**Annex IV**

**Analysis of data of Annex III with SAS/STAT® 15.1.**

Cumulative germination through time (data from Piacco, 1954) are analysed (see Annex III for the dataset) as an example. These data were recorded at uneven time spacings. This is sensible if it is done so that more frequent observations are performed at the beginning, as well as during steeper changes, of the germination time-course, to obtain a better delineation of the progress curve and its inception. In the present case, the adoption of closer observation times appears to have been a bit later and pointlessly protracted.

It should be noted that germination data through time can be arranged as either different dependent variable responses obtained at diverse times, i.e. a multivariate response (data arrangement 1, in Annex III), or as a single response variable (i.e. univariate) whose outcome depends on the time of observation (data arrangement 2, in Annex III), which is therefore considered a factor for a longitudinal study. Data must be arranged differently, according to the kind of analysis. For multivariate analysis, the germination responses at every time have to be organized into separate columns.

The comments to the statistical analysis presented here are based on the SAS/STAT® 15.1 User's Guide (2018), as well as on Littell et al. (2006) and Stroup et al. (2018), to which the readers should refer for a more in-depth exposition of the matter.

Although 'temp' (the temperature used for soaking seeds prior to the germination test) is a numerical variable, GLM uses linear regression to model continuous variables as covariates and thereby it assumes linearity of response to factor variables, but this is not guaranteed in this case. It can be tested with:

```
proc SGSCATTER data=reffile; /* 'reffile' is the data file, not to be specified if it is the current input*/
plot (d0 d1 d2 d3 d3_5 d4 d4_5 d5 d5_5 d6 d6_5 d7)*temp / pbspline;
run;
```

RESULT:

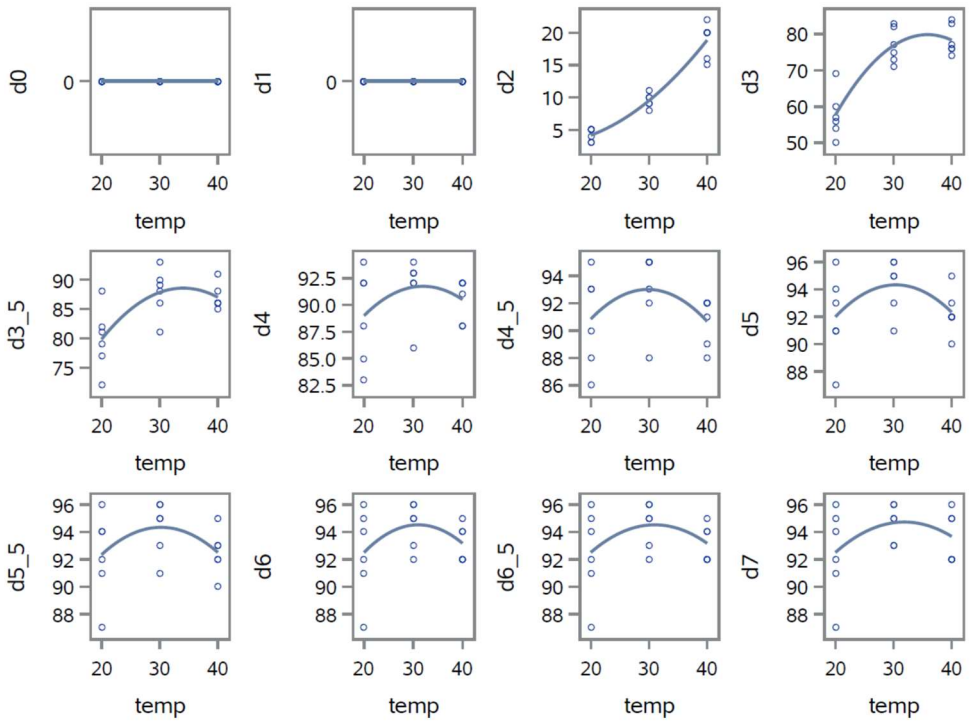



Figure 1

In Figure 1, it is evident that, from d2 to d7 (i.e., from day 2 to day 7), the germination response to increasing temperature is not linear in the range tested. Specifically, a power-like increase is displayed at d2, whereas a maximum at 30 °C is apparent starting from d4. An intermediate situation takes place at d3 and d3_5 (note that the underscore is used in place of the dot to separate decimals because the latter is not accepted in the variable name by the SAS software). Hence, the response to 'temp' is not linear and must be dealt with as a classification independent variable in an ANOVA model.

Homogeneity of variances is tested (by means of Levene's test). This requires the use of the GLM procedure and a multivariate organization of germination responses (data arrangement 1, in Annex III). Counts are used in this multivariate analysis, but proportions/percentages work as well.

```
/*MANOVA model*/
proc GLM;
class temp;
model d0 d1 d2 d3 d3_5 d4 d4_5 d5 d5_5 d6 d6_5 d7 = temp;
repeated time;
means temp / Tukey hovtest=Levene;
run;
```

With respect to a typical multivariate analysis of variance (MANOVA), the SAS GLM procedure offers a 'repeated' statement that performs MANOVA analysis but applies an adjustment to account for the correlation of measurements taken on the same experimental unit (or subject), i.e. on the same plate, by performing a repeated measures analysis of variance. The 'repeated' statement requests to perform a repeated-measures MANOVA wherein multiple response values on the same line in the data file are considered sequential measurements taken on the same subject through time. Note that time was not previously introduced as a variable (though nothing prevents introducing it in the 'class' statement), neither it appears in the data table, and is therefore identified here as the implicit variable across which repeated measurements were taken. MANOVA (not ANCOVA, since time is managed as a classification variable) is thus elicited to test both the effect of time and its interaction with the modelled factor. In addition, one-way ANOVAs are automatically performed for every multivariate response (that is, each observation time). The 'Tukey' option of the 'means' statement requests that means of the three levels of the 'temp' factor within each observation time are compared with Tukey's studentized range test (HSD). The 'hovtest=Levene' option invokes the Levene's test for the homogeneity of variances within each timepoint.

RESULTS (excerpts):

Table 1. Dependent Variable: d2

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| **Model** | 2 | 661.3333333 | 330.6666667 | 105.16 | <.0001 |
| **Error** | 15 | 47.1666667 | 3.1444444 | | |
| **Corrected Total** | 17 | 708.5000000 | | | |

Table 2. Dependent Variable: d3

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| **Model** | 2 | 1593.444444 | 796.722222 | 29.04 | <.0001 |
| **Error** | 15 | 411.500000 | 27.433333 | | |
| **Corrected Total** | 17 | 2004.944444 | | | |

Table 3. Dependent Variable: d3_5

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| **Model** | 2 | 232.1111111 | 116.0555556 | 6.97 | 0.0072 |
| **Error** | 15 | 249.6666667 | 16.6444444 | | |
| **Corrected Total** | 17 | 481.7777778 | | | |

Table 4.

| MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no time Effect<br>H = Type III SSCP Matrix for time<br>E = Error SSCP Matrix<br><br>S=1 M=3.5 N=2.5 | | | | | |
|---|---|---|---|---|---|
| **Statistic** | **Value** | **F Value** | **Num DF** | **Den DF** | **Pr > F** |
| **Wilks' Lambda** | 0.0001545 | 5032.25 | 9 | 7 | <.0001 |
| **Pillai's Trace** | 0.9998455 | 5032.25 | 9 | 7 | <.0001 |
| **Hotelling-Lawley Trace** | 6470.0336917 | 5032.25 | 9 | 7 | <.0001 |
| **Roy's Greatest Root** | 6470.0336917 | 5032.25 | 9 | 7 | <.0001 |

Table 5.

| MANOVA Test Criteria and F Approximations for the Hypothesis of no time*temp Effect<br>H = Type III SSCP Matrix for time*temp<br>E = Error SSCP Matrix<br><br>S=2 M=3 N=2.5 | | | | | |
|---|---|---|---|---|---|
| **Statistic** | **Value** | **F Value** | **Num DF** | **Den DF** | **Pr > F** |
| **Wilks' Lambda** | 0.01488516 | 5.60 | 18 | 14 | 0.0011 |
| **Pillai's Trace** | 1.64662320 | 4.14 | 18 | 16 | 0.0032 |
| **Hotelling-Lawley Trace** | 21.74020886 | 7.90 | 18 | 8.5 | 0.0022 |
| **Roy's Greatest Root** | 19.45604490 | 17.29 | 9 | 8 | 0.0002 |
| **NOTE: F Statistic for Roy's Greatest Root is an upper bound.** | | | | | |
| **NOTE: F Statistic for Wilks' Lambda is exact.** | | | | | |

Table 6. The GLM Procedure for Repeated Measures Analysis of Variance - Tests of Hypotheses for Between Subjects Effects.

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| **temp** | 2 | 747.0648148 | 373.5324074 | 9.58 | 0.0021 |
| **Error** | 15 | 584.5833333 | 38.9722222 | | |

Table 7. The GLM Procedure for Repeated Measures Analysis of Variance - Univariate Tests of Hypotheses for Within Subject Effects

| Source | DF | Type III SS | Mean Square | F Value | Pr > F | Adj Pr > F | |
|---|---|---|---|---|---|---|---|
| | | | | | | G - G | H-F-L |
| **time** | 11 | 307002.5926 | 27909.3266 | 5848.29 | <.0001 | <.0001 | <.0001 |
| **time*temp** | 22 | 1854.4907 | 84.2950 | 17.66 | <.0001 | <.0001 | <.0001 |
| **Error(time)** | 165 | 787.4167 | 4.7722 | | | | |

Table 8.

| Levene's Test for Homogeneity of d2 Variance ANOVA of Squared Deviations from Group Means | | | | | |
|---|---|---|---|---|---|
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| **temp** | 2 | 111.5 | 55.7284 | 5.01 | 0.0216 |
| **Error** | 15 | 167.0 | 11.1309 | | |

Figure 2.

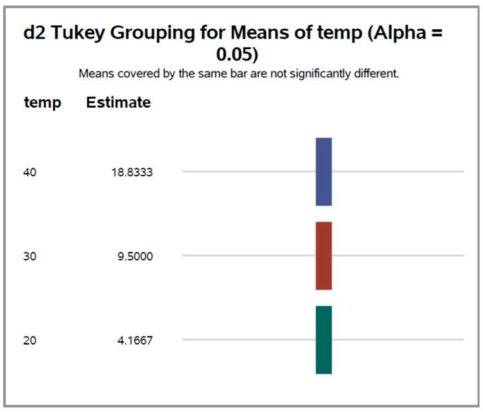


Figure 3.

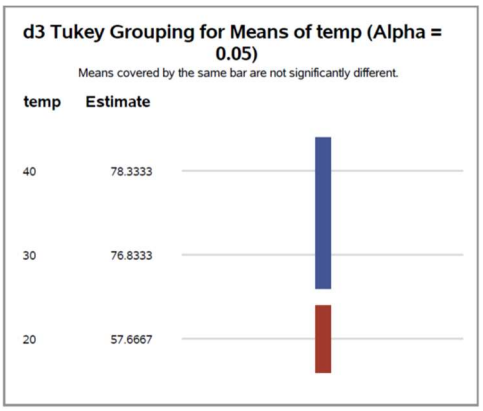
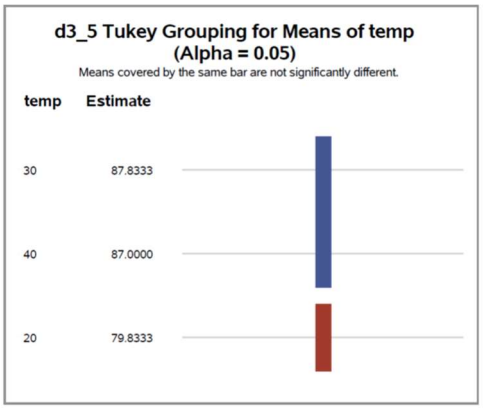



Figure 4.




Within single timepoints, one-way ANOVAs indicate a significant effect of 'temp' at times 2 d, 3 d and 3.5 d (Tables 1, 2 and 3), where 'significant' is intended as showing evidence of statistical significance at the $P \leq 0.05$ level. No significant effect of 'temp' was detected at the other timepoints (not shown). Note that the analysis is not possible for germination at days 0 and 1, whose mean value is zero. Table 4 shows that, as obvious, MANOVA found a highly significant effect of time on the germination progress. Table 5 adds that the germination progress through time is significantly different across soaking temperatures ('temp'), as shown in Figure 1. Accordingly, both the between-subjects (where plates are the subjects on which the repeated measures were performed) effect, that is 'temp' (in fact, different plates were subjected to diverse temperatures), as well as the within-subjects effects (i.e., time and its interaction with 'temp', since repeated measures through time were taken on the same plate), had a significant effect (Tables 6 and 7). To account for the correlation through repeated measurements, probabilities of the *F* tests were also corrected according to Greenhouse-Geisser (G - G) or Huynh-Feldt-Lecoutre (H-F-L) adjustments to numerator and denominator degrees of freedom (Table 7). Finally, Table 8 displays the Levene's test for homogeneity of variances at 'd2' (day 2), which is the only significant test result shown because all the other timepoints hint to non-significant heteroskedasticity. Note that these tests compare variances across the diverse levels of 'temp', not through different timepoints. Strong heteroskedasticity is expected to occur among different timepoints during the time-course of germination, but MANOVA cannot evidence it. Neither this is a problem, if the effect of treatments is separately contrasted within timepoints: a one-way ANOVA for every observation time is performed at every timepoint, allowing to identify the timepoints at which the progress of germination is different across the 'temp' levels. At 2 d heteroskedasticity was significant, but, given the evident diversity of the values among temperature levels at this timepoint, heteroskedasticity does not seem

to have altered the inference drawn from ANOVA. Unless the group variances (i.e., the between-replicates variances) are extremely different, the usual ANOVA test is relatively robust when the contrasted groups are all about the same size (Zar, 1999). The Tukey test for multiple pairwise comparisons was performed to identify which of the three levels of the 'temp' variable is different from each other. Figures 2-4 show a different effect for each of the three temperatures at 2 days, but only for 20 °C with respect to the other two temperatures at days 3-3.5.

Although a multivariate analysis is a reasonable approach to repeated measures, GzLMM allow a conceptually better analysis of germination time-courses. Specifically, when observations at a given timepoint are strongly affected by the treatment, the means, and therefore the variances, are largely different as well. Strong heteroscedasticity can bias the statistical analysis, essentially when some means are much more different than the others, and GzLMM are then the best statistical approach, especially because heteroscedasticity is associated with non-normal errors (Stroup, 2015).

The GLIMMIX procedure is therefore used because data are binomial and clustered. A GzLMM longitudinal (through time) analysis requires a univariate response variable, 'germ' (cumulative germination), which must therefore be listed in a single column (data arrangement 2, in Annex III).

```
proc GLIMMIX method=Laplace;
class plate temp time;
model germ/n = temp time temp*time / link=probit;
random plate(temp);
*nloptions tech=nrridg;
/*Note that the above 'nloptions' statement can be used to solve convergence and estimation problems. In
the present instance it has not been used, and it will be discussed when introducing marginal models*/
run;
```

As we'll see, this model has some troubles, and it converges to a solution only if fitted by true maximum likelihood, and this requires the marginal distribution to be numerically approximated by the Laplace method ('method=Laplace' option) or by adaptive Gauss-Hermite quadrature (which, anyway, would require some re-arrangement of the syntax). This is anyway a useful solution because applying one of these integral approximations is recommended for the initial evaluation of the model since they consent the use of true likelihood and thus the obtainment of unbiased estimation of overdispersion and of diagnostic tests for the variance/covariance structure (Stroup et al., 2018). This aspect is of much greater concern here than it was for end-of-test data analysed in Annex II because fitting the right variance/covariance structure is compelling in the analysis of longitudinal data.

The link function, as already mentioned in Annex II, is set to probit because this transformation is theoretically, and practically, more suitable to germination data.

Since the germination progress is not linear to time, not either on the linked scale (as we'll see), even 'time' must be considered a categorical variable. Time can be modelled as a continuous variable by, for example, introducing suitable polynomial/logarithmic terms that linearize the relationship; a possible solution in this sense will be considered subsequently.

Plates represent clusters of seeds and are subjected to sequential observations. To take into account that plates are independent blocks, a 'random' statement has been introduced that considers plates as representing a random effect within each temperature condition; that is, the germination response is assumed to randomly deviate from the seed population average because plates represents random samples of the population. This deviation is assumed to be constant throughout the germination time-course, and the plate effect is therefore modelled as if it represents a set of random intercepts normally distributed around

the general linear model. As plates are numbered 1-6 for every between-subjects level of the 'temp' factor, nesting is made explicit to indicate that plates coded with the same number but containing seeds pre-treated at different soaking temperatures are different entities (i.e., different subjects).

RESULTS (excerpts):

Table 9.

| Model Information | |
|---|---|
| **Data Set** | WORK.REFFILE |
| **Response Variable (Events)** | germ |
| **Response Variable (Trials)** | n |
| **Response Distribution** | Binomial |
| **Link Function** | Probit |
| **Variance Function** | Default |
| **Variance Matrix** | Not blocked |
| **Estimation Technique** | Maximum Likelihood |
| **Likelihood Approximation** | Laplace |
| **Degrees of Freedom Method** | Containment |

Table 10.

| Class Level Information | | |
|---|---|---|
| **Class** | **Levels** | **Values** |
| **plate** | 6 | 1 2 3 4 5 6 |
| **temp** | 3 | 20 30 40 |
| **time** | 12 | 0 1 2 3 4 5 6 7 3.5 4.5 5.5 6.5 |

Table 11.

| | |
|---|---|
| **Number of Observations Read** | 216 |
| **Number of Observations Used** | 216 |
| **Number of Events** | 14668 |
| **Number of Trials** | 21600 |

Table 12.

| Dimensions | |
|---|---|
| **G-side Cov. Parameters** | 1 |
| **Columns in X** | 52 |
| **Columns in Z** | 18 |
| **Subjects (Blocks in V)** | 1 |
| **Max Obs per Subject** | 216 |

Table 13.

| Optimization Information | |
|---|---|
| **Optimization Technique** | Dual Quasi-Newton |
| **Parameters in Optimization** | 37 |
| **Lower Boundaries** | 1 |
| **Upper Boundaries** | 0 |
| **Fixed Effects** | Not Profiled |
| **Starting From** | GLM estimates |

Table 14.

| Fit Statistics | |
|---|---|
| **-2 Log Likelihood** | 803.45 |
| **AIC (smaller is better)** | 877.45 |
| **AICC (smaller is better)** | 893.25 |
| **BIC (smaller is better)** | 910.40 |
| **CAIC (smaller is better)** | 947.40 |
| **HQIC (smaller is better)** | 882.00 |

Table 15.

| Fit Statistics for Conditional Distribution | |
|---|---|
| **-2 log L(germ \| r. effects)** | 763.13 |
| **Pearson Chi-Square** | 58.53 |
| **Pearson Chi-Square / DF** | 0.27 |

Table 16.

| Covariance Parameter Estimates | | |
|---|---|---|
| **Cov Parm** | **Estimate** | **Standard Error** |
| **plate(temp)** | 0.01026 | . |

Table 17.

| Type III Tests of Fixed Effects | | | | |
|---|---|---|---|---|
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **temp** | 2 | 15 | 0.00 | 1.0000 |
| **time** | 11 | 165 | 341.95 | <.0001 |
| **temp*time** | 22 | 165 | 4.25 | <.0001 |

Table 9 shows that the Variance matrix is not subdivided in blocks (=not blocked), because subjects are not explicitly identified, even though plates are indeed subjects of repeated observations. Not having data processed in terms of subjects leaves the default Containment Method for calculating the Degrees of Freedom. As the model is conditional, that is, the plate effect is considered a random factor, maximum likelihood can be used as estimation technique thanks to the Laplace Likelihood Approximation. Table 10 displays the classification effects and their levels. It can be noted that time levels are not numerically ordered, since time is considered a classification effect. Table 11 provides the number of aggregate observations (namely, 6 plates x 3 temperatures x 12 times = 216) as well as of individual Bernoulli trials (for a total of 216 observations x 100 seed/plate = 21600 trials, which gave 14668 response events; note that this would really be so only if an independent plate had been used for each observation). Table 12 indicates that even though there are three groups of plates (i.e., there are six plates for each of the three temperature levels), a single random parameter is estimated (G-side Covariance Parameters = 1), namely, the variance of the 'plate(temp)' effect, because, if not otherwise specified, the default variance/covariance structure is 'variance components', and it amounts to estimating the variance of the response among plates, assuming this variance is the same across treatments. In the same table it is shown that there are 52 columns in the X matrix (the matrix for the fixed effects), corresponding to a column for the intercept, three columns for the levels of the 'temp' effect, 12 for the 'time' effect, and 36 columns for their interaction. There are 18 columns in the Z matrix (the matrix for G-side random effects) for this model, corresponding to the 18 plates (six plates for each of the three temperature levels). On the other hand, there is a single subject (Blocks in V, where the V matrix includes all the random effects, both on the G-side and R-side), since data are not processed by subjects (which would otherwise mean that the V matrix would be subdivided in as many blocks as the number of subjects), and the whole matrix is therefore considered a single block. Thus, the maximum number of observations per subject is 216 (as there are 216 observations for a single block). The default optimization technique for GzLMM with binomial data, the Dual Quasi-Newton method, is used (Table 13). Because a Laplace approximation of maximum likelihood is used to compute the objective estimation function, the fixed effects are not profiled from the optimization (as otherwise typically done), thus that both fixed effects and covariance parameters participate in the optimization, and a total of 37 parameters is therefore indicated to be in optimization, corresponding to 36 fixed effect parameters and one random variance component (which

has a lower bound of 0). Also, as the fixed effects are part of the optimization, a few GLM iterations are initially performed to obtain starting values for the optimization of the fixed effect parameters. It might be noticed that 36, the number of fixed effect parameters, is obtained because all fixed effects are classification variables and each categorical predictor with *L* levels is coded into *L*−1 dummy variables, so that the estimated fixed parameters are: the model intercept, two dummies for temperature, 11 dummies for time and 2 x 11 dummies for their interaction (as can be checked by adding the 'solution' option into the 'model' statement). The Fit Statistics, which can be used to compare different models that use estimation techniques based on true likelihood, are shown in Table 14, whereas Fit Statistics for Conditional Distribution are displayed in Table 15. In the latter table, the Pearson Chi-Square / DF statistics indicates that there is not overdispersion in the model, since this ratio is largely below 1. All the relevant factors should therefore have been considered. The estimate of the Covariance Parameter, that is, the variance among plates, is given in Table 16 as 0.01026 on the probit scale. Variances of random factors are modelled as constant effects on the linked scale and correspond to variable spans on the percentile data scale, with a maximum value in the middle range, i.e. 50%. A rough method to obtain an idea of the magnitude of these variances is to consider the amplitude of the corresponding standard deviations (on the probit scale, the square root of a variance represents the standard normal deviates, *z*, from the mean of the standard normal distribution) once transposed onto the data scale around a fixed percentile, with 50% being the best reference. This can be seen by, for example, using the cumulative NORM.S.DIST function of Excel and considering that the mean of the standard normal distribution is centred to zero on the probit scale, which corresponds to 50 % germination on the percentile scale. For these data, this naïvely computed value corresponds to a maximum standard deviation of 4 % around 50 % on the percentile scale. Such variability is low, and its standard error could not be calculated (unless the 'nloptions' statement is used), hinting to some computational problem related to the fact that some means are on the boundary of the parameter space since they have values of zero. Finally, the type III tests of fixed effects are displayed in Table 17, and indicate a significant effect for time and temperature x time interaction but not for temperature.

As said, the above-mentioned missing standard error of the random variance suggests that there is some trouble in the calculations. Moreover, 'temp', which was shown to be significant by MANOVA (Table 6), is not significant in this analysis. What is most probably troubling the analysis is that some averages, notably those observed at the initial times of observations (days 0-1), have a value of zero. In generalized linear (mixed) models, means are analysed on the linear scale, and a mean of zero cannot be defined on a probit scale (as well as on a logit scale). These means were already an issue for the multivariate analysis, as they prevented performing the one-way ANOVA for data at days 0 and 1 (in such case, the problem was that an *F* ratio cannot be defined if the error variance is null), though the analysis of the subsequent non-zero means was not prevented. Since in GzLMM all the longitudinal means are modelled together, the inclusion of zero means makes the computation difficult to manage. Nevertheless, it ought to be noted that these zero means correspond to the initial lag time, required for seed imbibition and metabolism re-activation, when the seeds cannot yet attain visible germination. There is, therefore, a strong conceptual difference between this stage and the subsequent germination time-course. Thus, a separate analysis of the non-zero means can be a sensible approach, and it can be obtained with a 'BY' statement, which utilizes the 'stage' variable to define the initial lag period when the germination progress curve has not yet started. To identify, and to separate, the initial lag in a given experiment makes sense because a lag time is almost always required to start germination, and, of course, no difference in the germination progress can occur before germination is observed and all the data are zeros. This actually corresponds to exclude the lag data from the analysis because they cannot be properly analysed with ANOVA and linear models, at least without more sophisticated approaches.

It is worth mentioning that the lag data might be simply deleted by means of the following program code:

```
DATA new;
set reffile;
if germ=0 then delete;
run;
```

This 'DATA' command reads the data set 'work.reffile', and then removes all observations that do not meet the condition specified in the 'if' statement. The result is stored in the data set 'work.new'. Modelling can then be done processing this latter file, so that the 'by' statement is no longer necessary. It must be noted, however, that, in this way, any observation with 'germ=0' would be deleted, and even though this is not a trouble for the present dataset, it can bias estimates because it would erase also individual replications with values of zero for means that are not zero themselves (since some other replications would then have non-zero values). It could also be unnecessary, because, whereas means are modelled on the linked scale, where zeros cannot be transposed, the single observations are modelled on the data scale, where zeros are not of trouble. This would require, anyway, that true likelihood and not pseudo-likelihood is used as estimation technique, because pseudo-likelihood involves a transposition of the whole dataset on the linked scale. Nevertheless, due to shrinkage of BLUPs of random effects toward their mean, means with only some data having a value of zero can still be dealt with even if pseudo-likelihood is utilized, but standard errors can hugely increase. Pseudo-likelihood is applied by default for models containing random factors, or an R-side random effect (i.e., a variance/covariance structure), for non-Gaussian data, and it can be substituted with true likelihood, if there are not R-side effects, using an integral approximation by Laplace or adaptive quadrature method (Stroup et al., 2018).

As the Laplace integral approximation allows proper overdispersion diagnostics and covariance parameters testing (Stroup et al., 2018), a 'covtest' statement is included to test significance of the inclusion of the plate random effect, which was previously shown to be quite small:

```
proc GLIMMIX method=Laplace;
by stage;
class plate temp time;
model germ/n = temp time temp*time / link=probit;
random plate(temp);
covtest zeroG;
run;
```

RESULTS (excerpts):

Table 18: stage=progress

| | |
|---|---|
| **Number of Observations Read** | 180 |
| **Number of Observations Used** | 180 |
| **Number of Events** | 14668 |
| **Number of Trials** | 18000 |

Table 19: stage=progress

| Dimensions | |
|---|---|
| **G-side Cov. Parameters** | 1 |
| **Columns in X** | 44 |
| **Columns in Z** | 18 |
| **Subjects (Blocks in V)** | 1 |
| **Max Obs per Subject** | 180 |

Table 20: stage=progress

| Fit Statistics | |
|---|---|
| **-2 Log Likelihood** | 803.45 |
| **AIC (smaller is better)** | 865.45 |
| **AICC (smaller is better)** | 878.86 |
| **BIC (smaller is better)** | 893.06 |
| **CAIC (smaller is better)** | 924.06 |
| **HQIC (smaller is better)** | 869.26 |

Table 21: stage=progress

| Fit Statistics for Conditional Distribution | |
|---|---|
| **-2 log L(germ \| r. effects)** | 763.14 |
| **Pearson Chi-Square** | 58.53 |
| **Pearson Chi-Square / DF** | 0.33 |

Table 22: stage=progress

| Covariance Parameter Estimates | | |
|---|---|---|
| **Cov Parm** | **Estimate** | **Standard Error** |
| **plate(temp)** | 0.01026 | 0.004422 |

Table 23: stage=progress

| Type III Tests of Fixed Effects | | | | |
|---|---|---|---|---|
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **temp** | 2 | 15 | 7.17 | 0.0065 |
| **time** | 9 | 135 | 417.94 | <.0001 |
| **temp*time** | 18 | 135 | 5.20 | <.0001 |

Table 24: stage=progress

| Tests of Covariance Parameters Based on the Likelihood | | | | | |
|---|---|---|---|---|---|
| **Label** | **DF** | **-2 Log Like** | **ChiSq** | **Pr > ChiSq** | **Note** |
| No G-side effects | 1 | 838.61 | 35.16 | <.0001 | MI |

*MI: P-value based on a mixture of chi-squares.*

This model properly finds a null variance among plates for the lag stage, as well as non-significance of both the fixed effects and the random factor, which is obvious in absence of germination. On the other hand, it provides a better analysis of the germination progress stage. Table 18 shows that, for the progress stage, the number of aggregate observations is 180 (namely, 6 plates x 3 temperatures x 10 times = 180), with a corresponding number of Bernoulli trials. Table 19 indicates that there are 44 columns in the X matrix (the matrix for the fixed effects), corresponding to a column for the intercept, three columns for the levels of the 'temp' effect, 10 for the 'time' effect, and 30 columns for their interaction. There are 18 columns in the Z matrix (the matrix for G-side random effects), corresponding to the 18 plates (six plates for each of the three temperature levels), and the maximum number of observations per subject is 180 (as there are 180 observations for a single block in the V matrix). A total of 31 parameters were under optimization (not shown), corresponding to 30 fixed effect parameters and one random variance component (which has a lower bound of 0). The number of fixed effect parameters is obtained as: the model intercept, two dummies for temperature, 9 dummies for time and 2 x 9 dummies for their interaction (as can be checked by adding the 'solution' option into the 'model' statement). The Fit Statistics (Table 20) have all declined (i.e., improved), except the -2 Log Likelihood (which is not penalized for the number of parameters), with respect to Table 14, for the sole reason that the number of parameters has diminished, while the data are the same but for the null means of the lag stage. The Fit Statistics for Conditional Distribution (Table 21) are the same, as they derive from the -2 Log Likelihood, but the overdispersion parameter has slight increased (though still largely < 1), because of the smaller degrees of freedom after the two lag timepoints were analysed apart. The estimate of the Covariance Parameter (Table 22), that is, the variance among plates, still is 0.01026 on the probit scale, but its standard error has now been estimated. The type III tests of fixed effects are displayed in Table 23, and now they indicate a significant effect ($P \leq 0.05$) for both time and temperature as well as for their interaction. Finally, the Tests of Covariance Parameters (Table 24) demonstrates a significant effect of the plate random effect. The previously highlighted troubles have now been addressed.

The analysis can then be further implemented by plotting the residuals and providing estimates of means and multiple comparisons. Least-square means (LS-means) ought to be computed for the highest interaction

effect (of time with the other fixed factors) present in the model, i.e. temp*time in this instance, and multiple contrasts are usually requested within each timepoint, although this depends upon the scope of the analysis.

```
/*Conditional model with categorical time and integral approximation*/
proc GLIMMIX method=Laplace order=data plots=(residualpanel(ilink marginal) studentpanel(conditional));
by stage;
class plate temp time;
model germ/n = temp time temp*time / link=probit;
random plate(temp);
lsmeans temp*time / cl ilink plot=meanplot slice=time slicediff=time adjust=smm;
run;
```

Residual panels have been described in Annex II. The 'lsmeans' statement prescribes that: least-squares means are estimated for the levels of the 'temp*time' effect (on the linked scale) with their confidence limits (cl), whereas the 'ilink' option requests that the estimated means and confidence limits are also reported on the scale of the data. The 'plot=' option ensures that a plot of the estimates (on the linked scale) is displayed. Significance of the between-subjects effect is assessed through time points with the 'slice' option, and multiple contrasts are performed at each timepoint ('slicediff=time') with the studentized maximum modulus adjustment for multiplicity ('adjust=SMM'), which, for longitudinal data, protects the overall Type I error rate better than the Tukey's adjustment. The SMM adjustment can be useful in longitudinal data analysis to reduce the risk that small stochastic changes in the germination progress curve are identified as significant. The 'order=data' option was introduced in the GLIMMIX statement because, by default, the SAS software sorts the levels of classification variables according to an alphanumeric order that, as shown in Table 10, is not the same as if the variable were considered numerically continuous. A correct display of the plot of the estimates (requested by the 'plot=meanplot' option) requires that timepoints are arranged in increasing order. If the dataset is already ordered according to increasing time, then adding the 'order=data' option will do the job.

RESULTS (excerpts):

Figure 5: stage=progress

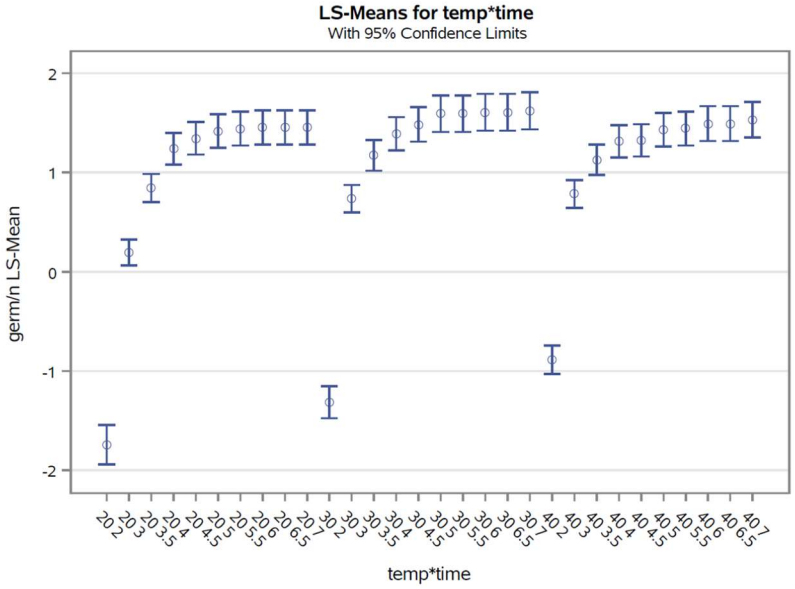


Table 25: stage=progress

| Tests of Effect Slices for temp*time Sliced By time | | | | |
|---|---|---|---|---|
| time | Num DF | Den DF | F Value | Pr > F |
| 2 | 2 | 135 | 24.78 | <.0001 |
| 3 | 2 | 135 | 23.48 | <.0001 |
| 3.5 | 2 | 135 | 5.88 | 0.0036 |
| 4 | 2 | 135 | 0.85 | 0.4317 |
| 4.5 | 2 | 135 | 1.02 | 0.3617 |
| 5 | 2 | 135 | 1.14 | 0.3243 |
| 5.5 | 2 | 135 | 0.90 | 0.4096 |
| 6 | 2 | 135 | 0.75 | 0.4742 |
| 6.5 | 2 | 135 | 0.75 | 0.4742 |
| 7 | 2 | 135 | 0.87 | 0.4223 |

Table 26: stage=progress

| Simple Effect Comparisons of temp*time Least Squares Means By time Adjustment for Multiple Comparisons: SMM | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Simple Effect Level | temp | temp | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Adj P | Alpha | Lower | Upper | Adj Lower | Adj Upper |
| time 2 | 20 | 30 | -0.4279 | 0.1299 | 135 | -3.29 | 0.0013 | 0.0038 | 0.05 | -0.6849 | -0.1710 | -0.7419 | -0.1140 |
| time 2 | 20 | 40 | -0.8560 | 0.1239 | 135 | -6.91 | <.0001 | <.0001 | 0.05 | -1.1012 | -0.6109 | -1.1556 | -0.5565 |
| time 2 | 30 | 40 | -0.4281 | 0.1093 | 135 | -3.92 | 0.0001 | 0.0004 | 0.05 | -0.6443 | -0.2120 | -0.6922 | -0.1640 |
| time 3 | 20 | 30 | -0.5413 | 0.09634 | 135 | -5.62 | <.0001 | <.0001 | 0.05 | -0.7318 | -0.3507 | -0.7741 | -0.3084 |
| time 3 | 20 | 40 | -0.5904 | 0.09680 | 135 | -6.10 | <.0001 | <.0001 | 0.05 | -0.7819 | -0.3990 | -0.8243 | -0.3565 |
| time 3 | 30 | 40 | -0.04916 | 0.09954 | 135 | -0.49 | 0.6222 | 0.9456 | 0.05 | -0.2460 | 0.1477 | -0.2897 | 0.1914 |
| time 3.5 | 20 | 30 | -0.3298 | 0.1060 | 135 | -3.11 | 0.0023 | 0.0068 | 0.05 | -0.5394 | -0.1201 | -0.5860 | -0.07357 |
| time 3.5 | 20 | 40 | -0.2853 | 0.1052 | 135 | -2.71 | 0.0075 | 0.0224 | 0.05 | -0.4933 | -0.07729 | -0.5395 | -0.03112 |
| time 3.5 | 30 | 40 | 0.04448 | 0.1097 | 135 | 0.41 | 0.6859 | 0.9687 | 0.05 | -0.1725 | 0.2615 | -0.2207 | 0.3096 |
| time 4 | 20 | 30 | -0.1513 | 0.1164 | 135 | -1.30 | 0.1959 | 0.4784 | 0.05 | -0.3816 | 0.07895 | -0.4327 | 0.1301 |
| time 4 | 20 | 40 | -0.07496 | 0.1146 | 135 | -0.65 | 0.5142 | 0.8844 | 0.05 | -0.3016 | 0.1517 | -0.3519 | 0.2020 |
| time 4 | 30 | 40 | 0.07638 | 0.1179 | 135 | 0.65 | 0.5184 | 0.8873 | 0.05 | -0.1569 | 0.3096 | -0.2087 | 0.3614 |
| time 4.5 | 20 | 30 | -0.1407 | 0.1212 | 135 | -1.16 | 0.2476 | 0.5721 | 0.05 | -0.3804 | 0.09892 | -0.4335 | 0.1521 |
| time 4.5 | 20 | 40 | 0.01942 | 0.1170 | 135 | 0.17 | 0.8684 | 0.9977 | 0.05 | -0.2120 | 0.2509 | -0.2634 | 0.3022 |
| time 4.5 | 30 | 40 | 0.1601 | 0.1207 | 135 | 1.33 | 0.1869 | 0.4608 | 0.05 | -0.07863 | 0.3989 | -0.1316 | 0.4519 |
| time 5 | 20 | 30 | -0.1735 | 0.1264 | 135 | -1.37 | 0.1720 | 0.4307 | 0.05 | -0.4234 | 0.07640 | -0.4789 | 0.1319 |

| Simple Effect Comparisons of temp*time Least Squares Means By time<br>Adjustment for Multiple Comparisons: SMM | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Simple Effect Level** | **temp** | **temp** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Adj P** | **Alpha** | **Lower** | **Upper** | **Adj Lower** | **Adj Upper** |
| **time 5** | 20 | 40 | -0.01344 | 0.1215 | 135 | -0.11 | 0.9120 | 0.9993 | 0.05 | -0.2537 | 0.2268 | -0.3070 | 0.2801 |
| **time 5** | 30 | 40 | 0.1601 | 0.1267 | 135 | 1.26 | 0.2087 | 0.5027 | 0.05 | -0.09054 | 0.4107 | -0.1462 | 0.4663 |
| **time 5.5** | 20 | 30 | -0.1500 | 0.1270 | 135 | -1.18 | 0.2394 | 0.5581 | 0.05 | -0.4011 | 0.1011 | -0.4569 | 0.1568 |
| **time 5.5** | 20 | 40 | -0.00177 | 0.1224 | 135 | -0.01 | 0.9885 | 1.0000 | 0.05 | -0.2439 | 0.2403 | -0.2976 | 0.2941 |
| **time 5.5** | 30 | 40 | 0.1483 | 0.1270 | 135 | 1.17 | 0.2452 | 0.5680 | 0.05 | -0.1029 | 0.3995 | -0.1587 | 0.4552 |
| **time 6** | 20 | 30 | -0.1519 | 0.1278 | 135 | -1.19 | 0.2367 | 0.5534 | 0.05 | -0.4046 | 0.1008 | -0.4607 | 0.1569 |
| **time 6** | 20 | 40 | -0.03810 | 0.1241 | 135 | -0.31 | 0.7593 | 0.9859 | 0.05 | -0.2836 | 0.2074 | -0.3381 | 0.2619 |
| **time 6** | 30 | 40 | 0.1138 | 0.1288 | 135 | 0.88 | 0.3787 | 0.7586 | 0.05 | -0.1410 | 0.3686 | -0.1976 | 0.4251 |
| **time 6.5** | 20 | 30 | -0.1519 | 0.1278 | 135 | -1.19 | 0.2367 | 0.5534 | 0.05 | -0.4046 | 0.1008 | -0.4607 | 0.1569 |
| **time 6.5** | 20 | 40 | -0.03810 | 0.1241 | 135 | -0.31 | 0.7593 | 0.9859 | 0.05 | -0.2836 | 0.2074 | -0.3381 | 0.2619 |
| **time 6.5** | 30 | 40 | 0.1138 | 0.1288 | 135 | 0.88 | 0.3787 | 0.7586 | 0.05 | -0.1410 | 0.3686 | -0.1976 | 0.4251 |
| **time 7** | 20 | 30 | -0.1690 | 0.1283 | 135 | -1.32 | 0.1901 | 0.4669 | 0.05 | -0.4228 | 0.08478 | -0.4791 | 0.1411 |
| **time 7** | 20 | 40 | -0.07996 | 0.1253 | 135 | -0.64 | 0.5246 | 0.8916 | 0.05 | -0.3278 | 0.1679 | -0.3828 | 0.2229 |
| **time 7** | 30 | 40 | 0.08903 | 0.1306 | 135 | 0.68 | 0.4965 | 0.8713 | 0.05 | -0.1692 | 0.3472 | -0.2265 | 0.4045 |

Figure 6: stage=progress

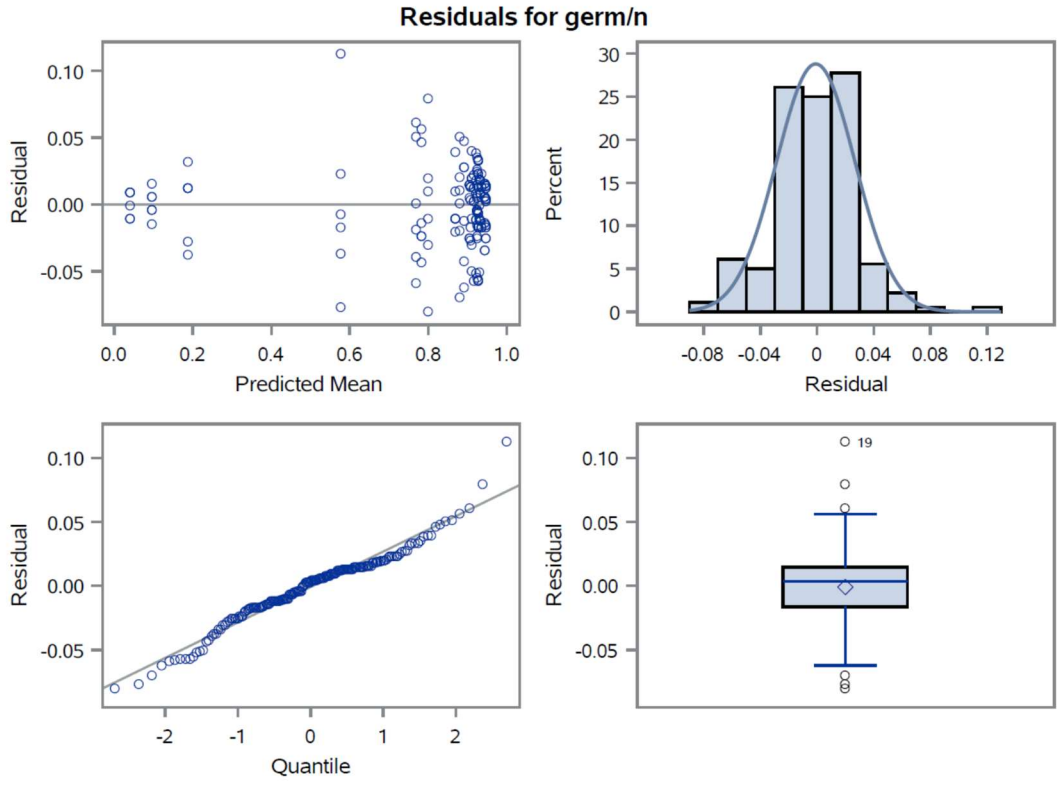
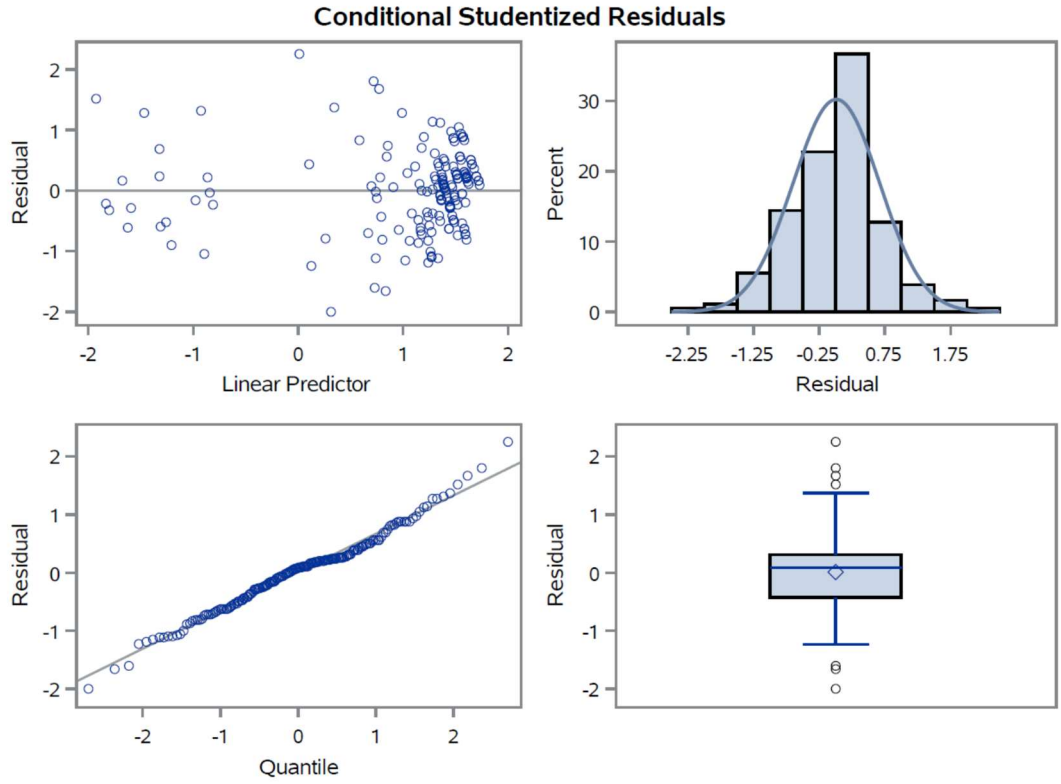


Figure 7: stage=progress




The full table of 'temp*time Least Squares Means' is not shown, but the plot of 'LS-Means for temp*time' is displayed with the estimated least-square means and their confidence interval with 95% limits on the linked scale (Figure 5). In this Figure, it may be noted that heteroscedasticity of means at day 2 (see Table 8) is modest, on the linked scale. Tests of significance for temperature effect sliced by timepoints (Table 25) show a significant effect of the tested temperature levels only for days 2-3.5, like found with one-way ANOVAs in MANOVA (Tables 1, 2 and 3). Multiple comparisons at each timepoint (Table 26) indicate a different effect for each of the three temperatures at 2 d, but only for 20 °C with respect to the other two temperatures at 3-3.5 d, like obtained with MANOVA (Figures 2-4). Residuals (Figure 6) are quite typical for binomial data with no evidence of anomalies. Platykurtosis is apparent (upper right plot in Figure 6) and linked to underdisperison (Table 21), because binomial errors of means close to the boundaries of the percentile range are skewed and are centred on the mean (if they could, hypothetically, be centred on the mode, neither platykurtosis nor underdispersion would presumably be found). Conditional studentized residuals (Figure 7) look better, as it should be. In this case, the Quantile plot (on the lower left) indicates a quite good fit of the conditional studentized residuals with the theoretical normally-distributed quantiles. In the present case, 100 seeds per plate and six plates for each mean are evidently enough to approximate a Gaussian distribution of conditional studentized residuals. These residuals are properly centred around the mean (lower right plot in Figure 7). Altogether, this conditional model appears correct, but it has some weakness, to wit, it might be too loose.

Although integral approximation is required for overdispersion diagnostics and covariance parameters testing, in the case of small-sized experiments (as common for germination tests), pseudo-likelihood is more accurate for the computation of confidence intervals and control of type I errors when using the Kenward-Roger method ('ddfm=KR2' option in the 'model' statement) for the adjustment of degrees of freedoms in *F* tests and of confidence intervals, available with pseudo-likelihood only (Stroup et al., 2018). Hence, once a conditional model has been satisfactorily identified, the integral approximation option should be removed to allow for the computation of pseudo-likelihood with the Kenward-Roger method (Stroup et al., 2018). In

addition, when *P* values are adjusted for multiplicity, the 'adjdfe=row' option in the 'lsmeans' statement ought to be specified to take into account the row-wise degrees of freedom in computing the adjusted *P* values of the contrasts as well. The model therefore becomes:

```
/*Conditional model with categorical time*/
proc GLIMMIX order=data;
by stage;
class plate temp time;
model germ/n = temp time temp*time / link=probit ddfm=KR2;
random plate(temp);
lsmeans temp*time / cl ilink plot=meanplot slice=time slicediff=time adjust=smm adjdfe=row;
run;
```

RESULTS (excerpts):

Table 27: stage=progress

| Model Information | |
|---|---|
| **Data Set** | WORK.REFFILE |
| **Response Variable (Events)** | germ |
| **Response Variable (Trials)** | n |
| **Response Distribution** | Binomial |
| **Link Function** | Probit |
| **Variance Function** | Default |
| **Variance Matrix** | Not blocked |
| **Estimation Technique** | Residual PL |
| **Degrees of Freedom Method** | Kenward-Roger2 |
| **Fixed Effects SE Adjustment** | Kenward-Roger2 |

Table 28: stage=progress

| Fit Statistics | |
|---|---|
| **-2 Res Log Pseudo-Likelihood** | -93.34 |
| **Generalized Chi-Square** | 69.47 |
| **Gener. Chi-Square / DF** | 0.46 |

Table 29: stage=progress

| Covariance Parameter Estimates | | |
|---|---|---|
| **Cov Parm** | **Estimate** | **Standard Error** |
| **plate(temp)** | 0.01286 | 0.005794 |

Table 30: stage=progress

| Type III Tests of Fixed Effects | | | | |
|---|---|---|---|---|
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **temp** | 2 | 15.53 | 5.99 | 0.0118 |
| **time** | 9 | 150 | 423.87 | <.0001 |
| **temp*time** | 18 | 150 | 5.27 | <.0001 |

Table 31: stage=progress

| Tests of Effect Slices for temp*time Sliced By time | | | | |
|---|---|---|---|---|
| **time** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| 2 | 2 | 119.3 | 23.34 | <.0001 |
| 3 | 2 | 57.71 | 21.39 | <.0001 |
| 3.5 | 2 | 80.72 | 5.45 | 0.0060 |
| 4 | 2 | 110.7 | 0.79 | 0.4561 |
| 4.5 | 2 | 122.9 | 0.97 | 0.3832 |
| 5 | 2 | 144.1 | 1.08 | 0.3436 |
| 5.5 | 2 | 147 | 0.85 | 0.4287 |
| 6 | 2 | 150 | 0.71 | 0.4932 |
| 6.5 | 2 | 150 | 0.71 | 0.4932 |
| 7 | 2 | 150 | 0.80 | 0.4532 |

Table 32: stage=progress

| Simple Effect Comparisons of temp*time Least Squares Means By time Adjustment for Multiple Comparisons: SMM | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Simple Effect Level** | **temp** | **temp** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Adj P** | **Alpha** | **Lower** | **Upper** | **Adj Lower** | **Adj Upper** |
| **time 2** | 20 | 30 | -0.4286 | 0.1333 | 150 | -3.22 | 0.0016 | 0.0048 | 0.05 | -0.6919 | -0.1652 | -0.7502 | -0.1069 |
| **time 2** | 20 | 40 | -0.8561 | 0.1274 | 141.2 | -6.72 | <.0001 | <.0001 | 0.05 | -1.1080 | -0.6042 | -1.1638 | -0.5483 |
| **time 2** | 30 | 40 | -0.4275 | 0.1132 | 88.04 | -3.78 | 0.0003 | 0.0009 | 0.05 | -0.6525 | -0.2025 | -0.7029 | -0.1522 |
| **time 3** | 20 | 30 | -0.5410 | 0.1008 | 55.32 | -5.37 | <.0001 | <.0001 | 0.05 | -0.7430 | -0.3391 | -0.7889 | -0.2931 |
| **time 3** | 20 | 40 | -0.5899 | 0.1012 | 56.27 | -5.83 | <.0001 | <.0001 | 0.05 | -0.7927 | -0.3872 | -0.8387 | -0.3411 |
| **time 3** | 30 | 40 | -0.04893 | 0.1038 | 62.27 | -0.47 | 0.6391 | 0.9520 | 0.05 | -0.2564 | 0.1586 | -0.3034 | 0.2055 |
| **time 3.5** | 20 | 30 | -0.3298 | 0.1100 | 78.45 | -3.00 | 0.0036 | 0.0108 | 0.05 | -0.5487 | -0.1108 | -0.5979 | -0.06164 |

| Simple Effect Comparisons of temp*time Least Squares Means By time<br>Adjustment for Multiple Comparisons: SMM | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Simple Effect Level** | **temp** | **temp** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Adj P** | **Alpha** | **Lower** | **Upper** | **Adj Lower** | **Adj Upper** |
| **time 3.5** | 20 | 40 | -0.2851 | 0.1092 | 76.24 | -2.61 | 0.0109 | 0.0321 | 0.05 | -0.5026 | -0.06765 | -0.5515 | -0.01875 |
| **time 3.5** | 30 | 40 | 0.04464 | 0.1136 | 89.21 | 0.39 | 0.6952 | 0.9713 | 0.05 | -0.1810 | 0.2703 | -0.2315 | 0.3208 |
| **time 4** | 20 | 30 | -0.1509 | 0.1201 | 111.3 | -1.26 | 0.2113 | 0.5073 | 0.05 | -0.3888 | 0.08697 | -0.4418 | 0.1400 |
| **time 4** | 20 | 40 | -0.07461 | 0.1183 | 104.9 | -0.63 | 0.5296 | 0.8947 | 0.05 | -0.3091 | 0.1599 | -0.3614 | 0.2122 |
| **time 4** | 30 | 40 | 0.07633 | 0.1215 | 116.9 | 0.63 | 0.5312 | 0.8960 | 0.05 | -0.1644 | 0.3170 | -0.2179 | 0.3706 |
| **time 4.5** | 20 | 30 | -0.1404 | 0.1247 | 129.3 | -1.13 | 0.2622 | 0.5965 | 0.05 | -0.3870 | 0.1063 | -0.4418 | 0.1610 |
| **time 4.5** | 20 | 40 | 0.01973 | 0.1206 | 113.5 | 0.16 | 0.8704 | 0.9978 | 0.05 | -0.2193 | 0.2587 | -0.2725 | 0.3119 |
| **time 4.5** | 30 | 40 | 0.1601 | 0.1242 | 127.7 | 1.29 | 0.1998 | 0.4858 | 0.05 | -0.08573 | 0.4059 | -0.1403 | 0.4605 |
| **time 5** | 20 | 30 | -0.1731 | 0.1297 | 150 | -1.33 | 0.1840 | 0.4552 | 0.05 | -0.4295 | 0.08320 | -0.4863 | 0.1400 |
| **time 5** | 20 | 40 | -0.01303 | 0.1250 | 130.7 | -0.10 | 0.9171 | 0.9994 | 0.05 | -0.2603 | 0.2342 | -0.3152 | 0.2891 |
| **time 5** | 30 | 40 | 0.1601 | 0.1301 | 150 | 1.23 | 0.2203 | 0.5243 | 0.05 | -0.09690 | 0.4171 | -0.1538 | 0.4741 |
| **time 5.5** | 20 | 30 | -0.1496 | 0.1303 | 150 | -1.15 | 0.2526 | 0.5809 | 0.05 | -0.4071 | 0.1078 | -0.4642 | 0.1649 |
| **time 5.5** | 20 | 40 | -0.00131 | 0.1259 | 134.5 | -0.01 | 0.9917 | 1.0000 | 0.05 | -0.2503 | 0.2476 | -0.3055 | 0.3029 |
| **time 5.5** | 30 | 40 | 0.1483 | 0.1304 | 150 | 1.14 | 0.2570 | 0.5882 | 0.05 | -0.1093 | 0.4059 | -0.1663 | 0.4630 |
| **time 6** | 20 | 30 | -0.1514 | 0.1311 | 150 | -1.15 | 0.2500 | 0.5764 | 0.05 | -0.4105 | 0.1077 | -0.4679 | 0.1651 |
| **time 6** | 20 | 40 | -0.03752 | 0.1275 | 141.7 | -0.29 | 0.7690 | 0.9876 | 0.05 | -0.2896 | 0.2146 | -0.3456 | 0.2705 |
| **time 6** | 30 | 40 | 0.1139 | 0.1321 | 150 | 0.86 | 0.3901 | 0.7718 | 0.05 | -0.1472 | 0.3750 | -0.2051 | 0.4328 |
| **time 6.5** | 20 | 30 | -0.1514 | 0.1311 | 150 | -1.15 | 0.2500 | 0.5764 | 0.05 | -0.4105 | 0.1077 | -0.4679 | 0.1651 |
| **time 6.5** | 20 | 40 | -0.03752 | 0.1275 | 141.7 | -0.29 | 0.7690 | 0.9876 | 0.05 | -0.2896 | 0.2146 | -0.3456 | 0.2705 |
| **time 6.5** | 30 | 40 | 0.1139 | 0.1321 | 150 | 0.86 | 0.3901 | 0.7718 | 0.05 | -0.1472 | 0.3750 | -0.2051 | 0.4328 |
| **time 7** | 20 | 30 | -0.1660 | 0.1316 | 150 | -1.26 | 0.2093 | 0.5040 | 0.05 | -0.4260 | 0.09411 | -0.4837 | 0.1517 |
| **time 7** | 20 | 40 | -0.07732 | 0.1287 | 147 | -0.60 | 0.5489 | 0.9075 | 0.05 | -0.3317 | 0.1770 | -0.3881 | 0.2334 |
| **time 7** | 30 | 40 | 0.08864 | 0.1338 | 150 | 0.66 | 0.5086 | 0.8805 | 0.05 | -0.1757 | 0.3530 | -0.2343 | 0.4115 |

Residual PL is now specified as estimation technique, with Kenward-Roger2 method for degrees of freedom and fixed effects SE adjustments (Table 27). The fit statistics are changed to PL-specific ones, with the Generalized Chi-Square / DF providing a rough, but, in this case, not so bad (compared to the proper estimate in Table 21), evaluation of overdispersion (Table 28). The estimates of the random deviance of the intercept due to plates and its standard error have slightly increased (Table 29; compare to Table 22) and should represent better estimates of this effect, likewise to overall *F* test probabilities (Table 30), timepoint *F* tests probabilities (Table 31) and multiple comparisons (Table 32). Inference has not changed, anyway. If the experimental design is simple, and all the effects have been properly modelled, the integral approximation with true likelihood is just a check.

It would be, at this point, also interesting to evaluate a model with continuous time, since, if properly modelled, a continuous variable contains more information than a classification one, and, if modelling the

time-course spares the number of parameters in optimization, this approach could require less degrees of freedom and therefore increase the power of the analysis. If 'time' is simply removed from the 'classification' statement (and the 'lsmeans' statement is removed because only fixed effects that are 'class' variables, or involve only 'class' variables, can, presently, be computed by this statement), convergence is faster (and the 'order=data' specification can be removed too, because time is now considered as a numerical variable and the ordering of timepoints is consequential), but same troubles promptly arise.

```
proc GLIMMIX method=Laplace;
by stage;
class plate temp;
model germ/n = temp time temp*time / link=probit;
random plate(temp);
run;
```

RESULTS (excerpts):

Table 33: stage=progress

| Dimensions | |
|---|---|
| **G-side Cov. Parameters** | 1 |
| **Columns in X** | 8 |
| **Columns in Z** | 18 |
| **Subjects (Blocks in V)** | 1 |
| **Max Obs per Subject** | 180 |

Table 34: stage=progress

| Optimization Information | |
|---|---|
| **Optimization Technique** | Dual Quasi-Newton |
| **Parameters in Optimization** | 7 |
| **Lower Boundaries** | 1 |
| **Upper Boundaries** | 0 |
| **Fixed Effects** | Not Profiled |
| **Starting From** | GLM estimates |

Table 35: stage=progress

| Fit Statistics | |
|---|---|
| **-2 Log Likelihood** | 2637.55 |
| **AIC (smaller is better)** | 2651.55 |
| **AICC (smaller is better)** | 2652.20 |
| **BIC (smaller is better)** | 2657.78 |

| Fit Statistics | |
|---|---|
| CAIC (smaller is better) | 2664.78 |
| HQIC (smaller is better) | 2652.40 |

Table 36: stage=progress

| Fit Statistics for Conditional Distribution | |
|---|---|
| -2 log L(germ \| r. effects) | 2599.60 |
| Pearson Chi-Square | 1975.37 |
| Pearson Chi-Square / DF | 10.97 |

Table 37: stage=progress

| Type III Tests of Fixed Effects | | | | |
|---|---|---|---|---|
| Effect | Num DF | Den DF | F Value | Pr > F |
| temp | 2 | 15 | 17.60 | 0.0001 |
| time | 1 | 159 | 3147.58 | <.0001 |
| time*temp | 2 | 159 | 14.50 | <.0001 |

The number of column in the X matrix (Table 33) has decreased to 8 (from 44, in Table 19) because the number of parameters in optimization (Table 34) was reduced to 7 (from 31). In fact, in linear models, a continuous variable is modelled, by default, as being linear, and a linear regression requires only two parameters, an intercept and a slope, rather than *L*-1 dummies for *L* levels of the continuous factor. Thus, the seven parameters correspond to one random variance component (which has a lower bound of 0) plus six fixed effect parameters (as can be checked by adding the 'solution' option into the 'model' statement). The latter are: the model intercept, two dummies for temperature, a slope for time (the overall model intercept is already present and therefore another intercept would be redundant), and two temperature dummies for their interaction, since only the slope parameter is modelled for time. This saves a lot of calculations and spares several degrees of freedom, but it comes at a big price: all the fit statistics burst higher (Table 35), evidencing a severe loss of fit in the model. In fact, the overdispersion parameter skyrocketed (Table 36). Nonetheless, the significance of effects remained high and even increased for 'temp' (Table 37). What has happened is that the smaller number of parameters has increased the power of significance tests, but the germination response to time was assumed to be linear, whereas it is not. The time effect is indisputable, whether it is considered as linear or not, so it is still highly significant, but the residuals were largely increased (higher overdispersion) because the actual curvilinear response is approximated with a straight line. What's going on with this model can be clarified by suitable graphs.

To report the actual data on the probit (linked) scale, a new column of probit-transformed data is introduced with a 'DATA' statement, and the model is run again adding an 'output' statement that produces an output file that automatically includes all variables already present in the original data set and also contains the

predicted values of the BLUPs on the scale of the link function (the linearized scale; here, the probit scale). A first plot is requested to show how the predicted values fit a two-step linear model, wherein the separate linear fits, as obtained according to the 'BY' statement, are evident for the 'lag' and 'progress' stages on the linked scale. A second plot then compares the same two-steps linear fitting to the original data previously reported on the linked scale.

```
DATA reffile;
set reffile;
p=probit(germ/n);
run;

proc GLIMMIX method=Laplace;
by stage;
class plate temp;
model germ/n = temp time temp*time / link=probit;
random plate(temp);
output out=gmxout pred=pred;
run;

proc SGPLOT data=gmxout;
loess y=pred x=time / group=temp nomarkers name="fit";
scatter y=pred x=time / group=temp;
keylegend "fit" / title="Temp";
run;

proc SGPLOT data=gmxout;
loess y=pred x=time / group=temp nomarkers name="fit";
scatter y=p x=time / group=temp;
keylegend "fit" / title="Temp";
run;
```

RESULTS (excerpts):

Figure 8

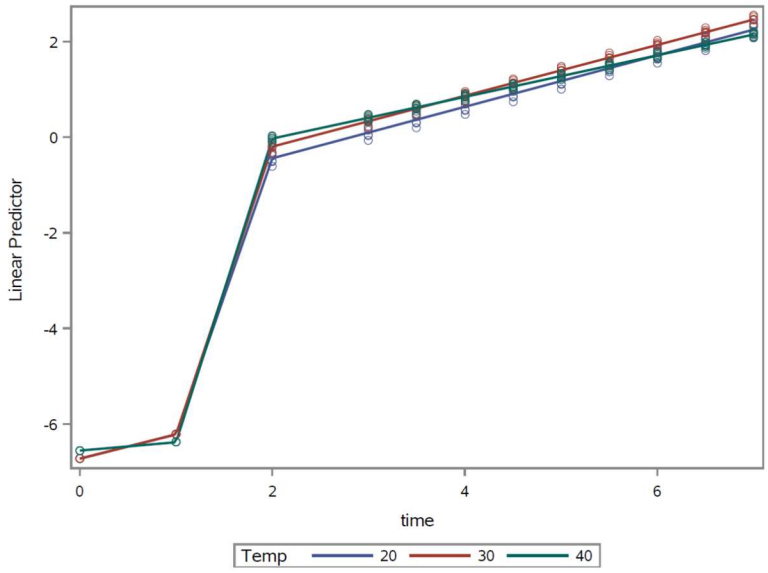
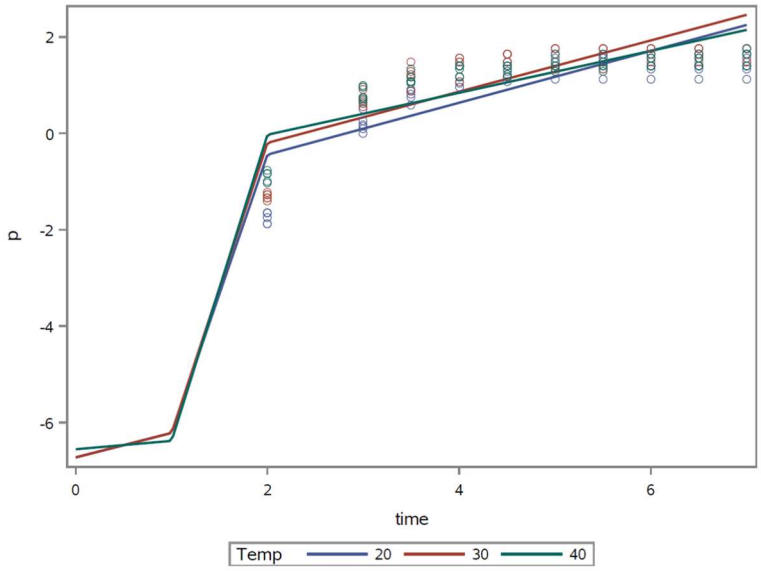


Figure 9




It is immediately evident that the predicted values have been modelled to adapt a linear trend (separately for the 'lag' and the 'progress' stages) on the linear scale (Figure 8), and that such two-steps linear trend does not really fit the actual data (Figure 9). The 'progress' data result thus widely overdispersed with respect to the linear fitting. Nevertheless, the effects of the modelled factors are even more clear in the forcedly linearized fitting, which explains the improved significances of the 'temp' factor for the 'progress' stage. As previously remarked, the germination progress is not linear through time, even on the linked scale. Also note that the model has fitted the germination responses of the 'lag' stage to positive predicted values, albeit very close to zero (thus that they do not result significantly different from zero; not shown), whereas germination values of zero cannot be represented on a probit scale.

To formally assess the real fitting of the model with time assumed as a linear covariate on the linked scale, time can be introduced, in the same model, as both a continuous and a categorical variable (using different names), and any additional significant effect of the categorical time variable over the assumedly linear continuous time variable can be established using a type I test of fixed effects (chosen with the 'htype' option in the 'model' statement), which performs sequential tests of the effects included in the model (Littell et al., 2006). If categorical time still has a significant effect after linear continuous time has been considered, then linear fitting does not explain the actual trend of the response through time.

```
DATA reffile;
set reffile;
xtime=time;
run;

proc GLIMMIX method=Laplace;
by stage;
class plate temp xtime;
model germ/n = temp time temp*time xtime/ link=probit htype=1;
random plate(temp);
run;
```

RESULTS (excerpt):

Table 38: stage=progress

| Type I Tests of Fixed Effects | | | | |
|---|---|---|---|---|
| Effect | Num DF | Den DF | F Value | Pr > F |
| temp | 2 | 15 | 7.73 | 0.0049 |
| time | 1 | 151 | 2390.51 | <.0001 |
| time*temp | 2 | 151 | 28.93 | <.0001 |
| xtime | 8 | 151 | 195.98 | <.0001 |

Type I test of fixed effects (Table 38) shows a highly significant effect of 'xtime' (categorical time) after 'time' (continuous time, modelled as linear on the linked scale) has already been considered, formally demonstrating, if Figure 9 was not enough, that germination progress is not linear to time, even on the linked scale. A model based on the assumption of linearity of germination progress through time (on the probit scale) has, therefore, to be rejected.

Modelling time as a nonlinear covariate requires a function that, either because of theoretical considerations or a good empirical fit, quantitatively describes the germination time-course. For example, this latter can be empirically fitted with a spline, that is, a continuous and smooth curvilinear interpolation based on a piecewise polynomial. To this aim, a constructed effect, 'spltime' is specified to model the germination time-course. As the shape of the germination time-course is typically different among treatments, this diversity is modelled in terms of the interaction between treatment and spline, i.e. 'temp*spltime' in the present instance. This interaction models separate curvilinear trends for the different temperature levels, whereas the 'temp' main effect instructs to model a separate intercept for each temperature. An overall 'spltime' main effect is therefore not needed (likewise to the 'time' main effect that was not needed in the previous model that assumed germination progress linear to time on the linked scale). The plots previously described are requested too.

```
DATA reffile;
set reffile;
p=probit(germ/n);
run;

proc GLIMMIX method=Laplace;
by stage;
effect spltime = spline(time);
class plate temp;
model germ/n = temp temp*spltime / link=probit;
random plate(temp);
output out=gmxout pred=pred;
run;

proc SGPLOT data=gmxout;
loess y=pred x=time / group=temp nomarkers name="fit";
scatter y=pred x=time / group=temp;
keylegend "fit" / title="Temp";
run;

proc SGPLOT data=gmxout;
```

```
loess y=pred x=time / group=temp nomarkers name="fit";
scatter y=p x=time / group=temp;
keylegend "fit" / title="Temp";
run;
```

RESULTS (excerpts):

Table 39: stage=progress

| Model Information | |
|---|---|
| **Data Set** | WORK.REFFILE |
| **Response Variable (Events)** | germ |
| **Response Variable (Trials)** | n |
| **Response Distribution** | Binomial |
| **Link Function** | Probit |
| **Variance Function** | Default |
| **Variance Matrix** | Not blocked |
| **Estimation Technique** | Maximum Likelihood |
| **Likelihood Approximation** | Laplace |
| **Degrees of Freedom Method** | Residual |

Table 40: stage=progress

| Optimization Information | |
|---|---|
| **Optimization Technique** | Dual Quasi-Newton |
| **Parameters in Optimization** | 22 |
| **Lower Boundaries** | 1 |
| **Upper Boundaries** | 0 |
| **Fixed Effects** | Not Profiled |
| **Starting From** | GLM estimates |

Table 41: stage=progress

| Fit Statistics | |
|---|---|
| **-2 Log Likelihood** | 804.38 |
| **AIC (smaller is better)** | 848.38 |
| **AICC (smaller is better)** | 854.82 |
| **BIC (smaller is better)** | 867.96 |
| **CAIC (smaller is better)** | 889.96 |
| **HQIC (smaller is better)** | 851.08 |

Table 42: stage=progress

| Fit Statistics for Conditional Distribution | |
|---|---|
| -2 log L(germ \| r. effects) | 764.07 |
| Pearson Chi-Square | 59.43 |
| Pearson Chi-Square / DF | 0.33 |

Table 43: stage=progress

| Covariance Parameter Estimates | | |
|---|---|---|
| Cov Parm | Estimate | Standard Error |
| plate(temp) | 0.01024 | 0.004415 |

Table 44: stage=progress

| Type III Tests of Fixed Effects | | | | |
|---|---|---|---|---|
| Effect | Num DF | Den DF | F Value | Pr > F |
| temp | 2 | 159 | 0.17 | 0.8449 |
| spltime*temp | 18 | 159 | 221.77 | <.0001 |

Figure 10.

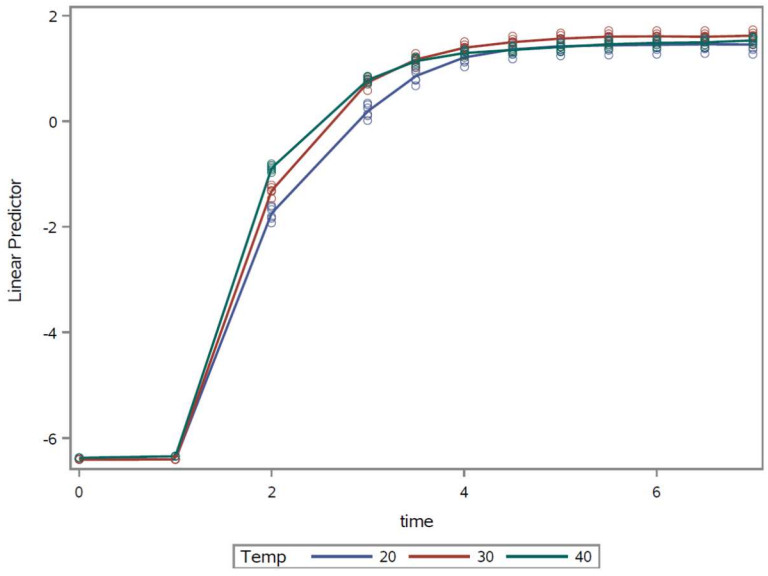
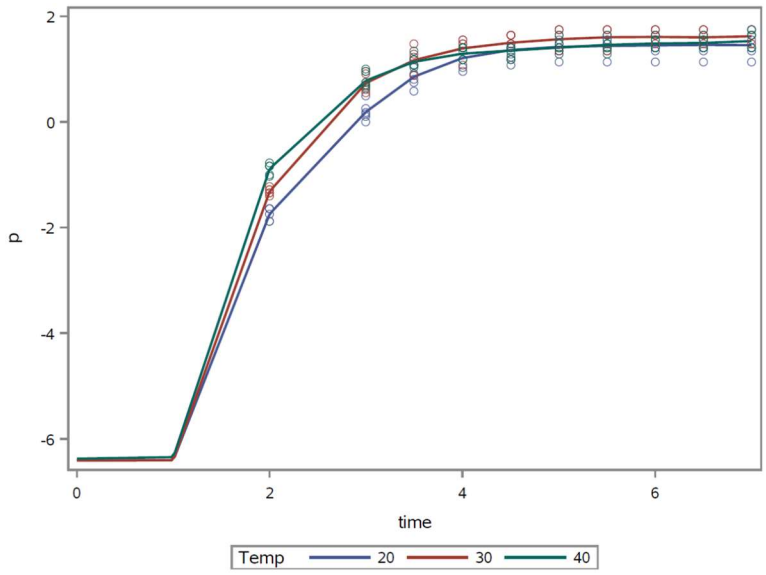


Figure 11.




The spline function allows to fit nonlinear time dependencies in an ANCOVA model. Table 39 shows that, whereas previous models used the 'Containment' Degrees of Freedom Method, the specification of a constructed effect, when an integral approximation is requested, causes the method to shift to 'Residuals'. Twenty-two parameters were optimized in this model (Table 40): one random variance component (which has a lower bound of 0), the model intercept, two dummies for temperature, and six spline parameters for each of the three temperature levels. Note that, by default, the GLIMMIX spline is a cubic B-spline with three equally spaced knots that yields seven design matrix columns, corresponding to the degree of the spline (three, for a cubic spline) plus the number of nodes, plus one (the intercept of the spline). As the temperature levels are modelled themselves as intercepts, the number of parameters is reduced accordingly. Table 41 shows that all the fit statistics have improved (i.e., decreased) for this model with respect to the model with categorical time (Table 20) but for -2 Log Likelihood, which is not penalized for the number of required parameters (31 in that case). The overdispersion parameter is estimated as 0.33 (Table 42), and the variance among plates (Table 43) is almost unchanged with respect to the model with categorical time (Table 20). Notably, the Tests of Fixed Effects (Table 44) indicate a non-significant effect for 'temp' and a highly significant effect for 'temp*spltime', which means that what the temperature causes in this experiment is not a vertical shift of the overall curvilinear trend, but only a change of its shape. This is not in contrast with previous findings of a significant 'temp' effect, because when time is modelled as a categorical variable, the ANOVA table considers a test of significance of the differences between mean response values of the 'temp' levels (even if modelled in terms of deviations from the intercept of the linear response), which are therefore averaged over the discrete time levels (on the linked scale), whereas when time is modelled as a continuous variable, the significance test compares mean response values for the 'temp' intercepts, that is, for values extrapolated at time=0.

The comparison of the fitting of the model to the linear predictors (BLUPs; Figure 10) with its fitting to the actual data (Figure 11), on the linear scale, provides three important hints: (a)- the fit of the actual data to the model is very good and, in accordance, the arrangement of the BLUPs is very close to the arrangement of the actual data; (b)- the lag-stage responses are still modelled as extremely low, but not zero, values; (c)- shrinkage of the BLUPs with respect to the observed data is evident. Shrinkage of BLUPs consists in computing estimates of the response variable for a given level of a random factor (i.e., predictors) that are closer to the between-subjects mean than the observed values. This occurs because, in the estimation, the difference between the observed value and the mean is shrunk (reduced) by an amount that depends on the estimated

non-systematic error variance components, such as the within-subject error variance across timepoints. Removing the erratic variability of the random effect is intended to provide a better evaluation of the intrinsic, systematic effect of each subject. In fact, extreme means are attenuated according to the underlying variability, thereby reducing the risk of misinterpreting the data because of noisiness and casual outliers (Stroup et al., 2018). BLUPs shrinkage explains the very small variance estimated for the plate effect, sharply lower than the error variance observed at each timepoint for the original data. It ought to be noted that more severe shrinkage is expected when the residual variability is large in comparison to the between-subjects variability, and whereas the former is essentially overlapping and confused with the latter in FGP data, it can be distinctly estimated in longitudinal experiments. Thus, smaller between-plates variance and, therefore, greater shrinkage, are to be expected in longitudinal studies than in FGP tests.

This model can therefore be considered satisfactory, and significance of the differences between responses to temperature at the different timepoints can be assessed. As mentioned above, LS-means cannot be automatically computed at the original observation timepoints because, once time has been modelled as a continuous variable based on those original timepoints, those same timepoints are then considered as discrete instances of a function whose levels are continuous, and the number of time levels at which to perform the comparisons is therefore no longer considered to be predefined. In the presence of a continuous variable, LS-means for classification variables are automatically estimated assuming an average value of the continuous variable. Nonetheless, they can be estimated for any other value of choice of the continuous time variable, including the original timepoints. As the continuous time is modelled according to an empirical spline function, I recommend not estimating LS-means at any other timepoints but the original ones. Of course, estimations must not be performed beyond the range of the originally observed timepoints: the chosen 'time=' in the 'lsmeans' statement must be within the modelled range, therefore also excluding the lag stage (obviously).

```
/*Conditional model with continuous time and integral approximation*/
proc GLIMMIX method=Laplace;
by stage;
effect spltime = spline(time);
class plate temp;
model germ/n = temp temp*spltime / link=probit;
random plate(temp);
lsmeans temp / at time=2 ilink adjust=smm;
lsmeans temp / at time=3 ilink adjust=smm;
lsmeans temp / at time=3.5 ilink adjust=smm;
lsmeans temp / at time=4 ilink adjust=smm;
lsmeans temp / at time=4.5 ilink adjust=smm;
lsmeans temp / at time=5 ilink adjust=smm;
lsmeans temp / at time=5.5 ilink adjust=smm;
lsmeans temp / at time=6 ilink adjust=smm;
lsmeans temp / at time=6.5 ilink adjust=smm;
lsmeans temp / at time=7 ilink adjust=smm;
run;
```

RESULTS (excerpts):

Table 45: stage=progress

| Differences of temp Least Squares Means Adjustment for Multiple Comparisons: SMM | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **temp** | **temp** | **time** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Adj P** |
| 20 | 30 | 2.00 | -0.4275 | 0.1299 | 159 | -3.29 | 0.0012 | 0.0037 |
| 20 | 40 | 2.00 | -0.8558 | 0.1239 | 159 | -6.91 | <.0001 | <.0001 |
| 30 | 40 | 2.00 | -0.4283 | 0.1093 | 159 | -3.92 | 0.0001 | 0.0004 |

Table 46: stage=progress

| Differences of temp Least Squares Means Adjustment for Multiple Comparisons: SMM | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **temp** | **temp** | **time** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Adj P** |
| 20 | 30 | 3.00 | -0.5457 | 0.09598 | 159 | -5.69 | <.0001 | <.0001 |
| 20 | 40 | 3.00 | -0.5919 | 0.09638 | 159 | -6.14 | <.0001 | <.0001 |
| 30 | 40 | 3.00 | -0.04622 | 0.09912 | 159 | -0.47 | 0.6416 | 0.9536 |

Table 47: stage=progress

| Differences of temp Least Squares Means Adjustment for Multiple Comparisons: SMM | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **temp** | **temp** | **time** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Adj P** |
| 20 | 30 | 3.50 | -0.3114 | 0.09710 | 159 | -3.21 | 0.0016 | 0.0049 |
| 20 | 40 | 3.50 | -0.2819 | 0.09621 | 159 | -2.93 | 0.0039 | 0.0116 |
| 30 | 40 | 3.50 | 0.02951 | 0.09968 | 159 | 0.30 | 0.7676 | 0.9873 |

Separately processing data according to the stage causes a large number of tables to be produced for the lag stage, wherein the analysis of germination data is borderline, if even possible, and progress of visible germination is already known not to take place. These tables are mostly useless, and removing these data from the dataset used for the analysis would simplify the output. With the present analysis settings, in fact, several LS-means tables are non-sense extrapolations of germination progress from the lag stage. Nevertheless, they help reminding that, though these data are not part of the visible germination progress, they are indeed data observed from the germination process, and must be recorded to ascertain the actual duration of the lag stage and the time at which germination starts. Apart from this caveat, Tables 45-47 show the only timepoints at which significant differences were found among temperatures. It appears that the inference accomplished with this model is the same as obtained with the MANOVA model (Figures 2-4) and the conditional model with categorical time (Table 26). It can be noted that, in the present model, the adjustment of probability for multiple tests is performed within each discrete timepoint, not overall. This reduces the risk that small stochastic changes in the germination progress curve be identified as significant

without diminishing the power of the test: whereas an overall (across all discrete timepoints) adjustment for multiple contrasts of longitudinal data may be sensible in experiments wherein specific timepoints are selected and modelled as categorical levels, in longitudinal germination studies an adjustment of probability across the whole set of timepoints is probably unnecessary when the very same replicate plates are observed through time and germination is modelled as a continuous function of time (in terms of a spline and/or a linear function, or either a non-linear function). In fact, on the one side, replication, non-independence and, overall, using a fitted function rather than the discrete original data in the tests, strongly reduce the risk of stochastic changes in the germination progress curve, and therefore an adjustment of probability across timepoints is much less of concern. On the other side, the power of the analysis could be treacherously diminished with an overall adjustment of probability: the means difference at a given timepoint could be declared significant or not depending on the number of timepoints. As the duration of the experiment and the frequency of observations can be chosen with quite broad freedom by the researcher, the analysis would not be consistent if this choice would determine the inference at a given timepoint. In this sense, it could be noted that the significance of differences at times 2-3.5 would decline (that is, adjusted probability of sameness would increase) if the experiment were further prolonged. In actuality, some timepoints are already redundant, as germination appear to have plateaued after about 5 days. So, in this model, an overall adjustment od probability is not advised. Adjustments are, indeed, recommended (Stroup et al., 2018) for unstructured comparisons (i.e., when several equipollent discrete levels are compared across a single effect), but not for structured comparisons (that is, if the contrasted levels are not equipollent). When time is modelled as a continuous variable, i.e. with a vectorial structure, adjusting probability for the number of timepoints is questionable. If required, anyway, contrasts of multiple means across a continuous variable (that is, across different timepoints) can be performed with an 'estimate' statement to obtain a family-wise adjustment of probabilities, but this is quite cumbersome.

It might be noticed that even temperature could be modelled as continuous, but, given it only has three levels, there is no advantage in modelling it as continuous rather than categorical. Usage of a spline to spare parameters for dummy variables can be convenient when the number of levels (*L*) of the variable is greater than seven: the spline requires six parameters in addition to the general intercept, whereas the number of dummies is *L*-1, thus that at least eight levels are needed to make continuous modelling a convenient choice.

As suggested (Stroup et al., 2018), the Laplace integral approximation is removed in the definitive model to consent usage of pseudo-likelihood with Kenward-Roger adjustments, which may be preferred once the best model has been selected.

```
/*Conditional model with continuous time*/
proc GLIMMIX;
by stage;
effect spltime = spline(time);
class plate temp;
model germ/n = temp temp*spltime / link=probit ddfm=KR2;
random plate(temp);
lsmeans temp / at time=2 ilink adjust=smm adjdfe=row;
lsmeans temp / at time=3 ilink adjust=smm adjdfe=row;
lsmeans temp / at time=3.5 ilink adjust=smm adjdfe=row;
lsmeans temp / at time=4 ilink adjust=smm adjdfe=row;
lsmeans temp / at time=4.5 ilink adjust=smm adjdfe=row;
lsmeans temp / at time=5 ilink adjust=smm adjdfe=row;
lsmeans temp / at time=5.5 ilink adjust=smm adjdfe=row;
lsmeans temp / at time=6 ilink adjust=smm adjdfe=row;
```

```
lsmeans temp / at time=6.5 ilink adjust=smm adjdfe=row;
lsmeans temp / at time=7 ilink adjust=smm adjdfe=row;
run;
```

RESULTS (excerpts):

Table 48: stage=progress

| Fit Statistics | |
|---|---|
| **-2 Res Log Pseudo-Likelihood** | -164.99 |
| **Generalized Chi-Square** | 70.37 |
| **Gener. Chi-Square / DF** | 0.44 |

Table 49: stage=progress

| Covariance Parameter Estimates | | |
|---|---|---|
| **Cov Parm** | **Estimate** | **Standard Error** |
| **plate(temp)** | 0.01285 | 0.005792 |

Table 50: stage=progress

| Type III Tests of Fixed Effects | | | | |
|---|---|---|---|---|
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **temp** | 2 | 159 | 0.17 | 0.8475 |
| **spltime*temp** | 18 | 159 | 224.73 | <.0001 |

Table 51: stage=progress

| Differences of temp Least Squares Means<br>Adjustment for Multiple Comparisons: SMM | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **temp** | **temp** | **time** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Adj P** |
| 20 | 30 | 2.00 | -0.4285 | 0.1333 | 159 | -3.22 | 0.0016 | 0.0047 |
| 20 | 40 | 2.00 | -0.8560 | 0.1274 | 141.2 | -6.72 | <.0001 | <.0001 |
| 30 | 40 | 2.00 | -0.4275 | 0.1132 | 88.09 | -3.78 | 0.0003 | 0.0009 |

Table 52: stage=progress

| Differences of temp Least Squares Means Adjustment for Multiple Comparisons: SMM | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **temp** | **temp** | **time** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Adj P** |
| 20 | 30 | 3.00 | -0.5448 | 0.1004 | 54.62 | -5.42 | <.0001 | <.0001 |
| 20 | 40 | 3.00 | -0.5910 | 0.1008 | 55.44 | -5.86 | <.0001 | <.0001 |
| 30 | 40 | 3.00 | -0.04622 | 0.1034 | 61.37 | -0.45 | 0.6565 | 0.9586 |

Table 53: stage=progress

| Differences of temp Least Squares Means Adjustment for Multiple Comparisons: SMM | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **temp** | **temp** | **time** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Adj P** |
| 20 | 30 | 3.50 | -0.3107 | 0.1015 | 56.83 | -3.06 | 0.0034 | 0.0100 |
| 20 | 40 | 3.50 | -0.2802 | 0.1006 | 55.03 | -2.78 | 0.0073 | 0.0218 |
| 30 | 40 | 3.50 | 0.03049 | 0.1039 | 62.57 | 0.29 | 0.7702 | 0.9876 |

Amid the PL-specific fit statistics, the Generalized Chi-Square / DF again provides a rough, but, in this case, not so bad (compared to the proper estimate in Table 42), evaluation of overdispersion (Table 48). The estimates of the random deviance of the intercept due to plates and its standard error have slightly increased (Table 49; compare to Table 43) and should represent better estimates of this effect, likewise to overall *F* tests probabilities (Table 50) and multiple comparisons (Tables 51-53). Inference did not change with respect to using the integral approximation with true likelihood.

Although both conditional models, with time either as a categorical or continuous variable, are satisfactory, in longitudinal studies, within-subject covariance parameters are more properly dealt with by modelling fixed effects errors by means of the R-side covariance matrix (Littell et al., 2006). Indeed, in generalized linear (mixed) models, modelling of longitudinal correlation can be achieved either indirectly, with a G-side random factor that considers the sharing of the same random effect level for responses from the same subject (i.e., plate), or directly, with an R-side covariance structure (Littell et al., 2006). These two modelling approaches can lead to different inferences because the random block effects are modelled on the linked scale whereas the effects of residuals are modelled on the data scale, unless data are analysed using a linearized pseudo variable, as in pseudo-likelihood (Littell et al., 2006). The two approaches tend therefore to provide closer inferences when pseudo-likelihood is used. Nevertheless, whereas the G matrix is typically used to model the variance of the between-subjects factors, for germination data, the covariance structure of the repeated measures, or within-subjects, model is more easily modelled with the R matrix.

Statistical analysis of germination data is typically quasi-marginal rather than conditional (that is, modelling of data correlation ought to be processed in terms of R matrix, not G matrix), because inference about the seed population, not the average plate, is of interest. Although an integral approximation, like the Laplace method, is required for overdispersion diagnostics and covariance parameters testing, such methods require

conditional independence without R-side overdispersion or covariance structure. In addition, the integral approximation prevents the usage of specific degrees of freedom adjustments that are important to compute unbiased significance of effects as well as confidence intervals (Stroup et al., 2018). These opposing requirements suggest that a conditional model with integral approximation should be initially tested to ascertain eventual overdispersion and the significance of random effects, but the marginal model that best reflects the experimental design ought to be ultimately used (Stroup et al., 2018). The model has therefore to be restructured in terms of R-side modelling. Thus, however, the Laplace integral approximation cannot be maintained.

To take into account the serial correlation (non-independency) among repeated measurements on the same plate, a 'random residual' statement is introduced to model the R-side correlation structure, in place of the previous 'random' statement. This, therefore, is a marginal model that models covariation (correlation, indeed, is just a standardized covariance) between residuals. If a proper subject factor is specified ('plate(temp)', in this case), the 'residual' keyword, in fact, elicits the modelling of the R matrix, that is, the matrix of the variance/covariance structure of the residuals (of the fixed effects). In a 'random residual' statement, the 'subject' option identifies the factor on which repeated measures are taken.

Some further changes are however needed to accomplish with the requirements of R-side error modelling. As some covariance structures assume correlation among contiguous data, the order of the data becomes relevant even when time is modelled as a categorical variable. By default, the SAS software sorts the levels of classification variables according to an alphanumeric order that, as shown in Table 10, is not the same as if the variable were numerically ordered. If the dataset already is ordered according to increasing time, then adding the 'order=data' option will do the job (otherwise, a more specific syntax can be used to make time an explicit ordering variable within a 'random' statement that anyway includes a 'residual' keyword, but this alternative solution is not described here).

As previously mentioned, the default covariance structure is 'variance components', which only considers a Gaussian error variance for each random effect subject, that is, for 'plate(temp)' in this case. A better covariance structure could then be considered, starting from a very general one, with the lower constraints and therefore the larger number of parameters, because the significance of R-side effects can be better evaluated by stepwise restricting the covariance structure (Littell et al., 2006). Thus, an unstructured covariance matrix can be introduced in the 'random residual' statement to ascertain whether the model can be improved. A completely general (unstructured) covariance matrix, parameterized in terms of variances and correlations, can be used to assess the widest effect of an R-side covariance structure.

Furthermore, it needs to be noticed that numerical convergence during the optimization of a model can sometimes be more easily achieved if the optimization technique makes use of a Hessian matrix, like the Newton-Raphson ('newrap') or Newton-Raphson with ridging ('nrridg') optimization techniques, which are more reliable than other techniques, i.e. the convergence criterion is satisfied with higher probability, but can require a longer processing time. The default optimization technique (in the presence of random factors), the Dual Quasi-Newton, is faster, but sometimes does not ensure convergence; whereas the 'nrridg' is better for binary data fitted with a pseudo-likelihood estimation method. The optimization technique can become relevant during the optimization of a repeated measures model because germination progress is not linear through time, even on the linked scale, and the covariance structure of the residuals, i.e. the parametrization of the R matrix, is nonlinear as well. Thus, model fitting may require a more reliable nonlinear optimization method, selected by means of the 'nloptions' statement (which may be generally advantageous for improving optimization convergence in marginal and quasi-marginal models for germination progress).

For completeness, time is first modelled as a categorical variable. All this considered, we can apply the following model:

```
proc GLIMMIX order=data;
by stage;
class plate temp time;
model germ/n = temp time temp*time / link=probit;
random residual / subject=plate(temp) type=unr;
nloptions tech=nrridg;
run;
```

This model does not converge to a solution for the stage=lag means, but it does for the germination progress stage.

RESULTS (excerpts):

Table 54: stage=progress

| Model Information | |
|---|---|
| **Data Set** | WORK.REFFILE |
| **Response Variable (Events)** | germ |
| **Response Variable (Trials)** | n |
| **Response Distribution** | Binomial |
| **Link Function** | Probit |
| **Variance Function** | Default |
| **Variance Matrix Blocked By** | plate(temp) |
| **Estimation Technique** | Residual PL |
| **Degrees of Freedom Method** | Between-Within |

Table 55: stage=progress

| Class Level Information | | |
|---|---|---|
| **Class** | **Levels** | **Values** |
| **plate** | 6 | 1 2 3 4 5 6 |
| **temp** | 3 | 20 30 40 |
| **time** | 10 | 2 3 3.5 4 4.5 5 5.5 6 6.5 7 |

Table 56: stage=progress

| Dimensions | |
|---|---|
| **R-side Cov. Parameters** | 55 |
| **Columns in X** | 44 |
| **Columns in Z per Subject** | 0 |
| **Subjects (Blocks in V)** | 18 |
| **Max Obs per Subject** | 10 |

Table 57: stage=progress

| Optimization Information | |
|---|---|
| **Optimization Technique** | Newton-Raphson with Ridging |
| **Parameters in Optimization** | 55 |
| **Lower Boundaries** | 55 |
| **Upper Boundaries** | 45 |
| **Fixed Effects** | Profiled |
| **Starting From** | Data |

Table 58: stage=progress

| Fit Statistics | |
|---|---|
| **-2 Res Log Pseudo-Likelihood** | -228.04 |
| **Generalized Chi-Square** | 150.00 |
| **Gener. Chi-Square / DF** | 1.00 |

Table 59: stage=progress

| Covariance Parameter Estimates | | | |
|---|---|---|---|
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** |
| **Var(1)** | plate(temp) | 0.2840 | 0.1037 |
| **Var(2)** | plate(temp) | 1.3833 | 0.5104 |
| **Var(3)** | plate(temp) | 1.1207 | 0.3649 |
| **Var(4)** | plate(temp) | 0.8413 | 0.3164 |
| **Var(5)** | plate(temp) | 0.9807 | 0.3581 |
| **Var(6)** | plate(temp) | 0.8000 | 0.2931 |
| **Var(7)** | plate(temp) | 0.4775 | 0.1500 |
| **Var(8)** | plate(temp) | 0.7745 | 0.2841 |
| **Var(9)** | plate(temp) | 0.7745 | 0.2841 |
| **Var(10)** | plate(temp) | 0.8321 | 0.3039 |
| **Corr(2,1)** | plate(temp) | 0.2313 | 0.2115 |
| **Corr(3,1)** | plate(temp) | 2.71E-35 | . |
| **Corr(3,2)** | plate(temp) | 0.4056 | 0.1847 |
| **Corr(4,1)** | plate(temp) | 1.24E-34 | . |
| **Corr(4,2)** | plate(temp) | 2.31E-33 | . |
| **Corr(4,3)** | plate(temp) | 0.2266 | 0.1664 |

| Covariance Parameter Estimates | | | |
|---|---|---|---|
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** |
| **Corr(5,1)** | plate(temp) | 3.25E-33 | . |
| **Corr(5,2)** | plate(temp) | -184E-34 | . |
| **Corr(5,3)** | plate(temp) | 1.37E-33 | . |
| **Corr(5,4)** | plate(temp) | 0.9212 | 0.06150 |
| **Corr(6,1)** | plate(temp) | 1.24E-24 | . |
| **Corr(6,2)** | plate(temp) | 5.02E-24 | . |
| **Corr(6,3)** | plate(temp) | -77E-25 | . |
| **Corr(6,4)** | plate(temp) | -144E-25 | . |
| **Corr(6,5)** | plate(temp) | 0.000833 | 0.01359 |
| **Corr(7,1)** | plate(temp) | -679E-25 | . |
| **Corr(7,2)** | plate(temp) | -966E-25 | . |
| **Corr(7,3)** | plate(temp) | -819E-25 | . |
| **Corr(7,4)** | plate(temp) | 1.92E-22 | . |
| **Corr(7,5)** | plate(temp) | 7.67E-23 | . |
| **Corr(7,6)** | plate(temp) | 0.8807 | 0.08780 |
| **Corr(8,1)** | plate(temp) | 1.44E-21 | . |
| **Corr(8,2)** | plate(temp) | -441E-23 | . |
| **Corr(8,3)** | plate(temp) | -502E-24 | . |
| **Corr(8,4)** | plate(temp) | 1.59E-21 | . |
| **Corr(8,5)** | plate(temp) | 3.54E-15 | . |
| **Corr(8,6)** | plate(temp) | -2.97E-9 | . |
| **Corr(8,7)** | plate(temp) | 0.4522 | 0.1677 |
| **Corr(9,1)** | plate(temp) | -536E-12 | . |
| **Corr(9,2)** | plate(temp) | -2.57E-8 | . |
| **Corr(9,3)** | plate(temp) | 9.173E-9 | . |
| **Corr(9,4)** | plate(temp) | 1.023E-7 | . |
| **Corr(9,5)** | plate(temp) | -2.2E-7 | . |
| **Corr(9,6)** | plate(temp) | 4.304E-7 | . |
| **Corr(9,7)** | plate(temp) | 6.578E-8 | . |
| **Corr(9,8)** | plate(temp) | 0.003389 | 0.01766 |
| **Corr(10,1)** | plate(temp) | 3.887E-6 | . |
| **Corr(10,2)** | plate(temp) | -0.00002 | . |
| **Corr(10,3)** | plate(temp) | -7.69E-7 | . |

| Covariance Parameter Estimates | | | |
|---|---|---|---|
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** |
| **Corr(10,4)** | plate(temp) | 9.719E-6 | . |
| **Corr(10,5)** | plate(temp) | -0.00003 | . |
| **Corr(10,6)** | plate(temp) | -0.00008 | . |
| **Corr(10,7)** | plate(temp) | -9.06E-6 | . |
| **Corr(10,8)** | plate(temp) | -0.00003 | . |
| **Corr(10,9)** | plate(temp) | 0.9819 | 0.009321 |

Table 60: stage=progress

| Type III Tests of Fixed Effects | | | | |
|---|---|---|---|---|
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **temp** | 2 | 15 | 22.13 | <.0001 |
| **time** | 9 | 135 | 983.84 | <.0001 |
| **temp*time** | 18 | 135 | 8.79 | <.0001 |

Table 54 shows that the Variance matrix is blocked by (i.e., subdivided in equal) blocks, one for each of the identified subjects (i.e., plates). As the residuals are modelled as an R-side random factor, the Residual Pseudo-Likelihood (REPL) is now used as estimation technique. Besides, the Between-Within method to calculate the degrees of freedom is automatically introduced because the residuals (on which the R-side random effects are modelled) are processed according to the presence of subjects: it implies that the residual degrees of freedom are divided into between-subjects and within-subject portions. In this way, a fixed effect whose levels change among observations taken on a given subject (that is, whose levels change within the subject), and modelled according to a structured covariance, is assigned within-subject degrees of freedom, otherwise it is assigned the between-subjects degrees of freedom. As an unstructured covariance was used here, all fixed effects are assigned the between-subjects degrees of freedom to provide for a better small-sample approximation. Time, though a categorical variable, is properly ordered (Table 55). Because of the unstructured covariance matrix, there are 55 R-side parameters and 44 columns in the X matrix (the matrix for the fixed effects), corresponding to a column for the intercept, three columns for the levels of the 'temp' effect, ten for the 'time' effect, and 30 columns for their interaction (Table 56). There are no columns in the Z matrix (the matrix for G-side random effects) for this model because the random effects are residual-type (R-side) only. Subjects (Blocks in V, where the V matrix includes all the random effects, both on the G-side and R-side) are 18 (six plates for each temperature), and the maximum number of observations per subject is 10 (one per every timepoint of the progress stage). The Newton-Raphson with ridging optimization technique is used, as requested, instead of the default Dual Quasi-Newton method (Table 57). The number of parameters in optimization is 55, and, as seen above, they are all covariance parameters because a residual likelihood technique is used to compute the objective function, so that the fixed effects are 'Profiled', that is, the parameters in the covariance matrix are optimized according to a likelihood function based on an iteratively-updated "profile" of fixed-effects parameters. Notice that the "profile" is initially obtained from the data by transposing them onto the linked scale and estimating the fixed-effects parameters with a linear mixed model, which utilizes, in subsequent fittings, the estimates of the parameters in the covariance matrix

as obtained by the optimization of the likelihood function, in a doubly iterative process based on the progressive adjustment of the response pseudo-data according to the updated estimates of random effects. A lower boundary constraint is established for every covariance parameter (namely, zero for variances and -1 for autocorrelation parameters) and an upper boundary constraint of +1 is set for estimates of correlations. It follows that ten variance components are optimized for the random effects, one for each timepoint in the progress stage, as well as 45 autocorrelation parameters, one for every pair of timepoints (i.e., 10·9/2). Presently, the -2 Residual Log Pseudo-Likelihood (Table 58) is the only usable fit statistic in the context of pseudo-likelihood. Note that, anyway, it is often not possible to properly compare values of this fit parameter across different statistical models, not even if one model is contained in the other in terms of fixed and/or random effects. This is because the pseudo-likelihood is based on pseudo-data iteratively updated according to the estimates of the random effects parameters, and it cannot be compared to a pseudo-likelihood statistic obtained for another model from pseudo-data that have most probably diverged (it would correspond to compare model fits corresponding to both different models and different datasets: the two differences are confounded). Only models that differ for the structure of correlation alone, and not for the variance structure (which modifies the BLUPs and therefore the pseudo-data), produce the same pseudo-data and can then be compared. For this reason, all the models' factors should be initially evaluated with a true likelihood, by means of an integral approximation in the presence of G-side random effects, and only thereafter the marginal model can be evaluated. Even the Generalized Chi-Square / DF is a measure of the residual variability in the marginal distribution of the data, and it provides only a rough estimate of overdispersion in this context: a value of 1 suggests there is no overwhelming overdispersion, and, indeed, we know that the present data, and therefore probably the modelled predictors, are underdispersed (Table 21). In fact, as the binomial distribution does not have a residual dispersion parameter but rather its variance is directly dependent upon the mean, the Generalized Chi-Square / DF ratio represents an estimate of the proportion of variance present in the data with respect to that explained by the model (i.e., an estimate of the multiplicative scale parameter), and therefore it can be used only as a rough indicator of overdispersion. Table 59 lists (in an ordinally coded sequence) the estimated variance/covariance parameters in the R matrix. The ten estimated variances are R-side normalized 'working' variances not amenable to interpretation (Stroup et al., 2018) and not comparable outside the same table. They indicate some possible difference in the variability among plates across timepoints. Noticeably, the correlation estimates are positive and non-null for contiguous observations only (note, however, that these are R-side 'working' correlations not real correlations between data; Stroup et al., 2018). Missing standard errors were on boundary where they are essentially equal to zero. All the fixed factors resulted highly significant (Table 60).

Given the considerations above, correlation can be modelled as much stronger for adjacent observations. The covariance structure can therefore be sharply simplified, saving degrees of freedom (and computation time): the high number of R-side parameters (55) could be reduced to 19 (10 variances, one for each observation time of the progress curve, plus 9 autocorrelation parameters, one for each subsequent pair of sequentially adjacent times) for an UNR(2) banded correlation structure without losing any fit of the model. The first-order ante-dependence structure, ANTE(1), requires the same 19 R-side parameters and is commonly used when time spacing between observations is not equal and the correlation structure changes over time (Littell et al., 2006; Stroup et al., 2018). As the exact covariance structure of the data is not known, however, it is recommended that marginal models utilize an 'empirical' procedure, or sandwich estimator, which is more robust to misspecifications of the covariance structure; the MBN correction is advised, specifically for small experiments (Stroup et al., 2018). The 'empirical' option is the preferred alternative to the use of Kenward-Roger degrees of freedom in R-side modelling (Stroup et al., 2018). This kind of models (characterized by estimating the R-side structure with a robust "sandwich" covariance estimator) was originally known as Generalized Estimating Equations (GEE), and is typically applied to longitudinal/clustered data analysis when the focus is on estimating the average response to changing one or more factor levels over the population rather than the effect on the average individual.

/*__Marginal model with categorical time__*/

```
proc GLIMMIX order=data empirical=mbn;
by stage;
class plate temp time;
model germ/n = temp time temp*time / link=probit;
random residual / subject=plate(temp) type=ante(1);
nloptions tech=nrridg;
lsmeans temp*time / cl ilink plot=meanplot slice=time slicediff=time adjust=smm;
run;
```

RESULTS (excerpts):

Table 61: stage=progress

| Optimization Information | |
|---|---|
| **Optimization Technique** | Newton-Raphson with Ridging |
| **Parameters in Optimization** | 19 |
| **Lower Boundaries** | 19 |
| **Upper Boundaries** | 9 |
| **Fixed Effects** | Profiled |
| **Starting From** | Data |

Table 62: stage=progress

| Fit Statistics | |
|---|---|
| **-2 Res Log Pseudo-Likelihood** | -322.96 |
| **Generalized Chi-Square** | 110.38 |
| **Gener. Chi-Square / DF** | 0.74 |

Table 63: stage=progress

| Covariance Parameter Estimates | | | |
|---|---|---|---|
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** |
| **Var(1)** | plate(temp) | 0.2806 | 0.1016 |
| **Var(2)** | plate(temp) | 1.3037 | 0.4837 |
| **Var(3)** | plate(temp) | 1.1029 | 0.4914 |
| **Var(4)** | plate(temp) | 0.8705 | 0.6514 |
| **Var(5)** | plate(temp) | 0.5803 | 0.3780 |
| **Var(6)** | plate(temp) | 0.3341 | 0.1081 |
| **Var(7)** | plate(temp) | 0.3086 | 0.08045 |
| **Var(8)** | plate(temp) | 0.2625 | . |
| **Var(9)** | plate(temp) | 0.2739 | 0.01112 |

| Covariance Parameter Estimates | | | |
|---|---|---|---|
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** |
| **Var(10)** | plate(temp) | 0.5785 | 0.3046 |
| **Rho(1)** | plate(temp) | 0.08797 | 0.2556 |
| **Rho(2)** | plate(temp) | 0.2585 | 0.2697 |
| **Rho(3)** | plate(temp) | 0.4884 | 0.4213 |
| **Rho(4)** | plate(temp) | 0.6520 | 0.2212 |
| **Rho(5)** | plate(temp) | 0.6903 | 0.3801 |
| **Rho(6)** | plate(temp) | 0.8356 | 0.08860 |
| **Rho(7)** | plate(temp) | 0.8207 | 0.4857 |
| **Rho(8)** | plate(temp) | 0.9950 | 0.002717 |
| **Rho(9)** | plate(temp) | 0.7254 | . |

Table 64: stage=progress

| Type III Tests of Fixed Effects | | | | |
|---|---|---|---|---|
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **temp** | 2 | 15 | 3.38 | 0.0613 |
| **time** | 9 | 135 | 420.23 | <.0001 |
| **temp*time** | 18 | 135 | 3.70 | <.0001 |

Figure 12: stage=progress

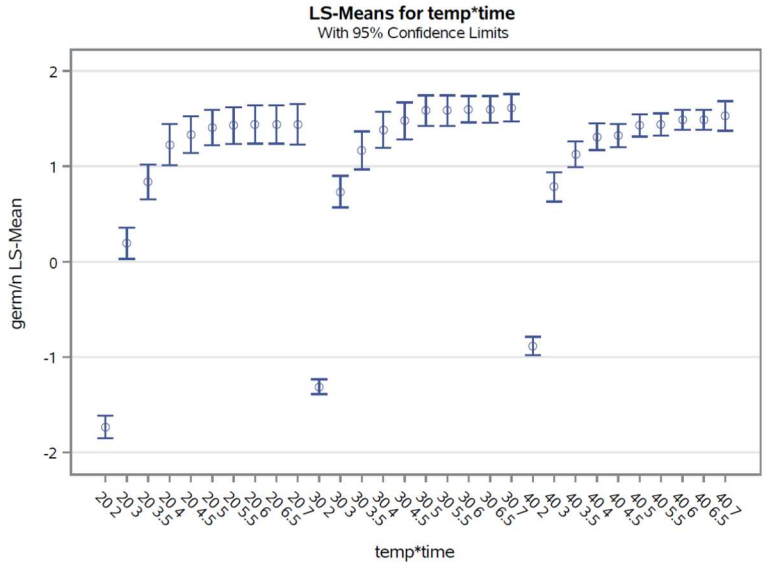


Table 65: stage=progress

| Tests of Effect Slices for temp*time Sliced By time | | | | |
|---|---|---|---|---|
| time | Num DF | Den DF | F Value | Pr > F |
| 2 | 2 | 135 | 61.61 | <.0001 |
| 3 | 2 | 135 | 15.89 | <.0001 |
| 3.5 | 2 | 135 | 3.95 | 0.0216 |
| 4 | 2 | 135 | 0.58 | 0.5598 |
| 4.5 | 2 | 135 | 0.95 | 0.3875 |
| 5 | 2 | 135 | 1.46 | 0.2354 |
| 5.5 | 2 | 135 | 1.20 | 0.3051 |
| 6 | 2 | 135 | 1.10 | 0.3347 |
| 6.5 | 2 | 135 | 1.09 | 0.3380 |
| 7 | 2 | 135 | 0.95 | 0.3888 |

Table 66: stage=progress

| Simple Effect Comparisons of temp*time Least Squares Means By time Adjustment for Multiple Comparisons: SMM | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Simple Effect Level | temp | temp | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Adj P | Alpha | Lower | Upper | Adj Lower | Adj Upper |
| time 2 | 20 | 30 | -0.4211 | 0.07156 | 135 | -5.88 | <.0001 | <.0001 | 0.05 | -0.5626 | -0.2796 | -0.5940 | -0.2482 |
| time 2 | 20 | 40 | -0.8476 | 0.07719 | 135 | -10.98 | <.0001 | <.0001 | 0.05 | -1.0003 | -0.6950 | -1.0341 | -0.6611 |
| time 2 | 30 | 40 | -0.4265 | 0.06257 | 135 | -6.82 | <.0001 | <.0001 | 0.05 | -0.5503 | -0.3028 | -0.5777 | -0.2753 |
| time 3 | 20 | 30 | -0.5400 | 0.1178 | 135 | -4.58 | <.0001 | <.0001 | 0.05 | -0.7730 | -0.3070 | -0.8247 | -0.2553 |
| time 3 | 20 | 40 | -0.5901 | 0.1141 | 135 | -5.17 | <.0001 | <.0001 | 0.05 | -0.8158 | -0.3645 | -0.8658 | -0.3144 |
| time 3 | 30 | 40 | -0.05013 | 0.1143 | 135 | -0.44 | 0.6617 | 0.9609 | 0.05 | -0.2762 | 0.1759 | -0.3264 | 0.2261 |
| time 3.5 | 20 | 30 | -0.3310 | 0.1368 | 135 | -2.42 | 0.0168 | 0.0495 | 0.05 | -0.6015 | -0.06055 | -0.6615 | -0.00052 |
| time 3.5 | 20 | 40 | -0.2907 | 0.1156 | 135 | -2.52 | 0.0131 | 0.0386 | 0.05 | -0.5193 | -0.06215 | -0.5700 | -0.01142 |
| time 3.5 | 30 | 40 | 0.04030 | 0.1217 | 135 | 0.33 | 0.7410 | 0.9824 | 0.05 | -0.2004 | 0.2810 | -0.2538 | 0.3344 |
| time 4 | 20 | 30 | -0.1565 | 0.1451 | 135 | -1.08 | 0.2828 | 0.6293 | 0.05 | -0.4434 | 0.1305 | -0.5071 | 0.1942 |
| time 4 | 20 | 40 | -0.08405 | 0.1306 | 135 | -0.64 | 0.5211 | 0.8892 | 0.05 | -0.3424 | 0.1743 | -0.3998 | 0.2317 |
| time 4 | 30 | 40 | 0.07242 | 0.1189 | 135 | 0.61 | 0.5435 | 0.9041 | 0.05 | -0.1627 | 0.3076 | -0.2149 | 0.3597 |
| time 4.5 | 20 | 30 | -0.1452 | 0.1378 | 135 | -1.05 | 0.2937 | 0.6459 | 0.05 | -0.4177 | 0.1272 | -0.4782 | 0.1877 |
| time 4.5 | 20 | 40 | 0.01006 | 0.1148 | 135 | 0.09 | 0.9303 | 0.9997 | 0.05 | -0.2170 | 0.2371 | -0.2674 | 0.2875 |
| time 4.5 | 30 | 40 | 0.1553 | 0.1152 | 135 | 1.35 | 0.1800 | 0.4469 | 0.05 | -0.07259 | 0.3832 | -0.1232 | 0.4337 |
| time 5 | 20 | 30 | -0.1783 | 0.1241 | 135 | -1.44 | 0.1530 | 0.3908 | 0.05 | -0.4237 | 0.06709 | -0.4782 | 0.1216 |

| Simple Effect Comparisons of temp*time Least Squares Means By time<br>Adjustment for Multiple Comparisons: SMM | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Simple Effect Level** | **temp** | **temp** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Adj P** | **Alpha** | **Lower** | **Upper** | **Adj Lower** | **Adj Upper** |
| **time 5** | 20 | 40 | -0.02278 | 0.1104 | 135 | -0.21 | 0.8368 | 0.9956 | 0.05 | -0.2411 | 0.1955 | -0.2895 | 0.2440 |
| **time 5** | 30 | 40 | 0.1555 | 0.1003 | 135 | 1.55 | 0.1234 | 0.3250 | 0.05 | -0.04286 | 0.3539 | -0.08689 | 0.3980 |
| **time 5.5** | 20 | 30 | -0.1555 | 0.1264 | 135 | -1.23 | 0.2205 | 0.5245 | 0.05 | -0.4054 | 0.09436 | -0.4609 | 0.1498 |
| **time 5.5** | 20 | 40 | -0.01168 | 0.1135 | 135 | -0.10 | 0.9182 | 0.9994 | 0.05 | -0.2362 | 0.2129 | -0.2861 | 0.2627 |
| **time 5.5** | 30 | 40 | 0.1439 | 0.09994 | 135 | 1.44 | 0.1523 | 0.3893 | 0.05 | -0.05379 | 0.3415 | -0.09766 | 0.3854 |
| **time 6** | 20 | 30 | -0.1587 | 0.1235 | 135 | -1.28 | 0.2011 | 0.4884 | 0.05 | -0.4029 | 0.08561 | -0.4571 | 0.1398 |
| **time 6** | 20 | 40 | -0.04879 | 0.1143 | 135 | -0.43 | 0.6702 | 0.9638 | 0.05 | -0.2749 | 0.1773 | -0.3250 | 0.2275 |
| **time 6** | 30 | 40 | 0.1099 | 0.08788 | 135 | 1.25 | 0.2134 | 0.5114 | 0.05 | -0.06393 | 0.2837 | -0.1025 | 0.3223 |
| **time 6.5** | 20 | 30 | -0.1587 | 0.1239 | 135 | -1.28 | 0.2024 | 0.4909 | 0.05 | -0.4036 | 0.08632 | -0.4580 | 0.1407 |
| **time 6.5** | 20 | 40 | -0.04879 | 0.1147 | 135 | -0.43 | 0.6712 | 0.9641 | 0.05 | -0.2756 | 0.1780 | -0.3259 | 0.2283 |
| **time 6.5** | 30 | 40 | 0.1099 | 0.08840 | 135 | 1.24 | 0.2161 | 0.5164 | 0.05 | -0.06496 | 0.2847 | -0.1038 | 0.3235 |
| **time 7** | 20 | 30 | -0.1738 | 0.1293 | 135 | -1.34 | 0.1812 | 0.4494 | 0.05 | -0.4296 | 0.08198 | -0.4864 | 0.1388 |
| **time 7** | 20 | 40 | -0.08785 | 0.1328 | 135 | -0.66 | 0.5093 | 0.8809 | 0.05 | -0.3504 | 0.1747 | -0.4087 | 0.2330 |
| **time 7** | 30 | 40 | 0.08598 | 0.1073 | 135 | 0.80 | 0.4246 | 0.8081 | 0.05 | -0.1263 | 0.2983 | -0.1734 | 0.3454 |

As said, this marginal model requires 19 R-side parameters for the progress stage: 10 variances and 9 autocorrelation parameters (Table 61). The -2 Residual Log Pseudo-Likelihood (Table 62) would suggest a better fit than the unstructured model, but the estimated random variances are not the same (Table 63) and therefore the pseudo-data do not correspond, thus the two fit statistics cannot be compared. In the same table it can be noted that the correlation between adjacent observations increases as the germination progress reaches a plateau (remind, however, that these are working correlations). In fact, during the germination surge, adjacent germination data largely differ, whereas they are very close at the plateau. The *F* tests indicate a highly significant effect for time and the temperature x time interaction, whereas the temperature effect is just above the 0.05 threshold (Table 64). The plot of the estimated least-square means and their confidence interval with 95% limits on the linked scale (Figure 12) is very similar to that obtained with the conditional model with categorical time (Figure 5). Analogously, differences among temperatures are significant at days 2-3.5 (Table 65), and multiple comparisons at each timepoint (Table 66) indicate a different effect for each of the three temperatures at 2 days, but only for 20 °C with respect to the other two temperatures at days 3-3.5, like inferred by previous models.

It ought to be noted that, even though the first-order ante-dependence structure, ANTE(1), is appropriate for the present model, the R-side variance/covariance structure could even be modelled with a first-order autoregressive structure, AR(1). AR(1) is, indeed, the most commonly covariance structure used in longitudinal studies, because, like ANTE(1), the correlation of the repeated measurements is assumed to decrease for observations that are farther away in time (Littell et al., 2006). However, AR(1), like several other time-series covariance structures, assumes equal time intervals, which ANTE(1) does not require (Littell et al., 2006). In fact, in AR(1) and similar simpler covariance structures, the correlation between adjacent within-subject errors is modelled as constant throughout time, that is, a single autocorrelation parameter is considered between adjacent observations throughout the germination time-course. The reason for this assumption is that changes in the data are supposed to be constant through time so that a single contiguous

correlation holds throughout the whole process. However, constant changes mean linearity of the response vs. time on the linked scale, which, we have seen, does not hold true for the germination time-course. In fact, as seen, correlation tends to increase through time because of plateauing. Nevertheless, in germination studies, a varying time interval is often adopted in the form of shorter time spacings in correspondence of faster changes in germination, thus that the shorter time spacings, partially compensate for larger changes in the response (this was not properly done in the original experiment from which these data were obtained, though). In this sense, AR(1) and similar covariance structures might still be sensible even in presence of this kind of uneven time spacing. The empirical MBN estimator helps increasing the robustness of this approach. Note that, in modelling the covariance structure of the residuals, the data are considered as ordered according to a categorical longitudinal variable (which is not made explicit, in this instance) even if time were modelled as continuous, because continuous effects are not allowed in R-side random effects, but for spatial covariance structures. Thus, considering time as continuous would not solve the problem of unequal time spacing.

In addition, it should be noted that when in the 'random residual' statements an unstructured covariance is specified, it accounts for both within-subjects as well as for between-subjects variances and covariances, whereas the AR(1) covariance structure only accounts for the within-subjects covariances (Littell et al., 2006). Thus, a 'random residual' statement specifying an unstructured-type covariance must not be accompanied by a separate 'random' statement to keep into account the between-subjects variance of the effect used as 'subject' in the 'repeated' statement; whereas such a separate 'random' statement, modelling a G-type random effect, is instead appropriate for the R-side AR(1) covariance structure (Littell et al., 2006). Such a G-type statement is advised against for the ANTE(1) covariance structure, because it can give convergence problems (Stroup et al., 2018). The simultaneous presence of G-side and R-side modelling leads to a quasi-marginal structure (Stroup et al., 2018) that includes the random intercept effect for the subjects together with an AR(1) structure of the residuals, and it can be therefore termed AR(1)+RE (Littell et al., 2006). Furthermore, when the AR(1) covariance structure is used, it could be advisable to add a 'group=' option that specifies the parameters of the covariance structure be varied according to the 'group' effect. This effect ought to be specified as the fixed factor other than time, if there is only one, or as the highest interaction between fixed factors apart from time. In this way, unequal variances and different correlations are modelled (both as R-side working scale parameters; Gbur et al., 2012) for the levels of the other fixed factor, or for the highest interaction between fixed factors apart from time, so that if they strongly affect the germination response, the R-side variance/covariance structure can better suit the different germination progress curves. Whether an ANTE(1) or AR(1)(with groups)+RE structure has to be used, depends on how much the progress curves differ in the experiment. If they are not very much diverse, like for the present data, the former structure is preferable, since it better suits the overall shape of the similar germination curves. If, on the other hand, the progress curves are of widely different shapes, like when seed samples with diverse dormancy intensities are compared, the latter structure can be more suitable because of better flexibility with respect to the widely different time-courses.

Finally, it needs to be noticed that the 'empirical' option requires that the data are grouped by subjects in every 'random' or 'random residual' statement, and that the specified 'subject=' effects are the same, or contain each other, across statements. Hence, in the 'random' statement, it must be made explicit that the between-plates effect is modelled in terms of random deviations from the general intercept. The following model is then fitted:

```
/*Quasi-marginal model with categorical time*/
proc GLIMMIX order=data empirical=mbn;
by stage;
class plate temp time;
model germ/n = temp time temp*time / link=probit;
```

```
random intercept / subject=plate(temp);
random residual / subject=plate(temp) type=ar(1) group=temp;
nloptions tech=nrridg;
lsmeans temp*time / cl ilink plot=meanplot slice=time slicediff=time adjust=smm;
covtest diagR;
run;
```

Note that the 'covtest' statement is used here to test independence of residuals by constraining off-diagonal elements in the R matrix to zero and thereby reducing the R-side covariance structure to the diagonal form. As the G-side structure is not modified, this test indicates whether modelling a correlation among errors of the repeated measures (i.e., considering conditional dependence of among-plates variances through time) improves the fit even if the effect of plates is already considered in terms of random deviations from the general intercept. This can be done since the pseudo-likelihoods with and without the correlation parameters are computed using the same pseudo-data, because all the variances are unmodified across the two models.

RESULTS (excerpts):

Table 67: stage=progress

| Model Information | |
|---|---|
| **Data Set** | WORK.REFFILE |
| **Response Variable (Events)** | germ |
| **Response Variable (Trials)** | n |
| **Response Distribution** | Binomial |
| **Link Function** | Probit |
| **Variance Function** | Default |
| **Variance Matrix Blocked By** | plate(temp) |
| **Estimation Technique** | Residual PL |
| **Degrees of Freedom Method** | Containment |
| **Fixed Effects SE Adjustment** | Sandwich - MBN(df,r=1,d=2) |

Table 68: stage=progress

| Dimensions | |
|---|---|
| **G-side Cov. Parameters** | 1 |
| **R-side Cov. Parameters** | 6 |
| **Columns in X** | 44 |
| **Columns in Z per Subject** | 1 |
| **Subjects (Blocks in V)** | 18 |
| **Max Obs per Subject** | 10 |

Table 69: stage=progress

| Optimization Information | |
|---|---|
| **Optimization Technique** | Newton-Raphson with Ridging |
| **Parameters in Optimization** | 7 |
| **Lower Boundaries** | 7 |
| **Upper Boundaries** | 3 |
| **Fixed Effects** | Profiled |
| **Starting From** | Data |

Table 70: stage=progress

| Fit Statistics | |
|---|---|
| **-2 Res Log Pseudo-Likelihood** | -193.92 |
| **Generalized Chi-Square** | 150.00 |
| **Gener. Chi-Square / DF** | 1.00 |

Table 71: stage=progress

| Covariance Parameter Estimates | | | | |
|---|---|---|---|---|
| **Cov Parm** | **Subject** | **Group** | **Estimate** | **Standard Error** |
| **Intercept** | plate(temp) | | 0.004386 | 0.004202 |
| **Variance** | plate(temp) | temp 20 | 1.1791 | 0.5069 |
| **AR(1)** | plate(temp) | temp 20 | 0.8355 | 0.07488 |
| **Variance** | plate(temp) | temp 30 | 0.6192 | 0.2296 |
| **AR(1)** | plate(temp) | temp 30 | 0.7039 | 0.1116 |
| **Variance** | plate(temp) | temp 40 | 0.3533 | 0.1110 |
| **AR(1)** | plate(temp) | temp 40 | 0.4732 | 0.1694 |

Table 72: stage=progress

| Type III Tests of Fixed Effects | | | | |
|---|---|---|---|---|
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **temp** | 2 | 15 | 2.42 | 0.1227 |
| **time** | 9 | 135 | 342.16 | <.0001 |
| **temp*time** | 18 | 135 | 3.03 | 0.0001 |

Figure 13: stage=progress

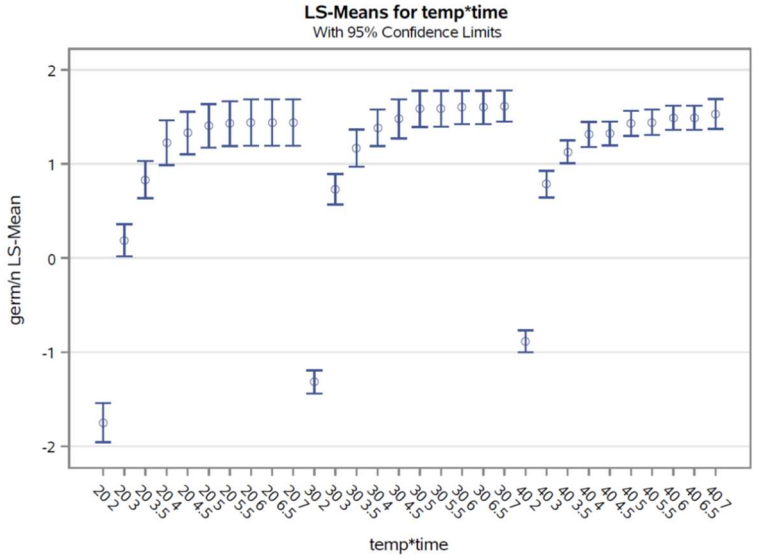



Table 73: stage=progress

| Tests of Effect Slices for temp*time Sliced By time | | | | |
|---|---|---|---|---|
| time | Num DF | Den DF | F Value | Pr > F |
| 2 | 2 | 135 | 29.69 | <.0001 |
| 3 | 2 | 135 | 15.98 | <.0001 |
| 3.5 | 2 | 135 | 3.72 | 0.0267 |
| 4 | 2 | 135 | 0.53 | 0.5899 |
| 4.5 | 2 | 135 | 0.83 | 0.4390 |
| 5 | 2 | 135 | 1.04 | 0.3568 |
| 5.5 | 2 | 135 | 0.84 | 0.4335 |
| 6 | 2 | 135 | 0.70 | 0.4986 |
| 6.5 | 2 | 135 | 0.70 | 0.4989 |
| 7 | 2 | 135 | 0.71 | 0.4926 |

Table 74: stage=progress

| Simple Effect Comparisons of temp*time Least Squares Means By time Adjustment for Multiple Comparisons: SMM | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Simple Effect Level** | **temp** | **temp** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Adj P** | **Alpha** | **Lower** | **Upper** | **Adj Lower** | **Adj Upper** |
| **time 2** | 20 | 30 | -0.4315 | 0.1222 | 135 | -3.53 | 0.0006 | 0.0017 | 0.05 | -0.6731 | -0.1898 | -0.7267 | -0.1362 |
| **time 2** | 20 | 40 | -0.8625 | 0.1202 | 135 | -7.17 | <.0001 | <.0001 | 0.05 | -1.1003 | -0.6247 | -1.1531 | -0.5719 |
| **time 2** | 30 | 40 | -0.4311 | 0.08551 | 135 | -5.04 | <.0001 | <.0001 | 0.05 | -0.6002 | -0.2620 | -0.6377 | -0.2244 |
| **time 3** | 20 | 30 | -0.5431 | 0.1188 | 135 | -4.57 | <.0001 | <.0001 | 0.05 | -0.7780 | -0.3081 | -0.8302 | -0.2559 |
| **time 3** | 20 | 40 | -0.5965 | 0.1121 | 135 | -5.32 | <.0001 | <.0001 | 0.05 | -0.8181 | -0.3749 | -0.8673 | -0.3257 |
| **time 3** | 30 | 40 | -0.05347 | 0.1081 | 135 | -0.49 | 0.6217 | 0.9453 | 0.05 | -0.2673 | 0.1603 | -0.3147 | 0.2078 |
| **time 3.5** | 20 | 30 | -0.3334 | 0.1413 | 135 | -2.36 | 0.0197 | 0.0578 | 0.05 | -0.6129 | -0.05400 | -0.6749 | 0.008014 |
| **time 3.5** | 20 | 40 | -0.2950 | 0.1173 | 135 | -2.51 | 0.0131 | 0.0387 | 0.05 | -0.5271 | -0.06294 | -0.5786 | -0.01143 |
| **time 3.5** | 30 | 40 | 0.03844 | 0.1173 | 135 | 0.33 | 0.7437 | 0.9830 | 0.05 | -0.1935 | 0.2704 | -0.2450 | 0.3219 |
| **time 4** | 20 | 30 | -0.1590 | 0.1548 | 135 | -1.03 | 0.3064 | 0.6645 | 0.05 | -0.4652 | 0.1472 | -0.5332 | 0.2152 |
| **time 4** | 20 | 40 | -0.08848 | 0.1377 | 135 | -0.64 | 0.5215 | 0.8895 | 0.05 | -0.3607 | 0.1838 | -0.4211 | 0.2442 |
| **time 4** | 30 | 40 | 0.07051 | 0.1182 | 135 | 0.60 | 0.5517 | 0.9091 | 0.05 | -0.1632 | 0.3042 | -0.2151 | 0.3561 |
| **time 4.5** | 20 | 30 | -0.1479 | 0.1553 | 135 | -0.95 | 0.3427 | 0.7143 | 0.05 | -0.4551 | 0.1593 | -0.5233 | 0.2275 |
| **time 4.5** | 20 | 40 | 0.006176 | 0.1309 | 135 | 0.05 | 0.9624 | 0.9999 | 0.05 | -0.2527 | 0.2650 | -0.3101 | 0.3225 |
| **time 4.5** | 30 | 40 | 0.1541 | 0.1226 | 135 | 1.26 | 0.2109 | 0.5068 | 0.05 | -0.08832 | 0.3965 | -0.1421 | 0.4503 |
| **time 5** | 20 | 30 | -0.1810 | 0.1514 | 135 | -1.20 | 0.2339 | 0.5485 | 0.05 | -0.4804 | 0.1184 | -0.5469 | 0.1848 |
| **time 5** | 20 | 40 | -0.02653 | 0.1346 | 135 | -0.20 | 0.8441 | 0.9962 | 0.05 | -0.2927 | 0.2397 | -0.3518 | 0.2988 |
| **time 5** | 30 | 40 | 0.1545 | 0.1178 | 135 | 1.31 | 0.1919 | 0.4705 | 0.05 | -0.07847 | 0.3875 | -0.1302 | 0.4392 |
| **time 5.5** | 20 | 30 | -0.1578 | 0.1543 | 135 | -1.02 | 0.3081 | 0.6669 | 0.05 | -0.4629 | 0.1473 | -0.5306 | 0.2150 |
| **time 5.5** | 20 | 40 | -0.01487 | 0.1383 | 135 | -0.11 | 0.9145 | 0.9994 | 0.05 | -0.2883 | 0.2586 | -0.3490 | 0.3192 |
| **time 5.5** | 30 | 40 | 0.1430 | 0.1183 | 135 | 1.21 | 0.2290 | 0.5398 | 0.05 | -0.09100 | 0.3769 | -0.1429 | 0.4289 |
| **time 6** | 20 | 30 | -0.1602 | 0.1537 | 135 | -1.04 | 0.2990 | 0.6538 | 0.05 | -0.4641 | 0.1437 | -0.5316 | 0.2112 |
| **time 6** | 20 | 40 | -0.05116 | 0.1407 | 135 | -0.36 | 0.7167 | 0.9770 | 0.05 | -0.3293 | 0.2270 | -0.3911 | 0.2888 |
| **time 6** | 30 | 40 | 0.1091 | 0.1107 | 135 | 0.99 | 0.3262 | 0.6924 | 0.05 | -0.1098 | 0.3279 | -0.1584 | 0.3765 |
| **time 6.5** | 20 | 30 | -0.1600 | 0.1538 | 135 | -1.04 | 0.3000 | 0.6552 | 0.05 | -0.4641 | 0.1441 | -0.5316 | 0.2116 |
| **time 6.5** | 20 | 40 | -0.05078 | 0.1408 | 135 | -0.36 | 0.7188 | 0.9775 | 0.05 | -0.3291 | 0.2276 | -0.3909 | 0.2894 |
| **time 6.5** | 30 | 40 | 0.1092 | 0.1107 | 135 | 0.99 | 0.3256 | 0.6916 | 0.05 | -0.1097 | 0.3281 | -0.1583 | 0.3767 |
| **time 7** | 20 | 30 | -0.1747 | 0.1507 | 135 | -1.16 | 0.2484 | 0.5735 | 0.05 | -0.4728 | 0.1234 | -0.5390 | 0.1895 |
| **time 7** | 20 | 40 | -0.09015 | 0.1486 | 135 | -0.61 | 0.5452 | 0.9051 | 0.05 | -0.3841 | 0.2038 | -0.4494 | 0.2691 |
| **time 7** | 30 | 40 | 0.08458 | 0.1165 | 135 | 0.73 | 0.4689 | 0.8491 | 0.05 | -0.1457 | 0.3149 | -0.1969 | 0.3660 |

Table 75: stage=progress

| Tests of Covariance Parameters Based on the Residual Pseudo-Likelihood | | | | | |
|---|---|---|---|---|---|
| **Label** | **DF** | **-2 Res Log P-Like** | **ChiSq** | **Pr > ChiSq** | **Note** |
| Conditional Independence | 3 | -139.87 | 54.05 | <.0001 | DF |

*DF: P-value based on a chi-square with DF degrees of freedom.*

The quasi-marginal model reverts to applying the 'Containment' Degrees of Freedom Method, but a Sandwich - MBN adjustment of Fixed Effects standard errors is now utilized (Table 67). The variance among plates (i.e., the random intercept effect) is computed as G-side variance/covariance parameter and it represents a column in the Z matrix for each of the 18 subjects (i.e., plates), whereas there are six R-side variance/covariance parameters (Table 68), namely one within-subject variance and one correlation for each temperature level. There are therefore seven parameters in optimization: the G-side between-subjects (among plates) variance, and three within-subjects variances with three within-subjects autocorrelation parameters on the R-side (Table 69). No large overdispersion is apparent, as expected (Table 70). Although the -2 Residual Log Pseudo-Likelihood cannot be compared between the marginal model with ANTE(1) structure and the current quasi-marginal model because random variances are modelled differently, thus that pseudo-data are diverse too, the overly lower value obtained with the former model (Table 62) with respect to the latter is in agreement with theoretical considerations suggesting that these data are better fitted by a covariance structure that allows the autocorrelation parameters to depend on the particular time period considered, rather than assuming a single overall autocorrelation parameter, owing to non-linearity of germination progression even on the linked scale. This, however, quite probably holds true because the germination time-courses at the three temperatures are very similar, thus that they can be modelled with a unique set of covariance parameters. Were the germination time-courses very different, a unique set of covariance parameters would presumably be unsatisfactory, and the quasi-marginal model with a different AR(1) correlation for each time-course could be more sensible. A marginal model with ANTE(1) structure with a diverse set of parameters for each time-course would, on the other hand, easily incur in overparameterization. In GzLMM, estimation convergence is a common practical issue, and non-convergence can result when there are fewer observations than parameters in the model that is being fit (Gbur et al., 2012). This is especially an issue for models with many covariance parameters, and it can be solved by fitting a simpler model (Gbur et al., 2012). It is therefore important to identify the simplest covariance model that adequately accounts for the correlation structure in the data (Gbur et al., 2012).

Table 71 shows that, in this model, the variance among plates is small and maybe non-significant (at least, it is numerically close to its standard error), while within-subjects variances and correlations (called AR(1) in the output table) seem to differ among temperature levels (remind that R-side parameters are 'working', not real, values; Stroup et al., 2018). A non-significant effect (assuming a significance threshold $P = 0.05$) is found for temperature (Table 72). The LS-means are similar to those of the previous models (Figure 13). As consistently inferred, significant differences are found at days 2-3.5 (Table 73), with every temperature level differing at day 2, 20 °C differing from the two other levels at day 3, but only from 40 °C at day 3.5, though the difference with 30 °C is almost significant (Table 74). As said, this model appears less suit to the present data, but it ought to be of greater value when the germination time-courses are strongly different. The test for conditional independence (Table 75) confirms that considering the correlation among errors of the repeated measures improves the fit even if the effect of plates is already considered in terms of random intercepts. Moreover, whereas conditional independence is highly significant, the estimated standard error

of the intercept variance, i.e. the among plate variance, is almost of the same size as the variance itself, suggesting it is probably not significant, but this cannot be tested exactly. As plates are physical elements of the experimental design, their effect is retained anyway (Stroup et al., 2018), though it is small (a value 0.004386 probits, from Table 71, corresponds to a standard deviation of 2.6 % around 50 % on the percentile scale). A predominant effect of conditional dependence with respect to random intercepts is confirmed even by a quasi-marginal model with ANTE(1) structure of residuals, which, though sometimes is computationally problematic (Stroup et al., 2018), converges to a reasonable solution in this instance (not shown).

Regarding the variability among plates, it should be noticed that, though it is low in the presently described experiment using neatly results from replicate plates of practically uniform seeds, it was much wider in the experiment described in Annex II, wherein seeds were infected with a fungus, thus that the germination response was consequent to the interaction of the fungus with the seeds. These biological interactions are subject to large stochastic fluctuations that result into stronger heterogeneity among plates, that is, a wider between-plates random effect. Differently from R-side working variances, G-side variances, modelled as constant on the linked scale, can be used to obtain an idea of the magnitude of variability among plates in percentage terms by naïvely computing the value that corresponds to a maximum standard deviation around 50 %, as previously mentioned. For example, a G-side between-plates variance of 0.2978 (Table 12 in Annex II) on the probit scale corresponds to a maximum standard deviation of about 21 % around a mean of 50 % on the percentile scale. This is a much larger variability, which can be expected when an important source of heterogeneity is present. In addition, as previously mentioned, greater shrinkage of BLUPs in longitudinal data also contributes to lowering the modelled random factor variance with respect to FGP tests.

As final improvement, a marginal model (given it better suits the present data than a quasi-marginal model) with time as continuous variable can be evaluated:

```
/*Marginal model with continuous time*/
proc GLIMMIX empirical=mbn;
by stage;
effect spltime = spline(time);
class plate temp;
model germ/n = temp temp*spltime / link=probit;
random residual / subject=plate(temp) type=ante(1);
nloptions tech=newrap;
lsmeans temp / at time=2 ilink adjust=smm;
lsmeans temp / at time=3 ilink adjust=smm;
lsmeans temp / at time=3.5 ilink adjust=smm;
lsmeans temp / at time=4 ilink adjust=smm;
lsmeans temp / at time=4.5 ilink adjust=smm;
lsmeans temp / at time=5 ilink adjust=smm;
lsmeans temp / at time=5.5 ilink adjust=smm;
lsmeans temp / at time=6 ilink adjust=smm;
lsmeans temp / at time=6.5 ilink adjust=smm;
lsmeans temp / at time=7 ilink adjust=smm;
run;
```

The ‘order=data’ option is no longer necessary since time is now considered a numerical variable. Given this is a marginal model, it is processed with REPL and ‘containment’ ddf with MBN correction, like the marginal model with categorical time. However, the ‘newrap’ optimization technique is used because the algorithm using the Newton-Raphson method with line search seems to achieve easier convergence for this covariance

structure in presence of continuous time. Whether to use ‘nrridge’ or ‘newrap’ is sometimes a matter of trying, though ‘newrap’ occasionally works better in the absence of G-side random factors.

RESULTS (excerpts):

Table 76: stage=progress

| Optimization Information | |
|---|---|
| **Optimization Technique** | Newton-Raphson |
| **Parameters in Optimization** | 19 |
| **Lower Boundaries** | 19 |
| **Upper Boundaries** | 9 |
| **Fixed Effects** | Profiled |
| **Starting From** | Data |

Table 77: stage=progress

| Fit Statistics | |
|---|---|
| **-2 Res Log Pseudo-Likelihood** | -640.18 |
| **Generalized Chi-Square** | 159.00 |
| **Gener. Chi-Square / DF** | 1.00 |

Table 78: stage=progress

| Covariance Parameter Estimates | | | |
|---|---|---|---|
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** |
| **Var(1)** | plate(temp) | 0.2827 | 0.1031 |
| **Var(2)** | plate(temp) | 1.3349 | 0.4739 |
| **Var(3)** | plate(temp) | 1.3715 | 0.4624 |
| **Var(4)** | plate(temp) | 1.0159 | 0.3103 |
| **Var(5)** | plate(temp) | 0.9402 | 0.2828 |
| **Var(6)** | plate(temp) | 0.4941 | 0.1361 |
| **Var(7)** | plate(temp) | 0.5203 | 0.1446 |
| **Var(8)** | plate(temp) | 0.4135 | 0.1095 |
| **Var(9)** | plate(temp) | 0.4135 | 0.1095 |
| **Var(10)** | plate(temp) | 0.4574 | 0.1208 |
| **Rho(1)** | plate(temp) | 0.1664 | 0.2473 |
| **Rho(2)** | plate(temp) | 0.4799 | 0.1989 |

| Covariance Parameter Estimates | | | |
|---|---|---|---|
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** |
| **Rho(3)** | plate(temp) | 0.7832 | 0.08769 |
| **Rho(4)** | plate(temp) | 0.9435 | 0.02831 |
| **Rho(5)** | plate(temp) | 0.8457 | 0.06812 |
| **Rho(6)** | plate(temp) | 0.9773 | 0.01083 |
| **Rho(7)** | plate(temp) | 0.9310 | 0.02992 |
| **Rho(8)** | plate(temp) | 1.0000 | . |
| **Rho(9)** | plate(temp) | 0.9647 | 0.01548 |

Table 79: stage=progress

| Type III Tests of Fixed Effects | | | | |
|---|---|---|---|---|
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **temp** | 2 | 15 | 1.06 | 0.3717 |
| **spltime*temp** | 18 | 90 | 259.88 | <.0001 |

Table 80: stage=progress

| temp Least Squares Means | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **temp** | **time** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Mean** | **Standard Error Mean** |
| 20 | 2.00 | -1.7308 | 0.05813 | 15 | -29.77 | <.0001 | 0.04175 | 0.005186 |
| 30 | 2.00 | -1.3075 | 0.03825 | 15 | -34.18 | <.0001 | 0.09552 | 0.006491 |
| 40 | 2.00 | -0.8800 | 0.04741 | 15 | -18.56 | <.0001 | 0.1894 | 0.01284 |

Table 81: stage=progress

| Differences of temp Least Squares Means<br>Adjustment for Multiple Comparisons: SMM | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **temp** | **temp** | **time** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Adj P** |
| 20 | 30 | 2.00 | -0.4233 | 0.06959 | 15 | -6.08 | <.0001 | <.0001 |
| 20 | 40 | 2.00 | -0.8508 | 0.07501 | 15 | -11.34 | <.0001 | <.0001 |
| 30 | 40 | 2.00 | -0.4275 | 0.06091 | 15 | -7.02 | <.0001 | <.0001 |

Table 82: stage=progress

| temp Least Squares Means | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| temp | time | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Mean | Standard Error Mean |
| 20 | 3.00 | 0.1988 | 0.07993 | 15 | 2.49 | 0.0252 | 0.5788 | 0.03126 |
| 30 | 3.00 | 0.7657 | 0.08365 | 15 | 9.15 | <.0001 | 0.7781 | 0.02489 |
| 40 | 3.00 | 0.8340 | 0.07622 | 15 | 10.94 | <.0001 | 0.7979 | 0.02148 |

Table 83: stage=progress

| Differences of temp Least Squares Means Adjustment for Multiple Comparisons: SMM | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| temp | temp | time | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Adj P |
| 20 | 30 | 3.00 | -0.5669 | 0.1157 | 15 | -4.90 | 0.0002 | 0.0006 |
| 20 | 40 | 3.00 | -0.6353 | 0.1104 | 15 | -5.75 | <.0001 | 0.0001 |
| 30 | 40 | 3.00 | -0.06838 | 0.1132 | 15 | -0.60 | 0.5547 | 0.9047 |

Table 84: stage=progress

| temp Least Squares Means | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| temp | time | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Mean | Standard Error Mean |
| 20 | 3.50 | 0.9058 | 0.08848 | 15 | 10.24 | <.0001 | 0.8175 | 0.02342 |
| 30 | 3.50 | 1.2303 | 0.07941 | 15 | 15.49 | <.0001 | 0.8907 | 0.01486 |
| 40 | 3.50 | 1.2801 | 0.06415 | 15 | 19.96 | <.0001 | 0.8998 | 0.01128 |

Table 85: stage=progress

| Differences of temp Least Squares Means Adjustment for Multiple Comparisons: SMM | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| temp | temp | time | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Adj P |
| 20 | 30 | 3.50 | -0.3245 | 0.1189 | 15 | -2.73 | 0.0155 | 0.0444 |
| 20 | 40 | 3.50 | -0.3744 | 0.1093 | 15 | -3.43 | 0.0038 | 0.0110 |
| 30 | 40 | 3.50 | -0.04983 | 0.1021 | 15 | -0.49 | 0.6325 | 0.9461 |

As previously seen, the marginal model with ANTE(1) variance/covariance structure requires 19 R-side parameters for the progress stage: 10 variances and 9 autocorrelation parameters (Table 76). The twenty-one parameters for fixed effects included in the optimization of the corresponding conditional model (Table 40) were not optimized by pseudo-likelihood in this model because they were profiled, that is, estimated with a linear mixed model onto the linked scale and thereafter used as pseudo-data for the pseudo-likelihood optimization. The -2 Residual Log Pseudo-Likelihood (Table 77) would suggest a better fit than the unstructured model (see Table 58), but the estimated random variances are not the same (compare Tables 78 and 59) and therefore the pseudo-data do not correspond, thus that the two fit statistics cannot be properly compared (though such a drop in this fit statistic, associated to very close estimations of the means, might hint to a real improvement). In the same tables, an increase of (working) correlation between adjacent observations is observed as the germination progress reaches a plateau. In fact, during the germination surge, adjacent germination data largely differ, whereas they are very close at the plateau. Notice that, in the R-side approach, off-diagonal terms of the R matrix act as working covariances embedded in the variance function, thus that these parameters do not have interpretations *per se*, but do account for serial correlation (Gbur et al., 2012). On the other hand, in the G-side approach, serial correlation embeds in the linear predictor and is assumed to follow a Gaussian random distribution (Gbur et al., 2012).

Like found for the conditional model with continuous time, the *F* tests (Table 79) indicate a highly significant effect for 'temp*spltime' but a non-significant effect for 'temp', which means that the different temperatures cause a change of the shape of the curvilinear trend and not a vertical shift of the spline. As noticed before, this is not in contrast with previous findings of a significant 'temp' effect when time is considered categorical, because when time is modelled as a continuous variable, the meaning of the fixed effects is changed. Like consistently found in all the well-suited models, differences among temperatures are significant at days 2-3.5, with estimated means close to those obtained with previous models. Multiple comparisons at each discrete timepoint indicate a different effect for each of the three temperatures at 2 days, but only for 20 °C with respect to the other two temperatures at days 3-3.5 (Tables 80-85), as inferred by the other well-suited models. Thus, essentially identical conclusions can be drowned from either analysis for this data set wherein 100 seeds are tested per each cluster/treatment/time combination. This is a noticeable accomplishment, as the limited number of Bernoulli trials normally utilized in germination studies could otherwise be expected to amplify discrepancies between the G-side and R-side approaches, though standard errors tend to be much more affected than the estimates (Gbur et al., 2012).

It may be worth mentioning that even in the case of a quasi-marginal model with AR(1) structure and continuous time, estimates would be similar to those obtained with categorical time (not shown).

As plates are nested within treatments, the effect of their interaction with such fixed factor is included in the estimate of their effect (Quinn and Keough, 2002). Nevertheless, as longitudinal studies correspond to partly nested experimental designs (Quinn and Keough, 2002), a plate x time interaction effect could be envisaged that could be managed with a random coefficients model (Littell et al., 2006) having both a random intercept and a random slope through time for each plate. However, if uniform and healthy seed samples are tested, plates are not heterogeneous subjects, like genetically diverse individuals or materially different places, rather they are quite uniform, almost alike to true replicates. Thus, it is not envisioned that they can introduce subject-specific effects that can affect the response of the single plate through time; that is, random slopes are expected to differ only negligibly, thereby representing a zero-variance component estimate that leads to boundary estimation troubles (Gbur et al., 2012). As seen, indeed, plates diversity, as measured in terms of deviations from the common intercept, is small, and through-time effects associated with individual plates are stochastic fluctuations. Furthermore, apart from the default 'variance components' structure, the plate x time interaction effect is already embedded into more complex covariance structures. Modelling of this interaction effect is therefore entrusted to the longitudinal variance/covariance structure.

Finally, it needs to be noticed that even though a thorough evaluation of the best variance/covariance structure would require testing several structures on the G-side with Laplace approximation and then choosing the best structure based on the smaller value of the AICC goodness-of-fit statistic (Gbur et al., 2012), this can be somewhat troublesome to apply to germination studies because of the hierarchical structure typical of their experimental designs, wherein plates represent subjects nested within treatment combinations, i.e. random effects within a fixed-effects framework, rather than being complete blocks over which fixed effects are replicated. Computational problems can, in fact, arise because nesting corresponds to an entirely unbalanced interaction (that is, the diverse treatments are applied onto different plates) and the nested plate effect is thereby completely confounded with its plate x treatment interaction (Quinn and Keough, 2002). As this interaction is embedded into covariance structure while the plate effect is modelled in terms of intercepts on the G-side, whereas on the R-side it isn't (at least, for ANTE(1) and AR(1) structures), this can cause some redundancy that leads to zero-variance estimates and therefore boundary estimation troubles when modelling on the G-side alone. Boundary estimation within the same matrix can ensue in assigning all the significance to either one or the other effect (usually to the interaction effect, which is evaluated first in type III ANOVA), especially when these effects are partially overlapping, and the interaction effect is very small or even negligible. This causes troubles when modelling the covariance structure of longitudinal data on the G-side alone. Small, partially overlapping effects of the plate random factor and of errors through time (wherein the plate x treatment interaction is embedded) are better managed by separately modelling them on distinct, dedicated matrices (chiefly because the G-side effect is evaluated prior to the R-side one, particularly in pseudo-likelihood). Thus, marginal and quasi-marginal models can provide a more robust convergence, at least for this kind of data and if the common ANTE(1) or AR(1) structures are utilized.

**References**


Gbur E.E., Stroup W.W., McCarter K.S., Durham S., Young L.J., Christman M., West M. and Kramer M. (2012). *Analysis of Generalized Linear Mixed Models in the Agricultural and Natural Resources Sciences*. American Society of Agronomy: Madison, WI, USA.

Littell R.C., Milliken G.A., Stroup W.W., Wolfinger R.D. and Schabenberger O. (2006). *SAS® for Mixed Models, Second Edition*. SAS Institute Inc.: Cary, NC, USA.

Piacco R. (1954). Ricerche sui metodi di analisi delle sementi di riso (*Oryza sativa* L.) nei riguardi della facoltà germinativa. *Annali della Stazione Sperimentale di Risicoltura e delle Colture Irrigue di Vercelli* 2:111-179.

Quinn G.P. and Keough M.J. (2002). *Experimental Design and Data Analysis for Biologists*. Cambridge University Press: Cambridge, UK.

Stroup W.W. (2015). Rethinking the analysis of non-normal data in plant and soil science. *Agronomy Journal* 107:811-827.

Stroup W.W., Milliken G.A., Claassen E.A. and Wolfinger R.D. (2018). *SAS® for Mixed Models: Introduction and Basic Applications*. SAS Institute Inc.: Cary, NC, USA.

Zar J.H. (1999). *Biostatistical Analysis*, 4th ed. Prentice Hall: Upper Saddle River, NJ, USA.