# DEGRADATION MODELS AND FAILURE PREDICTION

**Jerry Lawless**
**University of Waterloo**

**For ICRSA, May 21-24, 2003**

- Introduction
- Some Models for Degradation and Failure
- Failure Time Prediction
- Failure Prevention and Maintenance
- Comments

# INTRODUCTION: DEGRADATON AND FAILURE

- Events("failures") that have undesirable consequences and require corrective action occur in systems and processes, e.g.

  - catastrophic failure of a metal part due to fatigue-crack growth →
  - manufacturing process stoppages → lost productivity, cost of repair

- Occurrence of failures is related to the "condition" of a unit or system and to external factors (e.g. environment, usage)
- Degradation measures are variables that reflect deterioration in a unit or system
- By relating degradation to failure we hope to reduce or prevent failures, e.g., by monitoring degradation (planned maintenance) or reducing degradation (system improvement)

Suppose a degradation measure $Y(t)$ is associated with units of equipment at time $t \geq 0$, and that a counting process $\{N(t), t \geq 0\}$ records the number of failures over $(0, t)$.

- Defined failures: a failure is defined to occur at time $t$ when $Y(t)$ enters some region $D$, i.e. when $Y(t-) \epsilon \bar{D}, Y(t) \epsilon D$.
- Degradation-related failures: the failure intensity is

$$\lambda(t; H(t)) = \lim_{\Delta t \to 0} \frac{P(N(t) - N(t-) = 1 | H(t))}{\Delta t}$$

where $H(t)$ includes the degradation history $\bar{y}(t) = \{y(s) : 0 \leq s < t\}$ and previous failure history. External covariates $x(t)$ (e.g.stress, usage) may also be included.

- In some settings a failure time $T \geq 0$ is defined, measured from a specified time origin.

Some examples

- Fatigue-crack growth in metal (Lu and Meeker 1993)

$$
\begin{aligned}
Y(t) &= \text{crack length after } t \text{ load-cycles } (t \geq 0) \\
T &= \inf\{t : Y(t) \geq d\}
\end{aligned}
$$

- Ion source filament failures in semiconductor manufacturing (S. Crowder 1993)

$$
\begin{aligned}
Y(t) &= \text{current through filament (decreases as filament degrades)} \\
T &= \text{time to failure (breakage) of filament (hrs. of operation)}
\end{aligned}
$$

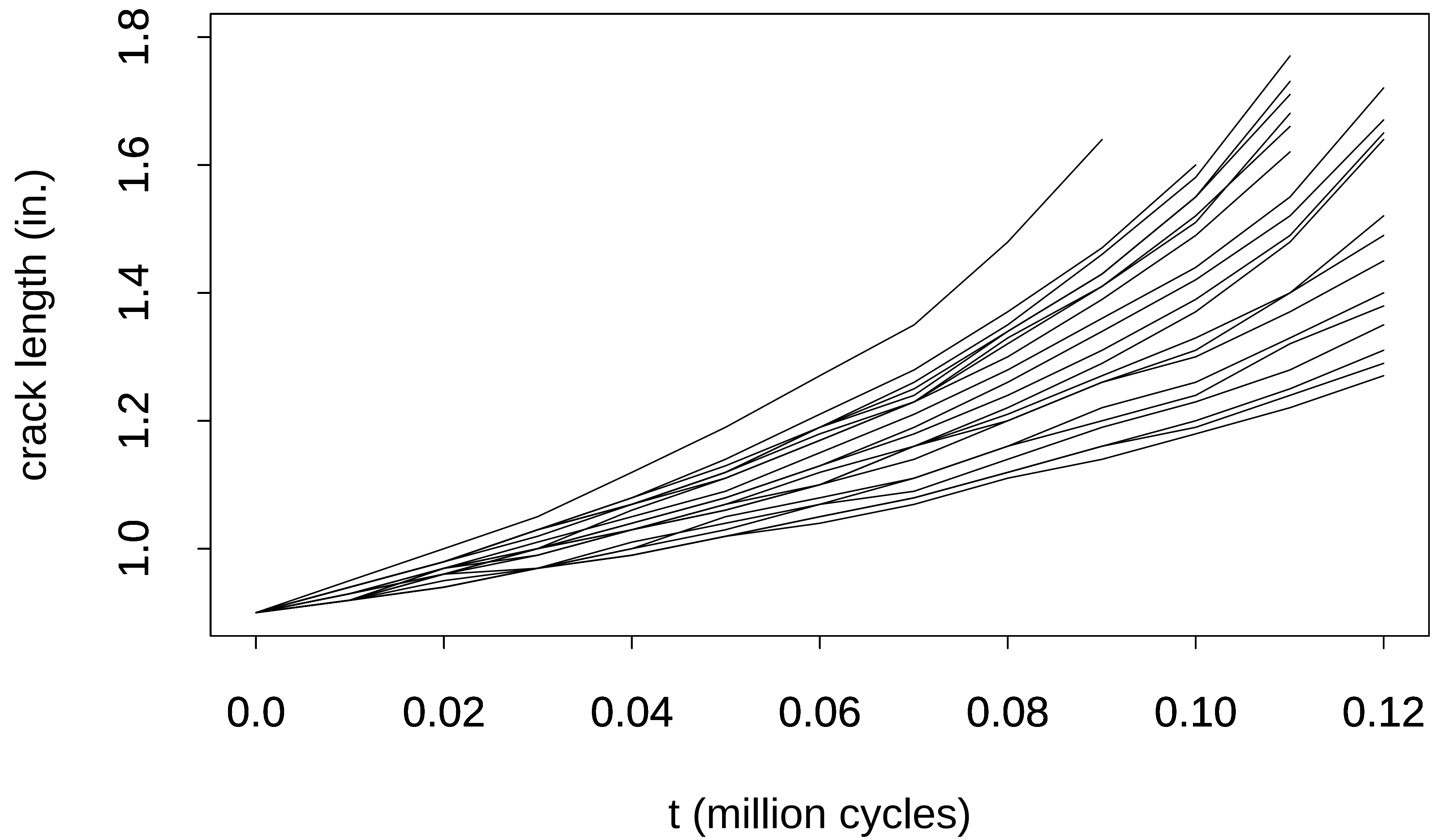
- Pump bearing failures (Banjevic et al. 2001)

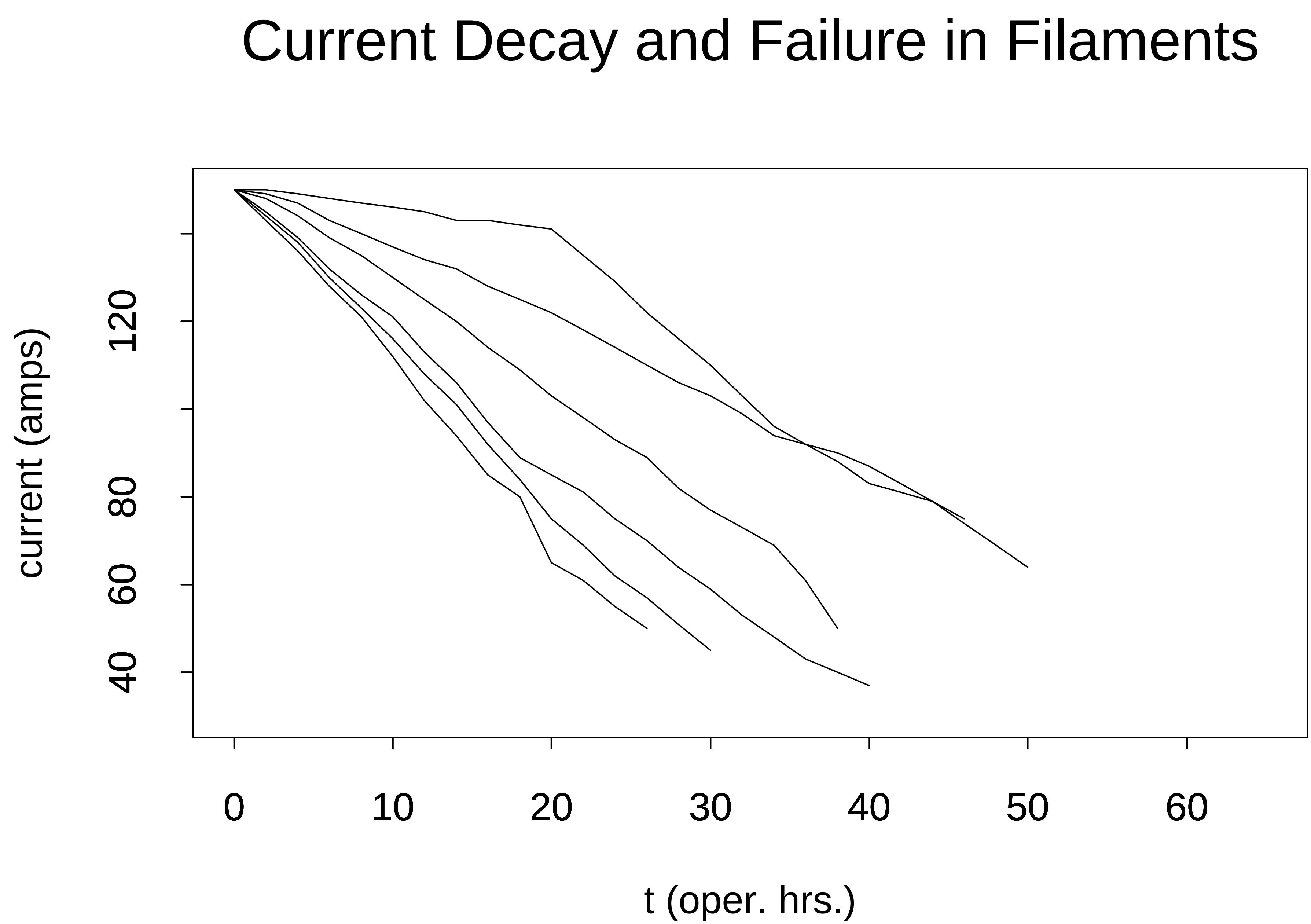
$$
\begin{aligned}
Y(t) &= \text{vector derived from frequency domain analysis of axial,} \\
&\quad\ \text{horizontal, and vertical vibration measurements.}
\end{aligned}
$$

- Persons with HIV (human immunodeficiency virus) disease

$$
\begin{aligned}
Y(t) &= \text{CD4 and viral load counts} \\
T &= \text{time to first major opportunistic infection}
\end{aligned}
$$

# Crack Growth for 21 Specimens




Current Decay and Failure in Filaments
current (amps)
40
60
80
120
0
10
20
30
40
50
60
t (oper. hrs.)

## Failure prediction

- Want to determine $P(\text{no failure in}(t, t+s)|H(t))$

- look at effect of including degradation history $\bar{y}(t)$ in $H(t)$

- For a single failure time $T$, compare

$$P(T > t+s|T > t, \bar{y}(t)) \text{ and } P(T > t+s|T > t).$$

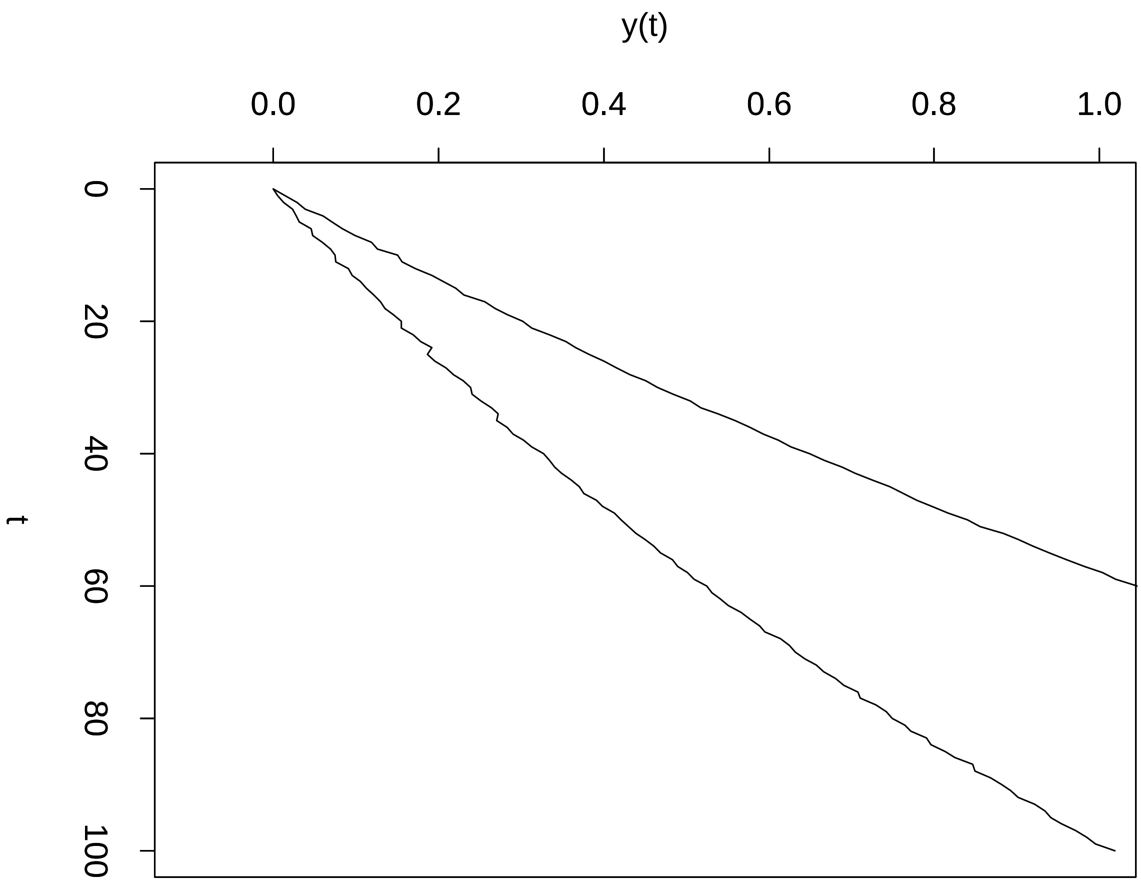
- A joint model for failure and $\{Y(t), t \geq 0\}$ is required.

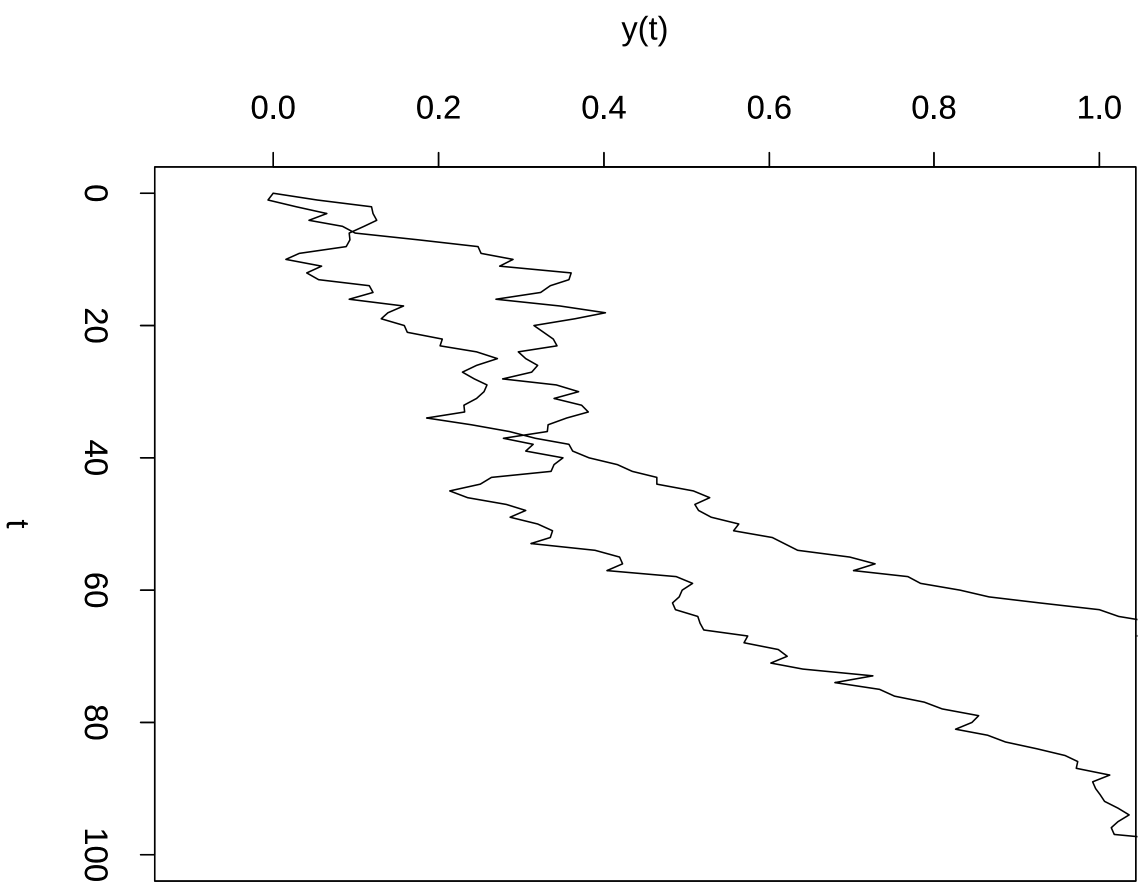
$$P(T > t+s|T > t, \bar{y}(t)) = E\{exp(-\int_t^{t+s} \lambda(u; \bar{y}(u))du)\}$$

- Features that affect the usefulness of degradation (or condition) measures include

- degree of dependence of failure on degradation

- the variability of degradation curves


y(t)
0.0
0.2
0.4
0.6
0.8
1.0
t
0
20
40
60
80
100


y(t)
0.0
0.2
0.4
0.6
0.8
1.0
t
0
20
40
60
80
100

# SOME MODELS FOR DEGRADATION AND FAILURE

**Defined Failures**

For simplicity consider scalar $y(t) \geq 0$ for $t \geq 0$ and failure threshold $y^F$, with

$$T = inf(t : Y(t) \geq y^F)$$

- Processes $\{Y(t), t \geq 0\}$ with independent increments

- Wiener processes, e.g. $\Delta Y(t) = Y(t + \Delta t) - Y(t) \sim N(\mu\Delta t, \sigma^2\Delta t)$
$T \sim$ inverse Gaussian

- Gamma processes, $\Delta Y(t) \sim Gamma(\alpha, \Delta\beta(t))$
Monotonic, so $P(T > t) = P(Y(t) < y^F)$

- Other Markov processes

- Monotonic processes are fairly straightforward, but not always suitable
Singpurwalla (1995 Stat. Sci), Aalen and Gjessing (2001 Stat. Sci)

- Need for random effects to accommodate unit-to-unit heterogeneity e.g. Wiener processes with random drift parameter $\mu$ or threshold $y^F$ (Aalen and Gjessing 2001)
  Gamma processes with random scale $\alpha$ or shape function $\beta(t)$ (Crowder and Lawless 2003)
- Random growth curves (e.g. Lu and Meeker 1993)

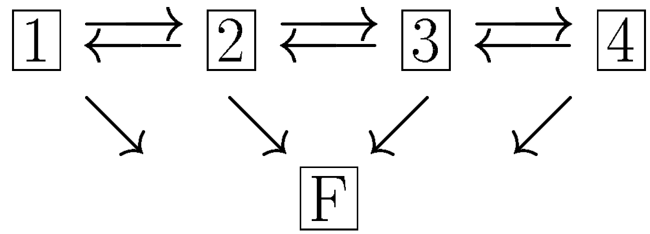
$$Y_i(t) = g(t; \theta_i) \qquad\qquad \theta_i \sim F(\cdot)$$

$$T_i \text{ is a function of } \theta_i$$

  Measurement error is sometimes added, so that $Y_i^*(t) = Y_i(t) + \epsilon_i(t)$ is observed instead of $Y_i(t)$.

# Degradation-related Failures

- Discrete-state models




- Multiplicative hazards $\lambda(t;\bar{y}(t)) = \lambda_0(t)g(y(t))$

Additive hazards $\lambda(t;\bar{y}(t)) = \lambda_0(t) + g(y(t))$

Tractability depends on the $Y(t)$ process, but failure time calculations are usually difficult. A few hazard-based models where $\{Y(t)\}$ or a function of it is Markov are moderately tractable.

e.g. $\lambda(t;\bar{y}(t)) = \lambda_0(t) + \beta g(y(t))$

$$P(T > t + \Delta t | T > t, \bar{y}(t)) = E\{e^{-\Delta\Lambda_0(t) - \beta \Delta V(t)} | T > t, \bar{y}(t)\}$$

where

$$\Delta\Lambda_0(t) = \int_t^{t+\Delta t} \lambda_0(u) du$$

$$\Delta V(t) = \int_t^{t+\Delta t} g(y(u)) du,$$

which gives, if $[\Delta V(t) | T > t, \bar{y}(t)] = [\Delta V(t) | V(t)]$,

$$e^{-\Delta\Lambda_o(t)} E\{e^{-\beta \Delta V(t)} | V(t)\}.$$

- This is a m.g.f. and is tractable for certain processes $\{V(t), t \geq 0\}$; numerical methods can also be employed
- Random effects on $V(t)$ can also be introduced.

## Shared Random Effects Models

- Unobservable random effect $Z_i$ for unit or system $i$:

$$T_i \perp \bar{Y}_i(\infty) | z_i \qquad \qquad Z_i \begin{array}{l} \nearrow \{Y_i(t)\} \\ \searrow T_i \end{array}$$

e.g.(various authors)

$$Y_i(t) \; = \; Z_{oi} + Z_{1i} t + \epsilon_i(t) \qquad \qquad Z_i = (Z_{oi}, Z_{1i}) \sim BVN$$

$$log \; T_i \; \sim \; N(\mu(Z_{oi}, Z_{1i}), \sigma^2)$$

- $P(T_i > t + \Delta t | T_i > t, \bar{y}_i(t))$

$= \int P(T_i > t + \Delta t | T_i > t, z_i) P(z_i | T_i > t, \bar{y}_i(t)) dz_i$

# Data and Model Selection

- Data often of form $\{y(t_j), j = 1, \ldots, n; T\}$ or $\{y(t_j), j = 1, \ldots, n; T > t_n\}$
- Experimental (off-line) data: can often observe a unit until failure, perhaps under accelerated stress conditions
- Observational (on-line) data: often observe very few failures.

\- May have substantial data on degradation of units but little indication as to when the failure intensity is high

# ILLUSTRATION: CRACK GROWTH FAILURES

- Experimental setting considered by Hudak et al (Lu and Meeker 1993)

$$\begin{aligned} Y_i(t) &= \text{crack length after } t \text{ test cycles for i'th unit } (i = 1, \ldots, 21) \\ T_i &= inf\{t : Y_i(t) > 1.6 \text{ inches}\} \\ -y_i(0) &= 0.9 \text{ for all units.} \end{aligned}$$

$Y_i(t)$ observed for $t = .01, .02, \ldots, .12$ (million cycles)

- Consider gamma proceses with random effects $z_i = a_i^{-1}$

$$\Delta Y_i(t) = Y_i(t + \Delta t) - Y_i(t) \sim \text{Gamma}(a_i, \Delta b(t))$$

where $E(\Delta Y_i(t)|a_i, \bar{y}_i(t)) = \Delta b(t) a_i^{-1}$ and

$$\Delta b(t) = b(t + \Delta t) - b(t)$$

- Used $Z_i \sim$ Gamma $(c, d)$ with $E(Z_i) = \mu_z = d/c$ and

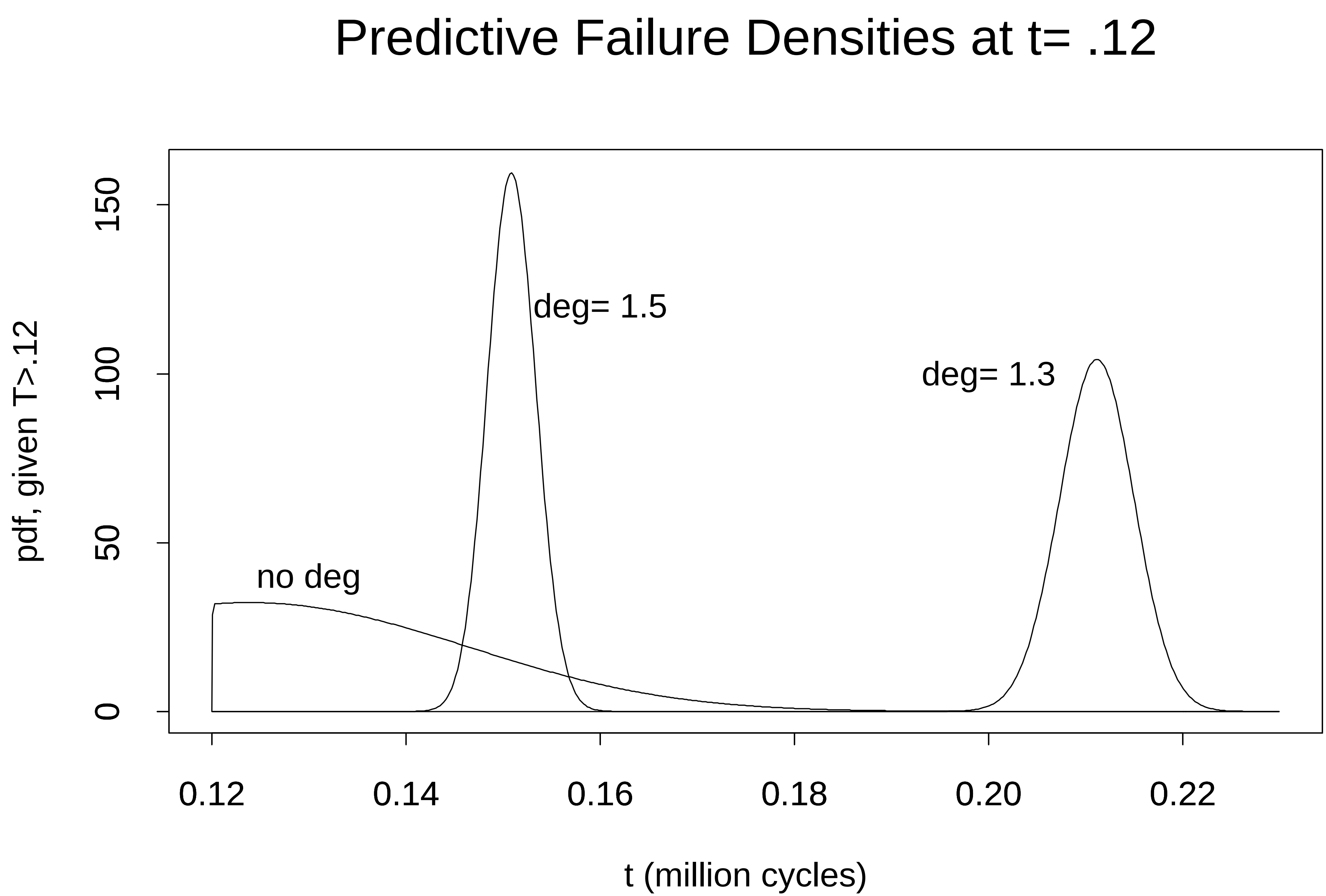
$$b(t) = e^{\beta_o}(e^{\beta_1 t} - 1)$$

- straightforward to fit
- Failure time prediction:

$$P(T_i > t + \Delta t | T_i > t, \bar{y}_i(t)) = F\left[\frac{\mu_z(1 + b(t)\gamma_z^2)(y^F - y(t))}{\Delta b(t)(y(t)\gamma_z^2 + 1)}\right]$$

where $\gamma_z = \sigma_z/\mu_z$ and $F \sim F_{2\Delta b(t), 2b(t)+2/\gamma_z^2}$

- Plot shows probability density functions for $T_i$, given $T_i > 0.12$, for two values of $y_i(.12)$ : 1.3 and 1.5, along with the p.d.f for $T_i$, given only that $T_i > 0.12$ ( no degradation measure)


Predictive Failure Densities at t= .12
pdf, given T>.12
0
50
100
150
no deg
deg= 1.5
deg= 1.3
0.12
0.14
0.16
0.18
0.20
0.22
t (million cycles)

# FAILURE PREVENTION AND MAINTENANCE

- Preventive maintenance and planned replacements (PM)
  - seek to minimize disruptions and loss of service from equipment or a system while maximizing profit (minimizing cost)
- Corrective maintenance (CM) following equipment or system failure
  - typically more costly to deal with than PM; loss of control
  - want to avoid failures, in general
- Maintenance planning and logistics is a complex issue, especially with large systems involving many parts. Focus here on individual parts or subsystems.

## Optimization of Maintenance Plans

- Huge mathematical literature but most of it is unused; few examples of models fitted to data.
- Age-based replacement policies

$$\begin{aligned} T &= \text{time to failure of part from last maintenance} \\ C_1 &= \text{cost of PM replacement} \\ C_2 &= \text{cost of CM replacement} \qquad C_2 > C_1 \end{aligned}$$

POLICIES: Choose a time $\tau$ and replace part at $min(T, \tau)$

Seek to minimize expected long run cost per unit time of replacements, which under renewal process assumptions is

$$\frac{C_2 P(T \leq \tau) + C_1 P(T > \tau)}{E\{min(T, \tau)\}}$$

- Reasons why optimal policy theory is not used much include model inadequacy, lack of good data on which to build models, changing conditions, and the need for flexibility in maintenance logistics.

## Condition-based Maintenance

- In a dynamic environment, use system condition measures $y(t)$ to monitor systems and to help plan maintenance. Widely applied in an informal way (engineering judgement).
- Covariates $x(t)$ related to usage, environment, previous maintenance can also be used.
- Optimality theory * here (e.g. Aven, Bergman, Makis and Jardine) also makes very strong renewal assumptions; its better to use an adaptive approach.

*Most policies replace a part at $min(T, \tau_d)$ where

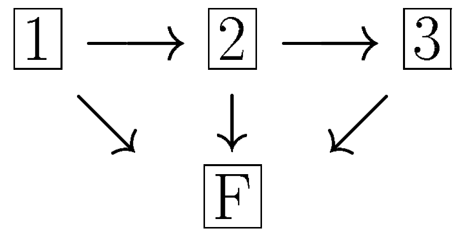
$$\tau_d = min(t : \lambda(t; \bar{y}(t)) > d)$$

- Need good data and simple models that can utilize them. Finding appropriate condition measures is crucial.
- Consideration of $P(T > t_{j+1} | T > t_j, \bar{y}(t_j), \bar{x}(t_j))$ at a sequence of monitoring times $t_1, t_2, \ldots$ is basic. This can be used to guide selection of the next monitoring time (if monitoring is not continuous) and decisions to do PM.
- In systems that have successfully avoided failure, there may be little direct empirical information about where the failure intensity $\lambda(t; \bar{y}(t), \bar{x}(t))$ is high.
- a model for $Y(t)$ (given $x(t)$) plus auxiliary assumptions about $\lambda(\cdot)$ can still be useful. "Control limit" policies where

$$\tau_d = min(t : y(t) > d)$$

are often used.

- Discrete-state models for $Y(t)$ are often appealing.




- "Delay-time" models

e.g. D. Banjevic et al. (2001) INFOR **39**, 32-50
Comments on application and a case study involving pump bearings and $y(t)$ based on vibration measurements
(PH-Markov Models)

P.A. Scarf (1997) Europ. J. Op. Res. **99**, 493-506
References to modelling and case studies.

## ADDITIONAL COMMENTS

- Reiterate need for “good” degradation or condition measures, simple models.
- Emphasize data collection in system monitoring
- Composite time scale methods (Kordonsky and Gertsbakh, Duchesne)