NC STATE UNIVERSITY

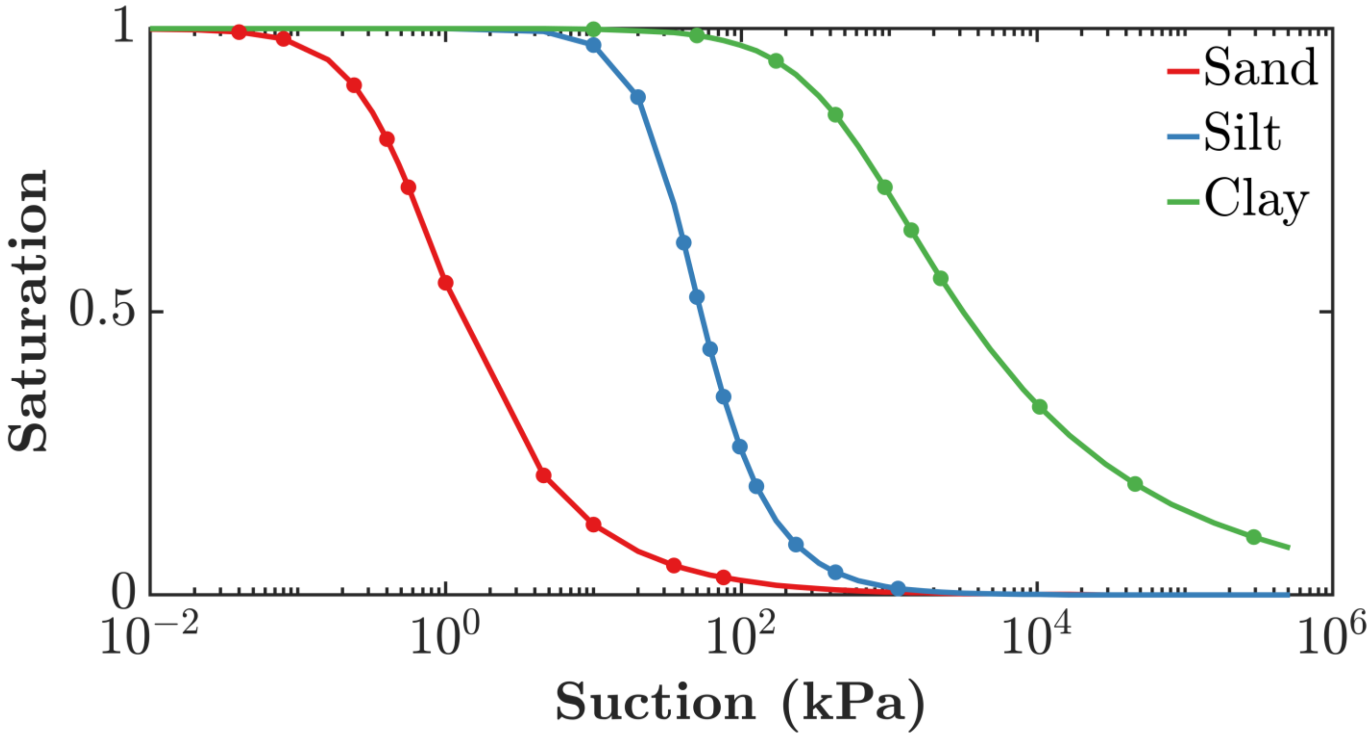
Bayesian Estimation of Soil-Water Characteristic Curves

Authors:
David Barba
Jungmin Lee
Sophia Andanje

ST540: Final Project
Presentation: 5/1/2024

What is SWCC?

- The Soil Water Characteristic Curve (SWCC) is an important prerequisite for studying the mechanical properties of unsaturated soil.



Why is Bayesian Analysis used for SWCC?

- The issues are:
  - The experimental measurement for the SWCC is time-consuming
  - Empirical methods suggested to estimate the SWCC have substantial uncertainties

- Suggest a Bayesian approach to estimate the SWCC from a limited number of tests:
  - take advantage of prior information: database of SWCC
  - define prior distribution of SWCC parameters based on soil properties from quick and low-cost lab tests (e.g. soil texture, void ratio, and effective grain size)

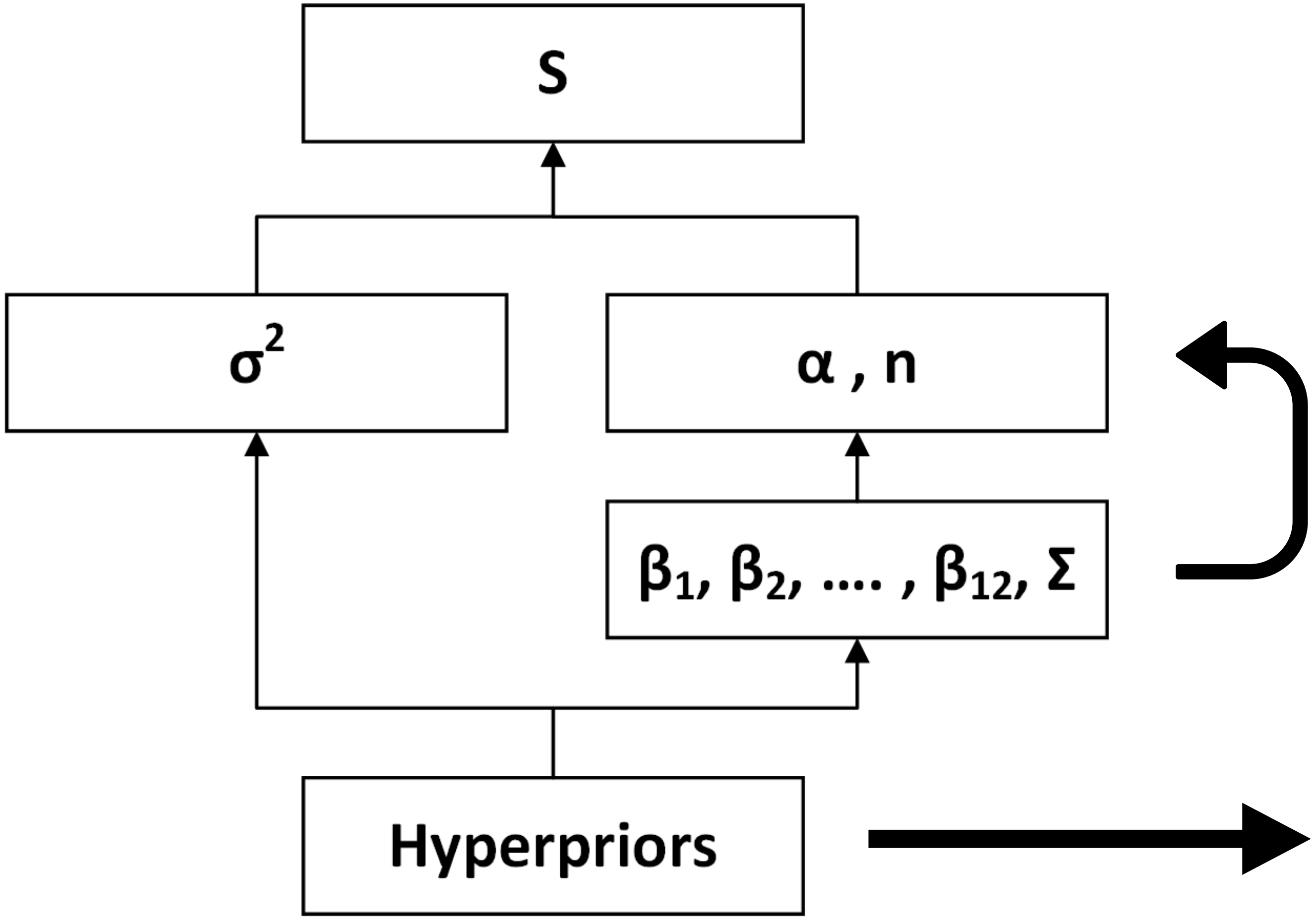
Data and Data Analysis

- Data Source: SoilVision Database
- SWCC model: van Genuchten (1980)

$$g_i = g(\psi_i|\alpha, n) = \frac{1}{[1 + (\alpha\psi)^n]^{1-\frac{1}{n}}}$$

- Parameter for the model:
  - α: the fitting parameter related to the air-entry value of soil (kPa)
  - n: the fitting parameter related to the slope of the SWCC
- Covariates:
  - Soil texture: percentage of clay (C_l) and coarse grain (C_g)
  - Void ratio (e)
  - Effective grain size (D_{10})

How to Establish our Model for SWCC?



$S_i \sim N(g_i, \sigma^2)$ for homoskedastic

$S_i \sim N\left(g_i, \sigma^2 g_i (1 - g_i)\right)$ for heteroskedastic

$$g_i = g(\psi_i|\alpha, n) = \frac{1}{[1 + (\alpha\psi)^n]^{1-\frac{1}{n}}}$$

(Zhang et al, 2021)

Multi Linear Regression (Database)

$$(\ln\alpha, \ln n)^T \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\mu_{\ln\alpha} = \beta_1 + \beta_2 X_1 + \beta_3 X_2 + \beta_4 X_3 + \beta_5 X_4$$

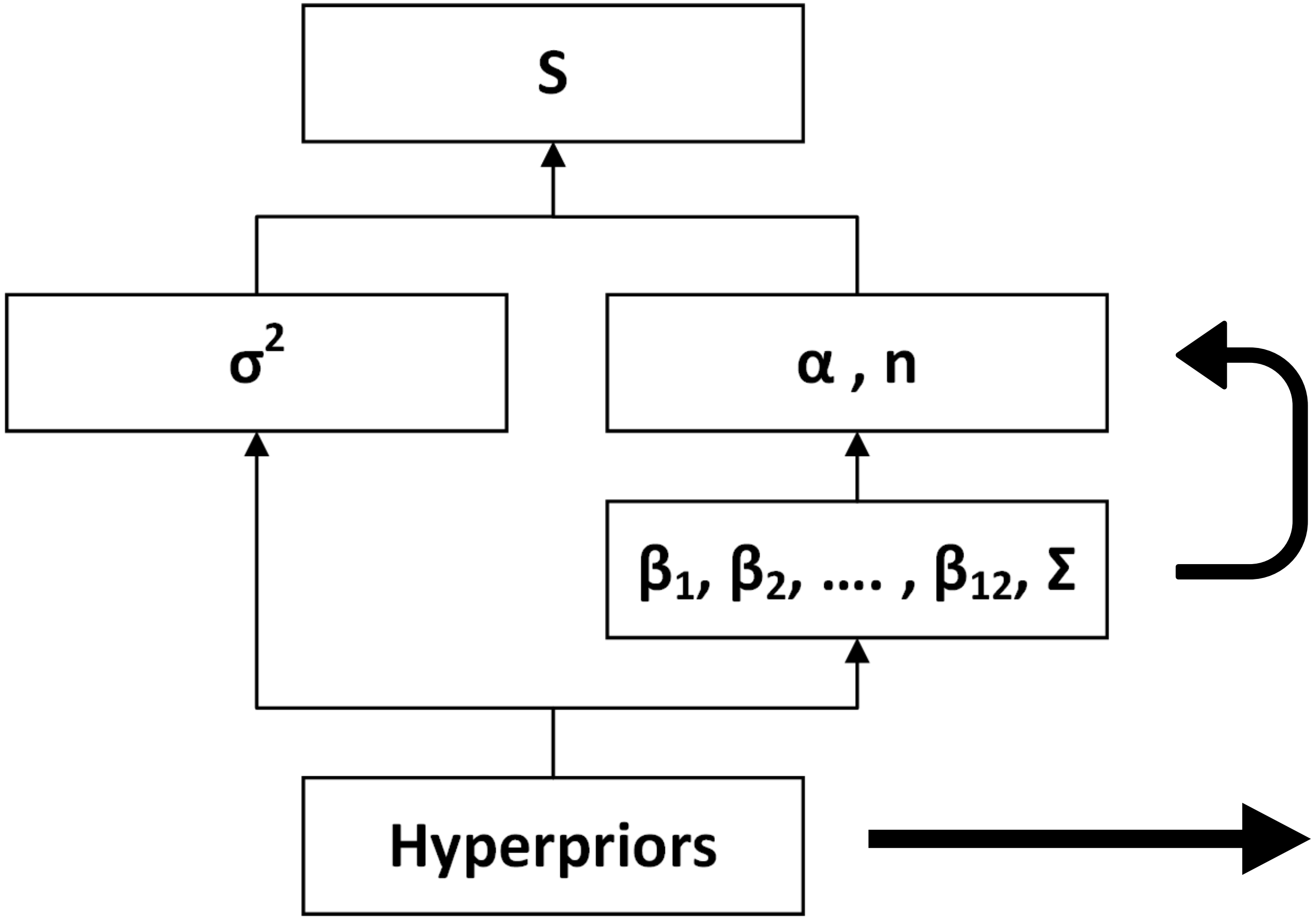
$$\mu_{\ln n} = \beta_6 + \beta_7 X_5 + \beta_8 X_6 + \beta_9 X_7 + \beta_{10} X_8$$

$\beta_{ij} \sim N(0, 0.001)$

$\boldsymbol{\Sigma} \sim InvWishart(2.1, \mathbf{I_2}/2.1)$

$\sigma^2 \sim InvGamma(2.366, 0.012)$ informative

How to Establish our Model for SWCC?



$S_i \sim N(g_i, \sigma^2)$ for homoskedastic

$S_i \sim N\left(g_i, \sigma^2 g_i(1-g_i)\right)$ for heteroskedastic

$$g_i = g(\psi_i|\alpha, n) = \frac{1}{[1+(\alpha\psi)^n]^{1-\frac{1}{n}}}$$

(Zhang et al, 2021)

Multi Linear Regression (Database)

$$(\ln\alpha, \ln n)^T \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\mu_{\ln\alpha} = \beta_1 + \beta_2 X_1 + \beta_3 X_2 + \beta_4 X_3 + \beta_5 X_4$$

$$\mu_{\ln n} = \beta_6 + \beta_7 X_5 + \beta_8 X_6 + \beta_9 X_7 + \beta_{10} X_8$$

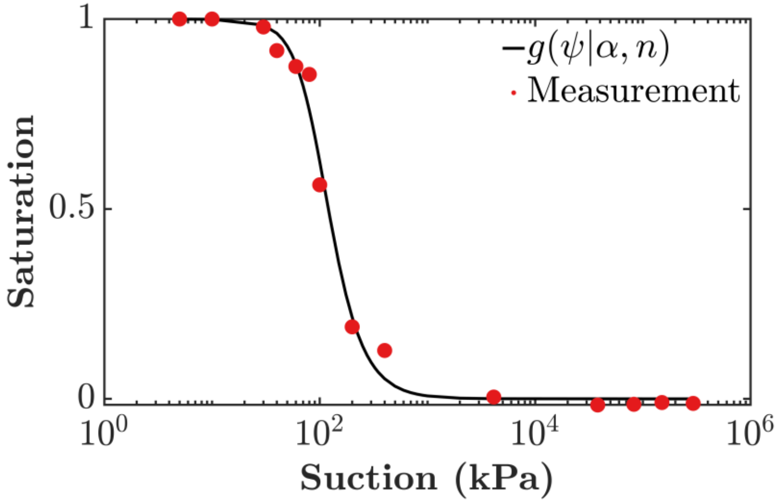
$\beta_{ij} \sim N(0, 0.001)$

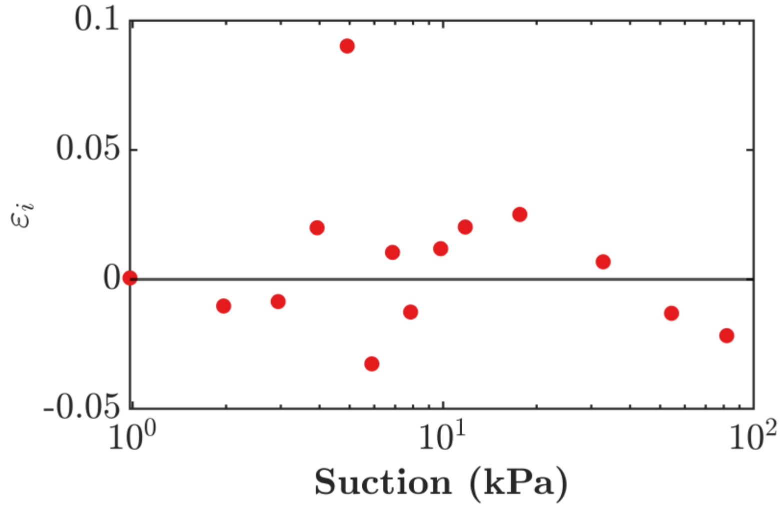
$\boldsymbol{\Sigma} \sim InvWishart(2.1, \mathbf{I_2}/2.1)$

$\sigma^2 \sim InvGamma(2.366, 0.012)$ informative

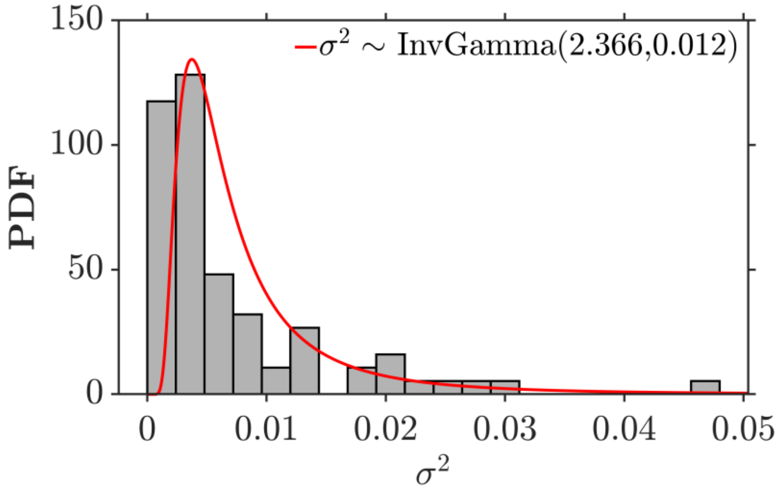
How to Establish our Model for SWCC?

- Heteroskedastic?





- Informative prior for σ^2?



Model Selection and Fitting

- Model selection:

Multi Linear Regression

Model	α	n	Min (ESS)	Dw	Penalty	DIC
A	C_i, C_g, e, D_{10}	C_i, C_g, e, D_{10}	26413.9	439.1	13.3	452.7
B	ln C_i, ln C_g, ln e, ln D_{10}	ln C_i, ln C_g, ln e, ln D_{10}	27290.1	436.7	13.34	450.1
C	C_i, C_g, e, D_{10}	ln C_i, ln C_g, ln e, ln D_{10}	27948.7	438.9	13.71	452.6

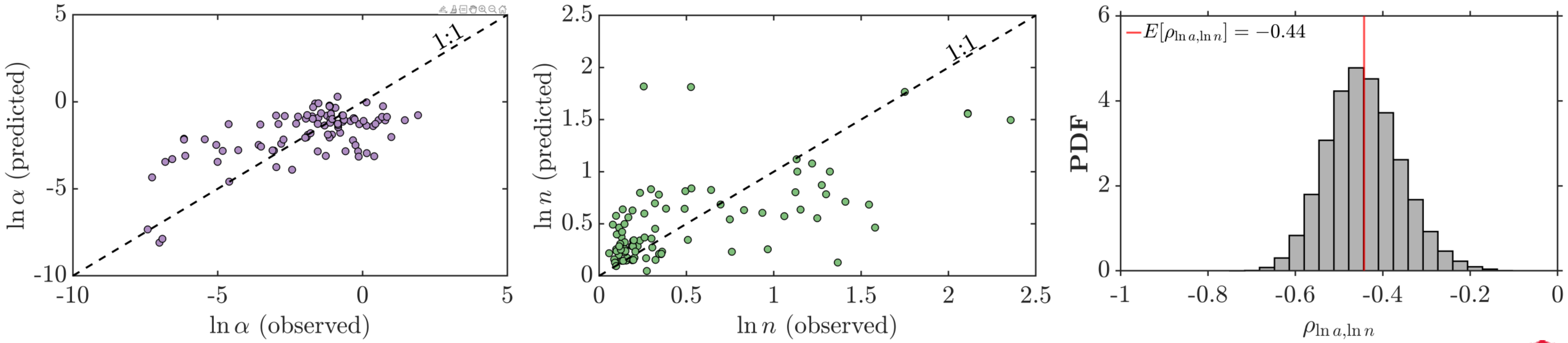
- Model fit:

Model	Min (α)	Min (n)	Max (α)	Max (n)	Range (α)	Range (n)
A	0.377	0	0.574	0.484	0.654	0.981
B	0.090	0	0.456	0.238	0.861	0.955
C	0.397	0	0.589	0.236	0.651	0.937

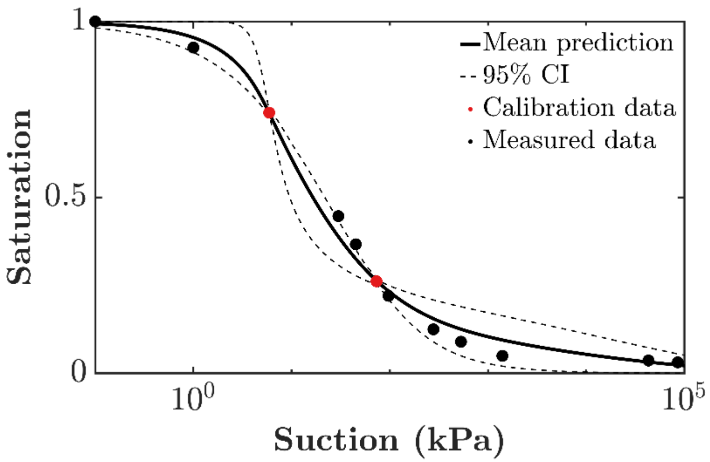
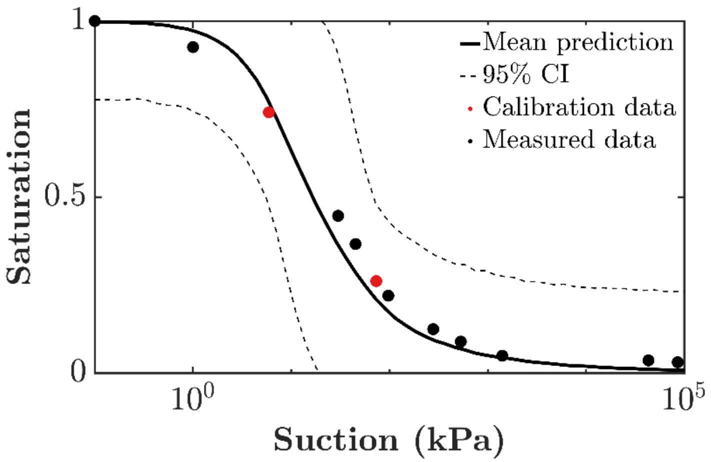
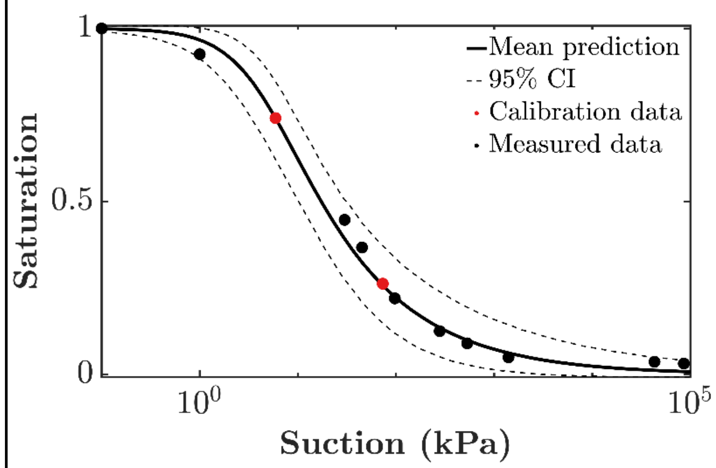
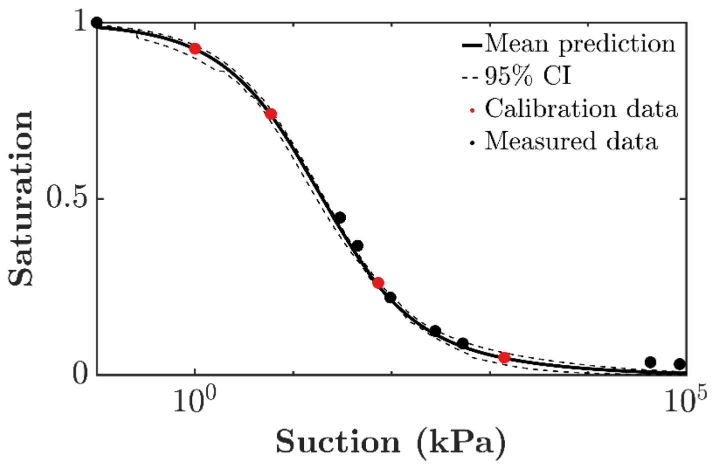
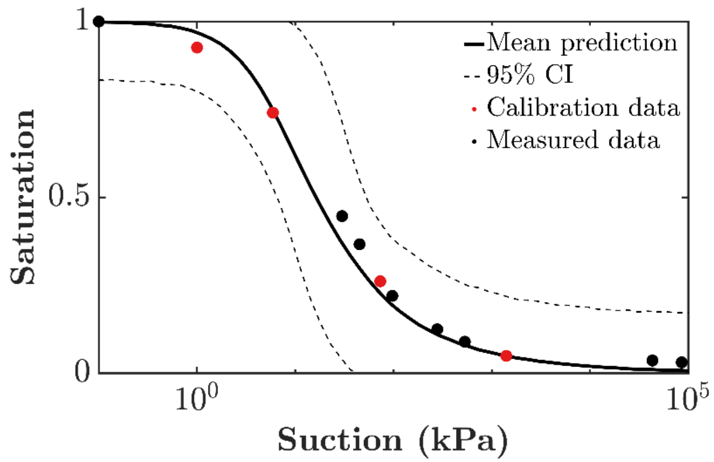
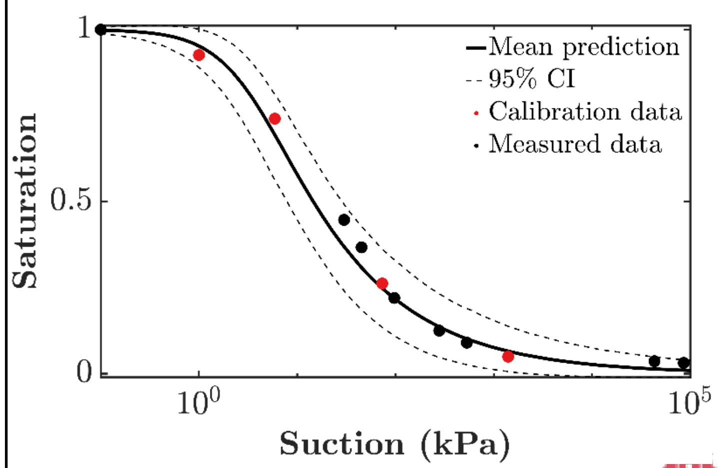
Cross Validation

- Five equally-sized groups / Predictions setting aside one group at a time and fitting the model with the observations of the remaining groups

Model	Parameter	Bias	MSE	MAD	COV	WIDTH
B	α	0.021	2.891	1.334	0.447	6.786
	n	-0.001	0.167	0.282	0.937	1.603



Results of Informative Model

<table>
<tr><th></th><th>Zhang et al. (2021)</th><th>Homoskedastic</th><th>Heteroskedastic</th></tr>
<tr><td>n = 2</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
<tr><td>n = 4</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
</table>

Conclusion

- Limitations and improvement of our model
  - The value of min = 0: the model did not fit the data sufficiently
    → add more test data (95 samples; more than 200 samples)
  - Using linear regression
    → use higher-order polynomial or a second-order polynomial regression
  - Limited covariates
    → use more covariates (currently use 4 covariates)

Reference

Rahimi, A., Rahardjo, H., Leong, E.-C. (2015). Effects of soil-water characteristic curve and relative permeability equations on estimation of unsaturated permeability function. *Soils and Foundations.* **55(6)**: 1400-1411.

Reich, B. J., Ghosh, S. K. (2019). Bayesian Statistical Methods.

van Genuchten, M.T. (1980). Closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil science society of America Journal*. **44**: 892 – 898.

Zhang, J., Yang, S., Zhang, L. L., Zhou, M. L. (2021). Bayesian estimation of soil-water characteristic curve. *Canadian Geotechnical Journal*. **59**: 569-582.

Thank you

ST540: Geotechnical Engineers