

## A simple and general change-point identifier<sup>‡</sup>

R. H. Riffenburgh<sup>1,\*</sup> and K. M. Cummins<sup>2</sup>

<sup>1</sup>*Naval Medical Center, San Diego, CA, U.S.A.*

<sup>2</sup>*University of California, San Diego, CA, U.S.A.*

### SUMMARY

In a not-necessarily-stationary time-series, a Moving  $F$  statistic can identify points of change in the nature of the series model (the estimate of the underlying data-generating process), in its parameters, in residual variability, or in any combination of these. In addition, it can uncover changes masked by a reduction in residual variability. Patterns in the forces giving rise to the data may often be perceived. To form the Moving  $F$ , a theory of the process or a regression method on a baseline sample estimates the series model and the residual mean square about it is calculated. This series model is extended past the baseline with residuals assumed to be normally distributed. The Moving  $F$  is calculated as the moving average of squared deviations about the series model in ratio to the baseline mean square. The Moving  $F$  crossing the critical  $F$  identifies a change in the series model, i.e. signals its presence and location. In our experience, this Moving  $F$  method is easier to use than other commonly employed change-point identifiers (CUSUM, EWMA, data-based bandwidth selection, MCMC) and has been found to work in several situations where some other identifiers fail. (MCMC is more general, but requires advanced statistical ability.) Examples given are monitored prostate specific antigen in a post-treatment prostate cancer patient and detection of Harold Shipman's medical murders. Moving  $F$  is 'simple and general' in the sense of both simultaneously; we have not found another relatively simple method to be as general. Published in 2005 by John Wiley & Sons, Ltd.

KEY WORDS: change-point; time-series; Moving  $F$ ; quality control; process control; MCMC

### 1. INTRODUCTION

Suppose we encounter a not-necessarily-stationary time-series, perhaps laboratory readings on a patient through time. If a change of some sort occurs in the state of this patient indicated by these readings, we wish to detect it and estimate the point at which it occurs (*identify* the change-point).

We will list the most common types of change for which a change-point identifier is needed. Then we will introduce and demonstrate a prospective analysis form of the Moving

---

\*Correspondence to: Robert H. Riffenburgh, Naval Medical Center, 34800 Bob Wilson Drive, Ste 5, San Diego, CA 92134-1005, U.S.A.

†E-mail: riff@sdsu.edu

‡This article is a U.S. Government work and is in the public domain in the U.S.A.

$F$ -statistic. (The original form could analyse only retrospective data.) Finally, we will note the most frequently used methods to identify changes.

## 2. TYPES OF CHANGE

We observe a time series of points in time,  $(y_i, t_i)$ ,  $i = 1, 2, \dots, L$ . These points arise from an unknown data-generating process with functional form estimated as the series model.  $y_i = f(t_i; \boldsymbol{\beta}) + e_i$ , where  $f$  is the series model,  $\boldsymbol{\beta}$  is a vector of parameters, and  $e_i$  is  $\text{NID}(0, \sigma_i^2)$ . ( $\text{NID} \equiv$  normally and independently distributed.) A change exists at  $t = t^*$ , where  $t^*$  is unknown.

Time may be measured in any units, milliseconds to years. While  $f$  may take any form, practicality constrains use to a small sample of data after a potential change, limiting  $f$  to a few simple forms.  $f$  may change (e.g. linear to exponential or quadratic), one or more components of  $\boldsymbol{\beta}$  may change (e.g. jumps, slope changes, changes in curvature), and/or the residual variability may change.

The term *regime*, introduced by Quandt [1], has been used to refer to the totality of characteristics influencing a time-dependent process, i.e.  $f$ ,  $\boldsymbol{\beta}$ , and  $\sigma^2$  simultaneously. Most of the change-point literature has addressed change in a single characteristic, primarily a jump, and occasionally an increase (not decrease) in residual variance. We have found no methods that address full regime changes except Moving  $F$  [2, 3] and Monte Carlo Markov chain (MCMC) methods [4].

One possible combination of events limiting all change-point methods that has not been addressed for prospective applications is the *masking of a change by a reduction in variance*. The increase in the triggering statistic of a change-point method caused by a change in  $f$  or  $\boldsymbol{\beta}$  may be offset by a decrease in  $\sigma^2$ . Although the triggering of the statistic gives evidence of a change, the absence of triggering is not conclusive evidence of non-change. Masked changes may be detected only by testing the regime, not just one parameter, accompanied by a decomposition of the regime into portions of the triggering statistic due to the various regime components.

## 3. THE MOVING $F$ METHOD

### 3.1. Example data

Prostate specific antigen (PSA) is related in a probabilistic manner to prostate cancer. According to the most frequently used clinical guidelines at the time of the example, a PSA level below 4 was normal, between 4 and 10 was equivocal, and above 10 indicated a high probability of cancer. We obtained data for patients whose prostate cancer had been treated by radiation and were at risk for recurrence. Figure 1 shows PSA readings from a man whose level was often in the equivocal range and so was monitored weekly. (For convenience, some missing data were supplied by resampling techniques, not a requirement for Moving  $F$ .) Treatment is more likely to be successful if delivered soon after a true upward turn in PSA. When PSA reach 10 at 42 months, it was accepted as adequate clinical evidence of a cancer recurrence, justifying aggressive treatment. Could this change-point have been identified sooner by statistical methods?

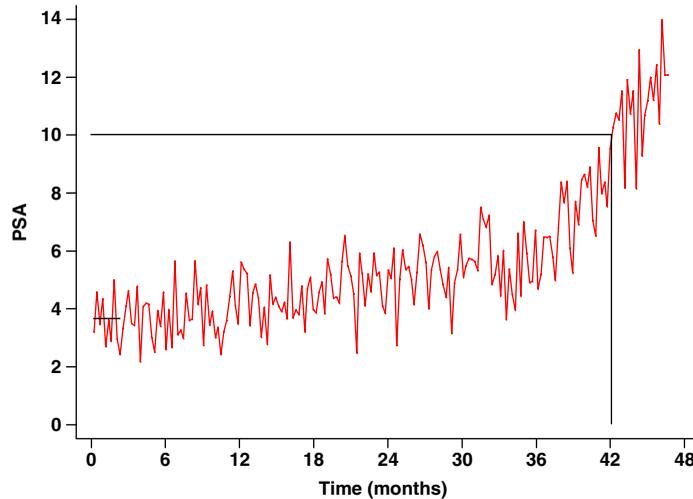


Figure 1. Weekly PSA readings on a patient with equivocal prostate condition. The clinician decided to intervene when the undulating PSA level broke 10 at month 42. (The short line segment on the beginning data represents the baseline fit to be used in the analysis.)

A reader may discern an increase in PSA level at 24 or 30 months, but that is from a retrospective view of the entire data set. To mimic a clinician's experience, the reader could move a covering piece of paper datum by datum, making a clinical decision after seeing each new datum. We believe the reader will find either a rather late decision to initiate treatment or a high false positive rate.

### 3.2. Concept of the Moving $F$ method

We estimate the nature of the series model from the first  $m$  data (baseline) of the time-series, usually from theoretical considerations or by a regression fit. The mean square of residuals about the baseline series model provides a baseline variance, say  $s_b^2$ . We choose the length of a moving sample,  $n$ , in a similar fashion. The moving average of  $n$  squared deviations from the series model provides a moving variance, say  $s_k^2$ , where  $k-n$  indicates the candidate change-point. With the assumption of normal residuals, the ratio of these variances  $s_k^2/s_b^2$  generates a Moving  $F$  with  $n-1$ ,  $m-1$  degrees of freedom. A significant change in the time series is signalled when the Moving  $F$  breaks a critical  $F$  value at some point  $t_k$  (when not significant at  $t_{k-1}$ ), detecting a change-point and estimating its location.

User defined parameters are the baseline size ( $m$ ) and moving sample size ( $n$ ), chosen from characteristics of the problem or by judgment. Assumptions are normality of residuals and correctness of the series model.

We describe Moving  $F$  as *general* because it can accommodate combinations of parameters changing simultaneously (regime changes) and can uncover masked changes.

We describe Moving  $F$  as *simple* because it can be handled by users outside the field of statistics using widely available software. For example, a medical investigator can perform a Moving  $F$  in a few steps on Excel software or on a statistics package, e.g. Stata.

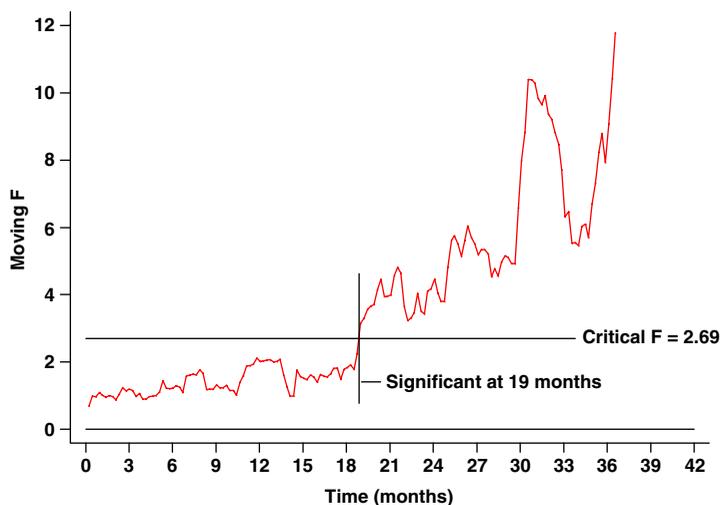


Figure 2. Moving  $F$  statistic on PSA levels (for regime as a whole). The critical  $F(12, 12, 0.95)$  is broken at 19 months, identifying a change 23 months prior to PSA level reaching 10.

### 3.3. Calculation of Moving $F$

(1) Starting with spreadsheet columns of time  $t_i$  and readings  $y_i$ , choose a baseline sample and identify  $f$ , the functional form of the baseline. The form might arise from theoretical grounds apart from statistics or, lacking theory, from a regression fit of the baseline data. Make a column of  $f$  values,  $f_i$ , if  $f$  is other than a constant. (2) Calculate  $s_b^2$ , the residual mean square of the baseline data about  $f$ . (3) Make a column  $(y_i - f_i)^2 / (n - 1)s_b^2$ . (4) The moving average of the column in step 3 is the Moving  $F$ . These steps are illustrated using data in Section 3.8.

### 3.4. Application of Moving $F$ to the example

We chose the length of the moving sample by mimicking the clinical decision making process. After prostate cancer treatment, the clinician will usually wait about three months before testing the PSA to allow the level to drop in response to the treatment, so we chose the first  $m = 13$  weekly readings as the baseline and chose a moving sample of size  $n = 13$ . Prior to recurrence, PSA level is expected to remain constant; after recurrence, it is expected to increase continually. A linear regression of the baseline data yields  $p = 0.886$  for the slope of the line with  $R^2 = 0.002$ , indicating that the PSA level was following a constant series model. We chose  $y_i = m_b + e_i$  as the model. The mean for the baseline is  $m_b = 3.6319$  and the regression residual mean square  $s_b^2 = 0.6802$ . (Urologists use only the first decimal place in PSA readings.)

We calculated a column of squared deviations from  $m_b$  divided by  $n - 1 = 12$  and divided each by  $s_b^2$ . We calculated the moving average of these elements to form the Moving  $F$ , shown in Figure 2. A horizontal line indicates the critical  $F(12, 12, 0.95) = 2.69$ . The Moving  $F$  can be seen to undulate about  $F = 1$  for about 18 months. At 19 months, the Moving  $F$  reaches 3.14, breaking the critical  $F$  to become significant for  $\alpha = 0.05$ . If this method had been used,

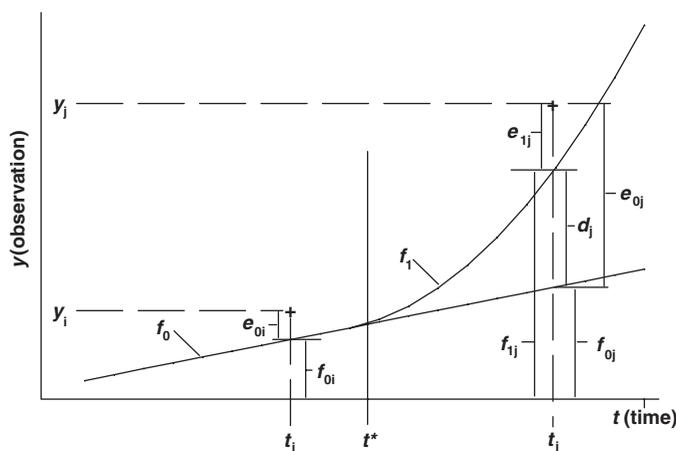


Figure 3. Relationship among observations, errors, and functional models. Datum  $(t_i, y_i)$  is a point prior to the change, datum  $(t_j, y_j)$ , after.

the clinical decision to intervene would have been made almost 2 years before the actual 42 months. The reader might note that the Moving  $F$  at 30 months gave  $p < 0.001$ . Not shown in Figure 2 is a calculated Moving  $F$  exceeding 65 by 42 months.

### 3.5. Development of the Moving $F$ method

Suppose there exists a time-dependent variable that may contain some change in regime elements, i.e. one or more of  $f$ , elements of  $\beta$ , and/or  $\sigma_i^2$ . We want to identify any of these changes. Smith's 1975 article [5] has described the problem in considerable generality. Let us denote by the subscripts 0 and 1 characteristics before and after the change-point, respectively. Symbolically

$$y_i = f_0(t_i; \beta_0) + e_{0i}, \quad t_i < t^* \quad \text{and} \quad (1)$$

$$y_i = f_1(t_i; \beta_1) + e_{1i}, \quad t_i \geq t^* \quad (2)$$

where  $e_{0i}$ ,  $e_{1i}$  are assumed  $\text{NID}(0, \sigma_0^2)$ ,  $\text{NID}(0, \sigma_1^2)$ , respectively. For shorthand, let  $f_0$ ,  $f_1$  at  $t_i$  be denoted  $f_{0i}$ ,  $f_{1i}$ , respectively, and  $d_i = f_{1i} - f_{0i}$ . Smith addressed cases with parameters known, partially known, and unknown for  $e_{0i}$ ,  $e_{1i}$  distributed as binomial and normal. In most applications treated in the literature,  $f_0$  and  $f_1$  are treated as the simplest case, i.e. constants. Moving  $F$ , by contrast, has taken  $f_0$  and  $f_1$  to be general linear models, which agrees with Smith's case 4.4. The estimates of the parameters in  $f_0$  and  $f_1$  are taken as the regression fits to the postulated functional forms. The relationship among the terms defined here may be visualized as in Figure 3.

We take a baseline sample of size  $m$  having sample variance

$$s_b^2 = \frac{1}{\eta_m} \sum_{i=1}^m (y_i - f_{0i})^2 \quad (3)$$

with  $\eta_m (= m - \text{number of parameters in } \beta_0)$  degrees of freedom (df). We take a moving sample of size  $n$  with rightmost observation  $k (\geq m+n)$  with similarly defined sample variances and df

$$s_{0k}^2 = \frac{1}{\eta_n} \sum_{i=k-n+1}^k (y_i - f_{0i})^2 \quad \text{any } t \text{ and} \quad (4)$$

$$s_{1k}^2 = \frac{1}{\eta_n} \sum_{i=k-n+1}^k (y_i - f_{1i})^2, \quad t \geq t^*$$

$k$  progresses  $m+n, m+n+1, m+n+2, \dots, L$ . The Moving  $F$  statistic is defined as

$$F_k = \frac{s_k^2}{s_b^2} \quad (5)$$

When  $t_k < t^*$ , both  $s_k^2$  and  $s_b^2$  estimate  $\sigma_0^2$ , and  $F_k$  differs from 1 only randomly. However, when  $t_k \geq t^*$ ,  $s_{0k}^2$  contains components of the regime based on  $f_1$  and may be written as

$$s_{0k}^2 = \frac{1}{\eta_n} \sum_{i=k-n+1}^k [(y_i - f_{1i}) + (f_{1i} - f_{0i})]^2 = s_{1k}^2 + \text{var}(d) + \text{cov}(f_1, d) \quad (6)$$

As the random error is uncorrelated with a change in the form of the functional model in expectation, the covariance term vanishes so that

$$E\{F_k\} = E\left\{\frac{s_{0k}^2}{s_b^2}\right\} = \frac{\sigma_1^2}{\sigma_0^2} + \frac{\sigma_d^2}{\sigma_0^2} \quad (7)$$

a change (if any) in the variability of the residual plus a change (if any) in the variability due to a model shift. The Moving  $F$  composes a moving test of the hypothesis:

$$H_0: \text{The regime components have not changed from baseline} \quad (8)$$

Indeed, when the Moving  $F$  method first appeared in 1970 in bi-directional form, it was called 'testimation', because the test consists of the estimation process: an existing estimate of a point of change implies rejection of  $H_0$ .

When a significant change occurs, the component(s) (such as mean, slope, curvature, etc. or residual variance) from which it arises is identified by analysis of the variance. The mean square due to each putative causal component is placed in ratio to the error mean square, forming  $F$  ratios. With a Moving  $F$ , the post- $t^*$  sample is subjected to a regression, providing causal components and their sample variances. The mean square due to each component and to the post- $t^*$  error mean square are placed in ratio to the pre- $t^*$  error mean square to form  $F$  ratios.

Furthermore, the change in expected  $F$  that occurs as  $t$  passes  $t^*$  is written as

$$E\{F_k - F_b\} = E\left\{\frac{s_{0k}^2(t \geq t^*)}{s_b^2} - \frac{s_{0k}^2(t < t^*)}{s_b^2}\right\} = \frac{\sigma_1^2}{\sigma_0^2} + \frac{\sigma_d^2}{\sigma_0^2} - \frac{\sigma_0^2}{\sigma_0^2} = \frac{\sigma_1^2 - \sigma_0^2}{\sigma_0^2} + \frac{\sigma_d^2}{\sigma_0^2} \quad (9)$$

We note that the change in the residual (error) variability ( $\sigma_1^2 - \sigma_0^2$ ) is either positive or negative. If this term remains unchanged while the series model changes significantly,  $F$  increases significantly. However, when the series model remains the same ( $\sigma_1^2 = 0$ ) while the residual

Table I. Set of possible expected mean square change combinations due to either the series model or residual variability or both.  $\sigma_0^2$  denotes residual mean square assuming pre-change model and  $\sigma_1^2$ , assuming post-change model.  $\sigma_\delta^2$  denotes mean square of model difference. Symbol  $\ll$  (or  $\gg$ ) denotes 'is significantly less (or greater) than'. (When this double symbol appears in two columns, only one need be significant.)

$\sigma_0^2$ versus $\sigma_1^2$	Size of $\sigma_\delta^2$	Change	Breaks significance
$\sigma_0^2 = \sigma_1^2$	$\sigma_\delta^2 \gg 0$	Change in a parameter of $f$	Upper $F$ bound
$\sigma_0^2 \ll \sigma_1^2$	$\sigma_\delta^2 = 0$	Increase in residual variability	Upper $F$ bound
$\sigma_0^2 \ll \sigma_1^2$	$\sigma_\delta^2 \gg 0$	Increase in variability <i>plus</i> a change in a parameter of $f$	Upper $F$ bound
$\sigma_0^2 \gg \sigma_1^2$	$\sigma_\delta^2 = 0$	Decrease in residual variability	Lower $F$ bound
$\sigma_0^2 \gg \sigma_1^2$	$\sigma_\delta^2 \gg 0$	Decrease in variability <i>plus</i> a change in a parameter of $f$	Either upper or lower or neither

variability about it changes significantly,  $F$  can *either increase or decrease* significantly, requiring a *two-tailed*  $F$  test. It follows that when the series model changes and the residual variability *decreases* simultaneously, the  $F$  either increases or decreases significantly or neither, depending on the relative amounts of change. It is possible that a significant change in series model occurs, but is *masked* by a decrease in residual variability. When masked significance occurs, it is uncovered graphically. The Moving  $F$  values due to error change  $(s_{1k}^2 - s_{0k}^2)/s_b^2 (= F_{ek}$ , say) and to model change  $s_{dk}^2/s_b^2 (= F_{dk})$  should be plotted along with the total regime Moving  $F(F_{rk})$ . The former, a two-tailed test, requires two critical  $F$  values, while the latter, only one. A masking is uncovered if  $F_{dk}$  is significant when  $F_{rk}$  is not. A summary of effects from relative changes in the component mean squares appears in Table I.

### 3.6. Application continued

Up to this point in the example, the Moving  $F$  has been a simple tool with obvious interpretation: calculate an  $F$  ratio and move it along until it triggers significance. Let us examine the significant change observed in the PSA data in Figure 2 for causal components and for masking.

While it seems apparent from inspecting the data that a change in the model and not in the variance is responsible for the observed significance in  $F$ , let us verify it. Figure 4 shows that  $F_{ek}$ ,  $k > t^*$ , does not show significance but  $F_{dk}$  does.  $F_{rk}$ , the Moving  $F$  shown in Figure 2, shows significance at 19 months, although it is not apparent in  $F_{dk}$  until 20 months. The significance at  $t = 19$  is masked by a change in error variability.

To decompose the Moving  $F$ , we introduce a model for  $f_1$  starting at  $t^* + 1$ . A linear regression fit to that sample yields  $p = 0.093$  for a test of a horizontal (stationary) model with  $R^2 = 0.236$  (adjusted to 0.166). These results suggest the beginnings of a slope, but one that is not yet statistically significant. The new mean is 4.32, with post-change  $s_1^2 = 0.5005$ .  $F_e = s_1^2/s_0^2 = 0.74$ , so the post-change variability is about  $\frac{3}{4}$  of that pre-change. This leaves  $F_d = F_r - F_e = 3.14 - 0.74 = 2.40$ , less than the critical  $F = 2.69$ . However,  $F_d$  is the ratio of mean square due to the model divided by  $s_0^2$ . Had this mean square been divided by the variability at the point of change,  $F_\delta$  would have been  $\frac{1}{0.75} = 1.33$  times as large, or 3.19, which would have exceeded the critical value. Thus the change in model at  $t^*$  is

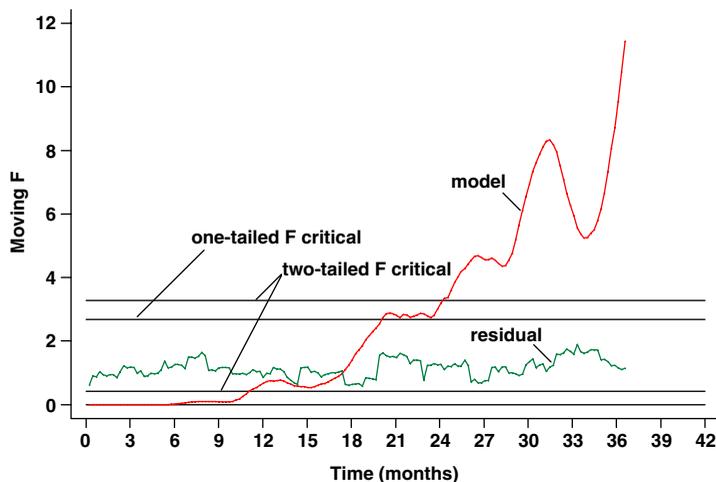


Figure 4. Moving  $F$  plots of the model and of residuals for the PSA-level data. The one-tailed critical  $F(12, 12, 0.95)$  is 2.69 and the two-tailed critical  $F$  values are 0.31 and 3.28.

significant at 19 months, but the significance has been masked by the reduction in the residual variance.

Under the (regression-satisfied) assumption of a constant as the model from 19 to, say, 34 months, the decomposition of  $F_r$  into  $F_d$  and  $F_e$  is not difficult. The components may be calculated as a moving average of the form  $[(y - m)^2 \times n/\eta]/s_0^2$  for each  $t_i > t^*$ , where  $\eta$  is the appropriate df,  $y$  is the moving average of PSA values with  $m$  being  $m_b$  for  $F_d$ , and  $y$  is the PSA value with  $m$  being the moving average of PSA values for  $F_e$ . A plot of the components thus calculated appears as Figure 4, with horizontal lines shown for the critical  $F$  values.

### 3.7. Multiplicity of tests

It could be argued that a moving sample test is actually a multiplicity of tests. However, unrecognized false positives and negatives are quite rare. A false positive would require a random spike, or even a small sequence of spikes, that would cause  $F$  to cross the critical line. Such an occurrence generates an obvious mesa (vertical-sided, flat-topped hill) in the plot if above the series model or an arroyo (its mirror image valley) if below the series model. A false negative would require a sustained reduction in error variance to mask an otherwise significant  $F$ , and has been already discussed.

### 3.8. Additional examples

Harold Shipman was a rural physician in Britain who was discovered in 1998 to have murdered a large number of his patients. From his recent article [6], Spiegelhalter provided the authors with data on the cumulative number of death certificates by sex that Shipman signed in excess of the local average, totalling 224. Starting from 1978 with the first excess, we pooled sexes and differenced to obtain total counts of excess certificates by year. These data, along with the following calculations to generate the Moving  $F$ , are given in Table II. We used the first

Table II. Number of death certificates signed by Dr Shipman by year in excess of the local average starting with the first excess in 1978, along with calculations providing a Moving  $F$ . The first 4 data are taken as baseline, with  $4.19 = s_b^2$ , the standard deviation about the average. The third column contains  $(y-f)^2/(n-1)s_b^2$ , where  $f$  is taken as 0 (null hypothesis: no excesses occur). The fourth column, a moving average of 3 of third column values, is Moving  $F$ .  $F(3, 2, 0.95) = 9.55$ .

Year	Excess	Squares	Moving $F$
1978	10.3		(1)
1979	10.3		(1)
1980	8.3		(1)
1981	6.0		(1)
1982	5.5	3.6098	2.43
1983	-0.8	0.0764	7.44
1984	12.5	18.6456	14.84
1985	14.7	25.7864	15.48
1986	4.1	2.0060	10.89
1987	6.4	4.8878	6.44
1988	10.2	12.4153	10.24
1989	10.6	13.4081	8.61
1990	0.1	0.0012	4.88
1991	3.2	1.2220	0.46
1992	-1.2	0.1718	10.65
1993	16.0	30.5489	11.77
1994	6.2	4.5871	41.79
1995	27.5	90.2446	68.86
1996	30.6	111.7375	109.60
1997	32.6	126.5210	90.48
1998	16.6	32.8831	79.85

four data as baseline, assuming a constant as a model for so small a sample and using the baseline variance as  $s_b^2$ . We calculated a Moving  $F$  of three data for the remainder, giving rise to Figure 5. It can be seen that the Moving  $F$  becomes significant in 1984. Had a Moving  $F$  death-certificate monitor been used and Shipman been stopped in 1984, 178 patients might have been saved.

An example of masked change was given by Riffenburgh [7] in which a model change is totally obscured. In this case, heart rate (HR) of trauma victims with extensive blood loss was compared for those given intrathecal *versus* the usual intravenous morphine, the latter known to impair the vascular system. Intrathecal morphine is seen to avoid raising HR. However, there is a change (reduction) in HR that was masked by a decrease in the residual mean square. When tested for masking, the reduction appeared, allowing the unanticipated conclusion that intrathecal morphine actually calmed the vascular system.

An example of decomposing  $F_{rk}$  for a non-stationary time series was given by Riffenburgh [2, 3]. If the model is known or postulated theoretically, that model can be used rather easily. However, this is often not the case in prospective analysis, where a moving regression fit to the moving sample will provide the model. This can be programmed into an analysis routine for rapid analysis, as in the case of monitoring patients in surgery or intensive care.

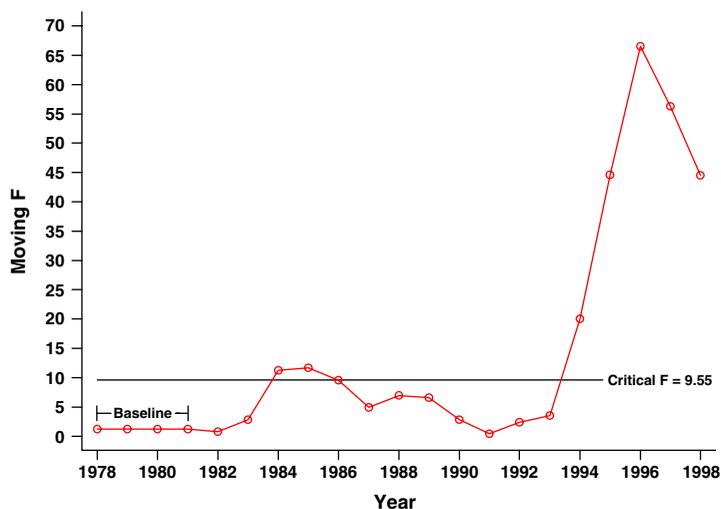


Figure 5. Moving  $F$  for Shipman's excess number of death certificates. The baseline used four points and the moving sample three. Critical  $F(3, 2, 0.95) = 9.55$ .

An additional benefit to the Moving  $F$  method, not as easily seen by other change-point methods, is the perception of various data patterns. Patterns for outliers were mentioned in Section 3.8. An example of relating a Moving  $F$  pattern to a cause may be seen in the size of the U.S. Navy Hospital Corps, where a large data burst occurs during World War II [8]. A different pattern can be seen in the PSA data in Figures 1, 2, and 4. Three high readings at about month 32 yield a sharp peak with sloped sides.

#### 4. MOVING $F$ COMPARED TO OTHER METHODS

It is convenient to categorize change identification approaches by sampling schemes: periodic, accumulating, smoothed, or moving samples.

*Periodic samples:* Periodic or sporadic sampling, e.g. quality or process control, may cost less than 100 per cent sampling and it allows destructive sampling, but it detects a change locating it only as having occurred since the last previous sample. Periodic sampling does not test regime changes nor detect masking.

*Cumulative samples:* In cumulative sampling, the analysis has the property of memory. An unusual deviation in the series is never forgotten—but so is the influence of an outlier or a past illness. Such memory is useful in some cases, but disadvantageous in others. The prominent methods are cumulative sum (CUSUM), originated by Page in 1954 [9], and exponentially weighted moving average (EWMA), originated by Roberts in 1959 [10]. The effect of memory is muted in EWMA, in which a fractional weight causes disadvantageous memory to diminish through time. Cumulative sampling can treat some subsets of regime change, but not general ones, nor does it detect masking.

*Smoothed samples:* Current-data methods, i.e. moving data not contrasted with baseline data, largely depend on detecting patterns that emerge from smoothing. Recently, Marron

and Chaudhuri [11] conceived of slope change in a moving regression fit to detect changes. Specifically, Kim and Marron (SiZer for jump detection, draft manuscript, personal communication, 2003) show that the slope forms a cusp at a change in a stationary series. Regime testing has not been developed, nor has detection of masking.

*Moving samples:* Moving samples compared to a baseline sample include not only Moving  $F$ , but also MCMC. MCMC uses a baseline to provide a Bayesian prior probability and a likelihood model. Basically, an appropriate model is posed [12], tested [13], and Monte Carlo integrated, and then carried along the time-series by Markov chain methods, specifically by the Metropolis–Hastings algorithm [14]. MCMC is capable of addressing all the specific cases documented by Smith. Moving  $F$  can also treat all of Smith's cases, except that it requires normality of the residuals. Both methods address full regime change analysis. MCMC is more general than Moving  $F$  in that it will solve a great many problems other than change-point identification. However, it requires a higher level of mathematical ability.

## 5. CONCLUSION

Moving  $F$  is, we believe, the simplest change-point method to use, and can be applied and understood by non-statisticians. Its generality is limited only by the assumption of normality of residuals about the series model, and arbitrary choices of moving sample sizes. We say simple and general in the sense of both; no other relatively simple method is as general.

Questions about Moving  $F$  remain that could be answered by further study include its robustness, e.g. sensitivity to choosing the wrong model, and its power relative to other change-point methods.

## REFERENCES

1. Quandt RE. The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the American Statistical Association* 1958; **53**:873–880.
2. Riffenburgh RH. Analyzing time-dependent characteristics when several data characteristics change at once. *Proceedings of the San Diego Biomedical Symposium* 1970; 17–22.
3. Riffenburgh RH. Regime estimation in a time sequence. *Technometrics* 1971; **13**:795–809.
4. Gilks WR, Richardson S, Spiegelhalter DJ (eds). *Markov Chain Monte Carlo in Practice*. Chapman & Hall: London, 1996.
5. Smith AFM. A Bayesian approach to inference about a change-point in a sequence of random variables. *Biometrika* 1975; **62**:407–416.
6. Spiegelhalter DJ, Best N. Shipman's statistical legacy. *Significance* 2004; **1**:10–12.
7. Riffenburgh RH. Detecting a point of any change in a time-series. *Journal of Applied Statistical Science* 1994; **1**:487–488.
8. Riffenburgh RH. *Statistics in Medicine*, Chapter 21. Academic Press: San Diego, CA, 1999; 481–490. (or 2nd Edition, Chapter 25. Elsevier: Amsterdam, 2005, in press.)
9. Page ES. Continuous inspection schemes. *Biometrika* 1954; **41**:100–115.
10. Roberts SW. Control chart tests based on geometric moving average. *Technometrics* 1959; **1**:239–250.
11. Marron JS, Chaudhuri P. Significance of features via SiZer. In *Statistical Modelling, Proceedings of 13th International Workshop on Statistical Modelling*, Marx B, Friedl H (eds), 1998; 65–75.
12. Gelfand AE. Model determination using sampling-based methods. In *Markov Chain Monte Carlo in Practice*, Gilks WR *et al.* (eds). Chapman & Hall: London, 1996; 145–161.
13. Raftery AE. Hypothesis testing and model selection. In *Markov Chain Monte Carlo in Practice*, Gilks WR *et al.* (eds). Chapman & Hall: London, 1996; 163–187.
14. Hastings WK. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika* 1970; **57**:97–109.