Detection of Methacholine with Time Series Models of Lung Sounds

Piet M. T. Broersen and S. de Waele

Abstract—A new method for the extraction of features from stationary stochastic processes has been applied to a medical detection problem. It illustrates a practical application of automatic time series modeling. Firstly, the model type and the model order for two time series prototype models are selected. The prototypes represent the lung noises of a single healthy subject, before and after the application of methacholine. Using the model error ME as a measure for the difference between time series models, new data can be divided into classes that belong to the prototype models for this person. The prototype models are obtained from a few expiration cycles under known conditions. This is sufficient to detect the presence of methacholine in new data of the same subject if he is able to maintain stationary conditions by following accurately the prescribed breathing pattern. It is not necessary to use the same model type and the same model order for the prototypes and for new data. Automatically and individually selected models for prototypes and data give a good detection of methacholine.

Index Terms—Detection, model error, prediction error, prototype model, spectral estimation.

I. INTRODUCTION

ASTHMA is a chronic disease with variable airway obstruction in which the airways are characterized by inflammatory changes. The association between lung sound characteristics and asthma has long been recognized. It is clear that heavy asthma can be detected rather easily by simply listening. The final goal of this research, however, might be the early detection of asthma by an automatic computer analysis of recorded lung sound signals. Using computerized phonopneumography, lung sounds have been analyzed in intensity and frequency content [1]. As expected, some differences are observed between healthy and asthmatic subjects. However, also variations between different subjects within one class are found that hamper the extraction of pure asthmatic features. By inducing acute airway narrowing using methacholine, an increase in sound intensity and frequency content has been found for a healthy subject, as well as a change in the wheezing of asthmatics [2]. Therefore, the characteristics of the lung sounds of a single person can be varied with methacholine in a way that simulates asthmatic features.

The purpose of this paper is to demonstrate the use of time series models for feature extraction in a medical context. This is also a validation of the use of theoretical results for time series which have been obtained for *stationary stochastic processes*

[3] on practical observations on humans. A first step toward detection of asthma is to investigate detection of the presence of methacholine in easily observable properties of lungs, for which the lung sounds have been used. The paper describes a method for the automatic detection of methacholine from measurements with a microphone placed at a precisely specified position on the chest. As test data, observations of the lung noises of one person with and without methacholine are available. New data come from the same person, both with and without methacholine, but recorded a few days later.

Autoregressive (AR) models have been used before to characterize lung sounds [4], [5]. They are one type of the wider class of time series models, including AR, Moving Average (MA), and combined ARMA models [6]. Theoretically, at least one of those three model types gives a good spectral model for unknown stationary stochastic processes [6]. By using robust methods in estimating the three model types, it is possible to select a single model **type**, AR or MA or ARMA [3]. Also the best model **order** can be selected from hundreds of candidate orders. This automatically selected model of finite order gives a description of the statistically significant details in the data [3], leaving out spurious details that are not significant.

The model error ME has been introduced as a measure for the accuracy of time series models [7], [8]. It compares results of different algorithms with the characteristics of the true process, as a criterion to evaluate estimation algorithms. This scalar measure can also be used to characterize the difference between the measured time series model of new data and previously determined prototype models. The new data model is obtained from microphone measurements under unknown conditions. Two prototype models are used; one prototype is the model in normal conditions and the second prototype is the model found after application of methacholine. The results show that the presence of methacholine is detected very well for the person in this test.

II. MEASUREMENTS

Records have been made from eight healthy nonsmoking normal male adults, between 19 and 28 years of age. All subjects underwent lung function tests before the actual recordings. From each subject, two similar data sets were recorded, at an interval of two or three days. The first measurement on each day was in the normal situation. The second measurement took place after applying methacholine, in a concentration such that the FEV, the Forced Expiration Volume, the flow of air forced out after maximal inhalation, was diminished by 25%.

The subjects had to follow a prescribed breathing pattern: 1.5 l/s during inspiration and expiration between fixed max-

imum and minimum volumes. Flow and volume are measured with a spirometer. Lung sounds were obtained from a piezo-electric air-coupled microphone, at the second intercostal mid-clavicular line. This signal was high-pass and low-pass filtered with 4th order Bessel filters at 100 Hz and 1500 Hz, respectively. Afterwards, the signal was digitized with a sampling frequency of 5000 Hz, with 12-bit resolution. A peculiar phenomenon was found in all digitized microphone signals, a very sharp peak at 1500 Hz, exactly the cutoff frequency of the low-pass filter. This could not be explained from a biological point of view. The small hardware error has mostly been left in the signals which were analyzed, because sometimes the removal by decimating the data turned out to be disadvantageous for detection. More-over, a comparison with previous studies of the same data [1], [2] remains possible. Synchronization of microphone data with flow measurements enables the detection of the transitions from expiration to inspiration.

The records have been analyzed before [1], [2]. It turned out to be rather difficult to discover reproducible effects of metha-choline, that can be used for its detection. The limited purpose of this paper is to investigate the feasibility of *automatically* de-tecting the presence of methacholine with time series models of lung sounds. Automatically means that no use is made of avail-able *a priori* knowledge about data properties from present or previous research [1], [2], [4], [5]. Only expiration data from one subject have been used. Without further detailed investiga-tions, it has been decided to use the first 1000 observations of the expiration phase. It is possible that other parts of the respira-tion cycle give better detection results. Also further filtering the data or resampling has not been studied in detail. Those opera-tions, if used optimally, may have a favorable influence on the possibilities of feature extraction and detection.

III. TIME SERIES MODELS

Estimation *theory* for time series often uses the notions *process* and *models* for the true signal properties and for the properties obtained from data, respectively. If only *practical* data are available, some adaptations will be necessary because the theoretical signal properties are not known then. Suppose that x_n are data which are generated by an ARMA(p, q) *process*, given by [6]

$$\mathbf{A(z)x_n = B(z)\varepsilon_n} \tag{1}$$

where ε_n is a white noise sequence with variance σ_ε^2 and the AR (autoregressive) polynomial $A(z)$ is defined as: $A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_p z^{-p}$. In this notation, $z^{-1} x_n = x_{n-1}$. Likewise, an MA (moving average) polynomial $B(z)$ is defined as a polynomial of order q in z^{-1}. The process is purely AR if $B(z) = 1$ and MA if $A(z) = 1$, so with $p = 0$. The process is stationary if $A(z)$ has no roots or poles for z on or outside the unit circle. Likewise, it is invertible if the roots of $B(z)$, denoted zeros, are all inside the unit circle. Theoretically, any stationary stochastic process can be written as an AR(∞) or an MA(∞) process [6], but also ARMA processes are possible. In practice, a suitably selected finite order model of one of the three model types will give a good description [3].

Models can be divided in two classes: models that are es-timated from the data at hand and models that come from an origin that is statistically independent from the present data. In the first class, the estimated parameters are statistically depen-dent on the data. With model parameters $\hat{A}(z)$ and $\hat{B}(z)$, obser-vations x_n can be filtered

$$\mathbf{\hat{B}(z)\hat{\varepsilon}_n = \hat{A}(z)x_n}. \tag{2}$$

This can be interpreted as filtering the measured data x_n to ob-tain $\hat{\varepsilon}_n$ as output signal. If the model parameters have been es-timated from the data x_n in (2), $\hat{\varepsilon}_n$ is called a residual; if the model is independent from the data x_n in (2), $\hat{\varepsilon}_n$ is called a prediction. Typically, the variance of $\hat{\varepsilon}_n$ is smaller than σ_ε^2 for residuals and greater than σ_ε^2 for predictions. The squared Pre-diction Error PE is defined as the variance of $\hat{\varepsilon}_n$ in (2), with the condition that x_n is independent of $\hat{A}(z)$ and $\hat{B}(z)$. For prac-tical data x_n, this variance can be estimated with (2) by filtering. However, if the true process is known as in (1), the output $\hat{\varepsilon}_n$ of the model with $\hat{A}(z)$ and $\hat{B}(z)$ can also be determined by sub-stituting the formula (1) for x_n in (2)

$$\mathbf{\hat{\varepsilon}_n = \frac{\hat{A}(z)}{\hat{B}(z)}x_n = \frac{\hat{A}(z)B(z)}{A(z)\hat{B}(z)}\varepsilon_n}. \tag{3}$$

Therefore, the relation between the output $\hat{\varepsilon}_n$ of the model (3) and the innovations ε_n that generated the true process is given by an ARMA$(p + q', p' + q)$ process. The variance PE of that process can easily be expressed in σ_ε^2, $A(z)$, $B(z)$, $\hat{A}(z)$, and $\hat{B}(z)$ [7]. Prototype models from test data are considered as in-dependent polynomials $\hat{A}(z)$ and $\hat{B}(z)$ in (2) and (3); new de-tection data are denoted $A(z)$ and $B(z)$. A point of investigation will be whether it is necessary to have the same type and order for prototypes and data models or that they can have different, selected, types and orders.

AR, MA, and ARMA models must be estimated for many or-ders before a choice can be made for new data [3]. Models of the unknown best order must be computed, but also models of higher orders. Only then, it is possible to conclude that lower orders are better and should be selected. However, maximum likelihood methods for MA and ARMA estimation are non-linear [6]. Maximum likelihood algorithms require initial con-ditions and must iterate for convergence. This is a known cause of problems in the practice of nonlinear time series estimation, especially for high-order models. For an AR(p) model, there is only one AR$(p + 1)$ model that gives the same fit. That is the model with parameter zero for order $p + 1$. However, for every ARMA(p, q) model that gives a good fit, there are infin-itely many ARMA$(p + 1, q + 1)$ models with the same fit: all models where an additional pole and zero cancel. Therefore, the variance of the parameters of those over-complete models will be ∞. The variance becomes so large that roots of estimated AR and/or MA polynomials can fall outside the stationary or invert-ible region, with the boundary being the unit circle. Some exam-ples of robust algorithms have been described [3], [9]; possible good choices for AR, MA, and ARMA algorithms are discussed briefly.

Burg's estimation method [10] can be recommended for AR models; CIC [11] can be used as a criterion for order selection. Other AR order selection criteria can give wrong orders if the candidate orders become higher than 0.1 N. Taking the actual finite sample behavior of the logarithm of the residual variance into account, the Combined Information Criterion CIC is based on the expectation and on the variance of the logarithm of the residual variance, as a function of the model order [11].

The Durbin's method for MA [12] and for ARMA [13] estimation consists of the use of the parameters of a long intermediate autoregressive model to compute MA parameters. In this way, nonlinear estimation is approximated by a sequence of linear operations. Durbin's algorithms always converge and produce an invertible solution. A recent improvement in the Durbin's method is that a new optimal order has been defined for the intermediate AR model. Hence, the performance of Durbin's algorithms in practice has been improved [14], [15]. In many examples with reasonable sample sizes, the Cramér-Rao lower bound for the parameter accuracy is approximated in estimated MA and ARMA models if this new intermediate AR order is used.

The Matlab routine ARMAsel [9] for the automatic evaluation of stationary stochastic data is as follows:

- AR models are estimated and an AR order p' is selected;
- MA models are estimated and MA order q' is selected;
- ARMA$(r, r-1)$ models are estimated and an order r' is selected;
- from those three selected models of different types, a single one is chosen with a criterion that is based on the estimate of the prediction error of the AR(p'), MA(q') and ARMA$(r', r'-1)$ models.

This final model, with *order and type selected* automatically for given observations, is denoted the ARMAsel model [9].

The prediction error PE is defined as the model fit to new and independent data, or the variance of the one step ahead error of prediction $\hat{\varepsilon}_n$ in (3). It is an obvious measure for the accuracy of time series models. Its asymptotical expectation for $p' \geq p$ and $q' \geq q$, for unbiased models which have at least all true nonzero parameters included, equals

$$\mathbf{PE} = \sigma_{\varepsilon}^2 \left(1 + \frac{\mathbf{p' + q'}}{\mathbf{N}} \right) \qquad (4)$$

where $p' + q'$ is the number of estimated parameters, N is the number of observations, and σ_{ε}^2 is the innovation variance of (1). The AR model with the minimum of PE also has an interpretation in the frequency domain as the model with the highest measure for the spectral flatness [16]. The scaling factor is a constant contribution σ_{ε}^2 which gives no relevant information about the quality of different models for the same process. Therefore, the model error ME is defined [7] as a scaled version of the excess prediction error due to the combination of model selection and parameter estimation

$$\mathbf{ME} = \mathbf{N}(\mathbf{PE} - \sigma_{\varepsilon}^2)/\sigma_{\varepsilon}^2 = \mathbf{N}(\mathbf{PE}/\sigma_{\varepsilon}^2 - 1) \qquad (5)$$

where σ_{ε}^2 is the variance of the innovations that are the excitations of the time series process (1). ME can easily be computed in the time domain [7], without needing an estimate for
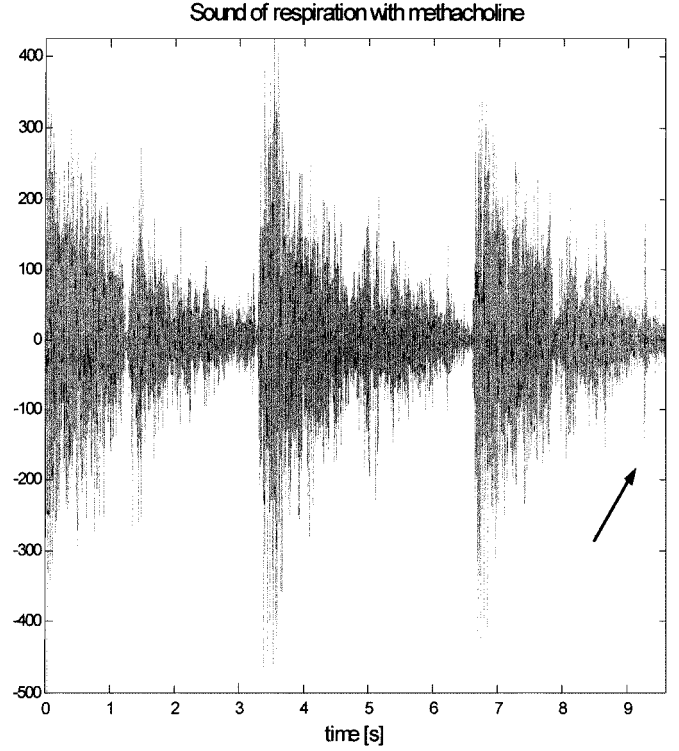

Fig. 1. Observations of three cycles of respiration sounds with methacholine, starting with inspiration that has a higher amplitude. The first transition to expiration is at 1.4 s, the second at 4.8 s.

σ_{ε}^2. The asymptotical expectation of ME for unbiased models is the number of estimated parameters, independent of the sample size. Hence, the asymptotical expectation of ME for an unbiased ARMA(p', q') equals $p' + q'$ and is independent of N, σ_{ε}^2 and of the true parameters [7]. The ME of unbiased models is independent of the properties of the true process, which has advantages for the interpretation. Only statistics determine what happens for models above the true process order, and not the characteristics of the true process, which will generally be unknown in applications.

The time-domain computation of ME requires only the parameters of two models that are to be compared. In the original derivation with the prediction error, it has been used for assessing the quality of estimation algorithms for known processes. Another new application of ME is to measure the difference between two models. For detection, ME will be computed with prototype models substituted for $\hat{A}(z)$ and $\hat{B}(z)$ and with the model parameters estimated from new data x_n for $A(z)$ and $B(z)$.

IV. FEATURE EXTRACTION WITH ARMA MODELS

Fig. 1 shows a registration of three respiration cycles. At first sight, it is clear that the variance of the sound is not constant during one cycle: the level is different for inspiration and for expiration. The variance also diminishes more or less gradually from the beginning until the end of the cycle. The arrow in the third cycle of Fig. 1 indicates the occurrence of a sudden increase in the sound level. This might also be present in the second cycle, but certainly not in the first cycle where a gradual
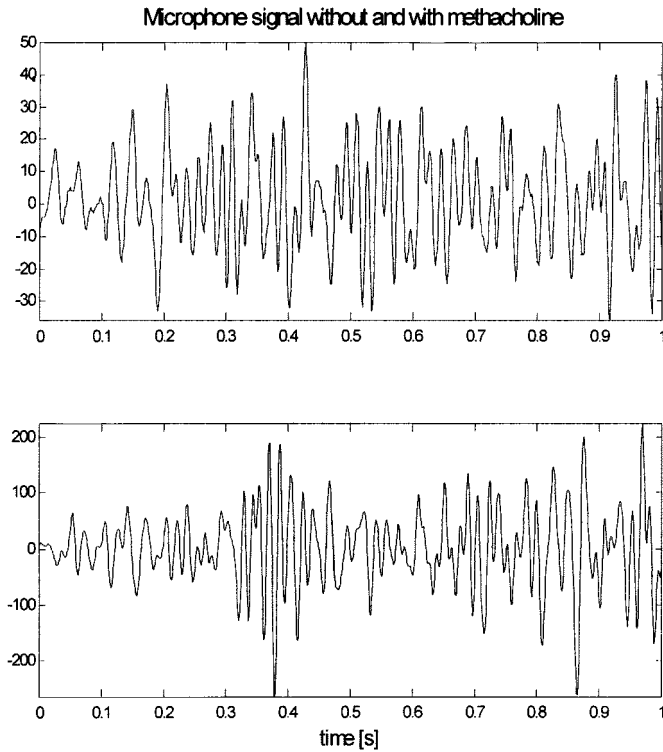
Fig. 2. Observations of the start of an expiration cycle for signals without and with methacholine, respectively. The ME between ARMAsel models is 951 if the model with methacholine is used as prototype.
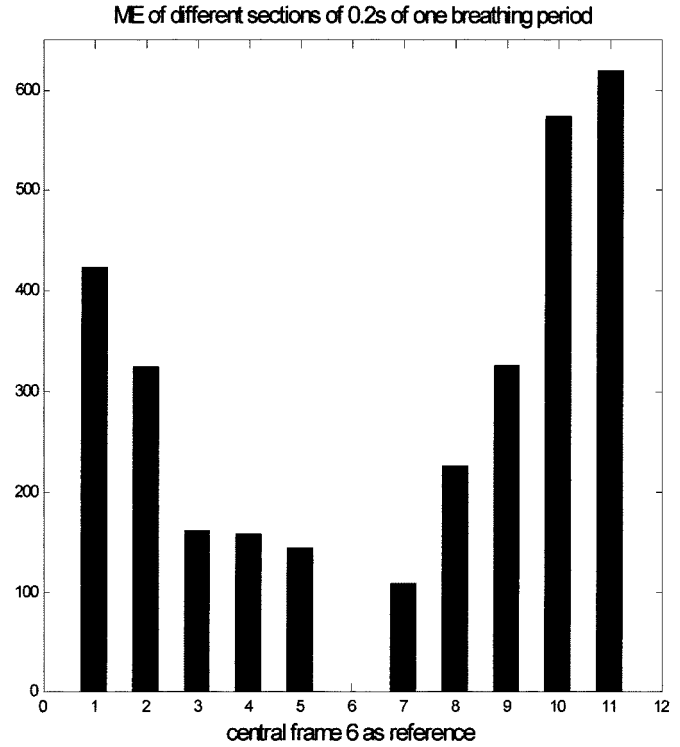

Fig. 3. The ME of different consecutive frames of 1000 observations of a single breathing period, with methacholine. Frame 1 starts at 3.8 s in Fig. 1. The first five frames are during inspiration, the last six frames during expiration.

increase of the sound level toward the end of the expiration phase is seen. Using isolated phenomena like this short burst for detection purposes has not been successful. It is clear that Fig. 1 is definitely not coming from a stationary process. As time series analysis has its theoretical background in *stationary* stochastic processes, it will be wise to use only a part of the data for detection. Most detection experiments reported in this paper have been based on the first 1000 observations of expiration; detection of the transition points was based on the separate flow measurements.

Fig. 2 shows time records of the first 1000 observations of lung noises during expiration phases, without and with the application of methacholine. At first sight it is seen that the level of the variance of the signals is different for this specific person, but this signal property has not been used for feature extraction because it varies for different subjects. Moreover, the level is not stationary, certainly not for the signal with methacholine. The value 951 for ME can be compared to the ME values for AR-MAsel models of different frames of the same signal. Fig. 3 uses the first 1000 expiration observations as a reference. It is clear that the statistical variations within one period are important, but they remain smaller than the difference caused by methacholine. Using the model without methacholine as a prototype in the data of Fig. 2 would give ME = 3022 instead of 951. Asymptotically, for small deviations, the ME is not sensitive to which signal is used for $A(z)$ and $B(z)$, and which signal yields $\hat{A}(z)$ and $\hat{B}(z)$, but for larger deviations this becomes important. The ME is a relative error in the frequency domain [8], which is sensitive for the exchange of numerator and denominator in (3) if their quotient is not approximately 1.
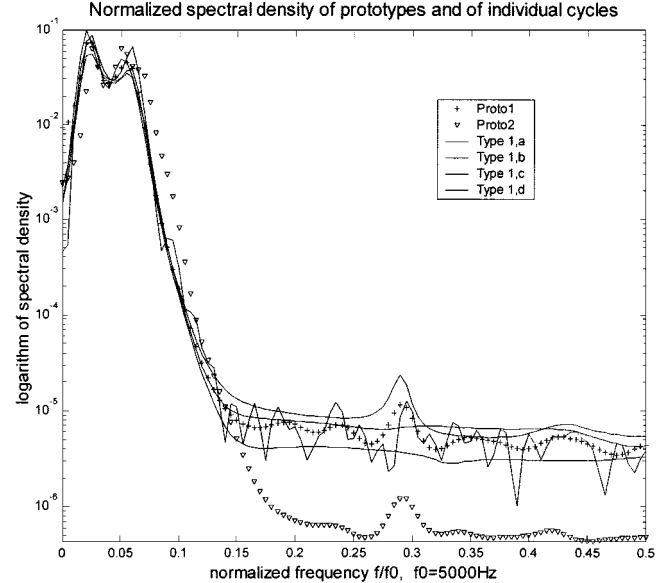

Fig. 4. AR(25) spectra of averaged prototype 1 and prototype 2 signals, together with four estimated spectra of type 1, each obtained from the first 1000 observations of a single expiration cycle with individual selection of model order and model type.

The spectral densities of both signals are given in Fig. 4; the integral of the spectral density is normalized to 1. The average AR(25) model of the first 1000 observations of four consecutive expiration cycles without methacholine is denoted Prototype 1 in Fig. 4; likewise, Prototype 2 is the average of four times 1000 observations with methacholine. For prototype 1, the spectra of the four contributing cycles have also been presented

separately, each with type and order selected from the data. The first 1000 observations of four sequential expiration cycles without methacholine have been used to estimate AR, MA, and ARMA models. First, the best model order has been selected for each model type. Afterwards, a single model type has been selected with ARMAsel [3], [9]. The selected models with the AR-MAsel algorithm were ARMA(8,7), ARMA(8,7), ARMA(7,6), and MA(53) for those four signals. However, it turned out that the spectra of AR(25) models are very close to the selected ones for all signals, because the ME of the AR(25) model was rather small when compared with the model with selected type and order. This shows that all three model types can give a good representation of the data, if the order is selected properly. As an example, the ME values between the preferred ARMA(8,7) model for the first signal of 1000 observations and for this signal selected AR(10) model, the selected MA(23) model and the chosen AR(25) model were 24, 47, and 19, respectively. This is harmless if the difference in ME between the prototypes is much greater. Experiments will provide information about those ME values. They should be much greater than the variations of ME for different well-fitting models of the same signal. Fig. 4 shows that the normalized prototype spectra are similar at low frequencies. However, the prototype 2 spectrum with methacholine is about ten times lower for higher frequencies. The variance of the signal with methacholine is about ten times greater for this subject. If the variance would be used as integral of the spectral density, the difference at low frequencies would become a factor of ten and at higher frequencies the spectra would be at the same level. It may seem attractive to use the accidental spectral difference at high frequencies in Fig. 4 for an ad hoc detection strategy for the presence of methacholine. Previous research [2] and inspection of the data of other subjects indicate that will not be successful. Moreover, the part of the spectrum above 1000 Hz is strongly modified by the low-pass filter at 1500 Hz.

The aim of the first experiment is to investigate whether it is possible to recognize the presence of methacholine by comparing *average* time series models of the same model type and order. It has been decided to take the average of four AR(25) models of the first 1000 observations of four consecutive expiration cycles as the model for the prototypes. Without any claim of optimality, the four AR(25) models have been transformed into the first 25 points of their normalized correlation functions. The four correlations have been averaged, and the AR(25) model belonging to the average correlation is taken as prototype 1. Likewise, four cycles with methacholine give prototype 2. The two prototype models P1 and P2, without and with methacholine, respectively, have been derived from the first four cycles of the first day. The average models of four further cycles of the first day are the test data T1a and T2a; the second day has the data T1b, T1c without and T2b, T2c with methacholine. The model error ME has been used to measure the difference of the test data with the two prototypes. The ME between P1 and P2, using P2 as prototype for $\hat{A}(z)$ and $\hat{B}(z)$, is 609. (Using P1 as prototype would yield 1827). The result of this in Table I is clear: the models obtained from type 1 data are close to P1 and far from P2. On the other hand, the models from type 2 data are close to P2 and far from P1. The larger the ratio between the ME values for wrong prototype/correct prototype, the better

TABLE I
THE ME BETWEEN AVERAGE AR(25) MODELS OF THE TEST DATA T1 AND T2 AND TWO PROTOTYPES P1 AND P2, WITHOUT AND WITH METHACHOLINE, RESPECTIVELY

	ME P1	ME P2
T1a	32	1020
T1b	43	617
T1c	71	588
T2a	463	37
T2b	481	96
T2c	411	95

TABLE II
THE ME VALUES FOR AR MODELS OF 1000 OBSERVATIONS OF EXPIRATION DATA AS A FUNCTION OF THE FIXED AR ORDER. PROTOTYPES AND TEST DATA HAVE THE SAME FIXED AR ORDER

AR order	P1/P2	T1/P1	T1/P2	T2/P1	T2/P2
1	0.7	1.4	0.1	4.6	0.9
2	34	48	22	166	51
4	815	43	319	718	201
8	908	74	543	770	225
16	961	81	774	709	225
25	933	105	813	723	228
32	979	126	930	740	245
64	1104	196	1003	922	414
128	1386	355	1208	1173	616
256	2352	848	1924	2115	1255
512	8019	3106	5355	7130	5001

detection is possible. The best ratio is found with the test data T1a and T2a, measured on the same day as the prototypes. The smallest ratio is found to be somewhat greater than four for T2c, measured two days later; but the difference with the two prototypes is still enough to support detection. Hence, the presence of methacholine can easily be detected with this method applied to one person.

After the conclusion that it is possible to detect the presence of methacholine, the next questions are:

- What are the best model type and order for detection?
- How can it be averaged or is it not necessary to take averages, and can single signals give sufficient information?

Various fixed-order AR models have been tried in Table II, the same for prototype and for test data. The prototypes are the AR models of 1000 observations recorded on the first day. Four test models have been used, each on the first 1000 observations of a single expiration cycle on the second day of the measurements. The results for test data versus prototypes are the averages of the ME of those four models. With AR models of orders 1 and 2, detection is impossible. For AR orders 4 and higher, the ratio was gradually becoming worse; but for all AR orders between 4 and 32, the ME for T2 with P1 was greater than three times the ME of T2 with P2. This is only slightly worse than in the last line of Table I, where prototypes and data models were based on the *average* of four data sets. Finally, the detection becomes almost impossible for AR order 512. Hence, the AR order should be not too low and not too high if AR models of fixed order are used.

The success of fixed-order AR models calls for investigation of the influence of selected orders. Table III gives results for the same data, but for every signal a separate AR order has been selected using the CIC criterion, for increasing values of the maximum candidate order for selection [11]. Results remain con-

TABLE III
THE ME VALUES FOR AR MODELS OF 1000 OBSERVATIONS OF EXPIRATION
DATA. PROTOTYPES AND TEST DATA HAVE INDIVIDUALLY SELECTED
AR-ORDERS AS A FUNCTION OF DIFFERENT MAXIMUM CANDIDATE AR
ORDERS FOR SELECTION

AR max order	P1/P2	T1/P1	T1/P2	T2/P1	T2/P2
1	0.6	1.5	0.1	4.4	0.6
2	31	49	18	160	32
4	822	53	311	651	212
8	912	83	529	744	235
16	1005	109	865	701	252
32	951	140	879	704	244
64	951	140	879	704	244
512	951	140	879	704	244

TABLE IV
THE ME VALUES OF EXPIRATION DATA AS FUNCTION OF THE NUMBER OF
OBSERVATIONS N. PROTOTYPES AND TEST MODELS WITH SELECTED
AR-ORDERS FOR MAXIMUM AR ORDERS $N/2$

N	P1/P2	T1/P1	T1/P2	T2/P1	T2/P2
8	30	7	20	16	159
16	50	3	21	36	127
32	11	20	15	21	13
64	28	18	77	36	13
128	61	101	313	51	54
256	152	47	205	96	63
512	470	76	355	307	151
1024	951	139	900	712	237
2048	1405	357	911	1839	771
4096	1835	384	1029	2610	833
8192	2547	573	1618	3649	848

TABLE V
THE ME VALUES FOR 1000 OBSERVATIONS OF EXPIRATION. PROTOTYPES AND
TEST DATA WITH SELECTED ARMA$(p, p - 1)$ ORDERS FOR DIFFERENT
MAXIMUM ARMA ORDERS FOR p

ARMA max	P1/P2	T1/P1	T1/P2	T2/P1	T2/P2
2	182	88	89	223	53
4	946	82	753	634	195
8	872	91	722	632	198
16	951	104	793	650	253
32	951	104	808	651	255
64	951	104	808	651	255

TABLE VI
THE ME VALUES FOR 1000 *DECIMATED* OBSERVATIONS OF EXPIRATION.
PROTOTYPES AND TEST DATA COMPUTED WITH THE ARMAsel ALGORITHM

	P1/P2	T1/P1	T1/P2	T2/P1	T2/P2
ARMAsel	951	113	809	663	293
decimate 2	751	73	299	629	375
decimate 4	73	326	274	292	165

ARMAsel results, with automatic selection of model type and model order, are comparable in quality to the best fixed-order AR results in Table III. It suggests that automatic time series analysis and selection with [9] can be used for feature extraction and detection. It is not necessary to use the same model type and model order for both prototypes and test data. Also the influence of decimation is shown in Table VI using ARMAsel models. Decimation with a factor of 2 uses an eighth-order Chebyshev low-pass filter with cutoff frequency f0/4 before resampling the data at the double interval. The detection performance after decimation becomes somewhat worse if all frequencies above f0/4 or 1250 Hz are filtered out, indicated by decimate 2. If all frequencies above f0/8 or 625 Hz are removed, like in decimate 4, detection is no longer possible. The explanation follows from Fig. 4. After removing all spectral contributions above f0/8 in the figure, the high-frequency part with the very significant difference between the two prototype models is filtered out. The remaining differences are too small for detection in this example where stationarity of the data is a problem.

V. DISCUSSION

It is not a consequence of theoretical arguments that models of selected orders perform well in detection and feature extraction. As an example, let us imagine two AR(100) processes that have 99 identical reflection coefficients and differ in the first one. In this case, the order 100 will be the best model order found with order selection criteria if sufficient data are available. But the best model order for the detection of two classes will probably be AR(1), because all differences between the two AR processes are concentrated in that single reflection coefficient. It is clear that this artificial example will not be representative of practical data. Therefore, data from lung sounds have been used to show the application of time series to detection in practice. The results of the Tables II and III show that selected model orders are suitable for the detection.

Not all eight subjects give the same results. The detection of methacholine with the method described is successful for five subjects and not for three other subjects [17]. There are several reasons why this method failed there. The detection

stant for maximum orders greater than 32. Obviously, the selected AR order is always less than 32. This behavior has the advantage that a high maximum candidate order can be used for unknown data, minimizing the probability of missing good models. If a lower order is sufficient, that will be selected.

A last question concerning AR models is the influence of the number of observations. Always, the first N observations of expiration have been used, from which an AR model is selected with a maximum candidate order of $N/2$. It may be expected that demands for stationarity are less strictly obeyed if the sample size approaches the whole duration of the expiration. Table IV shows that the detection is not possible for a sample size less than 256. The best quality is obtained for about 1000 observations. Many reasons can be the cause that the detection quality is not improved if longer samples than about 1000 are considered, e.g., lack of stationarity or the features of importance are concentrated in the beginning of the expiration cycle. The timing within the expiration cycle also could be varied, but that effect has not been studied.

The experiments of Table III have been repeated for selected ARMA$(p, p - 1)$ models with a variable maximum candidate order for p. Table V gives those results. Again, the results do not change anymore after the maximum candidate order has become high enough. The ME values for the data T1 without methacholine give a slightly better detection ratio than in Table III, but the ratio is somewhat worse for the data T2 with methacholine.

The final Table VI gives results for ARMAsel, the routine that automatically selects one single model from 250 AR models, 100 MA models, and 50 ARMA models [9]. It is remarkable that

of methacholine is disturbed by other phenomena that had more influence on the spectral respiration characteristics. The length of the respiration cycle varied too much for those three subjects, and the flow could not be maintained at a prescribed level [17]. Moreover, the characteristics, both with and without methacholine, were very different on the two recording days. For those three subjects, the presence of methacholine could not be detected from the experiments with time series, but the day on which the data were recorded. As a result, subjects should have some training to produce better stationary records in breathing.

VI. CONCLUDING REMARKS

Feature extraction with time series models can be applied automatically. First, prototype models are determined for subjects that belong to a known category. Afterwards, models obtained from unknown subjects can be classified into the category of the desired prototype. The use of time series models has the advantage that selection of model order and type is equivalent with the selection of the details that are statistically significant. This method of feature extraction can be applied to all measurement data that can be considered as realizations of stationary stochastic processes. The detection of the presence of methacholine in one subject is only a first step toward the detection of asthma by using prototype models of healthy and asthmatic subjects.

ACKNOWLEDGMENT

REFERENCES

[1] H. J. W. Schreur, Z. Diamant, J. Vanderschoot, A. H. Zwinderman, J. H. Dijkman, and P. J. Sterk, "Lung sounds during allergen-induced asthmatic responses in patients with asthma," in *Amer. J. Respir. Crit. Care Med.*, 1996, vol. 153, pp. 1474–1480.
[2] H. J. W. Schreur, J. Vanderschoot, A. H. Zwinderman, J. H. Dijkman, and P. J. Sterk, "The effect of methacholine-induced acute airway narrowing on lung sounds in normal and asthmatic subjects," *Eur. Respir. J.*, vol. 8, pp. 257–265, 1995.
[3] P. M. T. Broersen, "Facts and fiction in spectral analysis," in *Proc IMTC '98 Conf.*, vol. 2, St. Paul, MN, pp. 1325–1330.
[4] J. Vanderschoot, N. G. J. Kappeyne van de Coppello, and H. J. W. Schreur, "AR model orders of lung sounds," in *Proc. 14th Ann. Int. Conf. IEEE Engineering in Medicine and Biology Soc.*, 1992, pp. 2531–2532.
[5] J. Vanderschoot and H. J. W. Schreur, "Selection of regressors for $AR(q, v)$ modeling of normal lung sounds," in *Proc. 15th Ann. Int. Conf. IEEE Engineering in Medicine and Biology Soc.*, San Diego, CA, 1993, pp. 371–372.
[6] M. B. Priestley, *Spectral Analysis and Time Series*. London, U.K.: Academic, 1981.
[7] P. M. T. Broersen, "The quality of models for ARMA processes," *IEEE Trans. Signal Processing*, vol. 46, pp. 1749–1752, 1998.
[8] ——, "The performance of spectral quality measures," in *Proc. IMTC'99 Conf.*, Venice, Italy, pp. 751–756.
[9] ——, "Robust algorithms for time series models," in *Proc. ProRISC/IEEE CSSP'98 Conf.*, Mierlo, The Netherlands, pp. 75–82.
[10] J. P. Burg, "Maximum likelihood spectral analysis," in *Proc. 37th Meeting Soc. Exploration Geophysicists*, Oklahoma City, OK, 1967, p. 6.
[11] P. M. T. Broersen, The ABC of autoregressive order selection criteria, in Preprints Sysid '97 Conf., Kitakyushu, Japan, pp. 231–236, July 8–11, 1997.
[12] J. Durbin, "Efficient estimation of parameters in moving average models," *Biometrika*, vol. 46, pp. 306–316, 1959.
[13] ——, "The fitting of time series models," in *Revue Inst. Int. de Stat.*, 1960, vol. 28, pp. 233–243.
[14] P. M. T. Broersen, "The best order of long autoregressive models for moving average estimation," in *Signal Processing VIII, Proc. Eusipco Conf.*, Trieste, Italy, 1996, pp. 799–802.
[15] ——, "On orders of long AR models for ARMA estimation," in *Proc. ECSAP '97 Conf.*, Prague, Czech Republic, June 24–27, 1997, pp. 83–86.
[16] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1997.
[17] K. Rijnhout, "Detectie van methacholine met behulp van gemodelleerde longgeluiden," M.Sc. thesis (in Dutch), Dept. Appl. Phys., Delft Univ. Technol., Delft, The Netherlands, June 1999.
[18] P. M. T. Broersen, IEEE Trans. Instrum. Meas., vol. 49, Aug. 2000, to be published.

Piet M. T. Broersen was born in Zijdewind, The Netherlands, in 1944. He received the M.Sc. degree in applied physics in 1968 and the Ph.D. degree in 1976, both from the Delft University of Technology, Delft, The Netherlands.

He is currently in the Department of Applied Physics, Delft University. His research interests are the selection of order and type of time series models and the application to spectral analysis, model building, and feature extraction.

Stijn de Waele was born in Eindhoven, The Netherlands, in 1973. He received the M.Sc. degree in applied physics in 1998 from Delft University of Technology, Delft, The Netherlands, where he is currently pursuing the Ph.D. degree. His research is aimed at the development and application of new time series analysis algorithms and its application to radar processing.