Chapter 4

Sequential Methods for Multistate Processes

MICHAEL I. BARÓN
University of Texas at Dallas, Richardson, Texas
IBM Research Division, Yorktown Heights, New York

4.1. INTRODUCTION AND THE GENERAL SCHEME

This paper justifies the use of sequential methods for non-sequential problems that arise in the analysis of multistate processes.

In a variety of applications, the observed dynamical system switches between different modes. Thus, the observed process is non-stationary, but it consists of homogeneous segments separated by mode switch times, or *change points*. Each mode corresponds to a particular distribution, however, the change points and the modes are usually unknown.

Examples of such models are found in quality control (in-control and out-of-control modes, [16], [18]), economics (growth and recession), energy finance (one regular state and many spike states, [5], [12] [22]), climatology (global trends like cooling and warming, [8]), developmental psychology (different phases during development, learning, problem solving, [4], [15], [19]), epidemiology (regular, pre-epidemic, epidemic, post-epidemic periods, [3], [7]), etc.

We will refer to the general class of such stochastic processes as *multistate processes.* A subclass of multistate processes with the additional assumption of a Markov chain that governs the mode switching is called *hidden Markov chains.* The latter are widely used in speech recognition, biological sequences alignment, biochemistry, genetics, and signal processing ([10], [11], [20]).

In general, a multistate process generates a sequence $\mathbf{X} = \{\mathbf{X}_j\}_{j=1}^{\lambda}$ with

$$\begin{cases} \mathbf{X}_1 = (X_{\nu_0}, \ldots, X_{\nu_1}) & \sim \ f_1 \\ \mathbf{X}_2 = (X_{\nu_1+1}, \ldots, X_{\nu_2}) & \sim \ f_2 \\ \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ \mathbf{X}_\lambda = (X_{\nu_{\lambda-1}+1}, \ldots, X_{\nu_\lambda}) & \sim \ f_\lambda \end{cases}$$

where $\mathbf{X}_j$ are single-mode segments, $f_1, \ldots, f_\lambda$ are the corrseponding known or unknown densities, ν_i, $i = 1, \ldots, \lambda - 1$, are change points (with a convention that $\nu_\lambda = N$ is the total sample size and $\nu_0 = 1$), and λ is the unknown number of switches. The modes may be repeated, but any two successive segments are assumed to belong to different modes, i.e., there are no fake change points. In the simplest situation, the process oscillates between only two possible modes. As another extreme, all segments may be generated from different densities, and each density is selected from the family according to a prior distribution.

In different applications, one may be interested in

(a) segmenting the observed sequence into homogeneous subsequences by estimating all the change points and all the modes (this is known as the *decoding problem*),

(b) estimating the first change point ν_1 only, as the time when the process went out of control, or

(c) estimating ν_1 sequentially, in order to detect the disorder "as soon as possible".

Whereas the last problem is intensively studied in literature (see [6], [16], and [25] for the survey of existing algorithms), only a few methods have been proposed for problems (a) and (b).

In applications to climatology and energy finance, described in this paper, there are infinitely many possible states. In this case, none of the proposed *off-line* algorithms seems to estimate the change points adequatly.

The overall maximum likelihood (ML) procedure ([13]) and the related *Viterbi algorithm* [20]) are likely to detect false change points. Indeed, even in the simplest case of Bernoulli variables, adding a change point forces the likelihood function to increase if and only if the sample contains at least two different observations. For the same reason, a restricted ML procedure with conditions on the number of change points or the maximum distance between them ([17]) is likely to report the maximum allowed number of change points.

A conceptually different *binary segmentation* procedure ([24]) initially assumes exactly one change point in the data set and estimates its location. It divides the observed sequence into two parts. If it is found significant, the procedure is applied to each part, etc. The deficiency of the binary segmentation scheme is that during each step, it assumes only one change point and disregards all other possible change points in the same data. Thus, a point separating two dissimilar patterns may not divide the whole data set into two dissimilar parts. This happens, for instance, in rather practical ABABA patterns, where all A-segments come from close distributions, and so do all B-samples.

Clearly, the difficulties in applying off-line estimation procedures to heterogeneous data sets are caused by the need of artificial assumptions about the number of change points.

These problems can be resolved by a suitable *sequential scheme*, that can resample the data sequentially and handle one homogeneous segment at a time. At each step, a sequential change-point detection tool is applied to detect the occurrence of the next change point. If detected, the location of the change point is estimated, after which the algorithm is applied to the post-change data. The initially obtained set of change point estimates can be refined by re-estimating each change point based on the preceeding and succeeding segments only. Significance of each change point is naturally varified by a two-sample test comparing the two adjacent segments. Segments are merged if the change point separating them is found insignificant.

This scheme does not require any restriction on the possible number of segments. Depending on the stopping boundaries at each sequential step, it may result in frequent change points. However, a change point that separates two short segments generated by close distributions will not pass the test of significance and will be deleted. Thus, the scheme is able to detect changes either sufficiently large segments provide enough evidence of a change, or when a change is dramatic. In the last case, it will be detected even from short segments (like spikes in electricity prices below). There is a possibility of reporting zero change points in the observed data, which may also be a viable result.

After the homogeneous segments of the multistate process are identified, they can be compared in terms of similarities. Similar segments can then be clustered into the same class for further analysis and more detailed calibration. For example, the processes of learning and development typically go through a sequence of stages that are rarely repeated, whereas the prices of electricity between the spikes always return to the same "inter-spike", or regular mode.

The described approach can also be used to solve problem (b) of estimating the first change point that may be followed by more complicated patterns and additional change points.

We elaborate the proposed scheme for the analysis and modeling of electricity prices in the North Atlantic region of the US and also apply it to detect global climate changes and main economic patterns from historical data.

4.2. APPLICATION: ANALYSIS AND PREDICTION OF ELECTRICITY PRICES

Deregulation and restructuring of American electricity markets caused major changes in the U.S. energy industry. Because of the recent development and particular features of power economy (lack of storage, transmission constraints, dependence on economic and climatic factors, occasional shutouts and surges in prices), the industry needs delicate statistical analysis that would result in

(a) a working stochastic model for the process of electricity prices that can be used for the Monte Carlo simulation study leading to proper valuation of energy derivatives, contracts, and physical assets;

(b) a forecast in the form of the predictive distribution of electricity prices for any given day.

Electricity prices contain a clear trend that includes seasonal, weekly, and daily variations. Also, analysis of detrended hourly electricity prices reveals nonstationary patterns, see Figure 4.2.1. In the extreme situation, during a season of peak demand, a period of consistently high temperatures, abrupt increase in the demand of energy, and/or closure or maintenance of a power plant, electricity prices experience a sudden and very significant increase that forms *a spike*. Spikes last from several hours to several days, but the price of electricity can increase tenfold during this short period. This makes the application of traditional analysis of energy finance particularly difficult.

Recently proposed models ([9], [12]) attempted to model the two-stage behavior of electricity markets by adding a Poissonian jump term

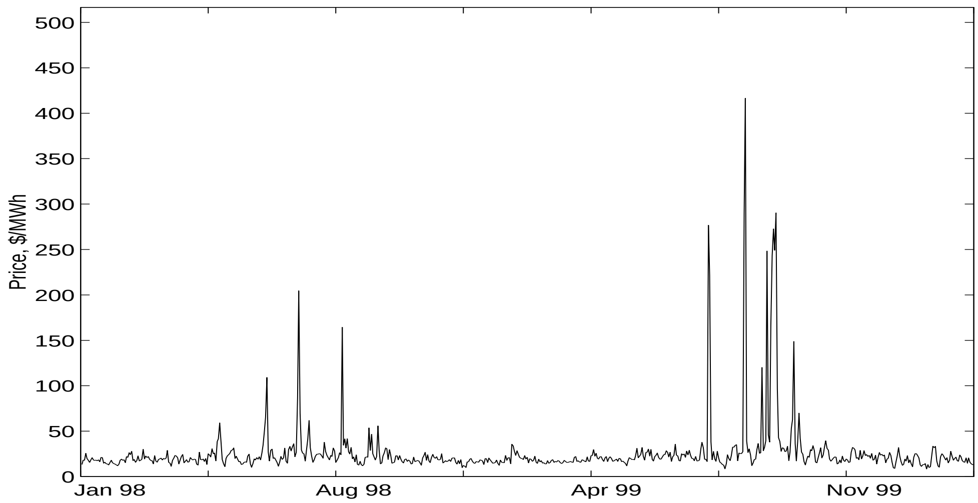

Figure 4.2.1: *Price of electricity in Pennsylvania–New Jersey–Maryland region in 1998–2000*

to the typically used mean reverting process for electricity prices. This term is intended to reflect abrupt spikes. Such models have been criticized for their insensitivity to weather, market conditions, and other factors. For example, spikes can only occur during a peak season, and only during a day, whereas according to these models, they may occur at any time.

In fact, due to the difference in supply and demand curves during spikes and inter-spike periods, the prices behave differently during these two phases of the process. Thus, they can not be effectively described by a single equation. A more realistic model is a *multistate process* where spikes and inter-spike periods form homogeneous segments, separated by *change points*. Analyzing such a model, one needs a sequential scheme described in Section 1 to identify the homogeneous segments that can then be analyzed separately.

Inter-spike (regular) mode. Analysis of data in different energy markets of the US supports the following model for electricity prices P_t between the spikes,

$$\log(P_t) = \alpha + \beta t + \gamma(m_t) + \delta(w_t) + X_t, \qquad (4.2.1)$$

where t is the day, $t = 1, 2, \ldots$, $m(t)$ is the month corresponding to t, $w(t)$ is the day of the week, γ and δ are their respective effects ($\sum_1^{12} \gamma(m) = \sum_1^7 \delta(m) = 0$), and residuals X_t form an autoregressive process

$$X_t = \phi X_{t-1} + \sigma \varepsilon_t \tag{4.2.2}$$

with respect to a normalized white noise sequence ε_t. Parameters ϕ, σ, α, β, $\gamma(m)$ and $\delta(w)$, $m = 1, \ldots, 12$, $w = 1, \ldots, 7$, may vary from one region to another (although the slope β is usually the same). They should be estimated separately for each market.

Spike modes. Although the spikes are short, electricity prices during spikes can differ very significantly (Figure 4.2.1). It is clear that different spikes are generated by different underlying distributions. Thus, the underlying multistate process has one regular mode and many spike modes, where detrended prices X_t form a sample from a lognormal distribution

$$f(x \,|\, \mu, \tau) = \frac{1}{\sqrt{2\pi} x \tau} \exp\left\{ -(\log x - \mu)^2 / 2\tau^2 \right\}.$$

Parameters μ and possibly τ are different for different spikes, and the Bayesian model with suitable prior distributions of μ and τ is appropriate. Thus, X_t are conditionally independent and identically distributed, given μ and τ. Unconditionally, they form a multistate process, and they are dependent within each spike.

This agrees with the intuition. For example, if during the price reaches \$200/MWh during a spike, then prices will be around this mark until the end of it, however, they will not affect the next spike.

Further, the prior mean of the distribution of μ varies in time,

$$\theta = \boldsymbol{E}(\mu) = \theta(t),$$

and it is high during the peak season, reducing gradually during the off-peak months.

Transitions. A two-phase Markov model ([5], [22]) differentiates between the spike and inter-spike phases, realizes the change of phases according to a non-stationary Markov chain, and associates parameters of distributions with different factors. The change of phases occurs according to a Markov chain with the transition probability matrix

$$\Pi = \begin{pmatrix} 1 - p(t) & p(t) \\ q & 1 - q \end{pmatrix}, \tag{4.2.3}$$

where $p(t)$ is the probability for a spike to start during day t and q is the probability. The frequency of spikes increases during a peak season, and the possibility of a spike is negligible during weekends. Therefore, the transition probability $p(t)$ varies depending on a day. At the same time, durations of spikes are barely predictable based on t, m_t, and w_t, and q can be assumed constant.

In the next section, we apply the proposed sequential scheme to divide the sequence of electricity prices into spikes and inter-spike segments, from which all the model parameters can be estimated.

4.3. METHODOLOGY AND IMPLEMENTATION

As proposed in Section 1, we construct a sequential change-point detection algorithm and apply it repeatedly to detect change points and isolate all the spikes. Before this algorithm is applied, the data are detrended, and possible information about the states is collected.

Detrending. Plain least squares estimates are inappropriate because they are significantly influenced by the presence of spikes. We recommend the weighted least squares method with significantly reduced weight for the most influential observations. For example, a small weight of 0.01 assigned to all observations with the studentized residual of $\log(P_t)$ exceeding 2.0 practically eliminates the effect of spikes on estimates. The resulting parameter estimates in (4.2.1) for the electricity prices in North Atlantic region are $\hat{\alpha} = 2.95$, $\hat{\beta} = 2.29 \cdot 10^{-4}$, and the month effects are given in the table,

m	1	2	3	4	5	6	7	8	9	10	11	12
$\hat{\gamma}(m)$	(.09)	(.18)	(.05)	(.02)	.05	.05	.41	.29	.02	(.07)	(.22)	(.19)

with $R^2 = 0.29$. All the interaction terms are insignificanti with p-values between 0.39 and 0.99. Also, we found no significant differences between prices on Saturday and Sunday, and during different weekdays, so that $\hat{\delta}(w) = 0.056$ for the weekdays and $\hat{\delta}(w) = -0.139$ for the weekends.

Estimation of parameters in "regular" mode. In practical change-point problems, the pre-change "in-control" distribution is usually known whereas the after-change "out-of-control" distribution is often unknown. Not having or not using the information about the "in-control" distribution leads to the loss of power of the sequential change-point detection procedure ([1], sect. 4). For the preliminary estimation, one can use the off-season (detrended) electricity prices, because, no spikes occurred during an off-peak season in 1998-2000. Fitting an AR(1) model (4.2.2) to the obtained residuals leads to the following estimates: $\hat{\phi} = 0.464$, $\hat{\sigma} = 0.178$.

Sequential segmentation. To simplify notations, at each step, we re-enumerate the data so that X_1 is the first observation after the most recently detected and estimated change point. Thus, every time, we detect the next change point based on $X_1, X_2, \ldots$.

Cusum algorithm is particularly popular for the sequential detection of change points. Various optimal features of the cusum scheme are discussed in [2], [6], and [21]. Our situation is featured by (a) nuisance parameters of distribution during spikes, and (b) autocorrelation of data during the resular mode. Thus, a modified cusum scheme is needed. Different mechanisms are used to detect transitions from the regular mode to spikes and vice versa, i.e., to detect the beginning and the end of a spike.

Detecting the beginning of a spike. In the regular mode, for each potential change point k, we use the joint distribution of detrended (pre-change) log-prices $\boldsymbol{x}_{1:k} = (x_1, \ldots, x_k)$

$$f(\boldsymbol{x}_{1:k}) = (2\pi \det \Sigma)^{-k/2} \exp\left\{-\boldsymbol{x}'_{1:k}\Sigma^{-1}\boldsymbol{x}_{1:k}/2\right\}$$

with $\Sigma_{ij} = \hat{\sigma}^2\hat{\phi}^{|i-j|}/(1-\hat{\phi}^2)$, according to the AR(1) model (4.2.2). The distribution of detrended (after-change) log-prices during spikes is normal(μ, τ^2). The mean μ differs from one spike to another, and it is assumed to have a conjugate normal(θ, η) distribution, while τ can be assumed constant for all the spikes. Under these conditions, the unconditional distribution of X_t during spikes is normal with nuisance parameters θ and $\eta + \tau^2$. As in [1], the nuisance parameters are estimated for each potential change point k. Then, the cusum stopping rule for detecting the beginning of a spike is

$$T = \inf\left\{ n \,|\, \max_{k<n} f(\boldsymbol{x}_{1:k}) \prod_{j=k+1}^{n} g_{kn}(x_j)/f(\boldsymbol{x}_{1:n}) > h \right\},$$

where

$$g_{kn}(x) = (2\pi s_{kn}^2)^{-1/2} \exp\left\{-(x-\bar{x}_{kn})^2/2s_{kn}^2\right\}, \tag{4.3.4}$$

$\bar{x}_{kn} = \frac{1}{n-k}\sum_{k+1}^{n} x_j$, and $s_{kn}^2 = \frac{1}{n-k}\sum_{k+1}^{n}(x_j - \bar{x}_{kn})^2$. The threshold h is usually chosen from the desired type I and type II error probabilities

p_I and p_{II}. For example, one can choose $h = (1 - p_{II})/p_I$, as in [14], sect. 2.2. In our setting, p_I is the probability of a false alarm, and p_{II} is the probability of failing to detect beginning of a spike.

Detecting the end of a spike. Detecting a change from the spike mode to the regular mode is quite different because all the spikes have different distributions, so there is no good estimate for the pre-change (spike) parameters. In this case, the modified cusum scheme has a considerable chance of failing to detect a change point ([1]). Thus, it is necessary to use the geometric(q) prior distribution of spike durations that follows from (4.2.3) and forces the corresponding Bayes rule to detect a change point before long.

One possibility is to construct a stopping rule as a result of a sequence of Bayes tests ([2], sect. 2),

$$T' = \inf \{n : \pi_n(\mathbf{X}) > 1 - p_I\}, \tag{4.3.5}$$

where

$$\pi_n(\boldsymbol{x}) = \boldsymbol{P}\{\nu \leq n | \boldsymbol{x}_{1:n}\} = \frac{\sum_{k=1}^{n} (1-q)^k f(\boldsymbol{x}_{k+1:n}) \prod_1^k g_{0k}(x_j)}{\sum_{k=1}^{\infty} (1-q)^k f(\boldsymbol{x}_{k+1:n}) \prod_1^k g_{0k}(x_j)}$$

is the posterior probability of a change point at $t = n$. If exact parameters of the (pre-change) spike distribution g are used in (4.3.4) instead of estimates $\bar{x}_{0k}$ and s_{0k}^2, then (4.3.5) becomes the Bayes sequential rule under the risk function $R(T, \nu) = \lambda \boldsymbol{E}(T - \nu)^+ + \boldsymbol{P}\{T < \nu\}$ ([23], sect. 4.3, [2], sect. 3).

A similar rule can be used to detect the beginning of spikes too, however, an expression for the transition probability function $p(t)$ should then be obtained first. Conversely, estimation of q is rather simple, $\hat{q} = (\text{average spike length})^{-1}$.

Refinement and final parameter estimation. During the sequential scheme, we used the minimum data to detect each change point. After the first set of estimated change points $(\hat{\nu}_1, \ldots, \hat{\nu}_{\hat{\lambda}-1})$ is obtained, each change point ν_k can be re-estimated based on $\boldsymbol{x}_{\hat{\nu}_{k-1}+1:\hat{\nu}_{k+1}}$. This subsample contains only one change point, and the maximum likelihood estimator is typically used here. However, when it follows sequential detection, the maximum likelihood estimator is *not distribution consistent*, i.e., it fails to estimate the change point with no error when the change is very significant ([4]), like the change from the regular mode to a spike. A distribution consistent estimator here is the minimizer of the p-value of the likelihood ratio test, comparing subsamples $\boldsymbol{x}_{\hat{\nu}_{k-1}+1:j}$

	Total	1998	1999
Mean spike duration $(1/q)$	2.3636	2	2.8
Mean interspike period (during the peak season)	18.7778	23	13.5
mean spike effect on log-prices θ	1.5473	1.1611	2.0107
within-spike variance τ^2	0.3730	0.1313	0.2693
between-spike variance η	0.0765	0.0648	0.1079
Transition probabilities			
$P_{peak}\{spike \to control\}$	0.4231	0.5000	0.3571
$P_{peak}\{control \to spike\}$	0.0533	0.0435	0.0740
$P_{off-peak}\{control \to control\}$	1	1	1

Table 4.3.1. Estimated parameters of the model.

and $\boldsymbol{x}_{j+1:\hat{\nu}_{k+1}}$ for $\hat{\nu}_{k-1} < j < \hat{\nu}_{k+1}$. Naturally, this scheme is repeated until the set of estimated change points stabilizes.

Results.

Applied to the electricity prices on Figure 4.2.1, the described scheme detected 11 spikes during 700 days. Spikes occurred on 05/19-21, 06/25-27, 07/21-23, 08/25, 09/15, and 09/22 in 1998, and on 06/08-09, 07/06-07, 07/24, 07/27-08/01, and 08/12-14 in 1999. The estimated parameters and other characteristics are summarized in Table 4.3.1.

Application of results. The proposed model with estimated parameters can be used to forecast the price of electricity on any given day. Examples of predictive densities are depicted on Figure 4.3.1.

Results of our analysis are also used in a Monte Carlo study that generates sequences of future electricity prices in order to estimate values of derivatives and financial deals.

4.4. OTHER APPLICATIONS

This section briefly mentions applications of the described general scheme in economics and climatology. Depending on the situation, the details of each step are changed. In economics and climatology, there are no spikes. Conversely, all the segments are rather long, and the patterns are rarely repeated.

It is interesting to determine different phases of economy based on

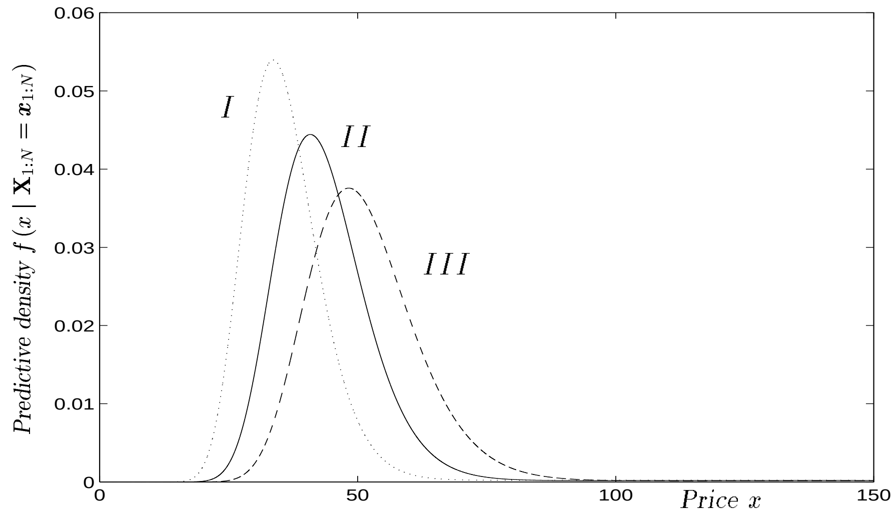

Figure 4.3.1: Predictive densities for a weekend (I) and a weekday (II) price three years ahead, and for a weekday price five years ahead (III).

the proportional daily change in the Dow Jones index. One would expect positive changes during the periods of growth and negative changes during recessions.

The proposed algorithm can also be used to detect the global climate changes. Analysis of the average temperatures in the US from 1895 till 2001 uncovers three change points, significant at the 0.002 level. The data set along with the estimated change points is on Figure 4.4.1. Although the differences among the obtained four segments are not obvious from Figure 4.4.1, the summary statistics (Table 4.4.1) shows that each transition was marked by warming, or the increase of average temperatures.

The data sets analyzed in this article are available from PJM Interconnection, L.L.C. (http://www.pjm.com/), National Climatic Data Center in Asheville, North Carolina (http://lwf.ncdc.noaa.gov/), and Yahoo! Finance (http://finance.yahoo.com/).

4.5. CONCLUDING REMARKS

A sequential scheme is proposed for the effective analysis of multi-state processes. The described scheme detects one change point at a time and divides the heterogeneous sequence into homogeneous segments that can then be analyzed separately. The general idea is to use

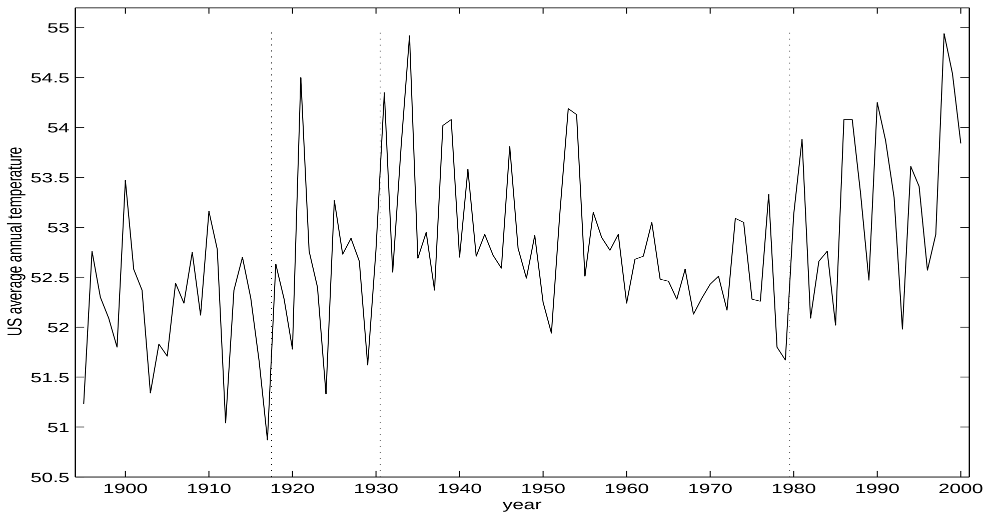

Figure 4.4.1: *Average annual temperatures in the US and three significant change points.*

Years	Temperatures		
	mean	median	standard deviation
1895–1917	52.1696	52.2900	0.6619
1918–1930	52.5869	52.6600	0.7963
1931–1979	52.8649	52.7100	0.7009
1980–2001	53.3214	53.3400	0.8466

Table 4.4.1. Summary statistics of temperatures for the obtained four periods of time.

a sequential change-point detection tool at each step, followed by re-estimation and refinement of the obtained set of change points. Details of each step, distributions of states and transition patterns are elaborated for each particular application. Estimating the parameters and understanding the mechanism that governs the change of phases sheds light on the behavior of a complicated multistate process and allows its modeling and forecasting.

REFERENCES

[1] M. Baron. Nonparametric adaptive change-point estimation and on-line detection. *Sequential Analysis*, 19:1–23, 2000.

[2] M. Baron. Bayes stopping rules in a change-point model with a random hazard rate. *Sequential Analysis*, 20:147–163, 2001.

[3] M. Baron. Bayes and asymptotically pointwise optimal stopping rules for the detection of influenza epidemics. In *C. Gatsonis, R. E. Kass, A. Carriquiry, A. Gelman, D. Higdon, D. K. Pauler and I. Verdinelli, Eds., Case Studies of Bayesian Statistics*, volume 6, New York, 2002. Springer-Verlag.

[4] M. Baron and N. Granott. Consistent estimation of early and frequent change points. In *Y. Haitovsky, H. R. Lerche, Y. Ritov, eds., Foundations of Statistical Inference, Proceedings of the Shoresh Conference 2000*, pages 181–192, Springer, 2003.

[5] M. Baron, M. Rosenberg, and N. Sidorenko. Automatic spike detection for the modeling and prediction of electricity pricing. *Energy & Power Risk Management*, 2001.

[6] M. Basseville and I. V. Nikiforov. *Detection of Abrupt Changes: Theory and Application*. PTR Prentice-Hall, Inc., 1993.

[7] C. B. Bridges, A. G. Winquist, K. Fukuda, N. J. Cox, J. A. Singleton, R. A. Strikas. Prevention and Control of Influenza. *Morbidity and Mortality Weekly Reports, Centers for Disease Control and Prevention*, 49:1–38, 2000.

[8] W. H. Calvin. *A Brain for All Seasons: Human Evolution and Abrupt Climate Change.* University of Chicago Press, Chicago, IL, 2002.

[9] L. Cleslow and C. Strickland. *Energy Derivatives: Pricing and Risk Management.* Lacima Publications, 2000.

[10] P. Cloth and R. Backofen. *Computational Molecular Biology.* Wiley, Chichester, England, 2000.

[11] S. Durbin, S. Eddy, A. Krogh, and G. Mitchison. *Biological sequence analysis. Probabilistic models of proteins and nucleic acids.* Cambridge University Press, 1998.

[12] R. Ethier and G. Dorris. Don't ignore the spikes. *Energy & Power Risk Management*, July/August, 1999.

[13] Y.-X Fu, R. N. Curnow. Maximum likelihood estimation of multiple change points. *Biometrika*, 77:563–573, 1990.

[14] Z. Govindarajulu. *The Sequential Statistical Analysis of Hypothesis Testing, Point and Interval Estimation, and Decision Theory.* American Sciences Press, Columbus, Ohio, 1987.

[15] N. Granott and J. Parziale. Microdevelopment: A process-oriented perspective for studying development and learning. In *N. Granott & J. Parziale (Eds). Microdevelopment: Transition processes in development and learning*, pages 1–28. Cambridge: Cambridge University Press, 2002.

[16] T. L. Lai. Sequential changepoint detection in quality control and dynamical systems. *J. R. Statist. Soc. B*, 57:613–658, 1995.

[17] C.-B. Lee. Nonparametric multiple change-point estimators. *Statist. Probab. Letters*, 27:295–304, 1996.

[18] D. C. Montgomery. *Introduction to Statistical Quality Control, Third Ed.* Wiley, New York, 1997.

[19] J. Piaget. Piaget's theory. In *P. H. Mussen, Carmichael's manual of child psychology*, pages 703–732, New York, 1970. Wiley.

[20] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *IEEE Proceesings*, 77:257–285, 1989.

[21] Y. Ritov. Decision theoretic optimality of CUSUM procedure. *Ann. Statist.*, 18:1464–1469, 1990.

[22] M. Rosenberg, J. D. Bryngelson, N. Sidorenko; M. Baron. Price spikes and real options: transmission valuation. In *E. I. Ronn, ed., Real Options and Energy Management*, pages 323–370, London, 2002. Risk Books.

[23] A. N. Shiryaev. *Optimal Stopping Rules.* Springer-Verlag, New York, 1978.

[24] L. Ju. Vostrikova. Detecting "disorder" in multidimensional random processes. *Sov. Math. Dokl.*, 24:55–59, 1981.

[25] S. Zacks. Survey of classical and Bayesian approaches to the change-point problem: Fixed sample and sequential procedures in testing and estimation. In *Recent Advances in Statistics*, pages 245–269, Academic Press, 1983.