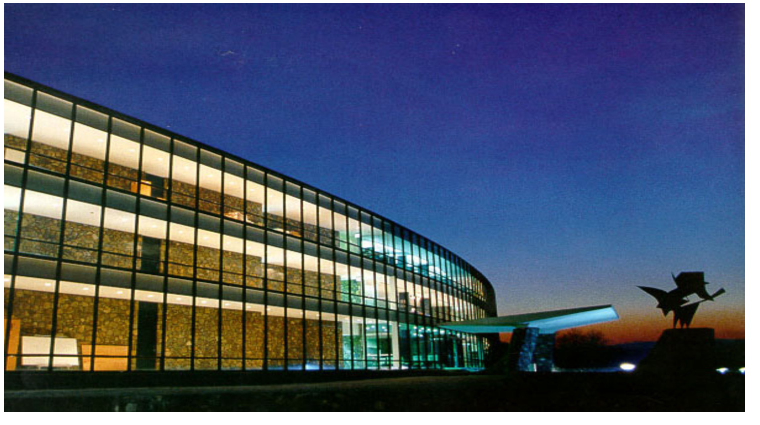
Computational and Monte-Carlo Aspects of Systems for Monitoring Reliability Data

Emmanuel Yashchin
IBM Research, Yorktown Heights, NY



COMPSTAT 2010, Paris, 2010

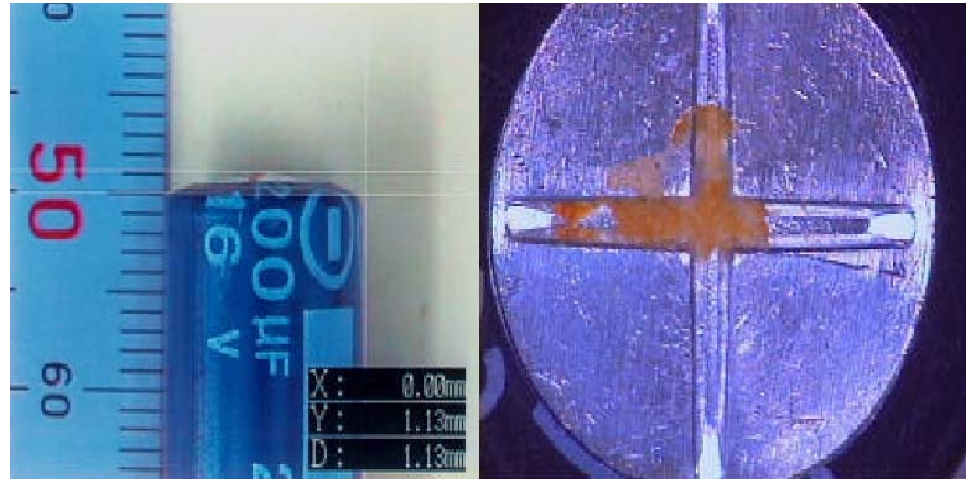
Outline

- Motivation
- Time-managed lifetime data
- Key issues in the design of a monitoring system
- Design of monitoring schemes Dynamically Changing Observations (DCO)
- Failure rate monitoring
- Wearout monitoring
- Computational and Monte Carlo Issues
- Conclusions

Reliability degradation of PC's caused by faulty capacitors

Bulging capacitors

Venting capacitor (top view)



Root cause: *Temperature – driven chemical reaction (unexpected failure mode)*

Potential for early detection: *High*

Business Day

TUESDAY, JUNE 29, 2010

The New York Times

Suit Over Faulty Computers Highlights Dell's Decline

By ASHLEE VANCE

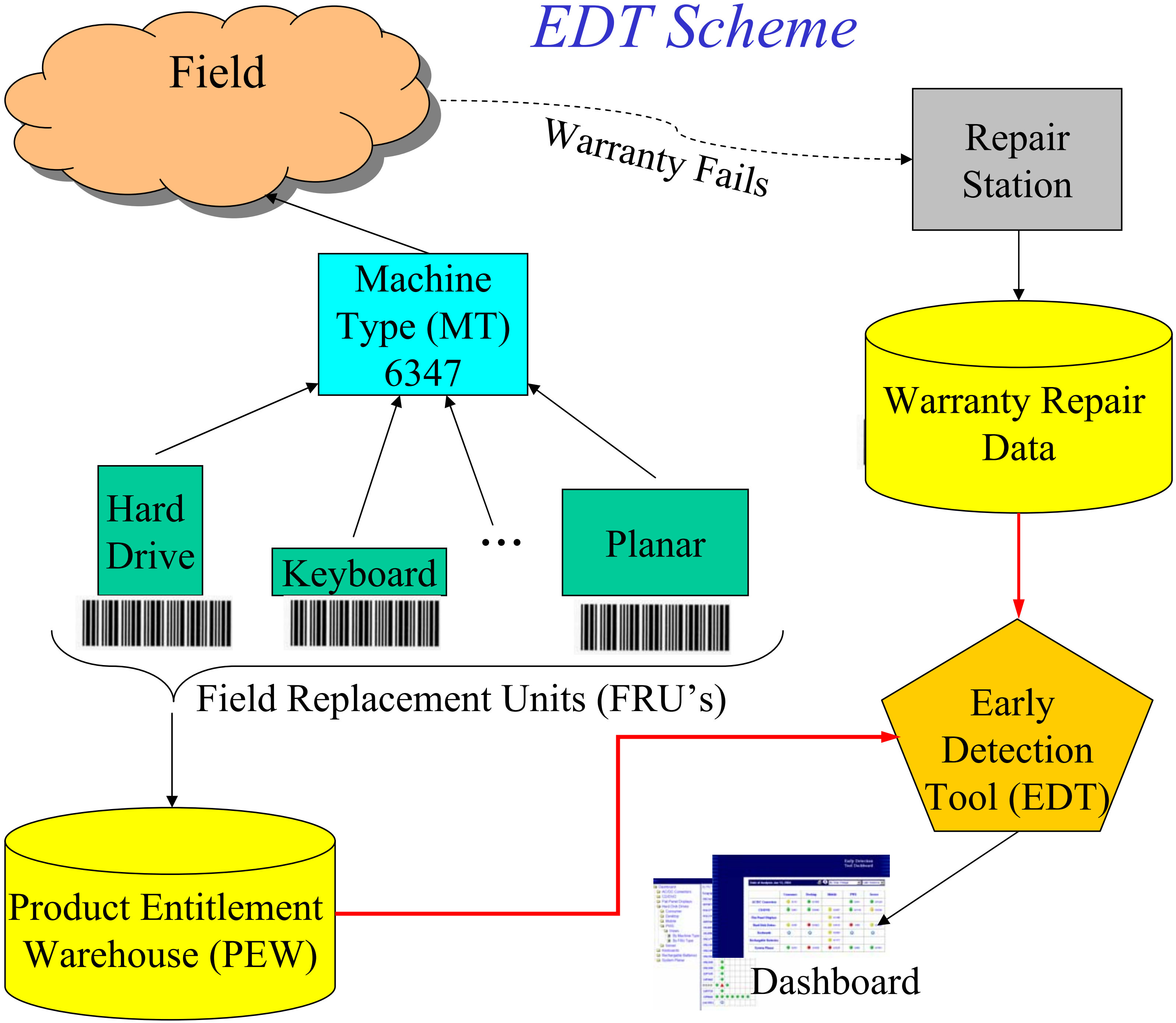
After the math department at the University of Texas noticed some of its Dell computers failing, Dell examined the machines. The company came up with an unusual reason for the computers' demise: the school had overtaxed the machines by making them perform difficult math calculations.

Dell, however, had actually sent the university, in Austin, desktop PCs riddled with faulty electrical components that were leaking chemicals and causing the malfunctions. Dell sold millions of these computers from 2003 to 2005 to major companies like Wal-Mart and Wells Fargo, institutions like the Mayo Clinic and small businesses.

Introduction

- Typical monitoring application: static observations
- Motivation: analysis of warranty data
- Early detection: key opportunity
- Time-managed data
- Early Detection Tool (EDT) for Warranty Data

EDT Scheme


Field
Warranty Fails
Repair Station
Machine Type (MT) 6347
Warranty Repair Data
Hard Drive
Keyboard
...
Planar
Field Replacement Units (FRU's)
Early Detection Tool (EDT)
Product Entitlement Warehouse (PEW)
Dashboard

Sorting schemes

Analyses to be run based on sorting with respect to potential root cause

- Sorting by vintage:
  - Product ship
  - Component ship
  - Calendar time

Sorting schemes

Component

VINTAGES	VOLUME	M1		M2		M3	
20010817	2	2	0	2	0	0	0
20010820	14	14	0	13	0	0	0
20010824	160	149	0	149	0	0	0
20010901	380	349	1	349	1	0	0
20010904	102	102	0	0	0	0	0
20010907	140	136	0	0	0	0	0
20010908	501	473	1	0	0	0	0
20010912	202	191	1	0	0	0	0
20010912	1	1	0	0	0	0	0
20010913	252	235	0	0	0	0	0
20010913	4	4	0	0	0	0	0
20010914	431	406	1	0	0	0	0
20010915	172	172	0	0	0	0	0

Machine

SHIP VINTAGES	VOLUME	M1		M2		M3	
20010925	221	206	1	186	1	0	0
20010926	287	261	1	161	0	0	0
20010927	350	316	0	117	0	0	0
20010928	385	381	1	49	0	0	0
20011001	506	492	0	0	0	0	0
20011002	612	578	1	0	0	0	0

Early Detection Tool (EDT) for Warranty Data

A system for detecting unfavorable changes in reliability of components.

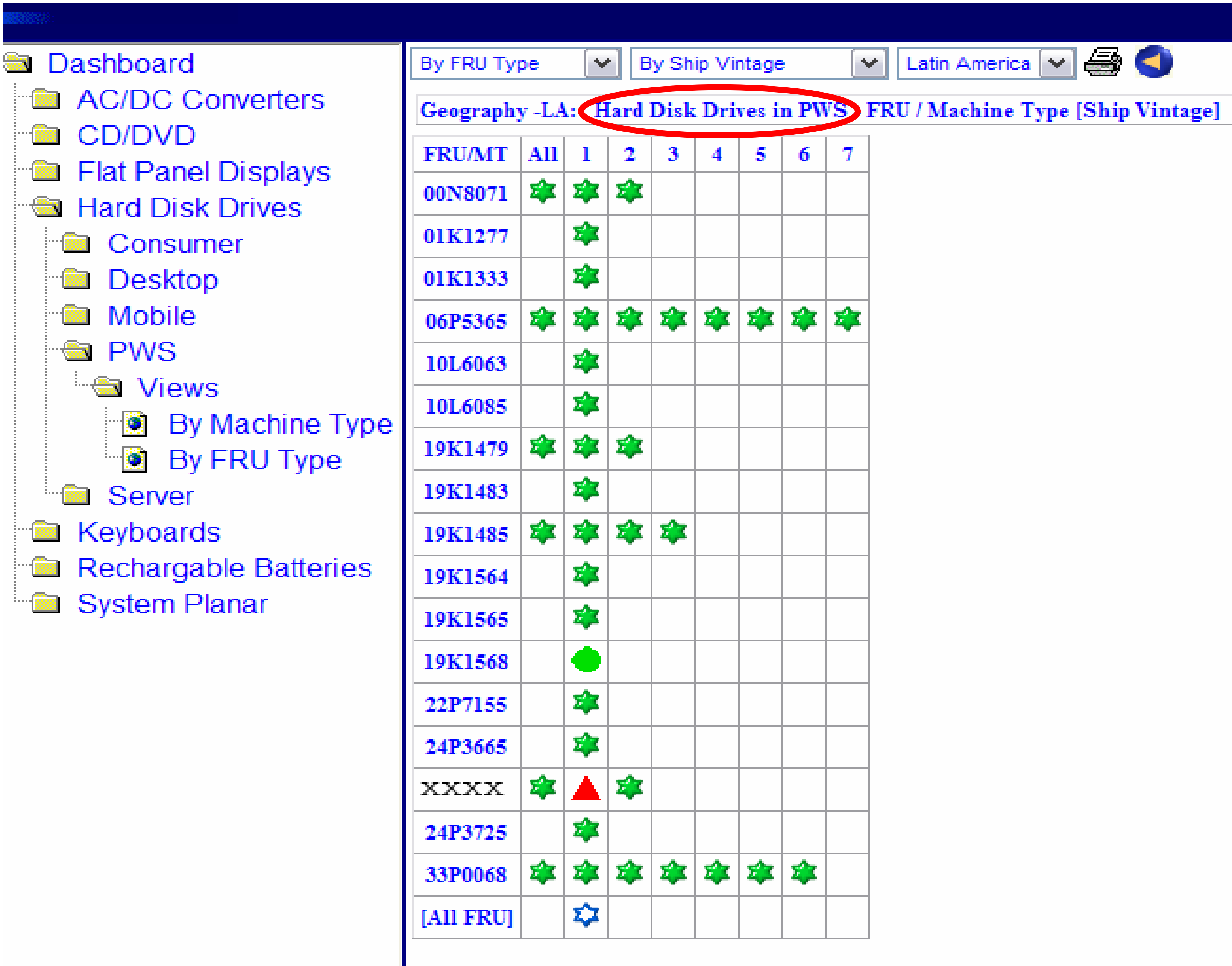
Multi-layer Dashboard:

Early Detection Tool Dashboard

Date of Analysis Jan 13, 2004 — By Ship Vintage — Latin America

	Consumer	Desktop	Mobile	PWS	Server
AC/DC Converters	0/13	0/165		0/41	0/124
CD/DVD	0/80	0/548	0/297	0/118	0/239
Flat Panel Displays			0/146		
Hard Disk Drives	0/38	8/382	0/410	1/55	0/165
Keyboards			0/250		
Rechargable Batteries			0/177		
System Planar	0/33	3/329	2/325	0/60	0/161

Nested (2-nd level) display:



Typical questions:

- Is the process of failures on target?
- If not, is the problem related to
  - *vendor's process? Assembly/Configuration process? Customer?*
  - *single Geo?*
  - *individual machine type? family of machine types?*
  - *individual Field Replacement Unit (FRU)? family of FRU's?*
  - *individual lot? sequence of lots?*
  - *stable process, but at unacceptably high replacement rate?*
  - *early fails?*
  - *increasing failure rate (wearout)?*
- What is the current state of the process?

Key Design Issues

1. ***Data***
   a. Multi-purpose, multi-stream
   b. Quality / Integrity
   c. Time managed, DCO
2. ***Alarms***
   a. False alarms vs. Sensitivity
   b. Believable and operationally (not statistically) significant
   c. Prioritization (severity, recentness, etc.)
   d. User control over the volume of alarms received
   e. Target setting
3. ***Modern statistical monitoring methodology***
   a. Reduce the Mean Time to Detection (MTTD) of unfavorable conditions
   b. Detect various types of changes (shifts, drifts, etc.)
   c. Detect intermittent problems
   d. Schemes designed using minimal level of user input

Key Design Issues (Cont)

4. ***Post-alarm activity***
   a. Facilitate diagnostics (incl. graphical analysis)
   b. Filtering
   c. Regime / Changepoint identification
   d. Actions
5. ***User interface***
   a. Multi-layer dashboards
   b. Reverse play
   c. Push / Pull / On-demand
   d. Communicate to users in a “human” language
6. ***Administration***
   a. Ease of use
   b. Training

New York Times

NEW YORK, SATURDAY, JULY 24, 2010

SIREN ON OIL RIG WAS KEPT SILENT, TECHNICIAN SAYS

WAKING CREW WAS ISSUE

Routine Practice Wasn't a Safety Oversight, Company Says

By ROBBIE BROWN

KENNER, La. — The emergency alarm on the Deepwater Horizon was not fully activated the day the oil rig caught fire and exploded, killing 11 people and setting off the massive spill in the Gulf of Mexico, a rig worker on Friday told a government panel investigating the accident.

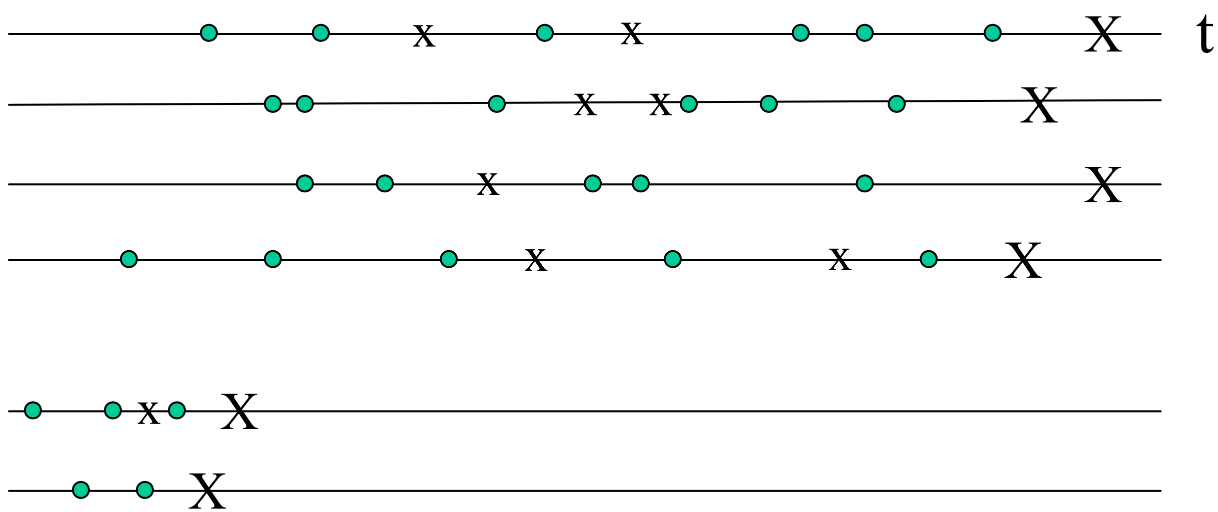
The worker, Mike Williams, the rig's chief electronics technician, said the general safety alarm was habitually set to "inhibited" to avoid waking up the crew with late-night sirens and emergency lights.

"They did not want people woke up at 3 a.m. from false alarms," Mr. Williams told the federal panel of investigators.

General data structure: sequence of life tests

E.g., current point in time: Aug 2, 2006. Current point affects data for all vintages, leading to dynamically changing statistics

Vintage	Sample	Lifetimes
2004-06-15	120	
2004-06-16	100	
2004-06-17	80	
2004-06-18	110	
……		
2006-07-20	95	
2006-07-21	110	

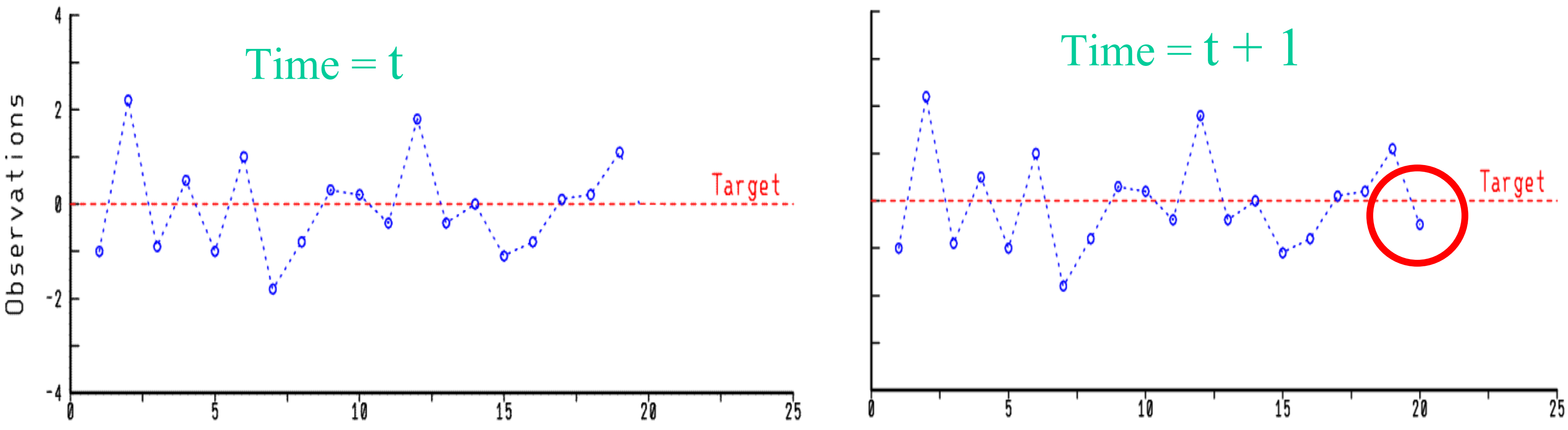


x – individually right-censored lifetimes

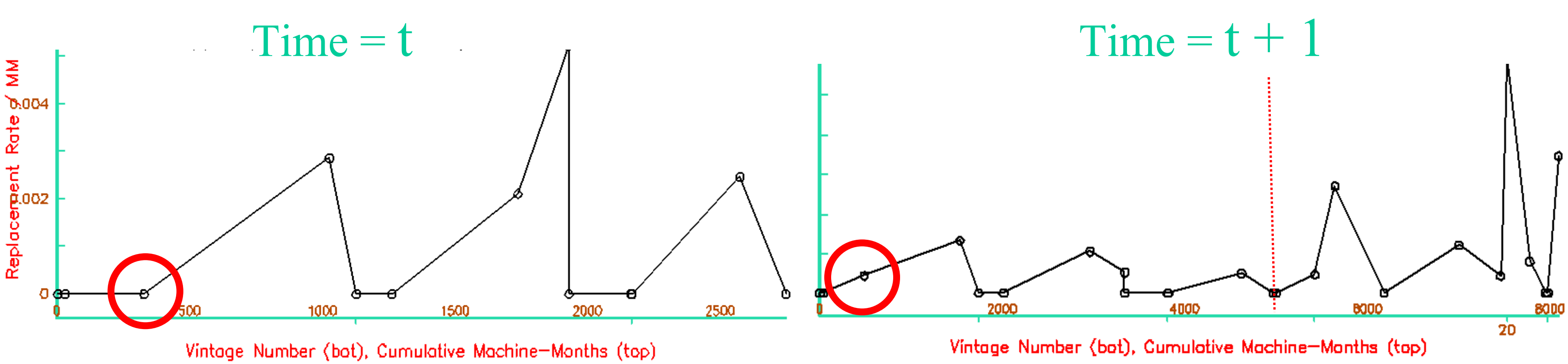
X – globally right-censored

Control charts with dynamically changing observations (DCO):

"Usual" control charts: Points observed earlier remain unchanged



DCO charts: Points observed earlier could change



Basic approach

- Sort data in accordance with vintages of interest
- Establish target curves for hazard rates.
- Transform time scale if necessary
- Characterize lifetime (possibly on transformed time scale) parametrically, e.g., Weibull
- For every parameter (say, λ), establish sequence of statistics $\{X_i , i = 1, 2, \ldots\}$ to serve as a basis of monitoring scheme; (e.g., assume $\lambda = E(X_i)$)
- Obtain weights $\{w_i , i = 1, 2, \ldots\}$ associated with $\{X_i\}$
- Establish acceptable & unacceptable regions $\lambda_0 < \lambda_1$
- Establish acceptable rate of false alarms
- Apply scheme to every relevant data set; flag this data set if out-of-control conditions are present

Main test: Repeated Page's scheme

Suppose that at time T we have data for N vintages

Define the set $\{S_i, i = 1, 2, \ldots, N\}$ as follows:

$$S_0 = 0, \quad S_i = \max[0, \gamma S_{i-1} + w_i(X_i - k)],$$

where $k \approx (\lambda_1 + \lambda_0)/2, \quad \gamma \in [0.7, 1]$

Define $S = max\,[S_1, S_2, \ldots, S_N]$;

Flag the data set at time T if $S > h$, where h is chosen via:

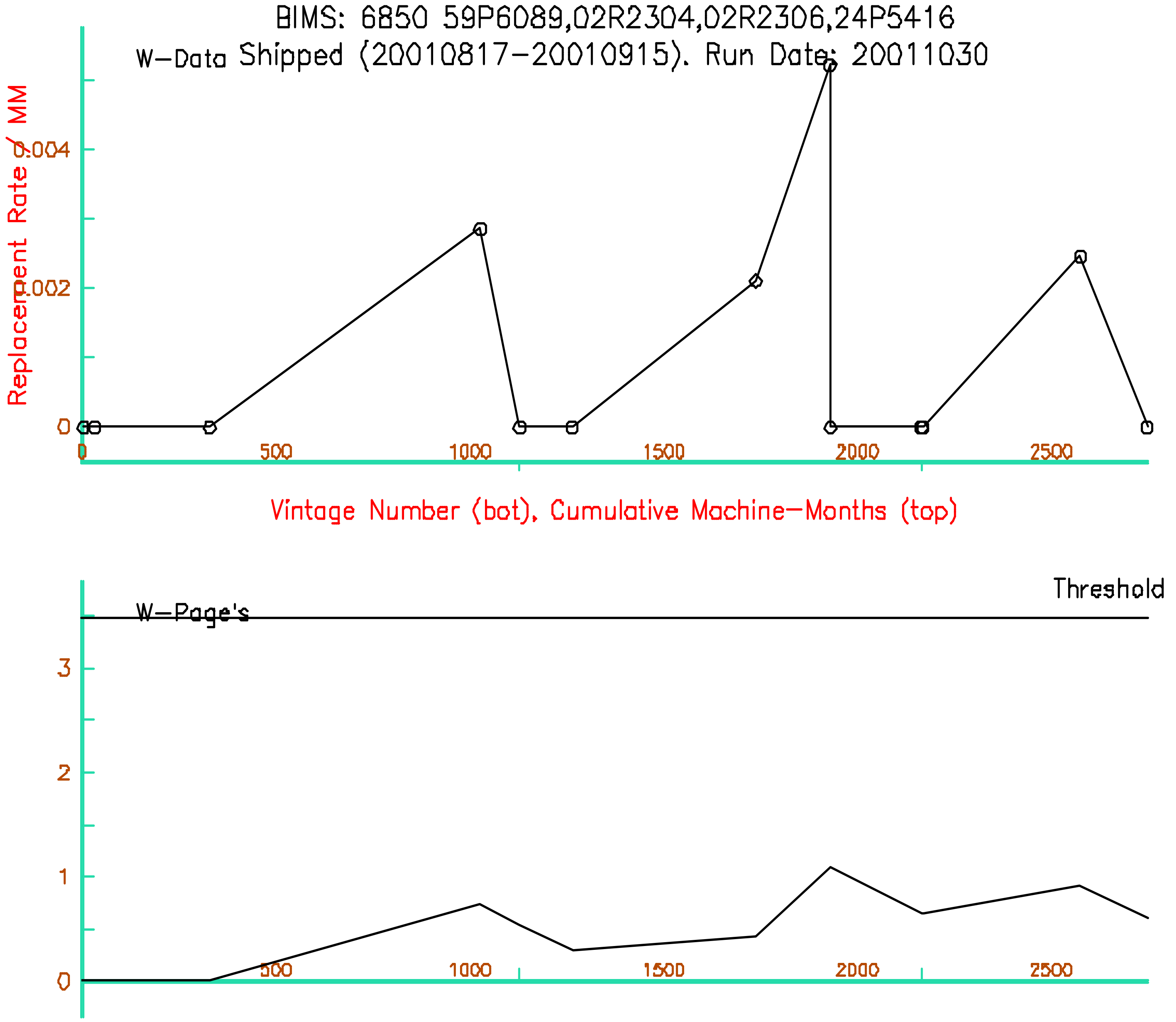
$$\text{Prob}\{ S > \text{h} \mid N, \lambda = \lambda_0\} = 1 - \alpha_0 \ \ (\text{e.g.} = 0.99)$$

Note: Average Run Length (ARL) is not used here!

Example1: Failure rate monitoring of a PC component
Monitoring Replacement Rate $\lambda = E(X_i)$

Data view of Oct 30 2001

OBS	DATES	WMONTHS	WFAILS	RATES
1	20010817	4	0	0
2	20010820	27	0	0
3	20010824	298	0	0
4	20010901	698	2	0.0029
5	20010904	102	0	0
6	20010907	136	0	0
7	20010908	473	1	0.0021
8	20010912	191	1	0.0052
9	20010912	1	0	0
10	20010913	235	0	0
11	20010913	4	0	0
12	20010914	406	1	0.0024
13	20010915	172	0	0

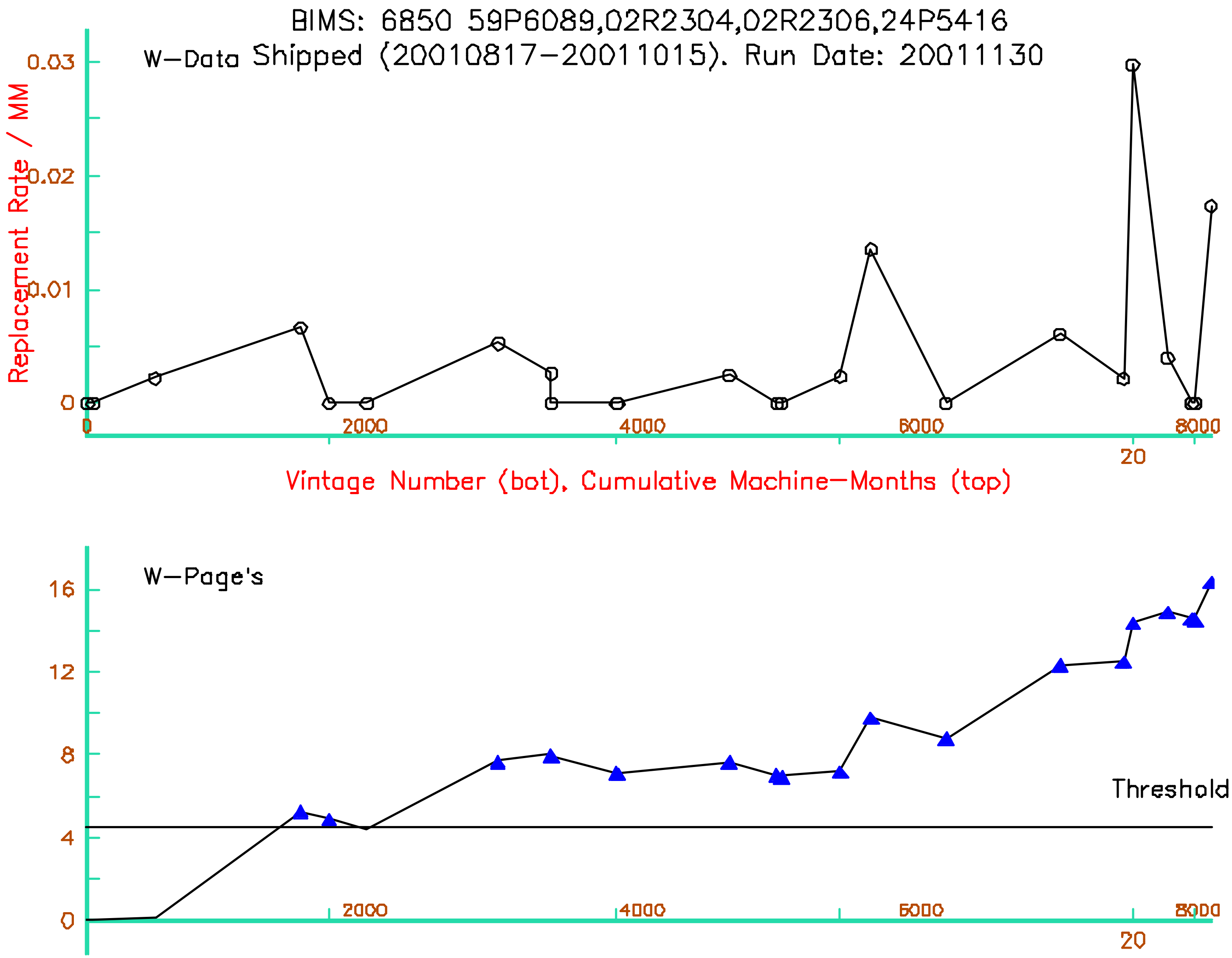


BIMS: 6850 59P6089,02R2304,02R2306,24P5416
W-Data Shipped (20010817-20010915). Run Date: 20011030
Replacement Rate / MM
0.004
0.002
0
0
500
1000
1500
2000
2500
Vintage Number (bot), Cumulative Machine-Months (top)
Threshold
W-Page's
3
2
1
0
500
1000
1500
2000
2500

Data view of Nov 30 2001

OBS	DATES	WMONTHS	WFAILS	RATES
1	20010817	6	0	0
2	20010820	40	0	0
3	20010824	447	1	0.0022
4	20010901	1047	7	0.0067
5	20010904	204	0	0
6	20010907	272	0	0
7	20010908	945	5	0.0053
8	20010912	381	1	0.0026
9	20010912	2	0	0
10	20010913	469	0	0
11	20010913	8	0	0
12	20010914	805	2	0.0025
13	20010915	341	0	0
14	20010919	36	0	0
15	20010928	420	1	0.0024
16	20010929	221	3	0.0136
17	20010930	540	0	0
18	20010930	821	5	0.0061
19	20011001	456	1	0.0022
20	20011007	67	2	0.0299
21	20011008	251	1	0.0040
22	20011009	173	0	0
23	20011013	1	0	0
24	20011013	22	0	0
25	20011015	1	0	0
26	20011015	115	2	0.0174

Now we have enough evidence to flag the condition:



Wearout Monitoring

Define Wearout Parameter: E.g. use shape parameter c of Weibull lifetime distribution

Establish acceptable/unacceptable levels: $c_0 < c_1$

Establish Data Summarization Policy: E.g. consolidate data monthly

Define the set $\{S_{iw}, i = 1, 2, \ldots, M\}$ as follows:

$$S_{0w} = 0, \quad S_{iw} = \max[0, \gamma_w S_{i-1,w} + w_{iw}(\hat{C}_i - k_w)],$$

where $k_w \approx (c_0 + c_1)/2, \quad w_{iw} =$ number of failures in vintage i

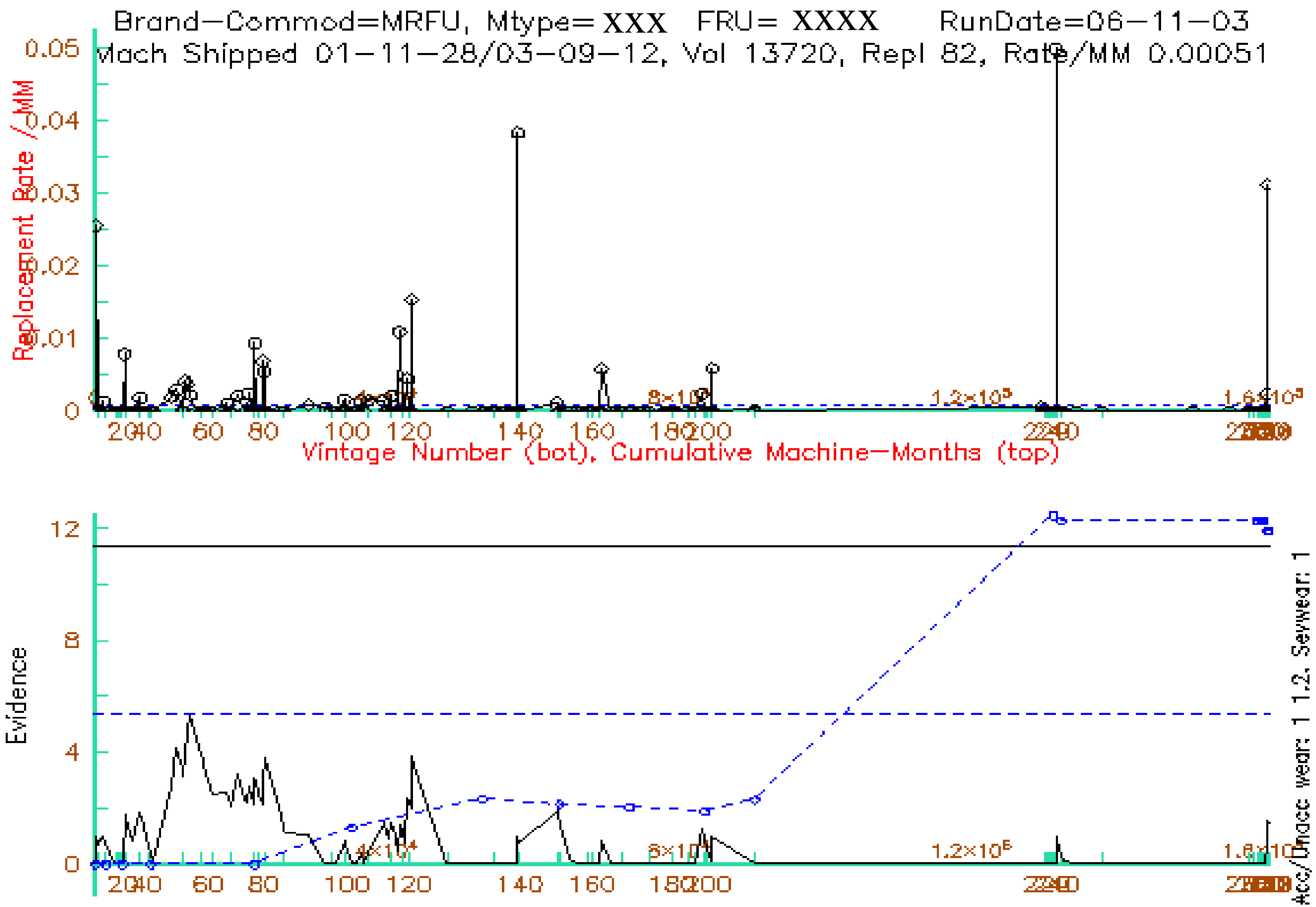
$\hat{C}_i =$ Bias - corrected estimate of c based on month i

Define $S_w = max\,[S_{1w}, S_{2w}, \ldots, S_{Mw}]$;

Flag the data set at time T if $S_w > h_w$, where h_w is chosen from:

$$\mathrm{Prob}\{S_w > h_w \mid M, c = c_0\} = 1 - \alpha_0 \ (\text{e.g.} = 0.99)$$

Example2: Joint Monitoring of Replacement Rate & Wearout



Accep/Unaccept levels: 0.0007 0.00123. Prob(no false alarm): 0.99. Severity: 0.6698

Some issues

Issue#1: for a wide enough window of vintages, the signal level *h* may get too high to provide desired level of sensitivity with respect to recent events

To address: - *enforce sufficient separation between acceptable & unacceptable levels, e.g. for* $\lambda = E(X_i)$ *require* $\lambda_1 / \lambda_0 > 1.5$

- *introduce supplemental tests. For example, define*

"active component" = component for which shipment record(s) are present within the last L days (L = active range). For such components use supplemental tests:

Test1 (based on last value of scheme): Flag the data set if $S_N > h_1$,

Test2 (based on failures within the active range): Flag if $X_{(L)} > h_2$, where $X_{(L)}$ = number of failures within active range

Issue#2: unfavorable changes in some parameters can show up "on the wrong chart"

To address: - *use special diagnostic procedures*

- *select different quantities to monitor (may affect interpretability)*

- *monitor model adequacy*

Computational & Monte Carlo Issues

1. Establish "on the fly" the thresholds for the tests, e.g., solve for h

$$\text{Prob}\{\, S > \text{h} \mid N, \lambda = \lambda_0\} = 1 - \alpha_0$$

where $S = max\,[S_1, S_2, \ldots, S_N]$;

(a) use parallel (vector) computations taking into account recursive nature of the process $S_1, S_2, \ldots, S_N$

(b) since the sequence of observed weights $\{w_i\}$ *is ancillary for* λ*, condition on them*

(c) use simulated replications of $S_1, S_2, \ldots, S_N$*, observe the set of maxima*

(d) use asymptotic result (requires existence of first to moments of X_i*)*

$$\text{Prob}\{\, S > \text{h} \mid N, \lambda = \lambda_0\} \sim \text{A} \bullet \exp[\, - \text{a} \bullet \text{h}], \quad \text{h} \to \infty$$

Scale: 100,000 data sets examined per week

Computational & Monte Carlo Issues (cont)

2. For active components, establish "on the fly" the thresholds for the main and supplemental tests, i.e., find suitable h, h_1, h_2

$$\text{Prob}\{\, S > \text{h or } S_N > \text{h}_1 \text{ or } X_{(L)} > \text{h}_2 \mid N, \lambda = \lambda_0 \} = 1 - \alpha_0$$

where S_N , $X_{(L)}$ = Supplemental statistics 1, 2

(a) involves policy for type-1 error allocation among tests

(b) use parallel simulation (conditioned on weights $\{w_i\}$)

(c) use asymptotic results for S and for S_{N}:

$$\text{Prob}\{\, S_N > \text{h} \mid N, \lambda = \lambda_0 \} \sim \text{A}_1 \bullet \exp[\, - \text{a}_1 \bullet \text{h}], \quad \text{h} \rightarrow \infty$$

Computational & Monte Carlo Issues (cont)

3. Establish "index of severity", so that flagged data sets could be ranked based on their "newsworthiness"

    Severity = combination of p-values (p_1, p_2, p_3) of the main and supplemental tests. E.g. Severity = $1 - \min\{p_1, p_2, p_3\}$

    Estimated via re-sampling techniques

4. Thresholds and severities for wearout index and for Weibull scale parameter
5. Predictions (e.g. of overall fallout) and related bounds
6. Estimation of filtered parameter values and confidence bounds
7. Regimes and change-points

Discussion

- Monitoring reliability characteristics in the presence of dynamically changing observations requires non-standard performance criteria and control schemes (e.g., repeated Weighted Cusum-Shewhart). Design and implementation of these schemes involves extensive use of MC methods.
- Practical applications are usually associated with a battery of tests (even for a single parameter), as several aspects of detection process need to be taken into account. Of special importance: failure rate and wearout characteristics
- Generalized approach: in terms of likelihood ratios
- System based on this approach deployed and proven useful in practice