

# Parameter Subset Selection for Mixed-Effects Models

Kathleen Schmidt, Ralph C. Smith, Jason Bernstein, Ana Kupresanin

August 7, 2018



LLNL-PRES-754903

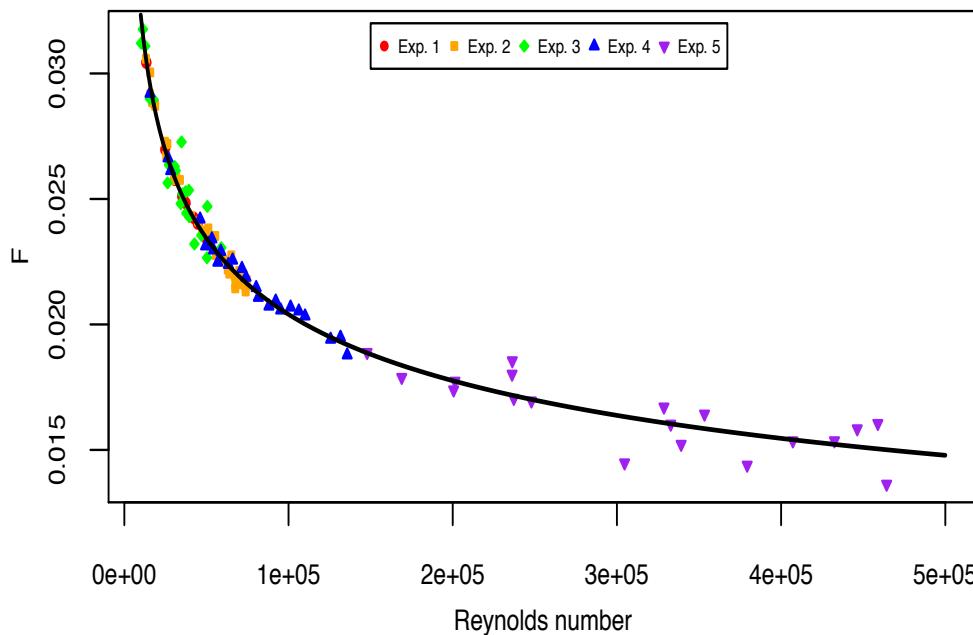
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

 Lawrence Livermore  
National Laboratory

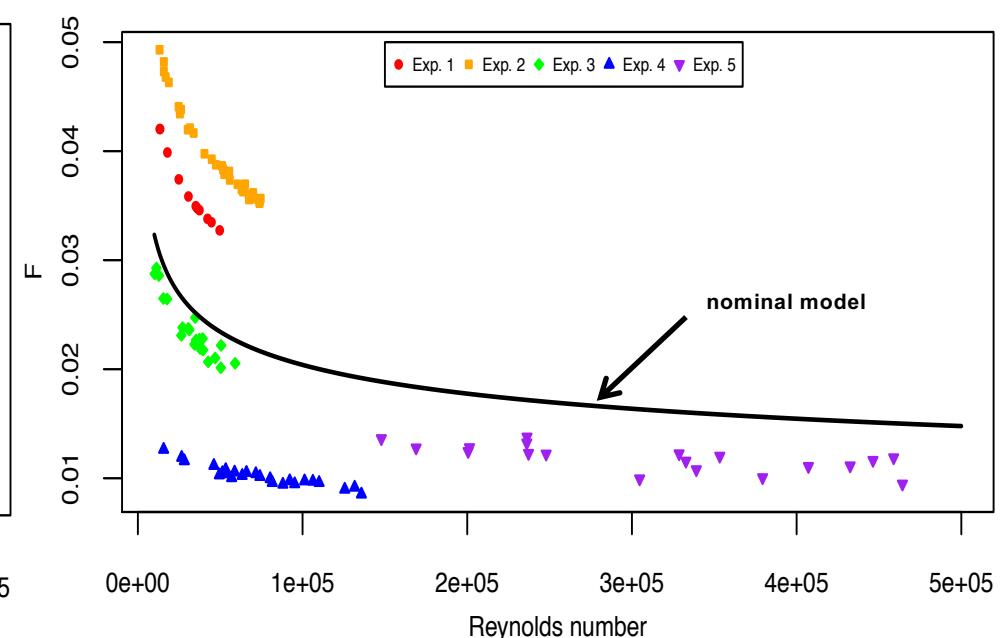
# Why Use Mixed-Effects Models?

- One set of parameters does not work for all groups
  - e.g., biological variation, differently manufactured materials, etc.

Desirable Situation



Reality



$$\text{McAdams Correlation: } F = \xi_1 Re^{\xi_2}$$

# Traditional vs Mixed-Effects Models

- “Traditional” Statistical Model

$$y_j = f(x_j; \theta) + \varepsilon_j, \quad \varepsilon_j \stackrel{iid}{\sim} N(0, \sigma^2)$$

- Mixed-Effects Model

$$y_{ij} = f(x_{ij}; \beta, b_i) + \varepsilon_{ij}$$

Fixed Effects      Random Effects

Different random effect parameters for each group

Assumptions:  $b_i \sim N(0, \Psi)$

$$\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma^2)$$



Lawrence Livermore National Laboratory

LLNL-PRES-754903



# Uncertainty Quantification for Mixed-Effects Models

---

- Bayesian inference
  - Estimate parameters and quantify uncertainty
  - Models in physics and biology can be high dimensional
    - Bayesian inference may not be feasible
- Reduce the number of model parameters
  - Sensitivity analysis, dimension reduction, model selection, etc.
    - Many techniques cannot be generalized to mixed-effects models

**Focus:** A method for selecting a parsimonious mixed-effect model



# Materials Strength Model: Stress-Strain Data

- 4 data sets for Tantalum

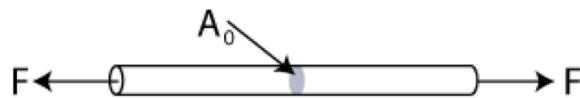
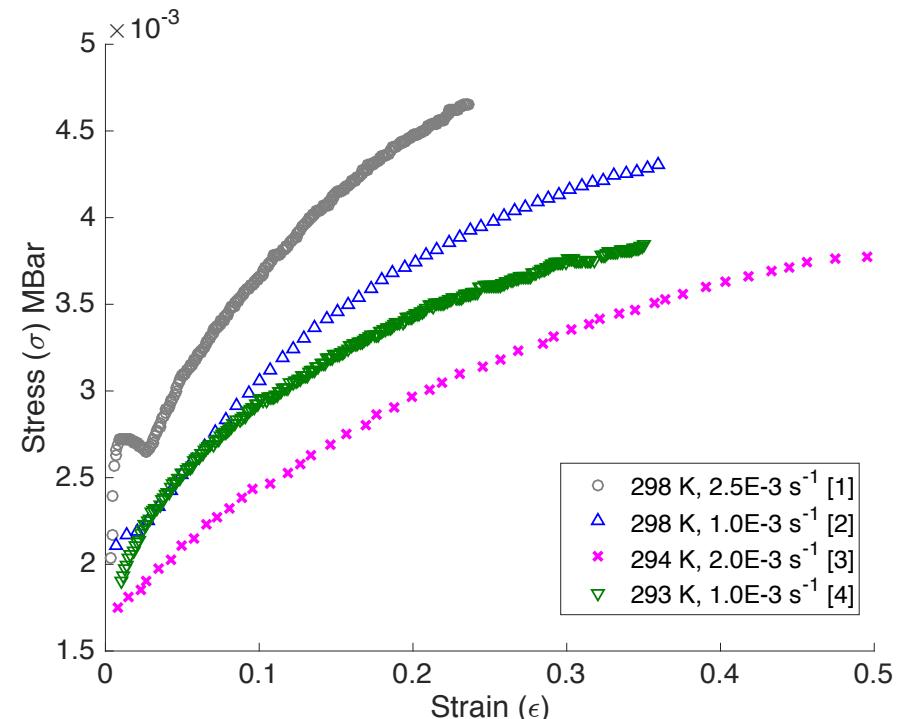
[1] (Murr, et al., 1995)

[2] (Gourdin, et al., 1994)

[3] (Johnson and Holmquist, 1989)

[4] (Perez-Prado, et al., 2001)

- Similar experimental conditions, very different responses



$$\text{Stress, } \sigma = \frac{\text{Force}}{\text{Cross-Sectional Area}} = \frac{F}{A_0}$$



Physics illustrations from (NDT Resource Center, 2008)

$$\text{Strain} = \frac{\text{Elongation}}{\text{Original Length}} = \frac{\Delta L}{L_0}$$



Lawrence Livermore National Laboratory

LLNL-PRES-754903



# Materials Strength Model: Example

- Johnson-Cook Model

$$\sigma = (A + B\varepsilon_p^n) (1 + C \ln(\dot{\varepsilon}_p^*)) \left(1 - T^{*^m}\right)$$

$$\dot{\varepsilon}_p^* := \frac{\dot{\varepsilon}_p}{\dot{\varepsilon}_{p0}} \quad \text{and} \quad T^* := \frac{T - T_0}{T_m - T_0}$$

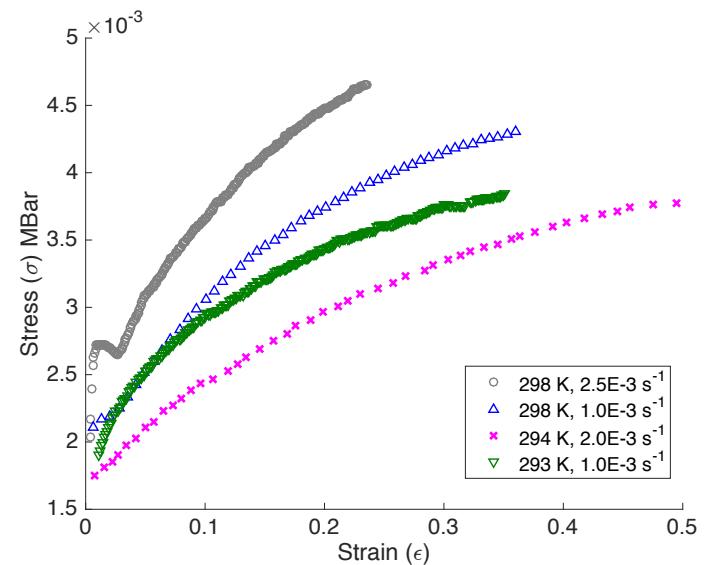
Model Term	Physical Interpretation
$\sigma$	Stress
$\varepsilon_p$	Strain
$\dot{\varepsilon}_p$	Strain Rate
$\dot{\varepsilon}_{p0}$	Reference Strain Rate ( $1.0 \text{ s}^{-1}$ )
$T$	Temperature
$T_m$	Melting Temperature (3250 K)
$T_0$	Room Temperature (290 K)
$A, B, C, m, n$	Model Parameters ( $m = 1$ )

- Statistical model

$$\sigma_j = (A + B\varepsilon_p^n) (1 + C \ln(\dot{\varepsilon}_p^*)) \left(1 - T^{*^m}\right) + \nu_j$$

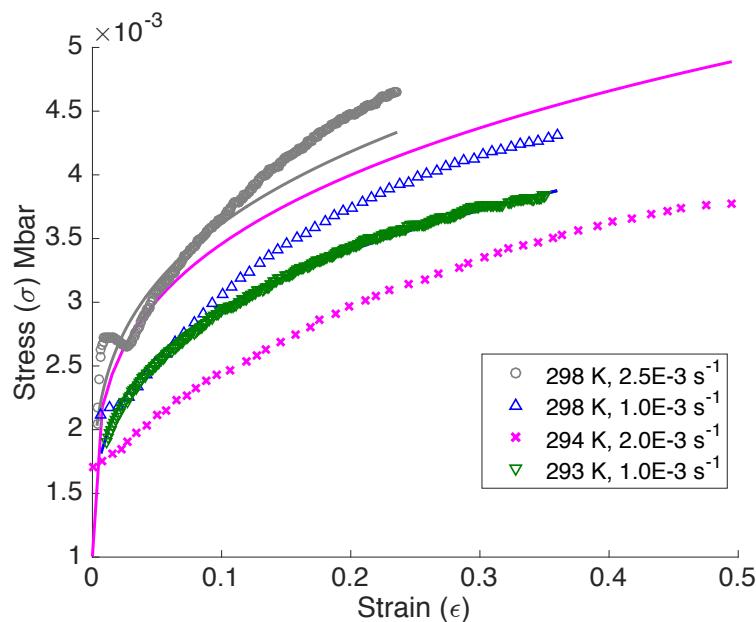
$$\nu_j \stackrel{iid}{\sim} N(0, \eta^2)$$

- Note that the same experimental conditions (temperature and strain rate) should produce the same response

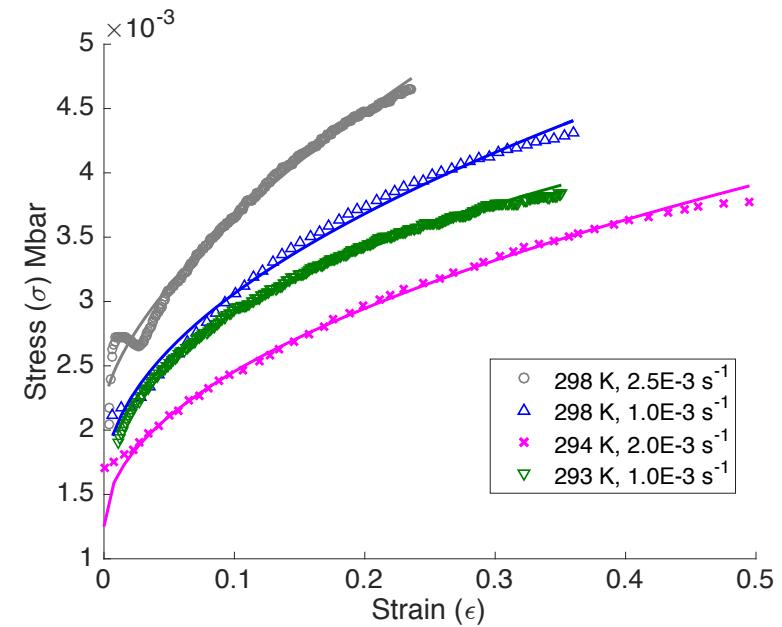


# Materials Strength Model: Fixed vs Mixed-Effects

Traditional Johnson-Cook



Mixed-Effects Johnson-Cook



$$\sigma_j = (\textcolor{teal}{A} + \textcolor{blue}{B}\varepsilon_p^{\textcolor{teal}{n}}) (1 + \textcolor{teal}{C} \ln(\dot{\varepsilon}_p^*)) (1 - T^*) + \nu_j$$

$$\nu_j \stackrel{iid}{\sim} N(0, \eta^2)$$

$$\sigma_{ij} = [(\textcolor{teal}{A} + \textcolor{red}{r}_{1i}) + (\textcolor{teal}{B} + \textcolor{red}{r}_{2i})\varepsilon_p^{(\textcolor{teal}{n}+\textcolor{red}{r}_{4i})}]$$

$$[1 + (\textcolor{teal}{C} + \textcolor{red}{r}_{3i}) \ln(\dot{\varepsilon}_p^*)] (1 - T^*) + \nu_{ij}$$

$$\textcolor{red}{r}_i \sim N(0, \Psi) , \quad \nu_{ij} \stackrel{iid}{\sim} N(0, \eta^2)$$



# Mixed-Effects Model Selection

- Traditional method: Information criteria
  - Score to assess quality of model (lower score = better model)
  - AIC or BIC

$$AIC = -2 \ln(L(\hat{\theta})) + 2p$$

$$BIC = -2 \ln(L(\hat{\theta})) + p \ln(N)$$

$L(\hat{\theta})$	Likelihood evaluated at optimal parameter estimate $\hat{\theta}$	Encourages goodness of fit
$p$	Number of parameters	Discourages overfitting
$N$	Number of data points	

- Prohibitive for models with a large number of parameters
  - Must consider  $2^{\{\# \text{ Fixed Effects} + \# \text{ Random Effects}\}}$  models

- Our method
  - Reduce the number of models/parameter sets (parameter subset selection)
  - Evaluate remaining models using information criteria

# Mixed-Effects Parameter Subset Selection (PSS) Algorithm

---

- Based on PSS Algorithm for purely fixed-effects models (Cintrón-Arias, et al., 2009)
- Our mixed-effects PSS algorithm
  - K. Schmidt and R. C. Smith, “A Parameter Subset Selection Algorithm for Mixed-Effects Models,” *International Journal for Uncertainty Quantification*, 6(5), 2016.
- We have made improvements to the published algorithm
  - Take into account how parameters work together

**Goal:** Reduce number of models to assess via information criteria



Lawrence Livermore National Laboratory

LLNL-PRES-754903



# Mixed-Effects Parameter Subset Selection (PSS) Algorithm

- 1) Construct a covariance matrix estimate.

$$Cov = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_p^2 \end{bmatrix}_{p \times p} \rightarrow SE_m = \sqrt{Cov(m, m)}, m = 1, 2, \dots, p$$

Square roots of diagonals are standard errors.

Smallest parameter uncertainty

- 2) For each subset of size  $n_p = 1, 2, \dots, p$ , choose the “best” parameter set

- Rate each of the  ${}_p C_{n_p}$  possible parameter sets (for each data set  $i$ )

**Assign selection scores  $\alpha$**

For each parameter set

$$\theta_k = [\theta_{1k}, \dots, \theta_{n_p k}]$$



$$\alpha_{k\textcolor{red}{1}} = |SE_{1k}/\hat{\theta}_{1k\textcolor{red}{1}}, \dots, SE_{n_p k}/\hat{\theta}_{n_p k\textcolor{red}{1}}|$$

⋮

$$\alpha_{k\#\textcolor{red}{Groups}} = |SE_{1k}/\hat{\theta}_{1k\#\textcolor{red}{Groups}}, \dots, SE_{n_p k}/\hat{\theta}_{n_p k\#\textcolor{red}{Groups}}|$$

For each data set  $i$ , find the smallest  $\alpha \Rightarrow$  Best parameter subset for the  $i$ th group



Lawrence Livermore National Laboratory

LLNL-PRES-754903



10

# Mixed-Effects PSS Algorithm (cont.)

- b) Determine the best parameter set over all data sets.

Assign the  $k$ th parameter set in the  $i$ th data set the selection index  $\gamma_{k_i}$

- Best parameter set  $\gamma_{k_i} = 1$
- Second best parameter set  $\gamma_{k_i} = 2$
- 
- 
- 
- Worst parameter set  $\gamma_{k_i} = p$

**Example:** Choose subset of size 2 from 4 parameters, 5 data sets

Ranks for parameter subsets of size 2

5 groups	4	3	2	1	5	6
	3	5	2	1	4	6
	3	4	1	2	6	5
	3	4	2	1	6	5
	3	5	2	1	4	6

- c) Calculate selection index sums  $\Gamma_k = \sum_i \gamma_{k_i}$

Smallest sum corresponds to the best set of  $n_p$  parameters over all data sets



Lawrence Livermore National Laboratory

LLNL-PRES-754903



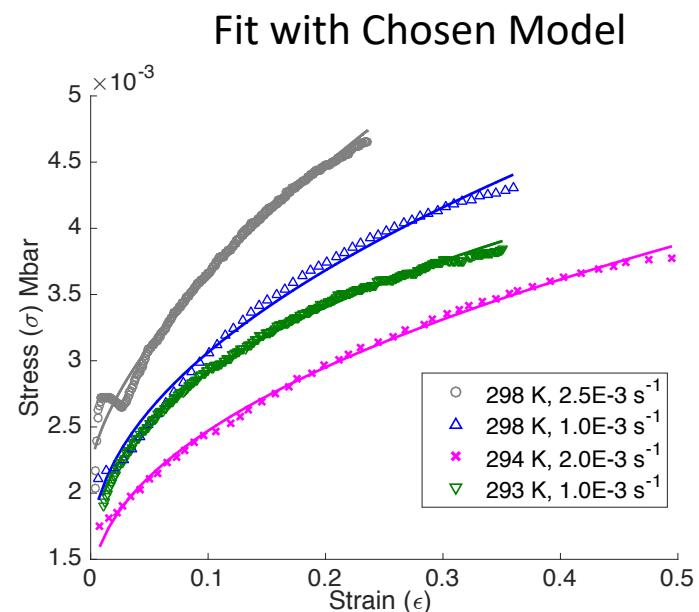
# Results

Size of Subset	Parameters								Information Criteria	
	A	B	C	n	r <sub>1</sub>	r <sub>2</sub>	r <sub>3</sub>	r <sub>4</sub>	AIC	BIC
1			X						-7815	-7810
2	X		X						-8240	-8231
3	X	X	X						-8238	-8225
4	X	X	X	X					-8236	-8218
5	X	X	X	X			X		-10738	-10741
6	X	X	X	X			X	X	-11216	-11220
7	X	X	X	X	X	X	X		-10734	-10739
“Full” Model	8	X	X	X	X	X	X	X	-11212	-11217

- Chosen Model:  $\sigma_{ij} = \left( A + B\varepsilon_p^{(n+r_{4i})} \right) [1 + (C + r_{3i}) \ln(\dot{\varepsilon}_p^*)] (1 - T^*) + \nu_{ij}$
- $$r_i \sim N(0, \Psi), \quad \nu_{ij} \stackrel{iid}{\sim} N(0, \eta^2)$$

# Conclusions and Future Work

- Mixed-effects modeling accounts for material variability
- Mixed-effects PSS paired with information criteria chose a reasonable parsimonious model to represent stress-strain data in the presence of material variability
- Future Work
  - Mixed-effects modeling for a range of experimental conditions
  - Different strength models



**Acknowledgements:** Thanks for Jeffrey Florando, Nathan Barton, and David Rivera for materials science insight and ongoing discussion.

**Disclaimer:** This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.



Lawrence Livermore National Laboratory

LLNL-PRES-754903



# Contact Information

---

Kathleen Schmidt  
Lawrence Livermore National Laboratory  
Applied Statistics Group  
E-mail: schmidt41@llnl.gov



Lawrence Livermore National Laboratory

LLNL-PRES-754903





**Lawrence Livermore  
National Laboratory**

# References

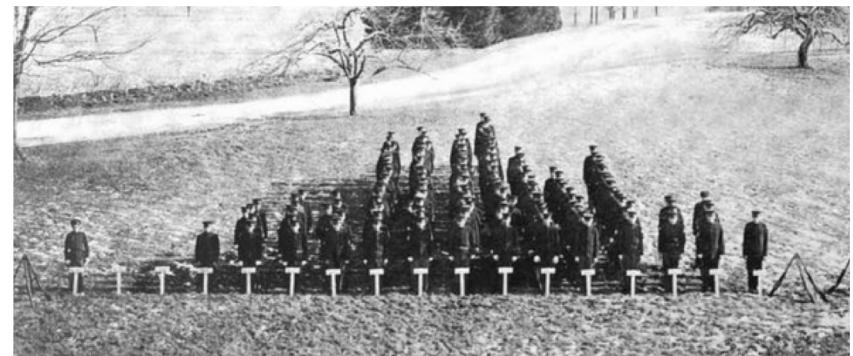
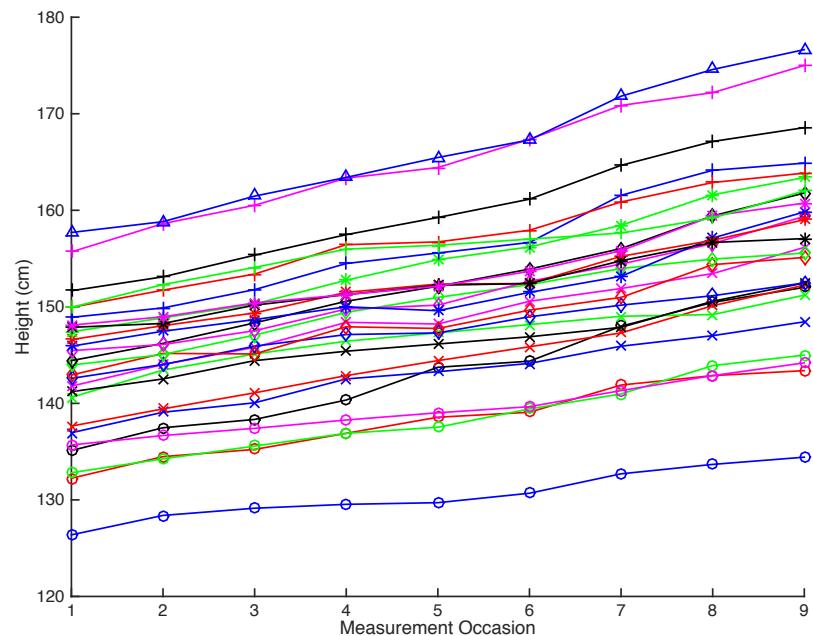
---

- A. F. Blakeslee, “Corn and men: The interacting influence of heredity and environment,” *Journal of Heredity*, 5(11), pp.511-518, 1914.
- A. Cintrón-Arias, H. T. Banks, A. Capaldi, and A. L. Lloyd. “A sensitivity matrix based methodology for inverse problem formulation,” *Journal of Inverse and Ill-Posed Problems*, 17(6), pp.545-564, 2009.
- L. E. Murr, M. A. Meyers, C. S. Niou, Y. J. Chen, S. Pappu, and C. Kennedy, “Shock-induced deformation twinning in tantalum,” *Acta materialia*, 45(1), pp. 157-175, 1997.
- W. H. Gourdin, D. H. Lassila, M. M. Leblanc, and A. L. Shields, “The influence of tungsten alloying on the mechanical properties of tantalum,” *Journal De Physique IV*, 4 (C8), pp. 207-212, 1994.
- G. R. Johnson and T. J. Holmquist, “Test data and computational strength and fracture model constants for 23 materials subjected to large strains, high strain rates, and high temperatures,” Los Alamos National Laboratory Report, LA-11463-MS, 1989.
- M. T. Perez-Prado, J. A. Hines, and K. S. Vecchio, “Microstructural evolution in adiabatic shear bands in Ta and Ta–W alloys,” *Acta materialia*, 49 (15), pp. 2905-2917, 2001.
- NDT Resource Center, “Introduction to Materials and Processes,” *NDT Resource Center*, 2008.  
url: [https://www.nde-ed.org/EducationResources/CommunityCollege/Materials/cc\\_mat\\_index.htm](https://www.nde-ed.org/EducationResources/CommunityCollege/Materials/cc_mat_index.htm)
- K. Schmidt and R. C. Smith, “A Parameter Subset Selection Algorithm for Mixed-Effects Models,” *International Journal for Uncertainty Quantification*, 6(5), 2016.
- L. Wu, *Mixed Effects Models for Complex Data*, CRC Press, Boca Raton, FL, 2010.



# Example of Mixed-Effects Model

- How do we model the heights of 26 boys measured over time?
  - Data from Wu (2010).



Distribution of student heights at Connecticut Agricultural College. Image from Blakeslee (1914).

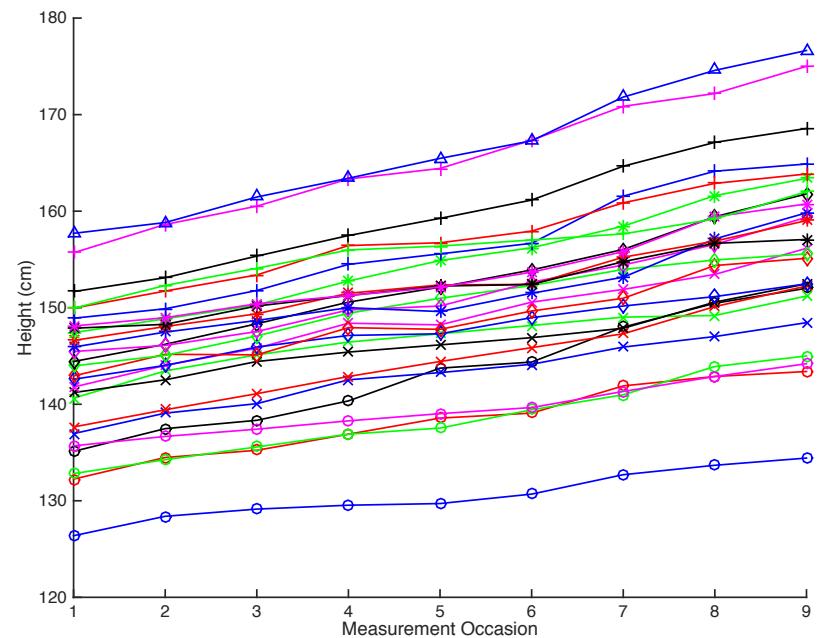
- Variation in starting heights and (lesser) variation in slope
- Population variation in height is typically normally distributed

# Example of Mixed-Effects Model

- Which model should we use?

$$y_{ij} = (\theta_1 + r_{1i})x_{ij} + (\theta_2 + r_{2i}) + \varepsilon_{ij}$$

$$r_i \sim N(0, \Psi) , \quad \varepsilon_i \sim N(0, \sigma^2 I)$$



# Example of Mixed-Effects Model

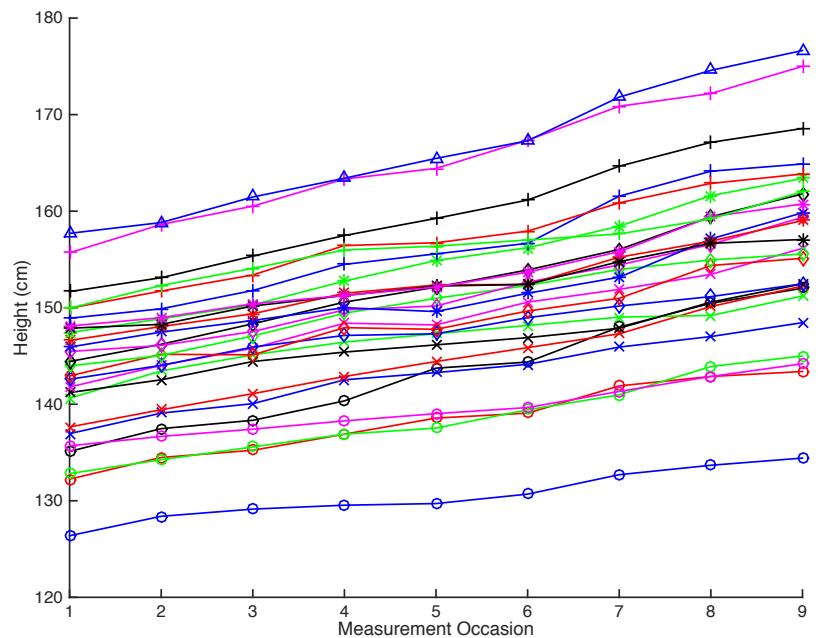
- Which model should we use?

$$y_{ij} = (\theta_1 + r_{1i})x_{ij} + (\theta_2 + r_{2i}) + \varepsilon_{ij}$$

$$r_i \sim N(0, \Psi) , \varepsilon_i \sim N(0, \sigma^2 I)$$

$$y_{ij} = \theta_1 x_{ij} + (\theta_2 + r_{2i}) + \varepsilon_{ij}$$

$$r_1 \sim N(0, \psi) , \varepsilon_i \sim N(0, \sigma^2 I)$$



# Sensitivity Analysis for Mixed-Effects Models

---

- Traditional sensitivity analysis techniques are ineffective
- Consider the model

$$y_{ij} = (\theta_1 + r_{1i})x_{ij} + (\theta_2 + r_{2i}) + \varepsilon_{ij}$$

$$r_i \sim N(0, \Psi) , \quad \varepsilon_i \sim N(0, \sigma^2 I).$$

- Note that

$$\frac{\partial y_{ij}}{\partial \theta_1} = \frac{\partial y_{ij}}{\partial r_{1i}} \text{ and } \frac{\partial y_{ij}}{\partial \theta_2} = \frac{\partial y_{ij}}{\partial r_{2i}}$$

- Fixed and random effects have same sensitivity measure



# Mixed-Effects Parameter Subset Selection (PSS) Algorithm

---

- 1) Check for identifiability issues with the purely fixed-effects model
  - e.g., Using Pearson correlation coefficients, pairwise scatter plots, or Bayesian joint parameter densities
  - Fix parameters as necessary
- 2) For  $n_p = 1, 2, \dots, p$ 
  - a) Construct an estimate  $Cov = \hat{s}^2 (\chi^T \chi)^{\dagger}$  of the covariance matrix containing the variances and correlations of the fixed and random effects
    - Pseudoinverse may be necessary
  - b) Determine the standard errors  $SE_m = \sqrt{Cov(m, m)} , m = 1, 2, \dots, p$
  - c) Calculate the selection scores for all  $i$  data sets.

There will be  ${}_p C_{n_p}$  possible parameter sets.

For the  $k$ th parameter set in the  $i$ th data set,  $\alpha_{k_i} = |\alpha_{1i}, \alpha_{2i}, \dots, \alpha_{({}_p C_{n_p})i}|$ ,

where  $\alpha_{\ell_i} = |SE_{m_\ell}/\hat{\theta}_{m_\ell i}| , \ell = 1, 2, \dots, {}_p C_{n_p}$ .

The smallest selection score for the  $i$ th data set corresponds to the “best” set of  $n_p$  parameters (with the lowest estimation uncertainty) for that data set.



# Mixed-Effects PSS Algorithm (cont.)

---

d) Determine the best parameter set over all data sets.

Assign the  $k$ th parameter set in the  $i$ th data set the selection index  $\gamma_{k_i}$

- Best parameter set  $\gamma_{k_i} = 1$
- Second best parameter set  $\gamma_{k_i} = 2$
- 
- 
- Worst parameter set  $\gamma_{k_i} = p$

e) Calculate selection index sums  $\Gamma_k = \sum_i \gamma_{k_i}$

- The smallest selection index sum corresponds to the best set of  $n_p$  parameters over all data sets

