Sequential Test Planning for Polymer Composites

Yili Hong[1]

FTC 2018
October 05, 2018

Joint work with I-chen Lee[2], Sheng-Tsaing Tseng[3], and Tirthankar Dasgupta[4]

[1]Department of Statistics, Virginia Tech, Blacksburg, VA

[2]Department of Statistics, National Cheng Kung University, Tainan, Taiwan

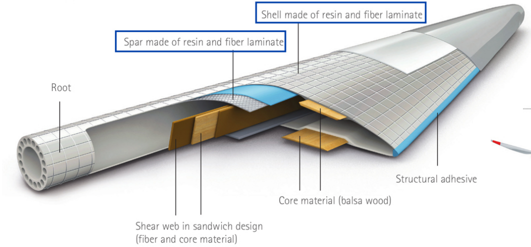
[3]Institute of Statistics, National Tsing Hua University, Hsinchu, Taiwan

[4]Department of Statistics and Biostatistics, Rutgers University, New Brunswick, NJ

Outline

Polymer composites

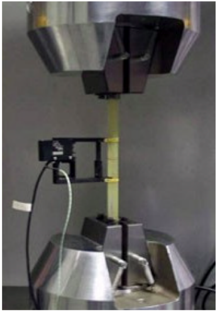
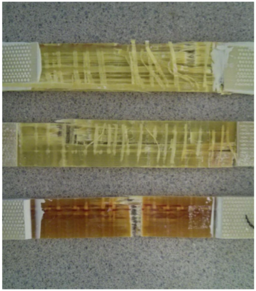
- A composite is any material made of more than one component with different physical or chemical properties.
- Polymer composites are made from polymers or from polymers along with other kinds of materials.
- Light weight, high strength, and long-term durability
- Wind turbine blades and aircraft are usually made of polymer composites.



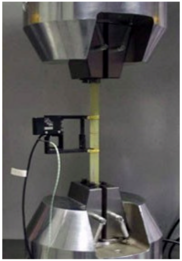
Source: http://3dprintingindustry.com/

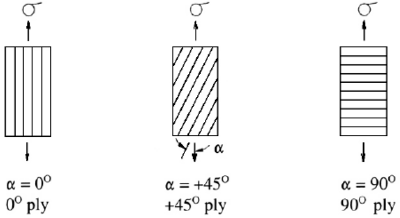
Fatigue tests for polymer composites

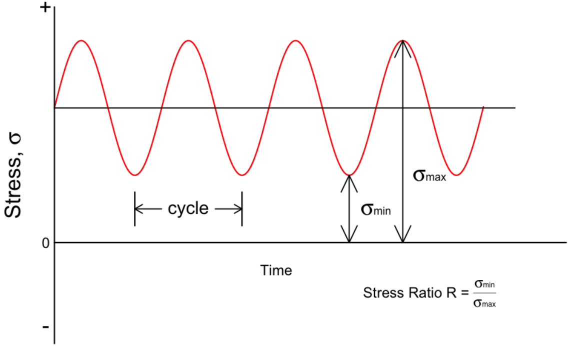
- The majority of testing is in accordance with the current standards provided in ASTM E739.
- The most common form of fatigue testing is cyclic constant amplitude fatigue testing.
- Fatigue occurs when the material is subject to varying levels of stress over a period of time.

Constant amplitude fatigue testing







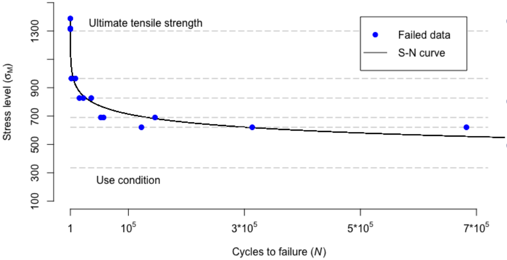
Objective

- One main goal of fatigue test is to demonstrate that a p proportion of the materials can last a certain number of cycles under the use stress level.

- It is related to the prediction of the quantile of the cycles to failure distribution at use condition.

- The objective of this talk is to develop a sequential test planning strategy to polymer composites fatigue testing.

Relationship between cycles and stress levels



- Three samples are tested for ultimate tensile strength σ_u.
- Four stress levels are used.
- S-N curve is fitted by an empirical model in Epaarachchi and Clausen(2003).

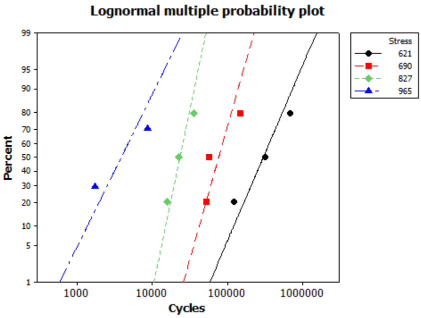
S-N curve

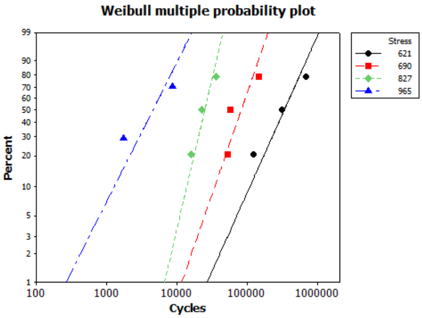
Epaarachchi and Clausen (2003) proposed the relationship as

$$N(\sigma_M) = \frac{1}{B} \log \left\{ 1 + \left(\frac{B}{A} \right) f^B \left(\frac{\sigma_u}{\sigma_M} - 1 \right) \left(\frac{\sigma_u}{\sigma_M} \right)^{\gamma(\alpha)-1} [1 - \psi(R)]^{-\gamma(\alpha)} \right\}.$$

- A is environmental effects on the material fatigue.
- B is effects from the material itself.
- σ_M and σ_m are the maximum and minimum strength during the test.
- f is the frequency of the cyclic testing.
- $\psi(R) = \begin{cases} R & -\infty < R < 1 \\ \frac{1}{R} & 1 < R < \infty \end{cases}$, where $R = \sigma_m / \sigma_M$.
- $\gamma(\alpha) = 1.6 - \psi |\sin(\alpha)|$ is a function of the smallest angle α.

Statistical model for cycles to failure





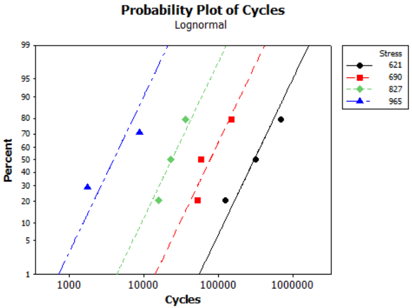
Assume the fatigue data are generated from the log-location-scale family of distributions, e.g., lognormal and the Weibull distribution.

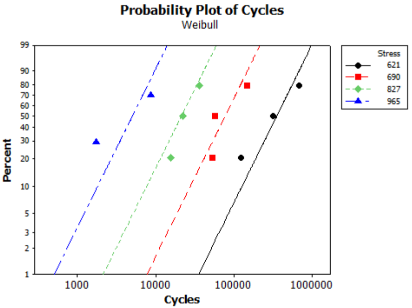
Statistical model for cycles to failure

- The log-location-scale family is used to describe the cycles to failure (T_i). That is,

$$F(t_i) = \Phi\left(\frac{\log(t_i) - \mu_{\sigma_{M_i}}(A, B)}{\nu}\right).$$

- The location parameter is $\mu_{\sigma_{M_i}}(A, B) = N(\sigma_{M_i})$.
- The scale parameter is ν.
- The unknown parameters are $\boldsymbol{\theta} = (A, B, \nu)$.





Likelihood function

- The likelihood function is

$$L(\boldsymbol{\theta}|data) = \prod_{i=1}^{n}\left[\frac{1}{\nu t_i}\phi\left(\frac{\log(t_i) - \mu_{\sigma_{M_i}}(A,B)}{\nu}\right)\right]^{(1-\delta_i)} \times \left[1 - \Phi\left(\frac{\log(t_i) - \mu_{\sigma_{M_i}}(A,B)}{\nu}\right)\right]^{\delta_i},$$

  where δ_i is the censoring indicator.
- The log-likelihood function is

$$l(\boldsymbol{\theta}|data) = \sum_{i=1}^{n}(1-\delta_i)\left[-\log\nu - \log t_i + \log\phi(z_i)\right] + \delta_i \log\left[1 - \Phi(z_i)\right],$$

  where $z_i = \left[\log(t_i) - \mu_{\sigma_{M_i}}(A,B)\right]/\nu$.

Asymptotic variance

- The Fisher information matrix is

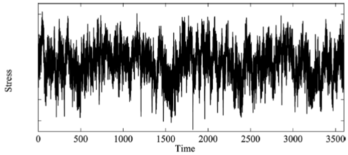
$$I(\boldsymbol{\theta}) = \mathrm{E}\left[-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right].$$

- The estimator of logarithm of the p quantile at stress s is $\log\left(\hat{\xi}_{p,s}\right) = \mu_s\left(\hat{A}, \hat{B}\right) + z_p \hat{\nu}$, where $\hat{A}$, $\hat{B}$, and $\hat{\nu}$ are estimators of A, B, and ν.
- The large-sample asymptotic variance of $\log\left(\hat{\xi}_{p,u_k}\right)$ at use condition u_k is derived as

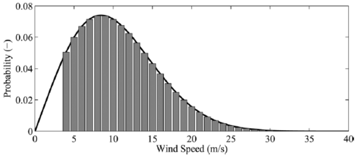
$$\mathrm{AVar}\left\{\log\left(\hat{\xi}_{p,u_k}\right)\right\} = \boldsymbol{c_k} \Sigma_{\boldsymbol{\theta}} \boldsymbol{c_k}',$$

where $\boldsymbol{c_k} = \left[\frac{\partial \mu_{u_k}(A,B)}{\partial A}, \frac{\partial \mu_{u_k}(A,B)}{\partial B}, z_p\right]$ and $\Sigma_{\boldsymbol{\theta}} = I^{-1}(\boldsymbol{\theta})$.

Distribution profile of use stress levels



Wind Turbine Blade Stress



Wind Speed Distribution

The weighted sum of the large-sample asymptotic variance of the estimator of the p quantile of the lifetime distribution at a vector of specified use levels can be expressed as

$$\text{AVar}\left\{\log\left(\hat{\xi}_{p,use}\right)\right\} = \sum_{k=1}^{K} w_k \text{AVar}\left\{\log\left(\hat{\xi}_{p,u_k}\right)\right\},$$

where w_k is the weight of the use level u_k and $\sum_{k=1}^{K} w_k = 1$.

Traditional optimum design

- Stress level
  - A fatigue test contains S stress levels.
  - Let $q_i = \sigma_{M_i}/\sigma_u$ and subject to $q_i \in [q_L, q_U]$, where $i = 1, ..., S$, q_L and q_U are the lower and upper bounds for stress levels.

- King et al. (2016) have discussed the optimum design on fatigue testings for composite material.
  - The results are obtained by minimizing $\text{AVar}\left\{\log\left(\hat{\xi}_{p,use}\right)\right\}$ under the assumption that values of true parameters are known.
  - The optimum design always ends up with a lower and a higher stress levels.

Difficulty

- Traditional optimum designs
  - Depending on values of parameters
  - Assume true values of parameters are known
  - Treat maximum likelihood (ML) estimators as the true values
  - Unreliable inference based on ML estimators for small sample size
- Limits of experiments
  - Experiments are time-consuming and costly
  - The number of samples can be tested at a time are limited

==============================

- (Sequential) Bayesian designs
  - Allow prior knowledge to be used in design
  - Deal with difficulties of traditional optimum designs
  - D-optimality (Dror and Steinberg, 2008; Roy et al., 2009; Zhu et al., 2014)

Bayesian test planning

- Bayesian D-optimality design (Chaloner and Verdinelli, 1995):

$$\eta^* = \arg\max_{\eta} \int_{\boldsymbol{\Theta}} \log\det\left(I\left(\boldsymbol{\theta}, \eta\right)\right) \pi\left(\boldsymbol{\theta}\right) d\boldsymbol{\theta},$$

  - η is all possible designs, η^* is the optimum design.
  - $I\left(\boldsymbol{\theta}, \eta\right)$ is the Fisher information matrix of $\boldsymbol{\theta}$ for the design η.
  - $\pi\left(\boldsymbol{\theta}\right)$ is the prior distribution of $\boldsymbol{\theta}$.
- Sequential Bayesian D-optimality design (Roy et al., 2009):

$$q^*_{D,n+1} = \arg\max_{q} \int_{\boldsymbol{\Theta}} \log\det\left(\sum_{i=1}^{n} I\left(\boldsymbol{\theta}, q_i\right) + I\left(\boldsymbol{\theta}, q\right)\right) \pi\left(\boldsymbol{\theta} | \mathbf{q}_n, \mathbf{t}_n\right) d\boldsymbol{\theta}.$$

  - $I\left(\boldsymbol{\theta}, q\right)$ is the Fisher information matrix for the stress level q.
  - $\pi\left(\boldsymbol{\theta} | \mathbf{q}_n, \mathbf{t}_n\right)$ is the posterior distribution for $\boldsymbol{\theta}$ given currently available data.

C-optimality

- For reliability analysts, it is important to obtain a precise prediction of a product's lifetime.
- The objective function is written as

$$\phi_n(q) = \int_{\boldsymbol{\Theta}} V(q, \boldsymbol{\theta})\, \pi(\boldsymbol{\theta}|\mathbf{q}_n, \mathbf{t}_n)\, d\boldsymbol{\theta},$$

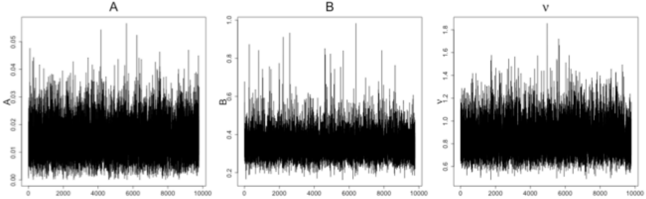
where $V(q, \boldsymbol{\theta}) = \left[\sum_{k=1}^{K} w_k \boldsymbol{c_k} \Sigma_{\boldsymbol{\theta}}(q)\, \boldsymbol{c_k}^T\right]$ and $\Sigma_{\boldsymbol{\theta}}(q) = \left[\sum_{i=1}^{n} I(\boldsymbol{\theta}, q_i) + I(\boldsymbol{\theta}, q)\right]^{-1}$.

- The optimum $(n+1)^{th}$ design is

$$q^*_{C,n+1} = \arg \min_{q \in [q_m, q_M]} \phi_n(q).$$

Algorithm

Step 1: MCMC draws, $\theta_n^{(j)}$, based on n observations (current data)



Step 2: evaluate the asymptotic variance of each draw for possible designs

Possible design (q)	$\theta_n^{(1)}$	$\cdots$	$\theta_n^{(J)}$	Average
q_L	$V\left(q_L, \theta_n^{(1)}\right)$	$\cdots$	$V\left(q_L, \theta_n^{(J)}\right)$	$\phi_n\left(q_L\right)$
$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
q_U	$V\left(q_U, \theta_n^{(1)}\right)$	$\cdots$	$V\left(q_U, \theta_n^{(J)}\right)$	$\phi_n\left(q_U\right)$

$$q_{C,n+1}^{*} = \arg\min\{\phi_n\left(q_L\right), \ldots, \phi_n\left(q_U\right)\}$$

Step 3: Add a new lifetime data under stress level $q_{C,n+1}^{*}$ from true distribution into current data

Example revisited

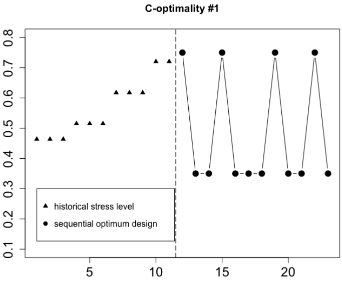
- The controlled settings are $R = 0.1$, $\alpha = 0$, and $\sigma_u = 1339.67$ MPa.
- Epaarachchi and Clausen (2003) provided the fitted values of A and B for different materials under different experimental settings.
- Assume the prior distributions are $A \sim$Unif$(0, 0.1)$, $B \sim$Unif$(0, 1)$, and $\nu^2 \sim$ Inv.Gamma$(4.5, 3)$.
- MCMC method and the Metropolis-Hastings algorithm are used to simulate the Markov chains.

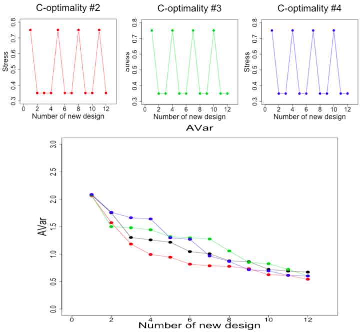
	$\hat{A}$	$\hat{B}$	$\hat{\nu}$
ML estimators	0.0157	0.3188	0.7259
Bayesian estimators	0.0145	0.3444	0.8413

- True values of parameters are set as $\hat{\boldsymbol{\theta}}_0 = (0.0157, 0.3188, 0.7259)$.

Proposed sequential Bayesian design

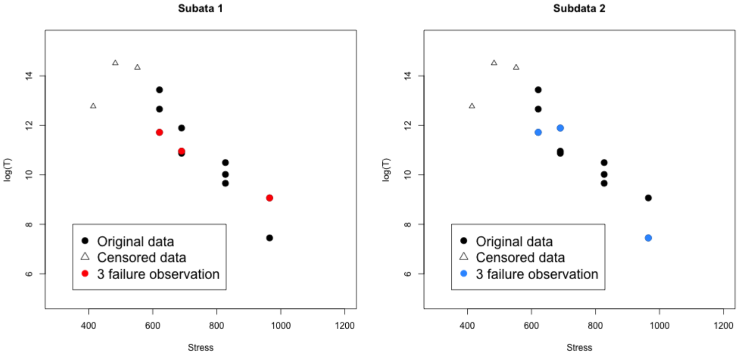
Let $q_L = 0.35$ and $q_U = 0.75$, and 12 new optimal design points are determined sequentially.





- The optimum stress levels can be divided into two groups: higher and lower levels.
- The proportion of higher and lower levels are 0.36 and 0.64.

Accurate estimation for small sample size?



Maximum likelihood estimators

	$\hat{A}$	$\hat{B}$	$\hat{\nu}$
$\hat{\theta}_0$	0.0157	0.3188	0.7259
$\hat{\theta}_1$	0.0005	0.7429	0.1658
$\hat{\theta}_2$	0.0162	0.3333	0.4044

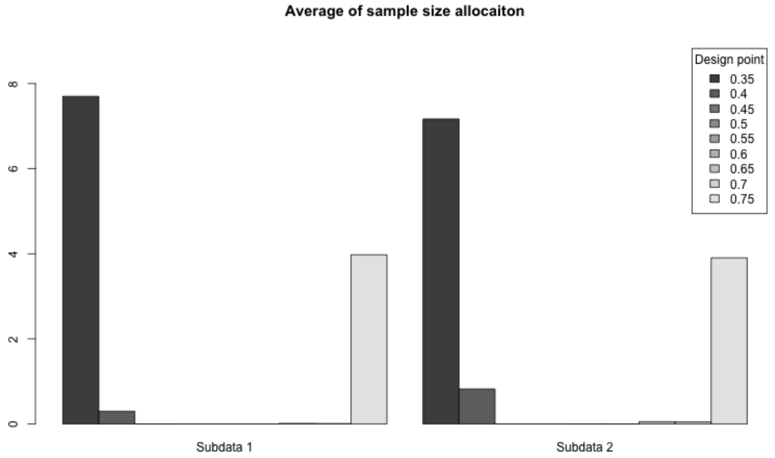
Comparison I

Assume the historical data is subdata 1 or subdata 2, and compare to the **two-level optimum design** (TOD, King et al. 2016).

Design	Stress levels	Sample size allocation
TOD based on $\hat{\boldsymbol{\theta}}_0$	(0.35, 0.75)	(8, 4)
TOD based on $\hat{\boldsymbol{\theta}}_1$	$(0.65, 0.75)$	$(11, 1)$
TOD based on $\hat{\boldsymbol{\theta}}_2$	(0.35, 0.75)	(8, 4)

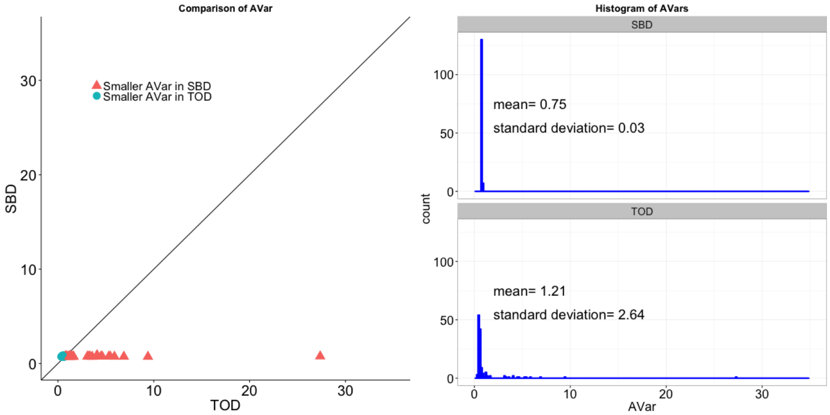
mean(AVar)	Subdata 1	Subdata 2
TOD by $\hat{\boldsymbol{\theta}}_0$	0.6236	0.4276
Sequential Bayesian design	0.7663	0.7170
TOD by ML estimators	4.0337	0.4240

Sample size allocation

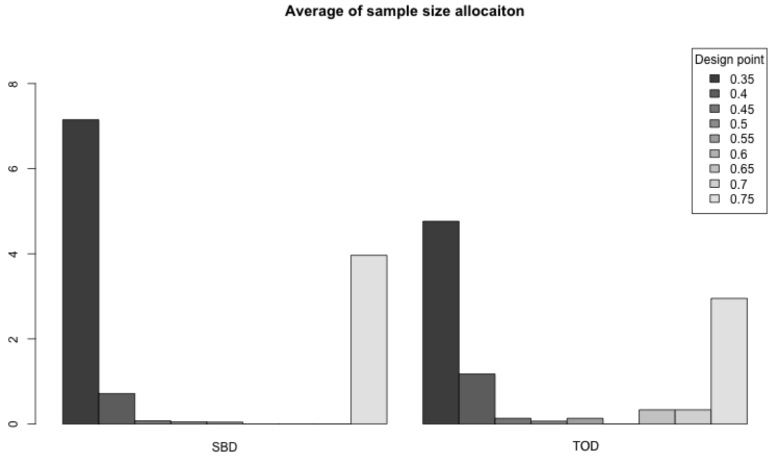


Comparison II

- Randomly taking 3 observations as the historical data, determine the traditional optimal design based on its ML estimators and the sequential Bayesion design (SBD).
- There are 165 combinations chosen in the simulation study.



Sample size allocation



Conclusion and Future Research

In sequential Bayesian designs,

- The optimal sequential Bayesian designs make new design points to either higher or lower stress levels.
- The sample size allocation is similar to the traditional two-level optimal design under the true values.
- The proposed method is more robust for small sample size.

Future research:

- Roy et al. (2009) proved the sequential Bayesian D-optimal design converges to the local D-optimal design for binary data.
- What are convergence properties of sequential Bayesian design for life test planning?

Reference

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science*, pages 273–304.

Dror, H. A. and Steinberg, D. M. (2008). Sequential experimental designs for generalized linear models. *Journal of the American Statistical Association*, 103(481):288–298.

Epaarachchi, J. A. and Clausen, P. D. (2003). An empirical model for fatigue behavior prediction of glass fibre-reinforced plastic composites for various stress ratios and test frequencies. *Composites Part A: Applied science and manufacturing*, 34(4):313–326.

King, C., Hong, Y., DeHart, S. P., DeFeo, P. A., and Pan, R. (2016). Planning fatigue tests for polymer composites. *Journal of Quality Technology (tentatively accepted)*.

Roy, A., Ghosal, S., and Rosenberger, W. F. (2009). Convergence properties of sequential bayesian d-optimal designs. *Journal of Statistical Planning and Inference*, 139(2):425–440.

Zhu, L., Dasgupta, T., and Huang, Q. (2014). A D-optimal design for estimation of parameters of an exponential-linear growth curve of nanostructures. *Technometrics*, 56(4):432–442.