Exploring Treatment By Trial Interaction Using Treatment Stability/Trial Dendrogram Plots

Peter Claussen, Steve Gylling

Gylling Data Management, Inc.

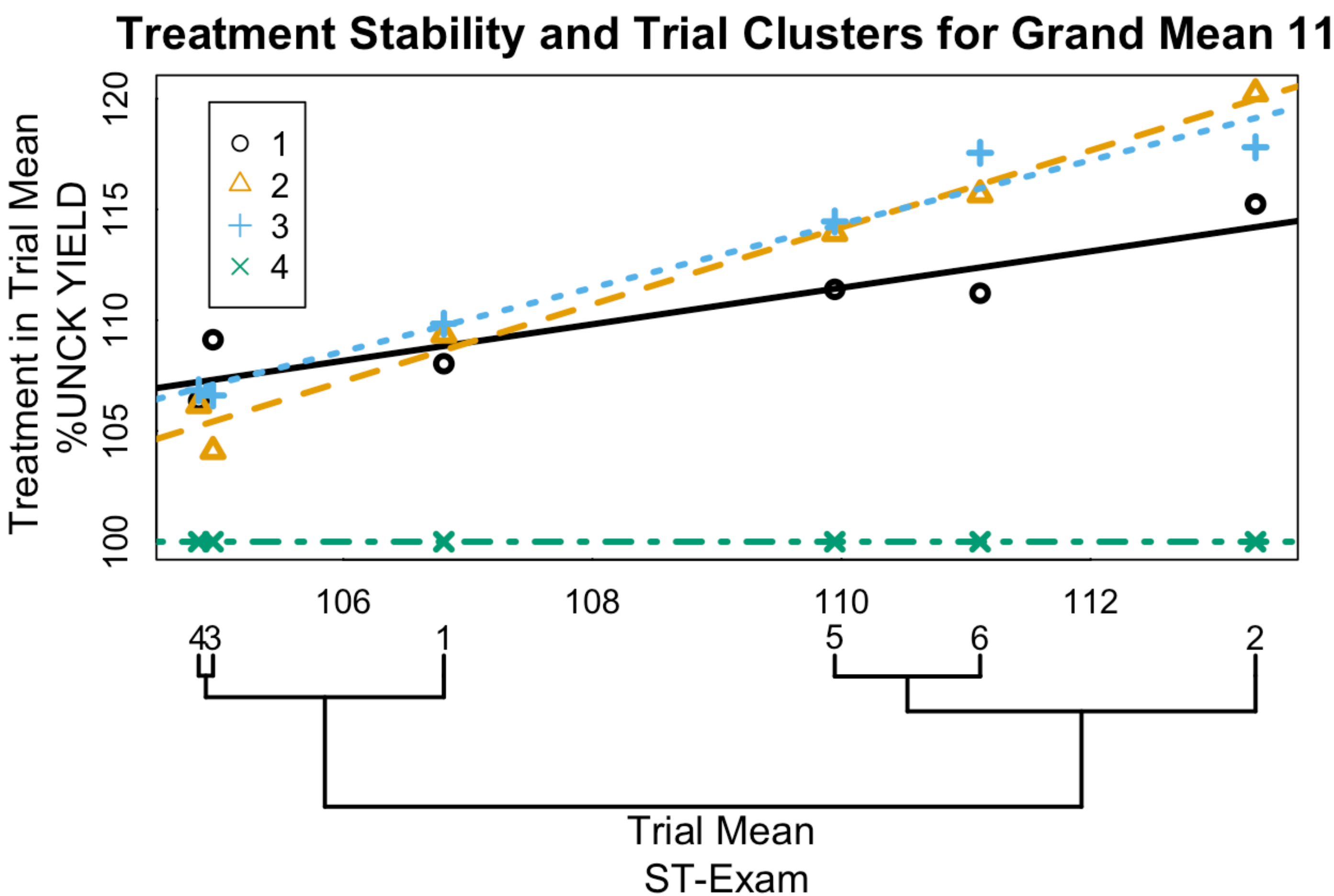
Booth #1020

Software Overview

- ARM
  - Single trial management and analysis
- ARM ST
  - Summarizes ARM trials across locations and years
  - Treatment x trial analysis
  - New in 2015 is the Treatment x Trial graph

Treatment x Trial Graph

Treatment Stability



Trial Dendrogram

(colorblind-safe palette – bconnelly.net/2013/10/creating-colorblind-friendly-figures/)

Trial Dendrogram

- Treatment x trial means represented as a matrix, with trials in rows.
  - Each trial is a single multivariate response comprised of individual treatment means within trials.

$$\begin{pmatrix} \vec{y}_1 \\ \vec{y}_2 \\ \vdots \\ \vec{y}_m \end{pmatrix} = \begin{pmatrix} \left(\overline{y}_{11}, \overline{y}_{12}, \ldots, \overline{y}_{1n}\right) \\ \left(\overline{y}_{21}, \overline{y}_{22}, \ldots, \overline{y}_{2n}\right) \\ \vdots \\ \left(\overline{y}_{m1}, \overline{y}_{m2}, \ldots, \overline{y}_{mn}\right) \end{pmatrix}$$

Trial Dendrogram

- Similarity between rows (trials) is computed using the R function `dist()`
  - The R default is Euclidean distance.

$$\left\|\vec{\bar{y}}_1-\vec{\bar{y}}_2\right\|=\left(\left(\bar{y}_{11}-\bar{y}_{12}\right)^2+\left(\bar{y}_{21}-\bar{y}_{22}\right)^2+\ldots+\left(\bar{y}_{n1\cdot}-\bar{y}_{n3}\right)^2\right)^{1/2}$$

  - The result is a matrix with pairwise similarity measures

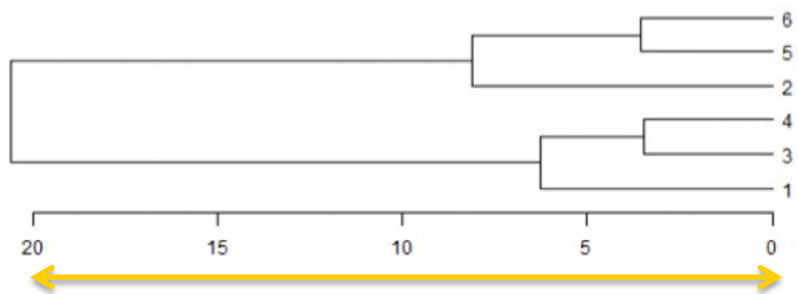
$$\begin{pmatrix}
 & \left\|\vec{\bar{y}}_1-\vec{\bar{y}}_2\right\| & \left\|\vec{\bar{y}}_1-\vec{\bar{y}}_3\right\| & \cdots & \left\|\vec{\bar{y}}_1-\vec{\bar{y}}_m\right\| \\
\left\|\vec{\bar{y}}_2-\vec{\bar{y}}_1\right\| & & \left\|\vec{\bar{y}}_2-\vec{\bar{y}}_3\right\| & \cdots & \left\|\vec{\bar{y}}_2-\vec{\bar{y}}_m\right\| \\
\left\|\vec{\bar{y}}_3-\vec{\bar{y}}_1\right\| & \left\|\vec{\bar{y}}_3-\vec{\bar{y}}_2\right\| & & \cdots & \left\|\vec{\bar{y}}_3-\vec{\bar{y}}_m\right\| \\
\vdots & \vdots & \vdots & & \vdots \\
\left\|\vec{\bar{y}}_m-\vec{\bar{y}}_1\right\| & \left\|\vec{\bar{y}}_m-\vec{\bar{y}}_2\right\| & \left\|\vec{\bar{y}}_m-\vec{\bar{y}}_3\right\| & \cdots &
\end{pmatrix}$$

Trial Dendrogram

- Hierarchical clusters produced using the R `hclust()`function.
  - Each item (leaf) is first given its own cluster
  - Proceeding iteratively, the nearest pair of clusters are joined until there is a single cluster
- Dendrogram is plotted by drawing lines linking clusters, where the distance of the link from the leaves is proportional to similarity between clusters

Trial Dendrogram

- Leaves are typically reordered to prevent crossing of branches.



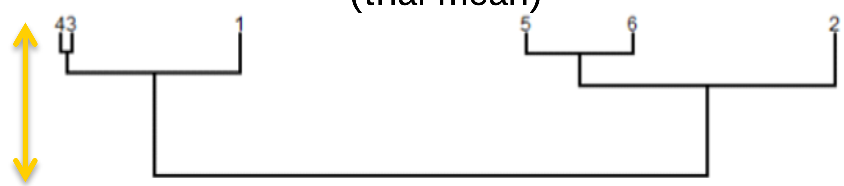
(trial number)

(distance)

- The ARM ST trial dendrogram constrains leaves to be aligned with trial means in the stability plot

(trial mean)

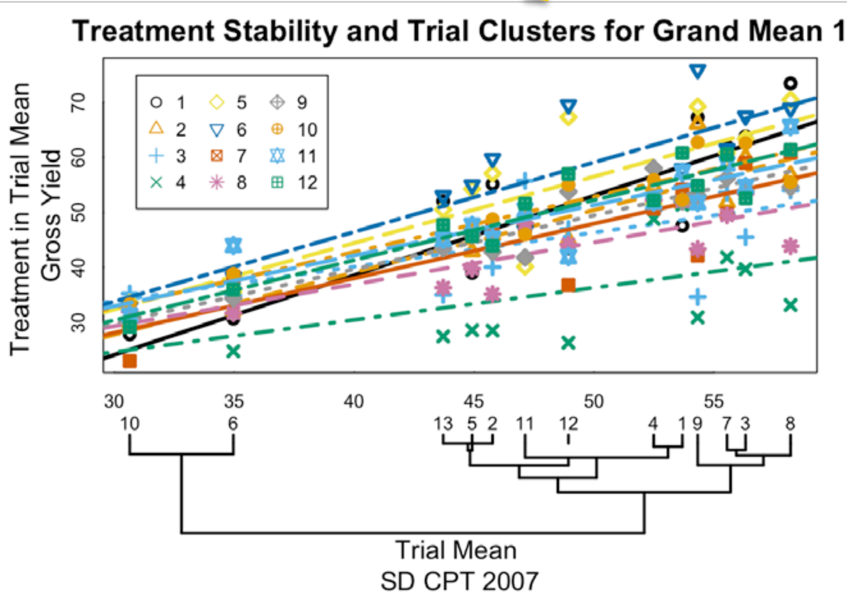
(distance)

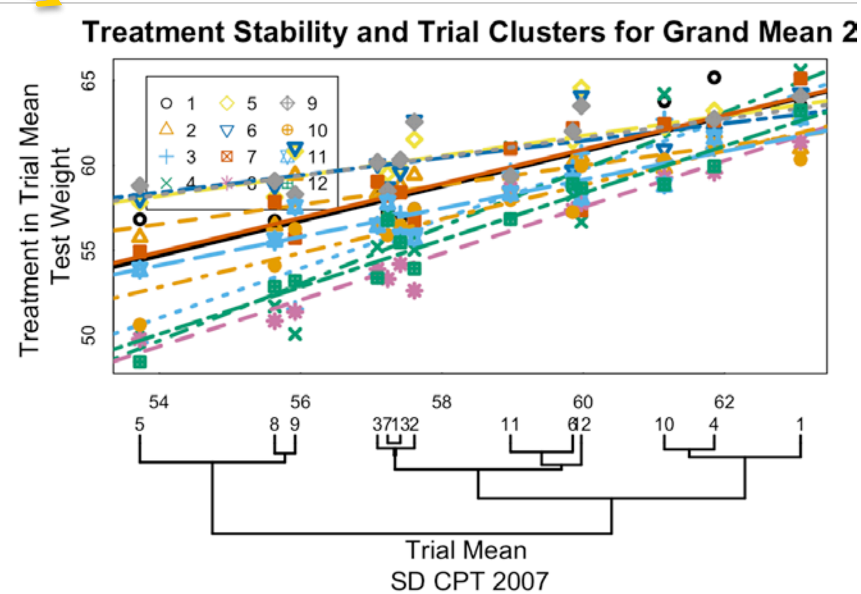


Types of interaction

How does the dendrogram relate to interaction?

- Some interactions arise when treatment differences tend to diverge or converge with increasing trial means

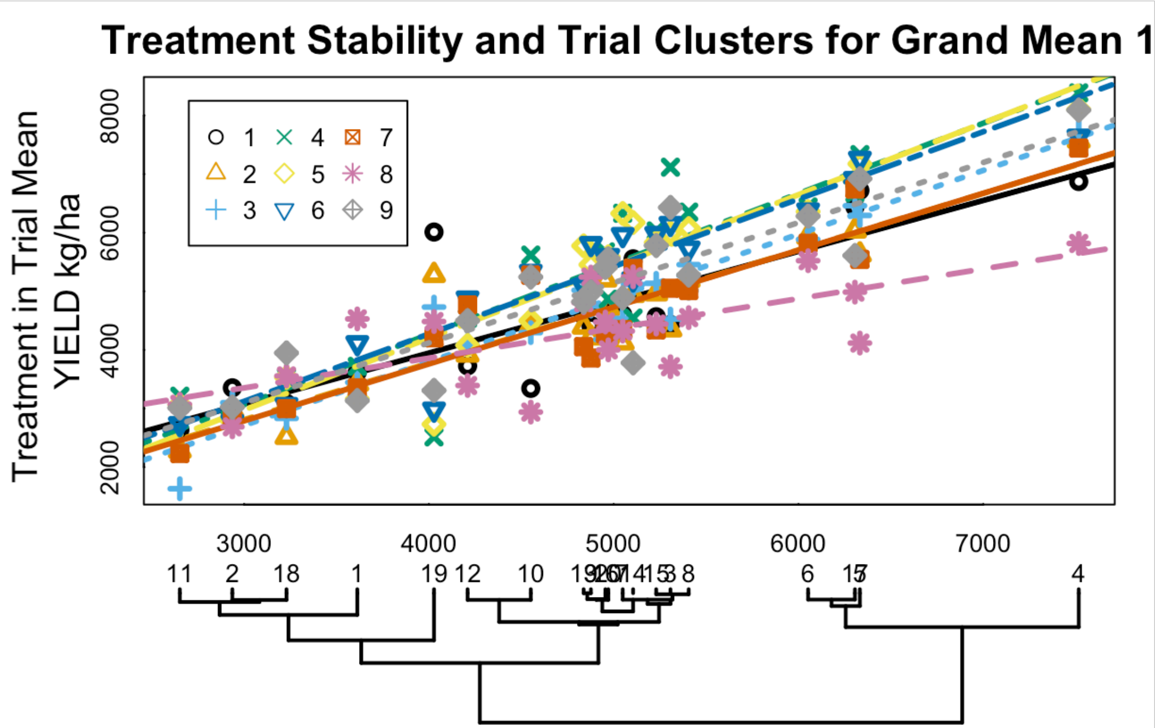




 (Data from South Dakota Crop Performance Trials, Winter Wheat, 2007)

Types of interaction

- Other interactions arise when some treatments perform inconsistently across trials



(Data from Cornelius et al., 1997)

Treatment x trial means

- Treatment x trial means can be written as the sum of

$$\overline{y}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\theta}_{ij}$$

  - ⌠ grand mean
  - 〈$_{\iota}$ treatment effect
  - ®$_{\varphi}$ trial effect
  - ($_{\iota\varphi}$ treatment x trial interaction

Trial similarity

- When distances are computed from treatment in trial means, treatment effects cancel and the distance is based on trial effect and treatment x trial interaction.

$$\vec{y}_{\cdot 1}-\vec{y}_{\cdot 2} = \left(\overline{y}_{11\cdot}-,\overline{y}_{12\cdot},\overline{y}_{21\cdot}-\overline{y}_{22\cdot},\cdots,\overline{y}_{n1\cdot}-\overline{y}_{n2\cdot}\right)$$

$$= \begin{pmatrix} \left(\hat{\mu}+\hat{\alpha}_1+\hat{\beta}_1+\hat{\theta}_{11}\right)-\left(\hat{\mu}+\hat{\alpha}_1+\hat{\beta}_2+\hat{\theta}_{12}\right), \\ \left(\hat{\mu}+\hat{\alpha}_2+\hat{\beta}_1+\hat{\theta}_{21}\right)-\left(\hat{\mu}+\hat{\alpha}_2+\hat{\beta}_2+\hat{\theta}_{22}\right), \\ \vdots \\ \left(\hat{\mu}+\hat{\alpha}_n+\hat{\beta}_1+\hat{\theta}_{n1}\right)-\left(\hat{\mu}+\hat{\alpha}_n+\hat{\beta}_2+\hat{\theta}_{n2}\right) \end{pmatrix} = \begin{pmatrix} \left(\hat{\beta}_1+\hat{\theta}_{11}\right)-\left(\hat{\beta}_2+\hat{\theta}_{12}\right), \\ \left(\hat{\beta}_1+\hat{\theta}_{21}\right)-\left(\hat{\beta}_2+\hat{\theta}_{22}\right), \\ \vdots \\ \left(\hat{\beta}_1+\hat{\theta}_{n1}\right)-\left(\hat{\beta}_2+\hat{\theta}_{n2}\right) \end{pmatrix}$$

Trial similarity, no interaction

- If interaction is absent, the similarity between two trials is proportional to the difference between trial effects only.

$$\vec{y}_{\cdot 1} - \vec{y}_{\cdot 2} = \begin{pmatrix} (\hat{\beta}_1 + 0) - (\hat{\beta}_2 + 0) \\ (\hat{\beta}_1 + 0) - (\hat{\beta}_2 + 0) \\ \vdots \\ (\hat{\beta}_1 + 0) - (\hat{\beta}_2 + 0) \end{pmatrix} = \begin{pmatrix} \hat{\beta}_1 - \hat{\beta}_2 \\ \hat{\beta}_1 - \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_1 - \hat{\beta}_2 \end{pmatrix}$$

$$\left\| \vec{y}_{\cdot 1} - \vec{y}_{\cdot 2} \right\| = \left(n \left(\hat{\beta}_1 - \hat{\beta}_2 \right)^2 \right)^{1/2}$$

Decomposition of Interaction

- Simple additivity
  - $\theta_{11} = \theta_{12} = \ldots = \theta_{nm} = 0$
- Random effect
  - $\theta_{ij} \sim N\left(0, \sigma_{\theta}^{2}\right)$
- Nonadditivity
  - $\theta_{ij} = \lambda \alpha_i \beta_j + \varepsilon_{ij}$

Tukey 1949

- Heterogeneous Slopes (Stability)
  - $\theta_{ij} = \lambda_i \beta_j + \varepsilon_{ij}$

Mandel 1961

- Other decompositions

Milliken and Johnson 1989, Cornelius, et al., 2001

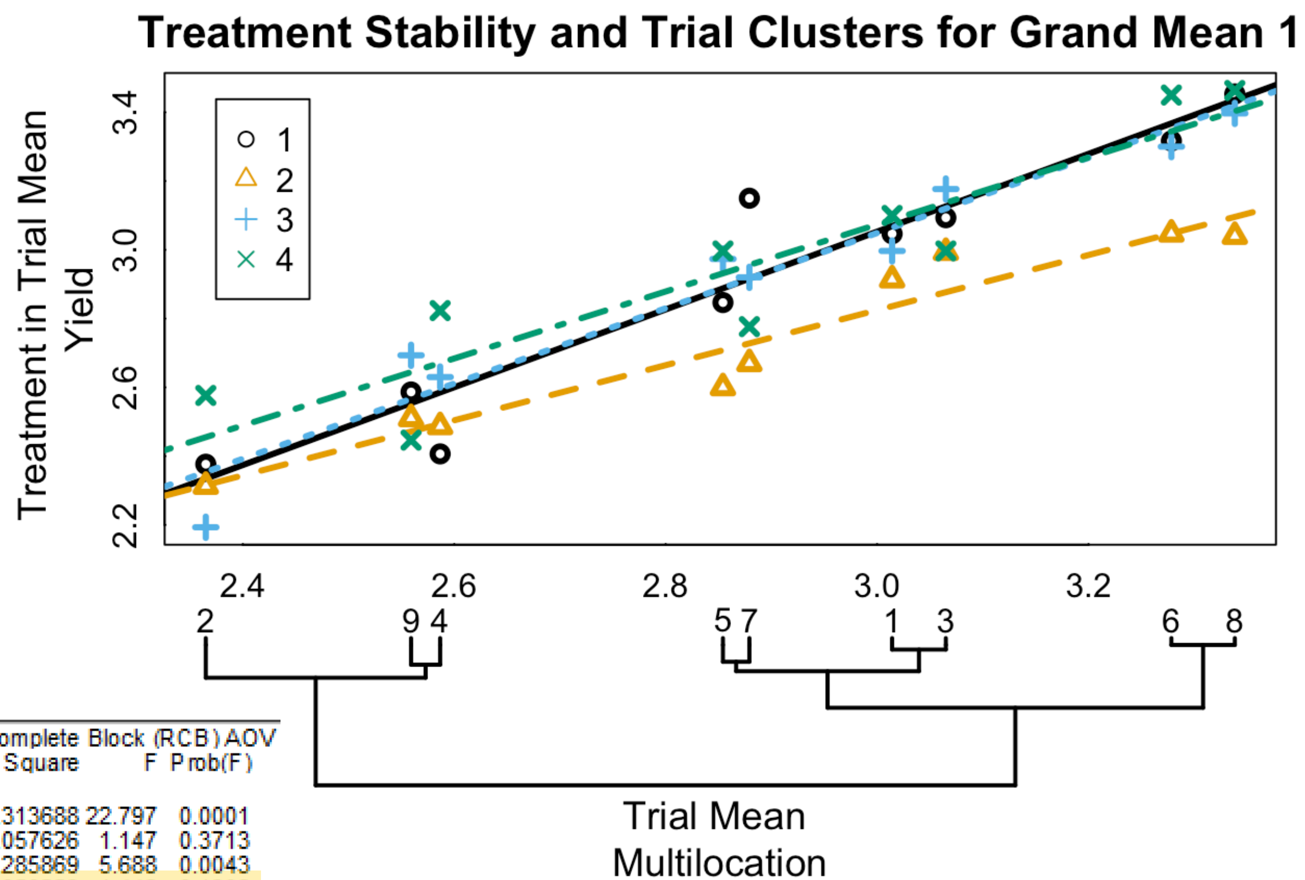
Reference Dendrogram

- To illustrate the effects on interaction on a treatment x trial dendrogram, produce a reference dendrogram based on additive treatment and trial effects only

$$\vec{y}_{\cdot 1} - \vec{y}_{\cdot 2} = \left(\overline{y}_{11\cdot} - \overline{y}_{12\cdot}, \overline{y}_{21\cdot} - \overline{y}_{22\cdot}, \dots, \overline{y}_{n1\cdot} - \overline{y}_{n2\cdot}\right)$$

$$= \begin{pmatrix} \left(\hat{\mu} + \hat{\alpha}_1 + \hat{\beta}_1\right) - \left(\hat{\mu} + \hat{\alpha}_1 + \hat{\beta}_2\right) \\ \left(\hat{\mu} + \hat{\alpha}_2 + \hat{\beta}_1\right) - \left(\hat{\mu} + \hat{\alpha}_2 + \hat{\beta}_2\right) \\ \vdots \\ \left(\hat{\mu} + \hat{\alpha}_n + \hat{\beta}_1\right) - \left(\hat{\mu} + \hat{\alpha}_n + \hat{\beta}_2\right) \end{pmatrix} = \begin{pmatrix} \left(\hat{\beta}_1\right) - \left(\hat{\beta}_2\right) \\ \left(\hat{\beta}_1\right) - \left(\hat{\beta}_2\right) \\ \vdots \\ \left(\hat{\beta}_1\right) - \left(\hat{\beta}_2\right) \end{pmatrix}$$

Simple Additivity

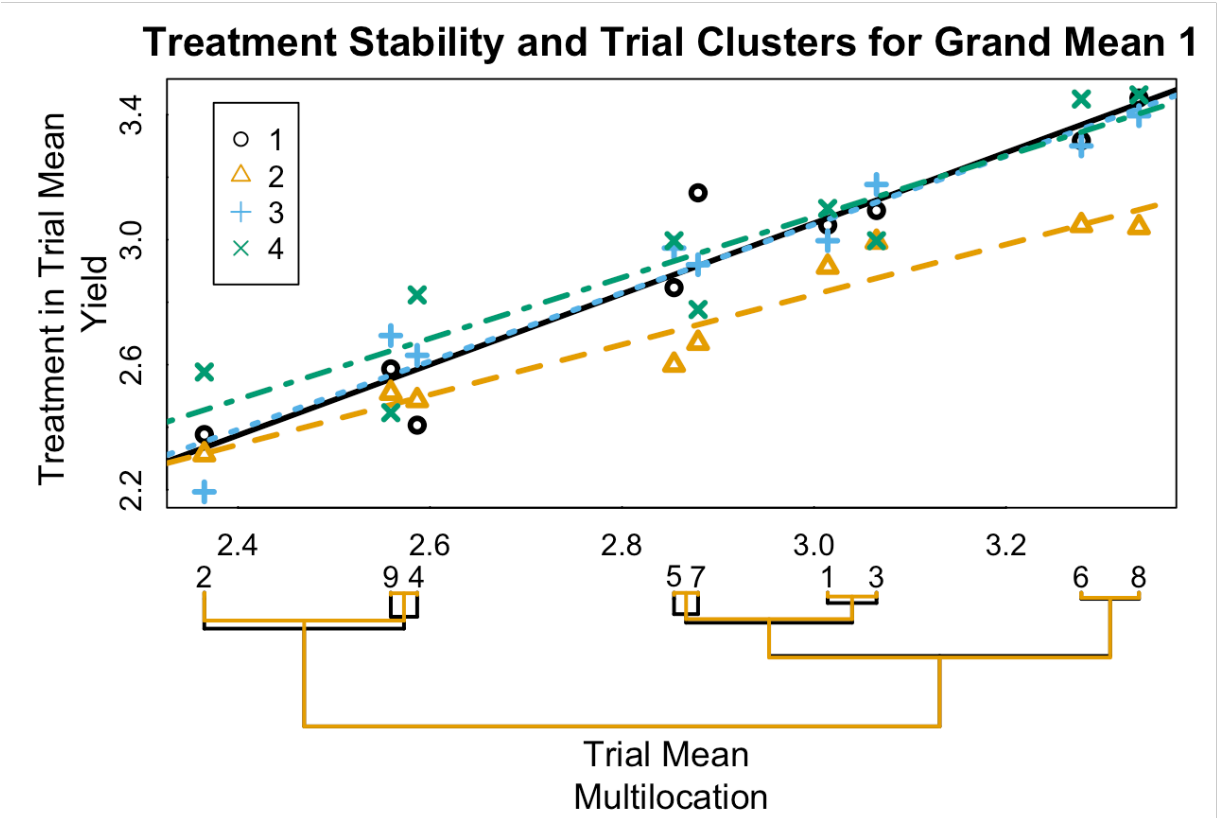


Advanced ST 8 (Fixed Treatment) Randomized Complete Block (RCB) AOV					
Source	DF	Sum of Squares	Mean Square	F	Prob(F)
Total		15.487067			
Trial	8	10.509500	1.313688	22.797	0.0001
Error = Rep in Trial	18	1.037267	0.057626	1.147	0.3713
Treatment	3	0.857607	0.285869	5.688	0.0043
Treatment x Trial	24	1.206226	0.050259	1.446	0.1306
Error/Residual	54	1.876467	0.034749		

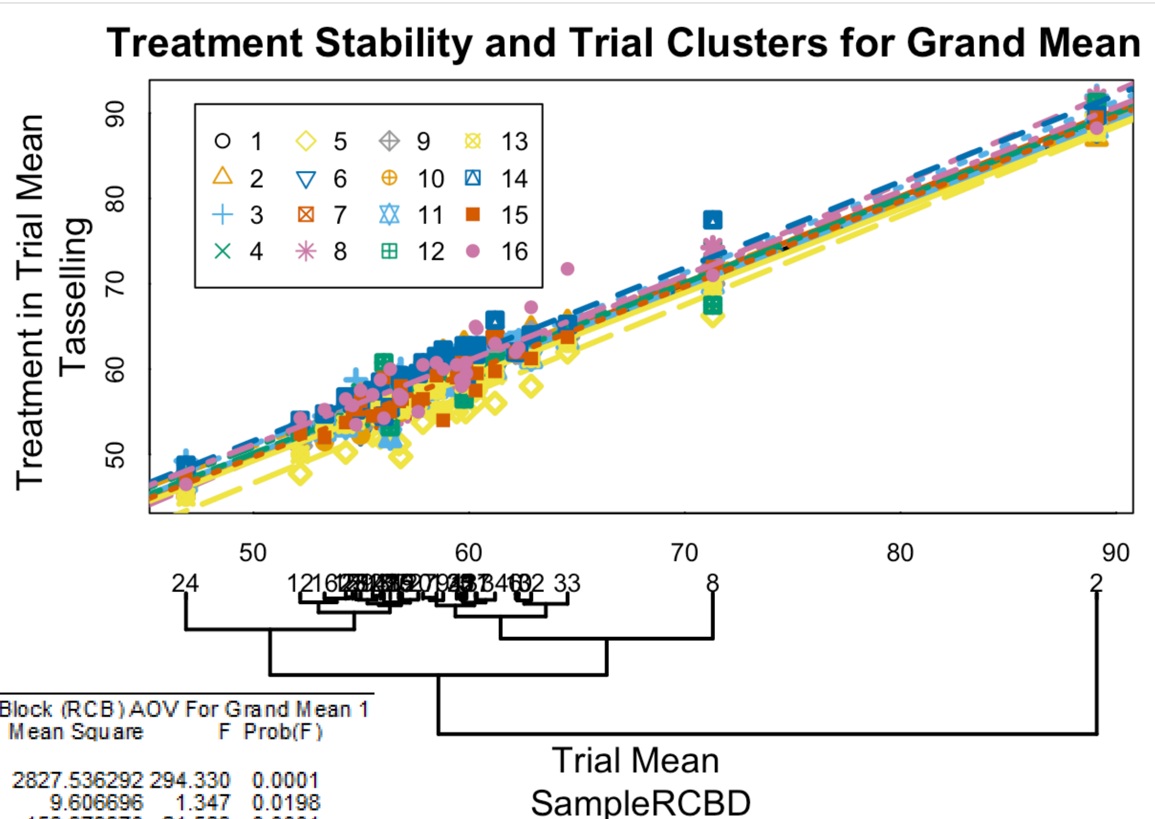
(Data from Littell et al., 1996)

Simple Additivity

Reference dendrogram is nearly identical to the treatment x trial dendrogram



Random Effect



Advanced ST 8 (Fixed Treatment) Randomized Complete Block (RCB) AOV For Grand Mean 1					
Source	DF	Sum of Squares	Mean Square	F	Prob(F)
Total		108926.749554			
Trial	34	96136.233929	2827.536292	294.330	0.0001
Error = Rep in Trial	105	1008.703125	9.606696	1.347	0.0198
Treatment	15	2309.699554	153.979970	21.583	0.0001
Treatment x Trial	510	3638.566071	7.134443	1.926	0.0001
~Treatment x Trial (nonadditivity)	1	5.212079	5.212079	0.730	0.3932
~Residual (nonadditivity)	509	3633.353992	7.138220	1.927	0.0001
=Stability of Slopes	15	72.091774	4.806118	0.667	0.8175
=Residual (nonstability)	495	3566.474297	7.204999	1.945	0.0001
Error/Residual	1575	5833.546875	3.703839		

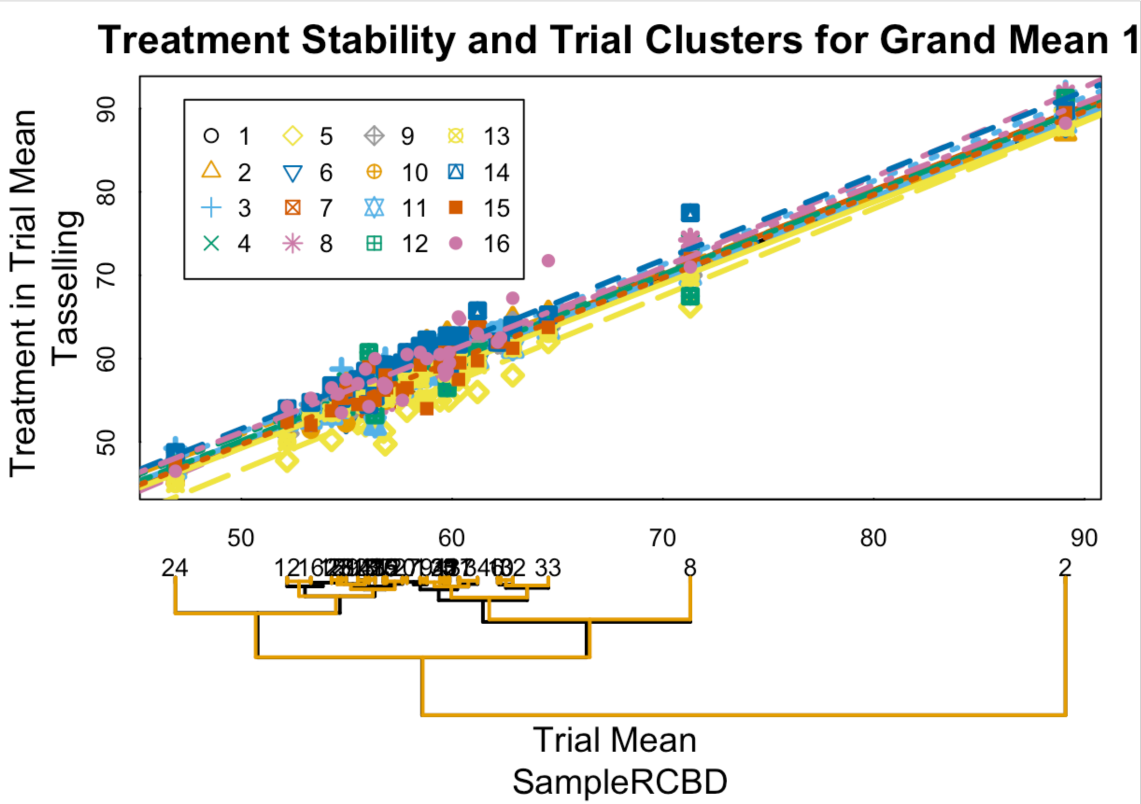
(Data from Vargas, 2013)

Random Effect

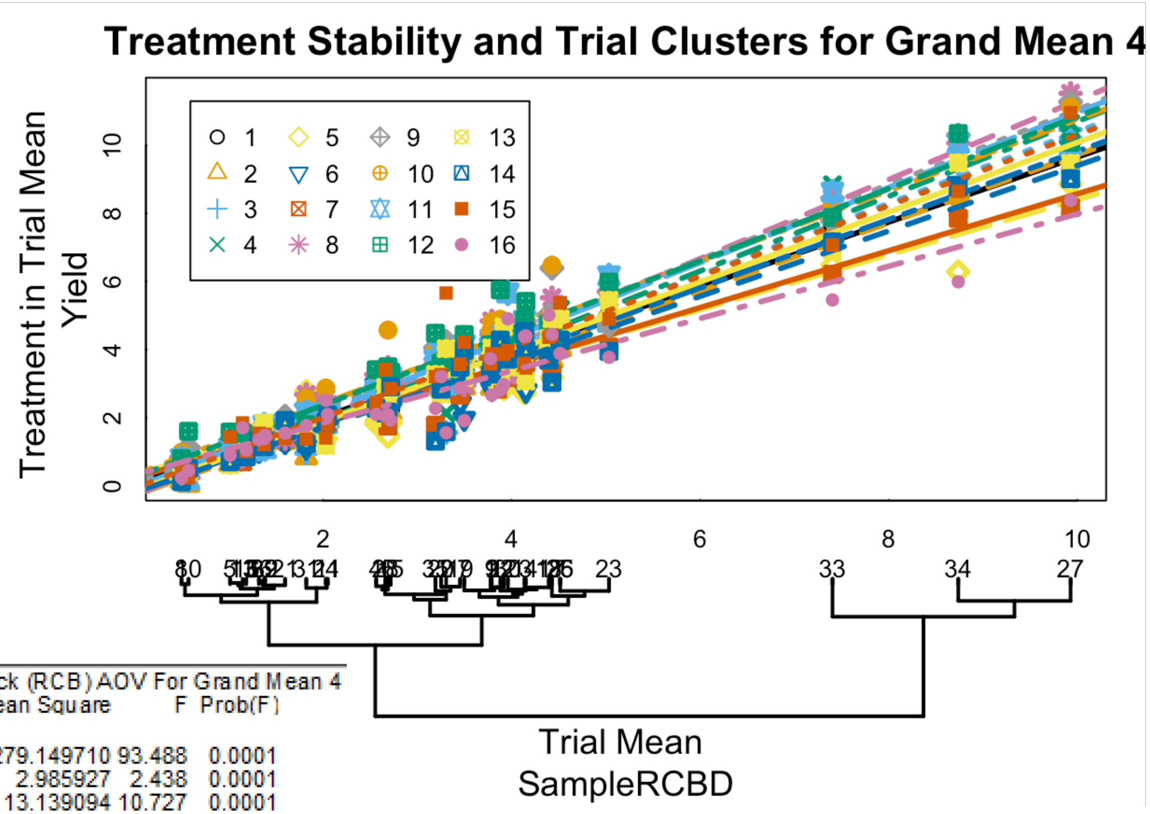
Interaction is present, but interaction effects cancel and dendrogram is preserved

Legend

	Trials	Means
24	SampleRCBDMokwa	46.9
12	SampleRCBDDamongo	52.2
16	SampleRCBDGwoza	53.3
15	SampleRCBDEjura	54.3
23	SampleRCBDKishi	54.6
28	SampleRCBDSaminaka	54.8
29	SampleRCBDSekou	55
14	SampleRCBDEjiba	55.5
18	SampleRCBDIkenne1	55.9
21	SampleRCBDIna	56
31	SampleRCBDWa	56.3
19	SampleRCBDIkenne2	56.8
25	SampleRCBDMokwa2	56.8
26	SampleRCBDNyankpala	56.8
35	SampleRCBDZaria2	56.8
7	SampleRCBDBagou	56.9
20	SampleRCBDIlorin	57.6
27	SampleRCBDOkeoyi	57.9
1	SampleRCBDKatibougou	58.5
9	SampleRCBDBirninKudu	58.8
4	SampleRCBDAngaradebou	59.4
30	SampleRCBDTsafe	59.6
22	SampleRCBDKadawa	59.7
3	SampleRCBDZuru	59.8
13	SampleRCBDDanja	59.8



Nonadditivity



Advanced ST 8 (Fixed Treatment) Randomized Complete Block (RCB) AOV For Grand Mean 4

Source	DF	Sum of Squares	Mean Square	F	Prob(F)
Total		12038.020315			
Trial	35	9770.239863	279.149710	93.488	0.0001
Error = Rep in Trial	108	322.480080	2.985927	2.438	0.0001
Treatment	15	197.086417	13.139094	10.727	0.0001
Treatment x Trial	525	643.032719	1.224824	1.795	0.0001
~Treatment x Trial (nonadditivity)	1	70.179638	70.179638	64.195	0.0001
~Residual (nonadditivity)	524	572.853081	1.093231	1.602	0.0001
=Stability of Slopes	15	114.296003	7.619734	7.350	0.0001
=Residual (nonstability)	510	528.736716	1.036739	1.520	0.0001
Error/Residual	1620	1105.181236	0.682211		

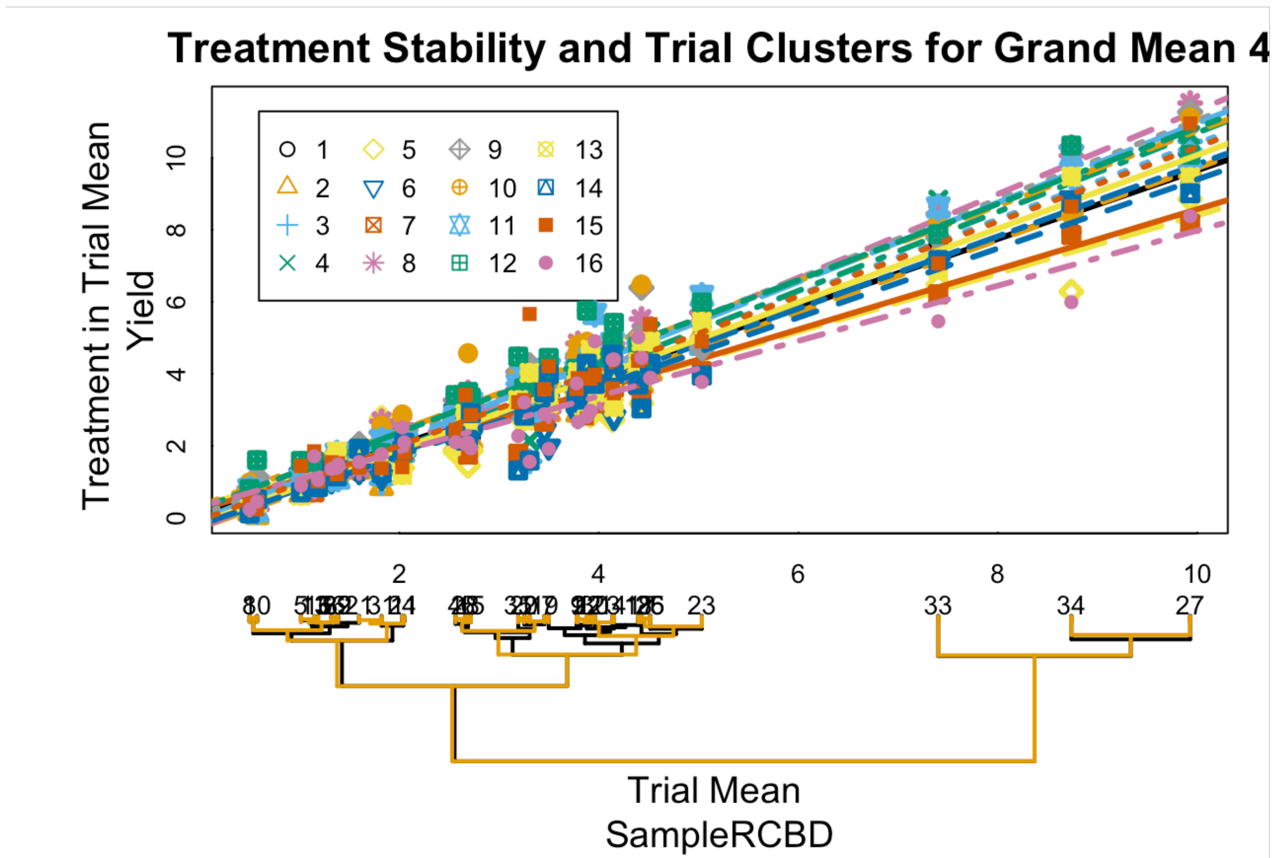
(Data from Vargas et al., 2013)

Nonadditivity

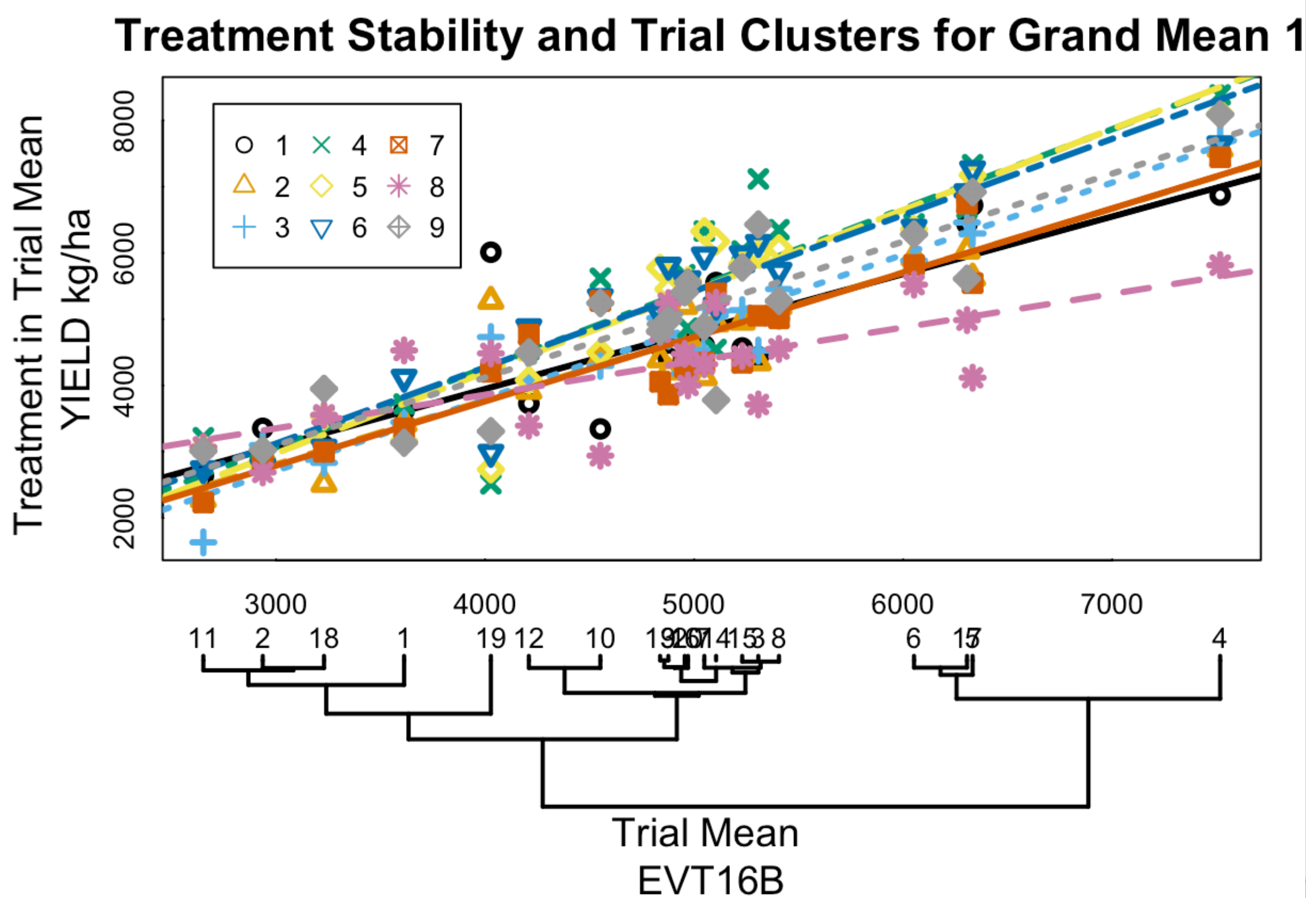
Major clusters, relative to trial means, are preserved

Legend

	Trials	Means
8	SampleRCBDBallah	0.5
10	SampleRCBDDalabani	0.6
5	SampleRCBDBadeggi	1
13	SampleRCBDDanja	1.1
16	SampleRCBDGwoza	1.2
36	SampleRCBDZaria2	1.3
29	SampleRCBDSekou	1.4
32	SampleRCBDWa	1.4
21	SampleRCBDIna	1.6
31	SampleRCBDTumu	1.8
11	SampleRCBDDamboa	2
24	SampleRCBDMokwa	2
4	SampleRCBDAngaradebou	2.6
6	SampleRCBDBagauda	2.7
15	SampleRCBDEjura	2.7
28	SampleRCBDSaminaka	2.7
35	SampleRCBDZaria1	3.2
2	SampleRCBDKetou	3.3
20	SampleRCBDIlorin	3.3
7	SampleRCBDBagou	3.5
19	SampleRCBDIkenne2	3.5
1	SampleRCBDKatibougou	3.8
9	SampleRCBDBirninKudu	3.8
12	SampleRCBDDamongo	3.9
22	SampleRCBDKadawa	3.9
30	SampleRCBDTsafe	4
14	SampleRCBDEjiba	4.1



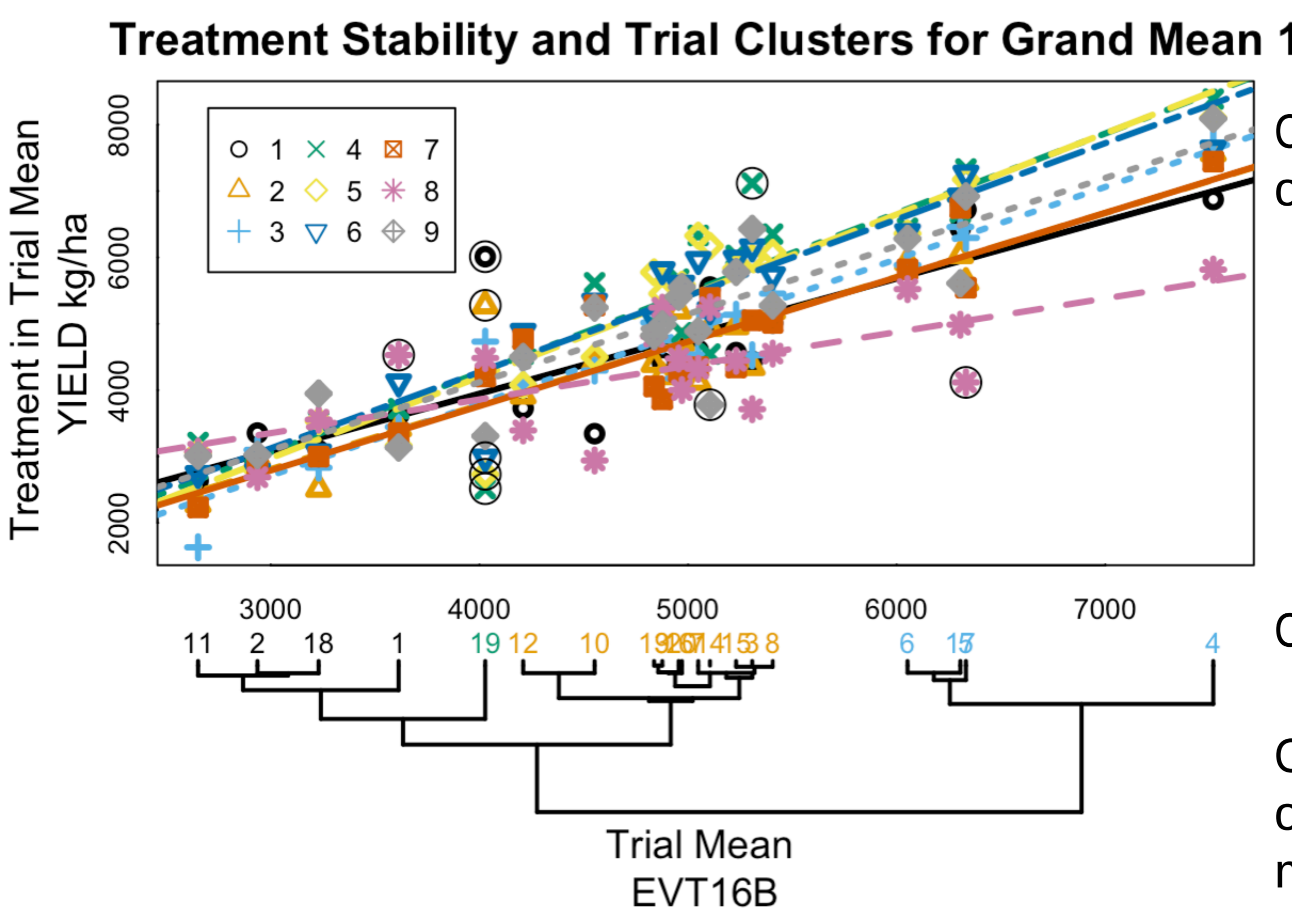
Complex Interaction



Legend		
	Trials	Means
11	EVT16B9	2652.22248744444
2	EVT16B3	2936.44473808888
18	EVT16B17	3229.44476738888
1	EVT16B1	3612.66702793334
19	EVT16B18	4027.77818055556
12	EVT16B11	4209.44486538888
10	EVT16B8	4551.44489958888
13	EVT16B12	4837.44492818888
9	EVT16B6	4876.77826545556
16	EVT16B15	4955.11160662222
20	EVT16B20	4970.66716373334
7	EVT16B4	5048.88939377778
14	EVT16B13	5105.11162162222
15	EVT16B14	5230.11163412222
3	EVT16B19	5307.22275294444
8	EVT16B5	5406.0005406
6	EVT16B2	6052.11171632222
17	EVT16B16	6305.44507498888
5	EVT16B7	6332.33396656666
4	EVT16B10	7516.66741833333

from Cornelius et al., 1997)

Complex Interaction

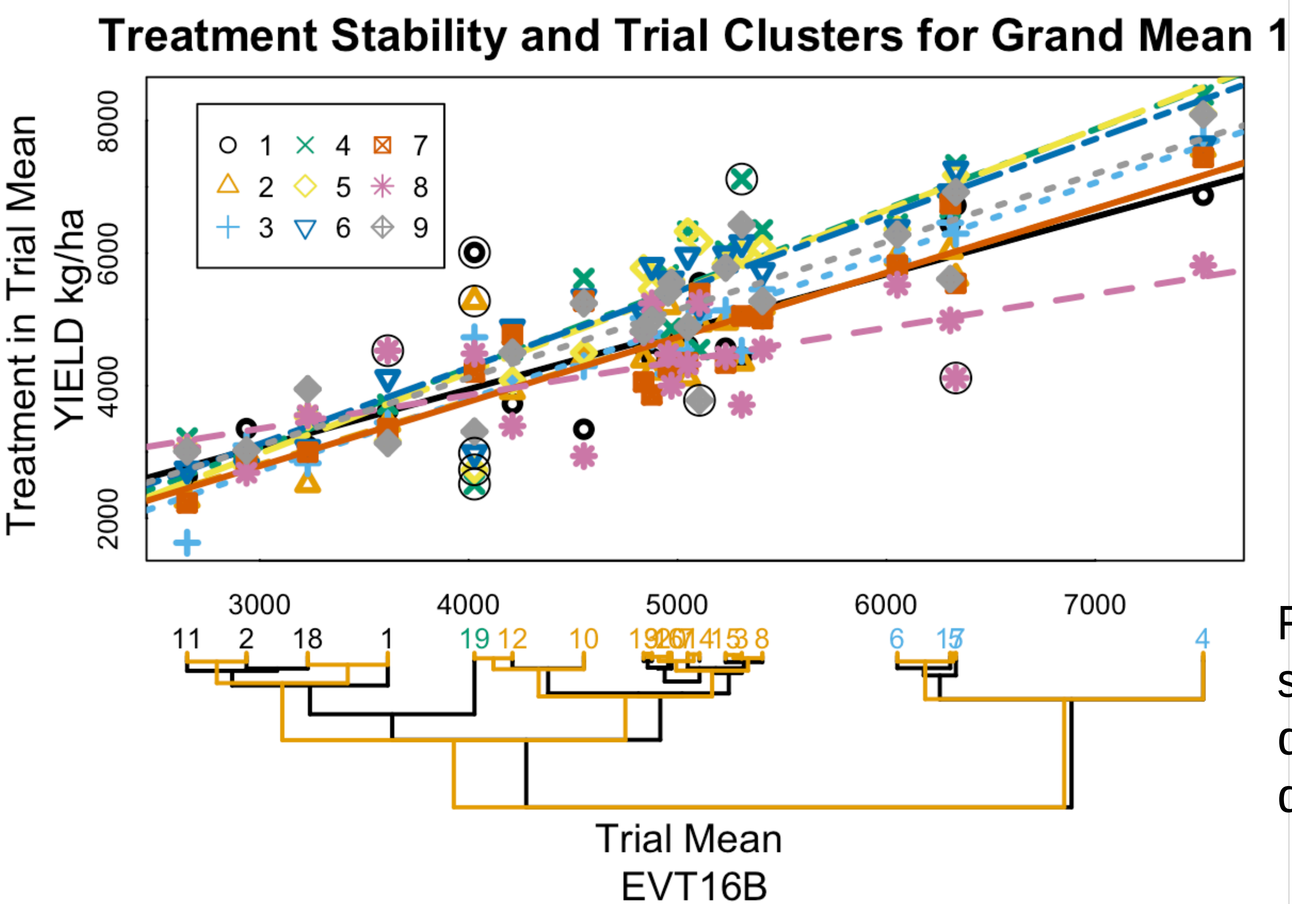


Circle the interaction outliers (2 s.d.)

Colors identify clusters

One trial that forms its own cluster contains most of the outliers

Complex Interaction



Reference dendrogram shows how interaction distorts treatment x trial dendrogram

Future Directions

- Different decompositions
  - Linear-bilinear models (i.e. AMMI, SHMM)
- Clustering methods and comparison of dendrograms
- Automated analysis of trial groups
- Covariate analysis
  - Use trial specific data to explain interaction
  - Geospatial maps

Thank You

- References
  - www.gdmdata.com
  - J Tukey. One degree of freedom for non-additivity. Biometrics, 5(3):232–242, 1949
  - J Mandel. Non-additivity in the two-way analysis of variance. Journal of the American Statistical Association, 56(296), 1961.
  - G Milliken, and D Johnson. *Analysis of messy data volume 2: nonreplicated experiments*. Vol. 1. CRC Press, 1989.
  - P Cornelius, J Crossa, M Seyedsadr, G Liu, and K Viele. Contributions to multiplicative model analysis of genotype-environment data. Proceedings of the Annual Meeting of the American Statistical Association, 2001
  - R Littell, G Milliken, W Stroup, and R Wolfinger. *SAS system for mixed models*. SAS Institute, Cary, NC., 1996.
  - P Cornelius, J Crossa, and M Seyedsadr. Statistical tests and estimators of multiplicative models for genotype-by-environment interaction. In. M Kang and H Gauch. *Genotype-by-environment interaction*. CRC Press, 1996.
  - M Vargas, E Combs, G Alvarado, G Atlin, K Mathews, and J Crossa. META: A suite of SAS programs to analyze multienvironment breeding trials. Agronomy Journal, 105(1):11, 2013
  - http://igrow.org/agronomy/wheat/winter-wheat-variety-trial-results/