

Kansas State University Libraries

New Prairie Press

Conference on Applied Statistics in Agriculture

2015 - 27th Annual Conference Proceedings

Best Linear Unbiased Prediction: an Illustration Based on, but Not Limited to, Shelf Life Estimation

Maryna Ptukhina

University of Nebraska - Lincoln, maryna@huskers.unl.edu

Walter Stroup

University of Nebraska-Lincoln, wstroup@unl.edu

Follow this and additional works at: <https://newprairiepress.org/agstatconference>



Part of the [Agriculture Commons](#), [Applied Statistics Commons](#), and the [Biostatistics Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](#).

Recommended Citation

Ptukhina, Maryna and Stroup, Walter (2015). "Best Linear Unbiased Prediction: an Illustration Based on, but Not Limited to, Shelf Life Estimation," *Conference on Applied Statistics in Agriculture*. <https://doi.org/10.4148/2475-7772.1091>

This Event is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact cads@k-state.edu.

Best Linear Unbiased Prediction: an Illustration Based on, but Not Limited to, Shelf Life Estimation

Marina Ptukhina¹, Walter Stroup²

¹ Department of Mathematics and Computer Science Whitman College, 345 Boyer Ave Walla Walla, WA 99362

² Department of Statistics University of Nebraska -Lincoln, 340 Hardin Hall North Wing Lincoln, NE 68583-0963

Abstract

Shelf life estimation procedures, following ICH guidelines, use multiple batch regression with fixed batch effects. This guidance specifically mandates estimates based on at least 3 batches. Technically, the fixed-batch model limits inference to the batches actually observed, whereas ICH requires resulting estimates to apply to all future batches stored under similar conditions. This creates a conflict between the model used and the inference space the model is intended to address. Quinlan, et al. [6] and Schwenke [8] studied the small sample behavior of this procedure. Both studies revealed large sampling variation associated with the ICH procedure, producing a substantial proportion of extremely low and extremely high estimates. Quinlan, et. al [5] also considered alternative approaches including mixed models with random batch effects. While this eliminated the conflict between model and intended inference space, there were still problems with the mixed model approaches Quinlan considered. We present a Bayesian augmented mixed model approach to shelf life estimation that takes advantage of the theoretical benefits of the mixed model and uses prior information about variance components to improve accuracy of shelf life estimation procedure.

Key Words: BLUP, shelf life estimation

1. Introduction

Accurate shelf life estimation is very important to a variety of applications. This paper specifically focuses on the pharmaceutical industry, where inaccurate estimation can lead to undesirable consequences. Overestimation of shelf life could lead to consumption of drugs that are no longer stable and effective, while underestimation can cause the consumer to discard good product prematurely. Thus accurate estimation of shelf life is essential to both consumers and producers. Recent research [3], [6], [9] suggests that shelf life estimation procedures used in the pharmaceutical industry are not always reliable: overestimation and underestimation are common. This paper explores new techniques that utilize prior information gained through previous stages during the development process, thus minimizing the required number of replications. These techniques are geared to provide consumers with accurate shelf life estimates while maintaining lower costs for the production facilities.

2. Shelf life

Suppose we have a product that deteriorates over time. In an agricultural setting, this could be a food product or pesticide applied to a field. In a pharmaceutical setting, this could be a drug or a vaccine. There are many possible ways the stability limiting characteristic may decrease. In the simplest case, we could imagine that a product's effectiveness decreases linearly over time, and its lifetime is determined by how long its measure of effectiveness remains within a defined acceptance region. Figure 1 shows a population of batches, where y is the stability limiting characteristic and x is storage time in months.

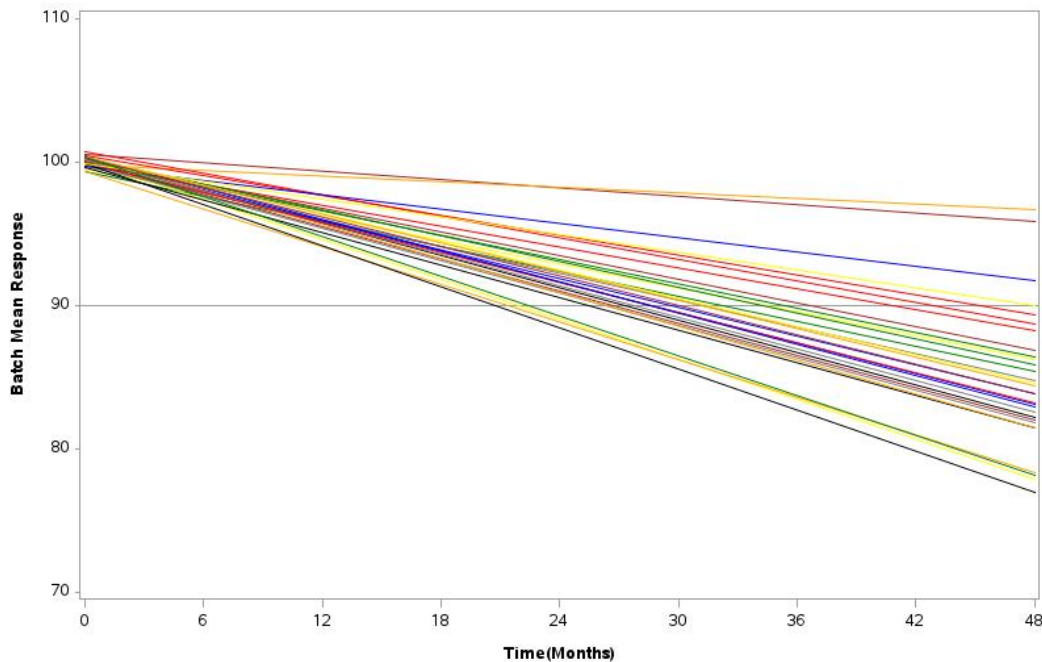


Figure 1: Distribution of batch mean responses over time

It is important to note that due to variation in batch responses over time, the intercepts and slopes will vary. Once the distribution of batch mean responses is specified, the distribution of shelf lives arises as a consequence. For the purpose of this discussion we will use the convention that the average stability limiting characteristic of interest at time zero is at 100%.

Over time, the stability limiting characteristic decreases to a point after which it is not considered acceptable anymore (below the acceptance criterion). When a batch intersects the acceptance criterion, we use $A = 90$ in this discussion, the value gets projected on the horizontal axis and the corresponding value on the horizontal axis represents the batch's shelf life.

When the process is repeated for each batch, the distribution of batch means produces the distribution of batch shelf lives on the horizontal axis. Figure 2 illustrates this relationship between the batch mean response distribution and the distribution of batch shelf lives.

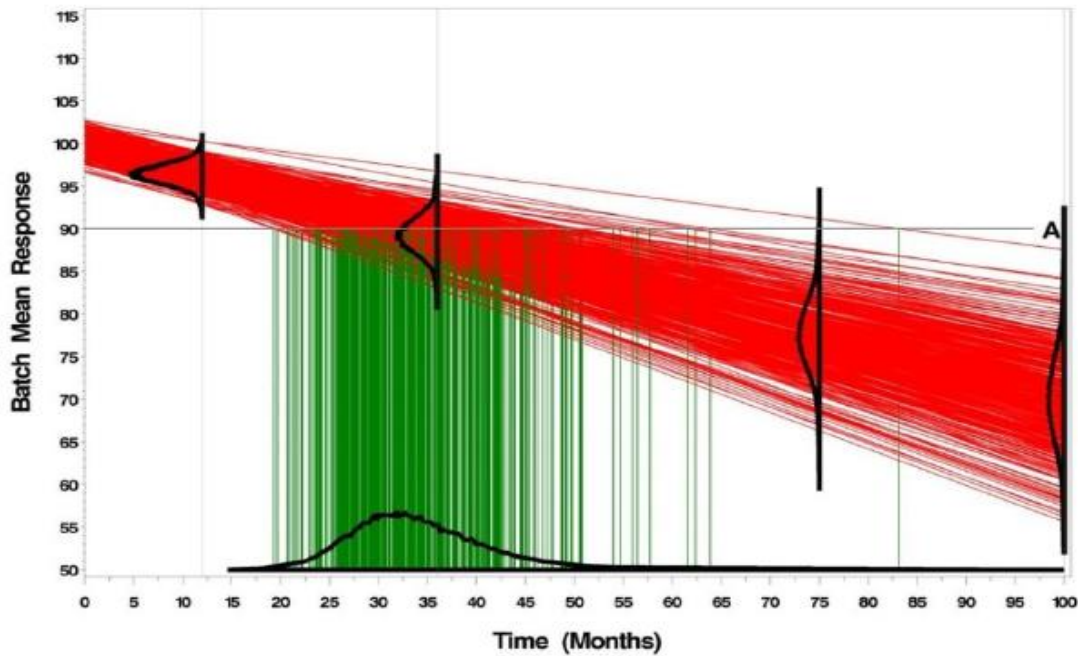


Figure 2¹: Distribution of batch mean responses over time and distribution of shelf lives

¹ Picture credited to Quinlan [7].

Within the shelf life determination frame, the statement that we would like to make is that any given batch's shelf life will meet or exceed the shelf life established for the product with acceptably high probability. For the purpose of this discussion let us assume that we would like that probability to be at least 95%. Therefore, in this paper the focus of shelf life determination is on the 5th percentile of the distribution of shelf lives.

2.1 Fixed batch effects model

A model typically used for this type of problem can be written as

$$y_{ij} = \beta_{0i} + \beta_{1i}X_j + e_{ij}, \quad (1)$$

where y_{ij} denotes the stability limiting response at the j^{th} time for the i^{th} batch, X_j denotes the j^{th} time (typically given in months of storage), β_{0i} denotes the intercept for the i^{th} batch, β_{1i} denotes the slope for the i^{th} batch and e_{ij} is the measurement error with the following distribution $e_{ij} \sim NI(0, \sigma^2)$. According to the ICH guidelines at least three batches should be used.

This model suggests unequal intercepts and unequal slopes for each batch, however it does not take into account the variability that naturally occurs among the batch intercepts and batch slopes.

Quinlan, et al. [6] and Schwenke [8] studied the small sample behavior of this procedure. Both studies revealed large sampling variation associated with the ICH procedure, producing a substantial proportion of unacceptably low and high estimates

2.2 Mixed model

ICH also requires resulting estimates to apply to all future batches stored under similar conditions. This creates a conflict between the fixed batch effects model used and the inference space the model is intended to address.

A random coefficient linear mixed model provides an alternative to the current approach, since it allows for inference to be applied to all future batches.

Random coefficient linear mixed model is given by:

$$y_{ij} = \beta_0 + b_{0i} + (\beta_1 + b_{1i})X_j + e_{ij}, \quad (2)$$

where y_{ij} denotes the stability limiting response at the j^{th} time for the i^{th} batch, X_j denotes the j^{th} time (typically given in months of storage), β_0 is the mean value of stability limiting characteristic at time zero, β_1 is the mean rate of decrease of the stability limiting characteristic, b_{0i} is random deviation on the mean value of stability limiting characteristics of the i^{th} batch and b_{1i} is the random deviation on the mean rate of decrease of stability limiting characteristic for the i^{th} batch that have the following distribution:

$$\begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix} \right),$$

$$e_{ij} \sim NI(0, \sigma^2).$$

where σ_0^2 is the variance among the intercepts of different batches, σ_1^2 is the variance among the slopes of different batches, σ_{01} is the covariance among the intercepts and slopes of different batches, e_{ij} is the measurement error that is independent of b_{0i} and b_{1i} , σ^2 is the variance of the measurement error.

Often in practice it is common to assume that σ_{01} is equal to zero. Specifically, we do so in this paper.

Quinlan, et. al [5] considered this mixed model with random batch effects. While this eliminated the conflict between model and intended inference space, there were still problems with the mixed model approaches that were considered. Specifically, while the variability of the mixed model approach was improved compared to ICH procedure, the mixed model approach produced a relatively large proportion of estimates that were too high.

The question motivating this study is: “What if we try a Bayesian approach to shelf life estimation?” In other words, would a use of plausible, somewhat informative priors improve the estimation procedure? If we have information about variance components from the previous stages of drug development, we could use it to inform the priors. To answer these questions, we performed a simulation study that is described in the next section. Before we describe the simulation and its results let us look at different ways of calculating shelf life.

2.3 How can we calculate shelf life?

Using the random coefficient linear regression mixed model (2) we could obtain the estimates for β_0 and β_1 . Estimated shelf life is then given by:

$$\hat{T}_1 = \frac{A - \hat{\beta}_0}{\hat{\beta}_1}, \quad (3)$$

where A is the acceptance criterion, $\hat{\beta}_0$ and $\hat{\beta}_1$ are the estimates for the intercept and slope.

Quinlan has tried this approach and showed that the resulting shelf life estimates were too high.

Instead of using the estimates for $\hat{\beta}_0$ and $\hat{\beta}_1$, we could get the confidence bounds for $\hat{\beta}_0$ and $\hat{\beta}_1$ using the mixed models estimation procedure and calculate shelf life based on the lower confidence bounds.

Estimated shelf life is then calculated as follows:

$$\hat{T}_2 = \frac{A - \hat{\beta}_{0,L}}{\hat{\beta}_{1,L}}, \quad (4)$$

where A is the acceptance criterion, $\hat{\beta}_{0,L}$ and $\hat{\beta}_{1,L}$ are the lower bound estimates of population-averaged intercept and slope, respectively. For example, we can use the 5th percentile of $\hat{\beta}_{0,L}$ and $\hat{\beta}_{1,L}$, which are the lower bounds of a one sided 95% confidence interval.

We could also get the shelf life based on the Best Linear Unbiased Predictors (BLUPs). To continue the discussion let us first define BLUPs more formally. Best Linear Unbiased Prediction (BLUP) is a method of estimating random effects. BLUP was developed by Henderson [2] and originally was used in animal breeding at the time of development. As we use more information obtained from the random effects the variance of the prediction error decreases, which is the beneficial property of BLUP.

Relating this to the shelf life situation, the BLUPs of i^{th} batch intercept and slope are $\hat{B}_{0i} = (\hat{\beta}_0 + \hat{b}_{0i})$ and $\hat{B}_{1i} = (\hat{\beta}_1 + \hat{b}_{1i})$ respectively. Using BLUPs, we can make batch-specific predictions taking into account the underlying distribution of random effects. This allows us to focus on the shortest-lived batch and obtain the shelf life based on the batch that reaches the acceptance criterion first.

Estimated shelf life is then calculated as follows:

$$\hat{T}_3 = \frac{A - (\hat{\beta}_0 + \hat{b}_{0i})}{(\hat{\beta}_1 + \hat{b}_{1i})}, \quad (5)$$

where A is the acceptance criterion, $(\hat{\beta}_0 + \hat{b}_{0i})$ and $(\hat{\beta}_1 + \hat{b}_{1i})$ are the estimates of the shortest-lived batch intercept and slope, respectively.

Alternatively, we could base the shelf life estimate on the lower confidence bounds of the batch-specific intercept and slope BLUPs of the batch with the shortest shelf life. Estimated shelf life is then given by:

$$\hat{T}_4 = \frac{A - (\hat{\beta}_0 + \hat{b}_{0i})_L}{(\hat{\beta}_1 + \hat{b}_{1i})_L}, \quad (6)$$

where A is the acceptance criterion, $(\hat{\beta}_0 + \hat{b}_{0i})_L$ and $(\hat{\beta}_1 + \hat{b}_{1i})_L$ are the lower bound estimates of the shortest-lived batch intercept and slope, respectively.

Figure 3 illustrates the difference between shelf life determined using the population-averaged batch response over time versus shelf life determined using the BLUP of the shortest-lived batch. Dashed lines on each graph represent the response based on the lower bound estimates. “SL Point Estimate” on each plot represents the shelf life obtained using the estimates of the slope and the intercept, “SL_LCL” is the shelf life obtained based on the lower bound estimates of the intercept and slope.

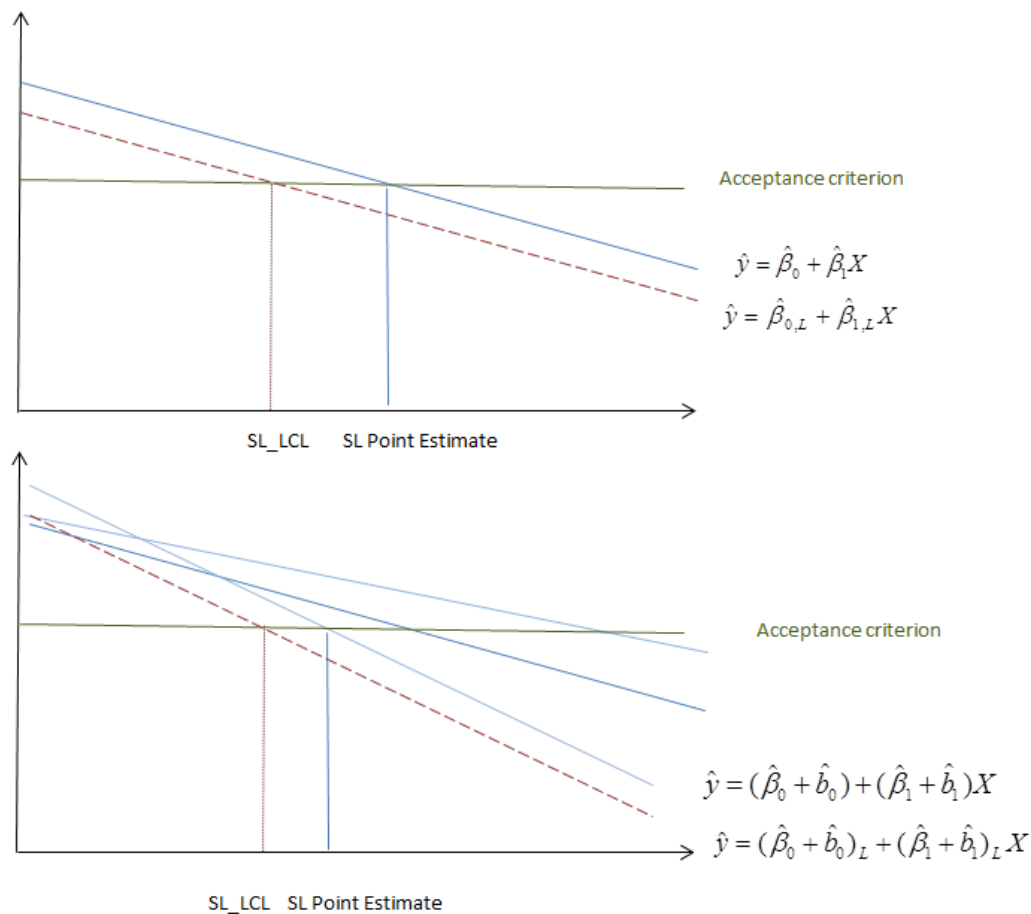


Figure 3: Population-averaged versus BLUP shelf life

3. Simulation

The idea of the simulation study is to generate the empirical distribution of true shelf life and use its characteristics as the standard for comparing the shelf life estimation procedures. Specifically, we focus on the 0 through 5th percentile of the empirical distribution of true shelf life and refer to this range of values as the target range. We chose this range because we would like the shelf life produced by each of the methods not to exceed the risk deemed acceptable (in this paper 5th percentile), but also not to be less than the effective minimum of the distribution.

We use the model described in section 2.2 as the starting point for obtaining the shelf life and then compare the distribution of shelf life obtained within each method to the target range. The performance of an estimation procedure is better if higher proportion of shelf lives fall within the target range. The percentage of shelf lives falling within the target range in this paper is also referred to as coverage.

As a result, two indicators of quality are of interest: a high percentage of calculated shelf lives below the 5th percentile of the empirical distribution of true shelf life, and a small spread of the distribution of calculated shelf lives.

3.1 Empirical shelf life distribution

The first step was to generate the empirical shelf life distribution. One thousand data sets were generated using a random coefficient regression model with 3 batches per trial. Observations were generated for each batch at times 0, 3, 6, 9, 12, 18 and 24 months.

Parameters used in simulation were similar to Quinlan's: $B_{0i} \sim NI(101, 1.5)$, $B_{1i} \sim NI(-0.33, 0.0015)$ and $e_{ij} \sim NI(0, 0.5)$, where B_{0i} and B_{1i} are intercept and slope effects of the i^{th} batch assumed to be uncorrelated, and e_{ij} is the measurement error. The acceptance criterion A was set to 90.

This set of parameters represents the cumulative knowledge for some product in the pharmaceutical industry. These parameters describe a stability limiting characteristic with an average of 101 at time zero, and 0.33-unit decrease per month, which corresponds to a 4 % annual rate of decrease of the stability limiting characteristic.

Simulations were performed in SAS 9.4 using PROC MCMC. The length of Markov chain used for each scenario was 2,000,000 (before thinning) with a burn-in of 100,000 and thinning of 1000.

A histogram of the empirical distribution of true shelf life is shown on Figure 4 below.

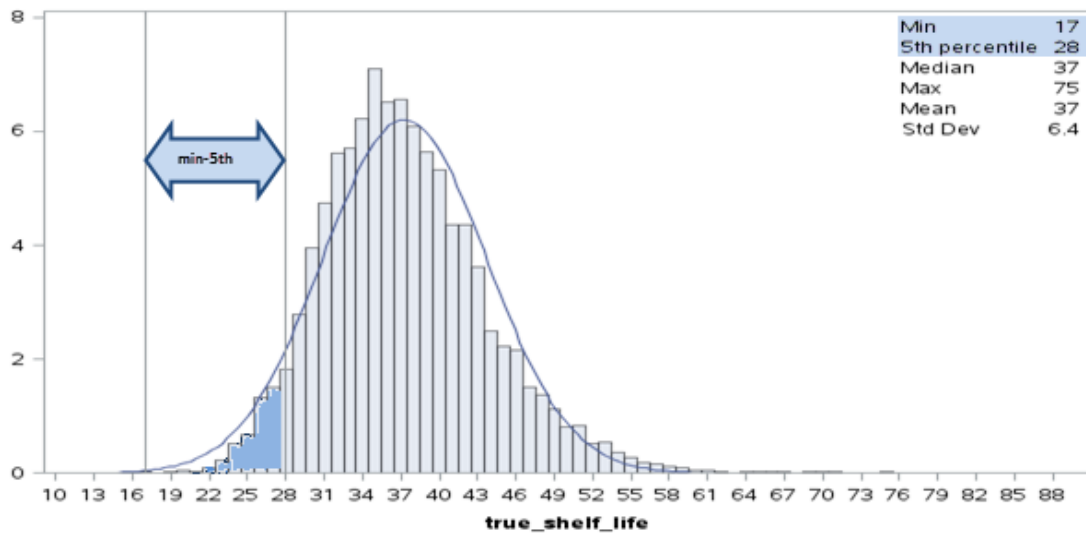


Figure 4: Empirical distribution of true shelf lives

From Figure 4 we can see that the empirical distribution of the true shelf life ranges from 17 to 75 months with a 5th percentile of 28 months and median of 37 months. Thus the target range of values between 0 and 5th percentile in this simulation corresponds to the range of shelf lives between 17 and 28 months. This range is shaded in Figure 4.

3.2 Estimation procedures

As mentioned earlier the idea behind a Bayesian augmented mixed model approach is to use the information that we have from the previous stages of product development and try different combination of priors to learn about the behavior of shelf life estimates for each scenario.

To explore this, we decided to look at four different scenarios shown in the Table 1 below. Technically, priors on σ_0^2 and σ_1^2 should be called hyperpriors, but for simplicity of this discussion they would be referred to as priors. Other names for the non-informative prior are vague, diffuse, and flat prior.

Table1: Four cases of priors

	Diffuse priors on β_0, β_1	Informative priors on β_0, β_1
Diffuse priors on $\sigma_0^2, \sigma_1^2, \sigma^2$	Case 1	Case 4
Informative priors on $\sigma_0^2, \sigma_1^2, \sigma^2$	Case 2	Case 3

For each scenario we tried different sets of diffuse and informative priors, the results presented for each case are typical results that are representative of each set of prior combinations. Results for each scenario are described below. In this section lower confidence bounds of 95% confidence interval of mixed model estimates $\hat{\beta}_{0,L}$ and $\hat{\beta}_{1,L}$ were used to obtain shelf life.

The guiding principle for selecting informative priors for each simulation study was that we assumed that we had available information about the most likely values of the parameters. The most likely values assumed for β_0 , β_1 , σ_0^2 , σ_1^2 and σ^2 were 100, -0.33, 1.5, 0.0015 and 0.5 respectively.

The following non-informative priors were used for intercept and slopes: $\beta_0 \sim NI(100,400)$, $\beta_0 \sim \text{general}(0)$, which is SAS notation for a flat prior and $\beta_1 \sim NI(0,4)$, $\beta_1 \sim \text{general}(0)$. The informative priors used for β_0 and β_1 were NI (100; 9) and NI (- 0:33; 0:05), respectively.

For the variance components the following inverse gamma priors were used: $\sigma_0^2 \sim IG(3,6)$, $\sigma_0^2 \sim IG(5,9)$, $\sigma_1^2 \sim IG(3,0.006)$, $\sigma_1^2 \sim IG(5,0.009)$, $\sigma^2 \sim IG(3,2)$, $\sigma^2 \sim IG(5,3)$ for informative and $\sigma_0^2, \sigma_1^2, \sigma^2 \sim IG(0.01,0.01)$ for non-informative priors.

The typical distributions of shelf life determined based on the lower 5th percentile of the population-averaged intercept and slope for each of the four cases are shown in Figure 5.

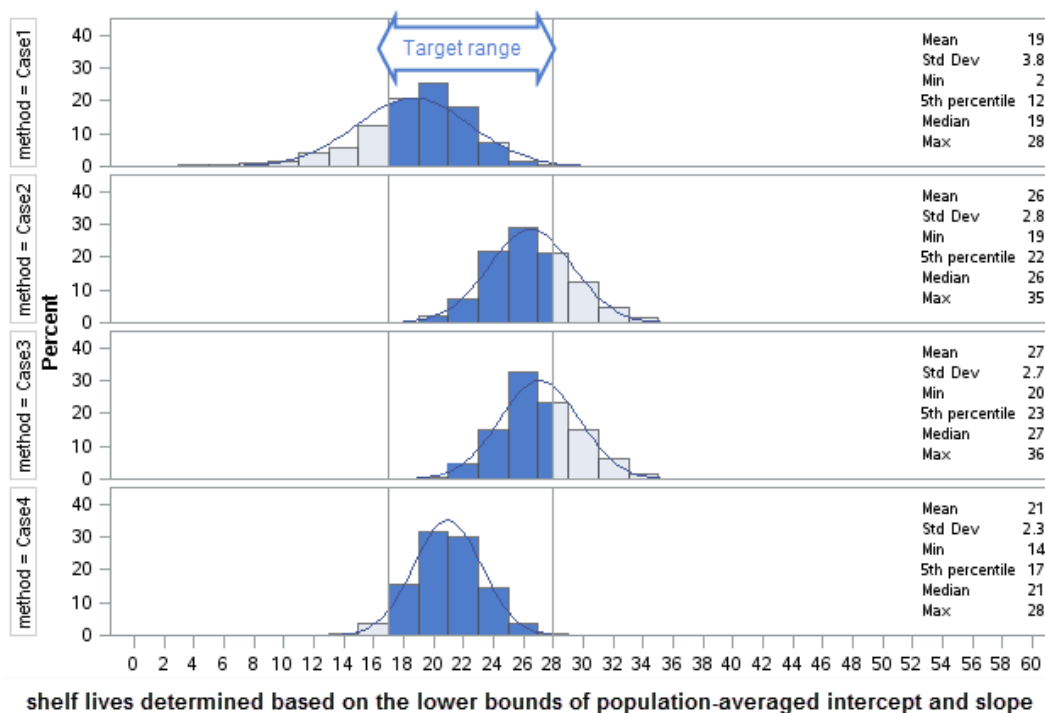


Figure 5: The resulting distributions of shelf life based on lower 5th percentile of population-averaged slope and intercept

As can be seen in Figure 5, the most promising results were achieved with the scenario that combined non-informative priors for variance components with informative priors for slope and intercept. This is because we are interested in calculated shelf lives that fall between the 0 to 5th percentile of the true shelf life, which in this simulation study translates into the target range of values between 17 and 28 months

The target range of values produced by each of the four cases is shaded in Figure 5. The first three cases of priors did not show the appropriate coverage of the target range, whereas the fourth combination of priors produced the coverage of approximately 90% of the target range.

Gelman [1] suggests using weakly informative priors that provide intentionally weaker information than the prior knowledge available for variance parameters. Kruschke [4] also recommends the use of mildly informed priors instead of non-informed priors or strongly informed priors. To explore these suggestions, we included weakly informative priors in both simulation studies.

Since the results of the fourth scenario showed the most promise, we decided to investigate this case further and see if making the priors for the variance components less vague, i.e. using weakly informative priors, will improve the results. The idea that we used to create a weakly informative prior is to use the most likely value of the parameter as the mode of a prior distribution, however allow for larger values compared to the informative prior

The first set of weakly informative priors: $\sigma_0^2 \sim IG(0.25, 1.875)$, $\sigma_1^2 \sim IG(0.25, 0.00375)$, $\sigma^2 \sim IG(0.25, 0.625)$ resulted in shelf life shown as Case 4a on Figure 6, the second set of weakly informative priors: $\sigma_0^2 \sim IG(0.5, 2.25)$, $\sigma_1^2 \sim IG(0.5, 0.00225)$, $\sigma^2 \sim IG(0.5, 0.75)$ resulted in the shelf life distribution shown as Case 4b on Figure 6.

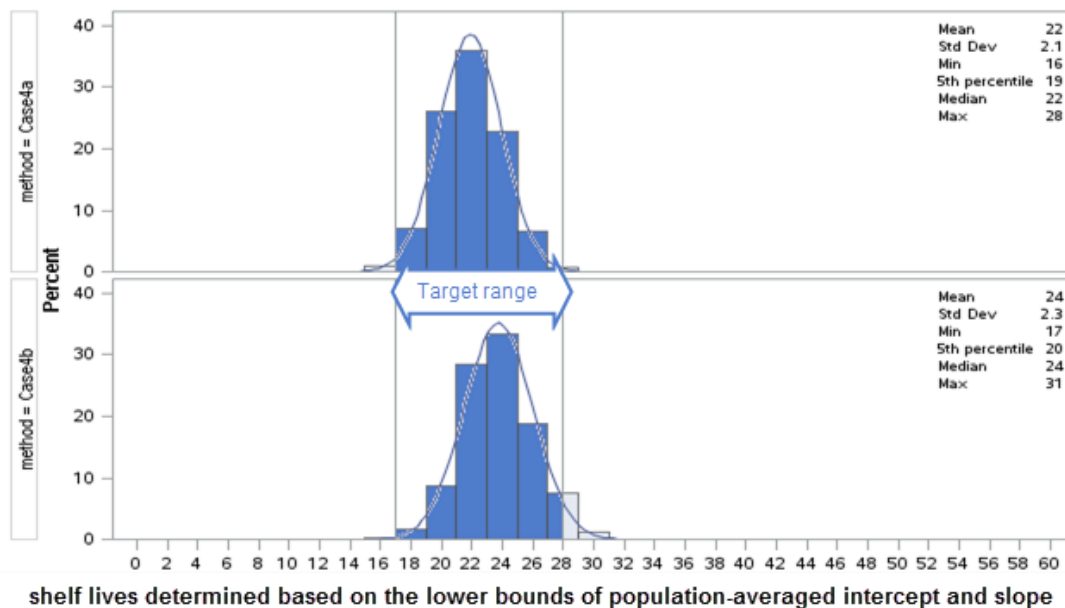


Figure 6: The resulting distributions of shelf life based on lower 5th percentile of population-averaged slope and intercept with weakly informative priors

New Prairie Press

<http://newprairiepress.org/agstatconference>

We can see in Figure 6 that the results of using weakly informative priors for the variance components look promising. The shaded area, 17 to 28 months, is symmetric and covers approximately 99% of the distribution of shelf life for case 4a and approximately 95% for case 4b.

4. Best Linear Unbiased Prediction (BLUP)

As described in section 2.3, using BLUPs of i^{th} batch intercept and slope are $\hat{B}_{0i} = (\hat{\beta}_0 + \hat{b}_{0i})$ and $\hat{B}_{1i} = (\hat{\beta}_1 + \hat{b}_{1i})$ respectively, we can make batch-specific predictions taking into account the underlying distribution of random effects. Doing so lets us focus on the shortest-lived batch and calculate the shelf life using the batch that reaches the acceptance criterion first.

Stroup and Quinlan [9] have explored this mixed model methodology further and found that shelf life estimates based on the lower BLUPs of batch specific slope and intercept of the shortest-lived batch were performing the best out of several other competing mixed model based approaches. However, all of the existing methods that use BLUPs for shelf life determination still need improvement, because with only three batches we do not have sufficient degrees of freedom to estimate variance components. Thus augmenting the BLUP approach with the Bayesian priors seems very reasonable.

4.1. Population average versus BLUP

In this section we will look at the result of using the first set of weakly informative priors described in the previous sections to obtain shelf lives based on BLUPs of the shortest-lived batch. Figure 7 shows the resulting distributions of the shelf life that are based on the median and the lower 5th percentile of the BLUPs of the shortest-lived batch with the first set of weakly informative priors on variance components.

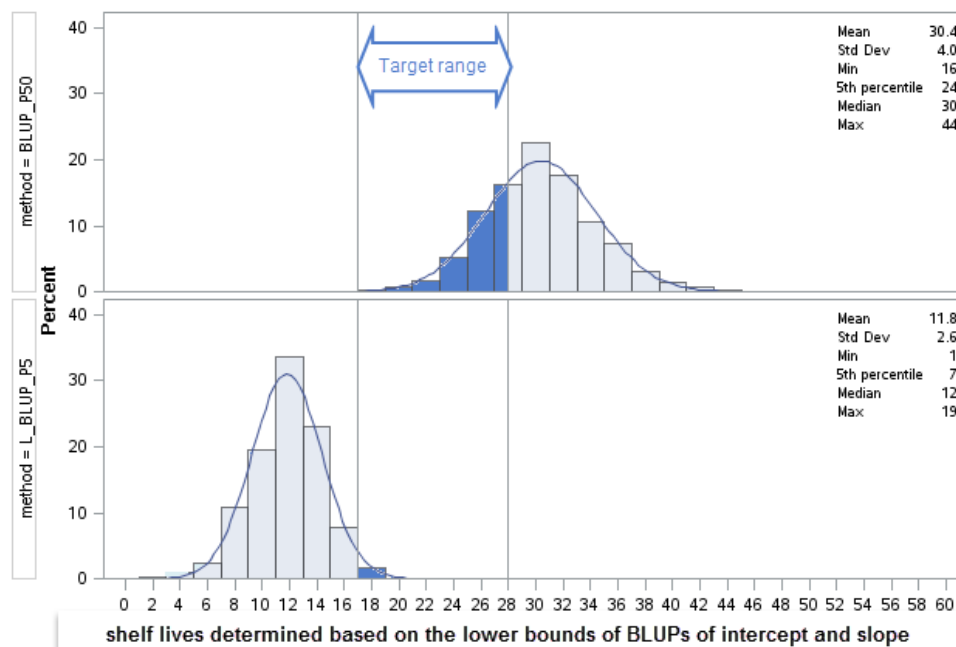


Figure 7: The resulting distributions of shelf life based on the median and lower 5th percentile of BLUP with weakly informative prior with weakly informative prior

We can see in Figure 7 that the shelf life values based on the posterior median of the BLUPs of the shortest-lived batch have a large proportion of values that are too high (unshaded area to the right of the target range of shaded values). On the other hand, the distribution of shelf life based on the lower 5th percentile of the BLUPs of the shortest-lived batch has a large proportion of values that are too low (unshaded area to the left of the target range of shaded values).

Therefore, it seems reasonable that if we use the lower bounds of BLUPs of intercept and slope between the 5th and 50th percentiles we could obtain the distribution of shelf life that has most of its values within the target range.

Figure 8 provides a visualization that supports the idea of why it makes sense to use the lower bounds of BLUPs of intercept and slope between the 5th and 50th percentiles. The vertical dashed line projected on the horizontal axis in Figure 8 shows the shelf life estimate based on the lower BLUPs of intercept and slope between the 5th and 50th percentiles.

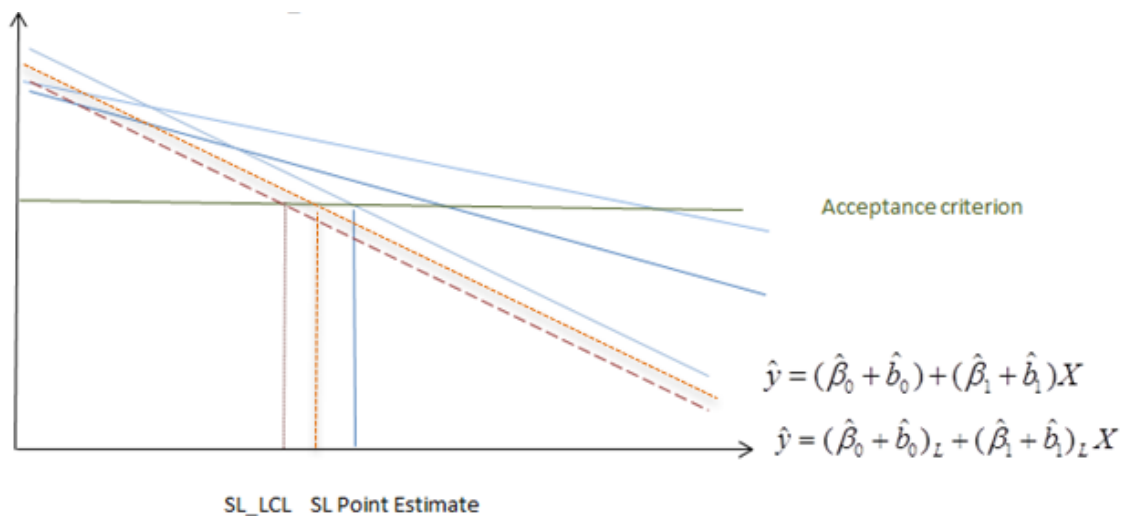


Figure 8: Shelf lives based on lower bounds of BLUPs of the shortest-lived batch

Continued trial and error investigation of the percentiles within the 5th and 50th range revealed that using the lower 23rd percentile of the BLUPs of the shortest-lived batch to determine shelf life produced desired results.

Figure 9 presents the distributions of shelf lives obtained based on the lower 23rd bounds of BLUPs of the intercept and the slope of the shortest-lived batch with the two sets of weakly informative priors introduced in section 3.2.

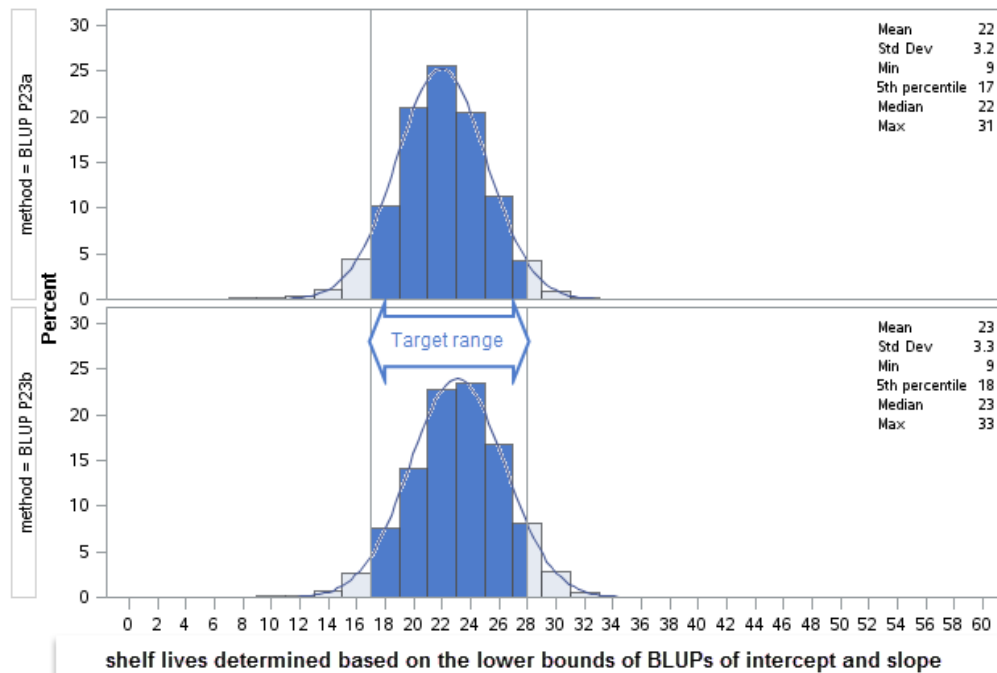


Figure 9: The resulting distribution of shelf life based on the lower 23rd percentile of BLUP with weakly informative priors.

The first set of weakly informative priors resulted in the distributions of shelf life shown as case BLUP P23_a in Figure 9, the second set of weakly informative priors resulted in the shelf life distribution shown as case BLUP P23_b in Figure 9. The shaded area, 17 to 28 months, is symmetric and covers approximately 90% of the target range of the distribution of shelf life for both cases. Thus we can see that using the approximate average of the distance between the 5th and the 50th percentiles, i.e. the 23rd percentile for the lower bound of the BLUPs of the intercept and slope of the shortest-lived batch substantially increases the coverage of the target range.

5. Summary and Conclusions

To summarize the results of this simulation study we can say that using Bayesian estimated BLUPs for shelf life determination shows promise. Using the lower 23rd percentile of the BLUPs of the shortest-lived batch to obtain shelf life in combination with the weakly informative priors produced the distributions of the shelf life that contain about 90% of values within the target range. However, to be able to recommend the use of the lower 23rd percentile of the shortest-lived batch as a basis for shelf life determination, additional theoretical justification needs to be developed.

Acknowledgements

We would like to thank Dr. Michelle Quinlan for her helpful advice, assistance and willingness to allow us to use her illustrations.

References

- [1]. A. Gelman et al., “Prior distributions for variance parameters in hierarchical models”, *Bayesian analysis* 1 (2006), no. 3, 515_534.
- [2]. C.R Henderson, “Best linear unbiased estimation and prediction under a selection model”, *Biometrics* (1975), 423_447.
- [3]. A. Kiermeier, A. Verbyla, and R. Jarrett, “Estimating a single shelf-life for multiple batches”, *Australian & New Zealand Journal of Statistics* 54 (2012), no. 3, 343_358.
- [4]. J. Kruschke, Tutorial: “Doing Bayesian data analysis with R and Bugs”, 2010.
- [5]. M. Quinlan, W. Stroup, J.Christopher , J. Schwenke “On the distribution of batch shelf lives”, 2013
- [6]. M. Quinlan, W. Stroup, J. Schwenke, J. Christopher “Evaluating the performance of the ICH guidelines for Shelf life estimation”, 2013
- [7]. M. Quinlan “On Issues in the Estimation of Quantiles in the Presence of Random Effects with Applications to Shelf Life Estimation”, PhD Dissertation. University of Nebraska, Lincoln, 2010.
- [8] J. Schwenke “Current practices in Shelf Life Estimation”, 2010.
- [9] W. Stroup, M. Quinlan, “Statistical consideration for stability and the estimating shelf life”, Chapter 21 in *non-clinical statistics for pharmaceutical and biotechnology industries*, Springer Press, 2015.