



PEAK  
PROCESS  
PERFORMANCE  
PARTNERS LLC  
*Stability by design...*



# An Empirical Comparison of Methods for the Statistical Design and Analysis of Accelerated Stability Experiments

---

William R. Porter, PhD

Peak Process Performance Partners LLC

[PPPP-LLC@comcast.net](mailto:PPPP-LLC@comcast.net)

Stan Altan, PhD & Jyh-Ming Shoung, PhD

Statistics & Decision Sciences, Janssen Pharmaceutical R&D

[saltan@its.jnj.com](mailto:saltan@its.jnj.com), [jshoung@its.jnj.com](mailto:jshoung@its.jnj.com)

# Outline

---

- Early period 1889 – 1950:
  - Accelerated stability assessment using Arrhenius models to estimate shelf life at room temperature.
- First transitional period 1950 – 1985:
  - Moving from accelerated studies to RT studies.
- Real time room temperature period 1986-2010:
  - 1987 FDA guidance and 2004 Q1E.
- Second transitional period 2010 – current:
  - Movement from real time studies back to accelerated studies with focus on degradants.
- Models, design space.
- Isoconversion paradigm.
- Comparison of modeling algorithms using simulated data.

# Early period 1889 – 1950

---

Svante Arrhenius (1889): Reaction rates are proportional to inverse absolute temperature.

$$k_{T_1} = k_{T_0} \exp\left(C\left(\frac{T_1 - T_0}{T_0 T_1}\right)\right) = k_{T_0} \exp\left(C\left(\frac{1}{T_0} - \frac{1}{T_1}\right)\right)$$

Usually expressed as:

$$k_T = A e^{-\frac{E_a}{RT}}$$

where  $k_T$  = Degradation rate

$A$  = Non-thermal constant (frequency factor)

$E_a$  = Activation energy (e.g., kJ mol<sup>-1</sup>)

$R$  = Universal Gas Constant (e.g., 8.314 × 10<sup>-3</sup> kJ mol<sup>-1</sup> K<sup>-1</sup>)

$T$  = Absolute temperature (e.g., kelvin)



# Early period 1889 – 1950

---

W.D. Bigelow (1921): Logarithm of death time (lifespan) of organisms shortens as temperature increases.

Higuchi et al. (1950): Plot of log(half-life) gives same activation energy as plot of log(reaction rate) vs. reciprocal absolute temperature.

$$\ln(t_{1/2,T}) = A + \left(\frac{\Delta H}{R}\right)\left(\frac{1}{T}\right) \quad \ln(k_T) = B - \left(\frac{\Delta H}{R}\right)\left(\frac{1}{T}\right)$$

# First transitional period 1950 – 1985

---

McBride & Villars (1954), Huyberechts et al. (1955): Fit Arrhenius model to data using two-stage linear regression *with* weighting. (*widely ignored*)

Garrett (1954): Fit Arrhenius model to data using two-stage linear regression *without* weighting. (*most popular*)

Garrett (1956): First published real time stability data with confidence band.

McLeod et al. (1958): Complete worked out numerical examples of Garrett method. (*useful validation tool*)

Toothill (1961): Uses common intercept, isoconversion design to estimate rates in first stage.

# First transitional period 1950 – 1985

---

Comer (1969): First reported use of mainframe computer to assign shelf life using real time stability data based on confidence limit.

Carstensen & Su (1971): First reported use of nonlinear regression algorithm to fit Arrhenius model.

Carstensen & Nelson (1974–1976): Developed confidence limit model for real-time shelf life assignment incorporated into 1987 FDA Stability Guidance.

Davies & Budgett (1980): Shelf life estimated using Arrhenius model is log normally distributed. First discussion of propagation of error for Arrhenius model.

# First transitional period 1950 – 1985

---

Buncher & Tsay, *Statistics In the Pharmaceutical Industry* (1981): First edition covers ONLY accelerated stability studies. NO MENTION of real time studies.

King, Kung & Fung (1984): Direct point and interval estimation of shelf life at normal storage temperature from accelerated stability assessment using nonlinear regression.

# Real time room temperature period 1986-2010

---

## 1987 FDA guidance and 2004 ICH Q1E

Buncher & Tsay, *Statistics In the Pharmaceutical Industry* (1991): Second edition covers BOTH accelerated AND real time stability studies.

Shao & Chow (1991): Bayesian approach. Random batch effects.

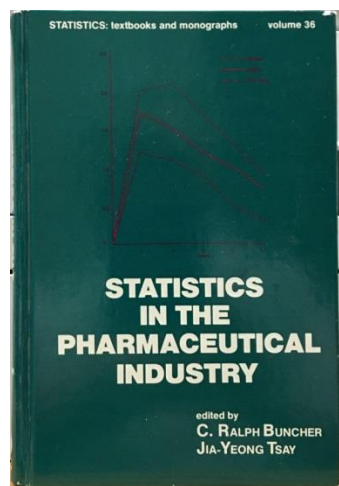
Shao & Chow (1994): Monte Carlo simulation using Bayesian Normal prior to evaluate posterior distribution of shelf life.

Su et al. (1994): Bayesian approach for Arrhenius model.

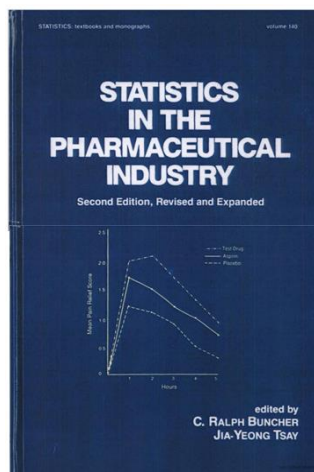


# Statistics in the Pharmaceutical Industry

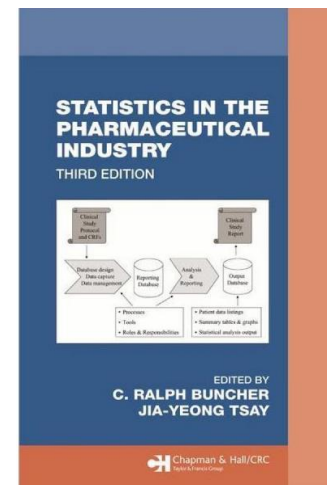
Edited by Charles Ralph Buncher, Jia-Yeong Tsay



**First Edition  
1981  
Accelerated Studies**



**Second Edition  
1991  
RT and Accelerated**



**Third Edition  
2006  
RT Studies**

# Second transitional period 2010 – current

---

Movement from real time studies back to accelerated studies with focus on degradants.

Humidity term incorporated to generate *extended Arrhenius model* for accelerated degradation of solids.

Focus shifts from assaying potency to measuring degradants.

# Solid state reaction mechanisms

Model <sup>a</sup>	Type	Integral form $g(\alpha) = kt$	$\alpha(t)$ Limit as $\alpha \rightarrow 0$
1-D Diffusion (D1)	Diffusion	$\alpha^2$	$(kt)^{(1/2)}$
2-D Diffusion (D2)	Diffusion	$[(1 - \alpha)\ln(1 - \alpha)] + \alpha$	$(2kt)^{(1/2)}$
3-D Diffusion-Jander (D3)	Diffusion	$[1 - (1 - \alpha)^{(1/3)}]^2$	$(9kt)^{(1/2)}$
Ginstling-Brounshtein (D4)	Diffusion	$1 - (2\alpha/3) - (1 - \alpha)^{(2/3)}$	$(9kt)^{(1/2)}$
Contracting line (R1)	Geometrical contraction	$\alpha$	$kt$
Contracting area (R2)	Geometrical contraction	$1 - (1 - \alpha)^{(1/2)}$	$kt/2$
Contracting volume (R3)	Geometrical contraction	$1 - (1 - \alpha)^{(1/3)}$	$kt/3$
Zero-order (F0)	Reaction-order	$\alpha$	$kt$
First-order (F1)	Reaction-order	$-\ln(1 - \alpha)$	$kt$
Second-order (F2)	Reaction-order	$(1 - \alpha)^{-1} - 1$	$kt$
Third-order (F3)	Reaction-order	$0.5 [(1 - \alpha)^{-2} - 1]$	$kt$
Power law (P2)	Nucleation	$\alpha^{(1/2)}$	$k^2 t^2$
Power law (P3)	Nucleation	$\alpha^{(1/3)}$	$k^3 t^3$
Power law (P4)	Nucleation	$\alpha^{(1/4)}$	$k^4 t^4$
Avrami-Erofeyev (A2)	Nucleation	$[-\ln(1 - \alpha)]^{(1/2)}$	$k^2 t^2$
Avrami-Erofeyev (A3)	Nucleation	$[-\ln(1 - \alpha)]^{(1/3)}$	$k^3 t^3$
Avrami-Erofeyev (A4)	Nucleation	$[-\ln(1 - \alpha)]^{(1/4)}$	$k^4 t^4$
Prout-Tompkins (B1)	Nucleation	$\ln[\alpha/(1 - \alpha)] + c^b$	$-\infty$

INFINITE DILUTION

Adapted from: Vyazovkin A & Wight CA. Kinetics in Solids *Annu Rev Phys Chem.* 48:127-1128 (1977)

# Models

## Pseudo-zero-order time-scaled model:

Porter (2013):

$$\lim_{[D] \rightarrow 0} [D]_{t,T,RH} = [D]_{t=0} + (ck_{T,RH}t)^m, \quad m \approx 1/2, 1, 2, 3, 4$$

## Effect of humidity (extended Arrhenius model):

Genton & Kesselring (1977):

$$\ln(k_{g,T,h}) = \ln(A_g) - E_a/RT + BRH$$

Clancy et al. (2017):

$$\ln(k_{g,T,h}) = \ln(A_g) - E_a/RT + n \ln(RH)$$

(mechanism dependent)



# Design space

All combinations of storage **time**, **temperature** and **relative humidity** where the extended Arrhenius model combined with pseudo-zero-order time-scaled reaction kinetic model adequately fits the experimental data.



*Stay in the space!*



Not too long!



Not too hot! Not too cold!



Not too damp! Not too dry!



No secondary degradation.

No phase changes.

- **Yes, it works for biologics!**



# Time-scaled extended Arrhenius model

Rauk et al. (2014):

Interchange shelf life and degradation rate:

$$k_{T,RH} = \frac{\sqrt[m]{[D]_{T,RH,t} - [D]_{t=0}}}{t}$$

$$t = \frac{\sqrt[m]{[D]_{T,RH,t} - [D]_{t=0}}}{k_{T,RH}}$$

Based on Genton & Kesselring (1977):

$$[D]_{t,T,RH} = [D]_{t=0} + \left\{ t \exp \left[ \left( \frac{1}{m} \right) \ln([D]_{fail} - [D]_{t=0}) - \ln(t_{fail,unstressed}) + \frac{E_a}{R} \left( \frac{1}{T_{unstressed}} - \frac{1}{T} \right) - B(RH_{unstressed} - RH) \right] \right\}^m + e_{t,T,RH}$$

Based on Clancy et al. (2017):

$$[D]_{t,T,RH} = [D]_{t=0} + \left\langle t \exp \left\{ \left( \frac{1}{m} \right) \ln([D]_{fail} - [D]_{t=0}) - \ln(t_{fail,unstressed}) + \frac{E_a}{R} \left( \frac{1}{T_{unstressed}} - \frac{1}{T} \right) - n[\ln(RH_{unstressed}) - \ln(RH)] \right\} \right\rangle^m + e_{t,T,RH}$$

# Isoconversion

---

Use in pharmaceutical research harks back to Higuchi et al. (1950)

Box & Lucas (1959) and Toothill (1961) showed isoconversion designs are statistically more efficient.

Vyazovkin & Wight (1988) and Vyazovkin & Sbirrazzuoli (2006) pointed out independence from assumed kinetic model (“model free”).

Waterman et al. (2007) popularized isoconversion designs in pharmaceutical industry.

# Isoconversion model

Eliminates time scale—just estimate shelf life under each stress condition:

$$\text{shelf life} = t_{fail,T,RH} = \frac{([D]_{fail} - [D]_{1,T,RH})(t_{2,T,RH} - t_{1,T,RH})}{[D]_{2,T,RH} - [D]_{1,T,RH}} + t_{1,T,RH} + f(e_{measurement})$$

Based on Genton & Kesselring (1977):

$$\ln(t_{fail,T,RH}) = \ln(t_{fail,unstressed}) + \frac{E_a}{R} \left( \frac{1}{T} - \frac{1}{T_{unstressed}} \right) - B(RH - RH_{unstressed}) + e_{\ln(t_{fail,T,h})}$$
$$e_{\ln(t_{fail,T,RH})} \sim N(0, [f(e_{measurement})]^2)$$

Based on Clancy et al. (2017):

$$\ln(t_{fail,T,RH}) = \ln(t_{fail,unstressed}) + \frac{E_a}{R} \left( \frac{1}{T} - \frac{1}{T_{unstressed}} \right) - n_1 [\ln(RH) - \ln(RH_{unstressed})] + e_{\ln(t_{fail,T,h})}$$
$$e_{\ln(t_{fail,T,RH})} \sim N(0, [f(e_{measurement})]^2)$$



# Models and initial degradant concentration

---

Traditional (Garrett, McBride & Villars) approach uses 2-stage linear regression but allows separate initial concentrations in first stage.

Toothill modification forces first stage to have common initial concentrations.

- But forcing common intercept introduces correlations requiring generalized least-squares regression in second stage.

King, Kung & Fung (and later, Rauk et al.) nonlinear regression also forces common initial concentrations.

Isoconversion shelf model only looks at end of shelf life and ignores initial concentrations.

# Comparing model-fitting algorithms

---

To assess accuracy, need simulated data derived from 2<sup>nd</sup> stage extended Arrhenius model with time-scaled pseudo-zero-order kinetic model for 1<sup>st</sup> stage with:

- known shelf life under unstressed conditions
- known activation energy
- known moisture sensitivity, and
- known error structure for measurement error.

We constructed an Excel workbook to generate simulated data assuming measurement error had two components:

- 1) constant Gaussian error independent of measured value, and
- 2) error proportional to measured value.

# Simulation (1)

<b>Actual storage temperature (°C):</b>	<b>50</b>	<b>50</b>	<b>60</b>	<b>60</b>	<b>70</b>	<b>70</b>
<b>Relative Humidity (%RH):</b>	<b>30.5</b>	<b>51.0</b>	<b>74.3</b>	<b>29.2</b>	<b>10.8</b>	<b>49.7</b>
<b>Sampling times <math>t</math> (in days):</b>	14	14	2	7	3	1
	28	28	3	14	7	2
	42	35	7	21	10	3

Extended Arrhenius model parameters:						
Time zero degradant level:	0.05 % API					
Degradant level specified limit:	0.5 % API					
Long term storage temperature:	30 °C					
Long term storage relative humidity:	50 %RH					
Long term storage desired shelf life:	2 years					
Implied Arrhenius activation energy:	29.95 kcal mol <sup>-1</sup>					
Moisture sensitivity: {	0.035 %RH <sup>-1</sup>	for Genton & Kesselring extended Arrhenius model				
	0.447	for Clancy et al. extended Arrhenius model				
Assay repeatability (noise):	0.005 %API					
Additional relative uncertainty:	5 ±% of [D]					

# Simulation (2)

---

One set of test data with randomly added uncertainty was generated for each of the six following conditions:

- 1) Diffusion model (time exponent  $m = \frac{1}{2}$ ) with Genton & Kesselring extended Arrhenius model.
- 2) Diffusion model (time exponent  $m = \frac{1}{2}$ ) with Clancy et al. extended Arrhenius model.
- 3) Geometric contraction/reaction order model (time exponent  $m = 1$ ) with Genton & Kesselring extended Arrhenius model.
- 4) Geometric contraction/reaction order model (time exponent  $m = 1$ ) with Clancy et al. extended Arrhenius model.
- 5) Induction model (time exponent  $m = 2$ ) with Genton & Kesselring extended Arrhenius model.
- 6) Induction model (time exponent  $m = 2$ ) with Clancy et al. extended Arrhenius model.

# Fitting Methods

---

- 1) Isoconversion
- 2) Two-stage common intercept OLS MLR OLS MLR Monte Carlo 1000 replicate datasets
- 3) Two-stage common intercept OLS MLR WLS MLR Monte Carlo 1000 replicate datasets
- 4) Two-stage common intercept OLS MLR OLS MLR single dataset used for other algorithms
- 5) Two-stage common intercept OLS MLR WLS MLR single dataset used for other algorithms
- 6) Two-stage common intercept NLS MLR WLS MLR
- 7) Two-stage common intercept NLS MLR GLS MLR
- 8) NLS modified King, Kung & Fung Genton-Kesselring (fixed  $m = \frac{1}{2}, 1$  or  $2$ )
- 9) NLS modified King, Kung & Fung Clancy et al. (fixed  $m = \frac{1}{2}, 1$  or  $2$ )
- 10) Naïve NLS modified King, Kung & Fung Genton-Kesselring fixed  $m = 1$  all datasets
- 11) NLS Raub et al. Genton-Kesselring  $m$  fitted (as per Clancy et al.)
- 12) NLS Raub et al. Clancy et al.  $m$  fitted (as per Clancy et al.)

- OLS = Ordinary least-squares regression
- MLR = Multiple linear regression
- WLS = Weighted least-squares regression
- GLS = Generalized least-squares regression
- NLS = Nonlinear least-squares regression

# Propagation of error in 2-stage models

---

A major defect of the original 2-stage methods is the failure to incorporate errors due to lack of fit in the 1<sup>st</sup> stage into parameter estimates obtained in the 2<sup>nd</sup> stage.

In the present study, this defect was remedied in one of two ways for the 2-stage common intercept models:

- The 2-stage process was replicated by generating additional pseudo-data (1000 replicates) from the original true (theoretical) data using the same assumed uncertainty algorithm (used in #2 OLS MLR OLS MLR MC & #3 OLS MLR WLS MLR MC methods). The median and average deviation of the 16<sup>th</sup> and 84<sup>th</sup> percentiles from the median 2<sup>nd</sup> stage results were reported.
- The variance-covariance matrix obtained from the 1<sup>st</sup> stage was used to generate weights to be used in the 2<sup>nd</sup> stage (used in #6 NLS MLR WLS MLR & #7 NLS MLR GLS MLR methods). The mean and standard error from the weighted 2<sup>nd</sup> stage results were reported.

# Results

---

After applying each fitting method, the percentage deviation of the point estimates of each fitted parameter to the true parameter values used to generate the simulated data (indicative of accuracy) were compared.

The relative standard errors of the deviations of the deviations of the parameter estimates (indicative of precision of fitted estimates) were also compared.

# Accuracy & Precision (1)

Simulation Results: Estimated Initial Degradant Concentration.												
Fitting Method	G&K $m = \frac{1}{2}$		G&K $m = 1$		G&K $m = 2$		Clancy $m = \frac{1}{2}$		Clancy $m = 1$		Clancy $m = 2$	
Isoconversion												
OLS MLR OLS MLR MC	-87.4%	± 11.9%	-68.6%	± 11.6%	-48.6%	± 9.2%	-87.6%	± 12.3%	-68.6%	± 11.4%	-49.1%	± 11.3%
OLS MLR WLS MLR MC	-87.4%	± 11.9%	-68.6%	± 11.6%	-48.6%	± 9.2%	-87.6%	± 12.3%	-68.6%	± 11.4%	-49.1%	± 11.3%
OLS MLR OLS MLR data	-19.8%	± 27.2%	-51.6%	± 20.0%	-5.8%	± 10.3%	10.8%	± 32.4%	30.2%	± 22.1%	-5.6%	± 21.5%
OLS MLR WLS MLR data	-19.8%	± 27.2%	-51.6%	± 20.0%	-5.8%	± 10.3%	10.8%	± 32.4%	30.2%	± 22.1%	-5.6%	± 21.5%
NLS MLR WLS MLR	-20.0%	± 28.0%	-52.0%	± 20.0%	-6.0%	± 10.0%	10.0%	± 34.0%	30.0%	± 22.0%	-6.0%	± 22.0%
NLS MLR GLS MLR	-20.0%	± 28.0%	-52.0%	± 20.0%	-6.0%	± 10.0%	10.0%	± 34.0%	30.0%	± 22.0%	-6.0%	± 22.0%
NLS King, Kung, Fung	-22.2%	± 24.4%	-54.1%	± 16.6%	-1.6%	± 9.3%	38.8%	± 36.7%	3.1%	± 25.2%	9.8%	± 20.3%
NLS Clancy et al.	-45.8%	± 33.9%	-30.0%	± 48.1%	-25.0%	± 11.5%	-7.4%	± 233.7%	29.0%	± 69.5%	8.3%	± 32.2%
Naïve NLS G&K $m=1$	281.7%	± 53.1%	-54.1%	± 16.6%	-368.3%	± 100.5%	404.5%	± 29.0%	17.8%	± 60.1%	-595.8%	± 173.5%
Naïve G&K $m$ fitted	-45.8%	± 33.9%	-30.0%	± 48.1%	-25.0%	± 11.5%	-107.3%	± 264.2%	-143.3%	± 222.9%	-18.2%	± 108.5%
Naïve Clancy $m$ fitted	-50.9%	± 110.4%	-91.0%	± 484.5%	254.6%	± 101.2%	-7.4%	± 233.7%	29.0%	± 69.5%	8.3%	± 32.2%

Simulation Results: Estimated Shelf Life at 30 °C 50 %RH.												
Fitting Method	G&K $m = \frac{1}{2}$		G&K $m = 1$		G&K $m = 2$		Clancy $m = \frac{1}{2}$		Clancy $m = 1$		Clancy $m = 2$	
Isoconversion												
OLS MLR OLS MLR MC	-9.0%	± 10.2%	-8.2%	± 6.2%	-6.2%	± 3.5%	-4.9%	± 10.1%	-4.1%	± 6.1%	-2.0%	± 3.6%
OLS MLR WLS MLR MC	-6.3%	± 11.8%	-3.9%	± 7.6%	-0.6%	± 5.7%	-3.8%	± 10.5%	-0.7%	± 9.0%	2.4%	± 9.0%
OLS MLR OLS MLR data	9.1%	± 5.4%	1.9%	± 2.5%	-4.2%	± 5.3%	4.9%	± 12.5%	6.7%	± 8.6%	-8.6%	± 10.4%
OLS MLR WLS MLR data	4.8%	± 7.2%	0.2%	± 2.3%	-1.2%	± 5.3%	20.4%	± 23.9%	-6.2%	± 10.7%	-4.5%	± 10.9%
NLS MLR WLS MLR	5.2%	± 6.6%	-1.6%	± 1.2%	-3.6%	± 1.8%	4.6%	± 17.3%	-0.6%	± 11.2%	-6.7%	± 4.4%
NLS MLR GLS MLR	4.1%	± 14.8%	-2.0%	± 5.7%	-3.2%	± 2.5%	10.4%	± 16.6%	-5.3%	± 6.2%	-6.1%	± 3.3%
NLS King, Kung, Fung	7.6%	± 10.0%	0.7%	± 4.3%	-1.0%	± 2.2%	10.1%	± 11.5%	1.2%	± 6.1%	-4.3%	± 3.5%
NLS Clancy et al.	9.9%	± 10.9%	0.7%	± 4.4%	-1.2%	± 1.9%	10.0%	± 11.9%	1.2%	± 6.3%	-4.3%	± 3.7%
Naïve NLS G&K $m=1$	-1.8%	± 24.3%	0.7%	± 4.3%	1.9%	± 23.6%	-10.0%	± 12.1%	-17.7%	± 11.2%	-20.2%	± 18.4%
Naïve G&K $m$ fitted	9.9%	± 10.9%	0.7%	± 4.4%	-1.2%	± 1.9%	-5.9%	± 9.1%	-15.2%	± 12.1%	-32.2%	± 7.4%
Naïve Clancy $m$ fitted	48.0%	± 54.5%	18.8%	± 37.8%	-15.6%	± 14.4%	10.0%	± 11.9%	1.2%	± 6.3%	-4.3%	± 3.7%



# Accuracy & Precision (2)

**Simulation Results: Estimated Arrhenius Activation Energy.**

Fitting Method	G&K $m = \frac{1}{2}$	G&K $m = 1$	G&K $m = 2$	Clancy $m = \frac{1}{2}$	Clancy $m = 1$	Clancy $m = 2$
Isoconversion	1.7% ± 2.6%	1.2% ± 0.8%	0.7% ± 0.5%	-2.4% ± 10.6%	7.5% ± 4.0%	-0.1% ± 0.7%
OLS MLR OLS MLR MC	-1.9% ± 2.5%	-1.8% ± 1.5%	-1.5% ± 0.8%	-0.7% ± 2.4%	-0.8% ± 1.5%	-0.5% ± 0.9%
OLS MLR WLS MLR MC	-1.2% ± 2.7%	-0.9% ± 1.6%	-0.3% ± 1.2%	-0.5% ± 2.5%	-0.3% ± 1.8%	0.2% ± 1.6%
OLS MLR OLS MLR data	2.9% ± 1.1%	0.7% ± 0.3%	-1.1% ± 0.5%	-0.2% ± 2.7%	0.7% ± 1.7%	-2.1% ± 1.0%
OLS MLR WLS MLR data	2.2% ± 1.4%	0.3% ± 0.2%	-0.5% ± 0.4%	1.5% ± 4.5%	-1.0% ± 2.6%	-1.2% ± 1.1%
NLS MLR WLS MLR	2.6% ± 1.2%	0.4% ± 0.2%	-0.6% ± 0.4%	-0.1% ± 3.2%	0.2% ± 2.3%	-1.3% ± 1.0%
NLS MLR GLS MLR	2.5% ± 2.2%	0.3% ± 1.0%	-0.5% ± 0.5%	0.2% ± 1.9%	-0.1% ± 1.1%	-1.2% ± 0.7%
NLS King, Kung, Fung	2.7% ± 2.0%	0.4% ± 0.9%	-0.4% ± 0.5%	0.4% ± 2.3%	0.1% ± 1.4%	-1.1% ± 0.8%
NLS Clancy et al.	3.2% ± 2.1%	0.4% ± 0.9%	-0.2% ± 0.4%	0.4% ± 2.4%	0.0% ± 1.4%	-1.1% ± 0.9%
Naïve NLS G&K $m=1$	-0.8% ± 5.2%	0.4% ± 0.9%	3.4% ± 4.9%	-5.4% ± 2.9%	-5.3% ± 3.0%	-3.3% ± 5.1%
Naïve G&K $m$ fitted	-3.3% ± 5.1%	-3.3% ± 5.1%	-3.3% ± 5.1%	-3.3% ± 5.1%	-3.3% ± 5.1%	-3.3% ± 5.1%
Naïve Clancy $m$ fitted	-3.3% ± 5.1%	-3.3% ± 5.1%	-3.3% ± 5.1%	-3.3% ± 5.1%	-3.3% ± 5.1%	-3.3% ± 5.1%

**Simulation Results: Estimated Moisture Sensitivity.**

Fitting Method	G&K $m = \frac{1}{2}$	G&K $m = 1$	G&K $m = 2$	Clancy $m = \frac{1}{2}$	Clancy $m = 1$	Clancy $m = 2$
Isoconversion	6.2% ± 3.5%	-1.2% ± 1.0%	0.4% ± 0.7%	-15.2% ± 35.9%	-7.2% ± 13.7%	-3.0% ± 3.6%
OLS MLR OLS MLR MC	-3.1% ± 4.0%	-2.4% ± 2.3%	-1.8% ± 1.3%	7.2% ± 10.2%	4.9% ± 6.0%	1.9% ± 3.6%
OLS MLR WLS MLR MC	-2.8% ± 4.2%	-2.0% ± 2.7%	-1.3% ± 1.7%	6.8% ± 11.3%	3.6% ± 6.9%	1.4% ± 4.4%
OLS MLR OLS MLR data	4.6% ± 1.7%	-1.7% ± 0.4%	-2.2% ± 0.8%	-0.1% ± 10.9%	-13.1% ± 7.0%	-2.3% ± 4.0%
OLS MLR WLS MLR data	5.1% ± 1.5%	-2.0% ± 0.2%	-2.0% ± 0.4%	-11.1% ± 13.5%	-4.9% ± 7.6%	0.0% ± 3.2%
NLS MLR WLS MLR	4.9% ± 1.7%	-2.0% ± 0.3%	-2.0% ± 0.3%	-3.5% ± 11.6%	-8.1% ± 7.7%	0.0% ± 3.2%
NLS MLR GLS MLR	4.6% ± 3.1%	-2.0% ± 1.4%	-1.7% ± 0.6%	-5.2% ± 7.4%	-5.7% ± 4.0%	-0.4% ± 2.3%
NLS King, Kung, Fung	4.6% ± 2.8%	-2.0% ± 1.2%	-1.8% ± 0.5%	-6.3% ± 9.1%	-6.4% ± 4.8%	-0.4% ± 2.6%
NLS Clancy et al.	5.7% ± 3.1%	-2.4% ± 1.4%	-0.7% ± 0.6%	-6.9% ± 10.0%	-5.4% ± 5.4%	-0.5% ± 3.0%
Naïve NLS G&K $m=1$	-4.5% ± 6.6%	-2.0% ± 1.2%	15.9% ± 6.5%	-56.1% ± 4.4%	-64.4% ± 4.9%	-72.3% ± 8.5%
Naïve G&K $m$ fitted	5.7% ± 3.1%	-2.4% ± 1.4%	-0.7% ± 0.6%	-62.5% ± 3.5%	-66.6% ± 5.6%	-59.9% ± 4.5%
Naïve Clancy $m$ fitted	182.6% ± 32.3%	172.2% ± 30.4%	242.5% ± 38.3%	-6.9% ± 10.0%	-5.4% ± 5.4%	-0.5% ± 3.0%

# Conclusions (1)

---

Modified two-stage methods gave the most accurate and precise predictions of initial degradant concentration, shelf life, activation energy and moisture sensitivity.

- OLS MLR OLS MLR MC & NLS MLR WLS MLR MC gave most *precise* estimate of initial concentration. Also performed well for estimating unstressed shelf life, activation energy and moisture sensitivity. (But note that these were median results from 1000 replicates.)
- OLS MLR OLS MLR data & NLS MLR WLS MLR data gave satisfactory point estimates of parameters that could be used as starting values for nonlinear models.
- NLS MLR WLS MLR & NLS MLR GLS MLR gave most *accurate* estimate of initial concentration. Also performed well for estimating unstressed shelf life, activation energy and moisture sensitivity. (But note that we used weighting in 2<sup>nd</sup> stage obtained from 1<sup>st</sup> stage variance-covariance matrix.)

Direct nonlinear regression methods less accurate and precise.

- But note that we used conventional non-linear modeling assuming all data have equal weight.

# Conclusions (2)

---

Assuming the wrong time exponent gave wildly inaccurate predictions:

- ALWAYS CHECK FIRST STAGE PLOTS FOR CURVATURE INDICATING CHANGE OF TIME SCALE IS NEEDED! Or, use NLS Raub time-scaled model to estimate time exponent, then select  $m = \frac{1}{2}$ , 1 or 2 and re-fit.

Try BOTH Genton & Kesselring AND Clancy et al. extended Arrhenius models!

- Fitting “wrong” model gives moisture sensitivity parameter estimates with larger standard errors than fitting “right” model.

# Questions

---

References and details available from the authors.



# References (1)

---

- Arrhenius SA. Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren. [On the reaction rate for the inversion of sucrose by acids.] *Z Physik Chem* 4:226–248 (1889). Partial translation: On the reaction velocity of the inversion of cane sugar by acids (excerpt). In: Back MH, Laidler KJ, eds. *Selected Readings in Chemical Kinetics*. Pergamon, Oxford (1967).
- Bigelow WD. The logarithmic nature of thermal death time curves. *J Infec Dis* 29:528–536 (1921).
- Box GEP, Lucas HL. Design of experiments in non-linear situations. *Biometrika* 46:77–90, (1959).
- Buncher CR, Tsay J-Y, eds. *Statistics In the Pharmaceutical Industry*. First Edition. Marcel Dekker (1981).
- Buncher CR, Tsay J-Y, eds. *Statistics In the Pharmaceutical Industry*. Second Edition. Marcel Dekker (1994).
- Buncher CR, Tsay J-Y, eds. *Statistics In the Pharmaceutical Industry*. Third Edition. CRC Press (2005).
- Carstensen JT, Nelson E. “Terminology regarding labeled and contained amounts in dosage forms”. *J Pharm Sci* 65(2):311–312 (1976).
- Carstensen JT, Su KSE. Statistical aspects of Arrhenius plotting. *Bull Parent Drug Assoc* 25(6): 287–302 (1971).
- Clancy D, Hodnett N, Orr R, Owen M, Peterson J. Kinetic Model Development for Accelerated Stability Studies. *AAPS PharmSciTech* 18(4): 1158–1176 (2017).
- Comer JA. “Processing of stability data for FDA submission”. The Dating of Pharmaceuticals Conference. Univ. Wisconsin Conference proceedings. 93–106 (1969).
- Davies OL, Budgett DA. Accelerated storage tests on pharmaceutical products: effect of uncertainty structure of assay and uncertainties in recorded temperature. *J Pharm Pharmacol* 32: 155–159 (1980).
- FDA. Guideline For Submitting Documentation for the Stability of Human Drugs and Biologics (1987).
- Garrett ER. Studies on the stability of fumagillin. III. Thermal degradation in the presence and absence of air. *J Amer Pharm Assoc Sci Ed* 43: 539–543 (1954).

# References (2)

---

- Garrett, E. (1956). "Prediction of stability in pharmaceutical preparations, Part III. Comparison of vitamin stabilities in different multivitamin preparations. *J Amer Pharm Assoc* 45(7):470–473 (1956).
- Genton D, Kesselring UW. Effect of temperature and relative humidity on nitrazepam stability in solid state. *J Pharm Sci*, 66(5): 676–680 (1977).
- Higuchi T, Havinga A, Busse L. The kinetics of the hydrolysis of procaine. *J Amer Pharm Assoc Sci Ed* 39(7): 405–410 (1950).
- Hyberechts MS, Halleux A, Kruys P. "Une application de calcul statistique a la cinetique chimique". *Bull. Soc. Chim. Belg.*, 64:203-209 (1955).
- ICH Harmonised Tripartite Guideline. Evaluation for stability data Q1E. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (2003).
- King S-YP, Kung M-S, Fung H-L. Statistical prediction of drug stability based on nonlinear parameter estimation. *J Pharm Sci* 73: 657–62 (1984).
- McBride WR, Villars DS. An application of statistics to reaction kinetics. *Anal Chem* 26: 901–904, (1954).
- McLeod, H.A., Pelletier, O. and Campbell, J.A. (1958). "The Prediction of Expiration Dates for Multivitamin Preparations by Accelerated Storage Tests." *Canadian Pharmaceutical Journal Scientific Section* 91: 173–180 (1958).
- Porter WR. Degradation of Pharmaceutical Solids Accelerated by Changes in Both Relative Humidity and Temperature and Combined Storage Temperature and Storage Relative Humidity (T×h) Design Space for Solid Products. *J Validation Technol* 19(2) online (2013).
- Rauk AP, Guo K, Hu Y, Cahya S, Weiss WF IV. Arrhenius time-scaled least squares: A simple, robust approach to accelerated stability data analysis for bioproducts. *J Phar Sci* 103: 2278–2286 (2014).

# References (3)

---

- Shao, J., & Chow, S. C. Constructing release targets for drug products: A Bayesian decision theory approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 40(3), 381-390. (1991).
- Shao, J., & Chow, S. C. Statistical inference in stability analysis. *Biometrics*, 753-763 (1994).
- Su, X. Y., Po, A. L. W., & Yoshioka, S. A Bayesian approach to Arrhenius prediction of shelf-life. *Pharmaceutical Research*, 11(10), 1462-1466 (1994).
- Tootill JPR. Slope-ratio design for accelerated storage tests. *J Pharm Pharmacol* 13: 75T–86T, (1961).
- Vyazovkin A, Wight CA. Kinetics in solids. *Annu Rev Phys Chem* 48: 125–149 (1977).
- Vyazovkin S, Wight CA. Isothermal and non-isothermal kinetics of thermally stimulated reactions of solids. *International Rev Phys Chem* 17(3): 407–433 (1998).
- Vyazovkin S, Sbirrazzuoli N. Isoconversional Kinetic Analysis of Thermally Stimulated Processes in Polymers. *Macromol Rapid Commun* 27: 1515–1532 (2006).
- Waterman KC, Carella AJ, Gumkowski MJ, Lukulay P, MacDonald BC, Roy MC, Shamblin SL. Improved Protocol and Data Analysis for Accelerated Shelf-Life Estimation of Solid Dosage Forms. *Pharmaceutical Research*, 24(4), 780–790 (2007).

---

Notes on historical models

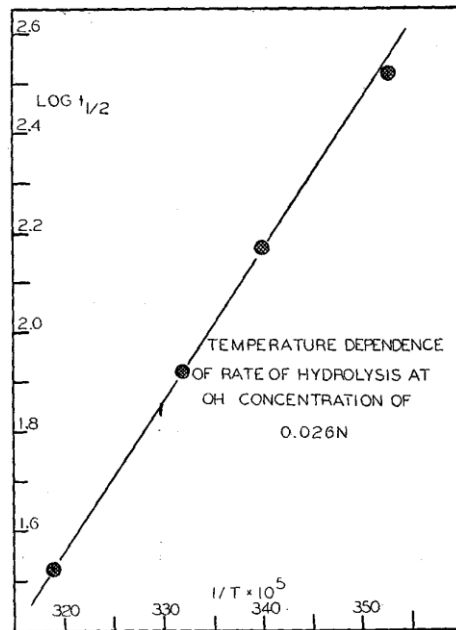
# BACKUP SLIDES



# Higuchi et al. half-life

Isoconversion was used.

A plot of the logarithm of the half-life vs. reciprocal absolute temperature gave the Arrhenius activation energy as a function of the slope of the line:



$$[C_{t,T}] = [C_{t=0,T}] \exp(-k_T t_T)$$

$$\ln([C_{t,T}]) = \ln([C_{t=0,T}]) - k_T t_T$$

$$t_{1/2,T} = \frac{\ln([C_{t=0,T}]) - \ln\left(\frac{[C_{t=0,T}]}{2}\right)}{k_T}$$

$$t_{1/2,T} = \frac{\ln\left(2 \frac{[C_{t=0,T}]}{[C_{t=0,T}]}\right)}{k_T} = \frac{\ln(2)}{k_T}$$

$$k_T = s \exp\left(\frac{-\Delta H_a}{RT}\right) = \frac{\ln(2)}{t_{1/2,T}}$$

$$t_{1/2,T} = \frac{\ln(2)}{s \exp\left(\frac{-\Delta H_a}{RT}\right)}$$

$$\ln(t_{1/2,T}) = \ln\left(\frac{\ln(2)}{s}\right) + \frac{\Delta H_a}{RT}$$

# McBride & Villars, Huyberegts et al.

---

## 2-Stage linear regression method

- Stage One:
  - Individual lines were fitted by least-squares regression to measured degradation at different times for each temperature.
  - Variance for each line at each temperature was estimated.
- Stage Two:
  - Weighted least-squares line was fitted to logarithmic form of Arrhenius model using weights  $w$  calculated for each temperature separately.

$$[D_{t,T_i}] = [D_{t=0,T_i}] + k_{T_i} t_{T_i} + e_{T_i}$$

$$k_T = A \exp\left(\frac{-E_a}{RT}\right)$$

$$\ln(k_T) = \ln(A) - \left(\frac{E_a}{R}\right)\left(\frac{1}{T}\right) + \frac{e_T}{\sqrt{k_T^2 \sum t_T^2}}$$

$$w_T^2 = k_T^2 \sum t_T^2$$

# Garrett (1954, 1956)

## 2-Stage linear regression method

- Stage One:
  - Individual lines were fitted by unweighted least-squares regression to appropriate kinetic model (concentration of active or degradant, pseudo-zero-order, pseudo-first order, or pseudo-second-order) at different times for each temperature.
- Stage Two:
  - Unweighted least-squares line was fitted to logarithmic form of Arrhenius model.

$$[C_{t,T}] = [C_{t=0,T}] - k_T t_T$$

*or*

$$\log_{10}([C_{t,T}]) = \log_{10}([C_{t=0,T}]) - k_T t_T$$

*or*

$$\frac{1}{[D_{t,T}]} = \frac{1}{[D_{t=0,T}]} + k_T t_T$$

$$\log_{10}(k_T) = P - \left( \frac{\Delta H_a}{\ln(10)R} \right) \left( \frac{1}{T} \right)$$

# Toothill

---

Toothill described a variation on the slope-ratio assay by Finney (**Finney DJ**. *Statistical Method in Biological Assay*, p 187, Charles Griffen, 1952).

- This method uses ANOVA to estimate slopes; it was not originally cast as a regression problem. We borrowed key concepts.
- All treatments start with the same initial value.
- The reciprocal absolute temperatures for the stress conditions must be selected to be equally spaced on the reciprocal absolute temperature scale. This forces the slopes of the decay lines to be in geometric progression. The experiment is designed so that each stress condition is followed until the same extent of degradation (isoconversion) has occurred. It also forces the slopes to have the same variance.
- The design permits fitting the Arrhenius model in one step.

# Carstensen & Su

---

It is possible to combine the two-stage approach in a single nonlinear model that can be fitted by nonlinear regression methods.

$$[C_{t,T}] = [C_{t=0,T}] - k_T t_T$$

*and*

$$\log_{10}(k_T) = \gamma - \left( \frac{E_a}{\ln(10)R} \right) \left( \frac{1}{T} \right)$$

*so*

$$[C_{t,T}] = [C_{t=0,T}] - t_T \left[ \gamma \times 10^{-\left( \frac{E_a}{\ln(10)R} \right) \left( \frac{1}{T} \right)} \right]$$

# King, Kung & Fung

$$\left. \begin{aligned} [C_{t,T}] &= [C_{t=0}] - k_T t_T \\ [C_{\text{target}}] &= [C_{t=0}] - k_{T_{\text{target}}} t_{\text{shelf-life}} \end{aligned} \right\} k_{T_{\text{target}}} = \frac{[C_{t=0}] - [C_{\text{target}}]}{t_{\text{shelf-life}}}$$

$$\left. \begin{aligned} k_T &= A \exp\left(-\frac{E_a}{RT}\right) \\ k_{T_{\text{target}}} &= A \exp\left(-\frac{E_a}{RT_{\text{target}}}\right) \end{aligned} \right\} A = k_{T_{\text{target}}} \left( \frac{1}{\exp\left(-\frac{E_a}{RT_{\text{target}}}\right)} \right)$$

$$[C_{t,T}] = \left[ [C_{t=0}] - t_T \left( \frac{[C_{t=0}] - [C_{\text{target}}]}{t_{\text{shelf-life}}} \right) \right] \exp \left[ \left( \frac{E_a}{R} \right) \left( \frac{1}{T_{\text{target}}} - \frac{1}{T} \right) \right]$$

By solving for the Arrhenius frequency factor  $A$  using the rate at the target temperature at which the shelf-life is to be predicted, we can estimate the shelf-life at the target temperature directly using nonlinear regression.

# Clancy et al.

---

Clancy and co-workers describe several alternate models other than the models presented here. Consult their published paper.

- The Raub et al. model captures several of the models described by Clancy et al.
- Clancy et al. define the “rate constant” differently. Their “rate” = our  $k^m$ . Beware!