PNNL-SA-138533

Slack-Variable versus Mixture Modeling for Mixture Experiments: A Definitive Comparison

Greg F. Piepel*, Dayton C. Hoffmann**, Scott K. Cooley

Pacific Northwest National Laboratory
Richland, WA

2018 Fall Technical Conference
October 4, 2018

* greg.piepel@pnnl.gov

** Summer intern, from BYU Idaho

A Comment About "A Definitive Comparison" in the Title

With tongue firmly in cheek …

- I've known Brad Jones and Chris Nachtsheim for a long time
- I thought that if they could self-proclaim "Definitive Screening Designs" …….
- Then I could self-proclaim

  "A Definitive Comparison"

  in the title of my talk!

Introduction

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- Over the years, several journal articles have advocated using slack variable (SV) models instead of mixture experiment (ME) models
- Recent articles (2015, 2016) claim SV models have advantages over ME models in collinearity & numerical stability, goodness-of-fit (GOF), and interpretation of component effects
- In conflict with a 2009 article that recommended ME models, which can
  - fit at least as well or better than SV models
  - be more appropriate for special kinds of blending effects

Introduction (cont.)

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- The conflicts and unaddressed topics in the literature prompted us to do a “definitive comparison” of the SV and ME modeling approaches
- This presentation (and a paper we are writing)
  - Reviews the literature on SV vs. ME modeling
  - Evaluates justifications and recommendations in the literature
  - Uses literature examples as part of our evaluations

ME Approach to Mixture Experiments

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- Experiments designed and data modeled using all of the mixture components with proportions

$$0 \le x_i \le 1 \quad (i = 1, \ldots, q) \qquad \sum_{i=1}^{q} x_i = 1$$

- May also have single-component constraints (SCCs)

$$L_i \le x_i \le U_i \qquad i = 1, \ldots, q$$

- and/or multiple-component constraints (MCCs)

$$C_k \le \sum_{i=1}^{q} A_{ki} x_i \;\text{ or }\; f(x_1, \ldots, x_q) \le D_k \qquad k = 1, \ldots, K$$

Mixture Experiment Models

- Linear and quadratic Scheffé ME models

$$\text{Linear}: \quad \eta(\boldsymbol{x}) = \sum_{i=1}^{q} \beta_i x_i$$

$$\text{Quadratic}: \eta(\boldsymbol{x}) = \sum_{i=1}^{q} \beta_i x_i + \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} \beta_{ij} x_i x_j \quad \text{(FSQM)}$$

- Reduced Scheffé quadratic mixture (RSQM) models contain a subset of the crossproduct terms
- Partial quadratic mixture (PQM) models contain a subset of squared and crossproduct terms (Piepel et al. 2002 JQT, Smith 2005 book)

SV Approach to Mixture Experiments

- One of *q* components is designated the SV (x_q)
  - Largest proportion values or
  - Largest range of proportions
  - Sometimes thought of as a “filler”
- Experiments designed and data modeled in terms of the remaining components (x_1, …, x_{q-1})
- In the design, SV proportion obtained by subtraction

$$x_q = 1 - \sum_{i=1}^{q-1} x_i$$

Slack Variable Models

- Linear and quadratic SV models (SV = x_q)

$$\text{Linear:} \quad \eta(x) = \alpha_0 + \sum_{i=1}^{q-1} \alpha_i x_i$$

(FQSV)

$$\text{Quadratic:} \quad \eta(x) = \alpha_0 + \sum_{i=1}^{q-1} \alpha_i x_i + \sum_{i=1}^{q-1} \alpha_{ii} x_i^2 + \sum_{i=1}^{q-2} \sum_{j=i+1}^{q-1} \alpha_{ij} x_i x_j$$

- Full linear & quadratic SV models are equivalent to full linear & quadratic ME models

- SV practitioners obtain reduced quadratic SV (RQSV) models by applying variable selection methods to quadratic terms

Main Justifications in the Literature for SV Modeling Approach

Literature review identified 4 claimed justifications for using SV models instead of ME models

1. SV: (i) large proportion values, (ii) largest range of proportions, and/or (iii) "filler" role
2. SV models can have less collinearity and numerical instability than ME models
3. RQSV models can fit better than RSQM models
4. Better interpretation of component effects on the response

Covered in several papers, but Kang et al. (2016, Tech.) covers all four.

Summary and Evaluation of Claimed Justifications for SV Modeling Approach

- In the presentation, we summarize and evaluate each of these 4 claimed justifications
  - "Evaluate" because there are some conflicting claims and recommendations on whether to use ME or SV models
- Evaluations based on
  - Previous literature results
  - New work to address previously unaddressed topics
  - 6 literature examples

6 Literature Examples

1. 4-Comp. Lubricant Blending (SCCs): Used in 4 papers, 2 comparing ME and SV models

2. 4-Comp. Drug Solubility (SCCs): Used in 3 papers comparing ME and SV models

3. 10-Comp. Waste Glass (SCCs & MCCs): Investigate collinearity and numerical instability for 55-term FSQM & RQSV models (also RSQM, RQSV, and PQM models) with data on highly constrained region

6 Literature Examples (cont.)

4. <u>3-Comp. Paint Tint Strength (Simplex)</u>: Used in 3 papers comparing ME and SV models

5. <u>3-Comp. Strawberry Mite Pesticide (Simplex)</u>: One powder component blends additively with other two liquid components

6. <u>4-Comp. Drug Efficacy (SCCs)</u>: Filler component has no effect

Justification 1: SV = Component w/ Largest Magnitude, Range, or "Filler" Role

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Literature Summary

- Snee (1973, Tech) mentioned SV approach when one component is large proportion (≥ 0.90) of the mixture
- Marquardt & Snee (1974, Tech) preferred ME approach, but allowed for SV approach when
  - one component is "vast majority" of the mixture
  - "trace" other components
- Kang et al. (2016, Tech.) mention SV approach when one component (i) large proportion of the mixture, (ii) largest range of proportions, (iii) "filler" role

Justification 1: SV Role, Choice
<u>Literature Summary (cont.)</u>

- Kang et al. (2016, Tech.) claim that “fillers” often
  - have effects that are different than the effects of other components and
  - do not “actively” affect properties of the mixture product
- They used sweet tea as an example, where
  - sugar is considered the “active” component that affects sweetness
  - water is the filler that affects sweetness negatively as a <u>solvent</u>
- This use of “active” differs from previous mixture literature where “inactive” means “no effect”

Justification 1: SV Role, Choice
Literature Summary (cont.)

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- For examples where there is no “natural” SV (e.g., with “filler” role), Kang et al. (2016, Tech.) proposed a criterion to choose the SV component *x* so that FQSV(*x*) has the smallest collinearity
- Then, model reduction of quadratic terms proceeds from FQSV(*x*)

Justification 1: SV Role, Choice Evaluation

- One component with
  - large proportion (or “vast majority”) of a mixture
  - largest range
  - “filler” role

  are alone insufficient reasons for SV modeling
- The “SV component” could have
  - no effect
  - additive effect
  - linear and nonlinear blending effects with other components

  There are ME (but not SV) modeling approaches appropriate for these situations, discussed later

Justification 2: Collinearity and Numerical Instability

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Literature Summary

- Mixture experiments are subject to collinearity (ill-conditioning) and possibly numerical instability because of
  - $\sum_{i=1}^{q} x_i = 1$
  - additional SCCs and possibly MCCs
  - esp. components with small proportions and/or ranges
- Remedies proposed (relevant to SV vs. ME modeling) include
  - SV model
  - transformation of component proportions

Justification 2: Collinearity, Num. Instability

Literature Summary (cont.)

- Khuri (2005, JApplStat), Cruz-Salgado (2015, IngInvTec), and Kang et al. (2016, Tech) compared collinearity diagnostics of FSQM and RSQM models versus FQSV and RQSV models
- They found:
  - FQSV model had less collinearity than the FSQM model
  - best-fitting RQSV model usually had less collinearity than the best-fitting RSQM model

Justification 2: Collinearity, Num. Instability

Literature Summary (cont.)

- Kang et al. (2016) investigated the effects on collinearity diagnostics of four different component transformations, two of which are
  - L-pseudocomponent (L-p)
  - [-1,+1]

  Latter reduced collinearity substantively more than the former
- Goal of smaller collinearity through component transformation is to “alleviate numerical instability” (Kang et al. 2016)

Justification 2: Collinearity, Num. Instability Evaluation

- One key issue in evaluating collinearity and numerical instability for ME and SV models, without or with transformation of components ….
- Is collinearity large enough to cause, or small enough to avoid, practical numerical instability (PNI)?
- PNI: Model parameters are incorrectly estimated such that predicted response values and GOF statistics (like RMSE and R^2) are incorrect
- Issue has not been studied in ME vs. SV modeling literature. So we studied it.

Justification 2: Collinearity, Num. Instability

Evaluation (cont.)

- Our prior experience was that PNI is unlikely to occur for ME models without component transformations (even with high collinearity) given modern
  - regression algorithms
  - computers (64- and 128-bit “words”)
  - software (e.g., double precision)
- 6 examples: We (i) did variable selection (reduced/ partial models) and (ii) fit all full/reduced/partial ME and SV models with three component transformations
  - None
  - L-p (N.A. for two simplex examples)
  - [-1,+1]

Justification 2: Collinearity, Num. Instability

Evaluation (cont.)

- For equivalent models, SV vs. ME and component transformations had no impact on
  - Quadratic terms selected in reduced or partial quadratic models (SV or ME)
  - RMSE, R^2, $R_A^{\,2}$, and $R_P^{\,2}$: Same to many decimal places
- This was true for all 6 examples, including the 10-component waste glass example when fitting 55-term FSQM and FQSV models
- Over all examples
  - CNs: 3.29 to 76,034
  - MaxVIF: 1.41 to 16,401,045

Maxs CN & MaxVIF not for 10-comp. example

Justification 2: Collinearity, Num. Instability Conclusions

- Despite sometimes (much) larger collinearity diagnostics, quadratic term selections and GOF statistics for ME and SV models were not affected by component transformations (i.e., no PNI)
- SV models do not always have smaller collinearity diagnostics than equivalent ME models
- With modern regression algorithms, software, and computers, we saw no practical impact on numerical stability of SV modeling or component transformations to reduce collinearity across 6 examples

Justification 3: Model Goodness-of-Fit

Literature Summary

- Cornell (2000, JQT), Khuri (2005, JApplStat), and Kang et al. (2016, Tech) compared RSQM and RQSV models for various examples to see which models had the best GOF statistics
- They concluded the best-fitting RQSV model (allowing each component to be the SV) can have better GOF statistics than the best-fitting RSQM model

Justification 3: Model Goodness-of-Fit
Literature Summary (cont.)

- Piepel & Landmesser (2009, QE) noted the class of PQM models is larger than the combined class of FSQM, RSQM, FQSV, and RQSV models
- Hence, the best-fitting PQM model will always fit as well or better than any FSQM, RSQM, FQSV, or RQSV model
- Kang et al. (2016) cited the Piepel & Landmesser (2009) paper, but did not include PQM models in their comparison of ME and SV models

Justification 3: Goodness-of-Fit Evaluation

- 6 examples illustrate
  - Best PQM model always fits better than best RSQM or RQSV
  - RSQM model can fit better than RQSV model (red)

Example	Model, RMSE (Smallest in blue/green font)			
	RSQM	RQSV	PQM	Other ME Models
1	0.0425	0.0410	0.0410	N/A
2	0.7990	0.8469	0.7990	N/A
3	0.1214	0.0953	0.0953	N/A
4	0.2831*	0.3573	0.2457	**0.2312** PQM-H3 hybrid
5	3.8204	3.7328	3.7328	**1.1830** Reduced H3
6	0.1642	0.1769	0.1440	**0.0970** FSQM 3-comp

* FSQM, all terms significant

Justification 4: Interpreting Component Effects

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Literature Summary

- Snee (1973, Tech), Snee & Rayner (1984, Tech), and Kang et al. (2016, Tech) note that for SV models
  - effects of changes in non-SV components are relative to offsetting changes in the SV
- However, none of the literature SV examples listed in Table 1 of Piepel and Landmesser (2009, QE) discussed this approach to interpreting component effects
  - Many practitioners ignore the SV as if it has no effect

Justification 4: Interpreting Component Effects: Literature Summary (cont.)

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- Kang et al. (2016) claimed this way of interpreting the effects of non-SV components is a "more straightforward way to understand the component effects for practitioners" compared to the methods typically used with the ME approach
- They also claimed interpretations of component effects using standard methods (e.g., Cox effect direction) are "very difficult for practitioners"

Justification 4: Component Effects <u>Evaluation</u>

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*

- <u>Dating back at least 200 years</u>, scientists in other disciplines (e.g., materials science) commonly used two “definitions” for the effect of a mixture component on a response

  A. Offset change in a component of interest with opposite change in one component (e.g., the SV)

  B. Add or subtract a change in the component of interest from the value in a starting mixture, …

     keeping the remaining components in the same relative proportions as in the starting mixture.

- Pros and cons to each approach for understanding effects of mixture components on a response

Justification 4: Component Effects Evaluation (cont.)

- Cannot say that one way of defining/interpreting component effects is better (or “more straightforward” or “very difficult”) than the other (both used for 200+ years)
  - Each is defensible and provides different information
- Either Definition A or B can be applied
  - experimentally to observe response values
  - by first fitting a model (ME or SV) and making predictions

  along the “effect direction” corresponding to the chosen Definition A or B
- Can plot the results of Definition A or B with a “response trace” or “profile” for each component

Justification 4: Component Effects Evaluation (cont.)

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- Referring to a SV model, Kang et al. (2016) state:
  - "The individual linear, quadratic, and interaction model terms can be interpreted in the same way as other classical regression models. The only difference is that the interpretations are with respect to the filler.
  - The significant linear effect β_i indicates (a) significant main effect of the corresponding component (with respect to the filler).
  - Similarly, a significant quadratic effect indicates that increasing the corresponding component while decreasing the filler the same amount, E(Y) would change in a nonlinear fashion.
  - The interactions can also be interpreted in the usual way."
- Vague--we are working to better understand

ME vs. SV Models for Different Kinds of ME Blending Behavior: Evaluation

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- The statistical literature comparing ME and SV models has focused on full and reduced versions of quadratic polynomial models
- The ME literature has proposed other model forms and methods when polynomial models are inadequate
  - Linear ME model reduction method for removing components with no effects and combining components with similar effects
  - Becker models (H1, H2, H3) for additive blending
  - Models with inverse terms when the response increases rapidly as components approach zero (or a lower limit, upper limit, or some other value)
  - Log-contrast models to assess additivity or inactivity of one or more components

ME vs. SV Models for Different Kinds of ME Blending Behavior: Evaluation (cont.)

- There are no corresponding models or methods that have been proposed for SV modeling
- Some of the models/methods would not work for SV model versions to the extent that "component effects relative to the SV" are built into the SV versions

Example: Cornell (2002, book) Pesticide Mixture with Additive Blending

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- A wettable-powder pesticide (x_3) that blends additively with two liquid pesticides (x_1, x_2)

Model	RSQM	RQSV(x_2)	PQM	RH3
# Terms	4	5	5	4
RMSE	3.82	3.73	3.73	**1.18**
R^2	0.962	0.967	0.967	**0.996**
CN	9.26	32.6	17.2	**4.04**
Max VIF	1.60	21.8	34.3	1.66

- RH3 = Reduced Becker H3 model
  = 3 linear terms + $(x_1 x_2)^{0.5}$

Example: Drug Efficacy with Filler Having No Effect (Inactive)

Pacific Northwest NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

- Two drugs, enhancer, and filler (f), which has no effect on drug efficacy (Piepel & Landmesser 2009)

Model	RSQM	RQSV(f)	PQM	NFSQM
# Terms	8	6	7	6
RMSE	0.164	0.177	0.144	**0.097**
R^2	0.927	0.899	0.939	**0.970**
CN	39,537	20,138	39,352	**41.7**
Max VIF	60,541	11.0	16,401,045	78.7

- RMSE, R^2, R_A^2, and R_P^2 are the same for each model with no, L-p, and [-1,+1] component transformations

Kang et al. (2016) Recommendations

- Recommend using SV models for all mixture experiments
  - Especially if one component is a filler
  - Effect interpretation when a change in a non-SV component is offset by a change in the SV
  - SV models tend to have smaller values of collinearity diagnostics than ME models
  - RQSV models can fit better than RSQM models
- Also recommend using a [-1, +1] transformation of the components with SV models
- We have different (definitive! ☺ ☺) recommendations based on our work

Our Recommendations

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

We recommend ME models

- Class of PQM models includes subclasses of FSQM, RSQM, FQSV, and RQSV models
  - Best-fitting PQM model must fit as well or better than best-fitting FSQM, FQSV, RSQM, or RQSV model
- With modern computers, software, and regression algorithms, we have not seen practical indications of numerical instability in examples with a wide range of values for collinearity diagnostics
- For 6 examples, same GOF results for full and reduced ME and SV models with: no, L-p, and [-1,+1] component transformations

Our Recommendations (cont.)

- To be safe, we recommend fitting ME models with "no" and [-1,+1] component transformations to see if PNI results from collinearity
  - PNI indicated by differences in GOF statistics
  - If no PNI, proceed with "no component transformation" ME models
- If PNI, see if L-p transformation avoids PNI
  - If so, proceed with ME models in L-p transformed components
  - If not, proceed with ME models in [-1,+1] transformed components
  - Unlikely with modern computing

Our Recommendations (cont.)

- ME models can be used to interpret component effects for both Definitions A and B
- Special ME methods and model forms for reducing linear ME models, additive blending, and inactive (no effect) components. No SV model counterparts.
- SV models have significant disadvantages vs. ME models if
  - SV or other components inactive (no effect)
  - SV or other components blend additively
  - SV has largest effect among all components (disadvantage for component effects relative to SV)
  - SV has significant quadratic blending effects

Acknowledgement, Closing

- Dayton Hoffmann was a student intern at PNNL during Summer 2018, funding from the Science Undergraduate Laboratory Intern program sponsored by the U. S. Department of Energy

- We are writing a paper on this work which is about 70% done
- Contact me at greg.piepel@pnnl.gov if you'd like a PDF of the slides or the paper when it is done
- Thank you for your kind attention!