One-step analysis of seed storage data and the longevity of Arabidopsis thaliana seeds

Fiona R. Hay1,4, Andrew Mead2, Kirstine Manger1 and Fiona J. Wilson3
1 Seed Conservation Department, Royal Botanic Gardens Kew, Wakehurst Place, Ardingly, West Sussex RH17 6TN, UK
2 Horticulture Research International, Wellesborne, Warwick CV35 9EF, UK
3 Nottingham Arabidopsis Stock Centre, Plant Science Division, School of Biosciences, University of Nottingham, Sutton Bonnington Campus, Loughborough LE12 5RD, UK

Received 29 August 2002; Accepted 20 November 2002

Abstract
Seeds of two ecotypes of Arabidopsis thaliana, NW20 and N1601, were aged over a range of saturated salt solutions at temperatures between 6 °C and 55 °C. For each ecotype, the results from 37 storage experiments were summarized using the Ellis and Roberts viability equations and a modified version of these equations which allows for a proportion of ‘non-respondents’. For both models, two approaches were taken in order to model the effect of moisture content (MC) and temperature on seed longevity. The first, a two-step approach, involved fitting individual survival curves and then multiple regression analysis of the fitted parameters on moisture content and temperature. For the second approach, the full viability models were fitted in one step, including the multiple regression for the effects of MC and temperature within the generalized linear model used to describe each survival curve. This one-step approach takes into account the full variability of the data and provides the best predictions of seed longevity based on the original assumptions of the Ellis and Roberts viability equations. As a consequence of taking into account all the variation, this one-step approach is more sensitive and thus more likely to detect changes due to reducing the number of parameters in the model as being significant. Whilst both approaches indicated that seeds from the two Arabidopsis ecotypes have the same response to MC and temperature, parameter values did differ between the approaches, with the one-step approach providing the better fit. The best model for these two ecotypes, from the one-step approach, confirmed a quadratic relationship between temperature and longevity, but the magnitude of the non-linearity is not as large as indicated by the universal value for the quadratic term.

Key words: Arabidopsis thaliana, seed longevity, viability equations.

Introduction
Within broad limits, reducing moisture content (MC) and temperature during storage increases the longevity of desiccation-tolerant (orthodox) seeds. The effects of moisture content and temperature on seed longevity are usually described using the viability equations of Ellis and Roberts (1980). These equations are commonly fitted sequentially. The first equation is based on an assumption that the time to death for individuals from a population of seeds stored at a constant MC and temperature follows a normal distribution. So, under constant storage conditions, probit viability, \( v \), is related to storage period, \( p \) (in days) as:

\[
   v = K_i - p/\sigma
\]

where \( K_i \) is the initial probit viability, and \( \sigma \) is the standard deviation of the normal distribution of seed deaths in time, equivalent to the time in days for probit viability to decrease by one. Thus, for any given storage experiment (constant storage environment) the parameters \( K_i \) and \( \sigma \) can be estimated by fitting this linear relationship within a probit analysis.
Whilst the initial viability, \( K_e \), is assumed to depend only on genotype and seed quality, the second Ellis and Roberts equation describes the effects of MC and temperature on the ‘rate’ at which seeds lose viability during storage, as given by the parameter \( \sigma \), as:

\[
\log_{10}(\sigma) = K_e - C_w \log_{10}(m) - C_h t - C_q t^2 \tag{2}
\]

where \( m \) and \( t \) are the storage MC and temperature, respectively, and \( K_e, C_w, C_h, \) and \( C_q \) are parameters to be estimated. The values of these parameters can be estimated from data from a series of storage experiments carried out at different temperatures and a range of MCs, either by fitting a logarithmic relationship between \( \log_{10}(\sigma) \) and MC for each (constant) temperature:

\[
\log_{10}(\sigma) = K - C_w \log_{10}(m) \tag{3}
\]

and a quadratic, relationship between \( \log_{10}(\sigma) \) and temperature for each (constant) moisture content:

\[
\log_{10}(\sigma) = C - C_h t - C_q t^2 \tag{4}
\]

or by the multiple linear regression of \( \log_{10}(\sigma) \) on \( \log_{10}(MC) \) and linear and quadratic components of temperature, as given in equation (2).

It was initially suggested that the values of the parameters in equation (2) would not differ between different seed-lots of a particular species (Ellis and Roberts, 1980). Once determined, it would therefore be possible to predict the longevity of any seed-lot of that species under any storage environment (within limits) once the initial viability of the seed-lot, \( K_e \), had been estimated. Indeed it was later proposed that the values of the temperature parameters, \( C_h \) and \( C_q \), were constant across all species (Dickie et al., 1990). However, more recent studies have shown that values of all four parameters may differ between seed-lots of the same species, depending, for example, on the time of harvest (Wechsberg, 1994; Hay et al., 1997) or on genotype (Zanakis et al., 1994; Tang et al., 1999). The parameter which appeared most likely to vary was \( K_e \), since the linear regressions of \( \log_{10}(\sigma) \) against \( \log_{10}(MC) \) for seeds stored at different temperatures could be constrained to have a common slope, i.e. a common estimate of \( C_w \) (Ellis et al., 1989; Hay et al., 1997).

Differences in the parameter values within a species cast some doubt on the reliability of predictions of longevity based on the viability equations. Tang et al. (2000) derived an alternative model for predicting seed longevity, which does not require estimation of \( K_e \). This alternative model predicts seed longevity for a given seed-lot as a function of the change in storage conditions (MC and temperature) compared with a reference set of conditions under which a storage experiment has been carried out. The ‘intercept’ parameter (\( K_e \)) is thus eliminated from this model, whilst the other moisture and temperature parameters are assumed to be constant within species.

There has also been some examination of the assumption that seed deaths follow a normal distribution over time (Moore and Roos, 1982; Wilson et al., 1989). Asymmetry in the survival curves, which may indicate a departure from the normal distribution, can also occur when there is an initial proportion of the population which are ‘non-respondents’, that is seeds which are empty, insect-infested or dormant (assuming that the proportion of dormant seeds remains constant during storage). These non-respondents are not the same as seeds that have already aged and died. Finney (1977) described the application of Abbott’s formula (Abbott, 1925) within probit analysis to correct for the proportion of insects in an insecticide bioassay which will die whatever the insecticide dose (the ‘control mortality’). Mead and Gray (1999) incorporated this same correction as an additional ‘control viability’ parameter in a modification of the viability equations of Ellis and Roberts (1980), to allow for an initial proportion of ‘non-respondents’.

The equivalent approach to that adopted by Finney is to work with the percentage of non-viable seeds rather than the percentage of viable seeds. This percentage of non-viable seeds, \( \%nv \), can be considered to consist of a proportion of the population, \( C_{nv} \), of non-respondents, and the remaining responding proportion \( (1-C_{nv}) \), which lose viability during storage, following the probit model proposed by Ellis and Roberts (1980):

\[
\%nv = 100 \times \left[ C_{nv} + (1-C_{nv}) \times \Phi\left(-\frac{K_i - p}{\sigma}\right) \right] \tag{5}
\]

where \( \Phi \) is the cumulative normal function, \( K_i \) is the initial viability of the responding part of the population in probits, and \( p \) and \( \sigma \) are storage time and the standard deviation of the normal distribution of deaths in time for the responding proportion of seeds, respectively, as in equation (1).

Rewriting equation (5) in terms of the percentage of viable seeds \( \%v \), and letting \( C_v (=1-C_{nv}) \) be the proportion of responding seeds (Mead and Gray, 1999), gives

\[
\%v = 100 \times C_v \times \Phi(K - p/\sigma) \tag{6}
\]

and including the effects of MC and temperature gives

\[
\%v = 100 \times C_v \times \Phi\left(K_i - \frac{p}{10K_e - C_w \log_{10} m - C_h t - C_q t^2}\right) \tag{7}
\]

Where the proportion of non-respondents, \( C_{nv} \), is equal to zero (i.e. \( C_v=1 \)), this equation is equivalent to the original Ellis and Roberts viability equations, though with
the cumulative normal function applied to the right-hand side of the equation rather than transforming the left-hand side to probits. Assuming the parameters $K_C$, $C_W$, $C_H$, and $C_Q$ are constant within species, it would therefore only be necessary to determine $K_i$ and $C_V$ for any seed-lot in order to use this equation for predicting the longevity of that seed-lot at a range of storage MCs and temperatures.

The effects of MC and temperature as described by equation (2) and incorporated into equation (7) appear to hold over a wide range of temperatures. At worst, the universal values for the temperature parameters $C_H$ and $C_Q$ are predicted to hold between $-13\, ^\circ C$ and $+90\, ^\circ C$ (Dickie et al., 1990). However, there are limits to the logarithmic relationship between $\sigma$ and MC. Equation (3) appears to apply only over the linear part of the moisture ad- or desorption isotherms, although there is some debate both as to where the exact cut-off point occurs at low MC and whether or not this point varies with temperature (Ellis et al., 1995; Vertucci et al., 1994; Buitink et al., 1998; Walters, 1998; Ellis, 1998). Moisture isotherms describe the relationship between seed MC and equilibrium relative humidity at a constant temperature during either ad- or desorption and are typically (for orthodox seeds) sigmoidal in shape. The exact shape of an isotherm depends on the species being studied, varying according to the composition of their seeds, and the temperature at which it is determined (Vertucci and Roos, 1993). If determined at the temperature at which the seeds are being stored, they may be used to indicate where the break in the linear logarithmic relationship between MC and $\sigma$ might occur (Buitink et al., 1998).

Despite the fact that Arabidopsis thaliana is the species most commonly found in plant science laboratories and that seeds are often stored for a number of years in the laboratory, the viability constants have never been determined for this species. This study was therefore initiated with the aim of determining the values of $K_E$, $C_W$, $C_H$, and $C_Q$ and thus compare the longevity of seeds from two different ecotypes of A. thaliana. These two ecotypes, Landsberg erecta (NW20) and Wassilewskija (N1601) are currently stored at the Nottingham Arabidopsis Stock Centre and are two of the background lines from which many other lines have been produced. They were selected for further storage experiments since, from an initial experiment, the longevity of their seeds appeared to differ.

As already described, the viability equations are usually fitted in two steps, using probit analysis to fit the individual survival curves and then multiple linear regression analysis to model the effects of MC and temperature on the rate of loss of viability, as represented by $\sigma$. Differences between seed-lots might then be assessed by looking for differences in the slope and/or intercept of individual survival curves (determined under the same storage conditions) and/or for differences in the relationships between $\sigma$ and MC or temperature.

In this paper, the results of fitting both the original and modified viability equations to data from individual storage experiments for the two A. thaliana ecotypes are shown. The combined models across storage conditions are then fitted using two approaches, the usual two-step process, and a one-step approach, using the general non-linear model fitting facilities in GenStat. Differences in longevity between the two ecotypes were assessed through joint analyses using both approaches, and the advantages and disadvantages of the two approaches are discussed.

### Materials and methods

#### Seed collection

Seeds from the two ecotypes of Arabidopsis thaliana (hereafter referred to as Arabidopsis), NW20 and N1601, were harvested on 16 November 1997 from approximately 150 plants grown in the NASC glasshouses. Watering was withheld after flowering and the plants left to dry in situ. The seeds and any abscised siliques were collected from plastic collars around the plant. The seeds, with initial MCs of 7.68% (NW20) and 7.55% (N1601), were sent to Wakehurst Place in waxed paper bags which were enclosed in 100 ml aluminium screw cap containers. Upon arrival, the containers were opened and placed in a dry-room maintained at 15–20% relative humidity and 15±1°C. Once the seeds had equilibrated (MC approximately 4.5%), the lids were replaced on the containers and opened periodically within the dry room for removal of seeds as required for experimentation.

#### Seed storage experiments

Seed storage experiments were set up between January 1999 and May 2000. Due to logistical constraints, it was not possible to set all the experiments up within a shorter period of time.
were placed over the relevant saturated salt solution in 60 mm (height)
raised `basket'. The baskets (diameter), each containing approximately 150 seeds were arranged,
were air-tight. They were then placed in dark incubators maintained
sealed and Nesco®lm wrapped around the seal to ensure the boxes
°
solutions giving a range of MCs between 2% and 14% (Tables 2, 3).
996
Experiment Salt $t$ (°C)$^a$ MC$^b$ Data$^c$ Model 1 Model 2 Model 3$^d$ Model 4

1 LiCl 55 2.60 21 0.91 76.37 As for model 1 As for model 1 As for model 1 75.15 72.89
2 CaCl$_2$ 55 3.04 23 1.01 61.27 As for model 1 As for model 1 As for model 1 63.48 61.65
3 MgCl$_2$ 55 4.40 24 0.98 24.07 As for model 1 As for model 1 As for model 1 24.58 23.86
4 K$_2$CO$_3$ 55 5.71 24 0.99 12.55 As for model 1 As for model 1 As for model 1 12.86 12.5
5 NaBr 55 6.17 24 0.92 6.47 As for model 1 As for model 1 As for model 1 6.38 6.18
6 NaN$_3$ 55 8.40 23 0.69 1.16 As for model 1 As for model 1 As for model 1 1.01 0.98
7 (NH$_4$)$_2$SO$_4$ 55 11.82 20 0.98 0.30 As for model 1 As for model 1 As for model 1 0.31 0.3
8 KCl 55 11.97 24 0.76 0.45 As for model 1 As for model 1 As for model 1 0.41 0.39
9 NaOH 45 2.31 17 0.98 212.13 As for model 1 As for model 1 As for model 1 219.35 211.24
10 LiCl 45 2.96 24 0.90 277.47 0.9825 0.94 271.67 269.81 260.89
11 CaCl$_2$ 45 3.41 22 1.11 186.60 0.9860 1.15 183.15 205.42 198.77
12 MgCl$_2$ 45 5.06 24 1.10 101.25 0.8659 1.78 78.52$^{**}$ 119.92 108.31
13 K$_2$CO$_3$ 45 6.04 22 0.99 56.73 0.9166 1.27 50.76 58.92 57.07
14 NaBr 45 6.87 21 1.01 29.81 0.9447 1.13 27.43 31.41 30.31
15 NaN$_3$ 45 7.48 37 0.85 8.94 As for model 1 As for model 1 As for model 1 8.49 8.24
16 KCl 45 8.59 20 0.85 3.34 As for model 1 As for model 1 As for model 1 3.18 3.09
17 KCl 45 11.79 22 0.87 127.86 1.03 263.30 282.57 271.89
18 (NH$_4$)$_2$SO$_4$ 45 11.95 23 0.99 1.19 As for model 1 As for model 1 As for model 1 1.22 1.19
19 MgCl$_2$ 36 5.25 24 1.12 312.50 0.8394 2.20 203.21$^{**}$ 350.36 338.11
20 K$_2$CO$_3$ 36 5.52 22 0.95 232.40 0.7819 2.49 125.87$^{**}$ 233.75 225.63
21 NaBr 36 6.65 22 1.02 91.68 0.9711 1.60 69.54 98.03 94.33
22 NaNO$_2$ 36 7.97 46 0.81 42.89 0.8356 1.32 34.23 38.79 37.46
23 NaNO$_3$ 36 8.47 18 0.88 12.48 As for model 1 As for model 1 As for model 1 11.94 11.54
24 (NH$_4$)$_2$SO$_4$ 36 10.73 24 0.70 8.00 0.8672 1.05 7.01 6.99 6.79
25 KCl 36 11.59 24 0.93 3.88 As for model 1 As for model 1 As for model 1 3.67 3.75
26 NaBr 26 7.25 21 0.99 239.92 0.8644 1.59 185.43 247.95 239.46
27 NaNO$_2$ 26 9.19 23 0.97 39.57 As for model 1 As for model 1 As for model 1 40.26 39.07
28 (NH$_4$)$_2$SO$_4$ 26 11.03 23 0.95 24.57 0.9202 1.20 22.02 24.46 23.71
29 KCl 26 12.48 23 0.88 13.66 0.9904 0.90 13.54 13.18 12.78
30 NaNO$_2$ 16 8.65 24 0.84 127.86 As for model 1 As for model 1 As for model 1 119.27 115.27
31 NaOH 16 9.65 23 1.00 112.88 As for model 1 As for model 1 As for model 1 118.33 114.03
32 (NH$_4$)$_2$SO$_4$ 16 10.95 22 0.99 77.52 0.9183 1.28 68.59 80.06 77.38
33 KCl 16 12.50 24 1.03 48.14 0.8611 1.56 39.28 50.64 49.04
34 NaNO$_2$ 6 8.80 20 0.94 261.85 0.8803 1.52 200.24 262.54 252.78
35 NaNO$_2$ 6 11.12 19 1.03 263.30 As for model 1 As for model 1 As for model 1 282.57 271.89
36 (NH$_4$)$_2$SO$_4$ 6 11.30 21 0.91 284.34 0.8387 1.63 197.36$^{**}$ 276.17 264.6
37 KCl 6 13.50 23 0.88 193.69 0.8422 1.50 146.69 185.25 178.89

Model r.m.d. (d.f.)$^c$
1.68 (788) 1.68 (751) 1.75 (823) 1.74 (822)

$^a$ Temperature; MC = moisture content (% f.wt.); data refers to the number of data points at the storage temperature and moisture content. $K_i$ (in NEDs) and $\sigma$ (in d) given in bold for models 1 or 2 indicate which of these models had the lower residual mean deviance for that particular experiment; * or ** indicate that the change in deviance was significant ($P < 0.05$; ** $P < 0.001$). Model r.m.d. (d.f.) refers to the overall residual mean deviance (pooled r.m.d. in the case of models 1 and 2) and degrees of freedom for the model. Model 3: $K_i$ estimated as 0.94. Model 4: $K_i$ estimated as 1.02; $C_V$ estimated as 0.9722.

Seed ageing was carried out at temperatures of 6, 16, 26, 36, 45, and 55 °C over a number of different saturated salt (AnalaR grade) solutions giving a range of MCs between 2% and 14% (Tables 2, 3). For each storage experiment (MC/temperature combination) and for each ecotype, 22 open glass bottles (32 mm height×11 mm diameter), each containing approximately 150 seeds were arranged, on their side, on a nylon-coated wire raised ‘basket’. The baskets were placed over the relevant saturated salt solution in 60×110×170 mm (height×width×depth) clear plastic boxes. The boxes were sealed and Nescofilm wrapped around the seal to ensure the boxes were air-tight. They were then placed in dark incubators maintained at the prescribed ageing temperatures. Not all MC/temperature combinations were included since the predicted longevity of the seeds would have been too short (h) for accurate sampling or too long (years).

Sampling was carried out at regular intervals of time depending on the ageing environment. At each sampling time, one or two bottles for each ecotype were removed and the lids placed on the bottles. The fresh weight of approximately 50 seeds was measured using preweighed nickle-coated, stainless steel crucibles on a Sartorius Ultramicro balance. The seeds were dried in a fan-ventilated oven at 103±2 °C for 17±1 h then allowed to cool over silica gel before determining the seed dry weight. MCs are expressed on a percentage fresh weight basis.
The remaining seeds from each sample bottle (100 seeds maximum) were sown on 1% distilled water agar in 90 mm Petri dishes. Each dish was placed in an incubator maintained at an alternating temperature of 26/16 °C (±2 °C) with a 12 h photoperiod, light being provided during the warm phase. Plates were regularly checked for germination over a 3-week period. Seeds were considered to have germinated when the emerged radicle was at least 5 mm long and a healthy epicotyl had been produced.

**Moisture sorption isotherms**

Moisture isotherms were determined at the storage temperatures by measuring the relative humidity (RH) of a sample of *Ranunculus sceleratus* seeds from a glass bottle (48 mm high×22 mm diameter) placed alongside the *Arabidopsis* bottles in the storage experiments. *R. sceleratus* seeds were used for RH determination due to the very small size of *Arabidopsis* seeds. At various intervals during storage the bottles containing the *R. sceleratus* seeds were removed from the saturated salt environment and the seeds poured into the measuring chamber of a water-activity measuring station [water activity measuring station WA-40 with Hygrolyt sensor used in conjunction with a BT-RS1 display unit (Rotronic Instruments UK Ltd., Crawley)]. This apparatus had been calibrated for each temperature and was placed inside the relevant incubator to equilibrate prior to measuring the RH. The seeds were allowed to equilibrate for at least 30 min after which the RH of the seed was recorded.
In order to determine a more complete isotherm for the two Arabidopsis ecotypes than that which would be provided from the storage conditions alone, further saturated salt solutions were prepared. For each ecotype, three bottles containing approximately 50 seeds, were arranged as before, inside these environments together with a bottle of $R.~scleratus$ seeds for RH measurement.

Fig. 1. Summary of the models used in the first (A) and second (B) stages of the two-step analysis and in the one-step analysis (C).
After equilibrating for up to 1 month, the Arabidopsis seeds were removed for moisture content determination and the R. sceleratus seeds removed for RH measurement.

Statistical analysis

Analysis of the isotherm data and the data from the storage experiments was carried out using the regression and generalized linear and non-linear modelling facilities (Lane and Payne, 2000) in GenStat for Windows, Version 5 (VSN International Ltd., Oxford, UK).

The FITNONLINEAR directive was used to fit the D’Arcy-Watt equation to the isotherm data. This equation is made up of three parts (Vertucci and Leopold, 1987; Sun et al., 1997):

\[ WC = \frac{K' \times K \times RVP}{\Gamma + (K \times RVP)} + \frac{(c \times RVP)}{1 - (k \times RVP)} \]  

where WC is water content [g H\textsubscript{2}O g\textsuperscript{-1} dry weight; WC=MC+(100–MC) where MC is expressed as % fresh weight], K’ and K are parameters that relate to strong water-binding sites, c relates to weakly-bound water, k’ and k relate to bulk water; RVP is relative vapour pressure (RVP=RH/100). A single model was fitted to the data for all temperatures of both ecotypes, with parameters then allowed to vary with either ecotype or temperature.

The data for each of the 37 individual storage experiments for each ecotype were initially analysed by separately fitting both the original Ellis and Roberts (1980) viability equation (model 1) [equation (1)] and the Mead and Gray modified viability equation (model 2) [equation (6)] (Fig. 1A), using the PROBITANALYSIS procedure (Payne, 2000). The improvements in fit provided by model 2 compared with model 1 were assessed using an approximate F-test of the change in residual deviance. In fitting this equation to the combined data set the fitting process would not converge whilst estimating parameter c and so this parameter was set to zero. A stepwise analysis allowing the remaining parameters to vary with ecotype gave no evidence of the need for separate curves for the two ecotypes (P >0.05). Although there was a significant reduction in the residual variance when allowing all four parameters to vary with storage temperature (P <0.001) (Table 1), the differences in the fitted curves were not large, and so only a common curve is presented (Fig. 2).

Results

Sorption isotherms

The moisture sorption isotherms follow the sigmoidal shape that is typical of orthodox seeds and which can be interpreted in relation to the types and amounts of water binding by fitting the D’Arcy-Watt equation (Fig. 2). In fitting this equation to the combined data set the fitting process would not converge whilst estimating parameter c and so this parameter was set to zero. A stepwise analysis allowing the remaining parameters to vary with ecotype gave no evidence of the need for separate curves for the two ecotypes (P >0.05). Although there was a significant reduction in the residual variance when allowing all four parameters to vary with storage temperature (P <0.001) (Table 1), the differences in the fitted curves were not large, and so only a common curve is presented (Fig. 2).

Individual survival curves

A total of 37 storage experiments were analysed for each ecotype, including at least four experiments for each storage temperature and at least two experiments for each salt solution (Tables 2, 3).

When fitting model 2 (the non-constrained individual survival curves) for ecotype NW20, the control viability parameter C\textsubscript{V} was estimated as 1.00 for a number of experiments (Table 2) and thus there were few differences between the fitted lines for models 1 and 2 for this ecotype (Fig. 3). Inclusion of the control viability parameter only improved the fit for 14 out of the 37 storage experiments (estimates shown in bold in the model 2 column of Table 2) and the reduction in residual deviance was significant (P <0.05) for only three experiments. When the individual survival curves were fitted with constraints on K\textsubscript{i} and C\textsubscript{V} (model 4), the estimated proportion of ‘non-respondents’ was low with C\textsubscript{V} estimated at 0.9722, i.e. the initial seed-lot contained less than 3% of non-respondents (Table 2). The reduction in residual deviance for model 4 compared with the constrained model without C\textsubscript{V} (model 3) was not significant (P >0.05). Again, generally there were few differences between the fitted lines for these models (Fig. 3).
For ecotype N1601, inclusion of the control viability parameter (model 2) resulted in a reduction in the residual mean deviance compared with model 1 for 25 of the 37 storage experiments (estimates shown in bold in model 2 column of Table 3). For example, including this parameter resulted in an improved fit for seeds aged at 6 °C and 10.87% MC, seen as an apparent asymmetry in the survival data (Fig. 4). The reduction in residual deviance between models 1 and 2 was significant ($P < 0.05$) for 16 experiments, including most of those at the lower storage temperatures (16 °C and 6 °C). The value of the control viability parameter ranged from 0.6911 to 0.8891 for those curves where this parameter improved the fit. Where the fit was not improved by inclusion of this parameter, the estimated value was greater than 0.89 and in a few cases estimated as 1.00. When $K_i$ and $C_V$ were constrained to common values across all experiments, $C_V$ was estimated as 0.8424 (model 4; Table 3). However, the reduction in residual mean deviance between this model and the constrained model without $C_V$ (model 3) was not significant ($P > 0.05$). Despite this, since $C_V$ was often estimated as considerably less than 1.00 in the non-constrained survival curve fitting for ecotype N1601 (model 2; Table 3), it was decided that the control viability parameter would be included in further model fitting for this ecotype, but not for ecotype NW20.

**Two-step fitting of the viability model**

The moisture sorption isotherm indicated that phase I water binding sites (Sun et al., 1997) were filled by c. 3–4% MC, although there is some scatter in the data in this range. There also appeared to be a break in the logarithmic relationship between $S$ and MC in this MC region (Figs 5, 6, 45 °C and 55 °C only). Thus, the data for the two (at 55 °C) or three (at 45 °C) lowest MCs were excluded from the subsequent regression analyses.

The fitted baseline regressions of $\log_{10}(S)$ on $\log_{10}(MC)$ for each storage temperature indicated the strong influence of estimates from experiments corresponding to ageing over NaNO$_2$, particularly at 26 °C, 16 °C, and 6 °C (Figs 5, 6). It has been suggested that NaNO$_2$ has a detrimental effect on seeds *per se*, in addition to the ageing effect of the RH which it maintains (Hong et al., 1999). Thus the regression analyses were repeated with the data for all the

---

**Fig. 2.** Moisture sorption isotherms for *Arabidopsis thaliana* ecotypes NW20 (hollow symbols, dotted square) and N1601 (solid symbols, cross) determined at 55 °C (squares), 45 °C (circles), 36 °C (triangles), 26 °C (inverted triangles), 16 °C (diamonds), and 6 °C (dotted squares, cross). The inset shows the data for the whole isotherm; high moisture levels are omitted from the main graph. The D’Arcy–Watt equation was fitted to all the data shown in the main graph (Table 1).
Fig. 3. Sample survival curves for seeds of *Arabidopsis thaliana* ecotype NW20 stored at the temperatures and MCs shown. Square symbols represent the observed data. Lines shown are the results of individual curve fitting: without $C_V$ (solid) (model 1, Table 2); with $C_V$ (dashed) (model 2, Table 2); without $C_V$, constraining $K_i$ to a common value across all storage experiments (dash-dot) (model 3, Table 2); constraining $K_i$ and $C_V$ to common values across all storage experiments (dotted) (model 4, Table 2).
Fig. 4. Sample survival curves for seeds of Arabidopsis thaliana ecotype N1601 stored at the temperatures and MCs shown. Square symbols represent the observed data. Lines shown are the results of individual curve fitting: without $C_V$ (solid) (model 1, Table 3); with $C_V$ (dashed) (model 2, Table 3); without $C_V$, constraining $K_i$ to a common value across all storage experiments (dash-dot) (model 3, Table 3); constraining $K_i$ and $C_V$ to common values across all storage experiments (dotted) (model 4, Table 3); the solid line is the result of individual curve fitting without $C_V$ and with constraining of $K_i$ (model 3, Table 2); the dotted line is the result of two-step fitting (Table 4); the dashed line is the result of one-step fitting (Table 5).
NaNO₂ experiments excluded. As a result of these omissions, the full Ellis and Roberts model [equation (2)] could be fitted without resulting in a significant increase in residual variance compared with the baseline model for either ecotype. In re-fitting the Ellis and Roberts model [equation (2)] using the universal temperature coefficients, there was a significant (P < 0.05) increase in variance for both ecotypes (Table 4, comparisons A and B).

Analysing the data for the two ecotypes concurrently, the model with common estimates of \( K_E \), \( C_W \), \( C_H \), and \( C_Q \) did not significantly increase the residual variance compared with the model with separate estimates for each ecotype, (P > 0.05; Table 4, comparison D). However, models allowing any one of the four parameters to vary between ecotypes did lead to a significant reduction in the residual variance compared with the common estimates model (P < 0.05, data not shown). A reduced model excluding \( C_Q \) (i.e. with just a linear effect of temperature) did not result in a significant increase in residual variance compared with the common estimates model (P > 0.05; Table 4, comparison E). However, there was a significant increase in residual variance when the universal values for \( C_H \) and \( C_Q \) were used in the model (P<0.01; Table 4, comparison F).

**One-step fitting of the viability model**

The one-step analysis was applied to the data for the two ecotypes both independently and concurrently. As with the two-step modelling, the data for storage experiments below 4% MC and for experiments corresponding to ageing over NaNO₂ were excluded from the analyses.

In the independent analyses, fitting the modified, control viability model resulted in estimates of the \( C_V \) parameter tending to 1.00 for both ecotypes (results not shown), therefore giving the same estimates for the other parameters as the analysis where \( C_V \) was not included in the model (Table 5). This was probably caused by the wide variation in initial viability [seen in the variation in individual estimates of \( C_V \) (Tables 2, 3)] and subsequent one-step analyses concentrated on the original Ellis and Roberts model. Fitted models with \( K_i \) (and \( C_V \)) set at the values estimated in the two-step approach (models 3 and 4) resulted in significant increases in residual deviance for both ecotypes in comparison with the best-fitting independent models (P <0.001; Table 5, comparisons A and C). This emphasizes the importance of using the one-step approach in obtaining the best-fitting model for the observed data. Fitting models for each ecotype with the temperature parameters set to the universal coefficients similarly resulted in significant increases in residual
deviance compared with the best-fitting models ($P < 0.001$; Table 5, comparisons B and D).

Fitting to data from both ecotypes concurrently, a best-fitting model was obtained with separate parameter estimates for each ecotype that were almost identical to those obtained from the independent analyses (Table 5). Models with common estimates of either $C_W$, $C_H$ or $C_O$ resulted in non-significant increases in residual deviance compared with this baseline model ($P > 0.05$, not all shown), with the model with a common estimate of $C_O$ providing the best fit (Table 5, comparison E). Similarly, models with common estimates of $C_W$ and either of $C_H$ or $C_O$ resulted in further non-significant increases in residual deviance ($P > 0.05$) with the model with common estimates of $C_W$ and $C_O$ providing the best fit (Table 5, comparison F). Constraining either $K_E$ or $K_H$ to be common across ecotypes did result in a significant increase in residual deviance ($P < 0.01$) (Table 5, comparison G), but constraining both did not result in any further increase ($P > 0.05$; Table 5, comparison H). Further constraining $K_i$ to be common across ecotypes resulted in a highly significant increase in the residual deviance ($P < 0.001$; Table 5, comparison J). Thus the model with separate estimates of $K_i$ but common estimates of the other four parameters was selected as providing an adequate description of the observed data.

Comparisons of this best-fitting model with alternative models such as including the universal temperature coefficients (Table 5, comparison K), excluding the quadratic temperature coefficient (Table 5, comparison L) or with parameter values fixed at those obtained from the two-step analysis (Table 5, comparison M) all resulted in highly significant increases in the residual deviance. Similarly, models with $K_i$ and $C_V$ fixed at the estimates from the two-step analysis, but with the other parameters estimated and either common or separate for ecotypes, also resulted in significant increases in the residual deviance (data not shown).

Despite the better fit obtained using the one-step analysis, the fitted lines for the individual survival curves derived from the one-step and two-step models were similar to each other. They were also generally close to those obtained using the estimates for $K_i$, $C_V$ (N1601 only), and $\sigma$ from the constrained individual survival curve fitting (Figs 7, 8).

Discussion

This paper describes two different approaches to the statistical analysis of a large set of seed storage data. There are some advantages and disadvantages to both approaches, which should perhaps be weighted up in relation to the specific aims of a seed storage study. Both methods of analysis also considered the inclusion of a control viability parameter.

Low critical moisture content

In determining the values of the species constants for use in the Ellis and Roberts predictive viability equations, it is often only the moisture-related parameters that are estimated, ageing seeds at a range of moisture contents but at only one or two storage temperatures. It is then
Table 5. One-step estimation of the unknown parameters (with standard errors) in the viability equation [equation (6)] for Arabidopsis thaliana ecotypes NW20 and N1601

Model fitting carried out for each ecotype independently or concurrently.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Ecotype</th>
<th>$K_i$</th>
<th>$C_V$</th>
<th>$K_E$</th>
<th>$C_W$</th>
<th>$C_H \times 10^{-3}$</th>
<th>$C_Q \times 10^{-5}$</th>
<th>Residual mean deviance (degrees of freedom)</th>
<th>$F$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent, without $C_V$; no constraints</td>
<td>NW20</td>
<td>0.56 (0.029)</td>
<td>n.a.</td>
<td>8.50 (0.089)</td>
<td>5.22 (0.063)</td>
<td>5.90 (0.283)</td>
<td>6.82 (3.96)</td>
<td>9.30 (590)$^{A,B}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N1601</td>
<td>0.81 (0.026)</td>
<td>n.a.</td>
<td>8.24 (0.065)</td>
<td>5.10 (0.047)</td>
<td>5.49 (0.198)</td>
<td>8.59 (2.85)</td>
<td>7.27 (623)$^{C,D}$</td>
<td></td>
</tr>
<tr>
<td>Independent; $K_i$ (and $C_V$) set</td>
<td>NW20</td>
<td>0.94</td>
<td>n.a.</td>
<td>8.25 (0.076)</td>
<td>5.15 (0.057)</td>
<td>5.56 (0.240)</td>
<td>10.23 (3.42)</td>
<td>11.89 (591)$^p$</td>
<td>165.29***</td>
</tr>
<tr>
<td></td>
<td>N1601</td>
<td>1.26</td>
<td>0.8424</td>
<td>8.15 (0.062)</td>
<td>5.09 (0.046)</td>
<td>5.50 (0.193)</td>
<td>7.98 (2.79)</td>
<td>8.02 (624)$^p$</td>
<td>65.52***</td>
</tr>
<tr>
<td>Independent without $C_V$; universal $C_H$ and $C_Q$</td>
<td>NW20</td>
<td>0.50 (0.031)</td>
<td>n.a.</td>
<td>8.18 (0.064)</td>
<td>5.17 (0.066)</td>
<td>3.29</td>
<td>47.8</td>
<td>11.31 (592)$^p$</td>
<td>64.91***</td>
</tr>
<tr>
<td></td>
<td>N1601</td>
<td>0.70 (0.029)</td>
<td>n.a.</td>
<td>8.11 (0.052)</td>
<td>5.11 (0.054)</td>
<td>3.29</td>
<td>47.8</td>
<td>10.26 (625)$^p$</td>
<td>129.73***</td>
</tr>
<tr>
<td>Concurrent, without $C_V$; no constraints</td>
<td>NW20</td>
<td>0.56 (0.027)</td>
<td>n.a.</td>
<td>8.50 (0.084)</td>
<td>5.22 (0.059)</td>
<td>5.90 (0.267)</td>
<td>6.88 (3.73)</td>
<td>8.26 (1213)$^E$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N1601</td>
<td>0.81 (0.028)</td>
<td>n.a.</td>
<td>8.24 (0.069)</td>
<td>5.10 (0.050)</td>
<td>5.50 (0.211)</td>
<td>8.53 (3.04)</td>
<td>8.02 (624)$^p$</td>
<td></td>
</tr>
<tr>
<td>Concurrent, without $C_V$; $C_V$; common $C_Q$</td>
<td>NW20</td>
<td>0.56 (0.027)</td>
<td>n.a.</td>
<td>8.42 (0.058)</td>
<td>5.15 (0.038)</td>
<td>5.57 (0.166)</td>
<td>7.76 (2.35)</td>
<td>8.26 (1215)$^{CG}$</td>
<td>2.30</td>
</tr>
<tr>
<td></td>
<td>N1601</td>
<td>0.81 (0.028)</td>
<td>n.a.</td>
<td>8.31 (0.055)</td>
<td>5.16 (0.039)</td>
<td>5.63 (0.166)</td>
<td>8.39 (2.36)</td>
<td>8.34 (1216)$^{EH}$</td>
<td>12.11**</td>
</tr>
<tr>
<td>Concurrent, without $C_V$; common $C_W$, $C_Q$</td>
<td>NW20</td>
<td>0.57 (0.027)</td>
<td>n.a.</td>
<td>8.36 (0.054)</td>
<td>5.15 (0.039)</td>
<td>5.63 (0.166)</td>
<td>8.39 (2.36)</td>
<td>8.33 (1217)$^{HJ,L,M}$</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>N1601</td>
<td>0.80 (0.028)</td>
<td>n.a.</td>
<td>8.34 (0.054)</td>
<td>5.15 (0.039)</td>
<td>5.63 (0.166)</td>
<td>8.39 (2.36)</td>
<td>8.33 (1217)$^{HJ,L,M}$</td>
<td></td>
</tr>
<tr>
<td>Concurrent, without $C_V$; common $K_E$, $C_W$, $C_Q$</td>
<td>NW20</td>
<td>0.59 (0.023)</td>
<td>n.a.</td>
<td>8.35 (0.054)</td>
<td>5.15 (0.039)</td>
<td>5.63 (0.166)</td>
<td>8.39 (2.36)</td>
<td>8.33 (1217)$^{HJ,L,M}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N1601</td>
<td>0.79 (0.024)</td>
<td>n.a.</td>
<td>8.35 (0.054)</td>
<td>5.15 (0.039)</td>
<td>5.63 (0.166)</td>
<td>8.39 (2.36)</td>
<td>8.33 (1217)$^{HJ,L,M}$</td>
<td></td>
</tr>
<tr>
<td>Concurrent, without $C_V$; all parameters common</td>
<td>NW20</td>
<td>0.68 (0.020)</td>
<td>n.a.</td>
<td>8.36 (0.056)</td>
<td>5.16 (0.040)</td>
<td>5.64 (0.172)</td>
<td>8.27 (2.45)</td>
<td>8.80 (1218)$^j$</td>
<td>68.64***</td>
</tr>
<tr>
<td></td>
<td>N1601</td>
<td>0.79 (0.024)</td>
<td>n.a.</td>
<td>8.36 (0.056)</td>
<td>5.16 (0.040)</td>
<td>5.64 (0.172)</td>
<td>8.27 (2.45)</td>
<td>8.80 (1218)$^j$</td>
<td></td>
</tr>
<tr>
<td>Concurrent, without $C_V$; universal $C_H$, $C_Q$ common $K_E$, $C_W$</td>
<td>NW20</td>
<td>0.50 (0.025)</td>
<td>n.a.</td>
<td>8.14 (0.041)</td>
<td>5.13 (0.042)</td>
<td>3.29</td>
<td>47.8</td>
<td>10.76 (1219)$^k$</td>
<td>178.37***</td>
</tr>
<tr>
<td></td>
<td>N1601</td>
<td>0.70 (0.026)</td>
<td>n.a.</td>
<td>8.14 (0.041)</td>
<td>5.13 (0.042)</td>
<td>3.29</td>
<td>47.8</td>
<td>10.76 (1219)$^k$</td>
<td></td>
</tr>
<tr>
<td>Concurrent, without $C_V$; common $K_E$, $C_W$, $C_H$</td>
<td>NW20</td>
<td>0.58 (0.023)</td>
<td>n.a.</td>
<td>8.47 (0.045)</td>
<td>5.19 (0.038)</td>
<td>6.21 (0.040)</td>
<td>n.a.</td>
<td>8.41 (1218)$^i$</td>
<td>12.00**</td>
</tr>
<tr>
<td></td>
<td>N1601</td>
<td>0.78 (0.024)</td>
<td>n.a.</td>
<td>8.47 (0.045)</td>
<td>5.19 (0.038)</td>
<td>6.21 (0.040)</td>
<td>n.a.</td>
<td>8.41 (1218)$^i$</td>
<td></td>
</tr>
<tr>
<td>Concurrent, $\approx C_V$; two-step parameter values</td>
<td>NW20</td>
<td>0.94</td>
<td>=0.8424</td>
<td>8.17</td>
<td>5.12</td>
<td>5.59</td>
<td>8.01</td>
<td>10.34 (1223)$^m$</td>
<td>50.00***</td>
</tr>
<tr>
<td></td>
<td>N1601</td>
<td>1.26</td>
<td>0.8424</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Experiments 1, 2, 9–11, 15, 22, 30, and 34 were excluded from the analyses for both ecotypes. ‘n.a.’ Indicates that the parameter was not included in the model. ‘$\approx$’ Indicates that values were fixed. * Superscript letters in the residual mean deviance (r.m.d.) column indicate which models were compared (i.e. model with superscript ‘A’ was compared with model with superscript ‘a’, etc.); asterisks in the column of $F$-values indicate whether the change in residual deviance between the two models being compared was significant using an approximate $F$-test ($**P < 0.01$; ***$P < 0.001$). Shaded row indicates the selected model.
Fig. 7. Sample survival curves for seeds of Arabidopsis thaliana ecotype NW20 stored at the temperatures and MCs shown. Square symbols represent the observed data; the solid line is the result of individual curve fitting without $C_v$ and with constraining of $K_i$ (model 3, Table 2); the dotted line is the result of two-step fitting (Table 4); the dashed line is the result of one-step fitting (Table 5).
Fig. 8. Sample survival curves for seeds of *Arabidopsis thaliana* ecotype N1601 stored at the temperatures and MCs shown. Square symbols represent the observed data; the solid line is the result of individual curve fitting with $C_V$ and with constraining of $K_i$ and $C_V$ (model 4, Table 3); the dotted line is the result of two-step fitting (Table 4); the dashed line is the result of one-step fitting (Table 5).
assumed that the universal values for $C_H$ and $C_Q$ are satisfactory for describing the effects of temperature. In some studies where seeds are stored at low moisture contents, a break in the logarithmic relationship between MC and $\sigma$ is apparent and the break point is then determined using a ‘broken stick’ model. Below this ‘low critical moisture content’ further drying has little or no effect on longevity (Ellis et al., 1988, 1989, 1995, 1996). This low critical MC seems consistently to coincide with the top of the first ‘knee’ on moisture isotherms, where most of the water that is present in the seeds is strongly bound (phase I water binding), despite the fact that the isotherms are often determined at a different temperature than that at which the ageing experiments are carried out. Ellis et al. justify this by the fact that the MC/ RH relationship does not vary much with temperature in this region of the isotherm. However, Vertucci and Roos (1993) have published isotherms determined at different temperatures which clearly show that the shape of the isotherms change with temperature. As temperature decreases, the isotherm shifts upward with less pronounced curvature at low RH. This shift with temperature is also seen for the Arabidopsis data, though the isotherm at 6 °C departs from this pattern (fitted curves not shown). However, the magnitude of this shift is not as great as might be anticipated (Vertucci and Roos, 1993), and so the single fitted curve was used to predict where phase I water binding is complete and the expected break in the logarithmic relationship between $\sigma$ and MC occurs.

Thus, using the isotherm and by examination of the plots of $\log_{10}(\sigma)$ versus $\log_{10}$(MC) (Figs 5, 6), it appears that the breakpoint occurs between 3% and 4% MC at 45 °C and 55 °C. This is the MC region where longevity reaches an optimum (at these temperatures) and compares well with optimum MCs at 50 °C of c. 2.7% and c. 4.8% predicted for seeds of peanut and soybean, respectively (Vertucci and Roos, 1993). The lipid content of Arabidopsis seeds is midway between these two species [typically within the range 30–43% (mean for 11 ecotypes 36.1%± 1.7; Pathak et al., 1994) compared with 20% and 45% for soybean and peanut, respectively (Vertucci and Roos, 1990)]. The isotherms for Arabidopsis should thus fall between the isotherms of these two species at any given temperature and, as indeed appears to be the case, the low critical MC should be intermediate to their predicted optima. Having established where this critical MC region occurs, the data corresponding to storage at MCs within and below this region were excluded from subsequent analyses.

**Two-step versus one-step model fitting**

In the classical, two-step fitting of the viability model, individual survival curves are fitted using probit analysis and then multiple regression analysis is used to look at the effect of temperature and MC on $\sigma$. The first step can therefore be considered to be a ‘smoothing’ step in the model fitting; the residual variance in the MC–$\sigma$ regression analysis taking no account of the variation for each survival curve. Thus the fitted MC–$\sigma$ regression is not affected by any lack-of-fit associated with a particular survival curve.

By contrast, the one-step approach is based on minimizing the combined variation about all survival curves simultaneously, so that the fitted model will attempt to improve the fit for poorly fitting data sets at the expense of those that individually fit well. This latter approach has only relatively recently been possible using generally available statistical software, however, it is clearly more correct statistically, taking full account of all the variability in the data.

For the two-step approach, the MC–$\sigma$ regression analysis is only based on the estimated values of $\sigma$ for each survival curve; the residual variation is therefore based on relatively few degrees of freedom (one degree of freedom per survival curve rather than typically, 20–25). As a consequence, parameter estimates will tend to be less precise, and so it may be easier to accept a model other than the best-fit model. The main-drawback of the two-stage fitting is, therefore, the potential to introduce interpretative bias when comparing models.

In the one-step model fitting, all the data are included in the estimation for each parameter. A consequence of having more observations upon which each estimate is based is that the standard errors are smaller. For example, for $C_W$, the standard error in the two-step fitting was approximately 0.2 compared with a typical value of about 0.05 in the one-step fitting. This will influence the outcomes of comparisons between different seed-lots, in this case seed-lots corresponding to different ecotypes. Seed-lot comparisons carried out using a one-step modelling process are more likely to give rise to statistically significant differences. Indeed, if the data is somewhat variable, samples from a single ‘seed-lot’ may lead to small differences in the estimates for these parameters. This may occur, in particular, in wild plant species. Here, in the one-step analysis, both $K_E$ and $C_H$ were found to be differing significantly between the two-ecotypes (Table 5); this was not the case in the two-step fitting (Table 4).

A possible disadvantage with one-step fitting of the viability model may be in allowing parameters, such as $K_i$, to vary between seed-lots and/or individual storage experiments, because of problems with the convergence of models with large numbers of parameters (in this study if $K_i$ was allowed to vary for each storage experiment there would be an additional 26 parameters per ecotype). Generally, however, the issue of allowing $K_i$ to vary between storage experiments is not of great concern and seed survival curves for the same seed-lot are often justifiably constrained to a common estimate of $K_i$ (Ellis et al., 1990). Pre-equilibration prior to storage, particularly at higher RHs could result in some viability loss, although
this is usually not significant. When storage experiments are set up over an extended period, 17 months in the case of the experiments described in this paper, this could be more worrying (depending on the storage conditions of the bulk seed-lots). Fortunately, when the initial viability for each storage experiment was plotted against time of start of experiment, there was not a significant relationship, indicating that a common estimate for $K_i$ was acceptable (data not shown). Although it is necessary for the one-step analysis to have sensible initial values of the unknown parameters ($K_i$, $C_V$, etc), this can be readily achieved through initial data exploration (such as fitting individual survival curves) or based on past experience.

The control viability parameter

In both analyses, a control viability parameter was initially included in the modelling. Single ageing experiments are routinely used to assess seed quality, usually by simply using $K_i$ as an index of the quality of the seed-lot. Including the control viability parameter in analysing such experiments could separate out those seed-lots that simply lost viability through ageing and those where some seeds are not germinating because, for example, they are empty or dormant. In this study ecotype N1601 appeared to have a greater proportion of non-responders than ecotype NW20, although the estimate for $C_V$ varied considerably. It also appeared to be a more significant parameter in the ageing experiments carried out at the lower temperatures (Table 3), perhaps due to the greater ‘accuracy’ in sampling times that are more spread out. It is probably this variability in $C_V$ that contributed to the failure of the control viability model to converge in the one-step model fitting. This may have been caused by undue weight from data sets showing less evidence of an initial proportion of non-viable seeds. This could be overcome by weighting the analysis to reduce the impact of those data sets considered to be less reliable. As with $K_i$, there was not a significant relationship between $C_V$ and the time of initiation of the individual storage experiments.

The estimates and conclusions arising from the concurrent one-step analyses differed slightly depending on whether $K_i$ was estimated with $C_V$ not included in the model or $K_i$ and $C_V$ were both included but set at the values determined from the constrained individual survival curve fitting. These differences might cease to exist with smaller data sets and/or if there is less inherent variation within the bulk seed-lots. The production and experimental methods which were used for these Arabidopsis seeds may not have removed some potential sources of variability such as varying maturity, ageing within the freshly collected seed-lot or varying proportions of empty seeds in samples removed from the bulk at different times. Since a one-step approach is advocated, including the estimation of $K_i$ in the modelling process is undoubtedly more satisfactory.

Parameter values

One of the assumptions of the Ellis and Roberts model is that $K_E$ and $C_W$ do not vary with temperature (certainly within the range used here). The two-step analysis supports this simplification as the increase in residual variance was not significant for this reduced model in comparison with a model of independent linear regressions of $\log_{10}(\sigma)$ versus $\log_{10}(MC)$ at each temperature. The one-step analysis did not allow different values of $C_W$ and $K_E$ for each temperature to be estimated, but additional parameters could be added to test whether their values change with temperature (although this could again lead to problems with convergence). Alternatively, an interaction term could be added, though the nature of the interaction might not be readily identifiable. In their development of a model for predicting rates of corn seed deterioration, Tang et al. (2000) found that the optimal model included ‘storage environment coefficients’ for the effects of moisture, temperature, temperature×moisture, and temperature×ln moisture.

When analysing multiple seed-lots, in this case corresponding to two different ecotypes, using the one-step approach it is more likely that it will not be acceptable to constrain parameters to a common value for all seed-lots. This may lead to the conclusion that the viability parameters differ subtly within species. For example in this study, if $K_i$ was estimated and $C_V$ omitted in the model fitting, a model with separate estimates of $K_E$ and $C_H$ could be accepted (Table 5). This conclusion is contrary to one of the generalizations of the viability model, i.e. that the parameters will not differ within species. Indeed, this generality probably arose because of the limitations of the statistical software available at that time. It may be desirable to re-analyse some historical data sets using the one-step method to test this assumption. The various analyses and model comparisons consistently indicated that the universal values for $C_H$ and $C_Q$ could not be accepted without giving a significant increase in residual variance/deviance. In particular the quadratic term, $C_Q$, when it was estimated, was a magnitude smaller than the ‘universal’ value of 0.000478. Nonetheless, in the one-step analysis it was still a significant parameter; when excluded, there were significant increases in residual deviance ($P <0.001$; Table 5). Over the temperature range of $5\,^\circ\text{C}$ to $55\,^\circ\text{C}$ (a reasonable range in seed storage experimentation) there is, therefore, a non-linear effect of temperature on $log_{10}(\sigma)$. However, the extent of this non-linearity is not as large as suggested by the universal values (Dickie et al., 1990).

Predictions of longevity

As the results show, both the one- and the two-step models appear to provide reasonable fits to the observed data, and the resulting survival curves are close to the curves
obtained from the individual survival curve fitting (Figs 7, 8). In some cases, the one-step model appeared to explain the observed data less well. Nonetheless, this approach does provide the best predictions based on the original assumptions of the Ellis and Roberts seed viability model.

The most important concern, to those involved in storing seeds, is how the different models affect predictions of seed longevity, which model should be used in practice? In the one-step fitting, there were significant increases in residual deviance when the parameters were set at the values from the two-step fitting, even when \(K_i\) and \(C_V\) were also set at the same values. Furthermore, there were differences in the estimated values depending on whether or not \(K_i\) (and \(C_V\)) were fixed. Although the differences are apparently very small, the effect of these differences will be magnified when these parameter estimates are used to make predictions of longevity for MCs and temperatures towards the limits of the experimental range used for this study or to extrapolate beyond these limits.

For example, for seeds stored at 10% MC and 35 °C (within the experimental temperature range), the estimates for \(\sigma\) are 9.4 d (with a 95% CI from 8.6–10.4 d) using the two-step parameter values and 10.4 d (95% CI from 10–10.9 d) using the one-step parameter values. These confidence intervals overlap and it can be reasonably certain that, under these storage conditions \(\sigma\) will be in the region of 9.5–10.5 d. It is also worth noting, however, that because all the variation is included in the one-step analysis, the confidence intervals are much smaller using this approach. By contrast, extrapolating to the low MC, low temperature conditions typical of long-term seed bank storage, for example, 5% MC and −20 °C, the predictions for \(\sigma\) are 2087 years (95% CI from 1348–3230 years) and 1908 years (95% CI from 1387–2430 years), using the two-step and one-step estimates, respectively. Differences in predictions of this magnitude may be of some concern, particularly in the case of short-lived species.

Compared with other species, Arabidopsis seeds do not appear to be particularly long-lived during dry-storage, although there are other species which are known to produce much shorter-lived seeds. For example, for seeds stored at 10% MC and 35 °C (predicted value of \(\sigma\) for Arabidopsis seeds 10.4 d), the predicted value of \(\sigma\) for maize (Zea mays L.) seeds is 85.5 d; under the same storage conditions, the predicted value of \(\sigma\) for seeds of Ranunculus sceleratus (L.) (an undomesticated species found in the UK) is 1.7 d.

Conclusions

One-step fitting, including all the data in the analysis, is more appropriate than two-step fitting of the viability model, which essentially involves a smoothing step. It is therefore recommended that one-step fitting of the viability model should be used routinely for analysing seed storage data, even if it is only the moisture relations of seed longevity which are being considered.

The effects of moisture content and temperature on seed longevity did not differ between the two ecotypes; the only parameter which differed significantly was the initial viability parameter, \(K_0\) (0.59 and 0.79 for these seed-lots of NW20 and N1601, respectively). The estimated values for \(K_E, C_W, C_H,\) and \(C_Q\) are 8.35, 5.15, 0.0563, and 0.0000839, respectively. The non-linear effect of temperature on seed longevity (\(\sigma\)) was confirmed, however, the magnitude of the quadratic relationship was not as large as that predicted by the universal temperature value for \(C_Q\) (0.000478).

The inclusion of the control viability parameter, correcting for non-respondents in the population, may be appropriate for some seed lots, though in this case was unable to be estimated within the one-step approach.

Having established the viability constants for these two ecotypes they can be used as standards to compare the longevity of other lines stored in the seed bank of the Nottingham Arabidopsis Stock Centre (or elsewhere). In particular, this may highlight lines which are particularly short-lived or which appear to be surviving longer than expected.

References


