# Dynamic Threshold Analysis of Daily Oxygen Saturation for Improved Management of COPD Patients

Malcolm Clarke, *Senior Member, IEEE*, Hulya Gokalp, Joanna Fursse, and Russell W Jones

*Abstract*—This study presents a novel dynamic threshold algorithm that is applied to daily self-measured $SpO_2$ data for management of chronic obstructive pulmonary disease (COPD) patients in remote patient monitoring to improve accuracy of detection of exacerbation. Conventional approaches based on a fixed threshold applied to a single $SpO_2$ reading to detect deterioration in patient condition are known to have poor accuracy and result in high false alarm rates. This study develops and evaluates use of a dynamic threshold algorithm to reduce false alarm rates. Daily data from four COPD patients with a record of clinical interventions during the period were selected for analysis. We model the $SpO_2$ time-series data as a combination of a trend and a stochastic component (residual). We estimate the long-term trend using a locally weighed least-squares (low-pass) filter over a long-term processing window. Results show that the time evolution of the long-term trend indicated exacerbation with improved accuracy compared to a fixed threshold in our study population. Deterioration in the condition of a patient also resulted in an increase in the standard deviation of the residual ($\sigma_{res}$), from 2% or less when the patient is in a healthy condition to 4% or more when condition deteriorates. Statistical analysis of the residuals showed they had a normal distribution when the condition of the patient was stable but had a long tail on the lower side during deterioration.

*Index Terms*—Remote patient monitoring, $SpO_2$, telehealth.

## I. INTRODUCTION

**T**ELEMONITORING can play a vital role in the management of patients with chronic conditions by using the data from long-term monitoring to identify trends and changes in the data that might be indicative of changes in the condition of the patient. Typically this includes monitoring $SpO_2$ data in patients with chronic obstructive pulmonary disease (COPD), or blood pressure and weight in patients with congestive heart failure (CHF). However, monitoring a patient over prolonged periods generates significant amounts of data, much of which will have little clinical relevance. The conventional approach to manage this data is to generate alerts regarding a change in condition based on a fixed threshold. However, this approach is known to result in a high number of false alarms, and leads

to distrust in the monitoring system with alerts being disabled or ignored. Moreover, there is a little or no evidence in the literature regarding the prognostic value of $SpO_2$ readings taken in the home. There is a need both to decrease the rate of false alarms and to demonstrate the value of the use of the data for long-term management of patients.

Pulse oximeters provide a noninvasive estimation of the arterial hemoglobin oxygen saturation ($SpO_2$) [1]–[3]. $SpO_2$ has been widely used in intensive care units, operating rooms, and primary care in the detection of hypoxemia (defined as $SpO_2$ < 90% in [4] and [5] and in this study). Pulse oximetry could have a vital role in improving the management of COPD patients by determining the need for long-term oxygen therapy and intervention in acute exacerbations [6], [7]. Exacerbations are caused by airway infections and result in worsening of the symptoms [8]. Frequent exacerbations and delayed intervention are associated with faster decline [9].

However, it is the case that oxygen saturation is not an absolute indication of condition, and patients can remain comfortable and stable at relatively low levels. This means that applying a simple threshold will result in many false alarms [10], [11]. There is also concern that measurements taken in the home by the patient may be of poor quality, with errors due to motion, poor perfusion, excessive light, venous pulsation, dyshaemoglobin, intravenous dyes, nail polish, artificial nails, pigmentation, and electromagnetic interference [1], [12]–[14]. This presents a challenge to discriminate whether a low reading is due to a measurement error, other causes of false reading or real [15]–[17].

Most work on analyzing and developing alerts for $SpO_2$ has been undertaken on data from the intensive care unit (ICU) or operating room (OR) [18]–[20]. However, the research is applied to continuous $SpO_2$ data and short-term effects, and the results and techniques are not applicable to home monitoring. Moreover, in ICU and OR, multiple physiological signals are measured and clinical staff is on hand to make an immediate decision on whether the alarm is true hypoxia or due to a technical fault, measurement error or other cause. In contrast, in home monitoring, there may only be a single physiological parameter being monitored and a small number of readings taken at infrequent intervals, such as once per day.

Our aim is to develop an approach that can accommodate the long-term fluctuations observed in $SpO_2$ that are tolerated by the patient and not indicative of exacerbation and determine the short-term changes that are, and thereby, reduce the number of false alarms and provide timely intervention.

TABLE I
PATIENT MONITORING INFORMATION

| Patient | Number of Days With Readings / Total Number of Monitoring Days | Reason for End of Monitoring | Number of Clinical Interventions |
|---|---|---|---|
| 1[1] | 127 / 134 (95%) | Died | 1[2] |
| 2 | 338 / 408 (83%) | Monitoring ended | 3 |
| 3 | 355 /465 (76%) | Monitoring ended | 1 |
| 4 | 310 /370 (84%) | Monitoring ended | 0 |
| 5 | 225 / 360 (63%) | Monitoring ended | 0 |
| 6 | 81 / 126 (64%) | Withdrew | 0 |
| 7 | 3 / 110 | Withdrew | 0 |

[1] Pulmonary fibrosis.

[2] $O_2$ therapy started on day 34 and continued until day 134.

## II. METHODS

### A. Patient Selection and Data Collection

$SpO_2$ data were collected from seven patients with chronic respiratory disease, who were recruited to participate in the inCASA project. All the patients were over the age of 65, were assessed as frail (using Edmonton frailty scale), had at least one chronic disease, and were living alone. Each patient was given a pulse oximeter and instructed to take at least one measurement each day; where possible first thing in the morning.

The pulse oximeter was connected wirelessly to a home gateway, which transmitted the data to a remote server for storage and to allow the clinical team to assess condition. The pulse oximeter was designed to be easy to use for the elderly; the user interface was a simple push button to start a reading and a green light when the reading was complete, at which point the sensor could be removed by the patient. The gateways were designed to be simple and self-contained for easy and rapid installation and unobtrusive for user acceptance. The gateway communication was based on GSM/GPRS for the WAN connection to the remote server and ZigBee Health Care Profile and IEEE 11073 for connection with the sensor [21].

Seven patients with chronic respiratory disease (six with COPD, one with pulmonary fibrosis) were enrolled and monitoring occurred between December 2012 and March 2014. Patients were enrolled at various times and for varying duration. Two patients (6 and 7) withdrew and their data were not included in the analysis. Compliance for taking daily readings was greater than 70% for four patients (patients 1–4 in Table I), and these have been selected for further analysis in this study. Missing readings were due to: the patient being absent from home; forgetting; or a technical fault. The first two were the most common.

Of the four patients, two had clinical interventions (including hospital admission), one had clinical intervention and died during the study, and one had no intervention. Clinical events were recorded in the patient electronic health record by the health care professionals caring for the patient. Detail on each patient is given in Table I.

### B. Multiple Daily Measurements

On occasion, patients would take repeated measurements on a single day. For our analysis, we require uniformly spaced data samples, one per day, and so a single representative value was derived to replace multiple daily measurements. Several options are available; including the median or maximum of the multiple measurements. The median has commonly been used in telemonitoring, whereas the maximum is often used by a clinician, who may repeat a measurement when suspecting it is incorrect and will select the higher reading (generally errors through motion, poor perfusion, excessive ambient light, dyshaemoglobin, result in a low reading).

Multiple measurements per day were seen in three of the four patients. For two patients, all the readings were above 90% (so both median and maximum value were above 90%) and so the method would not impact analysis. However, for one patient (patient 4), in 6 out of 10 days, one of readings was under 88% while the other was above 90%, which would lead to a difference in analysis. However, whenever a low reading was accompanied by a high reading on the same day, the higher was accepted by the clinician as representative, and so we adopted the maximum in our analysis.

## III. DATA ANALYSIS

### A. Signal Decomposition

Analysis of the statistical properties of the $SpO_2$ time series shows that it is nonstationary. Due to the stochastic nature of the $SpO_2$ data and its day-to-day variability, we have chosen to model the $SpO_2$ time series as the sum of a trend component (here, called the short-term trend) and a residual (the stochastic component), where the $SpO_2$ reading for day $i$ can be expressed as (1):

$$s_i = m_i + r_i \qquad (1)$$

where

$i$ denotes the time index (day of monitoring);
$s_i$ is the $SpO_2$ reading on day $i$ of monitoring;
$m_i$ is the short-term trend function;
$r_i$ is the residual.

The short-term trend, $m_i$, is further decomposed into two components: the long-term trend, $M_i$, and the difference between the short- and long-term trends, $y_i$, that represents the variation of the short-term trend about the long-term trend, where, $y_i = m_i - M_i$, giving

$$s_i = M_i + y_i + r_i. \qquad (2)$$

The full decomposition of the $SpO_2$ data for each time interval was according to (2).

### B. Trend Analysis

The trend represents the time evolution of the $SpO_2$ time-series data, distinct from short-term fluctuations. From our analysis, we have identified two distinct trends: long term and short term. We hypothesize that the long-term trend represents the underlying evolution of the condition over a prolonged period and may provide information on changes in the condition, such as deterioration, of the patient. The short-term trend will disclose
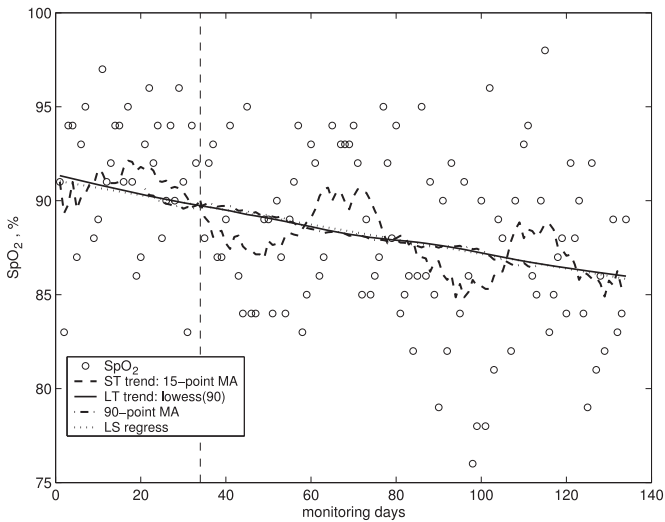
Fig. 1. SpO$_2$ time-series data for patient 1 (circles); 15-point MA filter (dashed-line); estimates of long-term trends with LOWESS (using 90 nearest neighbors), 90-point MA filter and LS regression (linear trend). The vertical line shows when the patient was placed on oxygen therapy.

rapid changes in condition that might accompany an exacerbation. The residual represents the expected fluctuations in daily readings. However, the variance may also indicate the condition of the patient; for example, there might be increased variance accompanying an exacerbation.

There are many methods to determine the trend components including least-square (LS) fitting, moving average (MA) filter, kernel smoothing, locally weighed least-squares smoother (LOWESS) [22], Bayesian smoother, Kalman filtering, or low-pass filtering [23]–[25]. In this study, we considered the MA filter, LOWESS, and LS regression.

*1) Short-Term Trend:* We evaluated 7-point and 15-point symmetric MA filters to determine the short-term trend component from the SpO$_2$ time-series data; and provide an estimate of parameter $m_i$ in (1).

The symmetric MA filter used in this study can be expressed as

$$\hat{m}_i = \frac{1}{N} \sum_{j=-k}^{k} s_{i+j} \quad , \quad N \leq 2k+1 \tag{3}$$

where $k = 3$ for 7-point and 7 for 15-point MA filter, $N$ is the number of readings within 15-day processing window and $N < 2k+1$ when there are missing readings in the processing window.

The width of the processing window is adjusted for the first and last seven data points in order to retain the properties of an MA filter and have an equal number of days of monitoring on each side.

The result from using the 7-point and 15-point MA filters was similar; both suppressed the day-to-day random variations on the readings but the greater smoothing provided by the 15-point MA filter was found preferable for this study. There were two further advantages: it is less sensitive to missing readings, and it allows a large class of trend functions (e.g., third degree

polynomials) to pass through without distortion. The result of the 15-point MA smoothing is given as the dashed line in Fig. 1.

*2) Long-Term Trend:* We considered three approaches for extracting the long-term trend: LS regression, MA filter, and LOWESS. For an MA filter and LOWESS, we used a sliding window approach and examined varying widths (60-day to 180-day) depending on the presence of sudden changes in the trend. The wider the processing window, the smoother is the estimate of the trend. However, narrow processing windows can better track sudden rapid changes in the trend that may be caused by an exacerbation. We only present results for 90-day processing window for the long-term trends.

Fig. 1 illustrates the results for the 90-point LOWESS (solid line), 90-point MA (dot-dash line), and LS regression (dotted line). In this example, the three approaches achieve similar results, i.e., an approximately linear declining trend, and are almost indistinguishable in Fig. 1. The LS method assumes the trend is a single linear function over the entire interval, and therefore, the data would have to be segmented into several time intervals if there were changes within the time series. In contrast, the MA and LOWESS filters do not assume the trend remains constant and adapt to the changes in the trend. The MA filter is quick to compute, which is an advantage for real-time applications. However, as there are no computation restrictions, we have chosen the LOWESS method.

Fig. 1 shows the analysis of the SpO$_2$ time series for patient 1, and includes the long-term and short-term trends. In this example, the long-term trend is almost linear, showing a gradual deterioration over the period of monitoring. The short-term trend shows the short-term variations about the long-term trend. In this example, this includes a number of excursions of increasing (days 35–60, 80–100) and decreasing (days 5–35, 60–75, 105–120) trend toward the hypoxia state. The variations of the short-term trend around the long-term trend resemble a cyclic component with a period of $\sim$47 days.

*3) Residuals and the Standard Deviation of Residuals:* Residuals are obtained by subtracting the estimated short-term trend, $\hat{m}$, from the SpO$_2$ readings, i.e., $r_i = s_i - \hat{m}_i$. Our expectation would be that the residual will have noise like properties; zero mean and will be stationary, that is, the variance (and standard deviation) will be consistent for all selected segments. However, the state of health might affect SpO$_2$, resulting in variation of the variance with condition.

The standard deviation of the residuals is a measure of rapid day to day fluctuations of the SpO$_2$ data about the short-term trend. This study investigates the standard deviation of residuals as a measure to detect a change in the condition of a patient. We use the notion of a moving estimator of the standard deviation, as we consider it is not stationary. The estimate is found using a sliding window that contains the residuals from the previous $N$ days. We investigated 30-day and 60-day windows, and determined that the 30-day window gave best results as it is better able to detect the changes characteristic of an exacerbation. In contrast, the 60-day window would fail to detect the onset of an exacerbation that followed on rapidly after another. Recovery time following an exacerbation can also be used to determine an appropriate duration of the window. For example, in [26] and
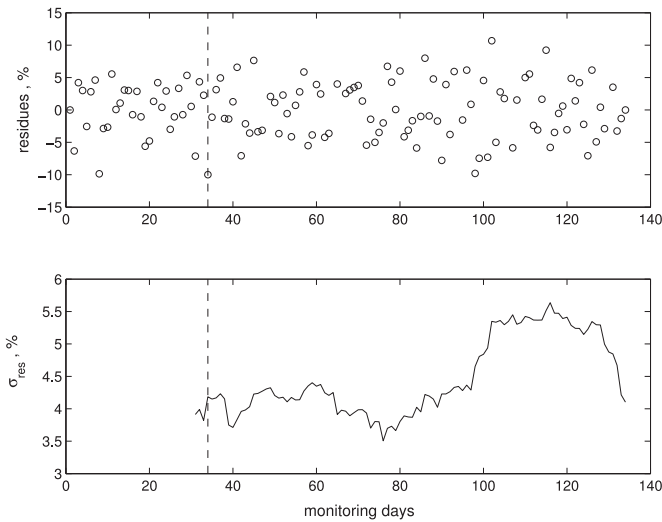
Fig. 2.    Residuals (top) and estimate of the standard deviation of the residuals (bottom) for patient 1.
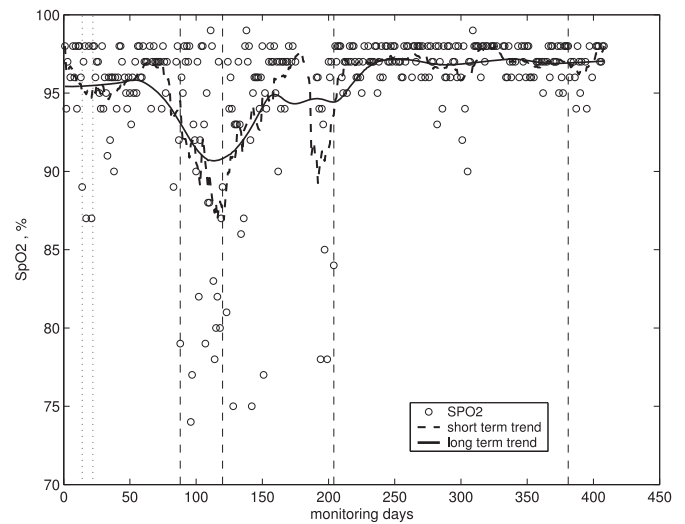


Fig. 3.    SpO$_2$ time-series data for patient 2; 15-point MA filter (dashed-line) and estimate of long-term trend with LOWESS (using 90 nearest neighbors).
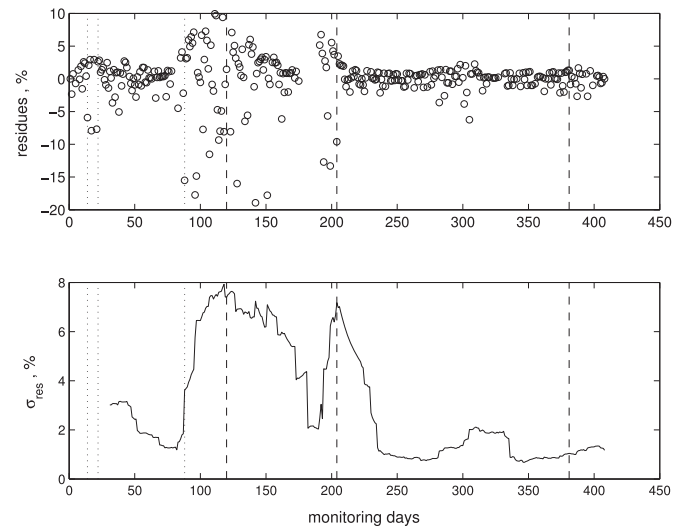


Fig. 4.    Residuals (top) and estimates of standard deviation of residuals (bottom) for patient 2.

[27], it was found that there is a high risk of recurrent exacerbations within 8 weeks of the onset of the previous. Therefore, it is highly likely that the onset of such recurrent exacerbations using a standard deviation estimated using a window in excess of 50 days will be missed; a shorter window is indicated for detecting exacerbations that are close in time.

The variance of the SpO$_2$ time series could be used as an indicator of the physiological state of COPD patients, but we have shown that the mean is not stationary. In contrast, the residual signal is a zero mean time series and its variance is preferred. Moreover, we have observed that the standard deviation of this residual signal is not stationary and has a continually varying value.

Fig. 2 shows the residuals (upper figure) and a windowed estimator (30 days) of the standard deviation ($\hat{\sigma}_{\text{res}}$). There is an increased value of the variance from 4% to 5% between days 100 and 120.

We tested whether the residuals were normally distributed using Kolmogorov–Smirnov (K–S) hypothesis test. We considered three periods for this test: whole monitoring period, first 90 days, last 44 days all the residuals. All of the periods passed the test at significance level of $\alpha = 0.05$, but the last two at a smaller value of the K–S statistics.

## IV. RESULTS

### A. SpO$_2$ Time-Series Data and Trends

*1) Patient 1:* Patient 1 had pulmonary fibrosis and commenced oxygen therapy on day 34 (vertical dashed line) and died on day 135. The long-term trend (see solid line in Fig. 1) is linear throughout the monitoring period and falls from 91.0% on day 1 to 85.8% on day 134, the last monitoring day. This trend would appear to indicate a steady decline in the condition of the patient. The short-term trend oscillates about the long-term trend with a period of approximately 47 days. On day 54, the nurses made a clinical note that the SpO$_2$ remained low, and this condition is clearly indicated in the short-term trend. The

short-term trend includes a succession of periods of increased and decreased level SpO$_2$, and this trend may be indicative of a change in the underlying physiological condition.

If a simple fixed threshold of 90% were to be applied to the raw data (a typical value for this type of work), then this would result in 70 separate alarms during the monitoring period. However, there is no specific mention in the clinical record of an exacerbation and these would all be deemed to be false alarms. However, the long-term trend clearly indicates a continued deterioration throughout the monitoring period, and the short-term trend indicates distinct periods of improvement and deterioration. The sudden increase in the value of the standard deviation of the residual around day 100 may also indicate an underlying deterioration in condition, where there is loss of physiological control of the condition, resulting in rapid fluctuations in the observations on a daily basis. Furthermore, although the patient
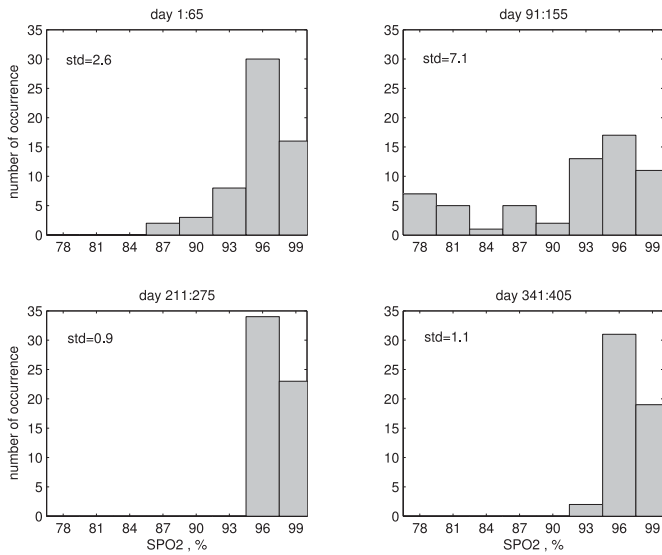
Fig. 5. Distribution of SpO₂ readings during periods of exacerbation and intervention.
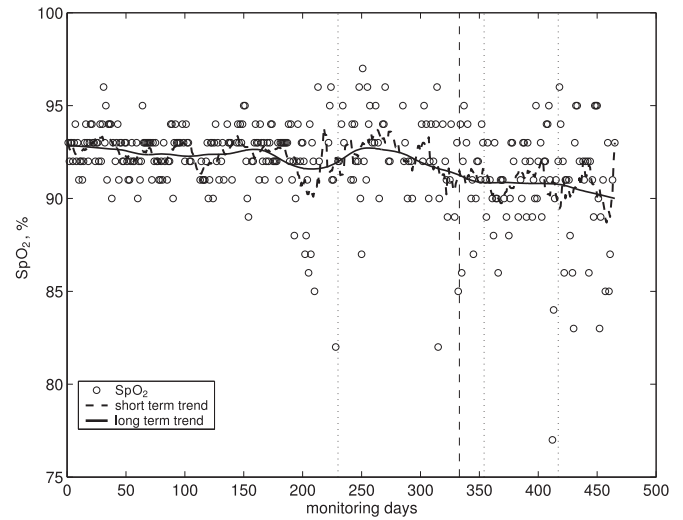


Fig. 6. SpO₂ time-series data for patient 3; 15-point MA filter (dashed-line); estimates of long-term trends with LOWESS (90).



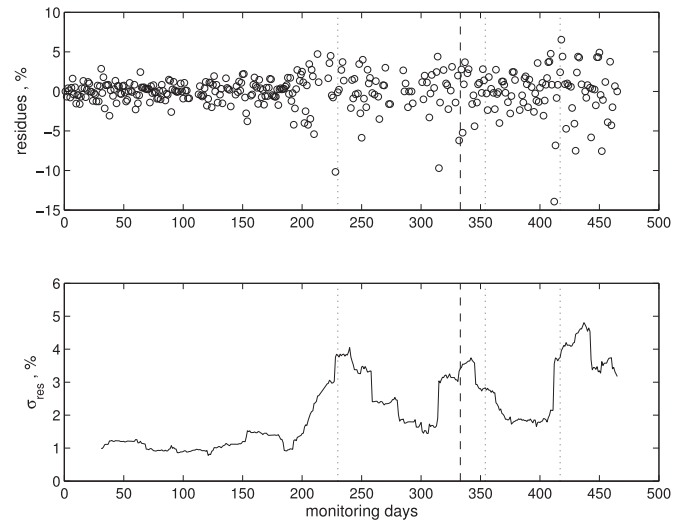Fig. 7. Residuals (top) and estimates of standard deviation of residuals (bottom) for patient 3.

commenced oxygen therapy on day 34 and reported that it was helping, there is no associated change in the SpO₂ data.

*2) Patient 2:* Patient 2 had three major recorded clinical interventions; hospital admission on day 120 for 3 days, chest infection on day 204, and an exacerbation on day 381 (heavy vertical dashed lines in Figs. 3 and 4). Patient 2 also had notes on clinical concerns around day 22, 30, and 88 (light vertical dashed lines in Figs. 3 and 4), which were investigated but no intervention made.

The occurrence of the first two major clinical events is clearly evident in the long-term trend, short-term trend, and the residuals. During periods of good health, the long-term trend and short-term trend of the SpO₂ remain between 96–98% and the standard deviation of the residuals is about 1%. However, during the first two major clinical events the long-term trend and the short-term trend fall to a level below 90% and the standard deviation of the residuals rises above 6%. In contrast, there are no significant changes in any of the long-term trend, short-term trend, or residuals accompanying the major third clinical event on day 381 to indicate an exacerbation.

In contrast to patient 1, patient 2 did not exhibit a progressive and steady decline throughout the monitoring period, instead there were a series of episodes, with the SpO₂ returning to a usual value following intervention. Neither was there any repetitive or cyclic behavior. Instead, the SpO₂ appears to follow the condition of the patient closely; with the onset of exacerbation and response to treatment clearly seen (e.g., intervention on days 120 and 204).

There was concern among the clinical team over the low readings around day 14, 22, and 88 (shown light vertical dashed lines). However, the patient did not report any change in condition on day 14 or 22; indeed the patient believed that their breathing had improved around day 14, and the long-term and short trend would support that although there was a change in underlying condition, it was not severe. On day 88, the patient was diagnosed with a cold, and the condition continued to

deteriorate until the patient was admitted to hospital on day 120 for 3 days. This is supported by the long trend, short-term trend, and residuals. The patient is further diagnosed with productive cough and chest infection on day 204, again supported by the long trend, short-term trend, and residuals.

We examined the distribution of the SpO₂ readings for periods in which there was an exacerbation and intervention and compared with other periods during the monitoring. Fig. 5 shows the distribution as a histogram for the periods: day 91 to 155 (exacerbation and intervention); day 1 to 65 (no treatment); day 211 to 275 (after treatment); and day 341 to 405 (just before the end of monitoring). It is clear that during the period of exacerbation, we observe a significant number of readings that are below a nominal threshold, e.g., 93%. However, although the condition of the patient during first 85 days of monitoring is considered stable and without exacerbation, there are also a significant number of readings that fall below that threshold and these would trigger false alarms. In contrast, using the long- and
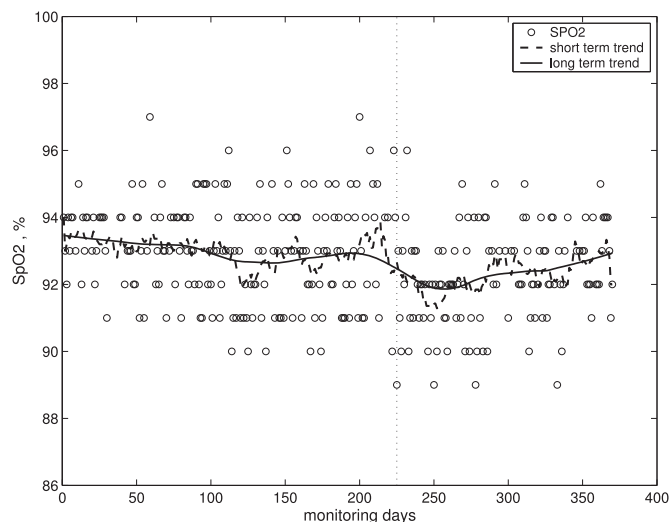
Fig. 8.   SpO$_2$ time-series data for patient 4; 15-point MA filter (dashed-line); estimates of long-term trend with LOWESS (90).



Fig. 9.   Residuals (top) and estimates of standard deviation of residuals (bottom) for patient 4.

short-term trends prevents such false alarms. It is evident that the distributions are not normal and in several instances have a long tail on the lower side.

*3) Patient 3:*   Patient 3 was reported to have shortage of breath on day 230 (light dotted vertical line), an exacerbation on day 333 (heavy dashed vertical line), feeling better on day 354 (light dotted vertical line), and assessed by respiratory team to determine if oxygen therapy was required on day 417 (light dotted vertical line).

Visual inspection of the SpO$_2$ data (see Fig. 6) would suggest that up to reporting of shortage of breath on day 230, the long-term trend had been relatively constant at about 92%, there were a few relatively minor short-term trend fluctuations (excepting the short-term trend at day 200) and the standard deviation of the residuals was low at 1% (see Fig. 7). Whereas after day 230, there is a steady decline in the long-term trend, falling to 90% by day 460, there are greater excursions in the short-term trend, and the standard deviation of the residuals rises to at least 2%, and over 3% for significant periods of time. Moreover, there is a clear excursion in the short-term trend, falling to 90% on day 205, shortly before the reported shortness of breath, and the standard deviation of the residuals increases from 1% on day 200 to 4% by day 230.

There are corresponding falls in the short-term trend and rises in the standard deviation of the residuals accompanying the exacerbation on day 333 and the introduction of oxygen on day 417 (see Fig. 7).

We further note that the long-term trend for patient 3 after day 330 begins to follow that of patient 1, with a continual linear fall and there is a sustained increase in the standard deviation of the residuals. We do not have the data to follow the trend to its conclusion.

*4) Patient 4:*   Patient 4 had no interventions during the monitoring period. The patient was seen after a particularly low reading on day 225 (89%), but it was decided that the patient was in good condition. In contrast, the long-term trend (see Fig. 8) would indicate a steady decline from 93.5% on day 1
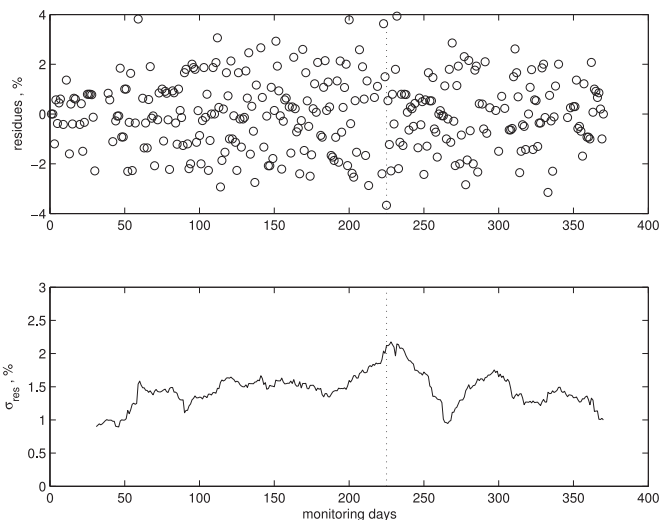
to 92% on day 275, and the standard deviation of the residuals (see Fig. 9) increases steadily from 1% to 2%. Thereafter, the long-term trend reverses and increases to nearly 93% and the standard deviation of the residuals falls to 1% by the end of monitoring.

*B.  Standard Deviation of Residuals*

*1) Results:*   We have proposed a new measure; the standard deviation of the residuals. The residual is derived by removing the short-term trend from the SpO$_2$ data, and the standard deviation of this signal is then found. As the signal is not stationary, we determine a moving windowed value.

When the patient is in a stable condition, the residuals are small and have a low standard deviation; whereas during exacerbation and decline in the condition, the standard deviation of the residuals rises significantly. From the patients presented in this paper, we have observed the standard deviation to be between 1% and 2% when the patient is in a stable condition, and during exacerbation or decline in the condition, the standard deviation of the residuals rises above 2%, with an increase above 5% in some cases. For example, patient 2 has a standard deviation of residuals of 2% during periods of stable condition and rises as high as 8% during exacerbation. In patients where there is a steady decline in condition, such as patient 1, the standard deviation of the residuals remains high throughout the monitoring period (4%), and increases from 4% to 5% as the condition deteriorates (see Fig. 2).

*2) Comparison With Long-/Short-Term Trends:*   From our small set of patient data, we have observed that the standard deviation of residuals appears to accompany the onset of exacerbation in each instance for each patient. Moreover, in the case of patient 3, the standard deviation of residuals indicates episodes of exacerbation (see Fig. 7) that are not indicated in the long-term or short-term trends (see Fig. 6). We observe that when patients are in a stable condition, the value of standard deviation of residuals is about 2% or less and the long-term
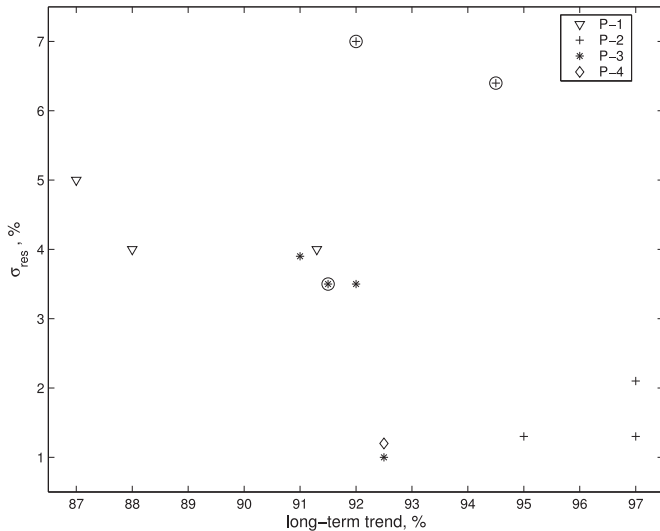
Fig. 10. Comparison of estimate of standard deviation of residuals against corresponding long-term trend value.

**TABLE II**
**Number of Alarms With Residual Method and Simple Threshold Method**

| Patient number | Day of Intervention | Number of Alarms | | Alarm Days With Residual Method |
|---|---|---|---|---|
| | | Residuals Method | Simple Threshold | |
| | | Residual $< -6\%$ | $SpO_2 < 90\%$ | |
| 1 | 34 | 10 (0) | 70 (0) | See 1 |
| 2 | 120;204;381 | 19 (17) | 28 (24) | See 2 |
| 3 | 333 | 7 (2) | 37 (7) | See 3 |
| 4 | 0 | 0(0) | 4 (0) | |

1: P-1 residual method: [2;8;31;34;42;90;98;99;101;125].
2: P-2 residual method: [88;96;97;102;107;114;115;118;123;128;134;142;151;162].
3: P-3 residual method: [228;315;332;412;413;430;452].

trend of $SpO_2$ is generally above 91%. During exacerbation the value of the standard deviation of residuals is above 3.5%, however, there is no unique threshold value for the long-term trend.

In Fig. 10, we compare the value of the estimate of standard deviation of residuals against the value for the long-term trend at the corresponding time for a number of representative points during monitoring for each of the patients to include good health and exacerbation. Data points that correspond to times when there was an exacerbation are encircled.

Note that in some cases, although there was concern regarding the health of the patient, there were no specific instances of exacerbation recorded. This includes patient 1 who had consistently low $SpO_2$, with 70 readings below 90% ($O_2$ therapy started on day 34) and patient 3 (* symbols in Fig. 10) with long-term trend consistently around 91–92. The values for the standard deviation of residuals for patient 1 and patient 3 were $\geq 4\%$ and $\geq 3.5\%$, respectively.

From Fig. 10, we do not see any correlation between the standard deviation of residuals and long-term trend and conclude that the long-term trend is not a reliable predictor of exacerbation. However, we note the continual fall in the long-term trend in some patients from which they do not recover and that suggest that the long-term trend could be an indicator of the underlying condition and eventual outcome.

*3) Approaches to Generating Alerts Using $SpO_2$:* Our aim is to develop accurate algorithms to determine exacerbations, deterioration, or other events of clinical significance. Inspection of the long-term trend of the $SpO_2$ in patients such as patient 1 and the later stages for patient 3 indicates that use of a fixed threshold is inappropriate, and any such use (e.g., 90%) will result in an unacceptable number of false positive detections. Instead, we must identify phenomena that exhibit a change that occurs when there is an event.

We considered the use of the short-term trend of the $SpO_2$, but this did not provide an indication in all events. Therefore, we

have investigated the use of the standard deviation of the residuals as a method to determine exacerbation. We have observed that standard deviation of the residuals remains below 2% when the patient has a stable condition and rises to above 3.5% when there is an exacerbation.

We have investigated how we can use the instantaneous value of the residuals as the criterion for detecting an exacerbation. Normally a value twice the standard deviation is used for detecting outliers in signals with normal distribution. However, we observe a long tail on the lower side and so we use a value three times the standard deviation as threshold. In this case, we set the residual threshold to determine an exacerbation to –6%. We call this the residuals method. For this study, we have chosen an empirical value for the threshold of three times the value of the standard deviation during the periods when the patients are in a stable condition. An optimum threshold value for the residuals can be determined from the distributions of the estimate of standard deviation of residuals for periods of stable condition and exacerbation.

Table II presents the number of alarms for each patient using the residual method compared with the simple method of a fixed threshold applied to the $SpO_2$ time-series data ($SpO_2 < 90\%$). For each method, we present the total number of points that were detected together with the number of those points (in brackets) that occurred immediately preceding, or within an exacerbation or treatment period. For example, for patient 2, 28 alarms were detected in total for the simple threshold method, but only 24 of those alarms occur between days 88 and 155 (as considered from the long-term trend), and thus, considered relevant.

For patients that have distinct episodes of exacerbation (patient 2 and 3), we observe that the residual method detects the episode several days in advance of the recording of the clinical event. There are also fewer false alarms.

Patient 1 did not have any clinically recorded events or distinct exacerbations, as in the case of patient 2 and 3, rather the patient had a condition in which there is expected to be continual deterioration. We observed that the long-term trend had a continual decline and also that the value of the standard deviation of the residual is consistently high in value throughout the monitoring period, with an increase in value at day 100 that

is maintained almost until the end. The clinical significance of this change is uncertain.

## V. DISCUSSION

We have investigated the decomposition of daily $SpO_2$ time-series data into three components; long-term trend, short-term trend, and the residual signal. We have then considered the physiological interpretation of each of the components and evaluated their use in understanding the underlying condition and detecting an exacerbation. The long-term trend has been found to provide information on the evolution of the patient condition. This is evident in the continual downward trend in patient 1, the downward trend in the latter stages of monitoring in patient 3 and the recovery from exacerbations in patient 2 and patient 3. The presence of these long-term trends highlighted that simple thresholds would not be an effective method for detecting exacerbation and would generate many false alarms if compensation for long-term trend were not applied. The short-term trend is designed to smooth the day-to-day variations and is considered to identify the rapid changes in the $SpO_2$ data that would accompany exacerbation, but there were not clear changes for all events and so there would be a number of false negatives and false positives depending on threshold. The estimate of the standard deviation of residuals was found to be the most accurate method to detect exacerbations and was also found to indicate deterioration in condition of the patient.

We have observed two types of behavior in the long-term trend. In some patients (e.g., patient 2), there is a rapid drop in the long-term trend that is accompanied by an increase in the estimate of standard deviation of the residuals and accompanies an exacerbation. There is usually a return to normal value. In the second behavior, (e.g., patients 1 and 3), there is a continual slow declining trend. This might be indicative of the underlying condition and could be used to determine long-term approach to management of the condition. This trend, when accompanied with high value for the estimate of standard deviation of the residuals, could have predictive value for increased risk of mortality (e.g., patient 1).

The standard deviation of residuals can indicate the status of the patient: low values ($\sim$2% or less) are seen to be associated with stable condition, whereas high values ($\sim$4% or greater) are seen to be associated with exacerbation or steady irreversible declining state. There are parallels with the findings of the study of [28] in which continuously measured $SpO_2$ signals from intensive therapy unit were analyzed. The signal was split into a series of time intervals during each of which the signal was approximated by a step-wise linear model (rising, falling, or steady) and the difference between the linear approximation and the signal gave the residuals. For best detection of exacerbation using standard deviation of residuals, we would need to determine optimum threshold values for onset and ending of the exacerbation. The threshold values should be chosen so that they are best able to distinguish exacerbations and deterioration from the stable condition. For this purpose, the distribution of the standard deviation of residuals of the two conditions can

be used. However, we had insufficient patients and examples to permit this analysis.

To be useful for timely detection of exacerbation, then the algorithm must be based on instantaneous residual values rather than their standard deviation. We have observed that the standard deviation remains below 2% when the condition of the patient is stable. The usual approach is to apply a threshold that is twice the standard deviation; however, during deterioration, the residuals are not normally distributed and have a long tail on the lower side. We therefore chose to use a threshold for the residuals that is three times the standard deviation ($-$6%).

Currently, we believe that the standard deviation of residuals provides insight to the status of the underlying regulatory mechanism, which is affected by the health of the patient. When in homeostasis, the feedback control mechanism will maintain the control variable within a well-defined range; in the normal case, breathing is regulated by the levels of oxygen and carbon dioxide in the blood to maintain well oxygenated levels close to 100%; and in the COPD patient, although lowered, the level of oxygen is maintained within well-controlled limits. However, when the feedback control is impaired, such as during an exacerbation, there is a loss of regulation and the controlled variable can vary wildly, giving rise to a signal with noise like properties. We detect this signal as the rise in standard deviation of the residual. In this case, when the patient has an exacerbation, blood oxygen levels fluctuate over a wide range in an uncontrolled fashion.

We investigated the impact of the choice of the window length for the MA filter that was used to determine the short-term trend (to remove the rapid day-to-day variations in the data). We determined that 15-point gave good results, being a compromise between reducing the day-to-day fluctuations and reacting to rapid changes in the underlying data.

We also investigated the type of filter to use for the long-term trend and found the LOWESS technique gave best results. However, we also determined that the results from an MA filter were acceptable. We investigate the window length (60–120 points). Again there is a compromise between short windows that will track rapid changes (such as patient 2) and determining underlying trend. When applied to real time data, MA filters have the advantage over LS regression: 1) they are simpler to implement, 2) they are a local smoothing technique and can, therefore, adapt to the changes in trend, and 3) they do not require data segmentation into time intervals over which regression is applied as in [29].

A significant limitation of this study is the small number of patients and the small number of clinically recorded events. This precludes statistical analysis of sensitivity and specificity. Although we cannot generalize the results or comment on statistical significance, we believe our approach is promising. We need to apply our approach to a larger dataset to determine the accuracy and predictive capability of the indicators that we have proposed. Future studies might consider other aspects including: are there combinations of indicators that can better predict exacerbation and underlying condition of the patient [30]; are there factors that affect day-to-day $SpO_2$ variability (e.g., $O_2$ therapy, changes in activity level, skin pigmentation, presence of co-morbidities such as tachycardia or other heart condition); the

role of self-measured $SpO_2$ in patient management (is it complementary or is it a better way of managing COPD patients?); how do missing data affect the accuracy of the suggested approaches?

## VI. CONCLUSION

Providing accurate notification of exacerbation, deterioration, and other clinical events in patients in remote patient monitoring is essential. Too often alerts are disabled because of an unacceptable level of false alarms, and daily screening of all patient data is too high a burden when going to scale. Improved algorithms that are accurate and can adjust automatically to the individual are required. We have presented approaches that are not only promising as a technique to determine exacerbation but also may provide information on the underlying condition and that may be useful in determining management of the patient.

We have undertaken retrospective analysis of $SpO_2$ data measured in patients with COPD as part of a long-term project to monitor frail elderly. Although we are unable to generalize our results or draw statistically significant conclusions, this exploratory work has identified several important parameters that can indicate exacerbation, deterioration, and provide information on the long-term management. These include the standard deviation of residuals (day-to-day $SpO_2$ variability) and the time evolution of the long-term trend. Using the standard deviation of residuals was found to be highly accurate in detecting exacerbation, and reduced the number of false alerts significantly compared to a simple threshold applied to the $SpO_2$ data. Moreover, we also observed that it could detect changes in patients where there was long-term deterioration.

## REFERENCES

[1] C. D. Hanning and J. M. Alexander-Williams, "Pulse oximetry: A practical review," *Brit. Med. J.*, vol. 311, no. 7001, pp. 367–370, 1995.

[2] A. Jubran, "Pulse oxymetry," *Critical Care*, vol. 3, no. 2, pp. 11–16, 1999.

[3] V. A. Potter, "Pulse oximetry in general practice: How would a pulse oximeter influence patient management?" *Eur. J. Gen. Pract.*, vol. 13, no. 4, pp. 216–220, 2007.

[4] M. C. Chambrin, P. Ravaux, D. Calvelo-Aros, A. Jaborska, C. Chopin, and B. Boniface, "Multicentric study of monitoring alarms in the adult intensive care unit (ICU): A descriptive analysis," *Intensive Care Med.*, vol. 25, no. 12, pp. 1360–1366, 1999.

[5] J. T. Moller, N. W. Johannessan, K. Espersan *et al.*, "Randomised evaluation of pulse oximetry in 20,802 patients:II. Perioperative events and postoperative complications," *Anesthesiology*, vol. 78, pp. 445–453, 1993.

[6] A. Pluddemann, M. Thompson, C. Heneghan, and C. Price, "Pulse oximetry in primary care: Primary care diagnostic update," *Brit. J. Gen. Pract.*, pp. 358–359, 2011.

[7] C. M. Roberts, J. Franklin, A. O'Neill, R.P. Roberts, J. Ide, M.L. Hanley, and J. Edwards, "Screening patients in general practice with COPD for long-term domiciliary oxygen requirement using pulse oximetry," *Respir. Med.*, vol. 92, no. 11, pp. 1265–1268, 1998.

[8] NICE clinical guideline. (2014, Oct.). Management of chronic obstructive pulmonary disease in adults in primary and secondary care. Working definition of COPD. [Online]. Available: http://www.nice.org.uk/guidance/cg101/chapter/working-definition-of-copd, (accessed Oct 2014)

[9] T. M. Wilkinson, G. C. Donaldson, J. R. Hurst, T. A. Seemungal, and J. A. Wedzicha, "Early therapy improves outcomes of exacerbations of chronic obstructive pulmonary disease," *Amer. J. Respir. Critical Care Med.*, vol. 169, no. 12, pp. 1298–1303, 2004.

[10] A. T. Rheineck-Leyssius and C. J. Kalkman, "Influence of pulse oximetry settings on the frequency of alarms and detection of hypoxemia: Theoretical effects of artefact rejection, alarm delay, averaging, median filtering or lower setting of alarm limit," *J. Clin. Monit. Comput.*, vol. 14, no. 3, pp. 151–156, 1998.

[11] N. E. Brown Connolly and M. Clarke, "A better way to evaluate remote monitoring programs in chronic disease care: Receiver operating characteristic analysis," *Telemed. e-Health*, vol. 20, no. 12, pp. 1143–1149, 2014.

[12] S. Fouzas, K. N. Priftis, and M. B. Anthracopoulos, "Pulse oximetry in pediatric practice," *Pediatrics*, vol. 128, no. 4, pp. 740–752, 2011.

[13] J. E. Sinex, "Pulse oximetry: Principles and limitations," *Amer. J. Emerg. Med.*, vol. 17, no. 1, pp. 59–66, 1990.

[14] M. Elliott, R. Tate, and K. Page, "Do clinicians know how to use pulse oximetry? A literature review and clinical implications," *Aust. Crit. Care*, vol. 19, no. 4, pp. 139–44, 2006.

[15] E. Bosque "Symbiosis of nurse and machine through fuzzy logic: Improved specificity of a neonatal pulse oximeter alarm," *Adv. Nurs. Sci.*, vol. 18, no. 2, pp. 67–75, 1995.

[16] D. L. Reich, A. Timcenko, C. A. Bodian, J. Kraidin, J. Hofman, M. DePerio, S. N. Konstadt, T. Kurki, and J. B. Eisenkraft, "Predictors of pulse oximetry data failure," *Anesthesiology*, vol. 84, no. 4, pp. 859–64, 1996.

[17] K. A. Carlson and J. S. Jahr, "An update on pulse oximetry. Part II: Limitations and future applications," *Anesthesiol. Rev.*, vol. 21, no. 2, pp. 41–6, 1994.

[18] A. T. Rheineck-Leyssius and C. J. Kalkman, "Advanced pulse oximeter signal processing technology compared to simple averaging. I. Effect on frequency of alarms in the operating room," *J. Clin. Anesthesia*, vol. 11, pp. 192–195, 1999.

[19] M. C. Chambrin, "Alarms in the intensive care unit: How can the number of false alarms be reduced?" *Crit. Care*, vol. 5, pp. 184–188, 2001.

[20] U. Gather, R. Fried, and M. Imhoff, "Online classification of states in intensive care," in *Data Analysis (Studies in Classification, Data Analysis, and Knowledge Organization)*. Berlin, Germany: Springer, pp. 413–428, 2000.

[21] M. Clarke, J. de Folter, C. Palmer, and V. Verma, "Building point of care health technologies on the IEEE 11073 health device standards," in *Proc. IEEE EMBS Spec. Topic Conf. Point-of-Care Healthcare Technol.*, Bangalore, India, Jan. 16–18, 2013, pp. 117–119.

[22] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *J. Amer. Statist. Assoc.*, vol. 74, no. 368, pp. 829–836, 1979.

[23] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and its Applications With R Examples*, 2nd ed. New York, NY, USA: Springer 2006.

[24] R. C. Brockwell and J. J. Andrews, "Understanding your anesthesia workstation," *ASA Refresher Courses Anesthesiol.*, vol. 35, no. 1, pp. 15–29, 2007.

[25] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1979.

[26] J. R. Hurst, Gavin C. Donaldson, Jennifer K. Quint, James J. P. Goldring, Ramin Baghai-Ravary, and Jadwiga A. Wedzicha, "Temporal clustering of exacerbations in chronic obstructive pulmonary disease," *Amer. J. Resp. Crit. Care Med.*, vol. 79, no. 5, pp. 369–374, 2009.

[27] E. Ozyilmaz, N. Kokturk, G. Teksut, and T. Tatlicioglu, "Unsuspected risk factors of frequent exacerbations requiring hospital admission in chronic obstructive pulmonary disease," *Int. J. Clin. Pract.*, vol. 67, no. 7, pp. 691–697, 2013.

[28] S. Charbonnier and S. Gentil, "On-line adaptive trend extraction of multiple physiological signals for alarm filtering in intensive care units," *Int. J. Adapt. Control Signal Process.*, vol. 24, pp. 382–408, 2010.

[29] S. Charbonnier and S. Gentil, "A trend-based alarm system to improve patient monitoring in intensive care units," *Control Eng. Pract.*, vol. 15, no. 9, pp. 1039–1050, 2007.

[30] M Clarke, "The need for an integrated approach to remote monitoring of physiological data and activity data," *J. Telemed. Telecare*, vol. 20, no. 3, pp. 159–160, Apr. 2014.