**Andrew Hillard**

**Data Analysis 2**

**Homework 5**

Problem 2

To fit the model using NLMIXED, we first need starting values for b1, b2, b3, resvar (residual variance), and uvar (random variable variance). The starting values for b1, b2, and b3, were estimated by using properties of the model function. The asymptote of the function is b1 and the time it takes to first reach the asymptote is b2. When the measurements for each plant are plotted against time an approximate asymptote appears to be 200 and the time it takes to reach that asymptote appears to be 1500. We then estimate the starting values as b1 = 200 and b2 = 1500. The starting value for b3 is found by plugging in the mean growth at time t. We chose time t equal to 450. The mean growth across plants at time 450 was around 53. Solving for b3, the starting value for b3 = 1029.27. Once I had approximate estimates for the starting values of b1, b2, and b3, I ran PROC NLIN on the model with no random effect u. The converged values of b1, b2, and b3 are b1=199.7, b2=797.8, and b3=300.7. We will use these values as the starting values. Using PROC NLIN, we estimate the starting value of resvar and uvar to be around 500 each.

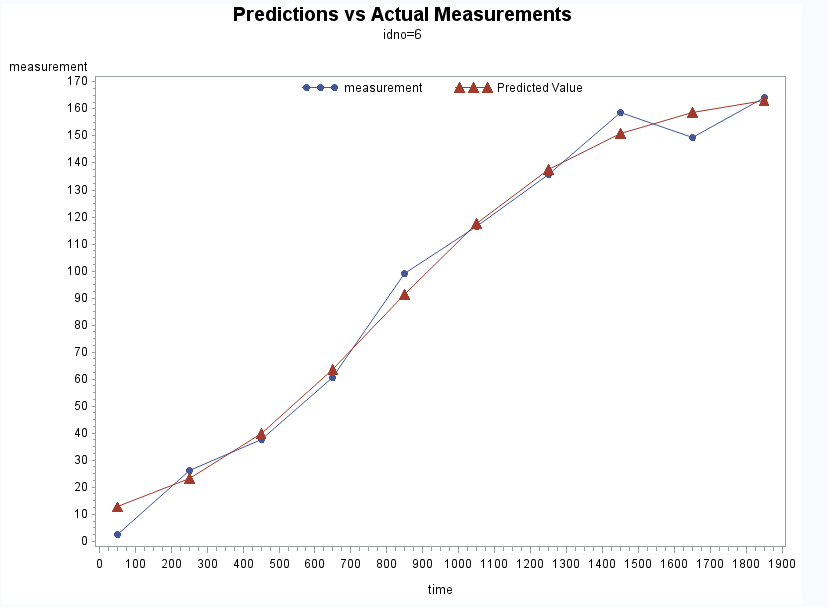
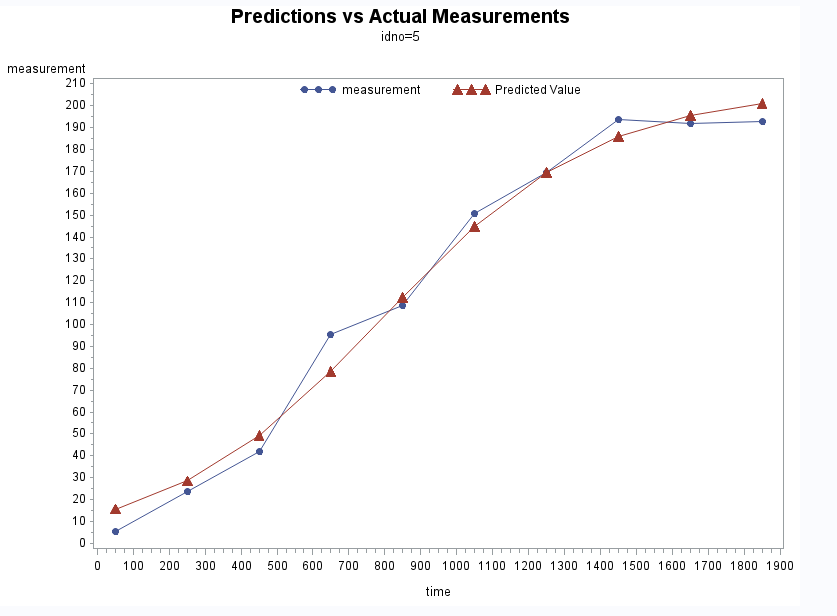
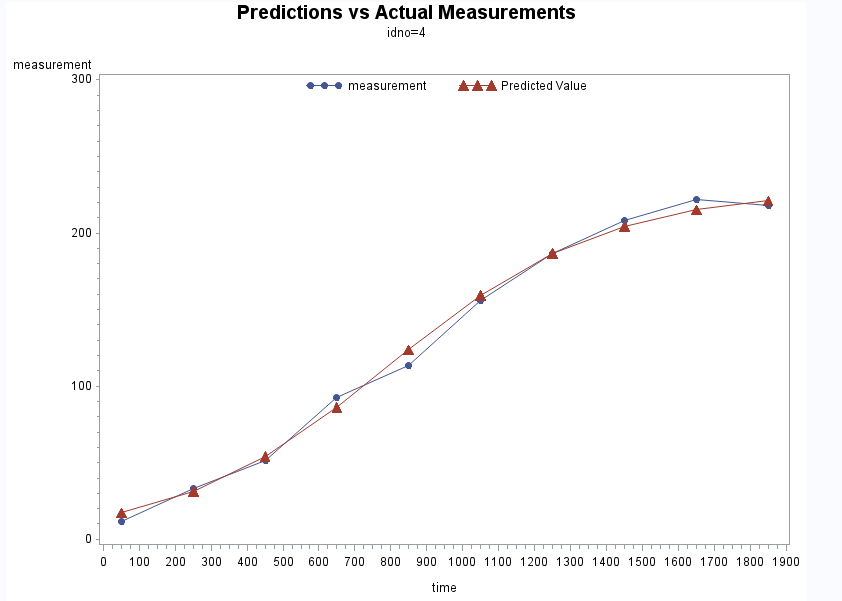
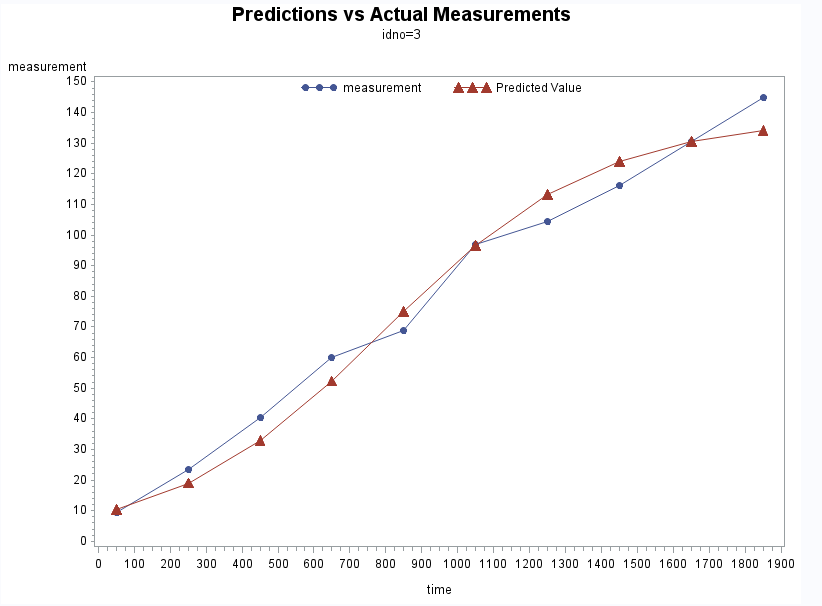
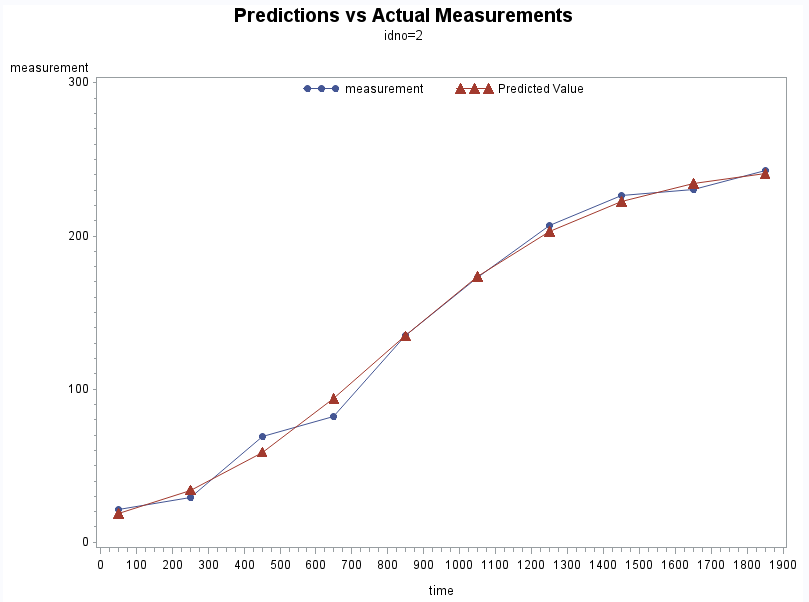
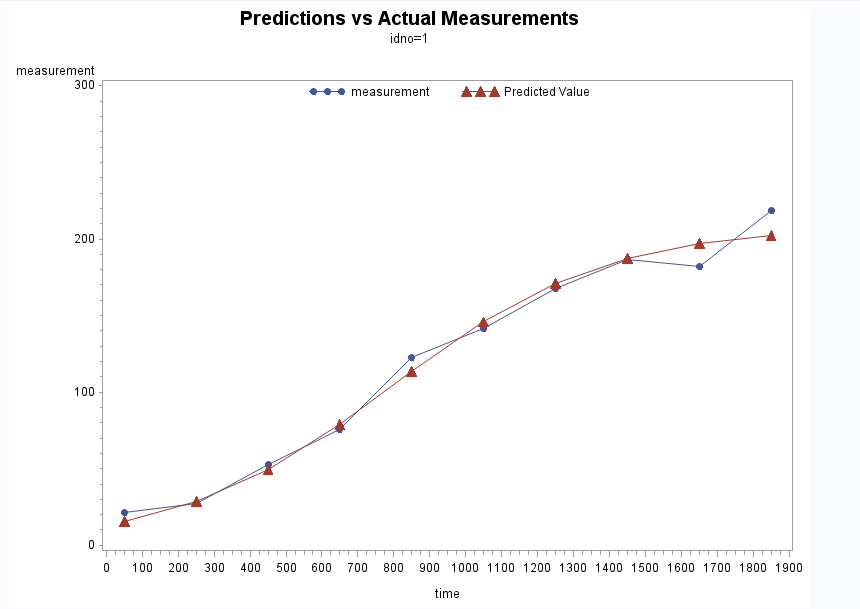
Once I had estimates for my starting values, I ran PROC NLMIXED and found the estimated model.

The estimated = 49.83 and = 1346.94.

The model was tested to see if the random effect was necessary by fitting a reduced model where the random effect equaled zero. The log likelihood ratio test between the full and reduced model was , which is greater than . Therefore, the random effect is useful and “necessary.”

The model was also tested for the null hypothesis that b3=350 by fitting a reduced model where b3=350. The log likelihood ratio test between the full and reduced model was , which is greater than . Therefore, the alternative hypothesis is supported that b3 does not equal 350.

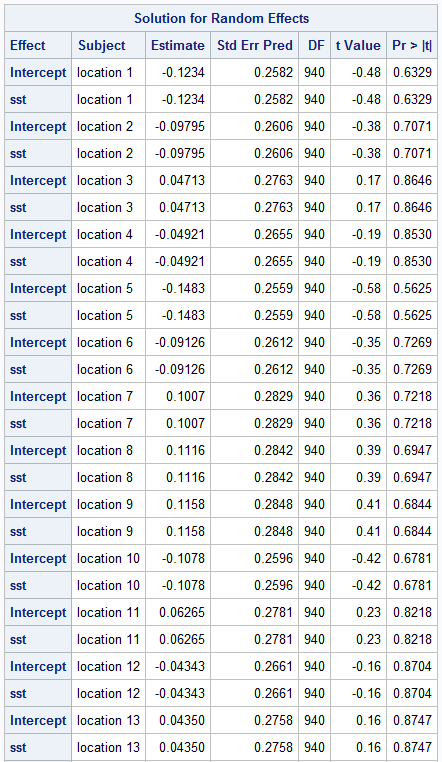
The plots of predicted values and actual measurements are given below. The predicted values are plotted with triangles and the actual values are plotted with circles. The plots are arranged in order, plant 1 to plant 6. The plots show that the model does a pretty good job predicting actual measurements.

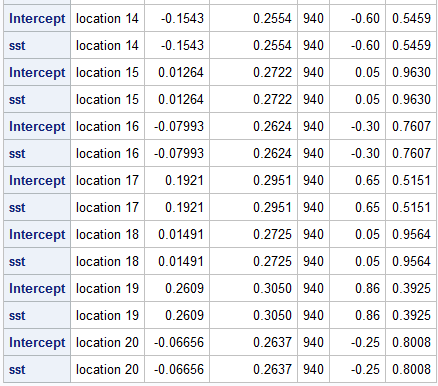


Problem 3

When and are modelled as spatially dependent random effects (Type=sp(exp) (long lat)), PROC NLMIXED results in the following model.

The model mean variance and standard error for the prediction of is 0.04712 and 0.05081. The mean variance for the spatial dependence is 1. The solution for random effect summary is given below.





The summary provides the p-value for the t-test of each predicted random effect, intercept and sst. There are no instances when the p-value is less than .05 for any random effect for a given location. Given these results, both the intercept and sst random effect are not significant.

The data is then modelled without random effects. To make sure there are no interactions, an interaction is also included in the model for sst and location. For the fixed effects, only the intercept and interaction term are significant at the 5% significance level. The data is then modelled accordingly and shown below.

Since there is a different estimate of for each location, the estimates are given in the table below. The conclusion is that sst by itself is not a significant predictor of tornado count, but the interaction of sst and location is a significant predictor. Thus, there are no intercept and sst random effects, but there is a significant fixed interaction between sst and location.

