A NEW BASIS FOR CHARACTERIZING GERMINATION

Robert K. Campbell and Frank C. Sorensen

Abstract
Procedures for characterizing seed lots (i.e., by mean and variation of days to germination, peak germination date, and other useful properties) are commonly complicated because frequency of days to germination does not have a normal distribution. A procedure is presented for minimizing this problem. Tests of a seed lot of Douglas-fir (Pseudotsuga menziesii (Mirb.) Franco) subjected to differing stratification periods and germination temperatures indicated that it is the rates at which seed embryos develop toward germination rather than days to germination that have a normal distribution. Consequently, by transforming cumulative percentages of germination to probits (a means for converting a normal sigmoid curve to a straight line) and fitting a straight line to probits regressed on rates rather than on days—or by plotting percentages on probit-scaled graph paper—the resulting intercept and regression coefficient can be used to easily calculate all the standard properties for characterizing a seed lot.

Additional index words: Douglas-fir, Pseudotsuga menziesii, stratification, germination rate, skewness.

Introduction
We recently started work on procedures for identifying provenance of Douglas-fir seeds of unknown origin. Germination characteristics of Douglas-fir seeds depend at least partly on stratification, germination temperature, light intensity and photoperiod (1, 14). Seeds from different provenance commonly react differently to treatment (1). Such reactions may reflect adaptation to climatic conditions at seed origin (19). Consequently, it may be possible to determine origin by establishing quantitative relationships between seed response and climatic or geographic factors of seed origin. To study these relationships, we needed to describe germination behavior in the fewest possible quantitative terms, in each of many seed lots. Herein we report a procedure for characterizing a seed lot with only two parameters, the intercept and regression coefficient of a linear regression equation. These coefficients can be estimated from the cumulative frequency distribution of germination counted over time.

1From Project Genetic Improvement of Douglas-fir and other Pacific Northwest Trees, Forestry Sciences Laboratory, Pacific Northwest Forest and Range Experiment Station, Portland, Oregon 97232.
2Principal Plant Geneticists.
Properties which have been suggested for describing germination behavior are: 1) germination capacity (total germination percentage) and 2) germination speed (germination energy) based on percentage germination at given times (18), or mean and variation in days to germination (11). Reported procedures for quantifying germination speed are based on frequency distributions for which time is measured in days to germination. The resulting distributions are usually skewed (12)—that is, the majority of seeds germinate in the first few days but a substantial part germinate much later. Consequently, the mode, the value of most frequent occurrence, is below the mean with a long tail above. The amount of skewness may vary from seed lot to seed lot depending on duration of stratification, germination temperature or other seed treatment. Recent quantitative models usually account for this by including one or more variable terms for describing skewness (5, 17). The resulting equations may be quite complex and may include several variables (9, 10). We also have observed skewness in days to germination in our tests, the amount varying with seed treatment and provenance.

Our method is based on the hypothesis that the embryo of each seed in a lot has a potential developmental rate depending on its genotype, pregermination treatment and the germination temperature. This rate represents the daily average rate of development towards germination. Thus, if it takes 25 days for a seed to germinate at a given temperature, the embryo developmental rate (r) is \( \frac{1}{25} = 0.04 \) units per day. The higher the rate the fewer the days to germinate. We further hypothesized that it is these potential rates rather than days to germination (t) that have a normal distribution. Specifically, we are proposing that the variable \( \frac{1}{t} \), where \( t \) is days (or other appropriate measure of time) to germination, is approximately normally distributed (3, p. 116). Consequently, if the time scale of germination is expressed in embryo developmental rates rather than in days, the shape of the distribution should be less affected by germination treatment and the resulting quantitative model less complex.

We report a test of this hypothesis using Douglas-fir seeds subjected to four treatments chosen to provide greatly different means and variabilities in days to germination.

Materials and Methods

Equal numbers of seed from eight trees of native Douglas-fir (Pseudotsuga menziesii (Mirb.) Franco) growing at low elevations in western Oregon made up our seed population. Seed samples were subjected to four treatments differing in stratification period (16 and 64 days) and germination temperatures (15 and 25 C, constant). Each treatment (analyzed as a separate experiment) was applied to 44 samples of 50 seeds.
Seeds were soaked in tap water at room temperature (ca. 22 C) in plastic vials for 24 hours, drained, and stratified in the same vials for the required number of days at 4 C. We then put stratified seed from each of the 176 samples onto moistened filter paper in a covered petri dish. The 44 dishes constituting an experiment (treatment) were placed in a germination chamber maintained at the assigned temperature.

At 11 scheduled times, we counted germination in four randomly chosen dishes in each experiment. Counting times were scheduled to provide equally spaced intervals on the embryo rate scale. Germination chambers were continuously lighted because counts came at odd hours, thus precluding partial photoperiods. For example, in one experiment, we counted seeds at 4.00, 4.55, 5.25, 6.26, etc., days to provide a rate scale of .25, .22, .19, .16, etc., units.

A seed was considered germinated if the radicle extended 1 mm beyond the seed coat. Remaining seeds were cut. Seeds with fully developed embryos and firm endosperms were added to the germination count to estimate potentially germinable seeds. For each counting time, percentage germination was calculated as (germinated seeds x 100) / (potentially germinable seeds).

We based statistical tests on properties of the probit transformation (p. 301-309). In brief explanation, assume that heights of men have a normal distribution. A small percentage will be shorter than 152 cm. As one progresses in cm increments to taller men, a larger and larger percentage will be shorter than the reference height. If the percentage falling below each step upwards in height is plotted against the cm value for the step, a sigmoid curve results. This curve, which is characteristic of the cumulative normal frequency distribution, can be changed to a straight line by changing percentages to probits.

The probit transformation is often applied to discontinuous data which by hypothesis are thought to be underlain by continuous data with normal distribution (3). The discontinuous data are visible, e.g., whether or not a seed has germinated. The underlying data (e.g., developmental rates of the embryos) are invisible but connected to the visible by a threshold, or discontinuity (e.g., germination). In this view, individual seeds with embryo developmental rates exceeding a certain value will be germinated, and individuals with developmental rates below the threshold will not. If, for example, the threshold rate (r) is 0.25 units/day, the proportion of seeds with rates greater than 0.25 can be estimated by counting germinated seeds at 4 days (i.e., 1/r days). By counting proportions of germinated seeds at a series of thresholds, chosen arbitrarily to sample the range of probable rates in the seed population, the underlying distribution of rates is described. If rates are distributed normally, the proportions of germinated seed when transformed to probits will plot as a straight line on the rate scale.
In each experiment, four petri dishes were randomly allocated to each rate threshold before the test. Corresponding to these thresholds, we counted seeds at 11 evenly separated times on the rate scale. Because we did not correctly guess the range of embryo developmental rates within our experiments, we did not efficiently allocate thresholds. Some were too high, and some were too low; and seeds in such lots were either all germinated or ungerminated at the chosen counting time. Consequently, rate distributions in the different experiments were sampled with different numbers of thresholds—from 5 to 9 instead of the designed 11.

To test the hypothesis that embryo developmental rates were normally distributed, we regressed proportions of germinated seeds (transformed to probits) against threshold rates (1/r days) by the method of maximum likelihood (7). Data from the four petri dishes sampled at each threshold were combined to provide adequate expected numbers at each threshold. We also tested the alternative hypothesis that days to germination (rather than rates) is normally distributed by regressing probits against days (retransformed from the rate thresholds).

Results

Length of stratification and temperature strongly influenced speed and variability of germination. With 16 days stratification and 15 C germination temperature, germination started at about 6 days and continued beyond 50 days. Peak germination occurred at about 9 days. With 64 days to stratification and 25 C temperature, germination occurred between 2 and 12 days with the peak at about 4 days.

Results suggest that it is the embryo developmental rates of individual seeds, rather than days to germination, that are distributed normally. Regardless of seed treatment, probits of germination percentage fitted the rate scale better than the day scale (Table 1). Mean squares for regression as a proportion of total mean squares (R²) were consistently higher for regression with rate than with days; and X² values for lack of fit were consistently lower with rate than with days. Evidence that germination rates might not always be distributed normally appeared only in experiments with high germination temperature and especially with long stratification (Table 1). There was, however, no consistency, no curvilinearity, in the deviations from regression to account for lack of fit (Figure 1). So deviations may reflect some irregularity in the test conditions (e.g., moisture in petri dishes) rather than skewness or kurtosis in the rate distribution.

If regression lines in Figure 1 represent the seed population, then frequency distribution of days to germination and of the underlying embryo developmental rates are as in Figure 2. With short stratifications and cool germination temperature, days to germination are strongly and positively skewed (Figure 2). Only with long stratification and with cool germination temperatures are days distributed almost normally. (See Table 1 also.)
Table 1. Mean squares and statistical tests of the fit of data from four experiments to the alternative models $y = a - bt$ versus $y = a - b(1/t)$, where $y =$ germination percentage transformed to probits, $t =$ days to germination and $1/t =$ germination rate.

<table>
<thead>
<tr>
<th>Stratification period</th>
<th>Germination temperature</th>
<th>16 days</th>
<th></th>
<th></th>
<th></th>
<th>64 days</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SOV</td>
<td>d.f.</td>
<td>days(t)</td>
<td>Rate(1/t)</td>
<td>d.f.</td>
<td>days(t)</td>
<td>Rate(1/t)</td>
<td>d.f.</td>
<td>days(t)</td>
<td>Rate(1/t)</td>
</tr>
<tr>
<td>Total</td>
<td>6</td>
<td>460.97</td>
<td>477.81</td>
<td>8</td>
<td>566.26</td>
<td>630.14</td>
<td>4</td>
<td>368.54</td>
<td>379.84</td>
</tr>
<tr>
<td>Regression</td>
<td>1</td>
<td>258.93**</td>
<td>465.73**</td>
<td>1</td>
<td>499.94**</td>
<td>621.99**</td>
<td>1</td>
<td>347.38**</td>
<td>379.42**</td>
</tr>
<tr>
<td>Lack of fit ($\chi^2$)</td>
<td>5</td>
<td>202.04**</td>
<td>12.08*</td>
<td>7</td>
<td>116.32**</td>
<td>8.14**</td>
<td>3</td>
<td>21.16**</td>
<td>0.43**</td>
</tr>
<tr>
<td>$R^2$ a/</td>
<td>0.56</td>
<td>0.97</td>
<td>0.80</td>
<td>0.99</td>
<td>0.94</td>
<td>1.00</td>
<td>0.82</td>
<td>0.95</td>
<td>0.95</td>
</tr>
</tbody>
</table>

*: ** Statistically significant at $p < 0.05$ and $p < 0.01$, respectively.

ns: Statistically nonsignificant.

a/ The proportion of total weighted sums of squares that is due to regression. Weighted sums of squares were calculated using means of the four samples at each rate-threshold.
Table 2. Relationships for calculating properties to describe germination behavior of seed lots. Relationships are based on equations of form $y = a - bx$ where $y =$ germination percentage transformed to probits, $a$ and $b$ are coefficients calculated from the data and $x =$ average rate of development of seed embryo to germination, or $x = 1/t$ where $t =$ days to germination.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean embryo development rate ($%$)</td>
<td>$(5 - a)b$</td>
</tr>
<tr>
<td>Standard deviation of embryo development rate ($\sigma$)</td>
<td>$-1/b$</td>
</tr>
<tr>
<td>Index: skewness of days to germination $^a$</td>
<td>$-1/(5 - a)$</td>
</tr>
<tr>
<td>Total germination percent</td>
<td>Transform $a$ from probit to percentage $^b$</td>
</tr>
<tr>
<td>Days to 10% germination</td>
<td>$b/(3.72 - a)$</td>
</tr>
<tr>
<td>Days to 90% germination</td>
<td>$b/(6.28 - a)$</td>
</tr>
<tr>
<td>Percent germination to a given day ($t$)</td>
<td>Solve $y = a - b(1/t)$, then transform $y$ from probit to percentage $^b$</td>
</tr>
</tbody>
</table>

$^a$/Not a standard statistical measure of skewness. It provides a "relative" measure of departure of days-to-germination from normality, i.e., the larger the coefficient of variation in embryo development rate and the greater the departure in days to germination, the larger the index.

$^b$/e.g., by a suitable table such as in Fisher and Yates' (1949, Table IX).
Table 3. Estimates of descriptive properties for seed lots germinated in four experiments with different germination temperatures and periods of stratification.

<table>
<thead>
<tr>
<th>Experimental conditions</th>
<th>Rate (mean and standard deviation)</th>
<th>Index skewness of days</th>
<th>Days to peak germination</th>
<th>Estimated total germination (percent)</th>
<th>Days to 10% germination</th>
<th>Days to 90% germination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stratification period (days)</td>
<td>Germination temperature (C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>15</td>
<td>0.078 ± 0.044</td>
<td>0.56</td>
<td>8.89</td>
<td>96.0</td>
<td>7.4</td>
</tr>
<tr>
<td>16</td>
<td>25</td>
<td>0.166 ± 0.073</td>
<td>0.44</td>
<td>4.64</td>
<td>98.0</td>
<td>3.9</td>
</tr>
<tr>
<td>64</td>
<td>15</td>
<td>0.135 ± 0.023</td>
<td>0.17</td>
<td>6.77</td>
<td>100.0a/</td>
<td>6.1</td>
</tr>
<tr>
<td>64</td>
<td>25</td>
<td>0.232 ± 0.064</td>
<td>0.28</td>
<td>3.80</td>
<td>100.0a/</td>
<td>3.2</td>
</tr>
</tbody>
</table>

\(^{a/}\) Estimated as > 99.99 percent.
Figure 1. Maximum likelihood fit of germination data from four experiments to equations of form \( y = a - bx \) where \( y \) = germination percentage transformed to probits and \( x \) = embryo developmental rate. Experiments varied in stratification period (SP = 16 and 64 days) and germination temperature (GT = 15 and 25°C).

Figure 2. Frequency distributions of days to germination (solid line) and of the underlying embryo developmental rate (broken line) for four experiments of contrasting stratification period (SP = 16 and 64 days) and germination temperature (GT = 15 and 25°C).
Once a regression equation has been calculated for a seed lot, mean germination rate and variability, total germination and other parameters are easily calculated by the relationships in Table 2. The frequency distribution curve of embryo developmental rate ($r$) such as in Figure 2 is based on calculation of ordinates for selected $r_i$ over the range of rates in the seed test. Ordinate values can be expressed as follows:

$$Y = \frac{1}{\sqrt{2\pi} \sigma} e^{-1/2 \left( \frac{r_i - \mu}{\sigma} \right)^2}$$

where:

- $\mu = \text{rate mean or } (5-a)/b$
- $\sigma = \text{rate standard deviation or } -1/b$
- $Y = \text{ordinate value at } r_i$

The frequency distribution curve of days to germination ($t$) is calculated similarly by determining ordinate values from:

$$Y = \frac{1}{t^2 \sqrt{2\pi} \sigma} e^{-1/2 \left( \frac{\ln t_i - \mu}{\sigma} \right)^2}$$

where:

- $t = \text{days or } 1/r$
- $Y = \text{ordinate value at } t_i$ and
- other symbols are as above

Days to peak germination ($t_{\text{max}}$) is calculated from:

$$t_{\text{max}} = \frac{\sqrt{\mu^2 + 8\sigma^2} - \mu}{4\sigma^2}$$

Computed values for our seed lots for several parameters useful in characterizing germination are given in Table 3.

**Discussion**

Means and variance of embryo developmental rates of individual seeds determine the shape of distribution of days to germination. Under normal conditions in the field, Douglas-fir seeds germinate in cool spring temperatures in 12- to 15-hour daylengths after a long period of natural stratification. We found an approximately normal distribution of days-to-germination only in the experiment which best approximated these conditions.

Embryo developmental rates are distributed normally over a wider range of treatments, but there probably are limits. We found evidence of non-normality in experiments with longest stratification and highest
temperature. These germination conditions are infrequently encountered during germination in nature or in forest nurseries. Stratification periods as long as 64 days are also uncommon in seed testing. A germination temperature of 22 C, however, is in the recommended range (2). Such deviations from normality in rates (and days to germination) that apparently occur in germination tests may partly account for the reported difficulties in correlating test results with nursery performance (16). We also have seen apparent deviations from normality in other experiments using unstratified seed, or seed stratified for very long periods, for example 124 days.

We tested the fit of developmental rates to the normal distribution by germinating seed in four petri dishes for each rate threshold. For hypothesis testing, we wished to eliminate the environmental correlations that might occur from sample to sample if germinants were counted at successive rate-thresholds within a single petri dish. Our procedure, however, was wasteful of seed and unnecessarily complex for the routine characterization of a seed lot. A simpler method is to periodically count the germinants within a test dish and, for each of the successive counting times \( t_1, t_2, \ldots, t_n \), accumulate germination as a percentage of total germination, transform the cumulative percentages to probits, and plot probits against the reciprocals of time \( 1/t_n \). We find little deviation from a straight line using 50 or more seeds per dish, providing pretreatments and germination temperatures are more or less within the natural range. For many practical purposes, mean embryo developmental rate, its standard deviation, and the regression equation can be scaled directly from a hand-fitted plot on probit-scale graph paper (4, p. 101 f). Germination need not be recorded at points evenly distributed on the rate scale. If statistical estimates are desired, procedures developed specifically for the cumulative curve are available (4).

With this system, all the usual properties for describing germination can be derived by estimating only two parameters, an intercept \( a \) and a regression coefficient \( b \). Furthermore, for most purposes the parameters can be estimated easily by graphical procedures. The reactions of Douglas-fir seeds to treatment appear to be similar to reactions of seeds in many other species. Therefore, the problems with normality which have been connected with fitting germination curves to days, whether by using probit transformation (12, 13), or logit transformation (15), are apparently avoided by using rates rather than days as the predicting variable.

**Literature Cited**
