Predicting Germination Response to Temperature. III. Model Validation Under Field-variable Temperature Conditions

STUART P. HARDEGREE*

USDA Agricultural Research Service, Northwest Watershed Research Center, 800 Park Blvd, Suite 105, Boise, ID 83712, USA

INTRODUCTION

McDonald (2002a), Hardegree (2006) and Hardegree and Winstral (2006) evaluated cardinal-temperature (CT) models, regression-based models and statistical gridding for predicting cumulative germination response to constant temperature, and concluded that empirical models with relatively few a priori shape assumptions provided the best model fit. Implementation of these models to estimate field-variable temperature response requires acceptance of several additional assumptions: subpopulations must retain their relative germination rate regardless of thermal conditions, and instantaneous germination rate must be unaffected by thermal pattern and thermal history (Garcia-Huidobro et al., 1982a, b; Benech Arnold et al., 1990; Probert, 1992). McDonald (2002b) noted that none of the models tested in his experiment adequately estimated germination response under conditions of diurnal temperature fluctuation. Hardegree and Van Vactor (1999) and Hardegree et al. (1999) obtained relatively accurate estimates of variable-temperature response but tested only a single model type that McDonald (2002a) suggested was computationally inefficient. The purpose of the present study was to evaluate eight thermal germination models, described previously by Hardegree (2006) and Hardegree and Winstral (2006), for predicting cumulative germination response under field-variable temperature conditions. Models were evaluated for predictive accuracy, computational efficiency and utility in generating metrics for seedlot comparison.

MATERIALS AND METHODS

Thickspike wheatgrass [Elymus lanceolatus (Scribn. and J.G. Smith) Gould], bluebunch wheatgrass [Pseudoroegneria spicata (Pursh) Löve], Sandberg bluegrass (Poa secunda Vasey) and bottlebrush squirreltail [Elymus elymoides (Raf.) Swezey] seedlots used in this study were the same as used by Hardegree (2006) and Hardegree and Winstral (2006) for constant-temperature experimentation and model development. The seedlots, treatments and procedures used in this experiment were as described previously (Hardegree, 2006) unless specifically mentioned otherwise.

Germination response to field-variable temperature regimes was evaluated in environmental chambers that were programmed to simulate thermal conditions in field-emergence plots at the Orchard Field Test Site in southern Ada County, Idaho (Hardegree and Burgess, 1995; Hardegree and Van Vactor, 2000) between 1 April and 17 June, 1996. Eighteen thermocouples, positioned at 1-cm depth, were monitored every 10 min (600 s), and hourly...
(3600 s) mean temperatures calculated for each sensor. Six thermocouples were positioned in each of three soil types (sandy-loam, loamy-sand and silt-loam) with each thermocouple determining the programmed temperature regime for a separate germination chamber. Five replicate germination vials containing seeds of a given species were loaded in each environmental chamber on the day simulating field conditions on 1 April, and every 10 d thereafter, for a total of six planting dates. Seeds were subsequently monitored for germination response every second day for 28 d. This procedure produced 108 field-variable temperature regimes (six planting dates × 18 environmental chambers) for each species. Temperature control failed in two chambers subsequent to run number four, reducing the total number of variable-temperature treatments to 104. Field-variable temperature experiments were conducted between runs two and three of the constant-temperature experiments described previously by Hardegree (2006) and Hardegree and Winstral (2006).

Germination counts were pooled by species within each environmental chamber and the within-box totals used for subsequent analysis. Cumulative germination was calculated for every species and temperature treatment for every day on which new germination was observed. Germination percentage values were scaled by dividing by the maximum germination percentage obtained in the optimal constant-temperature treatment for a given species (Covell et al., 1982; Hardegree, 2006). Seed populations were considered to be composed of subpopulations based on relative germination rate (Garcia-Huidobro et al., 1982a). Days required to achieve 5–95% germination were calculated for each species and treatment by linear interpolation (I-data) between measured germination percentiles from the cumulative germination curves (Covell et al., 1986; Hardegree, 2006). A second data set was derived directly from the measured data (M-data). On any day in which additional germination was observed, a single M-data point was recorded establishing germination time, germination rate, temperature and cumulative germination percentage (G).

Germination time (days) was estimated for every I-data point, and for every species and temperature treatment using the eight model formulations previously described by Hardegree (2006) and Hardegree and Winstral (2006). Models used to estimate germination time for these data were parameterized with I-data derived from the constant-temperature experiments (Hardegree, 2006; Hardegree and Winstral, 2006). Germination time was also estimated for every M-data point using the three-dimensional model formulations described by Hardegree and Winstral (2006). M-data estimates of germination time were obtained from models parameterized with both M- and IM-data from the constant-temperature experiment (Hardegree and Winstral, 2006).

All environmental chambers were monitored every 10 s and a mean temperature calculated for every 15-min (900 s) interval of every treatment. Per-day rate was estimated for every combination of seedlot, time interval, temperature treatment and model formulation using the techniques described by Hardegree (2006) and Hardegree and Winstral (2006). Per-day rate was divided by 96 to establish germination rate for a given 15-min interval. Germination for a given subpopulation was estimated to occur when the cumulative summation of rate estimates, subsequent to planting, reached a value of 1 (Hardegree and Van Vactor, 1999).

RESULTS
Variable-temperature treatments differed in both mean temperature and diurnal variability. Figure 1 shows the temperature pattern from two environmental chambers over the course of the variable-temperature experiment. The examples in this figure are representative of the minimum and maximum diurnal variability in temperature fluctuation among chambers. Planting dates 1–6 are indicated at the bottom of the figure. Figure 2 shows the mean, mean daily maximum and mean daily minimum temperatures across all treatments. These data were sorted in ascending order of mean temperature within each experimental run.

Figure 3 shows total germination percentage of all variable-temperature treatments superimposed upon the values derived previously for the constant-temperature experiment (Hardegree, 2006). Mean temperatures for the variable-temperature treatments were calculated only for the period up to the day on which maximum germination percentage was first achieved.

Figure 4 shows the mean residual error (±1 s.d.) in germination time predictions as a function of G and species for interpolated subpopulation data (I-data). Constrained (PCT, SCT) cardinal-temperature models are not represented in this figure but produced similar residual predictive errors to those shown for the unconstrained CT model (UCT). All model formulations showed relatively similar patterns of variability in residual predictive error as a function of G. Variability in residual predictive error increased for all models and species above a G of about 70%. All models except the iterative-probit-optimization (IPO) model tended systematically to overestimate germination time when parameterized with I-data. The magnitude of this error tended to be less than about 0-6 d except for some species at higher subpopulation values (Fig. 4). The IPO model systematically underestimated germination time for low subpopulation values and overestimated germination time for higher subpopulation values (Fig. 4). Linear regressions of IPO residual predictive error as a function of G, in the 0 and 70% subpopulation range, yielded significant slopes for all seedlots with $r^2$ values of between 0.80 and 0.90.

Figure 5 shows the residual predictive error in estimation of germination time of measured subpopulations (M-data) for models parameterized with IM-data. IM model parameterization was used to derive this figure because the 3-dimensional regression (3DR) model was previously shown to have extremely poor model fit when parameterized directly with M-data (Hardegree and Winstral, 2006). Residual predictive errors in SG and IPO model estimates with M-data were very similar to those shown in Fig. 5 for IM-data (data not shown). In the 0–70% subpopulation
range, mean and standard deviation of residual predictive errors in germination time for the 3DR model were 0.1 ± 0.4 d for *P. spicata*, 0.4 ± 0.6 d for *E. lanceolatus* and *E. elymoides*, and 0.6 ± 0.6 d for *Po. secunda*. The equivalent error rates for the SG model were 0.1 ± 0.3 d for *P. spicata*, 0.3 ± 0.6 d for *E. lanceolatus*, 0.2 ± 0.6 d for *E. elymoides* and 0.5 ± 0.5 d for *Po. secunda*.

Figure 6 shows base temperature ($T_b$) estimates as a function of $G$ and species for the unconstrained (UCT) and constrained (PCT, SCT) cardinal-temperature models described by Hardegree (2006), and for the IPO cardinal-temperature model described by Hardegree and Winstral (2006). The procedure used to optimize IPO model coefficients resulted in $T_b$ estimates that were lower than all UCT model estimates for three seedlots, and lower than all mean $T_b$ estimates derived for the PCT and SCT models (Fig. 6).

Figure 7 shows suboptimal thermal time ($\theta_1$) values derived for the UCT and IPO models as a function of seedlot and $G$. The relative differences among seedlots in this figure are very similar to those shown by Hardegree (2006) for PCT and SCT model formulations.

Figure 8 shows mean measured and mean predicted germination times as a function of subpopulation across all variable-temperature treatments. These data were plotted on a similar relative scale as the thermal time data in Fig. 7 to facilitate comparison of relative germination response. Differences in mean measured and mean predicted germination time among species (Fig. 8) were relatively smaller than those indicated by a comparison of thermal time coefficients (Fig. 7).

**DISCUSSION**

Two types of error were noted by Hardegree (2006) and Hardegree and Winstral (2006) when they compared predicted and measured germination response to constant temperature: random error, and systematic error associated with a priori assumptions of model shape. In these previous studies, an attempt was made to reduce random model errors by replicating germination vials within each chamber (6–8), replicating temperature treatments in
different environmental chambers (three) and using a relatively small temperature interval between treatments (1.75 °C). All CT-type models were shown to underestimate germination rate systematically, and overestimate germination time, at low constant temperature (Hardegree, 2006; Hardegree and Winstral, 2006). Two- and three-dimensional regression models also generated a consistent pattern of systematic residual model error across all species (Hardegree, 2006; Hardegree and Winstral, 2006). PWL and SG model formulations required the fewest shape assumptions and produced the lowest residual model error for both germination rate and germination time (Hardegree, 2006; Hardegree and Winstral, 2006).

Prediction of germination response to variable temperature requires adoption of additional assumptions associated with secondary thermal effects on subpopulation dormancy and viability (Garcia-Huidobro et al., 1982a; Hardegree et al., 1999). Temperature variation per se may induce dormancy release in some species (Thompson, 1974; Thompson and Grime, 1983; Murdoch et al., 1989; Benech Arnold et al., 1990). Seed exposure to supraoptimal temperatures may cause relatively irreversible degenerative processes such as membrane degradation and the denaturing of enzymes (Bewley and Black, 1994). Violation of thermal history assumptions may have contributed to predictive error in the variable-temperature study conducted by McDonald (2002b) for tropical and subtropical pasture legumes. The least accurate model predictions in the McDonald (2002b) study were for alternating temperature treatments that overlapped the supraoptimal temperature range. Alternating temperature regimes tend to maximize time spent at the upper temperature extreme where seed damage is most likely to occur (Hardegree et al., 1999). In the present experiment, and a previous variable-temperature experiment with the same seedlots (Hardegree and Van Vactor, 1999), model accuracy was relatively good (Figs 4 and 5). In these studies, maximum daily temperature was achieved for only a short period in
the warmest part of the day, and only reached supraoptimal temperature for later planting dates.

All of the I-data models in this study, except the IPO model, tended to overestimate germination time over the entire subpopulation range (Fig. 4). This error was generally less than about 0.6 d for subpopulations below about 70%. Model-specific systematic errors noted by Hardegree (2006) and Hardegree and Winstral (2006) were minimized in the current experiment as all treatments fell well within the suboptimal temperature range and generally stayed above 3°C (Fig. 3). One factor contributing to systematic overestimation of germination time may have been the 2-d measurement interval used in the variable-temperature experiment. The procedure used to generate I-data from the cumulative germination curves interpolated subpopulation values between days in which new germination was detected. A 2-d interval tended to increase the germination percentage noted on the first day with a detected count, and required interpolation of all lower subpopulations back to a value of zero measured 2 d previously. This extended time period may have decreased the interpolated germination time below what would have been estimated with a 1-d count interval. This error would have not been present in the M-data, which, indeed, showed a smaller relative predictive error for the 3DR and SG models below subpopulation values of about 70% (Fig. 5).
Hardegree and Winstral (2006) attributed systematic errors in the IPO model to the optimization procedure used to derive $T_b$ and $\theta_1$. They noted that the IPO model was more accurate at 3 °C than the two-dimensional CT models tested by Hardegree (2006) but at the expense of overestimating germination rate in the temperature range 5–10 °C. IPO model optimization also made it impossible to resolve single values for optimal germination temperature ($T_o$) as a function of $G$ (Hardegree and Winstral, 2006). Figure 6 shows that values for $T_b$ derived from the IPO optimization procedure were lower than mean values for $T_b$ derived by simple extrapolation of subpopulation data. The optimization procedure appears to have produced better overall model fit, but at the expense of converting cardinal-temperature values to empirical coefficients (Hardegree and Winstral, 2006). Figures 4 and 5 also show a consistent pattern of systematic error as a function of subpopulation for IPO model estimates at variable temperature.

All models showed relatively large and variable predictive errors for the highest seed subpopulations (Figs 4 and 5). This phenomenon was probably caused by a flattening out of the cumulative germination curve at high subpopulation values. In the highest subpopulation range, relatively small changes in $G$ result in relatively large changes in both germination time and germination rate, reducing the model sensitivity in this region. Similar lack of model sensitivity in the flat part of the cumulative germination curve may have contributed to increased predictive error seen for the very lowest subpopulations in the M-data (Fig. 5).

PWL and SG model formulations were deemed the most accurate of all two- and three-dimensional models tested by Hardegree (2006) and Hardegree and Winstral (2006). Predictive model accuracy was relatively high in the current study, however, for all models, except the IPO model, for subpopulations up to about 70% (Figs 4 and 5). Cardinal-temperature model formulations may be inappropriate for simulating germination response of colder thermal regimes earlier in the spring (Hardegree, 2006; Hardegree and Winstral, 2006). Additional testing of variable-temperature response at higher temperature may also be necessary to evaluate thermal history assumptions that were not adequately tested in the current study. Higher temperature treatments would not have much relevance for the species studied in this experiment, however, as supraoptimal thermal conditions on rangelands typically occupied by these species generally occur late enough in the spring that water is no longer available for germination (Hardegree et al., 2003).

There were relatively large differences in efficiency in both model development and model implementation for estimating variable-temperature germination response. McDonald (2002a) noted that superior model fit of the PWL model was offset by the need to generate as many as 11 regression equations for each subpopulation. With the smaller treatment interval used by Hardegree (2006), the number of PWL regression equations for each subpopulation increased to 20. Two-dimensional CT model formulations reduced the number of separate regressions to two per subpopulation. UCT model formulations, however, required tracking of individual values for $T_b$, PWL and SG model formulations were deemed the most accurate of all two- and three-dimensional models tested by Hardegree (2006) and Hardegree and Winstral (2006). Predictive model accuracy was relatively high in the current study, however, for all models, except the IPO model, for subpopulations up to about 70% (Figs 4 and 5). Cardinal-temperature model formulations may be inappropriate for simulating germination response of colder thermal regimes earlier in the spring (Hardegree, 2006; Hardegree and Winstral, 2006). Additional testing of variable-temperature response at higher temperature may also be necessary to evaluate thermal history assumptions that were not adequately tested in the current study. Higher temperature treatments would not have much relevance for the species studied in this experiment, however, as supraoptimal thermal conditions on rangelands typically occupied by these species generally occur late enough in the spring that water is no longer available for germination (Hardegree et al., 2003).

There were relatively large differences in efficiency in both model development and model implementation for estimating variable-temperature germination response. McDonald (2002a) noted that superior model fit of the PWL model was offset by the need to generate as many as 11 regression equations for each subpopulation. With the smaller treatment interval used by Hardegree (2006), the number of PWL regression equations for each subpopulation increased to 20. Two-dimensional CT model formulations reduced the number of separate regressions to two per subpopulation. UCT model formulations, however, required tracking of individual values for $T_b$, PWL and SG model formulations were deemed the most accurate of all two- and three-dimensional models tested by Hardegree (2006) and Hardegree and Winstral (2006). Predictive model accuracy was relatively high in the current study, however, for all models, except the IPO model, for subpopulations up to about 70% (Figs 4 and 5). Cardinal-temperature model formulations may be inappropriate for simulating germination response of colder thermal regimes earlier in the spring (Hardegree, 2006; Hardegree and Winstral, 2006). Additional testing of variable-temperature response at higher temperature may also be necessary to evaluate thermal history assumptions that were not adequately tested in the current study. Higher temperature treatments would not have much relevance for the species studied in this experiment, however, as supraoptimal thermal conditions on rangelands typically occupied by these species generally occur late enough in the spring that water is no longer available for germination (Hardegree et al., 2003).
\( T_m \) (maximum temperature), \( \theta_1 \) and \( \theta_2 \) (supraoptimal thermal time) for each subpopulation. PCT and SCT models could be implemented more efficiently than the UCT model as they used single values for \( T_m \) and \( T_h \) across subpopulations (Fig. 6) and yielded two equations for estimating \( \theta_1 \) and \( \theta_2 \) as a function of \( G \). The PCT model, however, required probit transformation of the data and additional log-transformations to optimize model performance (Hardegree, 2006). The SCT model was the most efficient of the CT models both to develop and to implement as it required fewer data transformations than the PCT model and yielded better model fit for the relationship between \( \theta \) and \( G \) (Hardegree, 2006). The most efficient two-dimensional model formulation proved to be the non-linear regression (NLR) model, which required calculation of a single equation for each subpopulation, accomplished automatically upon importation of the data using commercial computer software (Hardegree, 2006).

All two-dimensional model formulations required additional data manipulation to produce interpolated subpopulation values for model development. The use of interpolated subpopulation values may artificially enhance model fit as the data between adjacent measured values are not independent (Ellis et al., 1986; Hardegree and Winstral, 2006). Three-dimensional model formulations can be parameterized directly from measured subpopulation data and yield rate estimates for any subpopulation. The IPO model, however, was relatively less accurate than all other model formulations (Figs 4 and 5). 3DR model formulations were relatively accurate but only when parameterized with the combined IM-data set (Hardegree and Winstral, 2006). The PCT and SCT models were capable of predicting germination rate for measured subpopulation values, but required I-data parameterization and were relatively more difficult to implement than the SG model (Hardegree and Winstral, 2006). In order of relative efficiency, models parameterized from the I-data were ranked PWL << PCT < SCT < UCT < OPCT < NLR < 3DR < SG. The SG model was the only formulation that produced relatively accurate model estimates for variable-temperature response when directly parameterized from the M-data set.

One potential advantage of CT model formulations is that they yield coefficients that can be used directly for comparing seedlots and for screening germplasm (Covell et al., 1986; Ellis et al., 1986; Ellis and Butcher, 1988; Jordan and Haferkamp, 1989; Craufurd et al., 1996; Fidanza et al., 1996; Holshauser et al., 1996; Steinmaus et al., 2000; Phartyal et al., 2003). Figure 3 shows seedlot variability in \( T_m \) for different CT model formulations tested in the current experiment. Subpopulation variability in \( T_m \) may be higher than mean separation of \( T_h \) values between species, which limits the utility of this parameter for seedlot comparison. The utility for seedlot comparison, and physiological relevance of \( T_m \) is further degraded by the optimization procedure used by the IPO model (Fig 6).

Figure 7 shows the relative relationship between \( G \) and \( \theta_1 \) for UCT and IPO models tested in this experiment. The relative relationship of these parameters is very similar to measured and modelled values for \( \theta_1 \) derived for SCT and PCT model formulations by Hardegree (2006). These coefficients show relatively clear separation for \( P. \) spicata and \( P. \) secunda with intermediate thermal responses for \( E. \) lanceolatus and \( E. \) elymoides.

Hardegree et al. (2003) suggested an alternative method for comparing seedlots based on relative performance in potential field environments. Figure 8 shows the relationship between mean germination time and subpopulation, for all species for both measured and simulated data. In this comparison, it is more difficult to distinguish differences between seedlots other than for \( P. \) spicata. Hardegree et al. (2003) noted that the probability of significant time being spent at certain combinations of temperature and water availability may be relatively low. Seedlot comparisons based on the full range of potential germination response may not, therefore, be representative of performance under more realistic environmental scenarios. Ecological comparisons of this type require models that are both accurate and computationally efficient as the focus is on estimating germination response under a relatively large number of potential field conditions. An ecological index of this type may also require fewer data for model parameterization as microclimatic conditions outside the range expected in the field can be excluded from model development and analysis. Restricting the modelled temperature range in the constant-temperature experiment to only suboptimal temperatures would have significantly reduced model complexity for all models and improved model fit for NLR and 3DR model formulations. Alternatively, for the same effort, more treatment replication could have been applied in the suboptimal temperature range.

**CONCLUSIONS**

All models tested were relatively accurate in modelling variable-temperature germination response in the 0–70% subpopulation range except for the IPO model formulation, which had significant systematic errors as a function of \( G \). Model efficiency was greatest for the SG model, which did not require data transformation and could be parameterized directly from measured subpopulation data. Relative comparisons of seedlot performance between CT model indices and indices derived from field-variable temperature scenarios were slightly different. As models increase in complexity to account for other environmental factors, it will be harder to derive physiologically relevant model parameters. Predicted versus actual germination time, however, will continue to serve for all models as a criterion for judging predictive error.

**LITERATURE CITED**


Covell S, Ellis RH, Roberts EH, Summerfield RJ. 1986. The influence of temperature on seed germination rate in grain legumes. I. A


