

6-2012

# Scheduling Viability Tests for Seeds in Long-Term Storage Based on a Bayesian Multi-Level Model

Allan Trapp II

*Iowa State University, atrapp@iastate.edu*

Philip M. Dixon

*Iowa State University, pdixon@iastate.edu*

Mark P. Widrlechner

*United States Department of Agriculture, isumw@iastate.edu*

David A. Kovach

*United States Department of Agriculture*

Follow this and additional works at: [http://lib.dr.iastate.edu/ncrpis\\_pubs](http://lib.dr.iastate.edu/ncrpis_pubs)



Part of the [Agricultural Science Commons](#), [Agriculture Commons](#), [Biostatistics Commons](#), and the [Horticulture Commons](#)

The complete bibliographic information for this item can be found at [http://lib.dr.iastate.edu/ncrpis\\_pubs/30](http://lib.dr.iastate.edu/ncrpis_pubs/30). For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

---

# Scheduling Viability Tests for Seeds in Long-Term Storage Based on a Bayesian Multi-Level Model

## Abstract

Genebank managers conduct viability tests on stored seeds so they can replace lots that have viability near a critical threshold, such as 50 or 85 % germination. Currently, these tests are typically scheduled at uniform intervals; testing every 5 years is common. A manager needs to balance the cost of an additional test against the possibility of losing a seed lot due to late retesting. We developed a data-informed method to schedule viability tests for a collection of 2,833 maize seed lots with 3 to 7 completed viability tests per lot. Given these historical data reporting on seed viability at arbitrary times, we fit a hierarchical Bayesian seed-viability model with random seed lot specific coefficients. The posterior distribution of the predicted time to cross below a critical threshold was estimated for each seed lot. We recommend a predicted quantile as a retest time, chosen to balance the importance of catching quickly decaying lots against the cost of premature tests. The method can be used with any seed-viability model; we focused on two, the Avrami viability curve and a quadratic curve that accounts for seed after-ripening. After fitting both models, we found that the quadratic curve gave more plausible predictions than did the Avrami curve. Also, a receiver operating characteristic (ROC) curve analysis and a follow-up test demonstrated that a 0.05 quantile yields reasonable predictions.

## Keywords

After-ripening, Maize, Optimal threshold, ROC curve, Seed dormancy, Seed viability, Statistics

## Disciplines

Agricultural Science | Agriculture | Biostatistics | Horticulture | Plant Sciences

## Comments

This article is from *Journal of Agricultural, Biological, and Environmental Statistics* 17, no. 2 (June 2012): 192–208, doi: [10.1007/s13253-012-0085-y](https://doi.org/10.1007/s13253-012-0085-y).

## Rights

Works produced by employees of the U.S. Government as part of their official duties are not copyrighted within the U.S. The content of this document is not copyrighted.

# Scheduling Viability Tests for Seeds in Long-Term Storage Based on a Bayesian Multi-Level Model

Allan TRAPP II, Philip DIXON, Mark P. WIDRLECHNER, and David A. KOVACH

Genebank managers conduct viability tests on stored seeds so they can replace lots that have viability near a critical threshold, such as 50 or 85 % germination. Currently, these tests are typically scheduled at uniform intervals; testing every 5 years is common. A manager needs to balance the cost of an additional test against the possibility of losing a seed lot due to late retesting. We developed a data-informed method to schedule viability tests for a collection of 2,833 maize seed lots with 3 to 7 completed viability tests per lot. Given these historical data reporting on seed viability at arbitrary times, we fit a hierarchical Bayesian seed-viability model with random seed lot specific coefficients. The posterior distribution of the predicted time to cross below a critical threshold was estimated for each seed lot. We recommend a predicted quantile as a retest time, chosen to balance the importance of catching quickly decaying lots against the cost of premature tests. The method can be used with any seed-viability model; we focused on two, the Avrami viability curve and a quadratic curve that accounts for seed after-ripening. After fitting both models, we found that the quadratic curve gave more plausible predictions than did the Avrami curve. Also, a receiver operating characteristic (ROC) curve analysis and a follow-up test demonstrated that a 0.05 quantile yields reasonable predictions.

**Key Words:** After-ripening; Maize; Optimal threshold; ROC curve; Seed dormancy; Seed viability.

## 1. INTRODUCTION

Many plant germplasm collections no longer exist in their native environments. The genetic traits of these germplasm collections are valuable for crop improvement and new-

---

Allan Trapp II (✉) is PhD Candidate (E-mail: [atrapp@iastate.edu](mailto:atrapp@iastate.edu)) and Philip Dixon is Professor (E-mail: [pdixon@iastate.edu](mailto:pdixon@iastate.edu)), Department of Statistics & Statistical Laboratory, Iowa State University, Snedecor Hall, Ames, IA 50011-1210, USA. Mark P. Widrechner is USDA-ARS Horticulturist (retired) and Affiliate Associate Professor of Ecology, Evolution, & Organismal Biology and Horticulture (E-mail: [isumw@iastate.edu](mailto:isumw@iastate.edu)) and David A. Kovach is USDA-ARS Agronomist (retired) (E-mail: [David.Kovach@ars.usda.gov](mailto:David.Kovach@ars.usda.gov)), North Central Regional Plant Introduction Station, Iowa State University, Ames, IA 50011-1170, USA.

© 2012 International Biometric Society  
*Journal of Agricultural, Biological, and Environmental Statistics*, Volume 17, Number 2, Pages 192–208  
DOI: [10.1007/s13253-012-0085-y](https://doi.org/10.1007/s13253-012-0085-y)

product development and provide beneficial information to researchers. For example, the Germplasm Enhancement of Maize (GEM) Program uses different maize landraces from around the world to develop germplasm less vulnerable to crop pathogens, insects, and abiotic stresses (USDA-ARS 2010a). Institutions world-wide recognize the benefits of germplasm preservation, and the Agricultural Research Service of the United States Department of Agriculture (USDA-ARS) coordinates an extensive network of stations across the United States charged with obtaining, preserving, regenerating, distributing, enhancing, and maintaining the health of plant germplasm (USDA-ARS 2010b). Although material may be stored in many forms, e.g., seeds, tubers, whole plants, tissues, or roots, seed storage is the most common.

Stored seeds do not live forever and eventually need regeneration (Sharrock, Anishetty, and Fowler 1998). The life expectancy of stored seeds may be predicted by a seed-viability model. Some models incorporate storage temperature and seed-moisture content to predict germination as a function of seed age. Examples include one proposed by Ellis and Roberts (1980) and its modifications (Hay et al. 2003; Mead and Gray 1999; Tang et al. 2000). Other models, such as the Johnson–Mehl–Avrami kinetics model (Walters, Wheeler, and Grotenhuis 2005), consider temperature and moisture to be fixed and express germination only as a function of seed age. Because we are using data from a controlled environment seed-storage facility, we focus on this second model type.

The Avrami model's sigmoidal shape readily accounts for seed lots that maintain initially high viability levels over a long period of time and have abrupt decay rates (Walters, Wheeler, and Grotenhuis 2005). But it poorly describes seed lots with germination values that increase during early years of storage (e.g., first 7 to 10 years) or lots with germination values well below 100 % at the start of storage. When the Avrami model is fit to lots with either characteristic, the predicted viability curve may be concave up, and model predictions can be unrealistic. Examples of these poor model fits may be found in Section 5.

A suitable viability model can provide guidelines for genebank managers to schedule viability tests. An appropriate model will flag seed lots that are reaching a specified minimum germination level so they can be tested more frequently. A good model should also accurately identify seed lots that retain moderate to high viability levels. These could be tested less frequently.

We propose that seed viability can be modeled adequately as a quadratic function of seed age. The general 3-parameter quadratic regression is

$$y = \beta_0 + \beta_1 t + \beta_2 t^2 + \epsilon \quad (1.1)$$

where  $t$  is storage time and  $y$  the observed germination percentage. We assume  $\epsilon$ 's are independent and identically distributed as normal random variables.  $\beta_0$  represents the initial germination percentage at  $t = 0$ , while  $\beta_1$  and  $\beta_2$  are parameters describing how germination levels change over time. The quadratic model was efficiently fit to data on a large number of lots, each with a few observations, by assuming that the three regression coefficients for each lot follow a multivariate normal distribution (Laird and Ware 1982).

Seed managers need to know when the germination value for a seed lot is predicted to reach a critical value. Given predictions of the regression coefficients for a particular lot,

the time to reach a critical value is calculated by solving (1.1) for  $t$ . Our decision rule is based on the distribution of predicted  $t$  for each seed lot. We chose a Bayesian approach to estimate these distributions because it was considerably faster than a non-parametric bootstrap and more robust to misspecification of the random-effects distributions than a parametric bootstrap.

In the following pages, we argue that a multi-level model with Equation (1.1) as the first-level model can be used by genebank managers to predict viability and determine viability testing times for individual seed lots. Development and support of this model is derived from maize-viability testing data. Follow-up testing has been done to verify the accuracy of our model-based predictions.

## 2. DATA DESCRIPTION

Our maize data come from the USDA-ARS, North Central Regional Plant Introduction Station (NCRPIS) located in Ames, Iowa. There are 2,833 seed lots representing 2,314 unique accessions of maize with at least three viability tests conducted at different times, resulting in a total of 11,558 observations. This data set is based on regenerated and currently stored lots produced since 1948.

In any given seed lot, the expected initial germination value is not 100 %. Additionally, a living seed may not germinate during a viability test. A given seed lot may have a portion of dead seeds, and some seeds may remain dormant during testing. Seed dormancy is defined as a unique ordering of blocks to seed germination that has evolved over time to adapt plants to climate patterns and the abundance or scarcity of resources, such as water, oxygen, nutrients, and light (Finch-Savage and Leubner-Metzger 2006; Holdsworth, Bentsink, and Soppe 2008). During testing, the researcher may not provide conditions that release dormancy for all seeds. And, it is often difficult to distinguish with certainty between dormant and inviable seeds. Thus, we conducted analyses on what the seed-testing literature calls “normal germination percentages.” The fate of a seed in a viability test is classified as normal germination, abnormal germination, dormant, or dead. A seed has normal germination when the resulting seedling has essential structures that indicate it will develop into a mature plant (ISTA 2009).

Plots of trends in germination values of these 2,833 seed lots over time revealed the potential presence of after-ripening. Seed after-ripening refers to metabolic processes that must occur in otherwise mature seeds before germination can occur. After-ripening can be most readily observed when seeds are stored for a period of time at room-temperature conditions (Bewley 1997; Finch-Savage and Leubner-Metzger 2006; Leubner-Metzger 2003). But after-ripening at a reduced rate may also occur at colder temperatures (e.g. 4 °C) for some species (Chantre et al. 2009; Steadman, Crawford, and Gallagher 2003; Widrechner 2007), sometimes even at temperatures slightly below the freezing point (Sivakumar et al. 2006; Wang, Bai, and Tanino 2004). In other words, a slow after-ripening process in cold storage may lead to a gradual increase in normal germination percentage, which is eventually counteracted by a long-term decline in overall viability.

We have data on normal seed germination percentage and seed age. Seed age was measured in whole-year increments starting at 0.5 years. Seed germination was recorded in increments of a whole percent based on tests conducted on 200-seed samples. In each viability test, four groups of 50 seeds were tested on four separate towels grouped together. Consequently, all 200 seeds of a test were not subjected to a completely homogeneous test environment.

Much of the seed for this analysis has been stored under similar conditions over time. All 2,833 seed lots were dried, packed in clear, moisture-impermeable containers, and stored in a room held at 4 °C and 25 % relative humidity for the last 15 years. Due to historical changes in seed-storage conditions, some older seed lots were not continuously stored under these conditions, but instead were subjected to higher relative humidity levels. Also, different personnel conducted these germination tests over the past 61-year span. When considering the inconsistent storage conditions for older seeds and the execution of tests, we expect some overdispersion in the data.

The number of seed tests for each seed lot ranged from 3 to 7. 1,180 lots were tested 3 times; 914 lots were tested 5 times; 519 lots were tested 4 times; and the remaining 220 were tested 6 or 7 times. Interval length between any two consecutive tests of a seed lot varied. The median length between successive tests was 6 years, with 50 % between 5 and 7 years, reflecting past practices at the NCRPIS. There was one seed lot where 32 years passed before it was tested again.

### 3. MODELING SEED VIABILITY FOR MANY SEED LOTS

In this section, we outline two viability models that have different shapes of seed-viability curves. The first model is a parabola. Germination is described through a quadratic form of seed age. Parabolic viability curves can describe seeds lots with after-ripening. The second model relates seed age and germination through Johnson–Mehl–Avrami kinetics (Walters, Wheeler, and Grotenhuis 2005). These curves monotonically decrease over time and do not accommodate after-ripening. Both models include seed-lot specific random effects for each of the parameters in the respective model.

#### 3.1. QUADRATIC RANDOM COEFFICIENTS MODEL

Motivation for our model in Equation (1.1) comes from the data patterns. Many of the seed lots had germination patterns that followed 1 of 3 patterns (Figure 1(a)). Declines in seed viability for “high-viability” lots were relatively small or unobserved during the duration of testing. Declines in viability were observed in the “traditional” and “after-ripening” patterns. In a “traditional” seed lot, germination monotonically decreases over time. In an “after-ripening” lot, germination *increases* during the early years of storage, then decreases.

Figure 1(a) illustrates two important features of seed germination: the initial germination at time 0 is not always 100 %, and germination values may increase over time. Similar to a log-linear model of viability (Ellis and Roberts 1980) and variants thereof (Hay et al. 2003; Mead and Gray 1999; Tang et al. 2000), our model allows for different initial ( $t = 0$ )

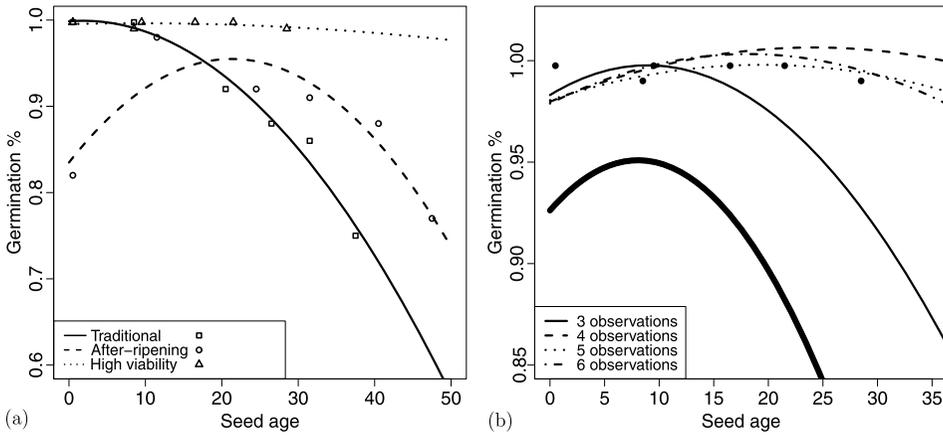


Figure 1. (a) Plot of 3 Prevalent Types of Seed-Lot Data Patterns: The points on the plot represent data from three separate seed lots. Curves are drawn by using ordinary least squares estimates of the parameters in Equation (1.1). (b) Shrinkage Plot: This plot shows how curves from one maize seed lot shrink to the overall curve (thick solid curve) of the maize seed collection when information about a seed lot is left out of the model. The first three testing observations were used to generate the thin solid curve; the first four were used to create the dashed curve; the first five were used to create the dotted line; all observations were used to create the dotted/dashed line.

germination values. Unlike many other models (Ellis and Roberts 1980; Hay et al. 2003; Mead and Gray 1999; Tang et al. 2000; Walters, Wheeler, and Grotenhuis 2005), our model allows for after-ripening. Figure 1(a) illustrates the flexibility of our quadratic model (the smooth curves) to conform to the three common data patterns.

We have viability data from thousands of maize seed lots, but few germination tests per seed lot. Biologically, we expect that seed-germination curves for different seed lots from a maize collection stored under common conditions will share some viability characteristics over time. Instead of fitting individual curves to each seed lot, which requires at least three observations per seed lot, we fit a multi-level model to the entire maize collection of 2,833 seed lots. This provides seed-lot-specific viability curves from which we estimate a seed-lot-specific testing age.

Using multi-level notation (Singer and Willett 2003), our model is as follows.

**Level I:**

$$y_{i,j} = \beta_{0,i} + \beta_{1,i}t_{i,j} + \beta_{2,i}t_{i,j}^2 + \epsilon_{i,j} \tag{3.1}$$

$$\epsilon_{i,j} \sim N\left(0, \frac{1}{\tau^2}\right) \tag{3.2}$$

**Level II:**

$$\beta_{0,i} = \beta_0 + \zeta_{0,i} \tag{3.3}$$

$$\beta_{1,i} = \beta_1 + \zeta_{1,i} \tag{3.4}$$

$$\beta_{2,i} = \beta_2 + \zeta_{2,i} \tag{3.5}$$

$$\begin{aligned} \begin{bmatrix} \zeta_{0,i} \\ \zeta_{1,i} \\ \zeta_{2,i} \end{bmatrix} &\sim \text{MVN}(\mathbf{0}, \Sigma) \quad \text{or} \\ \begin{bmatrix} \beta_{0,i} \\ \beta_{1,i} \\ \beta_{2,i} \end{bmatrix} &\sim \text{MVN}(\boldsymbol{\beta}, \Sigma) \end{aligned} \quad (3.6)$$

where

$i$  = the seed lot,  $1 \leq i \leq 2,833$ ,

$j$  = the testing occasion,  $1 \leq j \leq n_i$ , and  $n_i$  is the number of tests for seed lot  $i$ ,

$y_{i,j}$  = the germination level for the  $i$ th seed lot on testing occasion  $j$ ,

$t_{i,j}$  = the seed age of seed lot  $i$  on testing occasion  $j$ ,

$\tau^2$  = the precision of the conditionally independent and identically distributed  $\epsilon_{i,j}$ 's,

$\beta_0$  = the mean initial germination for the entire maize seed collection,

$\beta_1$  = the mean of the linear coefficient of the entire maize seed collection,

$\beta_2$  = the mean of the quadratic coefficient of the entire maize seed collection,

$\boldsymbol{\beta}' = [\beta_0, \beta_1, \beta_2]$ , and

$\Sigma$  = a 3 by 3 unstructured covariance matrix.

This hierarchical formulation garners information from all maize seed lots and uses it to generate a shrunk curve for each lot. Figure 1(b) shows how shrinkage occurs in our data. For a seed lot that has six actual germination tests, the thin solid curve was produced from the first three tests. If a line were fit only to these points (e.g. using ordinary least squares), the curve would be concave up. This curve would be implausible because the model would indicate that germination rises above 100 % after year 10. With a multi-level model, information from the maize collection shrinks these data effects into a curve that resembles the overall curve (thick solid line). However, as we incorporate more information about the lot into our model, the successive curves tend to follow the lot data more closely. This lot has 100 % germination at tests 3, 4, and 5. When the number of observations increases, the predicted viability curves begin to flatten and approach the seed-lot-specific's observed values.

### 3.1.1. Bayesian Estimation of the Quadratic Multi-level Model

The parameters in Equations (3.1), (3.2), and (3.6) were estimated through a Bayesian approach (Gelman and Hill 2007). The decision rule that we develop later is based on quantiles of the distribution of a non-linear function of the seed-lot-specific parameters in Equation (3.1). More details on these age predictions are covered in the Seed-Testing

Predictions section. Computations under the Bayesian approach were substantially faster than under a non-parametric bootstrap.

Since we had no a priori knowledge of the prior distributions of the parameters  $\tau^2$ ,  $\beta$ , or  $\Sigma$ , diffuse priors were specified. Note that the matrix in Equation (3.9) is an inverse variance-covariance matrix. We have

$$\tau^2 \sim \Gamma(0.001, 0.001) \quad (3.7)$$

$$\beta \sim \text{MVN} \left( \mathbf{0}, \begin{bmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 100 \end{bmatrix} \right) \quad (3.8)$$

$$\Sigma^{-1} \sim \text{Wishart} \left( \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, 3 \right). \quad (3.9)$$

We evaluated the sensitivity of the posterior distributions to the choice of prior distribution for  $\tau^2$  in (3.7). We considered three prior distributions:  $\Gamma(0.01, 0.01)$ ,  $\Gamma(0.001, 0.001)$ , and *Uniform*(0, 10000) (Gelman 2006). The posterior distributions of  $\tau^2$  coincided under (3.7) and *Uniform*(0, 10000) priors, which are more diffuse than a  $\Gamma(0.01, 0.01)$  prior. Also, the medians of the posterior distributions of all parameters in the Bayesian model closely resembled the estimates from a random coefficient model fitted by restricted maximum likelihood. This indicates that the prior assumptions, (3.7)–(3.9), minimally influenced the posterior distributions of the parameters (Gelman et al. 2004). We report results using  $\Gamma(0.001, 0.001)$  as the prior of  $\tau^2$ .

Estimation of the parameters was carried out through Markov Chain Monte Carlo (MCMC) methods in WinBUGS (Lunn et al. 2000). A burn-in period of 40,000 MCMC draws was selected by visually inspecting parameter trace plots and considering the effective number statistic,  $n_{\text{eff}}$ . To verify appropriate mixing of draws from the posterior distributions, three chains were initialized at disparate regions of the parameter space. Then we looked for  $\hat{R}$  statistics near a value of 1 and visually inspected trace plots again (Gelman et al. 2004). In total, there were 10,000 draws from each posterior distribution after burn-in, but we thinned each chain by a factor of 15 to avoid computer memory issues due to the large number of model parameters.

### 3.2. AVRAMI FORMULATION OF THE SEED-VIABILITY MODEL

Our multi-level quadratic model is not a commonly used seed-viability model. We compared it to a multi-level form of a commonly used viability model described by Walters, Wheeler, and Grotenhuis (2005). They modeled seed viability over time through Johnson–Mehl–Avrami kinetics (Williams et al. 1993). This model is customarily fit in its linearized form:

$$\ln(-\ln(y_{i,j})) = m_i * \ln(t_{i,j}) + c_{i,0} + \epsilon_{i,j} \quad (3.10)$$

where  $c_{i,0}$  is the initial germination and  $m_i$  is the slope on the complimentary log–log scale.  $t_{i,j}$  and  $y_{i,j}$  are defined in Section 3.1. Since there were seed tests that yielded 0 %

or 100 % germination, we modified them with Berkson’s empirical adjustment (Berkson 1953). Specifically, 0 % germinations values were replaced with  $1/(2(200)) = 0.0025$ , and 100 % germination values were replaced with 0.9975.

We assumed that  $\epsilon_{i,j} \sim N(0, 1/\tau^2)$  in Equation (3.10). As in Equation (3.6), we fitted a multivariate normal distribution with an unstructured covariance matrix to  $[m_i, c_{i,0}]'$ . Lastly, diffuse priors were assumed for the precision, the coefficients vector, and the variance of coefficients. These prior assumptions mirrored those outlined in Section 3.1.1. Procedures to determine appropriate burn-in and mixing were also the same.

#### 4. SEED-TESTING PREDICTIONS

After fitting the quadratic and Avrami models, we formulated seed-lot specific predictions of the age when viability drops below a critical amount, *CRV*. *CRV* is the lowest acceptable level of viability as specified by a genebank manager. Selection of a *CRV* germination value determines the test schedule for a seed lot. In this paper, we selected  $CRV = 0.50$ . The critical germination rate could easily be modified to meet other genebank standards.

Given a *CRV* and parameter estimates from either of the models in Sections 3.1 or 3.2, we inverted Equations (3.1) and (3.10) and solved for seed age. For our quadratic model, the predicted testing age of lot *i* at a posterior draw of the vector  $[\hat{\beta}_{0,i}, \hat{\beta}_{1,i}, \hat{\beta}_{2,i}]'$  is

$$\hat{t}_i^Q = \frac{-\hat{\beta}_{1,i} \pm \sqrt{\hat{\beta}_{1,i}^2 - 4\hat{\beta}_{2,i}(\hat{\beta}_{0,i} - CRV)}}{2\hat{\beta}_{2,i}} \tag{4.1}$$

The *Q* superscript signifies that  $\hat{t}_i^Q$  is a prediction from our quadratic viability model. The predicted curvature/concavity (evidenced through sign of  $\hat{\beta}_{2,i}$ ), prediction of initial germination ( $\hat{\beta}_{0,i}$ ), and location of the predicted curve’s apex all affect which root will be used as the predicted testing age for seed lot *i*. Also, there will be instances where no root exists. Appendix details when the positive or negative discriminant is used and how we handled situations where the roots do not exist. The inversion of Equation (3.10) at some posterior draw of the  $[\hat{m}_i, \hat{c}_{i,0}]'$  vector has a single root given by

$$\hat{t}_i^A = e^{\frac{\ln(-\ln(CRV)) - \hat{c}_{i,0}}{\hat{m}_i}} \tag{4.2}$$

The superscript *A* signifies that  $\hat{t}_i^A$  is a prediction from the Avrami model.

Sampling distributions of  $\hat{t}_i^Q$  and  $\hat{t}_i^A$  do not have analytical solutions, but they can be estimated by substituting realizations from either the posterior distribution of  $[\beta_{0,i}, \beta_{1,i}, \beta_{2,i}]$  into (4.1) or the posterior distribution of  $[m_i, c_{i,0}]$  into (4.2). These posterior distributions provide a range of reasonable seed ages that may be used to schedule future viability tests.

#### 5. COMPARISON OF AVRAMI AND QUADRATIC MODELS

We compare the Avrami and quadratic models by assessing their predictions. A prediction that a seed lot will not reach 50 % viability until a seed age of 800 years is contrary to

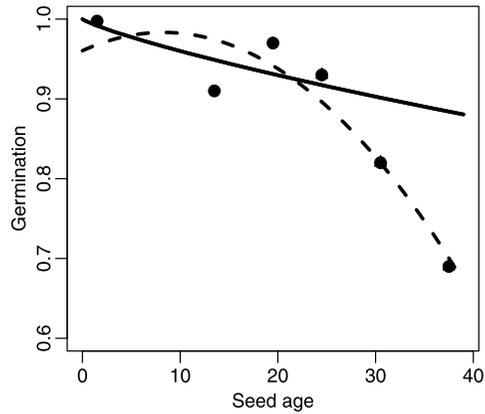


Figure 2. Comparison of Avrami (solid line) and Quadratic (dashed line) Curves for 1 Lot: Medians of the posterior distributions of the coefficients were used to create the curves. The quadratic curve captures the germination trend better than the Avrami curve does.

a wealth of past experience for crop seeds (Nagel and Börner 2010). We use the medians of the  $\hat{t}_i^Q$  and  $\hat{t}_i^A$  distributions as point predictions of the time to 50 % viability for the quadratic and Avrami models because they are “typical” values of the non-symmetrical posterior distributions.

Predictions of time to 50 % viability from the Avrami model fail our criteria. In that model, over 94 % of the lots have a test age, at  $CRV = 0.50$ ,  $>200$  years, with 75 % having predictions  $>864$  years and 23 % with predictions  $>10,000$  years. One lot has a predicted test age of 9,439,000 years. These predictions are unrealistic and impractical. In contrast, 75 % of the predictions of time to 50 % viability using the quadratic model are  $<65$  years.

To understand why predictions from the Avrami model are so large, we investigated predictions of individual seed lots with 6 or 7 historical tests. Figure 2 shows Avrami and quadratic curves using median parameter estimates for a seed lot that has six test occasions. The predicted germination values of the Avrami curve are calculated through Equation (4.2). The corresponding median  $\hat{t}_i^Q$  and  $\hat{t}_i^A$  are 46.5 and 296.9 years, respectively. 183 of the 220 lots with six or seven previous seed tests have similar concave-down data patterns as in Figure 2. Under the Avrami model, the predicted death rate, or the first derivative of germination with respect to age, for this seed lot *decreases* as age increases. A concave-up shape is observed. The data do not suggest this pattern. They show that the death rate increases as time increases. A quadratic model captures this essential concave-down seed feature. The Avrami model does not accurately represent the death rates with concave-down curvatures. As a result, many of its test-age predictions for seed lots are extreme. Further analyses in this paper are conducted using only the quadratic model.

## 6. SELECTION AND EVALUATION OF A $\hat{t}_i^Q$ QUANTILE

The purpose of fitting a model is to create practical predictions of when we expect a seed lot to drop below a  $CRV$  level. However, managers want to regrow a seed lot *before*

it reaches the *CRV*. Otherwise, there may not be a large enough portion of viable seeds to reproduce the lot and preserve its genetic profile. When one tests seeds at their predicted age, one of two negative consequences may happen. If the observed germination is above the *CRV*, money will have been spent to test seeds unnecessarily. If the observed germination is below the *CRV*, viability may no longer be acceptable. The costs of these consequences are not equal. The cost of an additional test is much less than the cost of unacceptable viability. Assuming our quadratic model is appropriate, predictions based on the posterior medians of  $\hat{t}_i^Q$  are equally likely to be less than or greater than the true age when *CRV* occurs. Instead of using medians, we suggest using a lower quantile. Choice of which lower quantile is investigated through a receiver operating characteristic (ROC) curve that summarizes the true positive and the false positive rates over a range of quantiles.

### 6.1. ROC CURVE DEVELOPMENT

We need to evaluate decision rules based on the posterior distributions of the  $\hat{t}_i^Q$ 's. This is difficult because the true age at which a lot reaches a *CRV*, the “gold standard,” is unknown. It is impossible to continuously record the true viability of a seed lot. We can only test *all* seeds once. In place of a gold standard, for a given seed lot  $i$ , we compare a prediction of its  $\hat{t}_i^Q$  distribution to its status at the last test,  $t_{\text{last},i}$ .

The data-based standard is appropriate here because the decision to regrow a seed lot will be based on the observed germination percentage at the last test. For this section, predictions were generated from the quadratic model refitted to the data set *excluding observations at the last test for each lot*. Comparisons between these predictions and  $t_{\text{last},i}$  are an out-of-sample (Harrell 2001) assessment of a model's ability to predict future performance. If an observed germination at  $t_{\text{last},i}$  is less than *CRV*, the correct decision is a predicted  $\hat{t}_i^Q$  that is earlier than  $t_{\text{last},i}$ . In other words, for a lot that is below *CRV* at  $t_{\text{last},i}$ , we want a prediction that schedules a germination test before that last testing age. If observed germination at  $t_{\text{last},i}$  is above *CRV*, then the correct decision is a predicted  $\hat{t}_i^Q$  that is later than  $t_{\text{last},i}$ . We want a prediction that schedules a germination test later than the most recent test.

The positive population,  $P$ , is those seed lots that needed to be regrown before  $t_{\text{last},i}$  because their germination values at  $t_{\text{last},i}$  are smaller than the *CRV*. There were 147 seed lots in  $P$ . These seed lots should have been regrown before  $t_{\text{last},i}$ . Table 1 outlines the two decisions associated with individuals in  $P$ . A true positive,  $TP$ , occurs if a seed lot from  $P$  has a predicted age that is before its  $t_{\text{last},i}$ , i.e., the model prediction is “regrow before  $t_{\text{last},i}$ .” A false negative,  $FN$ , is the case where a seed lot from  $P$  has a predicted test age after its  $t_{\text{last},i}$ , i.e. the model predicts “no need to regrow before  $t_{\text{last},i}$ .”

The negative population,  $N$ , is those seed lots that do not need to be regrown before  $t_{\text{last},i}$  because their germination values at  $t_{\text{last},i}$  are greater than the *CRV*. There were 2,686 sampled lots in  $N$ . A false positive,  $FP$ , is a member of  $N$  where its predicted test age is earlier than its  $t_{\text{last},i}$ . But a viability test has already been conducted at a later seed age,  $t_{\text{last},i}$ , and the germination was greater than the *CRV*, so the prediction is misleading and cannot be acted upon. The final class, true negative or  $TN$ , describes where a seed lot from

Table 1. Test Criteria and Decisions for a 50 % Critical Regrow Value:  $t_\alpha$  represents the  $\alpha$  quantile of a  $t$  posterior distribution.

Decision based on $t_\alpha$		
Predicted test age was <i>before</i> last observed germination test	Observed germination at last germination test	Predicted test age was <i>after</i> last observed germination test
False Positive ( <i>FP</i> ) Incorrect prediction	100 % ↑ ↑ 50 %	True Negative ( <i>TN</i> ) Correct prediction
Seed lot age: $t_0$ <span style="display: inline-block; border-bottom: 1px solid black; width: 60%;"></span> $t_z$		
True Positive ( <i>TP</i> ) Correct prediction	49 % ↑ ↑ 0 %	False Negative ( <i>FN</i> ) Incorrect prediction

$N$  has a predicted test age after its  $t_{\text{last},i}$ . Table 1 lays out all possible classifications that result from a prediction.

We use an  $\alpha$  quantile of the posterior distribution of  $\hat{t}_i^Q$ , calculated by omitting germination at each seed lot’s last test data, as the prediction for seed lot  $i$ . Different choices of  $\alpha$  lead to different sets of predictions. Therefore, the counts in Table 1 change. For example, predictions based on a large  $\alpha$  quantile such as  $\alpha = 0.9$  will have more *FN*’s than would predictions based on an  $\alpha = 0.2$  quantile. Paraphrased, if we had not “known” germination at  $t_{\text{last},i}$  and used  $\hat{t}_{0.9,i}$ ’s instead of  $\hat{t}_{0.2,i}$ ’s as the subsequent test ages, more seed lots would have dropped below *CRV* germination levels when tested. Those seed lots possibly would have been lost. Conversely, if  $\hat{t}_{0.2,i}$  ages were chosen, then we would commit more *FP*’s. We would have unnecessarily spent money on tests, in retrospect.

We summarize the relationship between the choice of  $\alpha$  and the true positive and false positive rates using an ROC curve, estimated by non-parametric smoothing (Krzanowski and Hand 2009). This provides a smooth, differentiable estimate of the ROC curve and facilitates estimation of optimal  $\alpha$ -quantile predictions (to be discussed in the next section). For each  $\alpha$ , we counted the number of *TP*’s in population  $P$  and the number of *FP*’s in population  $N$  and calculated the corresponding rates,  $tpr(\alpha)$  and  $fpr(\alpha)$ . Because a kernel smoother has problems near the boundaries,  $\alpha = 0$  and  $\alpha = 1$ , we estimated the density on the  $\text{logit}(\alpha)$  scale. We used a Berkson (1953) correction of  $1/(2(2,001))$  for the empirical  $\alpha_i$ ’s that were 0, i.e. all draws in lot  $i$ ’s posterior distributions of  $\hat{t}_i^Q$  were greater than the seed lot’s last test age. This happened with 1,057 lots. The density with regard to  $\text{logit}(\alpha)$  in the negative and positive populations,  $\hat{f}_N(x)$  and  $\hat{f}_P(x)$ , respectively, was estimated by using a biweight kernel smoother (Silverman 1986). To estimate  $fpr(\alpha)$  for a given  $\alpha$ , we integrated  $\hat{f}_N(x)$  for  $\text{logit}(x) \geq \text{logit}(\alpha)$ . For the same  $\alpha$ ,  $tpr(\alpha)$  was calculated by integrating the density  $\hat{f}_P(x)$ .

The points on the ROC curve in Figure 3(b) represent  $(fpr(\alpha), tpr(\alpha))$ ’s for various  $\alpha$  thresholds. The area under the ROC curve is 0.87. Our  $\alpha$  quantile rule can discriminate

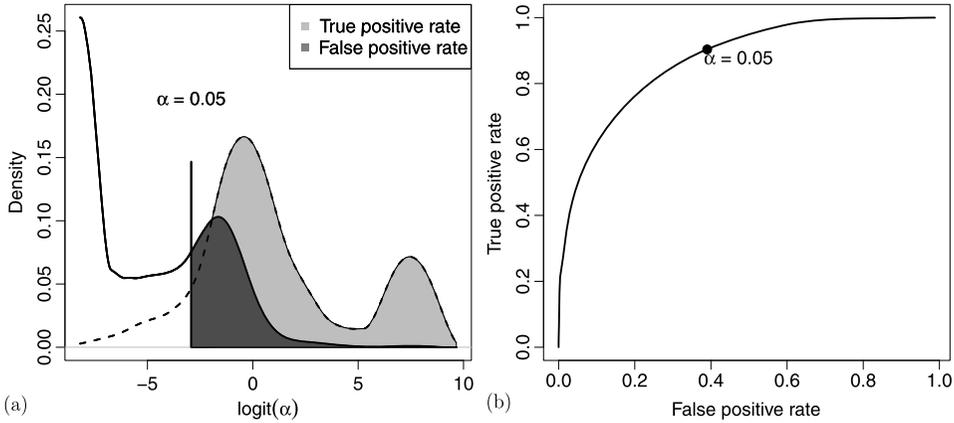


Figure 3. (a) Non-Parametric Density Estimates of  $\alpha$  with Respect to Negative (solid line) and Positive (dashed line) Populations: Density curves pertain to the logit transformation of  $\alpha$ . Areas shaded to the right represent the false positive and true positive rates corresponding to an  $\alpha = 5\%$  threshold. (b) ROC Curve: Point on the ROC curve corresponds to  $fpr$  and  $tpr$  shaded area in (a).

between the populations  $N$  and  $P$  (Krzanowski and Hand 2009). The point on Figure 3(b) corresponds to a threshold of 5%.

### 6.2. SELECTION OF AN OPTIMAL QUANTILE OF $\hat{t}_i^Q$ DISTRIBUTIONS

The ROC curve summarizes the error rates for various  $\alpha$  choices. However, the curve fails to account for the costs associated with misclassifications and the relative proportion of observations from  $N$  and  $P$ . These are crucial components in selecting an optimal threshold. The estimated relative proportion of lots in  $P$  is rather small at  $147/2,833$ . Consequently, we applied the cost function

$$C(\alpha) = q * (1 - tpr) * C(N|P) + (1 - q) * fpr * C(P|N) \tag{6.1}$$

as the criterion in choosing an optimal threshold. The relative proportion of seed lots in  $P$ ,  $q$ , is estimated from our data. In Equation (6.1), costs associated with an  $FN$  and an  $FP$  are represented as  $C(N|P)$  and  $C(P|N)$ , respectively. We believe that the cost of potentially losing a seed lot, an  $FN$ , is considerably higher than the cost of conducting a premature test. We consider cost functions where  $C(N|P)/C(P|N) > 1$ .

Under the criterion in Equation (6.1), the  $\alpha$  that minimizes the cost of misclassification corresponds to the point on the ROC curve with a derivative of

$$\frac{(1 - q)C(P|N)}{qC(N|P)} \tag{6.2}$$

Provided the slope is well defined at a given  $\alpha$ , the derivative of the ROC curve at  $\alpha$  is equal to the ratio of  $P$  and  $N$  densities evaluated at that  $\alpha$  (Krzanowski and Hand 2009).

A researcher must choose a suitable cost ratio to determine an optimal  $\alpha$  quantile. We estimate that the expense associated with potentially losing a lot to that of possibly conducting a premature test is approximately 30. For a  $C(N|P)/C(P|N) = 30$  (derivative of 0.61), we have an optimal threshold of  $\alpha = 0.05$ . However, if one believes that the cost

Table 2. Count Results from the 5-Stratum Follow-Up Study: The number of false negative and false positive cases are in columns *FN* and *FP*, respectively. The number of sampled lots from the negative and positive populations are in columns *N* and *P*, respectively. The Population column shows the number of lots out of the original 2,833 that are in the corresponding stratum. The Retest column indicates that we sampled 25 lots from each stratum.

Stratum	Size of the		Measured germination in 2009		Number of misclassifications	
	Population	Retest	<i>N</i> (>50 %)	<i>P</i> (<50 %)	<i>FP</i>	<i>FN</i>
1	550	25	15	10	15	0
2	181	25	24	1	20	0
3	208	25	24	1	18	1
4	164	25	25	0	2	0
5	308	25	25	0	1	0

of premature testing equals the cost of losing a seed lot ( $C(N|P)/C(P|N) = 1$ ), then  $\alpha = 0.99$  (derivative of 18.27) is suitable. In contrast, one may specify a high cost ratio of 100. This corresponds to an  $\alpha = 0.002$  (derivative of 0.18). In Section 6.1, it was noted that the densities of *P* and *N* are not unimodal. This results in derivatives that may not be well-defined. Graphical exploration suggests that unique derivatives do exist for cost ratios of 1, 30, and 100.

### 6.3. EVALUATION OF $\alpha = 0.05$ QUANTILE RULE

In this section, we investigate the effectiveness of predictions based on the  $\alpha = 0.05$  quantiles of the  $\hat{t}_i^Q$ 's. A stratified sample of 125 seed lots was taken from the original 2,833 lots and was retested in 2009. We calculated the observed *TP* and *FP* rates and compared them to their corresponding estimated rates from Section 6.1. Since a cost ratio of 30 is associated with an  $\alpha$  of 0.05, we expected to see more premature tests, *FP*'s, than late tests (i.e., tests with germination values below 50 %), *FN*'s.

We stratified the maize seed lots into 5 strata to better understand the decision rule's performance (Table 2). An early implementation of our model was used to assign seed lots to strata. All data points were used to generate predictions. Lots with predicted viability of 50 % before the year 2000 were assigned to stratum 1. Seed lots with predicted viability around 50 % between 2007 and 2012 were allocated to stratum 2. Strata 3 and 4 were composed of seed lots with predicted test times between 2012 and 2018. Stratum 3 included only lots with three or four previous tests whereas stratum 4 included lots with five or more previous tests. Stratum 5 was limited to seed lots with predicted test times at or beyond 2033. 25 seed lots were randomly sampled from each stratum.

In the retest data, *N* and *P* populations were defined with respect to observed germination values of the 125 seed lots in 2009. Lots were assigned to *N* if their observed germination values were >50 % in the retest. For a lot in *N*, an *FP* would result when the predicted age,  $\hat{t}_{\alpha,i}$ , would fall below the actual age of the lot in 2009. Table 2 displays the number of *FP* and *FN* cases for each of our sampled strata. The empirical *fpr*'s, calculated as  $FP/N$ , are rather high, especially for strata 1, 2, and 3 (1, 0.83, and 0.75, respectively).

In our study, only 1 *FN* occurred. The adjusted overall *fpr* is 0.55. The adjusted overall *fnr* is 0.035.

For  $\alpha = 0.05$ , we predicted an *fnr* of 0.097 and an *fpr* of 0.39. This corresponds to the point (0.39, 0.90) in Figure 3 when using the criterion in Equation (6.2). Our observed *fnr* is better than our predicted rate, but the predicted *fpr* is more optimistic than the observed value (difference is 0.16). Still, the  $\alpha = 0.05$  decision rule was able to identify nearly all lots that fell below 50 % viability (11 out of the 12 *P* lots).

## 7. DISCUSSION

We have shown that the Avrami seed viability model (Walters, Wheeler, and Grotenhuis 2005) fails to model patterns where germination increases over the initial years of storage. Biologically, after-ripening or an initial failure to break seed dormancy may result in such increases. To model these seed phenomena simply, we fitted a 3-parameter quadratic curve of germination at specified seed ages. Presumably, lots of our maize collection may decay in similar patterns over time. Thus, we fitted a Bayesian multi-level model. Curves of seed lots with only three previous tests were shrunk to a general maize viability curve, but curves of seed lots with a large number of historical tests were more individualistic. Since the goal was to estimate when to retest seed lots, we generated predictions of time to reach a critical germination percentage from our quadratic multi-level model and compared them to predictions based on the Avrami multi-level model. We found that our predictions were much more realistic than those based on the Avrami model.

From a physiological perspective, there is no reason to assume a symmetric viability curve. The pattern of increasing germination values over initial storage may not mirror the pattern of decreasing germination values over the later storage years. However, this is not a major concern. The goal is to fit a model that provides reasonable predictions of test ages. Our retest study shows that predicted test ages based on the 5 % quantile of  $\hat{t}_i^Q$  for 56 of the 125 retested lots are smaller than the actual ages when germination falls below 50 %. This suggests that our prediction model is on the safe side. The observed *fpr* is larger than the model-based estimate of *fpr* ( $0.55 > 0.39$ ) for our  $\alpha = 0.05$  rule. These retest results suggest that if there is bias in the predictions, then it is toward premature predictions.

Predictions in this paper were based on a *CRV* of 50 % and an  $\alpha$  quantile of 5 %. “Standard” values used by various genebanks can easily replace these quantities. If lot regeneration requires a more complete representation of a sample to ensure the conservation of its genetic profile, then a genebank manager may assign a larger *CRV*. If a genebank manager wishes to have a smaller proportion of premature tests, then one can choose an  $\alpha > 5$  %. Our model is flexible in meeting managers’ needs.

A key component of our model is its ability to pool information across similar seed lots. Although it was created for a collection of maize lots, our multi-level model should easily generalize to a collection of similar lots of other plant species. Our model is not limited to a single crop.

Lastly, our model’s predictions are dynamic. Every new viability-test data point will provide information on the current viability status of an individual lot *and* on the characteristics of the entire collection. Consequently, the entire collection’s viability test schedule

can be updated with every successive test. Our populations,  $N$  and  $P$ , change over time. Thus, it is appropriate that predictions change with each new viability assessment.

### APPENDIX: CALCULATION OF $\hat{t}_i^Q$

Knowing whether one adds or subtracts the discriminant in (4.1) is determined by a seed lot’s viability curve. A quadratic viability curve may be completely characterized by three seed features. We quantify these features with three numerical estimates. The value of the quadratic term,  $\hat{\beta}_{2,i}$ , estimates the shape/convexity of a lot’s curve.  $\hat{\beta}_{0,i}$  estimates the initial germination of a lot. Last, we estimate the location of a curve’s vertex. Biologically, this should be the age at which peak germination occurs (a local maximum). However, there are cases where the vertex is a local minimum of a curve. Examples include when a lot’s historical viability tests strictly increase over time or when they remain stagnant over time and vary little. Applying derivatives to Equation (3.1), we estimate the horizontal coordinate of the vertex as  $x^* = -\hat{\beta}_{1,i}/2\hat{\beta}_{2,i}$ .

Table A.1 enumerates the cases when one takes the negative discriminant of Equation (4.1); when  $\hat{t}_i^Q$  takes a value of 0 (e.g., original lot has large portion of dead seed); or when it is 10,000 (historical data do not provide evidence of a decline in viability). Lots that may or may not exhibit after-ripening and have a concave-down curve are estimated with the C1 and C3 curves. These curves differ in their predicted initial germination. Lots with consistently high germination rates are estimated with C7 and C8.

Table A.1.  $\hat{t}_i^Q$  Table: In our example, the *CRV* (critical value) corresponds to 0.50, but it may be adjusted to any desired value. The “negative” and “positive” in the discriminant column refer to the root that is used. For curves like C7, the predicted test time is infinity, but we capped it at 10,000 years; hence, there is a column for the value of  $\hat{t}_i^Q$ . Curve C9 has a test time upon the arrival of the seed lot ( $\hat{t}_i^Q = 0$ ).  $f^* = \hat{\beta}_{0,i} + \hat{\beta}_{1,i}x^* + \hat{\beta}_{2,i}(x^*)^2$  where  $(x^*, f^*)$  is the location of the curve’s vertex.

Curve type	$x^*$	$\hat{\beta}_{2,i}$	$\hat{\beta}_{0,i}$	$f^*$	Discriminant	$\hat{t}_i^Q$
C1	$x^* > 0$	Negative	$\hat{\beta}_{0,i} > CRV$	$f^* > CRV$	Negative	
C2	$x^* < 0$	Negative	$\hat{\beta}_{0,i} > CRV$	$f^* > CRV$	Negative	
C3	$x^* > 0$	Negative	$\hat{\beta}_{0,i} < CRV$	$f^* > CRV$	Negative	
C4	$x^* < 0$	Negative	$\hat{\beta}_{0,i} < CRV$	$f^* > CRV$		0
C5	$x^* > 0$	Negative	$\hat{\beta}_{0,i} < CRV$	$f^* < CRV$		0
C6	$x^* < 0$	Negative	$\hat{\beta}_{0,i} < CRV$	$f^* < CRV$		0
C7	$x^* > 0$	Positive	$\hat{\beta}_{0,i} > CRV$	$f^* > CRV$		10,000
C8	$x^* < 0$	Positive	$\hat{\beta}_{0,i} > CRV$	$f^* > CRV$		10,000
C9	$x^* > 0$	Positive	$\hat{\beta}_{0,i} < CRV$	$f^* < CRV$		0
C10	$x^* < 0$	Positive	$\hat{\beta}_{0,i} < CRV$	$f^* < CRV$		10,000
C11	$x^* > 0$	Positive	$\hat{\beta}_{0,i} > CRV$	$f^* < CRV$	Negative	
C12	$x^* < 0$	Positive	$\hat{\beta}_{0,i} > CRV$	$f^* < CRV$		10,000

## ACKNOWLEDGEMENTS

Journal paper of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 1018, was supported by USDA-ARS Research Project 3625-21000-053, Hatch Act and State of Iowa funds. Mention of commercial brand names does not constitute an endorsement of any product by the U.S. Department of Agriculture or cooperating agencies. We would also like to thank Dr. Candice Gardner for her input.

[Received February 2011. Accepted February 2012. Published Online March 2012.]

## REFERENCES

- Berkson, J. (1953), "A Statistically Precise and Relatively Simple Method of Estimating the Bioassay with Quantal Response, Based on the Logistic Function," *Journal of the American Statistical Association*, 48, 565–599.
- Bewley, J. D. (1997), "Seed Germination and Dormancy," *Plant Cell*, 9, 1055–1066.
- Chantre, G. R., Batlla, D., Sabbatini, M. R., and Orioli, G. (2009), "Germination Parameterization and Development of an After-Ripening Thermal-Time Model for Primary Dormancy Release of *Lithospermum arvense* Seeds," *Annals of Botany*, 103, 1291–1301.
- Ellis, R. H., and Roberts, E. H. (1980), "Improved Equations for the Prediction of Seed Longevity," *Annals of Botany*, 45, 13–30.
- Finch-Savage, W. E., and Leubner-Metzger, G. (2006), "Seed Dormancy and the Control of Germination," *New Phytologist*, 171, 501–523.
- Gelman, A. (2006), "Prior Distributions for Variance Parameters in Hierarchical Models," *Bayesian Analysis*, 1, 514–534.
- Gelman, A., and Hill, J. (2007), *Data Analysis Using Regression and Multilevel/Hierarchical Models*, New York: Cambridge University Press.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004), *Bayesian Data Analysis* (2nd ed.), Boca Raton, FL: Chapman and Hall/CRC Press.
- Harrell, F. E. Jr. (2001), *Regression Modeling Strategies With Applications to Linear Models, Logistic Regression, and Survival Analysis*, New York: Springer.
- Hay, F. R., Mead A., Manger, K., and Wilson, F.J. (2003) "One-Step Analysis of Seed Storage Data and the Longevity of *Arabidopsis thaliana* Seeds," *Journal of Experimental Botany*, 54, 993–1011.
- Holdsworth, M. J., Bentsink, L., and Soppe, W. J. J. (2008), "Molecular Networks Regulating *Arabidopsis* Seed Maturation, After-Ripening, Dormancy and Germination," *New Phytologist*, 179, 33–54.
- ISTA (2009), *International Rules for Seed Testing*, Bassersdorf: International Seed Testing Association.
- Krzanowski, W. J., and Hand, D. J. (2009), *ROC Curves for Continuous Data*, Boca Raton, FL: Chapman and Hall/CRC Press.
- Laird, N. M., and Ware, J. H. (1982), "Random-Effects Models for Longitudinal Data," *Biometrics*, 38, 963–974.
- Leubner-Metzger, G. (2003), "Functions and Regulation of  $\beta$ -1, 3-Glucanases During Seed Germination, Dormancy Release and After-Ripening," *Seed Science Research*, 13, 17–34.
- Lunn, D. J., Thomas, A., Best, N., and Spiegelhalter, D. (2000), "WinBUGS—A Bayesian Modelling Framework: Concepts, Structure, and Extensibility," *Statistics and Computing*, 10, 325–337.
- Mead, A., and Gray, D. (1999), "Prediction of Seed Longevity: A Modification of the Shape of the Ellis and Roberts Seed Survival Curves," *Seed Science Research*, 9, 63–73.
- Nagel, M., and Börner, A. (2010), "The Longevity of Crop Seeds Stored Under Ambient Conditions," *Seed Science Research*, 20, 1–12.
- Sharrock, S., Anishetty, N. M., and Fowler, C. (1998), "Discussion Paper on the Global Regeneration Need: Evidence Collected from Country Reports Prepared for the International Technical Conference on Plant Genetic Resources," in *Regeneration of Seed Crops and Their Wild Relatives*, eds. J. Engels and R. R. Rao, Rome: International Plant Genetic Resources Institute, pp. 86–104.

- Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, London: Chapman and Hall.
- Singer, J. D., and Willett, J. B. (2003), *Applied Longitudinal Data Analysis*, New York: Oxford University Press.
- Sivakumar, V., Anandalakshmi, R., Warriar, R. R., Tigabu, M., Odén, P. C., Vijayachandran, S. N., Geetha, S., and Singh, B. G. (2006), "Effects of Presowing Treatments, Desiccation and Storage Conditions on Germination of *Strychnos nux-vomica* Seeds, a Valuable Medicinal Plant," *New Forests*, 32, 121–131.
- Steadman, K. J., Crawford, A. D., and Gallagher, R. S. (2003), "Dormancy Release in *Lolium rigidum* Seeds is a Function of Thermal After-Ripening Time and Seed Water Content," *Functional Plant Biology*, 30, 345–352.
- Tang, S., TeKrony, D. M., Egli, D. B., and Cornelius, P. L. (2000), "An Alternative Model to Predict Corn Seed Deterioration During Storage," *Crop Science*, 40, 463–470.
- USDA-ARS (2010a), "Germplasm Enhancement of Maize," available online at: [http://www.public.iastate.edu/~usda-gem/GEM\\_Project/GEM\\_Project.htm](http://www.public.iastate.edu/~usda-gem/GEM_Project/GEM_Project.htm) (last accessed: October 20, 2010).
- (2010b), "USDA-ARS Station Facts and Purpose for Ames Plant Introduction Station," available online at: <http://www.ars.usda.gov/Main/docs.htm?docid=13482&page=3> (last accessed: October 20, 2010).
- Walters, C., Wheeler, L. M., and Grotenhuis, J. M. (2005), "Longevity of Seeds Stored in a Genebank: Species Characteristics," *Seed Science Research*, 15, 1–20.
- Wang, R., Bai, Y., and Tanino, K. (2004), "Effect of Seed Size and Sub-zero Imbibition-Temperature on the Thermal Time Model of Winterfat (*Eurotia lanata* (Pursh.) Moq.)," *Environmental and Experimental Botany*, 51, 183–197.
- Widrechner, M. P. (2007), "While They Were Asleep: Do Seeds After-Ripen in Cold Storage? Experiences with *Calendula*," *Combined Proceedings International Plant Propagators' Society*, 56, 377–382.
- Williams, R. J., Hirsh, A. G., Meryman, H. T., and Takahashi, T. A. (1993), "The High-Order Kinetics of Cytolysis in Stressed Red-Cells," *Journal of Thermal Analysis*, 40, 857–862.