Statistical Matters in the Context of GMO Application

Motivation and Goal

- A continued discussion among statistics experts
- Hear our concerns with the statistics review
- Understand EFSA's concerns
- Work together to produce a more effective and efficient submission and review process

Process

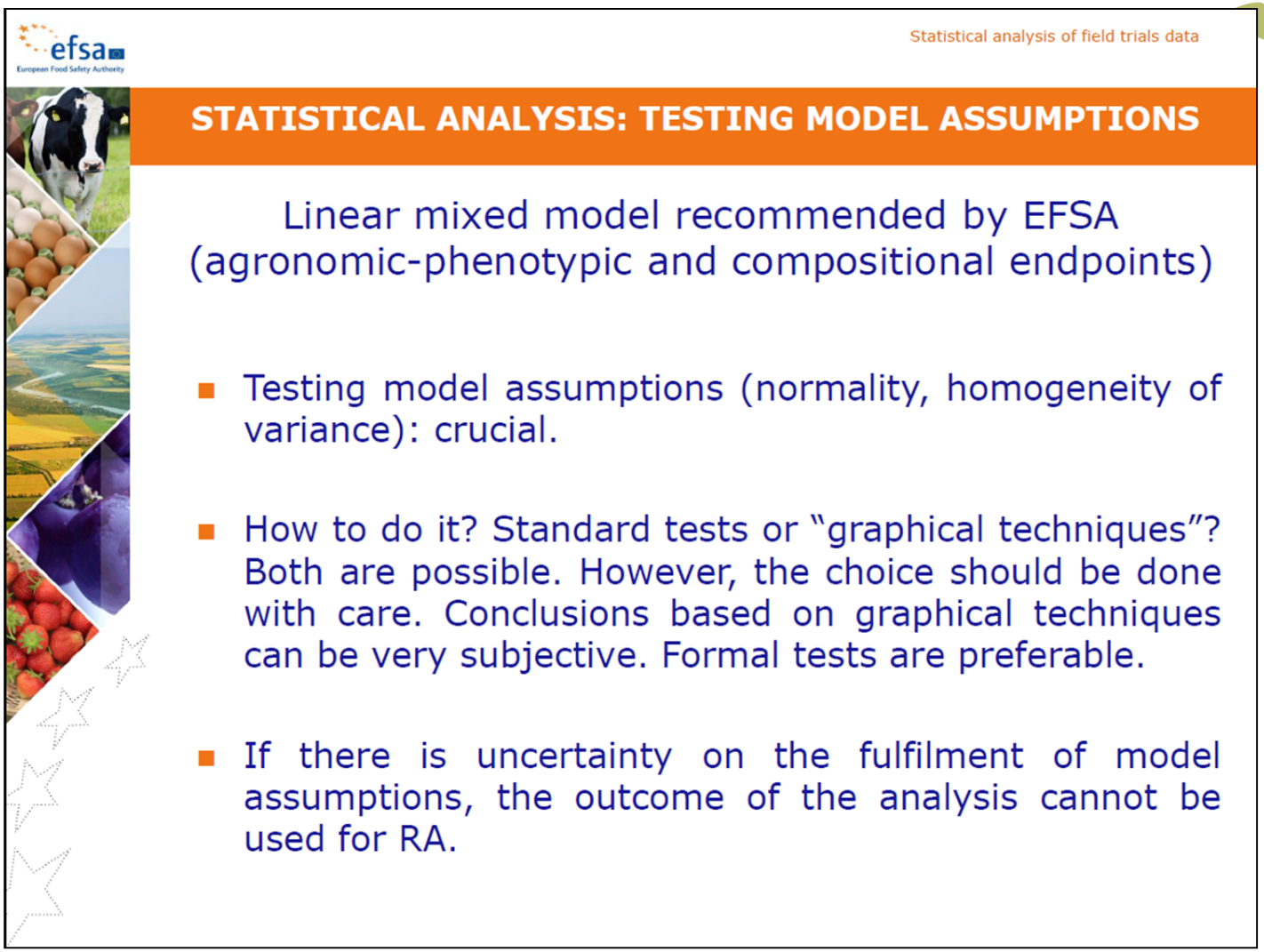
What is a Good Review Process?

- Promote valid statistics
- Promote common understanding of expectations
- Promote transparency and consistency
- Efficient

Issues with the Current Process

- Valid statistical approaches are being challenged
- Expectations are sometimes unclear
- Review questions have caused confusion
- Transparency and consistency have decreased

Unclear Expectations



Review Question Contradictory to Guidance

EFSA Guidance (2010)

Outlying observations can distort statistical analyses. Applicants should investigate whether this might be a problem. In general **graphical approaches are advised, e.g. by looking at residual plots.** Rejection of outliers is only allowed when there are biological/technical reasons….

Review Question (2014)

Visual inspection of the normal probability plots was used by the applicant to identify outliers in the compositional analyses of the comparative assessment. **To reduce the subjectivity** of the outlier detection procedure, the applicant **is requested to use a formal approach** for the detection of outliers (for example, studentized deleted residuals may be used with **cut-off values** that are based on their theoretical statistical reference distribution). The applicant is requested to perform the analysis with and without identified potential outliers.

Review Question Beyond Guidance

In the applicant's response document (XXXX-XX-XX), it is stated that … normality assumption was done by **only visual inspection** of normal probability plots of residuals. Although the EFSA scientific Opinion on 'Statistical considerations for the safety evaluation of GMOs' (2010) **does not explicitly prescribe the procedure to be applied, it requests the evaluation of normality**. The applicant is asked to clarify **why** the 'goodness-of-fit' evaluation on the model applied to check for any discrepancy between observed values and the values expected under the model **was not done by appropriate statistical tests** (e.g. Kolmogorov-Smirnov test).

Review Question Prescribing a Procedure

Question received in 2014

… The applicant is asked to **examine homoscedasticity** (e.g., using scatter plots of the residuals against predicted values and formal tests).

Question received in 2016

For 12 compositional endpoints and 2 agronomic-phenotypic endpoints, the applicant used a site-heteroscedastic variance model ('heterogeneous model') instead of the homogeneous-variance model ('homogeneous model') recommended by EFSA (EFSA GMO Panel, 2011). Quantitative support for this model choice should be provided.
For each of the 14 endpoints:

a) The applicant should carry out a formal test (e.g. Levene's) to check whether or not the assumption of variance homogeneity is violated.

b) If the assumption is violated, the applicant should compare the two competing models (homogeneous and heterogeneous) using standard methods (e.g. AIC).

c) In case the results of (a, b) do not support the choice of the heterogeneous model, the applicant should repeat the full statistical analysis, including difference/equivalence testing, using the homogeneous model.

The applicant should provide all results of **(a-c)** and software code used for the analysis.

How Can We Improve the Process?

- Start discussions between all parties involved to resolve technical disagreement
- Strengthen EFSA oversight for review questions
- Can we utilize the clarification teleconference during risk assessment to discuss the validity of questions?
- A review process that allows flexibility and expert judgment to be incorporated into the statistical analyses

Technical

Technical Summary

- Residual diagnostic plots are the primary tools used by statistics practitioners to evaluate model assumptions
- Formal tests and numerical cutoffs are counter-productive when evaluating model assumptions and identifying outliers
- Graphical assessments of model assumptions, coupled with expert judgment, better address inherent complexity

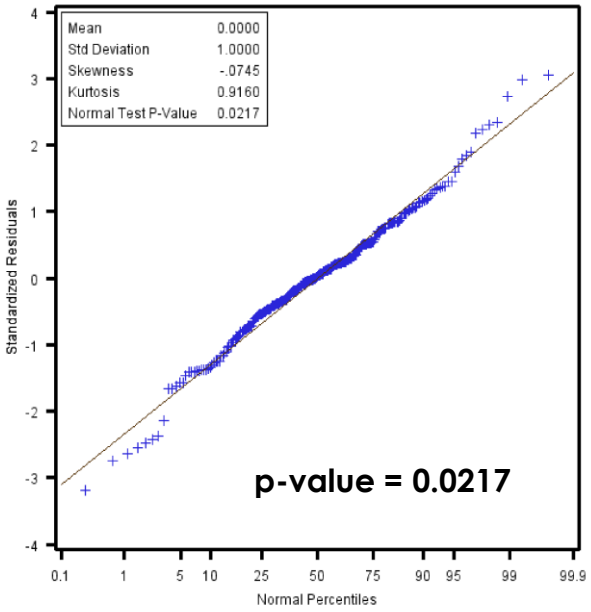
Formal Tests vs. Visual Assessment for Evaluating Model Assumptions

- Formal tests ignore robustness of linear mixed model analyses to minor deviations from assumptions
- With large sample sizes, formal tests are too sensitive and will unnecessarily exclude many endpoints from EFSA-style difference and equivalence testing
- Visual evaluation identifies important violations

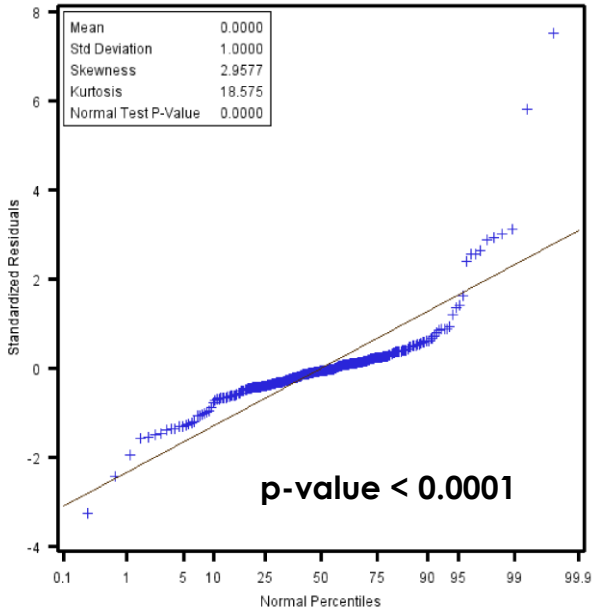
An Example of Using Normality Test

Examples of inconsequential (a) and more serious (b) deviations from normality

(a)



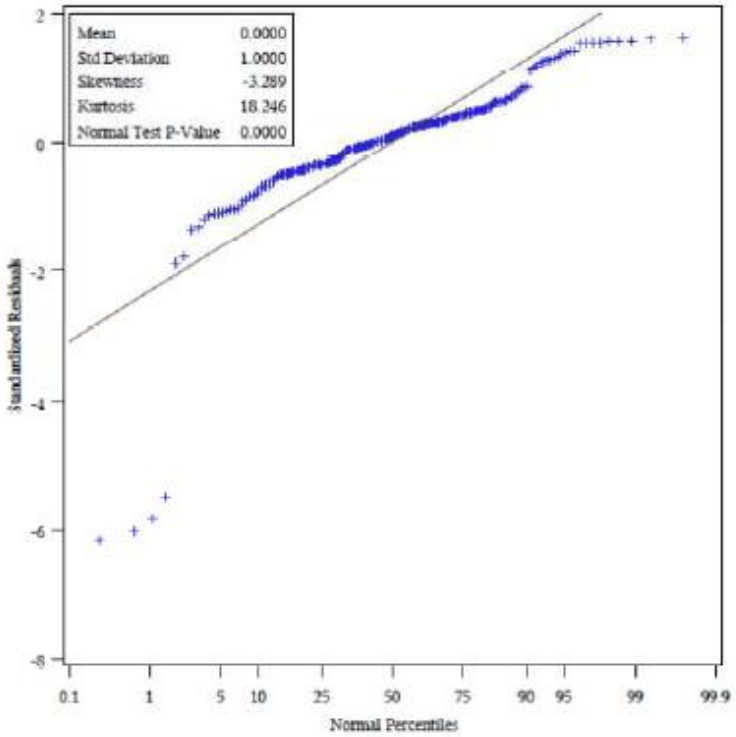
(b)



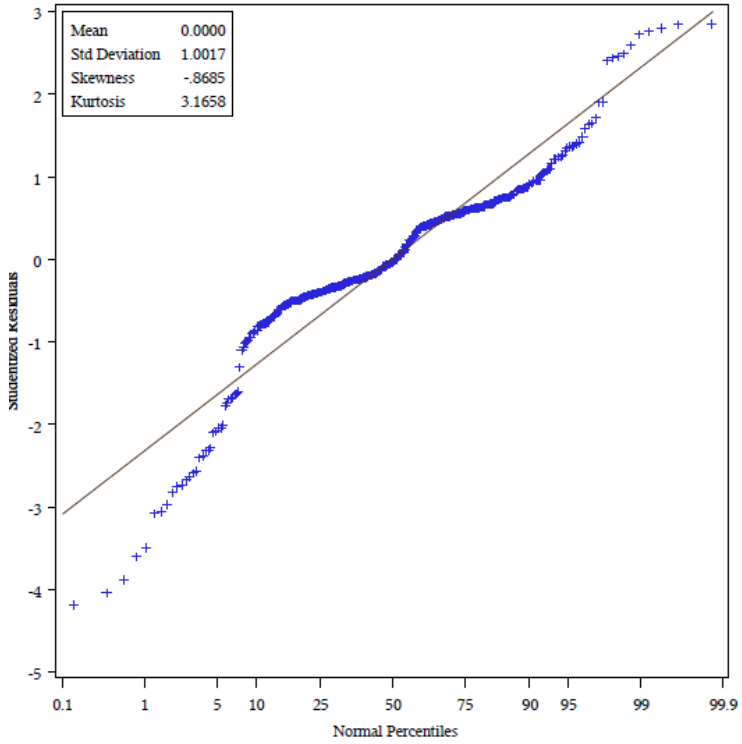
Hypothesis testing should not be used for model assumption evaluation.

Issues with Numerical Cut-offs

(a)



(b)

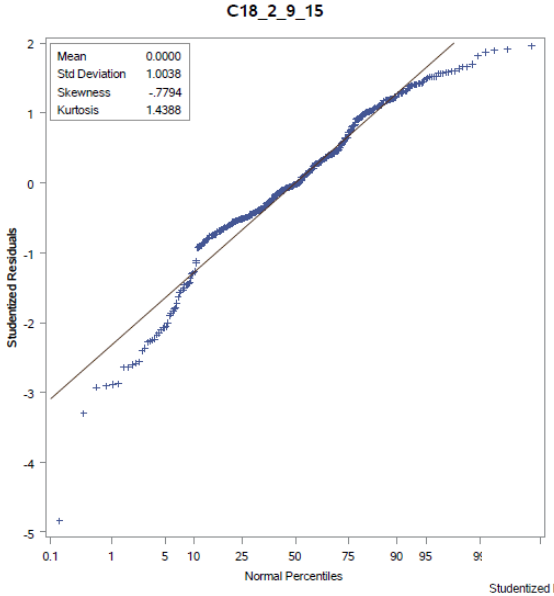


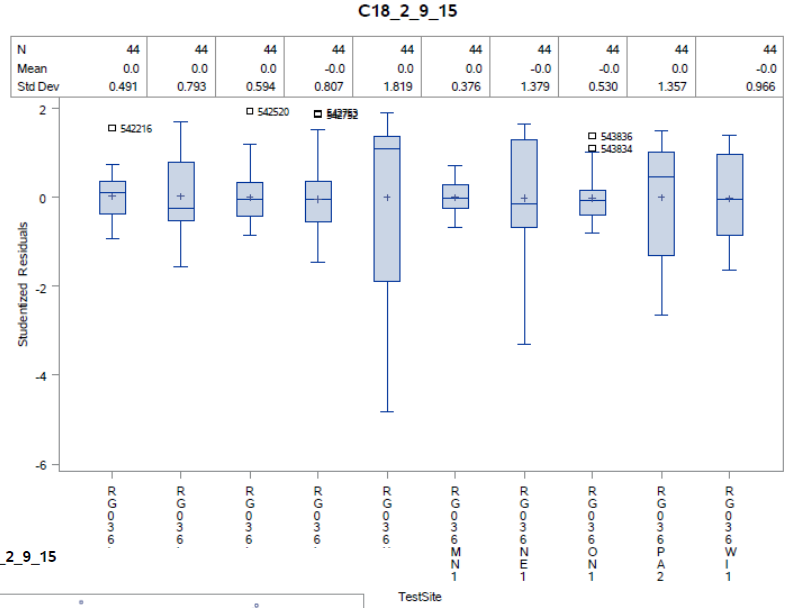
Arbitrary numerical cutoff should not be used for outlier identification.

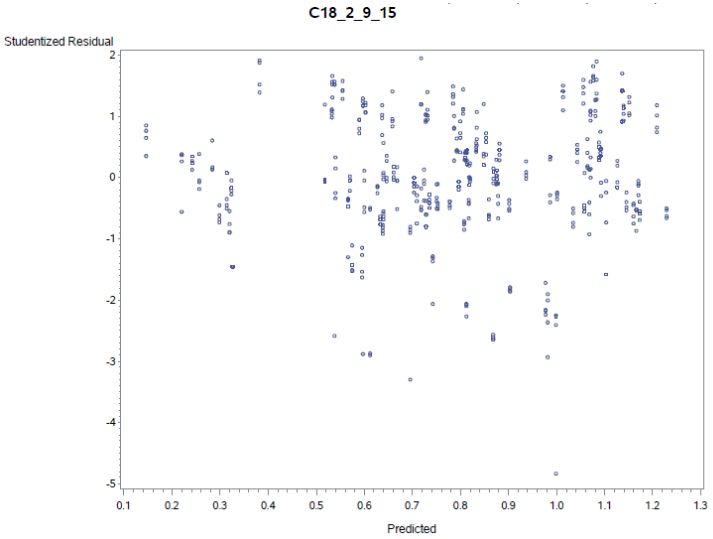
Formal Tests vs. Visual Assessment for Effective Diagnostics

- A formal test provides a single number (a p-value) that gives no information about the nature of the departures from model assumptions
- Visual evaluation deduces causes of violations and effectively parse interrelated issues

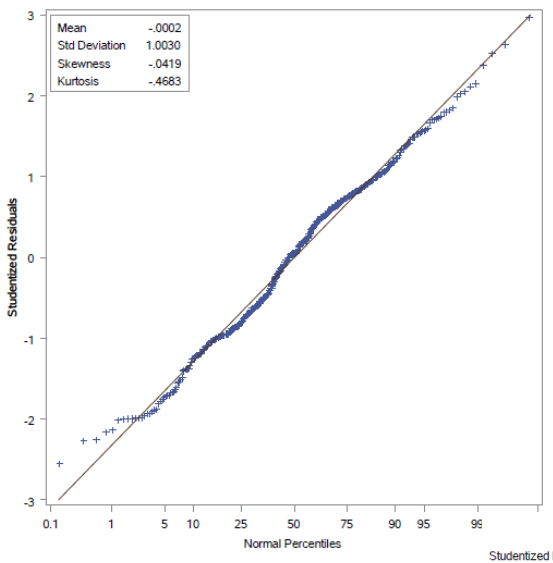
Visualizing Heteroscedasticity

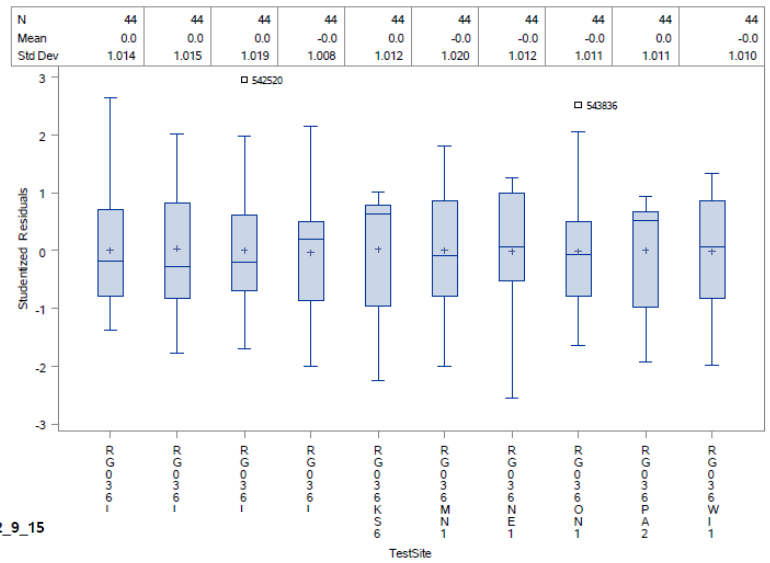


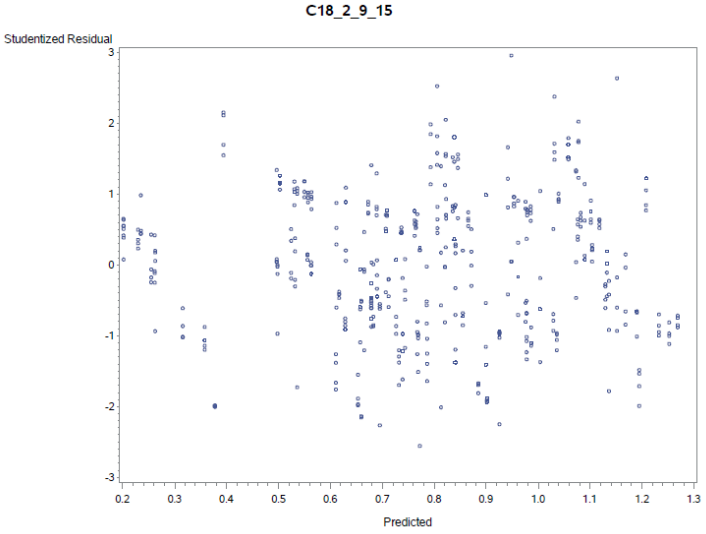




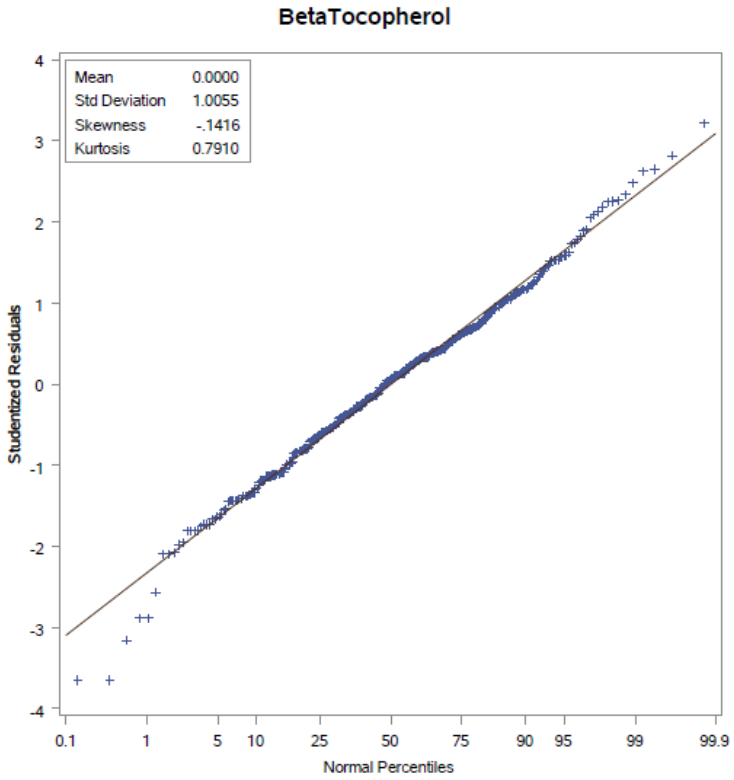
Modeling Heteroscedasticity

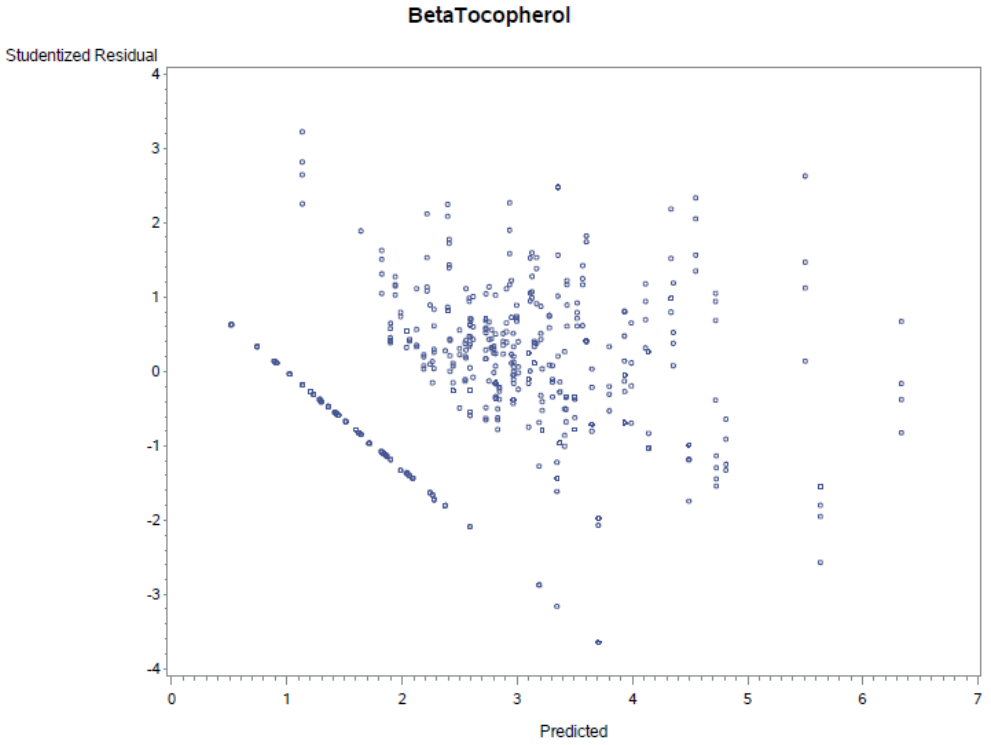




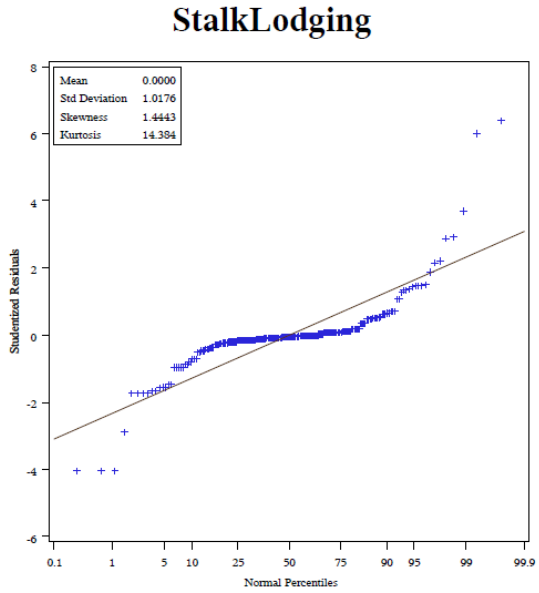


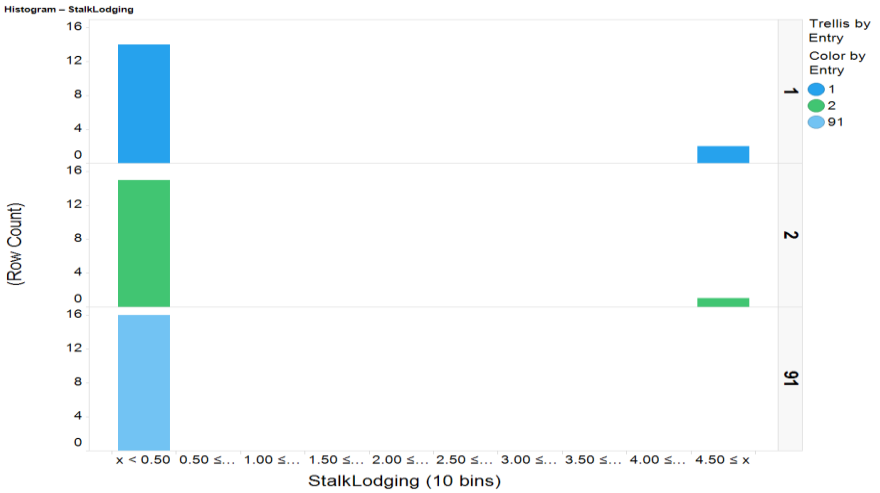
Visualizing Values below LLOQ

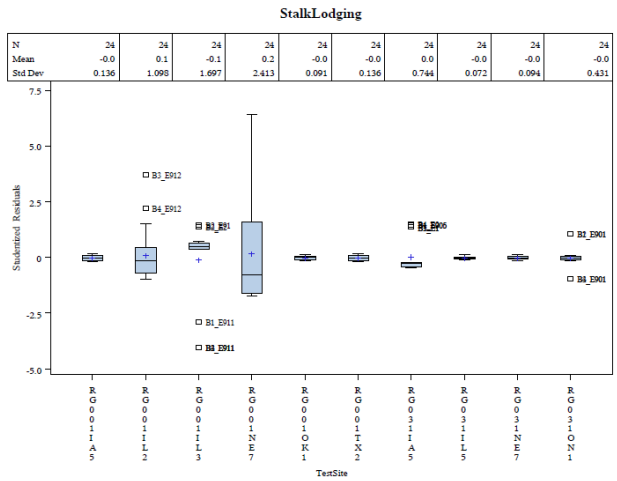




Visualizing Discrete Values



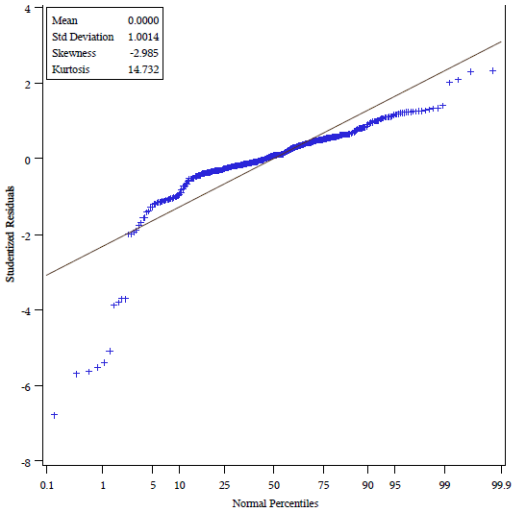


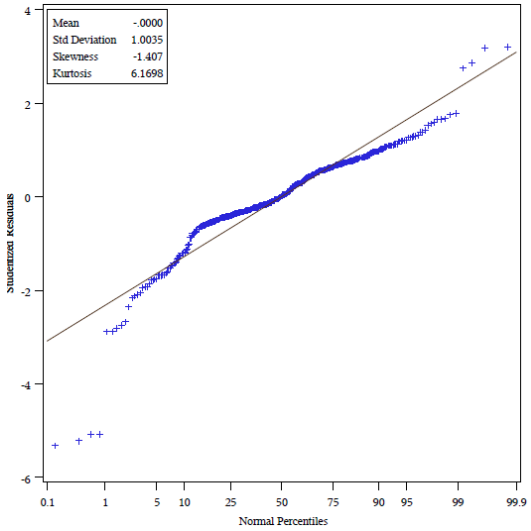


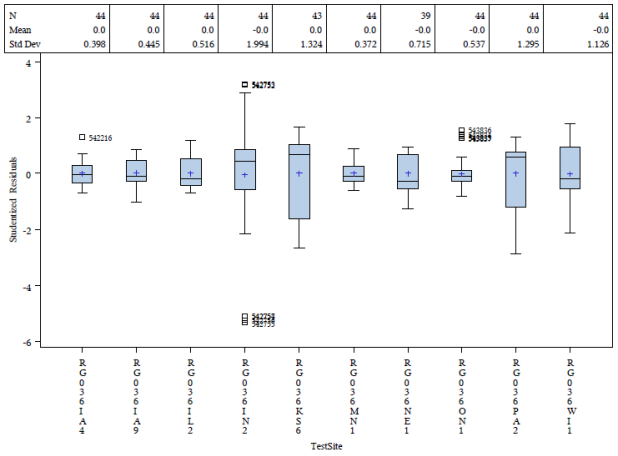
Not appropriate for mixed model analysis.

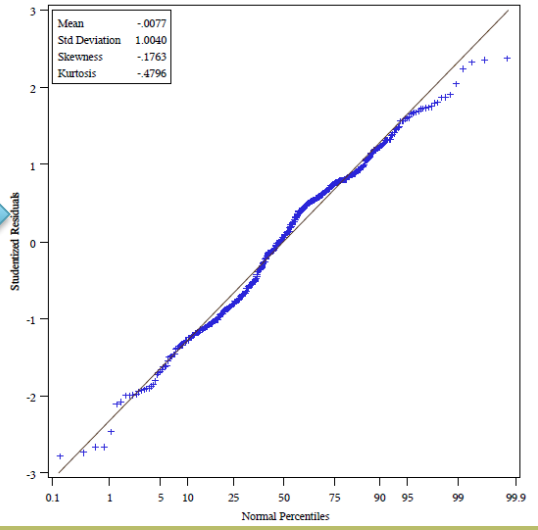
Not appropriate for any statistical testing.

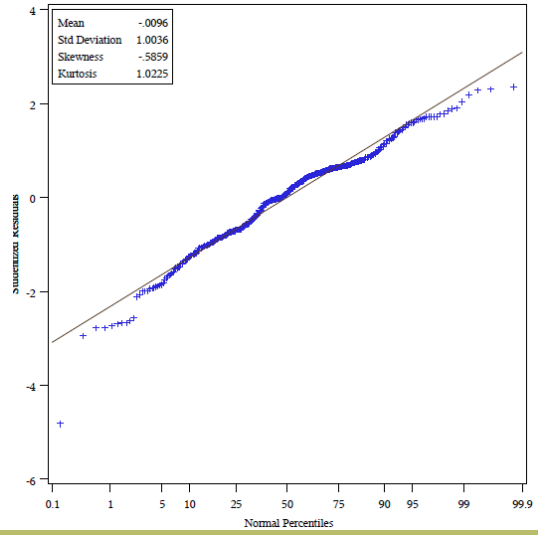
Interrelated Issues











More steps

Alternative sequences of steps are also appropriate because we are dealing with interrelated issues.

Summary

- The review process should allow flexibility and expert judgment to be incorporated into the statistical analyses
- Visual assessment and well executed expert judgment are more appropriate than formal testing of model assumptions
- Further discussions would be beneficial to resolve technical issues and improve review process