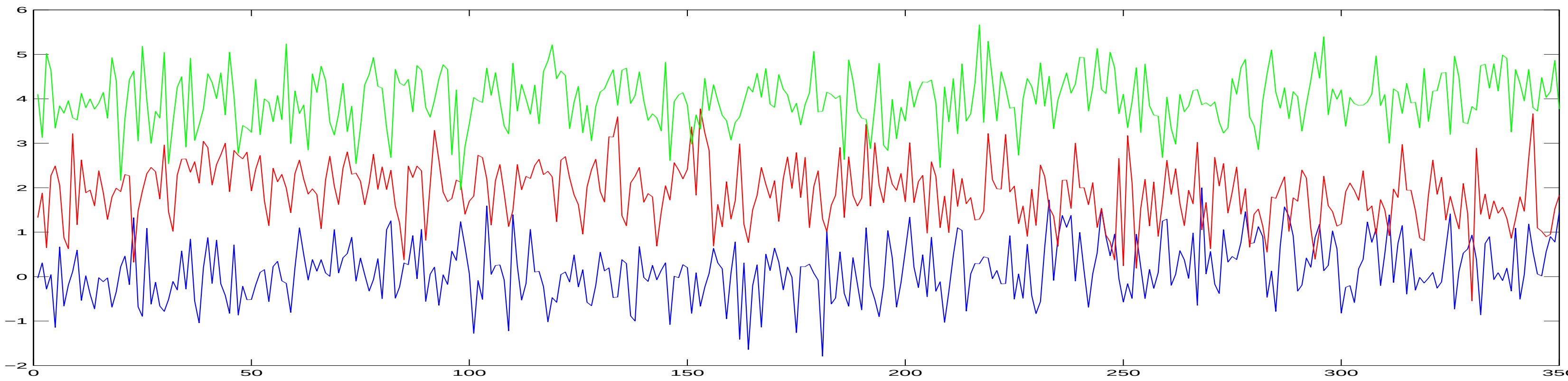
Bayesian and Asymptotically Optimal Change-Point Detection in Multivariate Time Series



Michael Baron

Department of Mathematical Sciences, University of Texas at Dallas

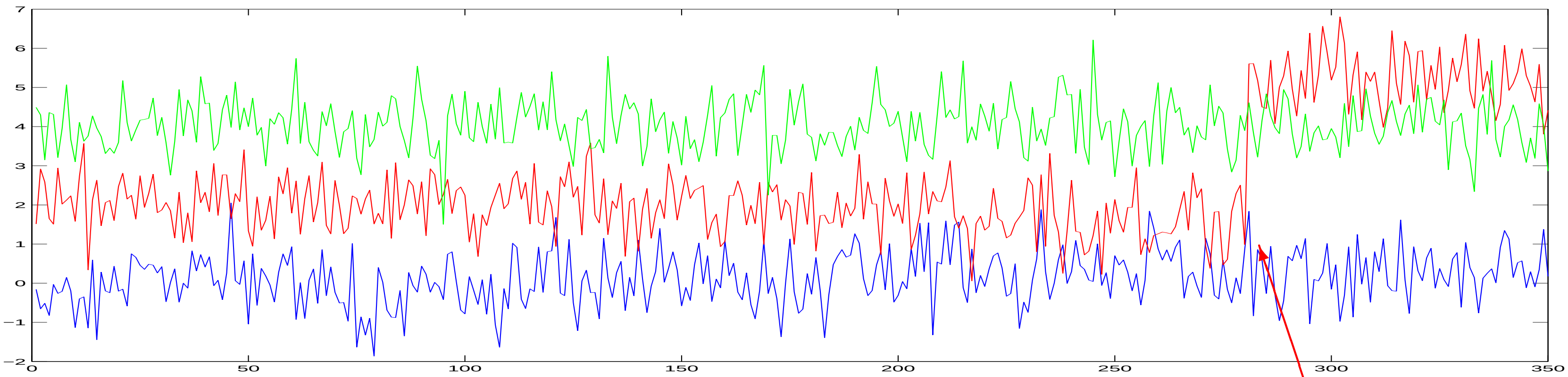
mbaron@utdallas.edu *www.utdallas.edu/∼mbaron/*

Supported by NSF grants DMS 1007775, DMS 1322353 and NSA grant H98230-11-1-0147

Recent Advances and Trends in Time Series Analysis: Nonlinear Time Series, High Dimensional Inference and Beyond. Banff Center, April 27 - May 2, 2014.

Objective

- Observe a multivariate time series sequentially



- A sudden change in one or several components may represent a threat, disorder, result of some event

Change-point (there may be a reason)

- Use statistical methods to detect such changes in real time (sequentially)

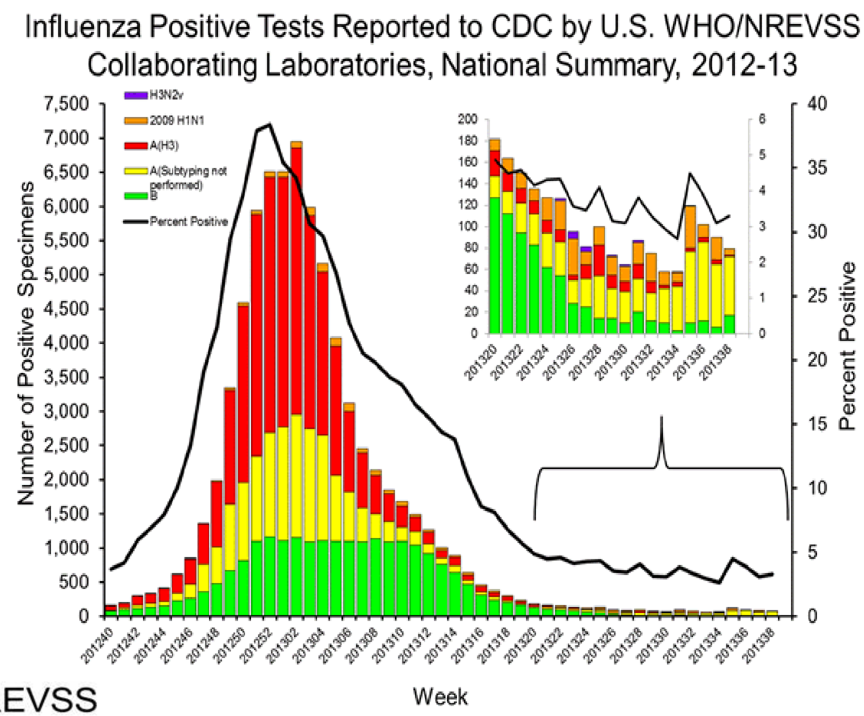
- Use the supplementary information (prior)

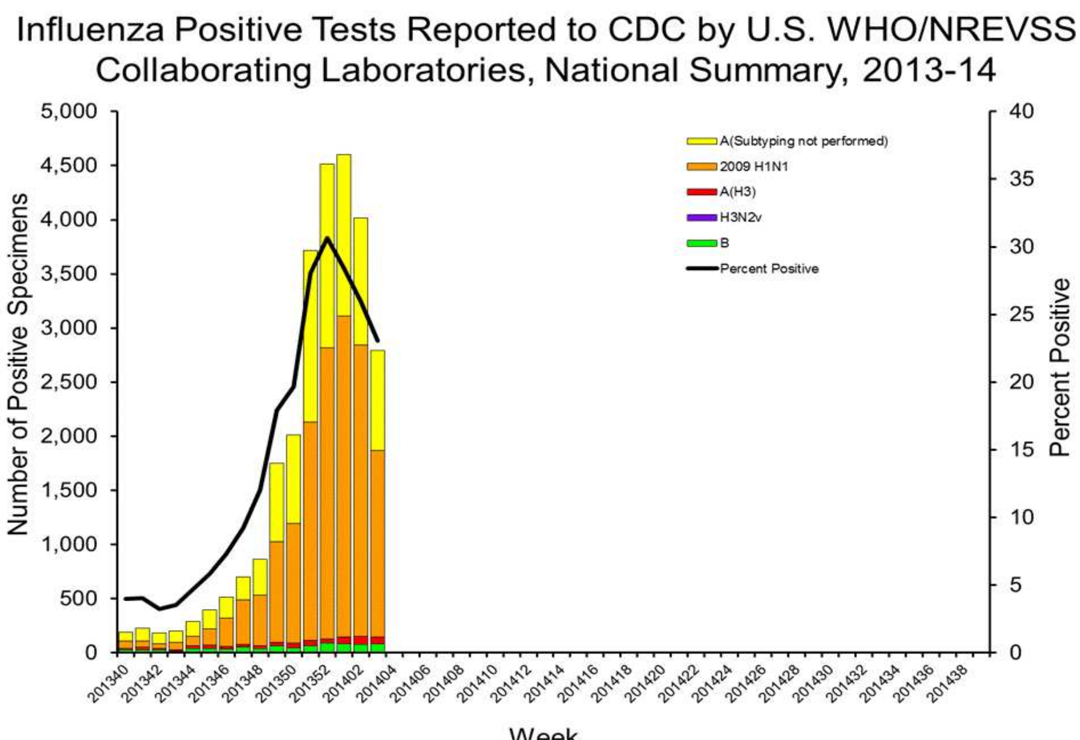
Examples

Example 1: Epidemics

More morbidity and mortality
in the U.S. than AIDS

Each year:	10,000 - 40,000 fatalities 200,000 hospitalizations 70 million work-loss days associated cost of $12 billion

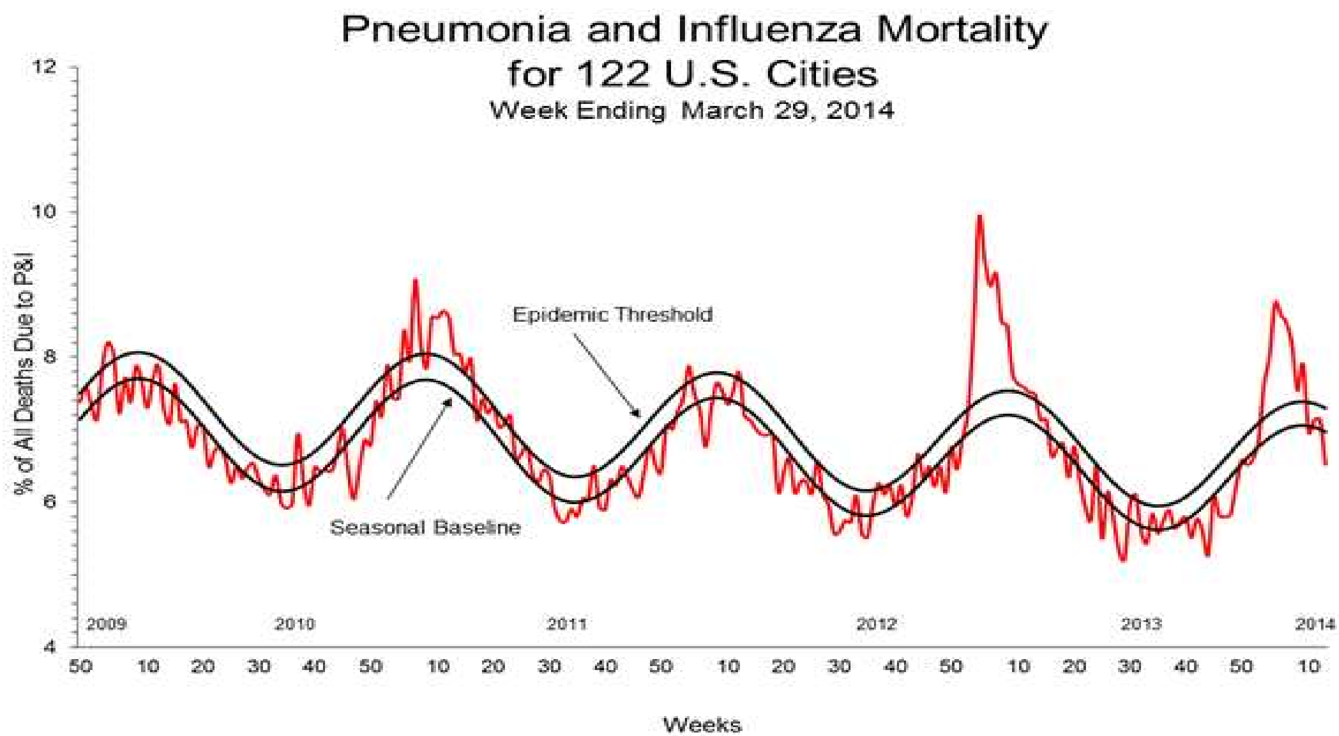




How early can we detect epidemics with 95% confidence?

Epidemics: detection and prevention

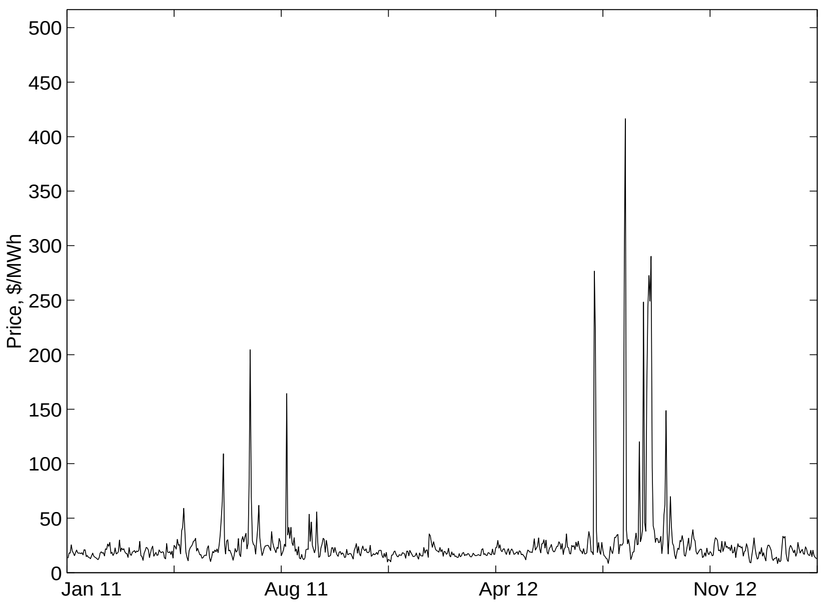
- Epidemics are official when the *epidemic threshold* is exceeded.
- Epidemic threshold = baseline + 1.645 SD



- Detect occurrence of a pre-epidemic trend
- Predict beginning of an epidemic

by solving the corresponding *change-point detection problem*

Example 2. Price of Electricity

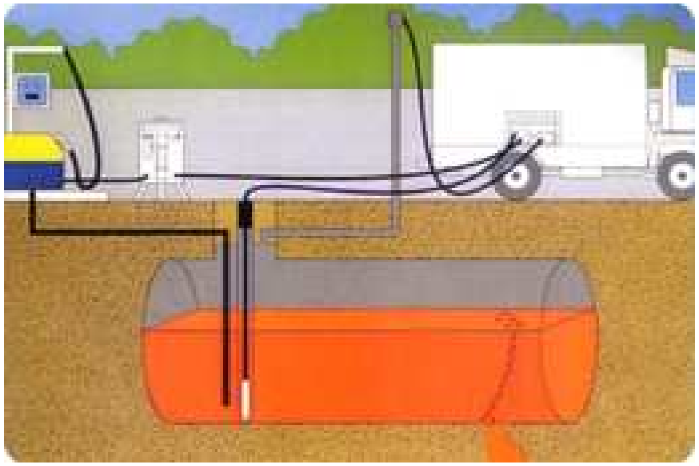


<u>Need:</u>

(a) working stochastic model
$\Rightarrow$ Monte Carlo simulation study
$\Rightarrow$ valuation of energy derivatives

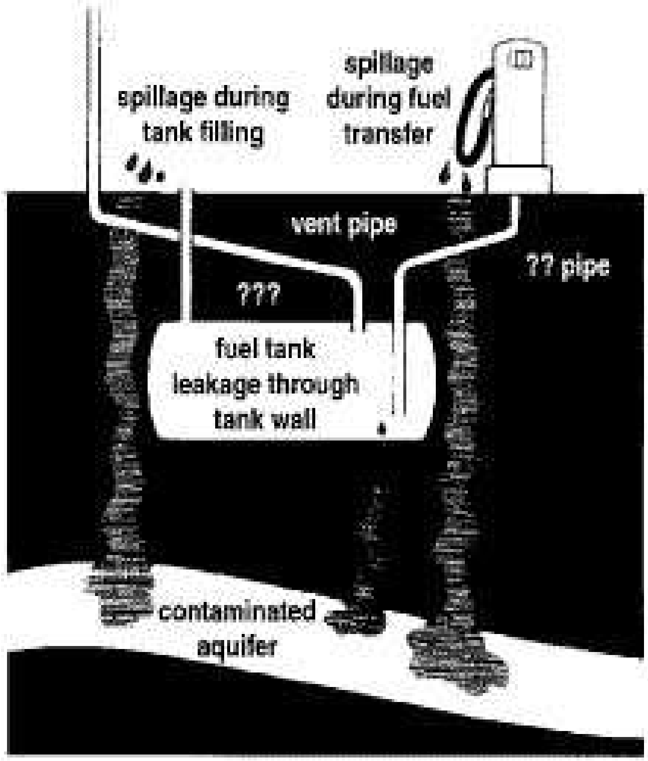
(b) forecast; predictive distribution of electricity prices for any given day

Example 3. Leaking Underground Storage Tanks



Underground storage...

leaking...



into water...

water
contamination

Current EPA requirements:

1. Detect leaks of 0.2 gallons per hour, with probability 95%
2. Maintain the probability of a false alarm of 5% per month.

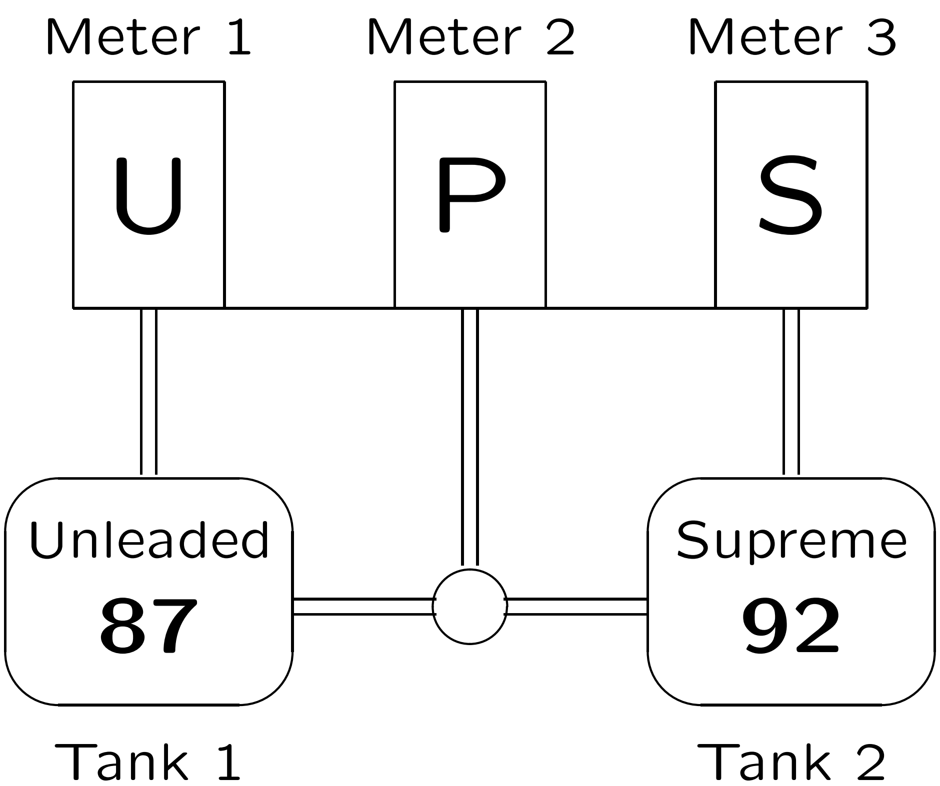
Problems for a statistician:

- Efficient detection of leaks
- Estimation of the time leaking started
- Estimation of the leaking amount

The Federal Register, 1988

Blended UST Leak Regression Model

$$
\begin{aligned}
y_{1j} &= \delta_1 + \beta_1 x_{1j} + \gamma \beta_2 x_{2j} + \epsilon_{1j} \\
y_{2j} &= \delta_2 + (1-\gamma)\beta_2 x_{2j} + \beta_3 x_{3j} + \epsilon_{2j}
\end{aligned}
$$



y_1 = displaced from Tank 1
y_2 = displaced from Tank 2

x_1 = dispensed through Meter 1
x_2 = dispensed through Meter 2
x_3 = dispensed through Meter 3

Intercepts δ_1, δ_2 = amount displaced from UST but not dispensed (leak rates)

Slopes $\beta_i = \frac{\Delta y}{\Delta x_i}$ are changes in $y = y_1 + y_2$ for a given change in x_i

Gill, Keating, Baron 2006

General Model and Assumptions

A multidimensional stochastic sequence $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots \in \Re^d$, possibly a time series

$(2d)$ families of joint distributions, $\mathcal{F}_j$ and $\mathcal{G}_j$, $j = 1, \ldots, d$

Before the change in channel j,	$\text{dist}(\boldsymbol{X}_{j,1:\nu_j})$	$\in$	$\mathcal{F}_j$
After the change,	$\text{dist}(\boldsymbol{X}_{j,\nu_j+1:t})$	$\in$	$\mathcal{G}_j$

Equivalent model: *there are (2d) stochastic sequences*, F_{θ_j} and G_{η_j}

$$\{U_{tj}\} \sim F_{\theta_j} \in \mathcal{F}_j \quad \text{and} \quad \{V_{tj}\} \sim G_{\eta_j} \in \mathcal{G}_j,$$

and

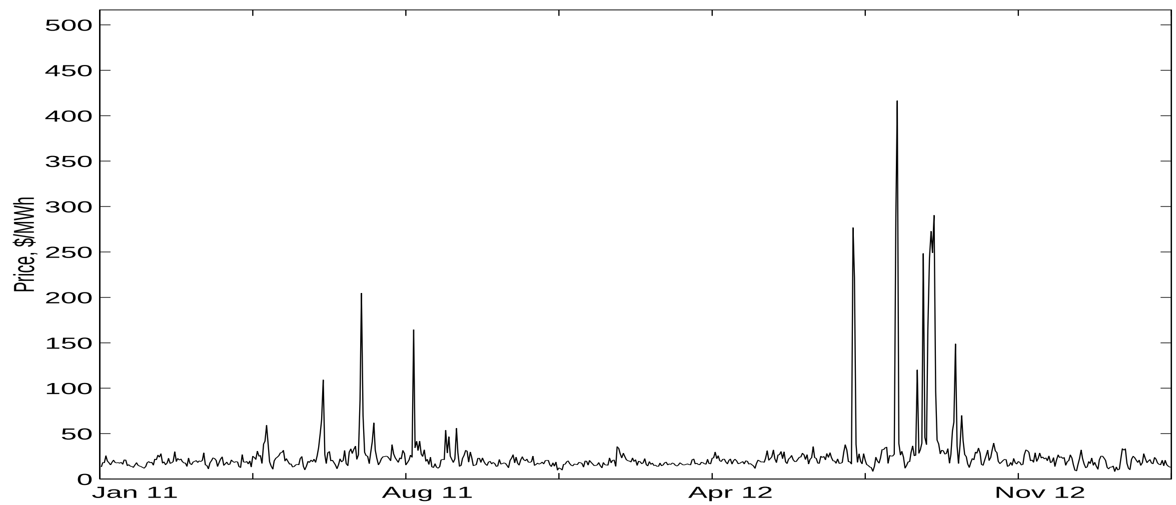
$$X_{tj} = \begin{cases} U_{tj} & \text{for} \quad t \leq \nu_j \\ V_{tj} & \text{for} \quad t > \nu_j \end{cases} ; \qquad \nu_j = \text{change-point in channel } j$$

Prior distributions: $(\nu_1, \ldots, \nu_d) \sim \pi_\nu$, $\theta_j \sim \pi_\theta$, $\eta_j \sim \pi_\eta$

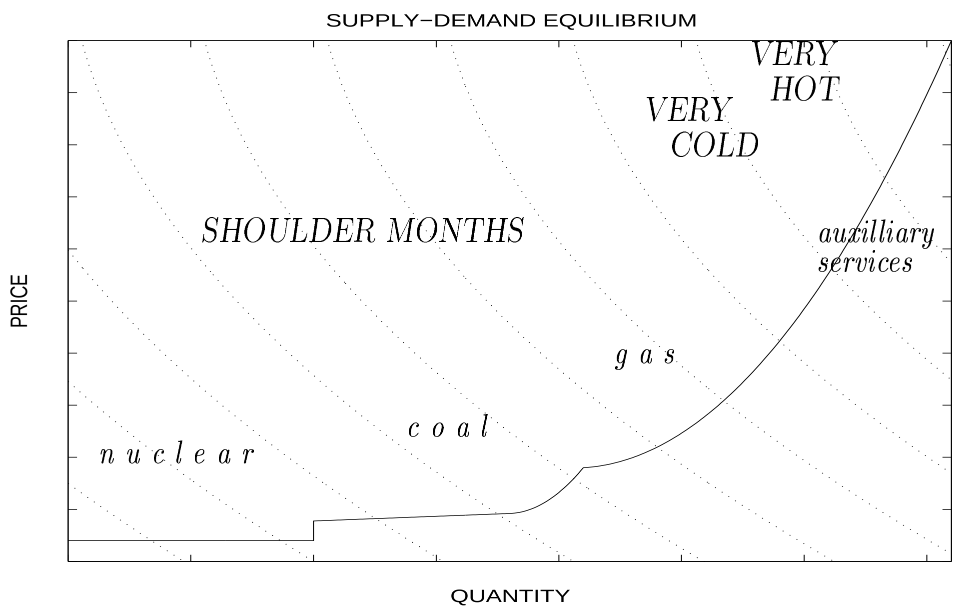
Prior distribution

Electricity prices

The data,
spot prices
of electricity,
show multiple
change-points



Reasons for change-points:



Spikes are likely to start:

- during a season of peak demand
- during a day
- during a weekday
- during unusually hot weather
- during closure or maintanence of a power plant

Prior distribution

Influenza epidemics

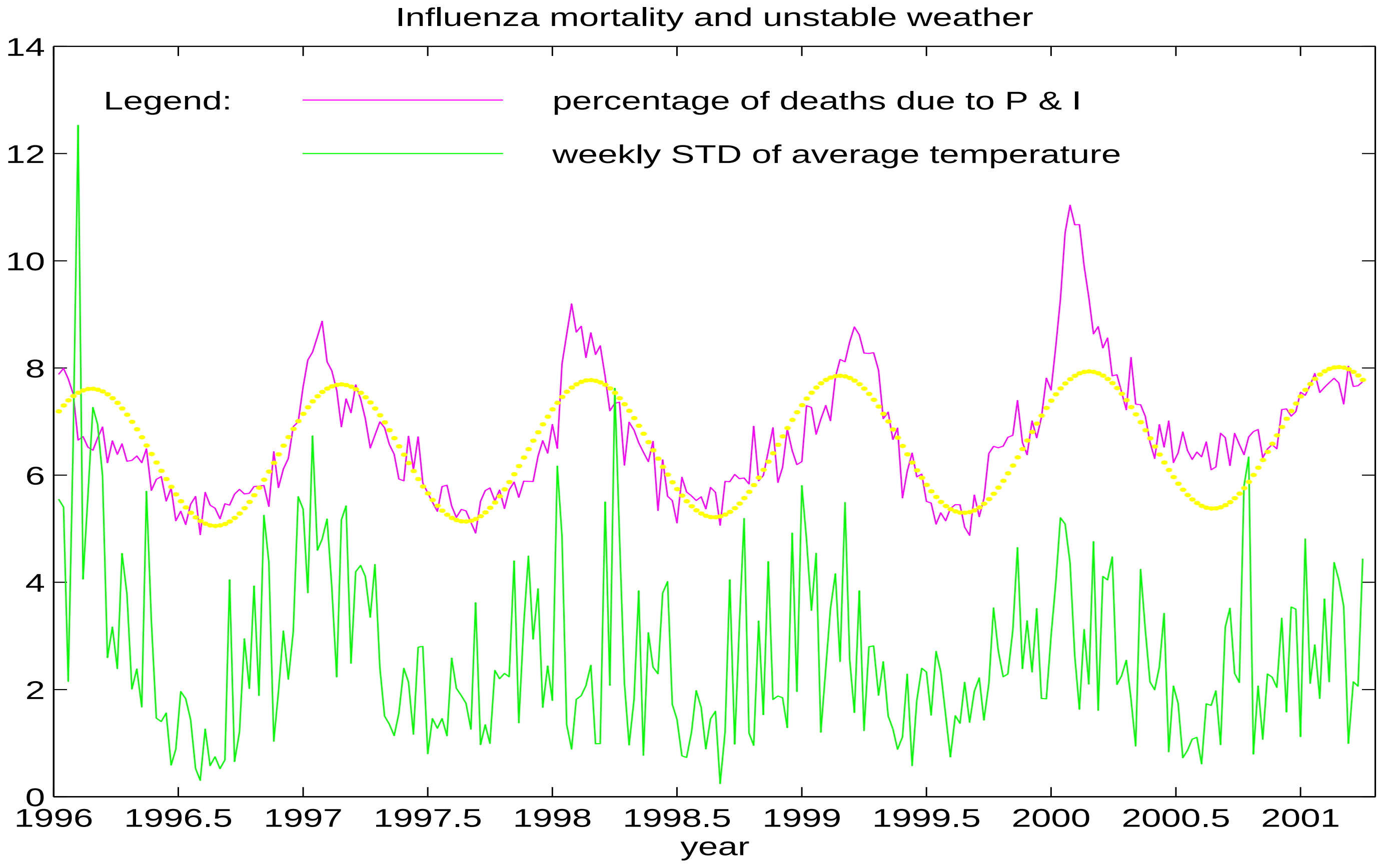
Significant factors:

- weather
- air pollution
- pollen
- ozone level
- others

Examples

cold front followed by high pressure $\Downarrow$ likelihood of epidemic $\uparrow$	high humidity, vigorous air movement $\Downarrow$ virus dies rapidly $\Downarrow$ likelihood of epidemic $\downarrow$

Influenza mortality and weather conditions



Bayes sequential change-point detection (theory)

Shiryaev (1978)

Under the loss function $L_1(T,\nu) = \lambda(T-\nu)^+ + I_{\{T<\nu\}}$

and the *Geometric* prior $\pi_n = \boldsymbol{P}\{\nu = n\} = pq^n$,

$$T^* = \inf\{n \geq 0 \mid \Pi_n \geq \pi^*\}$$

is *Bayes*, where

$$\begin{aligned}\Pi_n &= \boldsymbol{P}\{\nu \leq n \mid X_1, \ldots, X_n\}, \\ \pi^* &= \arg\min_{\pi} s(\pi),\end{aligned}$$

and

$$s(\pi) = \sup_{\{T\}} \left(-\boldsymbol{E}^{\boldsymbol{X},\nu}\{L_1(T,\nu) \mid \Pi_0 = \pi\} \right)$$

is *the payoff function*.

Ritov (1990)

CUSUM stopping rule $\quad T(h) = \inf\{n : W_n \geq h\}$,

$$W_n = \max_{k \in [0;n)} \log \frac{\mathcal{L}(X_1, \ldots, X_n \mid \text{change at } \nu = k)}{\mathcal{L}(X_1, \ldots, X_n \mid \text{no change so far})},$$

is *Bayes* under the loss function

$$L_2(T, \nu) = C_1 I_{\{T < \nu\}} + C_2 (T - \nu)^+ - C_3 \min\{T, \nu\}$$

and the prior

$$\boldsymbol{P}\{\nu = n \mid \nu \geq n, X_1, \ldots, X_n\} = p\left(1 - \frac{f}{g}(X_n) e^{W_{n-1}}\right)^+$$

Bayes stopping rules

Classical assumptions	In practice
• IID observations before and after the change point	• Observed time series, stochastic process of a known type
• Known distributions, or known family, or use nonparametrics	• Always there are nuisance parameters
• Geometric prior or some special prior	• Lots of prior information, side factors, complex prior

There is a need for weaker assumptions...

Hierarchical Bayes stopping rules

Assumption: $\phi_n = \boldsymbol{P}\{\nu = n \mid \nu \geq n\}$ = homogeneous Markov chain

(includes the cases of Shiryaev 1978 and Ritov 1990)

Result (applying theory of optimal stopping):

$\exists$ functions $\eta(\cdot)$ and $\zeta(\cdot)$ and a Markov sequence $\{U_n\}_{n=0}^{\infty}$, with

$$\boldsymbol{E}^{\boldsymbol{X},\nu} L(T,\nu) = \boldsymbol{E}\{\sum_{n<T} \eta(U_n) - \zeta(U_T)\},$$

then $T^* = \inf\{n \mid s(U_n) = \zeta(U_n)\}$ is Bayes,

where $s(u) = \sup_{\{T\}}(-Risk \mid U_0 = u)$ = payoff function

It solves the fixed-point equation

$$s(y) = \max\left\{\zeta(u),\ \boldsymbol{E}\left\{s(U_{n+1}) \mid U_n = u\right\} - \eta(u)\right\}$$

M.B. 2001

Implementation

Numerical solution: define operators

$$\begin{aligned} \mathcal{T}w(u) &= \boldsymbol{E}\left\{w(U_{n+1}) \mid U_n = (u)\right\} \\ Qw(u) &= \max\left\{w(u), \mathcal{T}w(u) - \eta(u)\right\}. \end{aligned}$$

Compute

$$\begin{aligned} Q^1\zeta(u) &= Q\zeta(u), \\ Q^2\zeta(u) &= QQ\zeta(u), \\ Q^3\zeta(u) &= QQQ\zeta(u), \\ Q^4\zeta(u) &= QQQQ\zeta(u), \end{aligned}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$s(u) = \lim_{N \to \infty} Q^N \zeta(u)$$

Large number of iterations on a lattice $\Rightarrow$ payoff $s(u)$.

Computing the critical probability: $\pi^* = \arg\min s(u)$

Computing the stopping time: $T^* = \inf\left\{n \geq 0 \mid \Pi_n \geq \pi^*\right\}$

Bayes rules - energy pricing

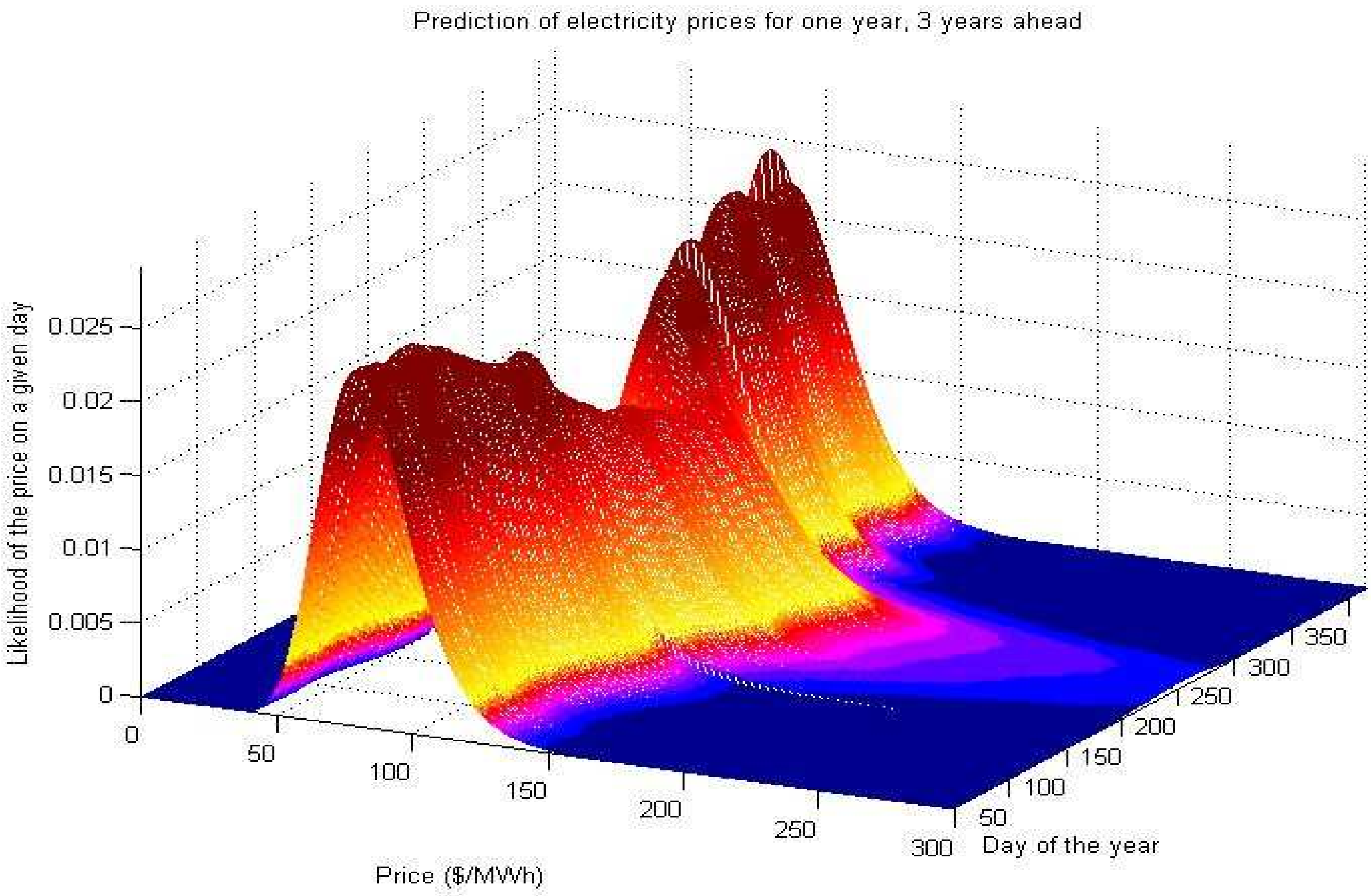
Trend $\log(P_t) = \alpha + \beta t + \gamma(m_t) + \delta(w_t) + X_t$

Regular mode $X_t = \phi X_{t-1} + \sigma\varepsilon_t$

Spikes $X_t \sim N(\mu(spike), \tau)$, where $\mu \sim N(\theta, \eta)$

	Estimates
Mean spike duration $(1/q)$	2.3636
Mean interspike period (during the peak season)	18.7778
mean spike effect on log-prices θ	1.5473
within-spike variance τ	0.3730
between-spike variance η	0.0765
Transition probabilities	
$P_{\text{peak}}\{spike \to control\}$	0.4231
$P_{\text{peak}}\{control \to spike\}$	0.0533
$P_{\text{off-peak}}\{control \to control\}$	1

Prediction for one year (predictive densities for each day)



Exact Bayes solution? Limitations of the method:

- Computation is complicated and slow.
- Practically, it is an approximate solution, by iterations
- No general algorithm of computing $s(u)$ accurately.
- The prior ϕ_t must form a *very simple* Markov chain
- Nuisance parameters are not allowed. We estimated them and "pretended" that they are true in the *naive empirical Bayes* way.

What are the alternatives?

Alternative 1:

Sequential testing

For $n = 1, 2, 3, 4, \ldots\ldots,$ test $H_0 : \underset{\text{(no change)}}{\nu > n}$ vs $H_A : \underset{\text{(}\exists\text{ change)}}{\nu \leq n}$

where $T(\alpha) = \inf\{n \ : \ H_A \text{ is rejected } \}$

- Sequential probability ratio test $\Rightarrow$ CUSUM

$$\begin{cases} W_n < h & \Rightarrow \text{ accept } H_0 \text{ and take } X_{n+1} \\ W_n \geq h & \Rightarrow \text{ reject } H_0 \text{ and stop sampling} \end{cases}$$

- Bayes test $\Rightarrow$?

- Bayes test $\Rightarrow$

  Reject H_0 when $\boldsymbol{P}\{\nu > n \mid X_1, \ldots, X_n\} \leq \alpha$, i.e., when $\Pi_n \geq 1-\alpha$

$$T(\alpha) = \inf \left\{ n : \sum_{k<n} \frac{\phi_k \rho(k,n)}{\prod_k^n (1-\phi_i)} > \frac{1-\alpha}{\alpha} \right\}$$

This is a level α test for each n. But until T, $\boldsymbol{P}\{$ false alarm $\} > \alpha$.

The Bayes stopping rule has the same form: $\inf\{n : \Pi_n \geq \pi^*\}$ in all the solved cases, and it *minimizes the Bayes risk*.

Alternative 2:

Asymptotically pointwise optimal rules

For the risk

$$R(T,\theta) = \boldsymbol{E}\left\{L(T,\theta,\delta) + cT\right\} = \boldsymbol{E}\left\{loss + cost\right\},$$

a stopping rule T is APO if

$$\limsup_{c \downarrow 0} \frac{\inf_\delta \boldsymbol{E}\left\{L(T,\theta,\delta) \mid \boldsymbol{X}_{1:T}\right\} + cT}{\inf_\delta \boldsymbol{E}\left\{L(U,\theta,\delta) \mid \boldsymbol{X}_{1:U}\right\} + cU} \leq 1$$

a.s. for any stopping rule U.

Bickel, Yahav 1967, 1968
Ghosh, Mukhopadhyay, Sen 1997 [sec. 5.4]

Asymptotically pointwise optimal rules

Theorem (Bickel, Yahav)

If $N^{\beta}\boldsymbol{E}\left\{L(N,\theta,\delta) \mid \boldsymbol{X}_{1:N}\right\} \to V$ a.s. for some $\beta, V > 0$, then

$$T = \inf\left\{n \;\middle|\; \frac{\boldsymbol{E}\left\{L(n,\theta,\delta) \mid \boldsymbol{X}_{1:n}\right\}}{n} \leq \frac{c}{\beta}\right\}$$

is APO.

Translating to multichannel change-point problems:

- Replace cT with the *delay term* $\sum_{j=1}^{d} \lambda_j \boldsymbol{E}(T - \nu_j)^+$

(penalty increases with each missed change-point)

- Add penalty for a *false alarm* $\mathrm{PFA} = \boldsymbol{P}\left\{T < \min_j \nu_j\right\}$

- *Risk* $R(T, \nu_1, \ldots, \nu_d) = \sum_{j=1}^{d} \lambda_j \boldsymbol{E}\left\{(T - \nu_j)^+ \mid \boldsymbol{X}_{1:t}\right\} - \log^{-1} S_X(T)$,

$S_X(t) = \boldsymbol{P}^\pi\left\{\bigcap_j (\nu_j > t) \mid \boldsymbol{X}_{1:t}\right\} =$ posterior survival function of $\min \nu_j$

Call stopping time T asymptotically pointwise optimal if

$$\limsup_{\lambda \downarrow 0} \frac{R(T, \nu_1, \ldots, \nu_d)}{R(U, \nu_1, \ldots, \nu_d)} \leq 1 \text{ a.s., for any stopping rule } U.$$

Theorem (known distributions)

Let

$r_t = -\log^{-1} S_X(t)$, "posterior expected loss",

$\rho_{1j} = \frac{g_j(X_{1j})}{f_j(X_{1j})}$, $\rho_{t+1,j} = \frac{g_j(X_{t+1,j}|\boldsymbol{X}_{1:t,j})}{f_j(X_{t+1,j}|\boldsymbol{X}_{1:t,j})}$, log-likelihood ratios,

and assume a strong law of large numbers

$$t^{-1} \sum_{k=1}^{t} \log \rho_{kj} \to K_j > 0, \text{ as } t \to \infty, \ G\text{-a.s., for } j = 1, ..., d$$

This condition holds for the IID case with $K_j = K(G_j, F_j)$, and also for finite-range dependent identically distributed variables, asymptotically quadrant independent or sub-independent sequences, invertible ARMA time series, L_p-mixingales, L_p-NED (near-epoch-dependent in L_p norm), strong mixing, quasi-stationary, or extended negatively dependent sequences, and other scenarios.

Then there exists an a.s. limit

(a) $\lim_{t\to\infty}(tr_t) = \dfrac{1}{K}$ if $t^{-1}|\log S(t)| \to 0$

(b) $\lim_{t\to\infty}(tr_t) = \dfrac{1}{K+L}$ if $t^{-1}|\log S(t)| \to L > 0$

(c) $\lim_{t\to\infty}(t^{\beta} r_t) = \dfrac{1}{L}$ if $t^{-\beta}|\log S(t)| \to L > 0$

for some $\beta > 1$, where $S(t) = \boldsymbol{P}\left\{\bigcap(\nu_j > t)\right\}$ is the prior survival function of the first change-point $\min \nu_j$.

Theorem (*the form of APO stopping rules*)

Under (a) or (b), the stopping rule

$$\tilde{T} = \inf \left\{ t \;\middle|\; S_X(t) \le \exp\left(-\frac{1}{t \sum \lambda_j}\right) \right\}$$

is *APO*. Under condition (c), the stopping rule

$$\tilde{T} = \inf \left\{ t \;\middle|\; S_X(t) \le \exp\left(-\frac{\beta}{t \sum \lambda_j}\right) \right\}$$

is *APO*.

The case of nuisance parameters

Assume $\theta_1, \ldots, \theta_d \in \Theta$ and $\eta_1, \ldots, \eta_d \in H$, unknown nuisance parameters, so that $F_j = F_{\theta_j} \in \mathcal{F}$ and $G_j = G_{\eta_j} \in \mathcal{G}$;

Priors: $\theta_j \sim \pi_\theta$, $\eta_j \sim \pi_\eta$, independently of ν.

Define marginal and conditional densities, free of nuisance parameters,

$$f_j^*(\boldsymbol{X}_{1:t,j}) = \int f_{\theta_j}(\boldsymbol{X}_{1:t,j} \mid \theta) d\pi_{\theta_j}(\theta), \;\; f_j^*(X_{t+1,j} | \boldsymbol{X}_{1:t,j}) = \frac{f_j^*(\boldsymbol{X}_{1:t+1,j})}{f_j^*(\boldsymbol{X}_{1:t,j})},$$

$$g_j^*(\boldsymbol{X}_{1:t,j}) = \int g_{\eta_j}(\boldsymbol{X}_{1:t,j} \mid \eta) d\pi_{\eta_j}(\eta), \;\; g_j^*(X_{t+1,j} | \boldsymbol{X}_{1:t,j}) = \frac{g_j^*(\boldsymbol{X}_{1:t+1,j})}{g_j^*(\boldsymbol{X}_{1:t,j})}.$$

Now detect changes from F_j^* to G_j^*, $j = 1, ..., d$

Consider $\rho^*_{1j} = \frac{g^*_j(X_{1j})}{f^*_j(X_{1j})}$, $\rho^*_{t+1,j} = \frac{g^*_j(X_{t+1,j}|\boldsymbol{X}_{1:t,j})}{f^*_j(X_{t+1,j}|\boldsymbol{X}^{1:t,j})} =$ marginal likelihood ratios

Theorem (APO, under nuisance parameters)

Assume SLLN

$$t^{-1} \log \rho^*_{tj} \to K_j > 0, \textit{ as } t \to \infty, \ G_\eta\textit{-a.s., for all } \eta \in H.$$

Then

(1) the stopping rule $\tilde{T}^* = \inf \left\{ t \mid S^*_X(t) \leq \exp\left(-1/t \sum \lambda_j\right)\right\}$ is APO under conditions (a) or (b);

(2) the stopping rule $\tilde{T}^* = \inf \left\{ t \mid S^*_X(t) \leq \exp\left(-\beta/t \sum \lambda_j\right)\right\}$ is APO under condition (c),

where $S^*_X(t) = \boldsymbol{P}\left\{\min \nu_j > t|\boldsymbol{X}_{1:t}\right\}$ is the *marginal* (parameter-free) posterior survival function of the earliest change point.

Multichannel change-point detection in multivariate time series with nuisance parameters

Multivariate AR process

$$\boldsymbol{X}_t - \boldsymbol{\mu}_t = \Phi(\boldsymbol{X}_{t-1} - \boldsymbol{\mu}_{t-1}) + \Sigma^{1/2}\boldsymbol{Z}_t, \qquad \boldsymbol{Z}_t \sim N(\boldsymbol{0}, I) \in \Re^d$$

The jth mean changes from $\mu_0^{(j)}$ to $\mu_0^{(j)} + \delta_j$ at time ν_j.

Let $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d)$ have a *Negative Multinomial* $(k_0; p_0, \ldots, p_d)$ prior distribution,

$$\pi_{\boldsymbol{k}} = \boldsymbol{P}\{\nu_1 = k_1, \ldots, \nu_d = k_d\} = \begin{pmatrix} \sum_{j=0}^{d} k_j - 1 \\ k_0 - 1, k_1, \ldots, k_d \end{pmatrix} \prod_{j=0}^{d} p_j^{k_j}$$

Verify assumptions of the Theorem

1. $\lim_{t\to\infty}\left(-\frac{\log S(t)}{t}\right) = \lim_{t\to\infty}\left(-\frac{\log \boldsymbol{P}\left\{\min \nu_j > t\right\}}{t}\right) = d\log\left(1+\frac{1}{d}\right)$ (case $p_0 = \ldots = p_d$)

2. $\frac{1}{n}\sum_{t=1}^{n} \log \rho_{[e]}(t) \to \boldsymbol{E}_G \log \rho_{[e]}(t) = K_{\boldsymbol{e}}$, as $t\to\infty$, with probability one, where

$$K_{\boldsymbol{e}} = \boldsymbol{e}' D(I-\Phi)'\Sigma^{-1}(I-\Phi)(\boldsymbol{\mu_0}+\boldsymbol{\delta}) - \left(\boldsymbol{\mu_0}+\frac{1}{2}D\boldsymbol{e}\right)'(I-\Phi)'\Sigma^{-1}(I-\Phi)D\boldsymbol{e}.$$

$D = Diag(\delta_1,\ldots,\delta_d)$, $\boldsymbol{e} = (e_1,\ldots,e_d)$, $e_j = 0$ if $\mu^{(j)} = \mu_0$, $e_j = 1$ if $\mu^{(j)} = \mu_0 + \delta_j$

Conditions are verified with $\beta = 1$. Then

$$T = \inf\left\{t : S_X(t) \le e^{-\left(t\sum\lambda_j\right)^{-1}}\right\}$$

is an APO stopping time for multichannel change-point detection.

Nuisance parameters

Assume that the pre- and post-change location parameters of the observed time series $\boldsymbol{X}_t$ are unknown, and its expected value $\boldsymbol{\mu}_t$ has a prior distribution,

$$\boldsymbol{\mu}_t \sim \textit{Multivariate Normal } (\boldsymbol{M} + D\boldsymbol{e}(t), \Xi);$$

$D\boldsymbol{e}(t)$ is a vector shift by the time t. Under these conditions, the marginal (parameter-free) log-likelihood ratio is

$$\log \rho^*_{[\boldsymbol{e}]} = \boldsymbol{y}'_{\boldsymbol{e}}(I - \Phi B)\boldsymbol{X}_t - c_{\boldsymbol{e}},$$

where

$$\begin{aligned} \boldsymbol{y}_{\boldsymbol{e}} &= \left[(I-\Phi)\Xi(I-\Phi)' + \Sigma\right]^{-1}(I-\Phi)D\boldsymbol{e}, \text{ and} \\ c_{\boldsymbol{e}} &= (\boldsymbol{\mu}_{\boldsymbol{0}} + D\boldsymbol{e}/2)'(I-\Phi)'\left[(I-\Phi)\Xi(I-\Phi)' + \Sigma\right]^{-1}(I-\Phi)D\boldsymbol{e}, \end{aligned}$$

and $\log \rho^*_{[\boldsymbol{e}]}(t)$ is a stationary time series with mean

$$K_{\boldsymbol{e}^*} = \boldsymbol{E}_{G^*} \log \rho^*_{[\boldsymbol{e}]} = \boldsymbol{y}'(I-\Phi)(\boldsymbol{M} + \boldsymbol{\delta}) - c_{\boldsymbol{e}}.$$

LLR $\log \rho^*_{[\boldsymbol{e}]}(t)$ satisfies the strong law of large numbers, and

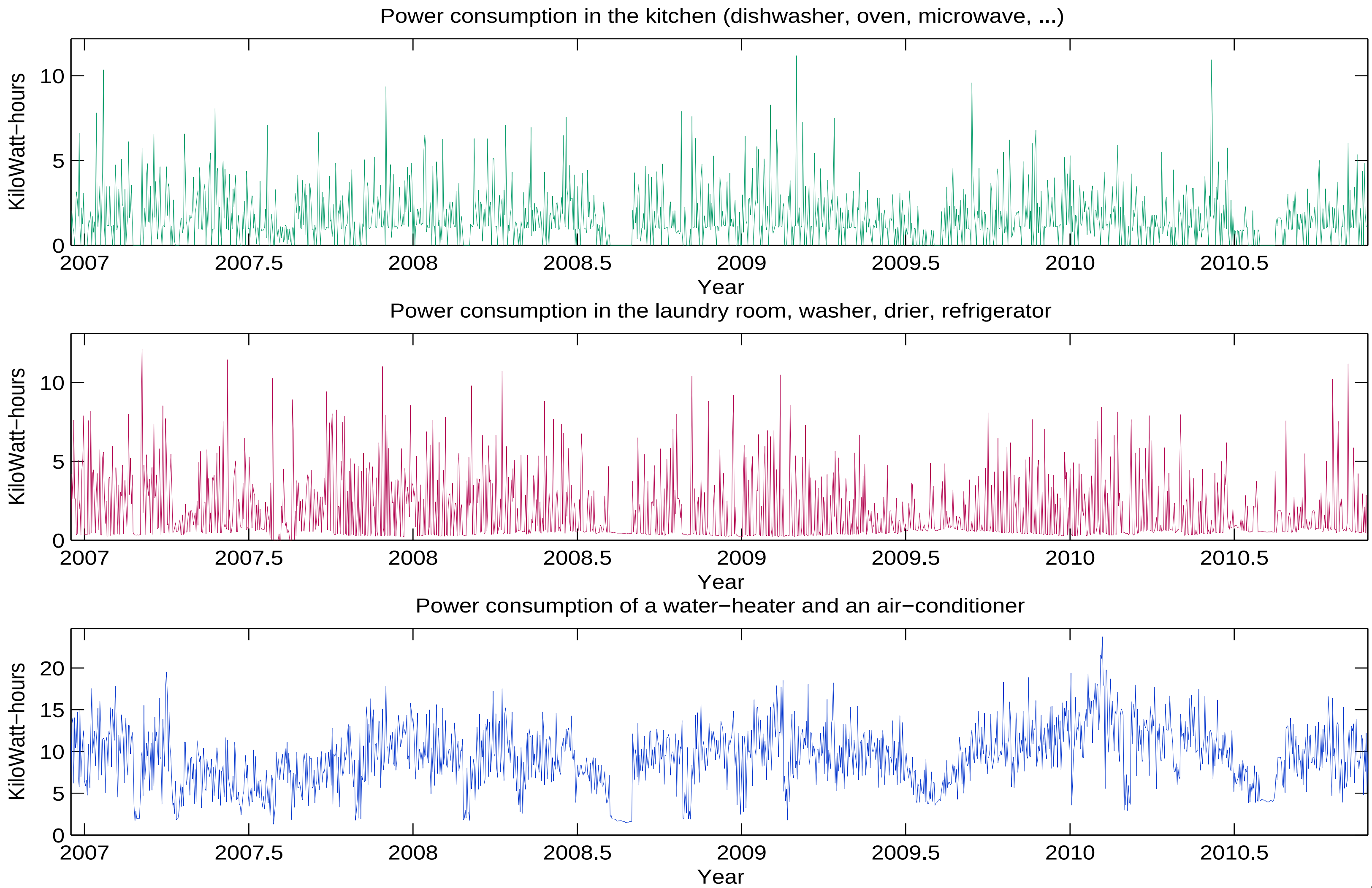
$$\frac{1}{n}\sum_{t=1}^{n} \log \rho^*_{[\boldsymbol{e}]}(t) \to K_{\boldsymbol{e}^*},$$

with probability one, as $n \to \infty$. Therefore,

$$T = \inf\left\{t : S_X(t) \leq e^{-\left(t\sum\lambda_j\right)^{-1}}\right\}$$

is an APO stopping time.

Example 4: Disaggregation of Power Consumption

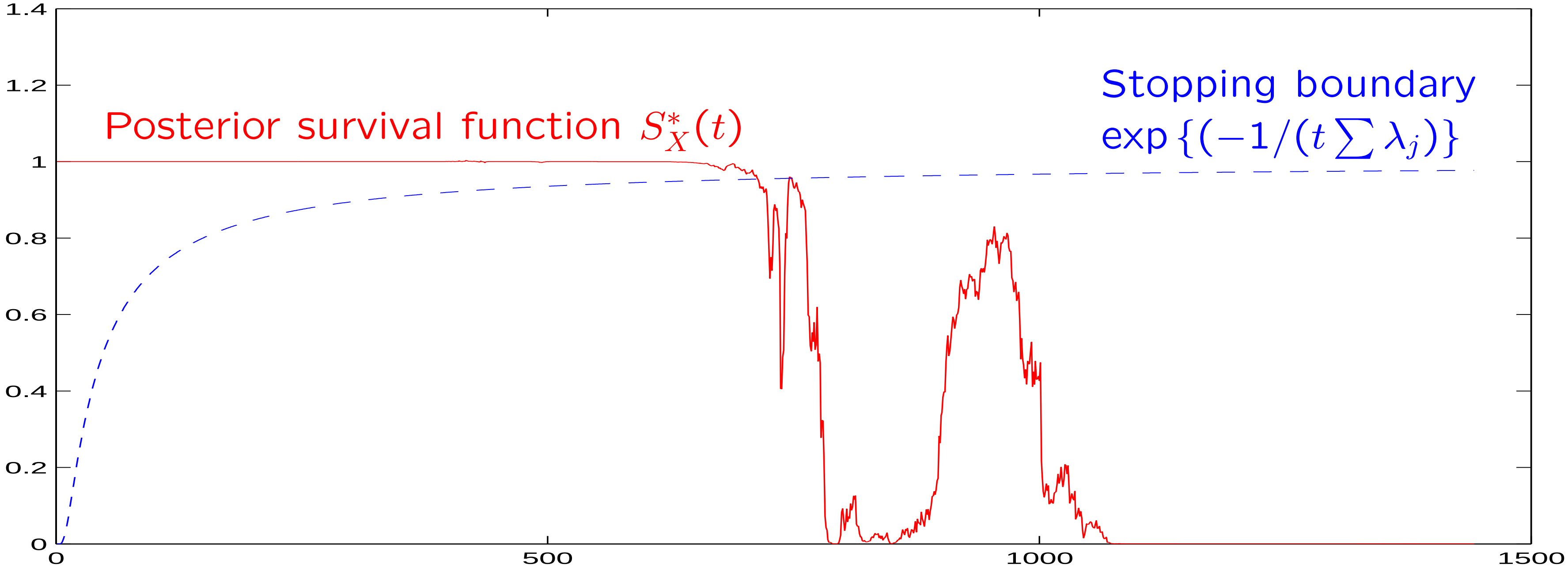


Energy Monitoring and Management Systems monitor and evaluate the long term energy impact on any changes in patterns due to renovations, aging appliances, or increasing use of green energy sources.

Reasons:

- Evaluate performance of electric appliances
- Identify energy saving opportunities

APO change detection



Results: detected at $T = 714$, the actual change-point is estimated to have occurred on day $\hat{\nu}_1 = 709$ (11/23/2008), and it was a 2.17 KWatt increase in the 3rd channel (0.77 standard deviations).

Further: another change-point on day $\hat{\nu}_2 = 1291$ (06/28/2010), a 0.86 KWatt reduction, also in the 3rd channel (0.31 standard deviations).

Key References

- M. Baron. Bayes stopping rules in a change-point model with a random hazard rate. *Seq. Anal.*, 20:147–163.

- M. Baron. Bayes and asymptotically pointwise optimal stopping rules for the detection of influenza epidemics. In *Case Studies in Bayesian Statistics*, 6:153–163, Springer-Verlag, NY, 2002.

- M. Baron. Sequential methods for multistate processes. In N. Mukhopadhyay, S. Datta, and S. Chattopadhyay, editors, *Applications of Sequential Methodologies*, pages 55–73. Marcel Dekker, Inc., New York, 2004.

- M. Baron. Asymptotically pointwise optimal change detection in multiple channels, 2014 (under review).

- M. Baron, M. Rosenberg, and N. Sidorenko. Electricity pricing: modeling and prediction with automatic spike detection. *Energy, Power, and Risk Management*, October 2001: 36–39.

- M. Basseville and I. V. Nikiforov. *Detection of Abrupt Changes: Theory and Application*. PTR Prentice-Hall, Inc., Englewood Cliffs, NJ, 1993.

- P. J. Bickel and J. A. Yahav. Asymptotically optimal Bayes and minimax procedures in sequential estimation. *Ann. Math. Stat.*, 39:442–456, 1968.

- J. Chen and A. K. Gupta. *Parametric Statistical Change Point Analysis: With Applications to Genetics, Medicine, and Finance*. Birkhäuser, Boston, MA, 2012

- H. Chernoff and S. Zacks. Estimating the current mean of a normal distribution which is subject to changes in time. *Ann. Math. Stat.*, 35: 999–1028, 1964.

- M. Ghosh, N. Mukhopadhyay, and P. K. Sen. *Sequential Estimation*. Wiley, New York, 1997.

- R. Gill, J. Keating, and M. Baron. Detecting abrupt leaks in blended underground storage tanks. *Comm. Statist.: Theory and Methods*, 35:4, 727–742, 2006.

- Z. Harchaoui and C. Lévy-Leduc. Multiple change-point estimation with a total variation penalty. *J. Amer. Statist. Assoc.*, 105(492):1480–1493, 2010.

- D. V. Hinkley. Inference about the change-point in a sequence of random variables. *Biometrika*, 57:1–17, 1970.

- E. S. Page. Continuous inspection schemes. *Biometrika*, 41:100–115, 1954

- H. V. Poor and O. Hadjiliadis. *Quickest Detection*. Cambridge University Press, Cambridge, UK, 2009

- Y. Ritov. Decision theoretic optimality of CUSUM procedure. *Ann. Stat.*, 18:1464–1469, 1990.

- A. N. Shiryaev. On optimal method in earliest detection problems. *Theory Probab. Appl.*, 8:26–51, 1963.
- A. N. Shiryaev. *Optimal Stopping Rules*. Springer-Verlag, New York, 1978.
- D. O. Siegmund, B. Yakir, and N. R. Zhang. Detecting simultaneous variant intervals in aligned sequences. *Ann. Appl. Statist.*, 5:645–668, 2011.
- A. G. Tartakovsky, I. V. Nikiforov, and M. Basseville. *Sequential Analysis: Hypothesis Testing and Change-Point Detection*. Chapman & Hall/CRC, 2014
- A. G. Tartakovsky, X. R. Li, and G. Yaralov. Sequential detection of targets in multichannel systems. *IEEE Trans. Information Theory*, 49(2):425–445, 2003.
- A. G. Tartakovsky and V. V. Veeravalli. Change-point detection in multi-channel and distributed systems with applications. In *Applications of Sequential Methodologies*, pages 339–370, New York, 2004. Marcel Dekker, Inc.
- Y. Xie and D. Siegmund. Sequential multi-sensor change-point detection. *Ann. Stat.*, 41:670–692, 2013.

Thank you!

• Y. Gel, R. Kulik, H. Yu	for	organizing and inviting
• NSF, NSA	for	sponsoring
• T. Wang, R. Gill, L. Khan, M. Rosenberg, A. Mustafa	for	collaborating
• Everyone	for	attending

mbaron@utdallas.edu
www.utdallas.edu/∼mbaron/