GEneralized **M**ultiplicative **AN**alysis **O**f **VA**riance (GEMANOVA)

of kill kinetics data of antibacterial agents

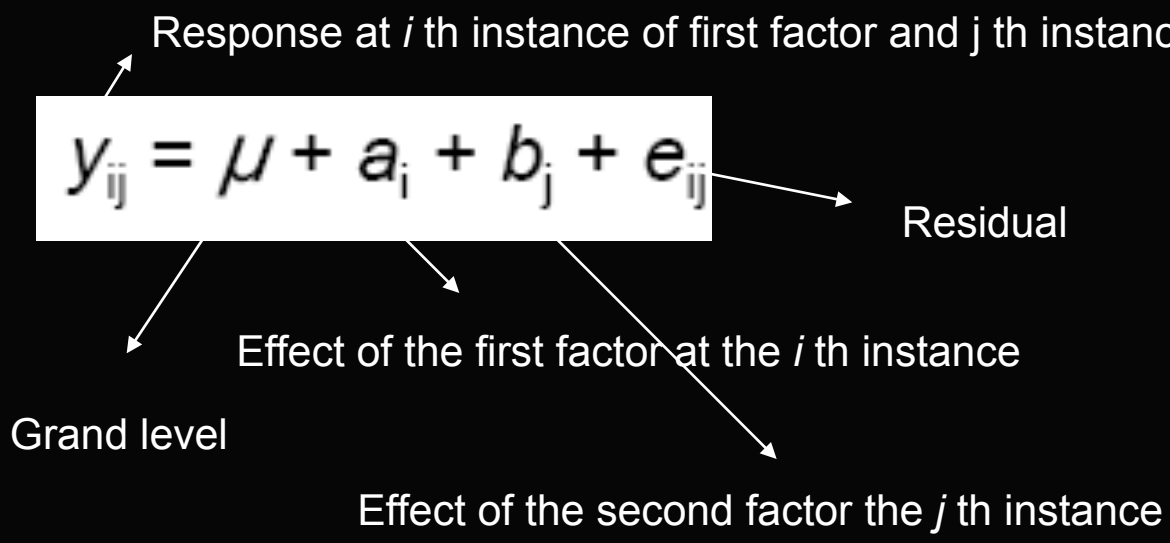
Professor D Brynn Hibbert

GEMANOVA

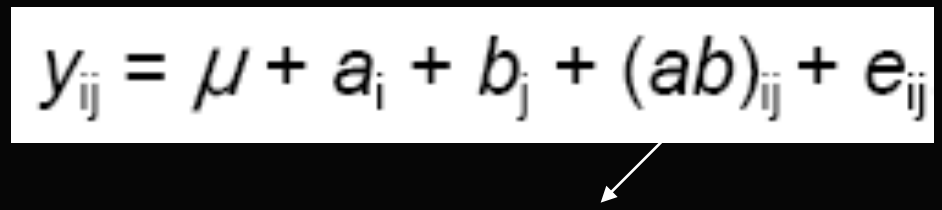
- Used to analyze the variance
- Very good for systems with complex interactions
- Works on experimentally designed data
- Uses a PARAFAC algorithm

✓ An example from the bactericidal effect of novel antibacterial agents

Qualitative ANOVA



1 + I + J parameters



1 + I + J + IJ parameters

Quantitative ANOVA

$$y_{ij} = b_0 + b_1 x_{1i} + b_2 x_{2j} + b_{12} x_{1i} x_{2j} + e_{ij}$$

Value of the first factor at the *i* th level

Value of the second factor at the *j* th level

x1	x2	x1x2	y
+1	+1	+1	√
+1	-1	-1	√
-1	+1	-1	√
-1	-1	+1	√

Only parameters b_0, b_1, b_2 and b_{12} are to be estimated

Qualitative *vs* Quantitative ANOVA

Qualitative ANOVA:

- Quite flexible therefore prone to overfit especially for interactions

Quantitative ANOVA:

- Very restricted

"Although both models can theoretically handle the interaction of any order and complexity, data that are **mainly generated by interactions** are difficult to be modeled adequately by ANOVA."

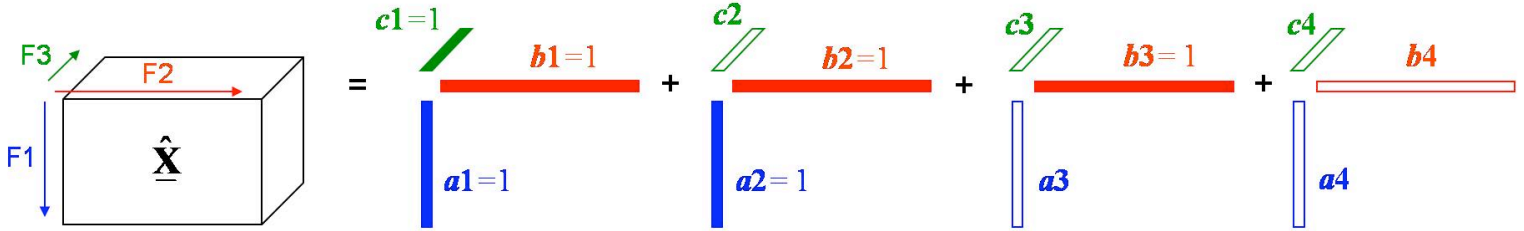
R. Bro, M. Jakobsen, J. Chemometr. 16 (2002) 294–304.

GEMANOVA: a constrained PARAFAC model

Three factors F1, F2 and F3 are modeled using GEMANOVA terms.

Unity vectors are introduced to obtain the terms of interest.

PARAFAC



GEMANOVA

$$\hat{x}_{ijk} = m + c2_k + a3_i c3_k + a4_i b4_j c4_k$$

$$x_{ij} = \mu + a_i + b_j + \sum_{r=1}^{R} c_{ir} d_{jr} + e_{ij}$$

GEMANOVA

$$y_{ij} = \mu + a_{i1} + b_{j1} + \sum_{f=1}^{F} c_{if} d_{jf} + e_{ij}$$

Main effects Interactions

Main effect:

a_i	ANOVA	qualitative
bx_i	ANOVA	quantitative
a_i	GEMANOVA	

Three-way interaction:

			No. of parameters
$(abc)_{ijk}$	ANOVA	qualitative	IJK
$bx_i x_j x_k$	ANOVA	quantitative	1
$a_i b_j c_k$	GEMANOVA	one-component	$I+J+K$
$\sum_{f=1}^{F} a_{if} b_{jf} c_{kf}$	GEMANOVA	F-component	

R. Bro, M. Jakobsen, J. Chemometr. 16 (2002) 294–304.

Bactericidal effect of natural gum active ingredients

5 substances of	*Bm*, *B*, *R*, *S* and *T*
in 2 structural shapes of	monomer and polymer
2 oxidation forms of	oxidized and non-oxidized
2 concentrations (MIC)	1 and 5 mg L^{-1}
on 3 genera of bacteria	*E. coli*, *H. pylori* and *S. aureus*

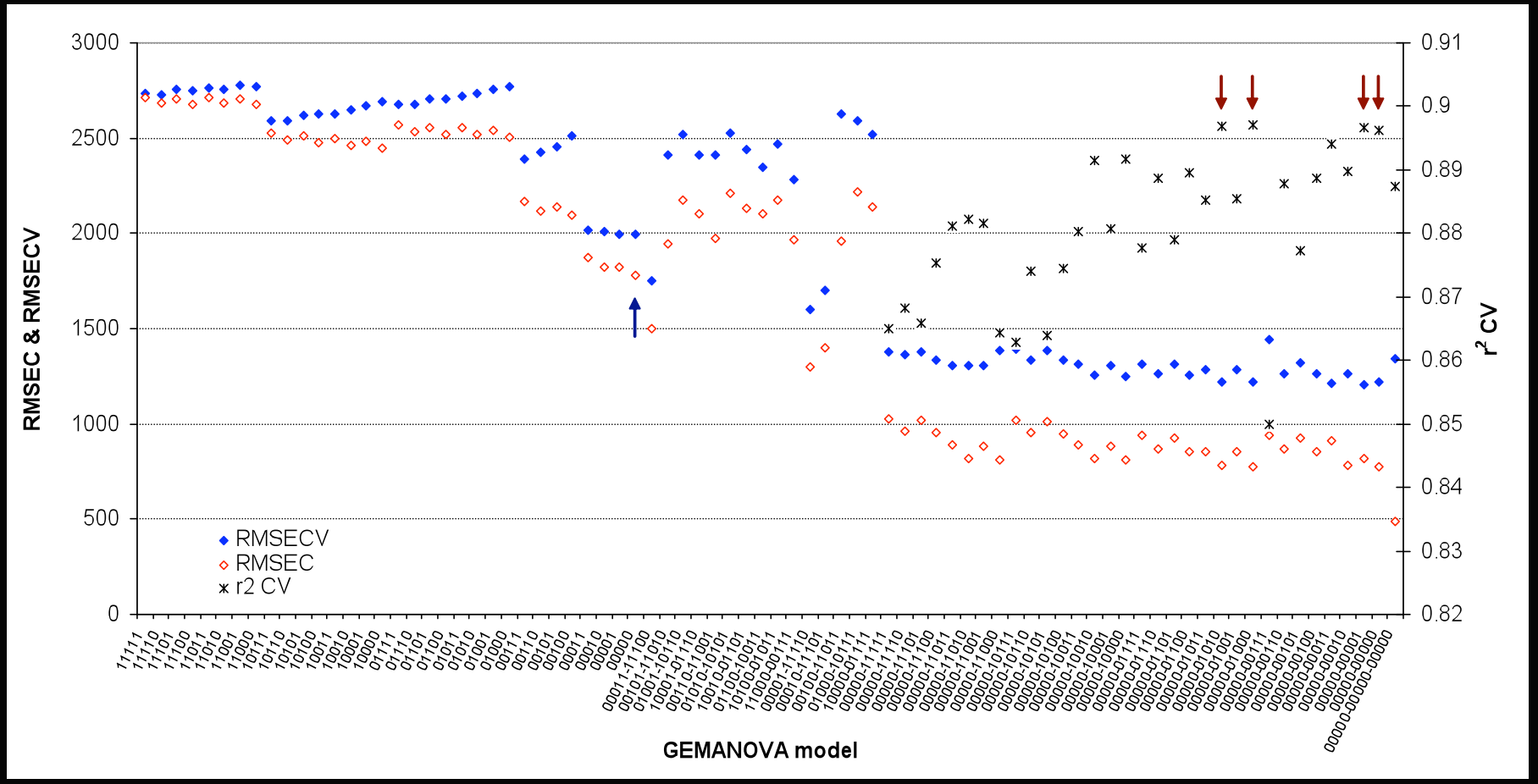
ANOVA

Source	No. of parameters in ANOVA
Substance	5
Bacteria	3
Structure shape	2
Oxidationform	2
Concentration	2
SubstanceBacteria	5 3
SubstanceStructure shape	5 2
SubstanceOx form	5 2
SubstanceConc	5 2
BacteriaStructure shape	3 2
BacteriaOx form	3 2
BacteriaConc	3 2
Structure shapeOx form	2 2
Structure shapeConc	2 2
Ox formConc	2 2
SubstanceBacteriaStructure shape	5 3 2
SubstanceBacteriaOx form	5 3 2
SubstanceBacteriaConc	5 3 2
SubstanceStructure shapeOx form	5 2 2
SubstanceStructure shapeConc	5 2 2
SubstanceOx formConc	5 2 2
BacteriaStructure shapeOx form	3 2 2
BacteriaStructure shapeConc	3 2 2
BacteriaOx formConc	3 2 2
Structure shapeOx formConc	2 2 2
SubstanceBacteriaStructure shapeOx form	5 3 2 2
SubstanceBacteriaStructure shapeConc	5 3 2 2
SubstanceBacteriaOx formConc	5 3 2 2
SubstanceStructure shapeOx formConc	5 2 2 2
BacteriaStructure shapeOx formConc	3 2 2 2
SubstanceBacteriaStructure shapeOx formConc	5 3 2 2 2
Total number of estimated parameters:	**647**

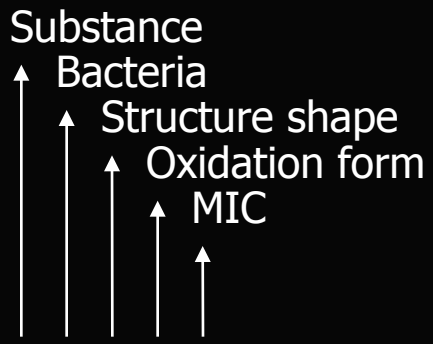
ANOVA results

Source	Replicates 1 & 2	Replicates 1 & 3	Replicates 2 & 3
Substance	Sig	Sig	Sig
Bacteria	Sig	Sig	Sig
Structure shape	Non sig	Non sig	Non sig
Ox form	Sig	Sig	Sig
MIC	Sig	Sig	Sig
Substance*Bacteria	Sig	Sig	Sig
Substance*Structure shape	Sig	Sig	Sig
Substance*Ox form	Sig	Sig	Sig
Substance*MIC	Non sig	Sig	Non sig
Bacteria*Structure shape	Sig	Sig	Sig
Bacteria*Ox form	Sig	Sig	Sig
Bacteria*MIC	Sig	Sig	Sig
Structure shape*Ox form	Sig	Sig	Sig
Structure shape*MIC	Non sig	Non sig	Non sig
Ox form*MIC	Sig	Non sig	Sig
Substance*Bacteria*Structure shape	Sig	Sig	Sig
Substance*Bacteria*Ox form	Sig	Sig	Sig
Substance*Bacteria*MIC	Non sig	Sig	Sig
Substance*Structure shape*Ox form	Sig	Sig	Sig
Substance*Structure shape*MIC	Non sig	Sig	Non sig
Substance*Ox form*MIC	Non sig	Sig	Sig
Bacteria*Structure shape*Ox form	Sig	Sig	Sig
Bacteria*Structure shape*MIC	Non sig	Sig	Non sig
Bacteria*Ox form*MIC	Sig	Sig	Sig
Structure shape*Ox form*MIC	Non sig	Non sig	Sig
Substance*Bacteria*Structure shape*Ox form	Sig	Sig	Sig
Substance*Bacteria*Structure shape*MIC	Non sig	Sig	Non sig
Substance*Bacteria*Ox form*MIC	Sig	Sig	Sig
Substance*Structure shape*Ox form*MIC	Non sig	Non sig	Sig
Bacteria*Structure shape*Ox form*MIC	Sig	Non sig	Sig
Substance*Bacteria*Structure shape*Ox form*MIC	Sig	Sig	Sig

GEMANOVA models



Best GEMANOVA models



	No. of parameters
$y_{ijklm} = a_i b_j c_k d_l f_m + g_i h_j s_k t_l u_m$	(5+3+2+2+2)+(5+3+2+2+2)=28
$y_{ijklm} = a_i b_j c_k d_l f_m + g_i h_j s_k t_l$	(5+3+2+2+2)+(5+3+2+2+0)=26
$y_{ijklm} = a_i b_j c_k d_l f_m + g_i s_k t_l u_m$	(5+3+2+2+2)+(5+0+2+2+2)=25
$y_{ijklm} = a_i b_j c_k d_l f_m + g_i s_k u_m$	(5+3+2+2+2)+(5+0+2+0+2)=**23**

GEMANOVA
Model parameters

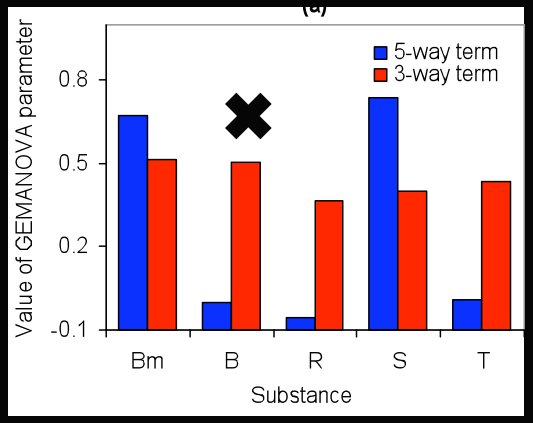
Estimated response due to the incubation of *H. pylori* with substance *B* in its polymer and oxidized form and in 5 MIC is:

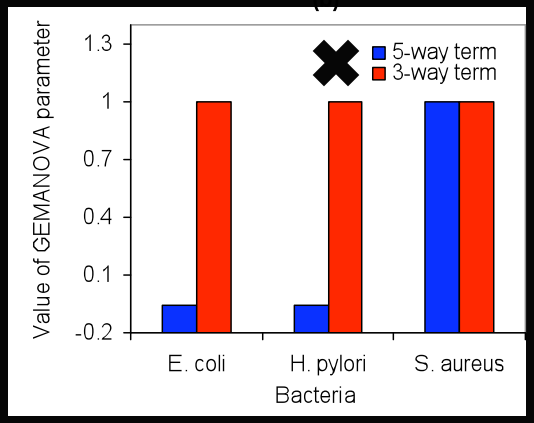
= (0.0011 × -0.0588 × 0.9996 × 0.5832 × 1.8368×10^{4}) + (0.5121 × 1.0000 × 0.5114 × 1.0000 × 0.5501×10^{4}) =

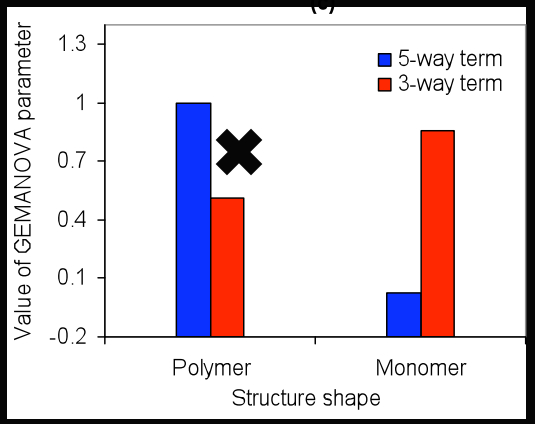
-0.6926 + 1415.0453 = 1414.3527

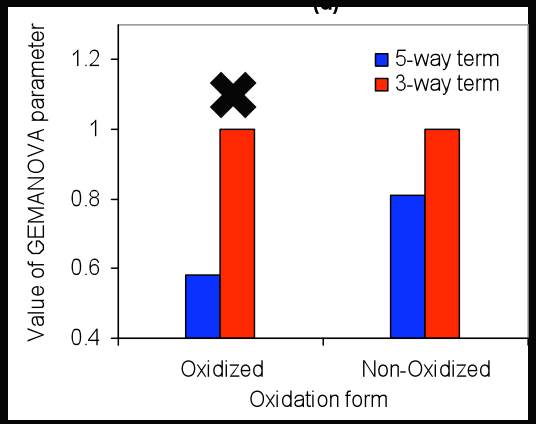
5-way << 3-way

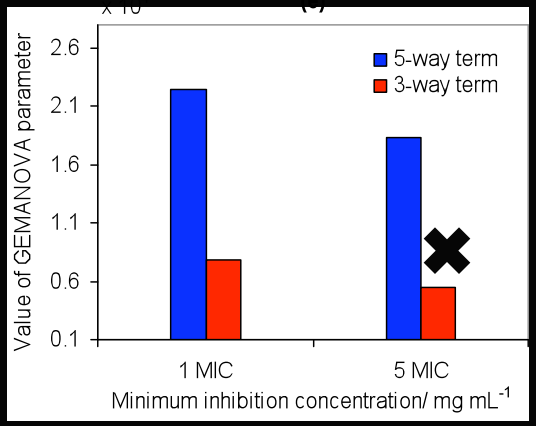
$$y_{ijklm} = a_i b_j c_k d_l f_m + g_i s_k u_m$$











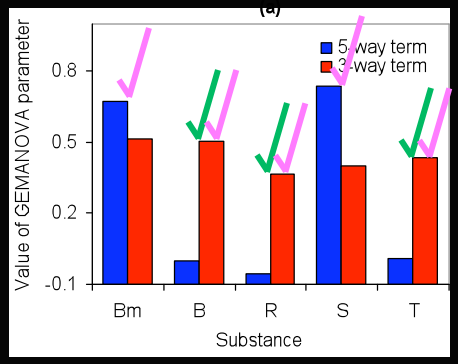
GEMANOVA model interpretation

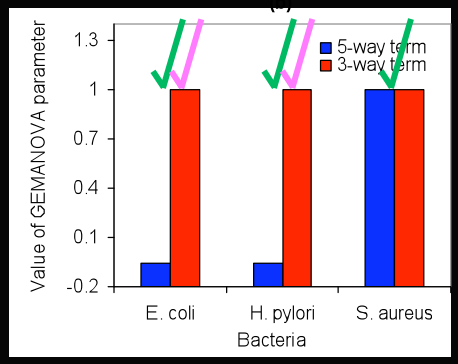
1. Substances ***B***, ***R*** and ***T*** kill all three bacteria better when they are in polymer shape and 5 MIC. Oxidation form does not have an impact on their activity.

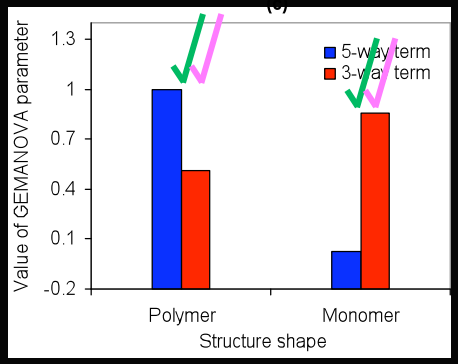
$$y_{ijklm} = a_i b_j c_k d_l f_m + g_i s_k u_m \quad (0)$$

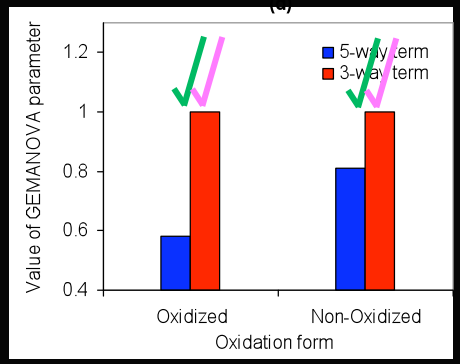
2. Bacteria of ***E. coli*** and ***H. pylori*** are more vulnerable when substances (all five) are in the polymer form and 5 MIC. *R* and *B* are the most and least effective substances, respectively for these two bacteria whose sensitivity are not affected by oxidation form of bactericidals.

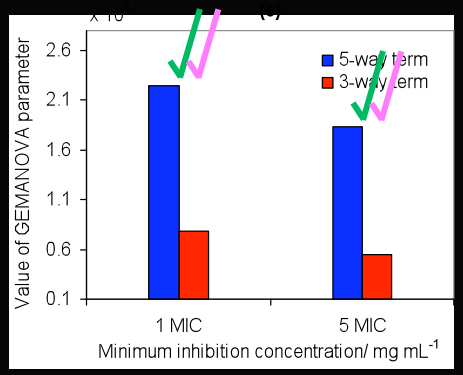
$$y_{ijklm} = a_i b_j c_k d_l f_m + g_i s_k u_m \quad (0)$$











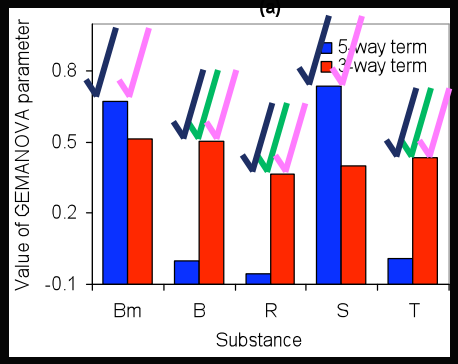
GEMANOVA model interpretation

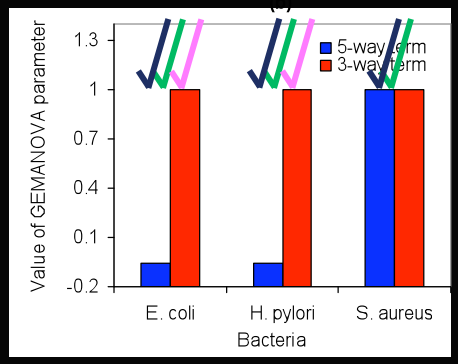
1. Substances ***B***, ***R*** and ***T***
2. Bacteria of ***E. coli*** and ***H. pylori***

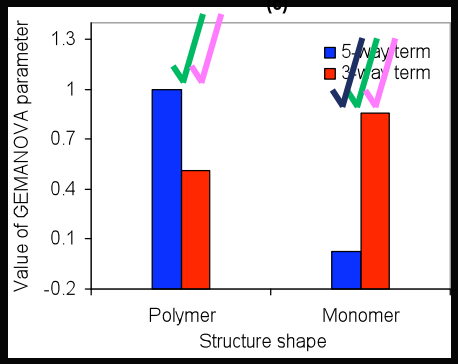
$$y_{ijklm} = a_i b_j c_k d_l f_m + g_i s_k u_m \quad (0)$$

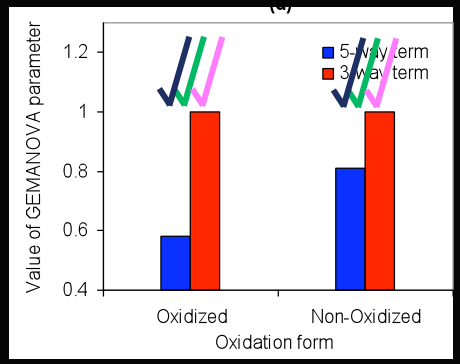
3. The **monomer shapes** of all substances exhibit the same activity on three studied bacteria. While oxidation form does not have an impact on the effectiveness of monomer bactericidal, the greater activity is observed in higher concentrations (5 MIC).

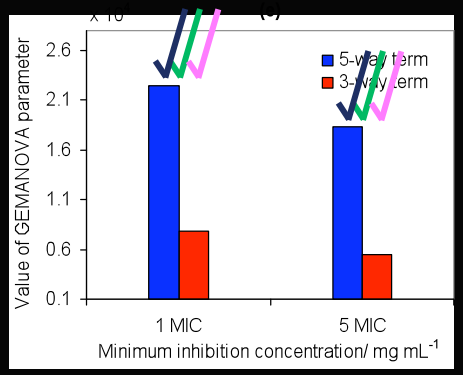
$$y_{ijklm} = a_i b_j c_k d_l f_m + g_i s_k u_m \quad (0)$$





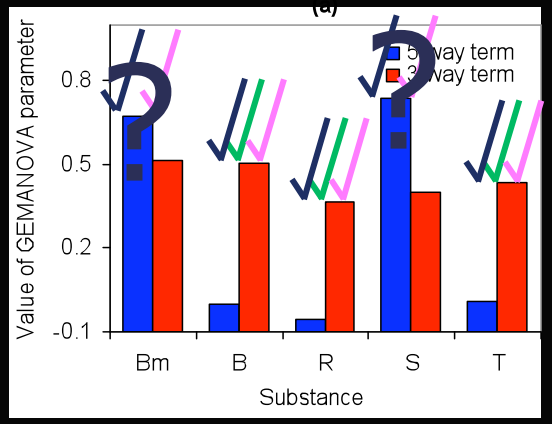


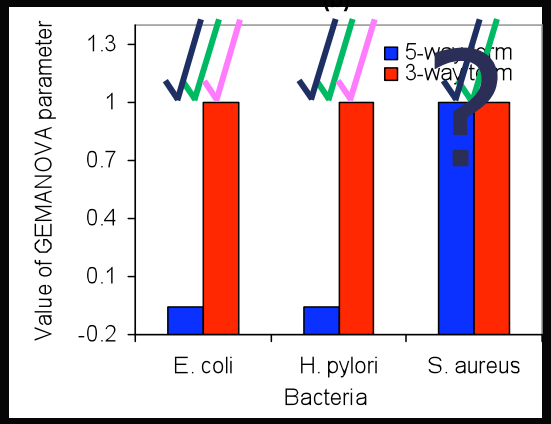


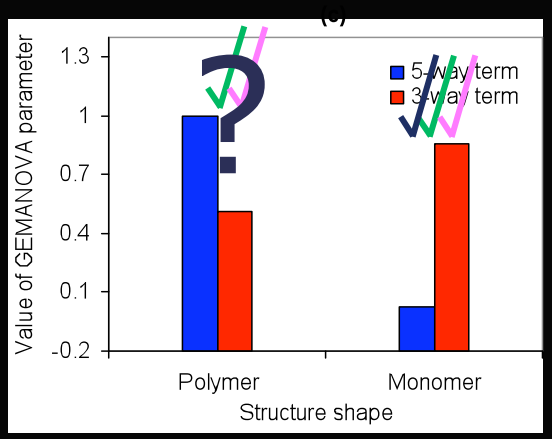


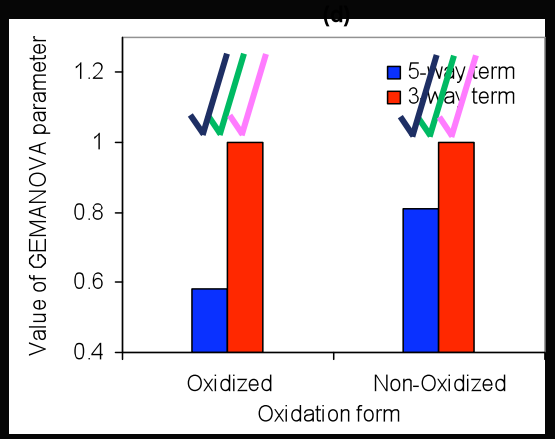
GEMANOVA model interpretation

The variance of data due to the effects of *B*, *R* and *T* in two oxidation forms and two MICs, the effects of *E. coli* and *H. pylori* and the effect of monomer is described by a three-way GEMANOVA term.

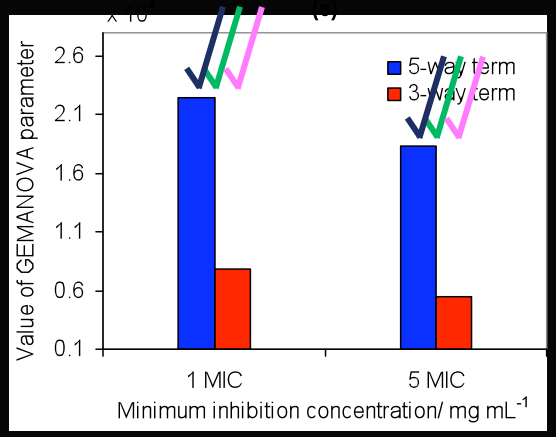








What is not explained solely by this three-way term is the effect of *Bm* and *S* in the polymer shape on *S.aureus* in two oxidation forms and two concentrations.

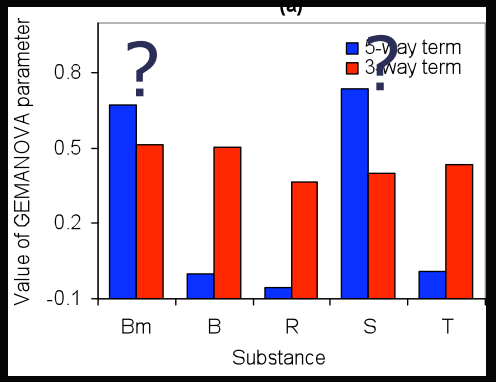


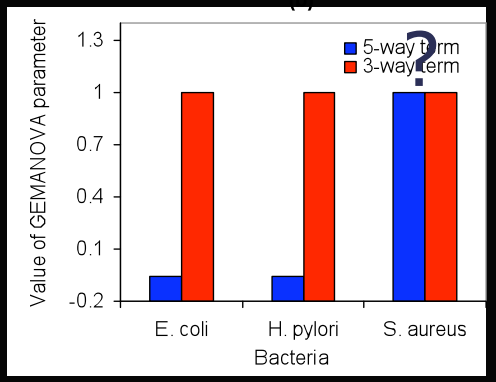
GEMANOVA model interpretation

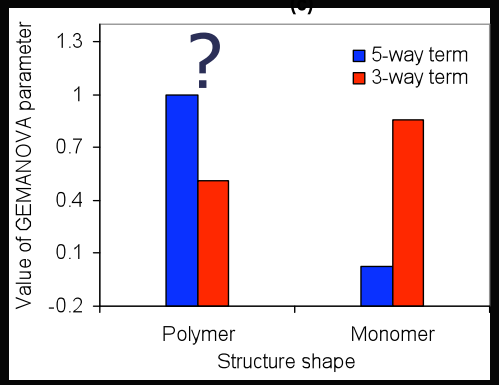
8 (2x1x1x2x2) data points out of 120 (5×3×2×2×2)

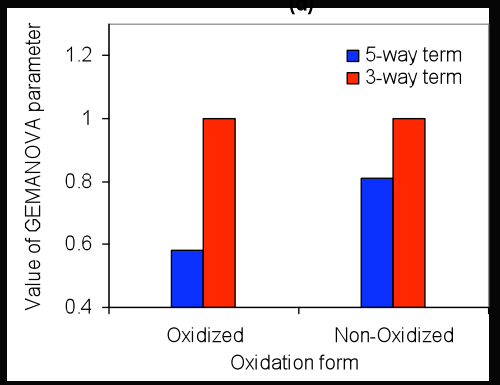
$$y_{ijklm} = a_i b_j c_k d_l f_m + g_i s_k u_m$$

5-way > 3-way

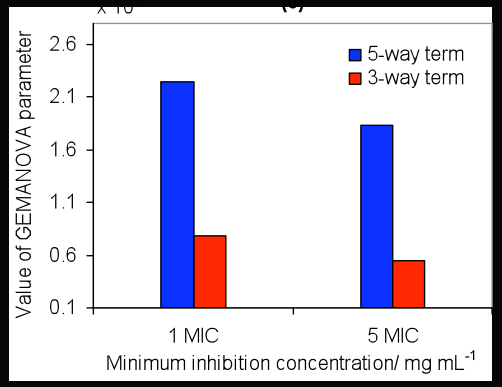








1. Monomer form of *Bm* and *S* kill *S.aureus* better
2. Oxidized form of *Bm* and *S* kills *S.aureus* better.
3. *Bm* and *S* kill *S.aureus* better at higher concentration (5 MIC).



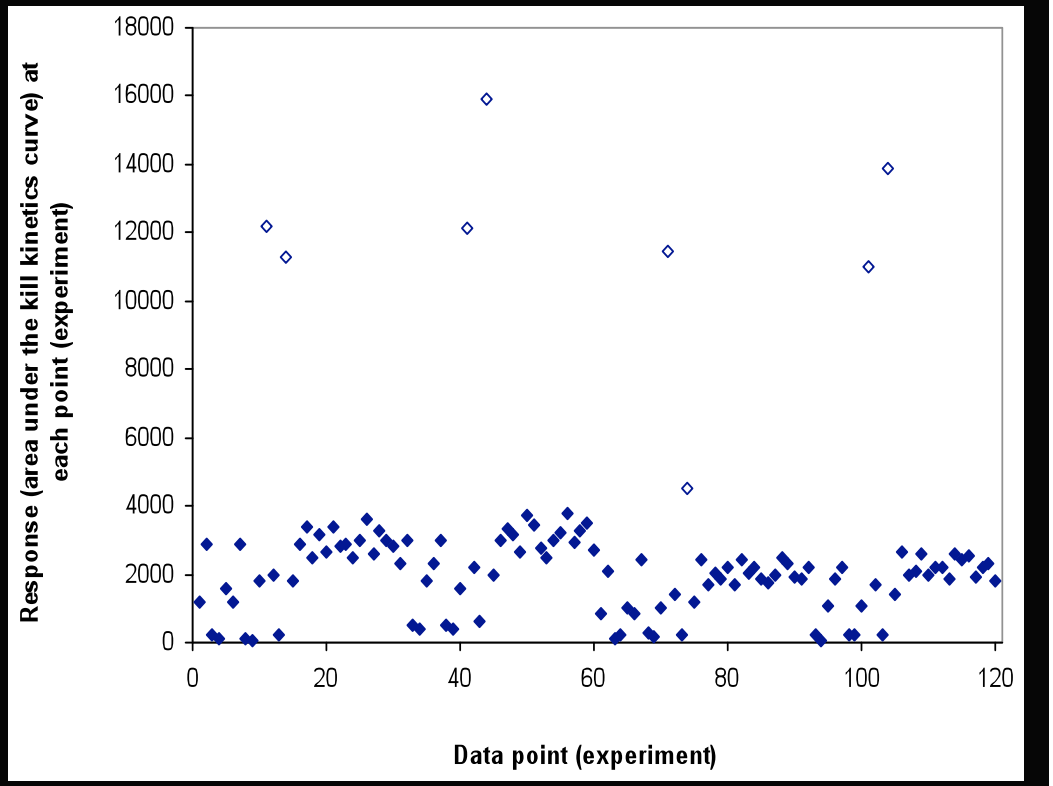
What is special about those eight data points?

$$y_{ijklm} = a_i b_j c_k d_l f_m + g_i s_k u_m$$

Polymer form kills gram positives less

$$y_{ijklm} = a_i b_j c_k d_l f_m + g_i s_k u_m$$

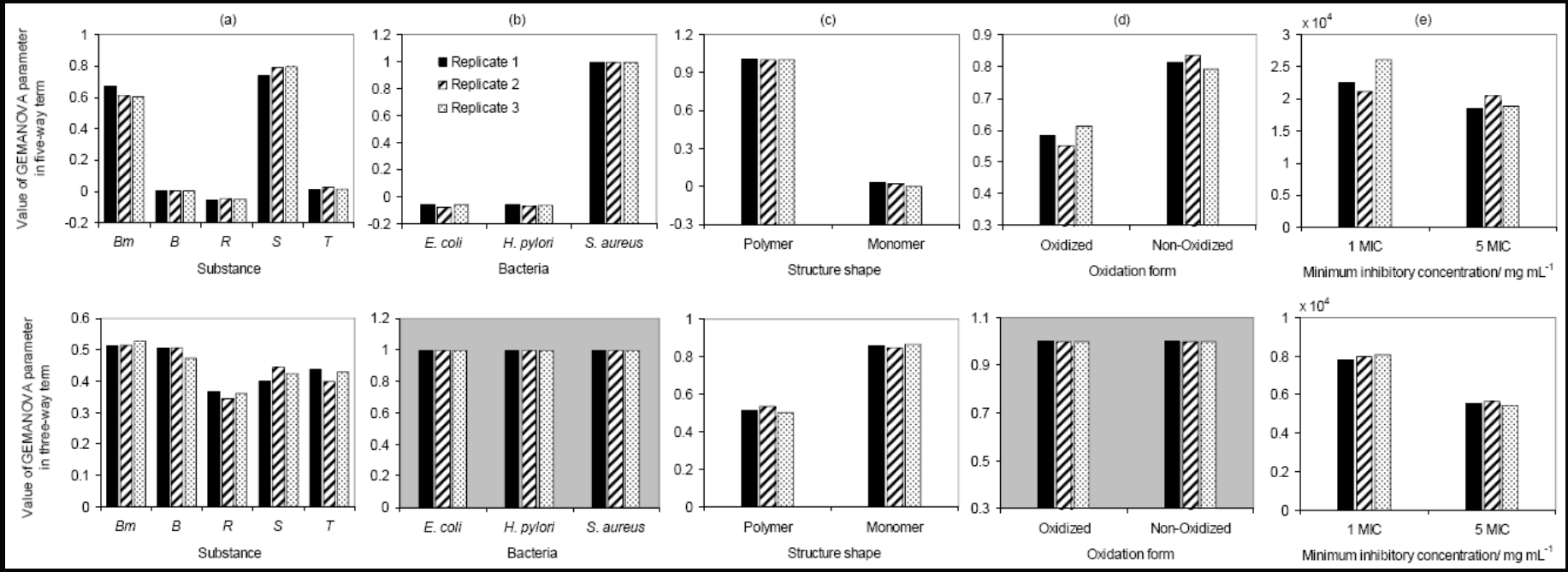
Polymer form kills gram negative and gram positives more



That is why the effect of **structure shape** was insignificant by ANOVA!

Robustness of GEMANOVA models

Three independent repeats with significant noise



GEMANOVA vs ANOVA

Model	No. of parameters	RMSE	RMSEP	r^2
ANOVA	442, 566, 540	480	801	0.93
GEMANOVA	23, 23, 23	912	1333	0.93

Replicates 763 0.97

Conclusion

- Multiplicative ANOVA can be a REAL replacement for traditional ANOVA in systems with complex interactions
- See details in Chemom. Intell. Lab. Syst 92 (2008) 101–109

Wanted:

- An easier way to find the optimum model
- Inclusion of replicate data
- A probability figure