensitive), if the variability of the process is small and vice versa.

Further research has to be done with multivariate time series analysis. The underlying processes for intensive care variables are probably multivariate or even high dimensional, because of the multitude of dependencies between physiologic variables. This could allow statements about the overall state of a patient or about an organ system. Automatic pattern detection in any situation and on-line application are not feasible at the moment. Moreover, AR models are especially highly interactive and both methods require a deep statistical knowledge in the user. On the other hand, autoregressive models and phase space models can be very helpful in the scientific off-line analysis of univariate intensive care monitoring data. They may support a more analytical and reproducible approach toward the evaluation of pathologic changes and therapeutic effects in the individual patient. The development of automatic methods for time series analysis would allow an instantaneous statistical analysis at the bedside. This would offer an option to the health care professional for a more reliable evaluation of the individual treatment.

Some time series analysis methods, such as the Kalman Filter, could also be used for on-line analysis of physiologic monitoring data. The generation of time series models including confidence intervals could enhance trend analysis, for not only could the slope of a trend be calculated but also outliers, which could represent clinically significant changes. Moreover, in the long run these techniques could be employed to generate smart alarms, that may be more reliable and less error prone than the simple limit alarms currently used.

Therefore, it appears that it may be rewarding to invest further efforts into the development of medical time series analysis techniques.

Appendix 1

Autoregressive models

A physiologic variable, e.g., heart rate, denoted by x_t, is observed at equidistant time points $t = 1, \ldots, N$. The set of observations $\{x_t\}_{t \in \aleph} = \{x_1, \ldots, x_N\}$ is called a time series. In the following, the data-generating process is modeled by an autoregressive (AR) process. An autoregressive process formally resembles a multiple regression. A stochastic process (X_t) is an autoregressive process of the order p [indicated by the notation AR(p)] where

$$X_t = \alpha_1 X_{t-1} + \ldots + \alpha_p X_{t-p} + \varepsilon_t, t \in \aleph, \alpha_1, \ldots, \alpha_p \in \Re$$

t are the time points of observation, $\alpha_1, \ldots, \alpha_p$ are weights measuring the influence of preceding values $X_{t-1}, \ldots, X_{t-p}$ on the value X_t, and ε_t is a white noise process. That means, ε_t is a sequence of uncorrelated variables from a fixed distribution with time dependent mean, usually assumed to be 0 and time invariant variance. Most applications of AR models assume ε_t to be normally distributed.

Conceptually, an autoregressive process is one with a "memory", in the sense that each value is correlated with preceding values. Following this interpretation, each value in an AR(p) process is determined by p preceding values, where older values will have a fading effect. Typically, low order AR processes ($p \leq 2$) are suitable to describe physiologic variables [8–10].

After completion of preliminary tests with classical interactive model selection [14], which showed that either first or second order models were statistically appropriate, second order autoregressive

Table 3 Results for AR and PS models. Pattern identification: rate of successful pattern detection. Estimation of sensitivity compared to clinical findings

	Results for autoregressive models					
	No change	Outlier	Level change		Trend	Total
			Temporary	Permanent		
Autoregressive models						
Pattern detected						
Yes/no	23/0	35/0	24/0	42/0	10/0[a]	134/0
Sensitivity						
Correct	16	27	17	32	10	102
Too high	7	8	7	10	0	32
Total	23	35	24	42	10	134
Phase space models						
Pattern detected						
Yes/no	23/0	35/0	20/4	37/5	0/10[b]	134
Sensitivity						
Correct	7	12	8	9	3	33
Too high	16	23	16	33	7	101
Total	23	35	24	42	10	134

[a] Indirect identification through analysis of the autocorrelation function of the residuals of the AR(2) model

[b] Mathematical theory does not allow trend detection (work in progress). But practical application can identify trend by the shape of the vector cloud

models were chosen for all time series. In cases where an AR(1) process is sufficient, an AR(2) process overdetermines the model. This leads to an unnecessary estimation of the parameter α_2, which will be close to 0. The additional computational effort is low, but it avoids an extensive model selection process. This model selection is also supported by earlier studies of time series analysis of intensive care data [9].

Each time series was split into two segments, an estimation period (observations $x_1, \ldots, x_n$) prior to the observed pattern and a prediction period including this pattern. A second order AR model (AR(2) model) was fitted to the data from the estimation period, for which the weights $\alpha_1, \ldots, \alpha_p$ had to be estimated. At the estimation stage the weights must be chosen according to some defined criterion. A criterion which requires relatively little computational effort is the least-squares criterion. Here, the values $\hat{\alpha}_1, \ldots, \hat{\alpha}_p$ that minimize the sum of squares

$$\sum_{t=p+1}^{n} [x_t - \alpha_1 x_{t-1} - \ldots - \alpha_p x_{t-p}]^2 = \sum_{t=p+1}^{n} [e_t]^2$$

are defined to be the least squares estimates of $\alpha_1, \ldots, \alpha_p$. The $e_t = x_t - \hat{\alpha}_1 x_{t-1} - \ldots - \hat{\alpha}_p x_{t-p}$ are called estimated residuals. They can be regarded as the estimates for the ε_t and they denote the deviation in the prediction of the actually observed value x_t with the estimated value $\hat{x}_t = \hat{\alpha}_1 x_{t-1} + \ldots + \hat{\alpha}_p x_{t-p}$. Obviously the cumulative deviations should be as low as possible, as expressed in the above formula. On the basis of the estimated weights, confidence intervals for the estimation period as well as for the prediction period can be constructed [14].

The estimated model was tested for goodness of fit with the plot of the autocorrelation function (ACF plot) of its residuals. The autocorrelation function describes the correlation between time delayed observations x_t and $x_{t-\tau}$, $t, \tau \in \aleph$:

$$ACF(\tau) = Corr(x_t, x_{t-\tau}) = \frac{\sum_{t=1}^{N-\tau} (x_t - \bar{x}_t)(x_{t-\tau} - \bar{x}_t)}{\sum_{t=1}^{N} (x_t - \bar{x}_t)^2}, t, \tau \in \aleph$$

where $\bar{x}_t = \sum_{t=1}^{N} x_t$ is the estimated mean of the time series. Because the errors ε_t and $\varepsilon_{t+\tau}$ are assumed to be independent, their estimators (the residuals) should be independent for all time lags ($ACF(\tau) = 0$). Other tests of autocorrelation in the residuals (Box–Ljung, Durbin–Watson and the root mean square error) were also applied [14].

In some cases a difference between the visual classification and the percentage of observations identified as outliers by the model occurred if a 95 % CI was chosen. One possible reason for this difference is a temporary violation of model assumptions like stationarity or the Gaussian distribution of the observations. The other important reason is that a fixed level for the CI cannot adapt to the extent of the process variability. If the level is fixed and the variability is small, the CI is small as well. Observations outside such a CI may be clinically irrelevant. To avoid false classifications, the chosen CI level had to be adjusted from an initial 95 % to 99, 99.9, 99.99, or in some cases to 90 %. The pattern identification was rerun with this adjusted CI.

Phase space models

Phase space models are based on a transformation of time series in a Euclidean space. This transformation is called Phase Space Embedding, a technique derived from the theory of nonlinear dynamic systems. Given a time series (x_t) with N observations, Packard et al. [34] and Takens [35] constructed so-called phase space vectors $\vec{x}_t$, which are defined by:

$$\vec{x}_t = (x_{t+m-1}, \ldots, x_{t+1}, x_t)', \vec{x}_t \in \Re, t = 1, \ldots, N - m + 1, m \in \aleph \backslash 0$$

Here, m is called embedding dimension. Numerous rules exist for choosing m in nonlinear models. In most cases the components of the phase space vectors are not neighboring observations, but rather then are separated by a time delay [36, 37]. Focusing on Gaussian processes, m is chosen in analogy to the order of an AR(p) model. The components of $\vec{x}_t$ are chronological observations with a time delay (lag) always of 1. This improves the identification of patterns in Gaussian processes as dependencies between consecutive observations are taken into consideration.

For the example of $m = 2$ the set of vectors $\vec{x}_t = (x_{t+1}, x_t)'$ can be plotted in a two-dimensional space. This vector cloud is called Phase Space Reconstruction. With this reconstruction it is possible to visualize properties of the underlying dynamic. In Fig. 2 a two-dimensional reconstruction of a time series is given. In this case, the underlying system is in a steady state. Where chronological observations are combined in order to show the movement through

space, the vectors form an elliptic cloud. The geometry of the time series reflects the dependency structure of the underlying process. This can be formalized. An ellipse, which is estimated with statistical methods from the (contaminated) data, is placed around the vector cloud [16].

The estimation of the ellipse is based on the assumption that the two-dimensional vectors $(X_{t+1}, X_t)'$, constructed from the process variables are bivariate normally distributed

$$(X_{t+1}, X_t)' \sim N(\vec{\mu}, \Sigma), \vec{\mu} = (\mu, \mu)', \sum = \begin{bmatrix} \sigma^2 & \gamma \\ \gamma & \sigma^2 \end{bmatrix}$$

Analogous to the univariate case the bivariate distribution has a mean vector $\vec{\mu}$. The parameter matrix Σ describes the dependency structure of the process (the elliptic form of the embedding). In practice, μ, σ^2, and γ are unknown and have to be estimated from the time series data. The most convenient way is to use classical estimators

$$\hat{\mu} = \bar{x}_t, \hat{\sigma}^2 = ACF(0), \hat{\gamma} = ACF(1)$$

These estimators can be replaced by robust estimators [16], so that outliers in the data have little influence on the parameter estimation.

It is obvious that the distance of the vector $\vec{x}_t$ from the mean vector (the steady state) $\vec{\mu}$ gives information about suspicious observations. Using the Euclidean norm $\|\vec{x}_t\| = \sqrt{x_{t+1}^2 + x_t^2}$ is not adequate here, because it does not consider the orientation of the observation from the mean vector. Thus a weighted distance, the Mahalanobis distance (MD) has to be used for identifying suspicious observations at time t and for estimating the ellipse

$$MD_t = \sqrt{(\vec{x}_t - \mu)' \sum^{-1} (\vec{x}_t - \mu)}$$

It is well known in the statistical literature that MD_t^2 is asymptotically chi-square distributed with two degrees of freedom ($MD_t^2 \sim \chi_2^2$). The set of vectors $\vec{x}_t$ given by

$$\{\vec{x}_t \mid (\vec{x}_t - \mu)' \Sigma^{-1} (\vec{x}_t - \mu) = \chi^2_{2,1-\alpha}\}$$

forms an imaginary ellipse around the mean vector $\vec{\mu}$, where $\chi^2_{2,1-\alpha}$ is the $1 - \alpha$ quantile of a chi-square distribution with two degrees of freedom for a given level α. In practice, μ, σ^2, and γ have to be replaced by their estimators. All observations inside this ellipse are in accordance to the steady state, whereas observations extruding from this ellipse are outliers or affected by level changes.

The size of the ellipse is determined by the given level α. This is the probability that the procedure falsely identifies one or more values as outlying observations, if model assumptions are valid. To ensure that this probability statement is valid, it is necessary to adjust the level, because every single observation has to be examined whether it extrudes from the ellipse. The adjusted level α_N is given by $\alpha_N = \dfrac{\alpha}{2(N - m + 1)}$, i.e., $\alpha_N = 0.000085$ with $\alpha = 0.01$ and $N = 60$ [16]. If a phase space vector lies outside the estimated ellipse the probability that it actually belongs to the steady state is smaller than $0.00085 \approx 0.0001$.

All observations of the time series of Fig. 2 lie inside such an estimated ellipse. In this case, it can be determined that the system is in a steady state and that no specific pattern is present. If one or more vectors leave this ellipse, a disturbance of the dynamic can be assumed. It is possible to distinguish several disturbances by the movements of the affected vectors. Figures 3–6 display typical examples for the patterns, which were investigated in this study.

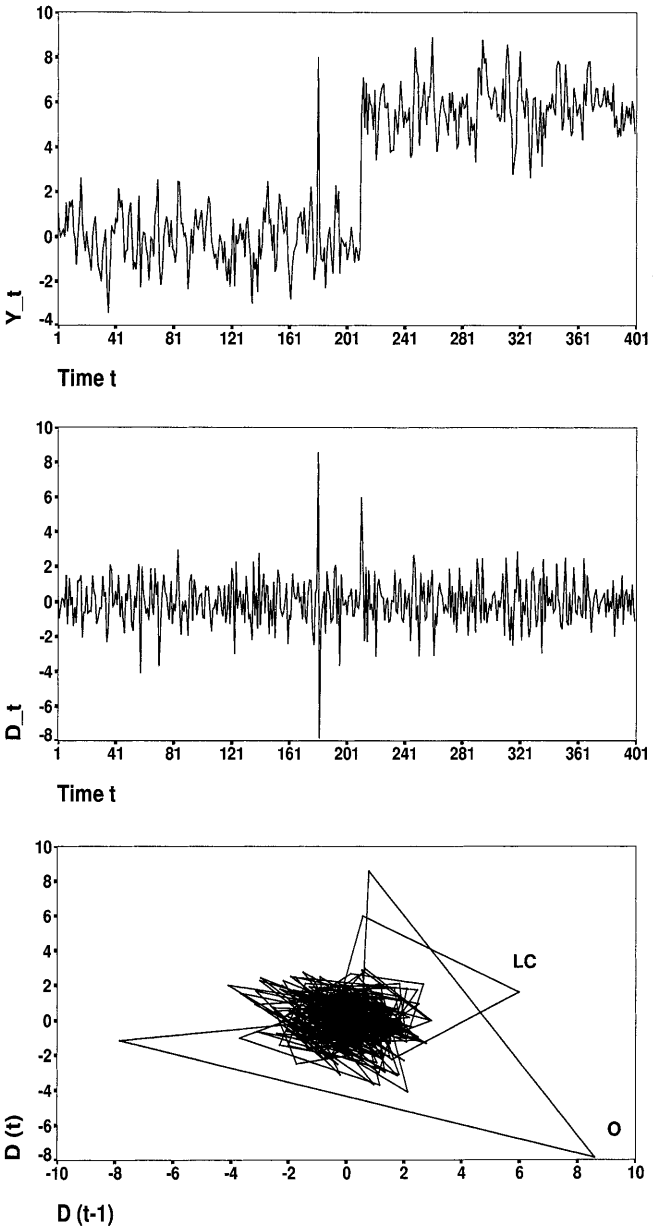

Fig. 7 Phase Space Embedding of a simulated time series: **Top** simulated AR(1) process with outlier and level change; **middle** differenced AR(1) process with outlier and level change; **Bottom** phase space embedding of the differenced series (*O* outlier, *LC* level change). Plot of d_n versus d_(n – 1). After a change occurred the phase space vectors fall back into the ellipse describing the steady state

The identification procedure that we developed uses the differenced time series d_t, which is $d_t = \gamma_t - \gamma_{t-1}$, $t = 1, \ldots, N$. In a differenced series an abrupt level change will be represented by one outlier. The procedure focuses on the identification of these observations. On the basis of the movement of the phase space vectors, which contain such observations, a discrimination between different patterns is done. The vectors $\vec{d}_t$, $t = 2, \ldots, N$ were analyzed in consecutive order whether they pointed into a distant region. If a

vector lies in a distant region, i.e., the vector extrudes from the cloud, it can be discriminated between different patterns after observing further values (a detailed description of the methodology is given in Bauer et al. [16]).

A graphic example may show the general approach of this methodology. In Fig. 7 a simulated time series following an AR(1) process, the differenced series, and the two-dimensional embedding of the differenced series with the corresponding estimated ellipse are shown. This example discriminates only between the patterns of outlier and abrupt level change. If at the time point t the vector $\vec{d}_t$ extrudes from the cloud, the decision between outlier and abrupt level change takes place at the time point $t + 1$. If the distance between $\vec{d}_t + 1$ and the fixed point LC in Fig. 7 is smaller than the distance between $\vec{d}_t + 1$ and the fixed point O, then a level change can be diagnosed, otherwise, an outlier is present. If more patterns are considered, more fixed points have to be determined and the time point at which the decision takes place is delayed.

Acknowledgements The authors wish to thank Mrs. Penelope Hood Greco and Mr. James Greco, Kittery, MA, USA, for their invaluable help with the English manuscript.

References

1. Sandefur JT (1990) Discrete dynamical systems, theory and applications. Clarendon Press, Oxford
2. Deistler M (1989) Linear system identification – a survey. In: J Willems (ed) From data to model. Springer, Berlin Heidelberg New York, pp 1–25
3. Oppenheim AV, Schafer RW (1995) Discrete-time signal processing. Prentice Hall, Englewood Cliffs, NJ
4. Gordon K, Smith AFM (1990) Modeling and monitoring biomedical time series. J Am Stat Ass 85: 328–337
5. Hill DW, Endresen J (1978) Trend recording and forecasting in intensive care therapy. Br J Clin Equipment: January, 5–14
6. Balvers RJ, Cosimano FF, McDonald B (1991) Predicting stock returns in an efficient market. J Finance 45: 1109–1128
7. Krämer W, Runde R (1996) Stochastic properties of German stock returns. Empirical Econ 21: 281–306
8. Imhoff M, Bauer M (1996) Time series analysis in critical care monitoring. New Horiz 4: 519–531
9. Imhoff M, Bauer M, Gather U, Löhlein D (1997) Time series analysis in intensive care medicine. Appl Cardiopulmonary Pathophysiol 6: 263–281
10. Lambert CR, Raymenants E, Pepine CJ (1995) Time-series analysis of long-term ambulatory myocardial ischemia: effects of beta-adrenergic and calcium channel blockade. Am Heart J 129: 677–684
11. Gordon K (1986) The multi-state Kalman Filter in medical monitoring. Comput Methods Programs Biomed 23: 147–154
12. Smith AFM, West M (1983) Monitoring renal transplants: an application of the multi-process Kalman filter. Biometrics 39: 867–878
13. Smith AFM, West M, Gordon K, Knapp MS, Trimble IM (1983) Monitoring kidney transplant patients. Statistician 32: 46–54
14. Box GEP, Jenkins GM, Reinsel GC (1994) Time series analysis, forecasting and control, 3rd edn. Prentice Hall, Englewood Cliffs, NJ
15. Schlittgen R, Streitberg BHJ (1994) Zeitreihenanalyse. R Oldenbourg, München
16. Bauer M, Gather U, Imhoff M (in press) Analysis of high dimensional data from intensive care medicine. In: Payne R (ed) Proceeding in computational statistics, Springer, Berlin Heidelberg New York
17. Lim KL, Muir K, Powell RJ (1994) Urine neopterin: a new parameter for serial monitoring of disease activity in patients with systemic lupus erythematosus. Ann Rheum Dis 53: 743–748
18. Gerber WD, Schellenberg R, Thom M, Haufe C, Bolsche F, Wedekind W, Niederberger U, Soyka D (1995) Cyclandelate versus propranolol in the prophylaxis of migraine – a double-blind placebo-controlled study. Funct Neurol 10: 27–35
19. Marrazzi MA, Markham KM, Kinzie J, Luby ED (1995) Binge eating disorder: response to naltrexone. Int J Obes Relat Metab Disord 19: 143–145
20. Hepworth JT, Hendrickson SG, Lopez J (1994) Time series analysis of physiological response during ICU visitation. West J Nurs Res 16: 704–717
21. Faustmann PM, Ganz RE (1994) Central cardio-autonomic disorganization in interictal states of epilepsy detected by phase space analysis. Int J Neurosci 78: 43–47
22. Garfinkel A, Spano ML, Ditto WL, Weiss JN (1992) Controlling cardiac chaos. Science 257: 1230–1235
23. Khadra LM, Maayah TJ, Dickhaus H (1997) Detecting chaos in HRV signals in human cardiac transplant recipients. Comput Biomed Res 30: 188–199
24. Fukunaga K (1990) Introduction to statistical pattern recognition, 2nd edn. Academic Press, New York
25. Kalman RE (1960) A new approach to linear filtering and prediction problems. J Basic Eng, Transactions ASME, Series D 83: 95–108
26. Daumer M, Falk M (1998) On-line change-point detection for state space models using multi-process Kalman filters. In: O'Leary D (ed) Proceedings of the International Linear Algebra Society Symposium on "Fast algorithms for control, signals and image processing". Elsvier, Amsterdam, pp 125–135
27. Stronegger WJ (1991) Kalmanfilter zur Online-Analyse von Verlaufskurven. In: Gilg U, Seeber H (eds) Multivariate Modelle. Springer, Berlin Heidelberg New York
28. Navendra KS, Parthasarathly K (1990) Identification and control of dynamic systems using neural networks. IEEE Transactions on Neural Networks 1: 4–27
29. Rashima YH, Ilguni Y, Adachi N (1997) An adaptive control design using a memory based learningsystem. Int J Control 68: 1085–1102
30. Chan EHP (1989) Application of neural-network computing in intelligent alarm processing. In Proceedings of the IEEE power industry computer application conference, Seattle, Washington, May 1–5, PICA '89. IEEE, New York, pp 246–251
31. Kirschen DS, Bruce F, Wollenberg BF (1992) Intelligent alarm processing in power systems. Proceedings of IEEE 80: 663–672
32. Gather U (1990) Modeling the occurrence of multiple outliers. Allg Statistisches Archiv 74: 413–428
33. Gather U, Kale BK (1992) Outlier generating models – a review. In: Venugopal N (ed) Contributions to stochastics. Wiley (Eastern), New Delhi, pp 57–85
34. Packard NH, Crutchfield JP, Farmer JD, Shaw RS (1980) Geometry from a time series. Phys Rev Lett 45: 712–716
35. Takens F (1980) Detecting strange attractors in turbulence. In: Rand DA, Young LS (eds) Dynamical systems and turbulence. (Lecture notes in mathematics, vol 898) Springer, Berlin Heidelberg New York, pp 366–381
36. Isham V (1993) Statistical aspects of chaos: a review. In: Barndorff-Nielsen OE, Jensen WS, Kendall WS (eds) Networks and chaos – statistical and probabilistic aspects. Chapman & Hall, London, pp 251–300
37. Tong H (1995) A personal overview of non-linear time series analysis from a chaos perspective. Scand J Statist 22: 399–445