Mixed-effect model analysis of ISTA GMO Proficiency Tests

ISTA GMO TF – ISTA Statistics Committee

Jean-Louis Laffont

Outline

PT-Round	Species	Event	spiking levels	# samples
PT01	Maize	T25 MON810	1.0	30
PT02	Maize	MON810	0.7, 1.4	10
PT03	Maize	T25 MON810	0.2, 2.0, 4.0	12
PT04	Soybean	GTS40-3-2	0.1, 0.5, 1.0	12
PT05	Soybean	GTS40 A2704	0.2, 0.5, 1.0, 1.5	12
PT06	Canola	GT73	0.3, 0.6	10
PT07	Maize	MON863 NK603	0.4, 0.8, 2.0	12
PT08	Soybean	GTS40-3-2	0.13, 0.5, 1.0, 2.3	14
PT09	Maize	MON863 NK603	0.1, 0.8, 90, 96	12+2

- Computation of measurement CV and quantification of flour sub-sampling variation
- Comparison of different methods regarding precision

… using mixed-effect modelling

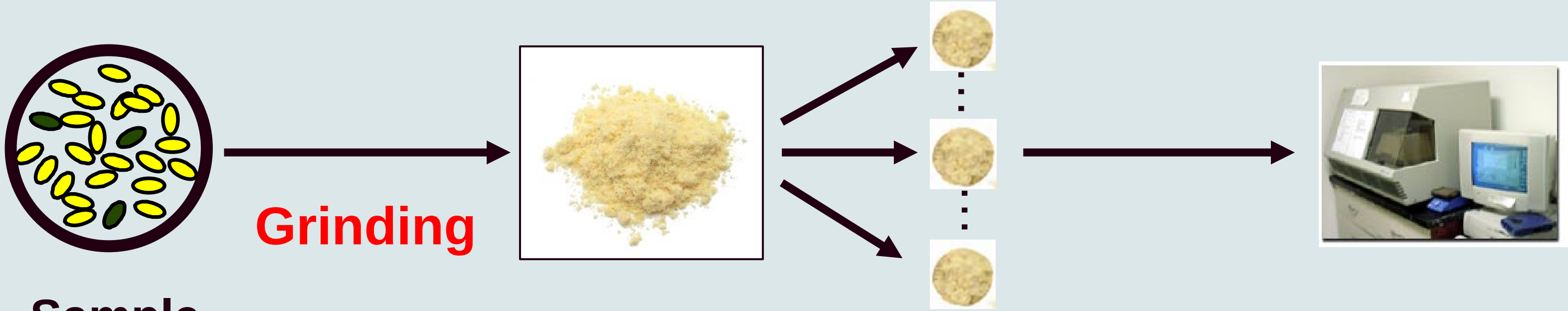
About mixed-effect models…

Analysis of variance (ANOVA) → *Collection of statistical models, and their associated procedures, in which the observed variance is partitioned into components due to different explanatory variables*

Mixed-effect models is one class of such models

→ Will be used here to quantify the different variations contributing to the total variation of a "%GMO" result.

The different sources of variation of a %GMO result from an ISTA PT



For some PTs, some laboratories kindly provided detailed information on their final result:

sample #	spiking level (%)	flour sub-sample #	measurement # (Replicates)	test results (%)
1	2	1	1	0.95
1	2	1	2	0.8
1	2	2	1	0.73
1	2	2	2	0.78
2	4	1	1	1.69
2	4	1	2	1.56
2	4	2	1	1.69
2	4	2	2	1.75
3	0.2	1	1	0.16
...				

Mixed-effect modelling

Mixed-effect model used to analyze laboratory detailed results

Each dataset corresponding to the results from a particular laboratory is analyzed using a heteroscedastic linear mixed effects model.

$$Y_{ijkl} = \mu_i + A_{j(i)} + B_{k(ij)} + E_{ijkl}$$

where:

. μ_i is the mean for the i^{th} spiking level

. $A_{j(i)}$ is the random effect of the j^{th} sample within spiking level i. The $A_{j(i)}$ are i.i.d. $N(0, \sigma^2_{sample})$, where i.i.d. is used to indicate that the observations are independently and identically distributed.

. $B_{k(ij)}$ is the random effect of the k^{th} flour sub-sample from sample j and spiking level i. The $B_{k(ij)}$ are i.i.d. $N(0, \sigma^2_{sample})$.

. E_{ijkl} are the measurement errors:

$$\begin{cases} E_{1jkl} \text{ are i.i.d. } N(0, \sigma_1^2) \\ E_{2jkl} \text{ are i.i.d. } N(0, \sigma_2^2) \\ \dots \\ E_{Ijkl} \text{ are i.i.d. } N(0, \sigma_I^2) \\ \mathrm{cov}(E_{ijkl}, E_{i'j'k'l'}) = 0 \text{ for } i \text{ different from } i' . \end{cases}$$

Measurement variability assumed to be different for each spiking level

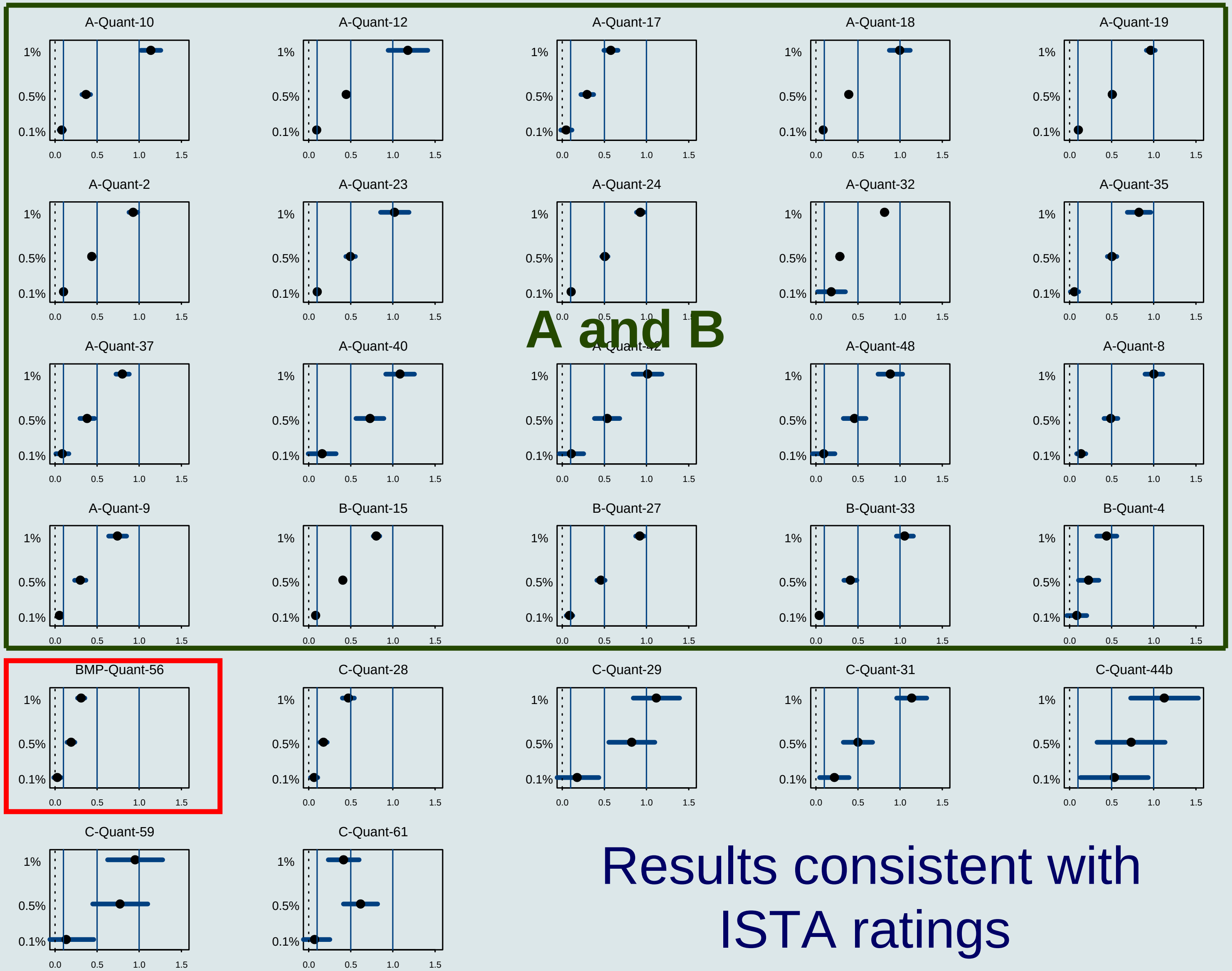
Analysis of laboratory detailed results
Output from mixed-effect model

Lab	Method	Rating	Spiking level (%)	J	K	L	Mean	95% Confidence interval	$s^2_{measurement}$	$CV_{measurement}$	s^2_{sample}	P-value (sample)	s^2_{flour}	P-value (flour)
10	Q	A	0.1	9	3	2	0.083	(0.052 ; 0.114)	0.000391	24%	0.000000	0.5317	0.002026	0.0013
			0.5	9	3	2	0.371	(0.319 ; 0.422)	0.008562	25%				
			1	9	3	2	1.138	(1.020 ; 1.256)	0.061285	22%				
12	Q	A	0.1	9	3	1	0.098	(0.074 ; 0.123)	0.000000	0%	0.000000	0.9984	0.000962	0.0191
			0.5	9	3	1	0.447	(0.422 ; 0.471)	0.000000	0%				
			1	9	3	1	1.179	(0.943 ; 1.415)	0.129417	31%				
15	Q	B	0.1	9	3	2	0.086	(0.070 ; 0.101)	0.000210	17%	0.000141	0.0658	0.000007	0.9292
			0.5	9	3	2	0.407	(0.384 ; 0.430)	0.001648	10%				
			1	9	3	2	0.804	(0.765 ; 0.843)	0.006250	10%				
18	Q	A	0.1	9	2	1	0.087	(0.072 ; 0.102)	0.000012	4%	0.000037	0.8221	0.000271	1.0000
			0.5	9	2	1	0.392	(0.375 ; 0.409)	0.000108	3%				
			1	9	2	1	0.997	(0.873 ; 1.120)	0.023503	15%				
19	Q	A	0.1	9	2	3	0.106	(0.093 ; 0.119)	0.000441	20%	0.000000	0.9983	0.000126	0.3762
			0.5	9	2	3	0.509	(0.480 ; 0.538)	0.003659	12%				
			1	9	2	3	0.966	(0.913 ; 1.019)	0.012920	12%				
2	Q	A	0.1	9	2	3	0.102	(0.092 ; 0.112)	0.000275	16%	0.000000	0.9975	0.000040	0.6247
			0.5	9	2	3	0.437	(0.418 ; 0.456)	0.001581	9%				
			1	9	2	3	0.928	(0.877 ; 0.980)	0.012244	12%				
...														

… a lot of information!

A portion of it will be mined in the following slides…

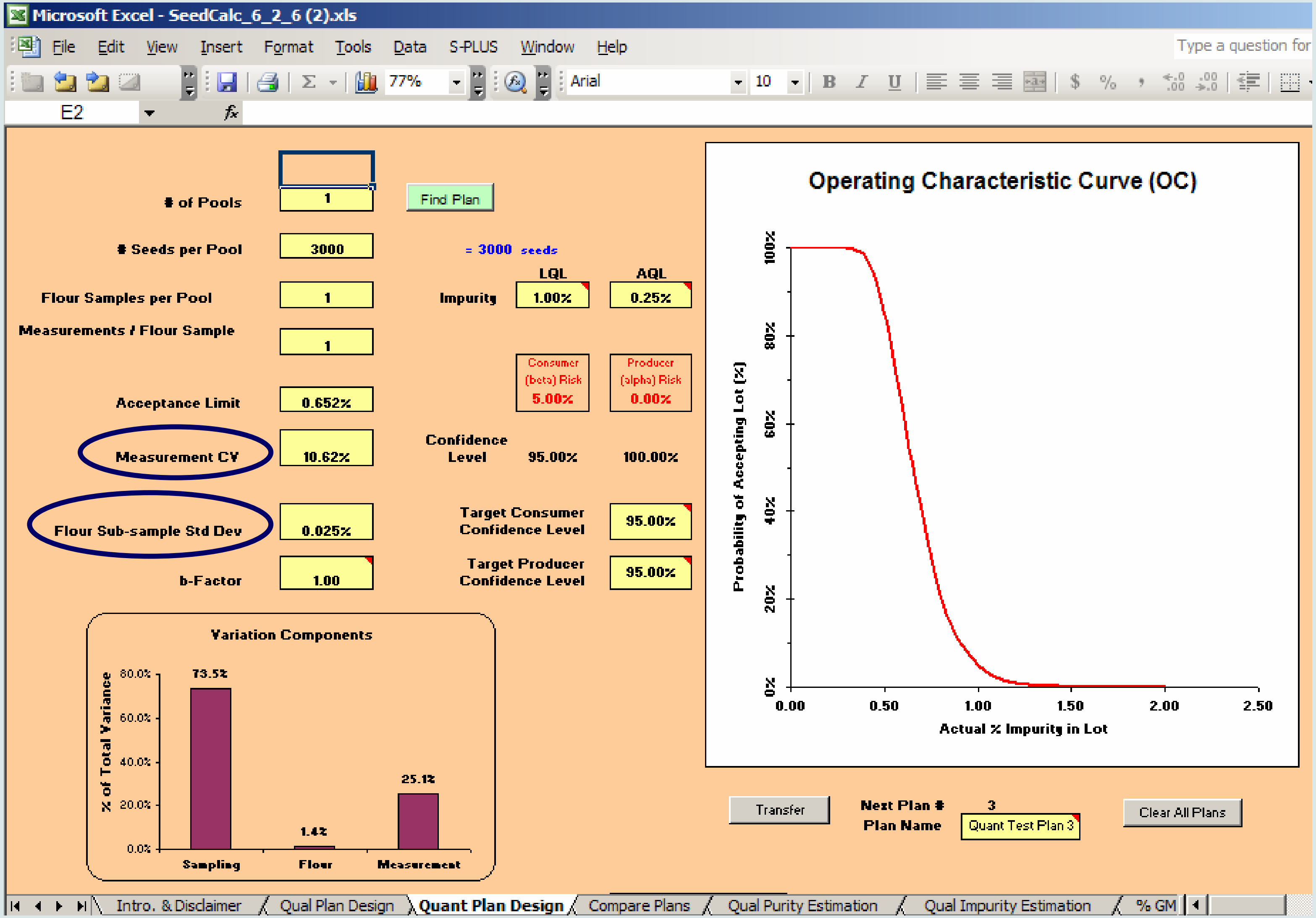
Analysis of laboratory detailed results – PT04 example
Mean estimates and associated 95%CI



Results consistent with ISTA ratings

Analysis of laboratory detailed results

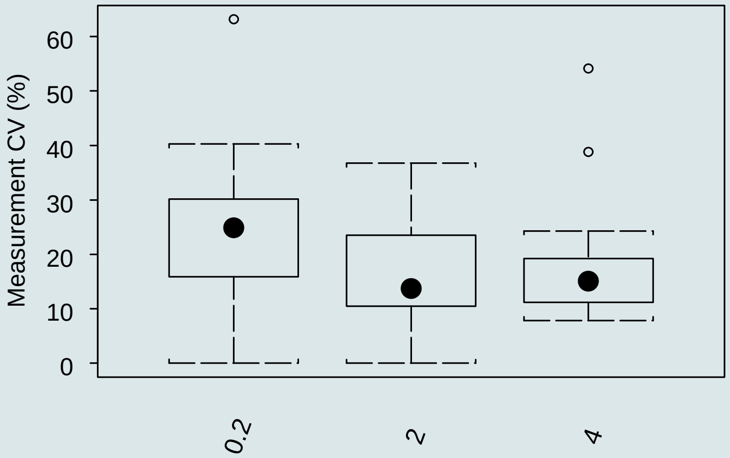
Seedcalc *Measurement CV* and *Flour Sub-sample Std Dev*



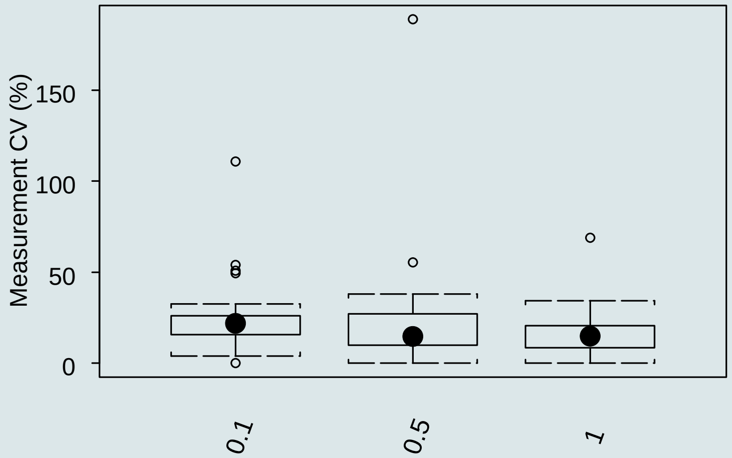
Analysis of laboratory detailed results

Measurement CV

PT03 - Measurement CV (%)

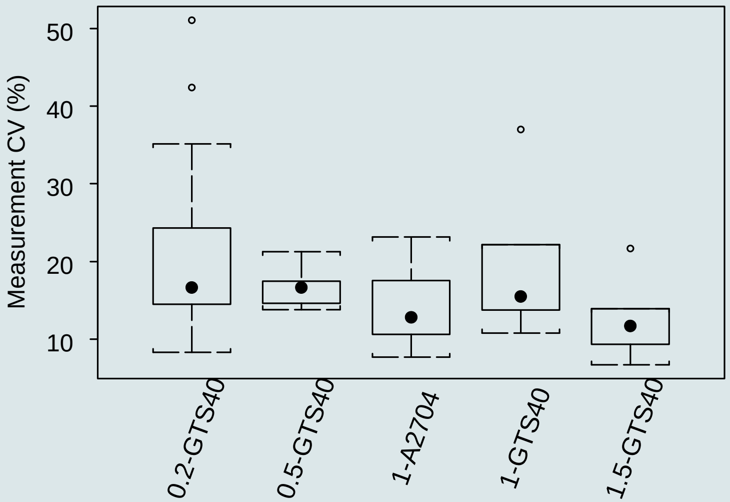


PT04 - Measurement CV (%)

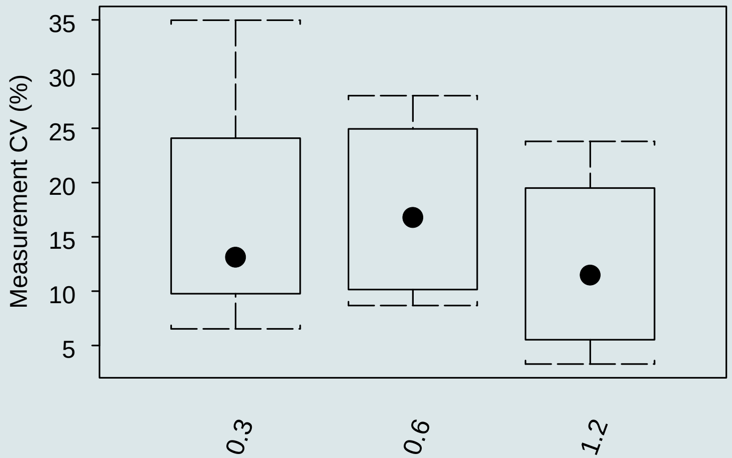


Average measurement CV: 20% - 0.75 quantile: 25%

PT05 - Measurement CV (%)

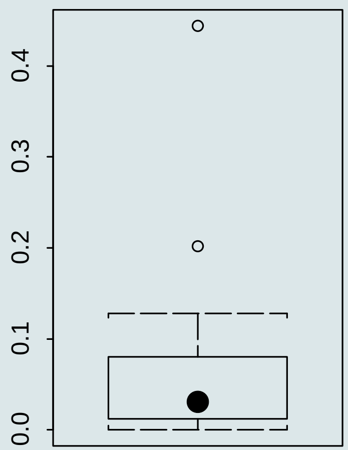


PT06 - Measurement CV (%)

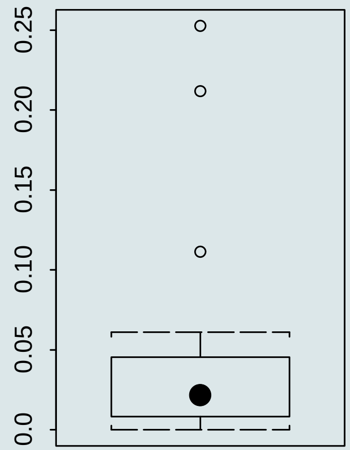


Analysis of laboratory detailed results
Flour Sub-sample Std Dev

PT03 - Flour sub-sample Std Dev

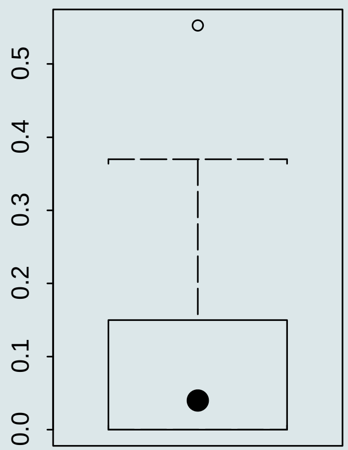


PT04 - Flour sub-sample Std Dev

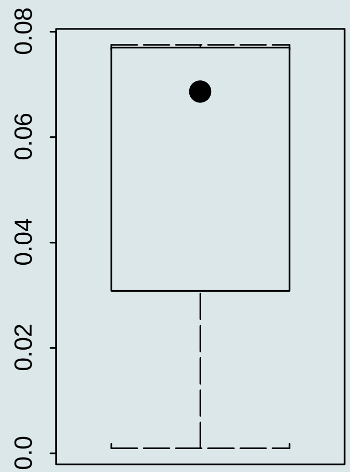


Average: 0.06% - 0.75 quantile: 0.08%

PT05 - Flour sub-sample Std Dev



PT06 - Flour sub-sample Std Dev



Mixed-effect model used to analyze laboratory submitted results

For PT07 and PT09, we have the information unit of measure (%DNA copies, %Mass, %Seed).
For the other PTs, we have the information about the method used: Quantitative vs Semi-Quantitative.

Each dataset corresponding to one type of unit of measure or one method and one PT (laboratories with BMP rating removed) is analyzed using an heteroscedastic mixed-effect model.

$$Y_{ijk} = \mu + \alpha_i + L_j + \left(\alpha L\right)_{ij} + E_{ijk}$$

where: . μ is the intercept.

. α_i is the fixed effect of the i^{th} spiking level.

. $B_{k(ij)}$ is the random effect of the j^{th} laboratory. The L_j are i.i.d. $N(0,\ \sigma^2_{lab})$.

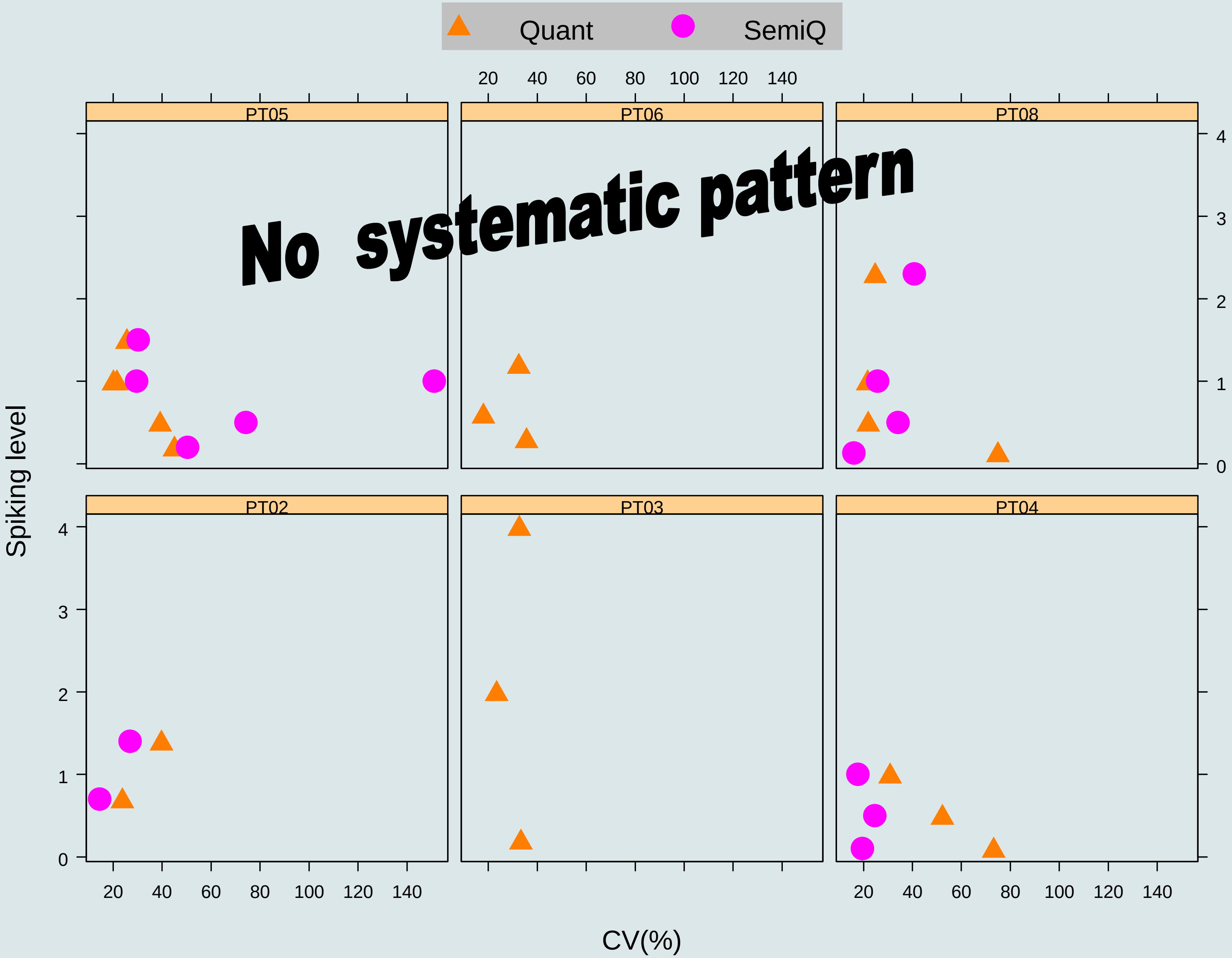
. $(\alpha L)_{ij}$ is the random interaction effect between the i^{th} spiking level and the j^{th} laboratory. The $(\alpha L)_{ij}$ are i.i.d. $N(0, \sigma^2_{spiking\_level \times lab})$.

. E_{ijk} are the residuals:

$$\begin{cases} E_{1jk} \text{ are i.i.d. } N(0, \sigma_1^2) \\ \ldots \\ E_{Ijk} \text{ are i.i.d. } N(0, \sigma_I^2) \\ \text{cov}(E_{ijk}, E_{i'j'k'}) = 0 \text{ for } i \text{ different from } i' . \end{cases}$$

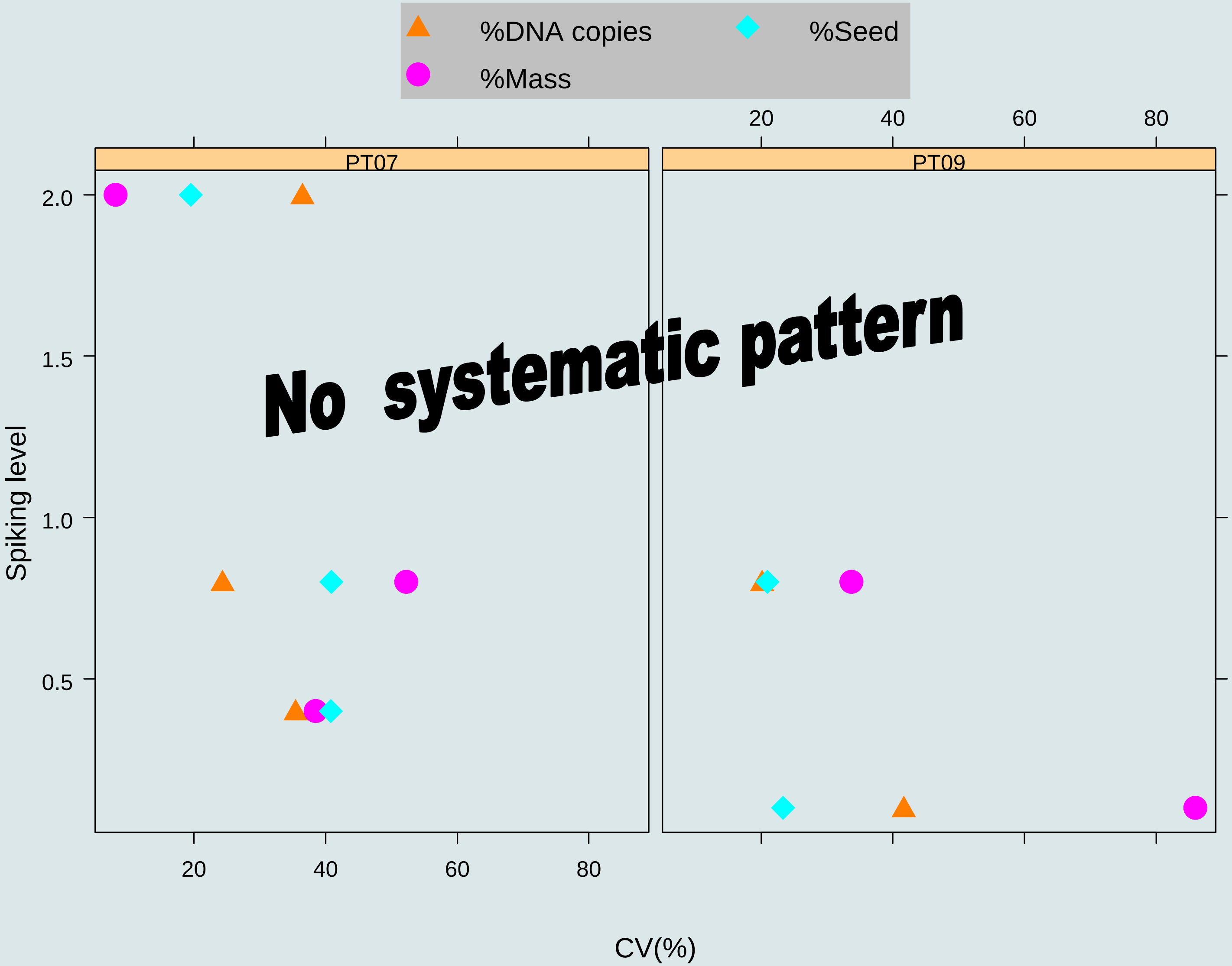
Analysis of laboratory submitted results

Residual variation expressed as $\hat{\sigma}_i / (\hat{\mu} + \hat{\alpha}_i)$



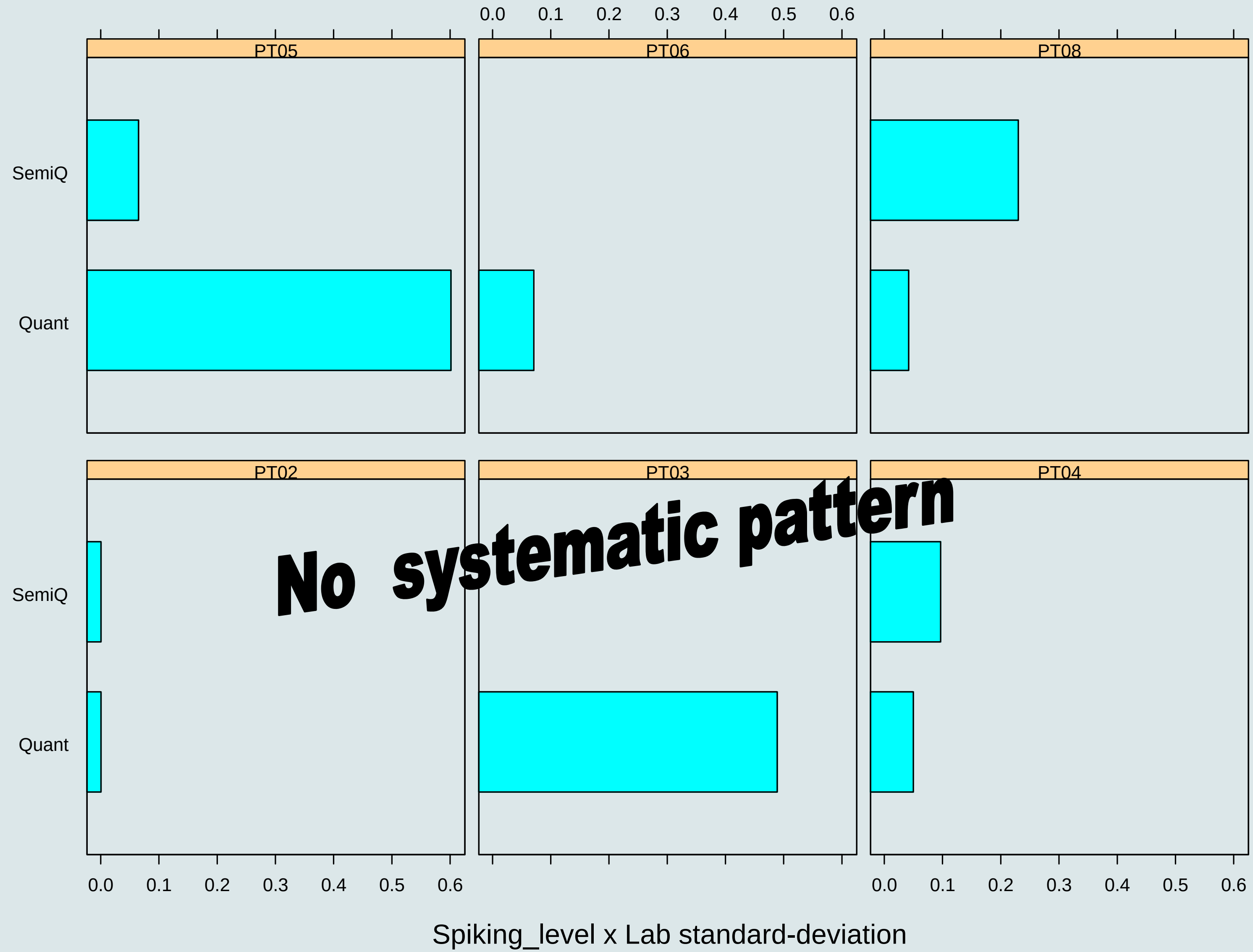
Analysis of laboratory submitted results

Residual variation expressed as $\hat{\sigma}_i / (\hat{\mu} + \hat{\alpha}_i)$



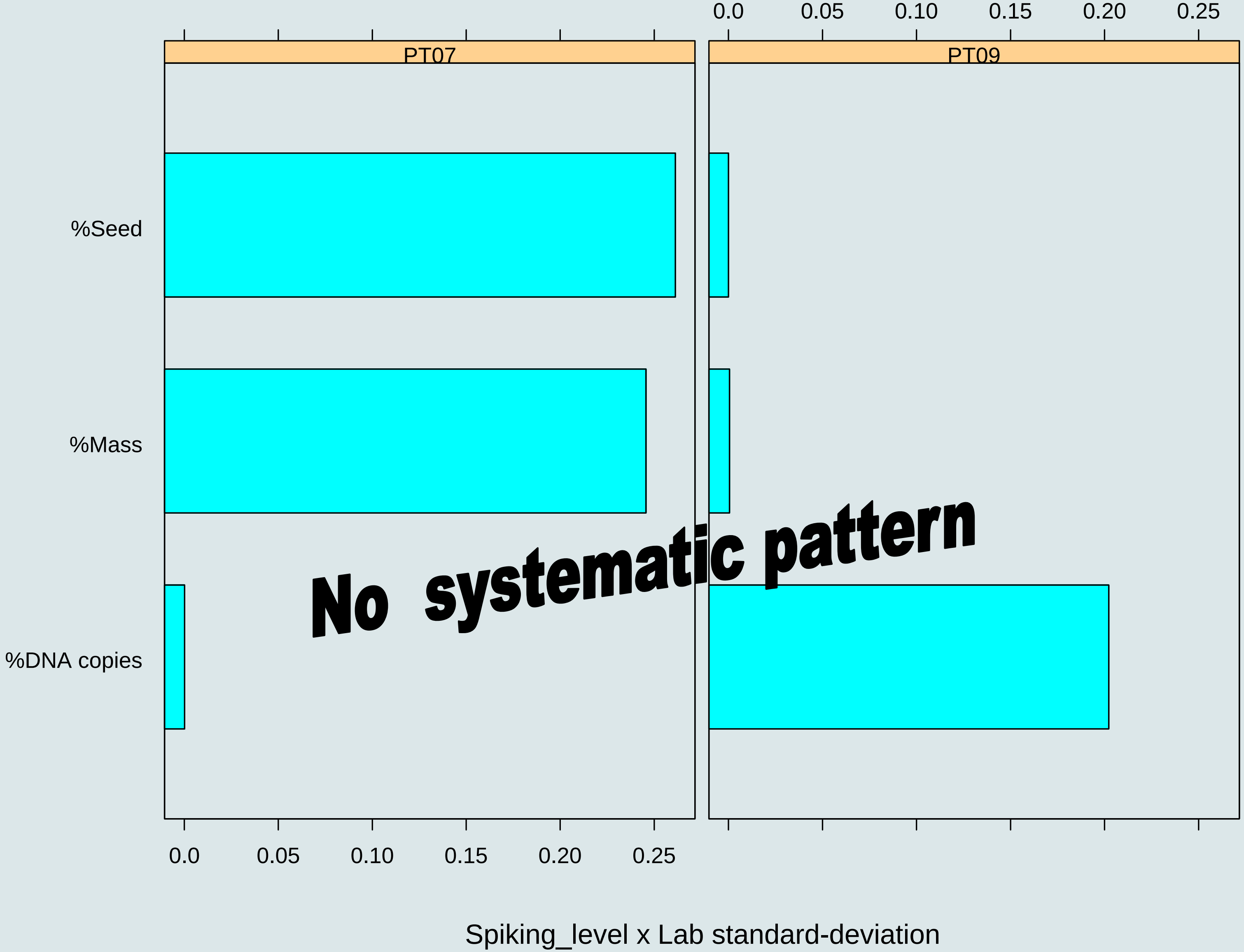
Analysis of laboratory submitted results

Spiking level x Lab interaction expressed as $\hat{\sigma}_{spiking\_level \times lab}$



Analysis of laboratory submitted results

Spiking level x Lab interaction expressed as $\hat{\sigma}_{spiking\_level \times lab}$



Summary

- The mixed-effect model analysis of ISTA GMO PTs allowed us to estimate some parameters useful for Quantitative Test Plan design with Seedcalc.

- The mixed-effect model analysis of ISTA GMO PTs did not reveal differences in precision (not accuracy) when different methods or units of measure are used for estimating GMO%.