Attachment Efficiency Analysis

Jun Chen and Tory Norton

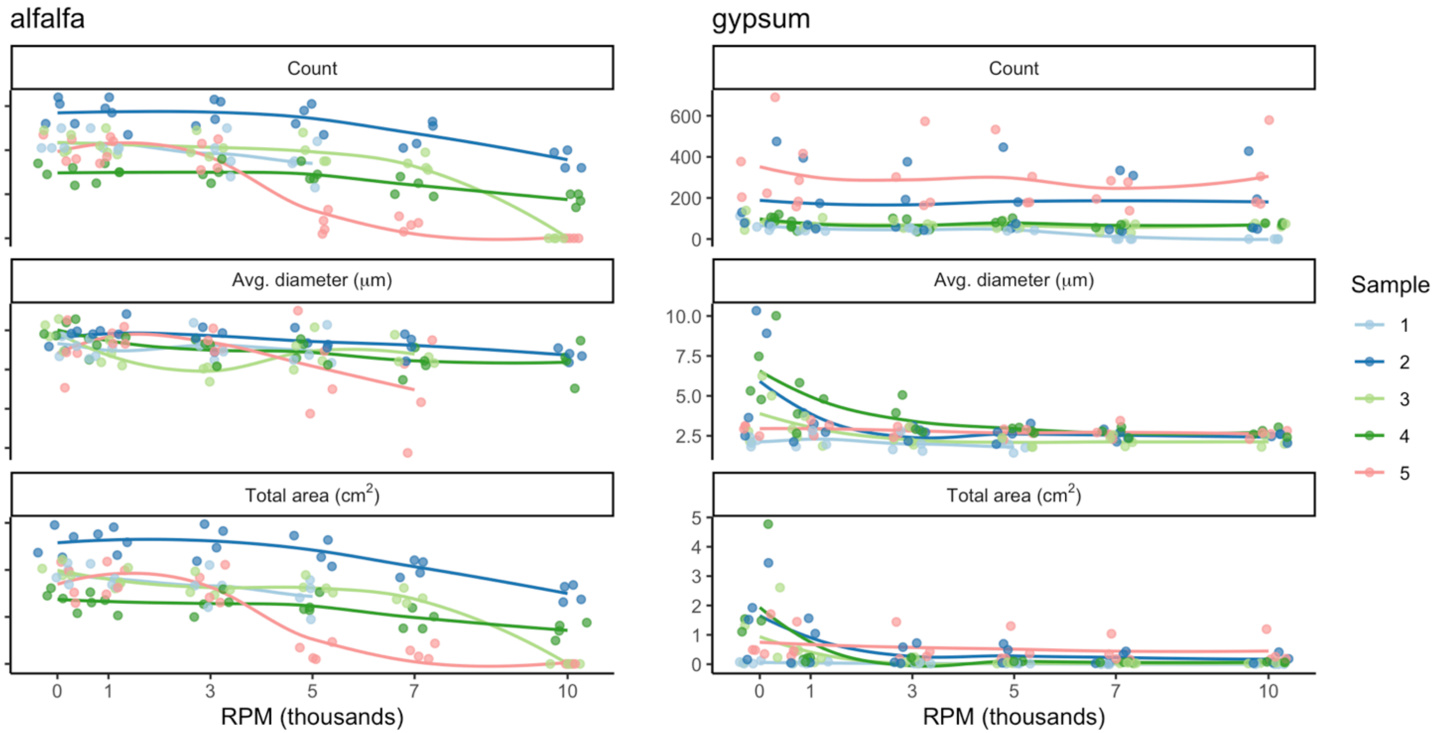
Background: The Soiling Problem

- Dust particles adhere to solar panels, reducing energy efficiency (*"soiling"*)
- To reduce soiling, we must understand the **adhesive properties** of dust particles
- Dust composition varies by region
- Understanding adhesion for different materials can:
  - Help identify region-specific challenges
  - Inform strategies for mitigation

Experimental Design

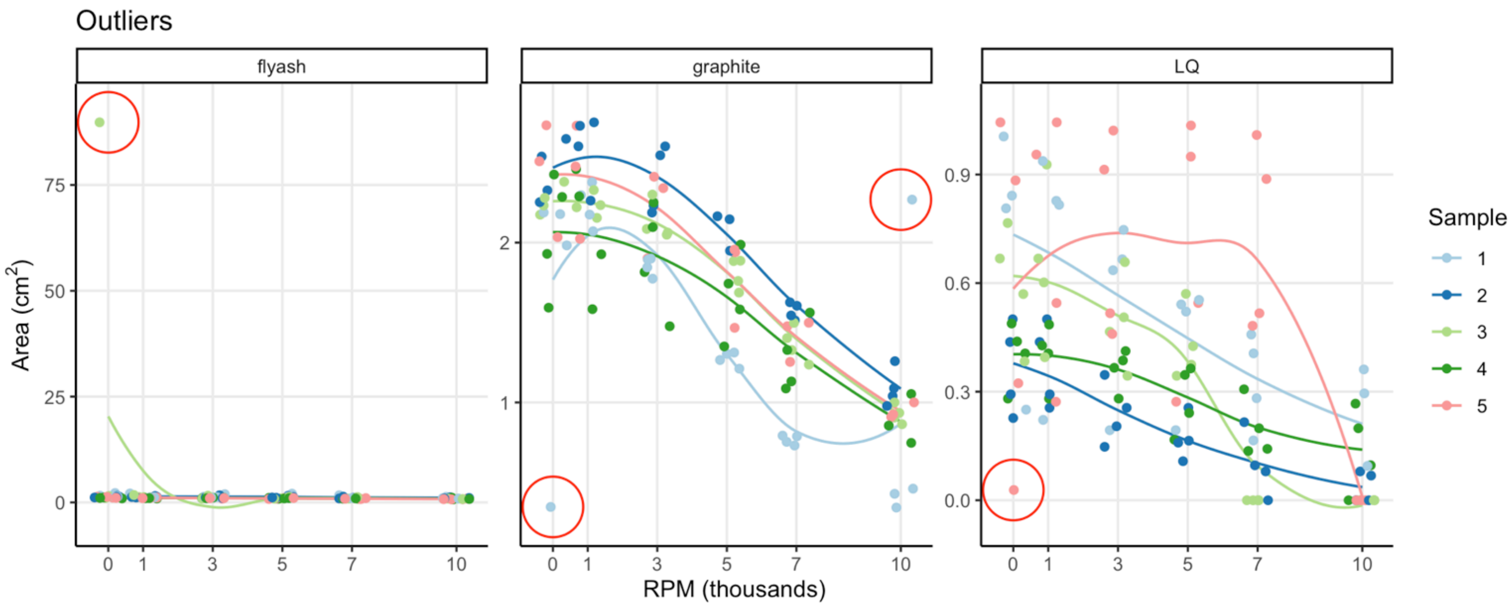
- **Materials:** alfalfa, fly ash, solosphere, silicon dioxide, graphite, gypsum
- Five glass slides were soiled with each material
- Each slide was imaged at four sites; particles were binned according to size and counted
- Slides were centrifuged at:
  - 1000, 3000, 5000, 7000, and 10000 rpm
- After each round of centrifuging, slides were re-imaged at the same four sites

Response: Particle area



- **Soiling metric:** Soiling reduces panel output by obscuring surface area, so we use total particle area as a natural metric.
- Area captures both particle breakup (e.g., gypsum: shrinking diameters) and particle removal (e.g., alfalfa: declining counts).

Data Cleaning



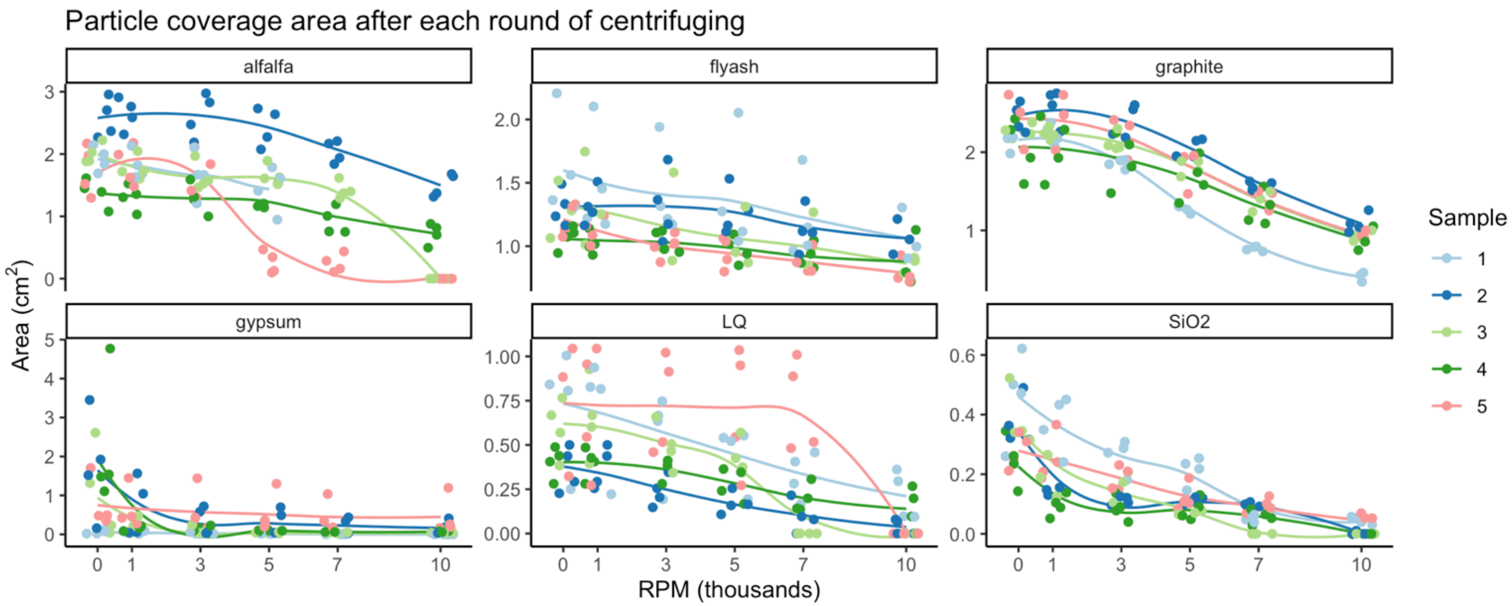
- **Outlier removal:** Four data points (from three sites) were removed due to clearly erroneous values (e.g., all pre-centrifuge particle counts lower than counts after 10000 rpm).
- **Scope of removal:** These sites show reasonable behavior at other rotation speeds, so only the affected speeds were removed.

Physical Properties

Material	Density	Work of Adhesion	Hamaker
solosphere	2200	0.0151	9.09e-20
SiO2	1522	0.0108	6.5e-20
gypsum	1121	0.0165	9.97e-20
LQ	1000	0.0123	7.43e-20
alfalfa	1000	0.0123	7.43e-20
flyash	641	0.0166	1e-19
graphite	641	0.0166	1e-19

- **Motivation:** We examine how material physical properties (e.g., density, work of adhesion) influence soiling behavior, both to explain observed trends and to guide interpretation for untested materials.
- **Variable selection:** Because work of adhesion and Hamaker constant are collinear, we exclude the Hamaker constant from the analysis.
- **Effect type:** Exploratory analysis suggested an intercept based on work of adhesion and an interaction term for density and rotation speed.

Correlated Data



- **Motivation:** Measurements come from sites nested within samples, with repeated measurements taken at the same site. Data from the same site/sample are likely correlated.

- Visual inspection shows that measurements from the same slide tend to cluster at each rotation speed. This suggests that within-sample correlation is significant.

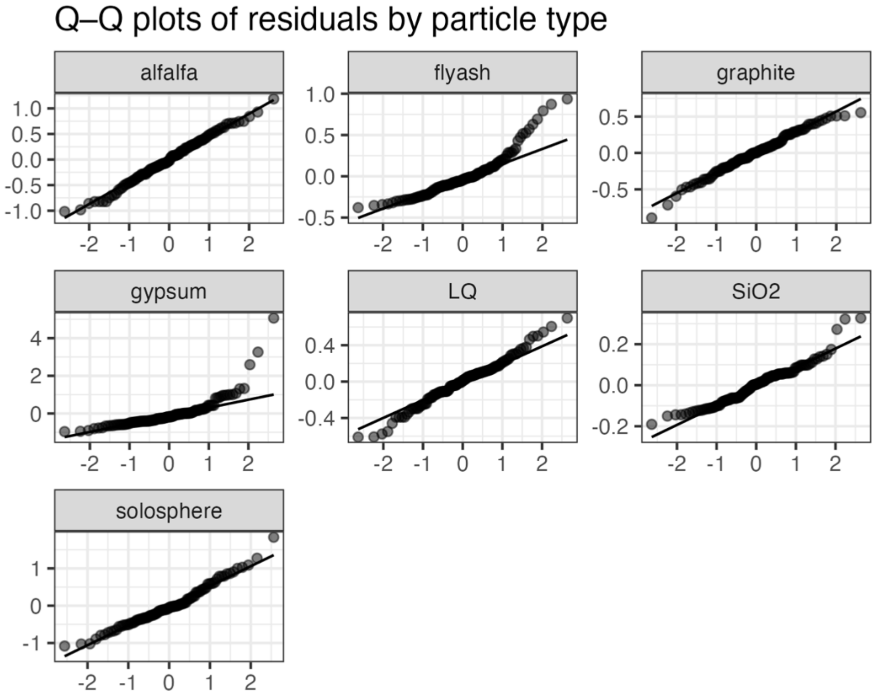
Frequentist Modeling - Model Specification

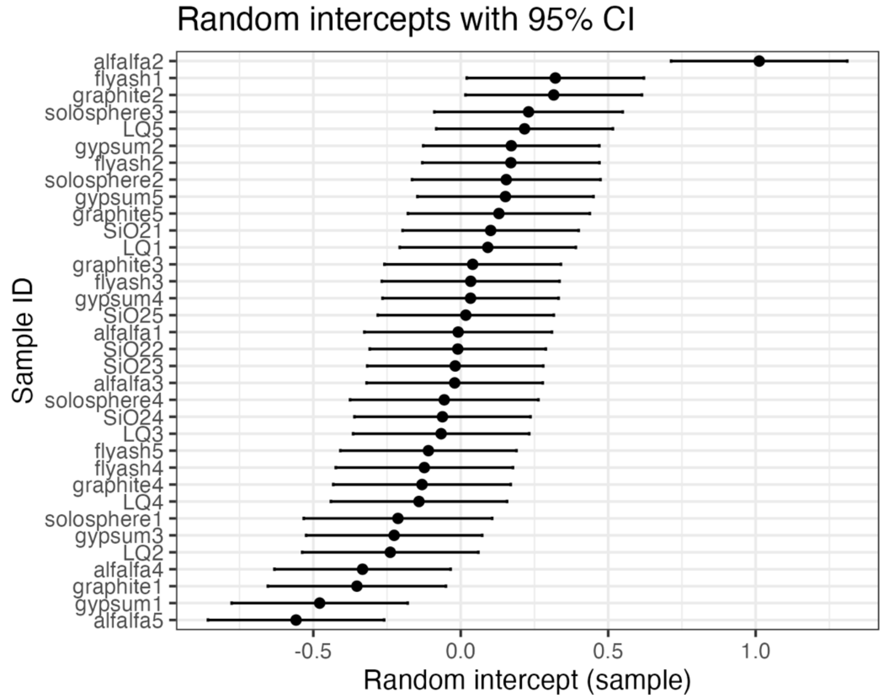
We fit a **linear mixed-effects** model (on centered and scaled area):

$$\text{total\_area} = \underbrace{\beta_1 \text{work}_{\text{type}} + \mu_{\text{type}} + \mu_{\text{samp}}}_{\text{intercept}} + \underbrace{\{\beta_2 + \gamma_{\text{type}}\}\,\text{rpm} + \beta_3\,\{\text{rpm} \times \text{density}_{\text{type}}\}}_{\text{slope}} + \epsilon$$

- **Fixed effects** capture:
  - Baseline slope
  - Effects of physical constants
- **Random slope:** Allows rate of change to vary by particle type.
- **Random intercept:** Accounts for correlation among measurements from the same slide.
  - A site-within-sample level intercept was considered, but model fit was similar to sample-level intercept. We concluded that sample-level effect may be sufficient to account for correlation.

Frequentist Modeling - Goodness-of-Fit

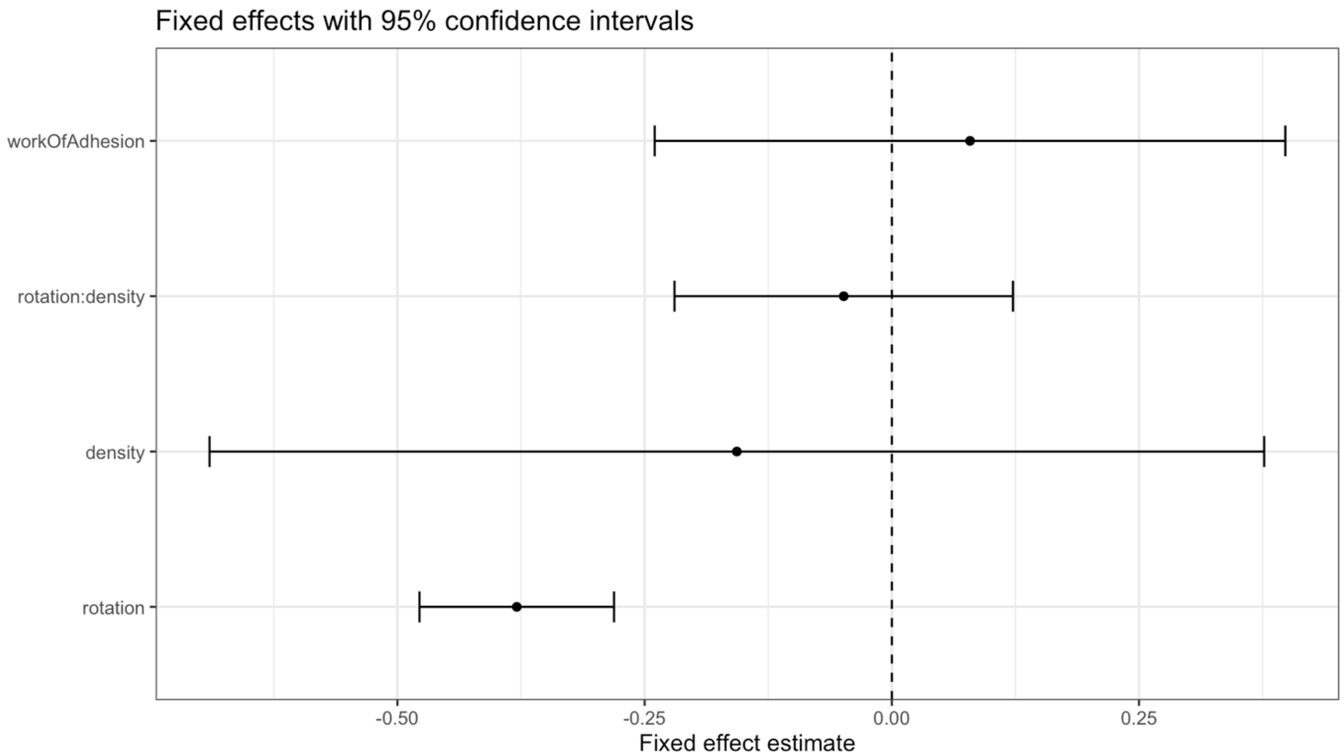




Residuals are approximately normally distributed within each particle type, with mild deviations in the tails.

Random intercepts by sample are relatively large, indicating that within-type variation is significant.

Frequentist Models - Results



- The 95% CIs for the effects of density and work of adhesion overlap zero, indicating that, after accounting for variability captured by the random effects, the data provide limited evidence that their population-average effects on adhesive behavior differ from zero.

Bayesian Hierarchical Model

Motivation:

1. The model selection is based on the AIC/BIC results among a collection of candidate LMMs
2. We are interested in how the total retained area depends on rotation, particle density and work of adhesion, and how particle type affects the slope of the rotation-area relationship
3. We adopt a Bayesian hierarchical framework to enable particle pooling across particle types and samples

Model:

$$y_{pij} = \beta_0 + \beta_1 \mathrm{WOA}_p + (\beta_2 + \gamma_p)\mathrm{rpm}_j + \beta_3(\mathrm{rpm}_j \times \mathrm{density}_p) + \mu_i + a_p + \epsilon_j$$

Here,

- y_{pij} is the total area in (particle, sample) cell, 797 observations
- WOA_p is the work of adhesion of particle type for type p
- rpm_j is the rotation speed (RPM) for each observation j
- $\mathrm{density}_p$ the density of particle type
- μ_i is the (type, sample)-level random intercept for (particle, sample) cell
- a_p is the type-level random intercept
- γ_p is the type-specific random slope representing the deviations from the mean rotation effect

Prior Specification & MCMC Initializations

Prior Assumptions

$$
\begin{aligned}
\beta_k &\sim \mathcal{N}(0, \sigma_\beta^2), && k = 0, 1, 2, 3, \\
\mu_i &\sim \mathcal{N}(0, \sigma_\mu^2), && i = 1, \ldots, S, \\
\gamma_p, a_p &\sim \mathcal{N}(0, \sigma_\gamma^2), && p = 1, \ldots, P, \\
\sigma_\ell &\sim \text{half-}t(0, 2.5), && \ell \in \{\beta, \mu, \gamma\}, \\
\varepsilon_{pij} &\sim \mathcal{N}(0, \sigma^2), \\
\sigma &\sim \text{half-}t(0, 2.5).
\end{aligned}
$$

- The half-t distributions provides weak-informative and heavy-tailed for data-driven purpose.

Initializations:

$$
\begin{aligned}
\beta_k^{(0)} &\sim \mathcal{N}(0, 1^2), && k = 0, 1, 2, 3, \\
\mu_i^{(0)} &\sim \mathcal{N}(0, 0.5^2), && i = 1, \ldots, S, \\
\gamma_p^{(0)}, a_p^{(0)} &\sim \mathcal{N}(0, 1^2), && p = 1, \ldots, P, \\
\sigma, \sigma_\mu, \sigma_\beta, \sigma_\gamma &\sim \text{Uniform}(0.2, 2).
\end{aligned}
$$

- We use over-dispersed but reasonable initial values for all parameter in our 4 MCMC chains for a better MCMC convergence diagnostics.

MCMC Sampling Process

Adaptive period: 500 iterations

Burn-in: 2,000 iterations

Sampling: 700,000 iterations each chain with thinning = 1

Rafery-Lewis Diagnostics for required sampling size: Among all fixed association parameters, the model requires at least 616 burn-in and 687,456 iterations/chain, indicating our MCMC setting is adequate for analysis

Effective Sample Size:

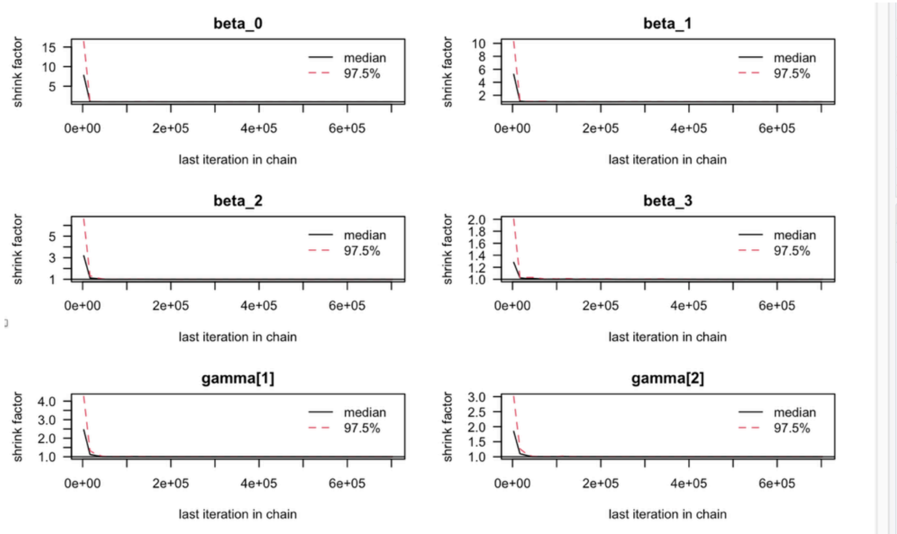
beta_0 (intercept)	beta_1 (WOA effect)	beta_2 (Mean Rotation Effect)	beta_3 (Rotation x Density Effect)
12877.23	8702.58	7572.06	13540.37

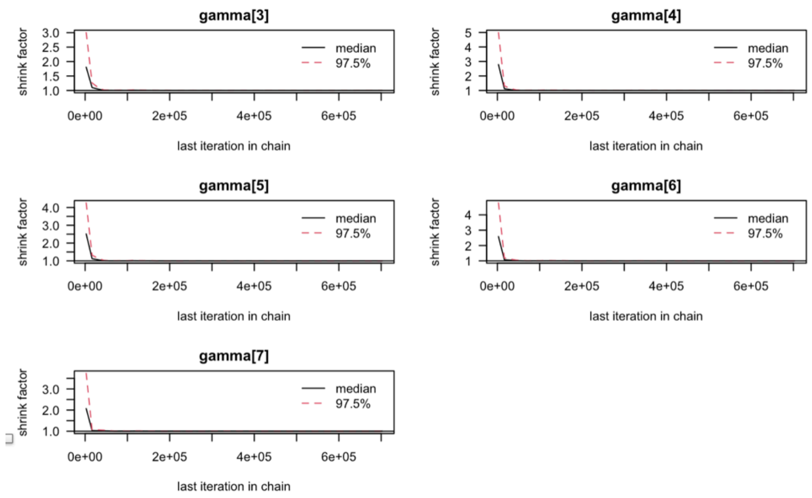
gamma_alfalfa	gamma_flyash	gamma_graphite	gamma_gypsum	gamma_LG	gamma_SiO2	ga
8491.48	10609.71	10493.57	8010.32	8385.73	8698.95	11

Convergence Analysis

Gelman-Rubin Factor: Gelam-Rubin $\hat{R} = 1$ for all fixed effects and random slopes.

Gelman-Rubin Plots:

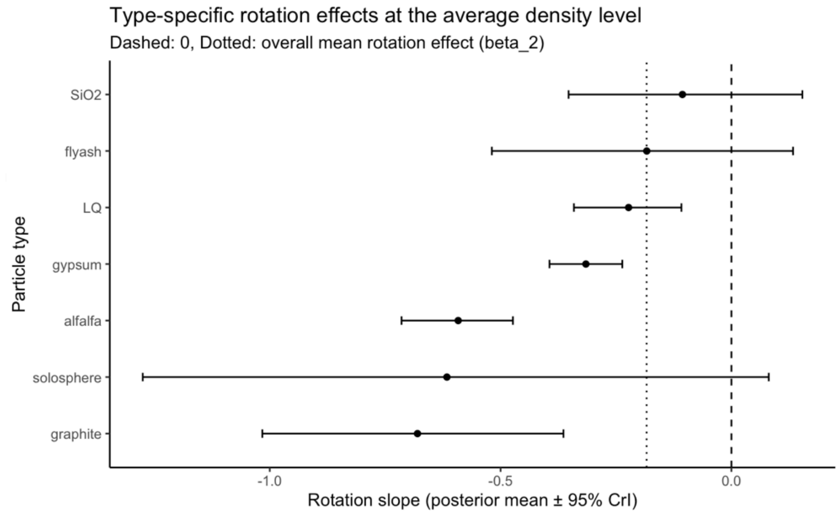




- All shrink factors shrink to one quickly.

Inference (Posterior Summary)

Parameter	Mean	2.5%	97.5%	P(>0)	P(<0)
Intercept	0.01	-0.31	0.34	0.53	0.47
WOA effect	0.15	-0.16	0.57	0.79	0.21
Mean rotation effect	-0.18	-0.61	0.11	0.16	0.84
Rotation × density	-0.03	-0.34	0.28	0.43	0.57
Rot. for alfalfa	-0.59	-0.71	-0.47	0.00	1.00
Rot. for flyash	-0.18	-0.52	0.13	0.10	0.89
Rot. for graphite	-0.68	-1.02	-0.36	0.00	1.00
Rot. for gypsum	-0.32	-0.39	-0.24	0.00	1.00
Rot. for LQ	-0.22	-0.34	-0.11	0.00	1.00
Rot. for SiO_2	-0.11	-0.35	0.15	0.17	0.83



Work of adhesion shows a insignificant positive effect, rotation tends to reduce retained area at the average level but remains insignificant, but the magnitude and significance of this reduction varies substantially by particle type, while the rotation–density interaction appears weak.

Conclusions

Key inference

- Work of adhesion has a **positive association** with retained area according to posterior mean but shows **insignificant**.
- Rotation tends to **reduce retained area on average** insignificantly, but the magnitude and significance of this effect **varies substantially by particle type**.
- The rotation–density interaction is **weak**, with posterior mass centered near zero.

Limitations

- Density effects are inferred indirectly through an interaction term and may be underestimated.
- The model focuses on linear effects and does not capture potential nonlinearities.
- Required number of sampling is large, leading to large computational cost.

Future work

- Explore more **nonlinear effects** of predictors.
- Allow different structured prior assumptions informed by material properties.
- Improving the computational efficiency.