Module 3: Drying of beech wood - a case study, Part I

3.1 Introduction

This module consists of the first part of a complete analysis of the beech wood data presented as an example in module 2. The aim is to show that the principles for data analysis and result summary for fixed ANOVA and/or regression models also apply for mixed models. And maybe some readers will find it helpful to have some of these principles reviewed.

For completeness we repeat here the description and initial factor structure considerations. To investigate the effect of drying of beech wood on the humidity percentage, the following experiment was conducted. Each of 20 planks was dryed in a certain period of time. Then the humidity percentage was measured in 5 depths and 3 widths for each plank:

depth 1: close to the top
depth 5: in the center
depth 9: close to the bottom
depth 3: between 1 and 5
depth 7: between 5 and 9

width 1: close to the side
width 3: in the center
width 2: between 1 and 3

So there are $3 \cdot 5 = 15$ measurements for each plank and all together 300 observations. The data is can be found as **planks** and is reproduced in the following table.

	Width 1 Depth					Width 2 Depth					Width 3 Depth				
Planks	1	3	5	7	9	1	3	5	7	9	1	3	5	7	9
1	3.4	4.9	5.0	4.9	4.0	4.1	4.7	5.2	4.6	4.3	4.4	4.8	5.0	4.9	4.2
2	4.3	5.5	6.2	5.4	4.7	3.9	5.6	5.7	5.5	4.9	4.0	4.7	4.5	3.9	4.0
3	4.2	5.5	5.6	6.3	4.5	5.4	6.2	6.1	6.4	5.2	4.5	4.9	4.9	4.9	4.4
4	4.4	6.0	7.1	6.9	4.6	4.6	6.1	6.6	6.5	4.7	4.9	5.9	5.8	6.4	4.7
5	3.9	4.7	5.2	5.0	3.7	4.2	5.2	5.4	4.8	3.9	4.0	4.4	4.4	4.1	3.5
6	4.6	5.9	6.3	5.8	4.8	5.9	7.3	6.9	6.9	4.4	5.2	5.7	6.6	6.0	4.0
7	3.9	5.6	6.0	5.3	5.0	4.9	6.9	7.1	6.1	4.5	4.3	5.4	5.9	5.5	4.2
8	3.9	4.5	5.3	5.6	4.7	3.7	4.9	4.8	4.9	4.3	3.8	4.5	5.4	4.8	4.0
9	3.6	4.1	4.0	4.4	3.7	3.8	5.1	5.0	4.6	3.3	3.0	3.9	4.7	4.9	3.8
10	6.5	8.7	9.5	7.9	6.6	6.9	8.9	7.4	7.0	6.9	5.8	7.5	7.7	7.3	5.9
11	3.7	5.2	5.5	5.9	4.4	4.7	5.8	5.7	4.9	4.2	3.7	5.0	6.3	5.2	4.3
12	4.3	5.8	6.2	5.2	4.4	4.8	6.7	7.0	6.1	5.2	5.1	5.7	5.9	6.4	5.1
13	6.5	8.8	9.1	8.9	6.0	5.9	7.5	8.4	7.9	5.7	4.0	4.2	4.9	4.6	3.5
14	4.4	6.2	6.7	6.4	4.3	5.7	7.0	7.4	7.3	5.5	4.6	6.2	6.8	5.8	4.9
15	5.5	7.1	7.5	6.9	5.4	6.4	8.4	8.9	8.1	6.1	6.5	8.4	9.1	9.2	7.5
16	5.2	6.0	6.2	6.6	5.3	6.6	7.6	7.8	7.7	5.8	5.9	6.7	6.7	5.0	3.9
17	3.7	4.5	5.0	4.5	3.7	3.7	4.4	4.8	4.4	4.3	3.7	4.5	4.7	5.3	3.9
18	6.0	7.4	7.8	7.5	5.7	6.9	8.6	8.8	7.5	5.4	5.1	6.1	5.2	5.4	4.7
19	3.8	4.6	4.8	4.4	3.8	3.7	4.7	4.7	4.3	3.7	3.3	3.5	3.7	3.4	3.2
20	6.1	7.4	7.7	6.7	4.6	4.7	6.3	7.1	6.5	5.1	4.7	6.0	6.0	6.3	4.2

In this experiment we have 3 factors apart from the trivial factors `I` and `0`. Let us use the factor names `plank`, `width` and `depth`. The factor `plank` has 20 levels, `width` has 3 and `depth` has 5 levels. For the ith measurement of humidity, $\texttt{plank}_i$ denotes the plank on which this measurement was performed. And correspondingly $\texttt{width}_i$ and $\texttt{depth}_i$ denotes the width and depth, respectively, of this ith measurement.

It would be natural to include the interaction between `width` and `depth` corresponding to the product factor `width` $\times$ `depth`. The product factor has in this case 15 levels.

A natural model would include `plank` as a block factor while `depth` and `width` enter together with their interaction. If Y_i denotes the humidity percentage corresponding to the ith measurement, the model with fixed block effect can be written as:

$$Y_i = \mu + \alpha(\texttt{width}_i) + \beta(\texttt{depth}_i) + \gamma(\texttt{width}_i, \texttt{depth}_i) + \delta(\texttt{plank}_i) + \epsilon_i, \quad (3.1)$$

where $i = 1, \ldots, 300$ and where the ϵ_is are independent and normally distributed random variables. Or similarly:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \delta_k + \epsilon_{ijk}$$

where Y_{ijk} is the kth measurement within the (i, j)th combination of the two factors, $i = 1, \ldots, 3$, $j = 1, \ldots, 5$ and $k = 1, \ldots, 20$. As pointed out in Module 1 the block (plank) effect should be considered as a random effect, leading to the mixed model:

$$Y_i = \mu + \alpha(\texttt{width}_i) + \beta(\texttt{depth}_i) + \gamma(\texttt{width}_i, \texttt{depth}_i) + d(\texttt{plank}_i) + \epsilon_i, \quad (3.2)$$

where $d(\texttt{plank}_i) \sim N(0, \sigma^2_{Plank})$ and $\epsilon_{ijk} \sim N(0, \sigma^2)$. This model corresponds to the factor structure diagram given in figure 3.1.

3.2 Initial explorative analysis

Having realized the complete structure of the data, it is time to do initial plotting/ explorative analysis. Throughout this module, figures and results are presented without showing SAS code or raw SAS output. This can be seen as a standard for reports in the course! Typically, numerous figures not entering a final project report should be studied, since this phase is explorative, and final figures to present the key results are chosen after the statistical analysis is completed. The SAS information may be followed in chronological order in **the sas section**.

The plotting of various average profiles is usually a helpful tool for data with several factors. In figure 3.2 four of these are presented. In the top left diagram the width humidity patterns for each plank is depicted by plotting the average humidity (taking the average of the five depths for each width and plank) against the widths.

It is immediately clear that there is extensive plank-to-plank variations in the level of humidity. The message about the width effect is less clear. In the top right the similar plot for the depth effect is seen. Here the message is much clearer: The humidity is high in the center (depth=5) and low at the top (depth=1) and at the bottom (depth=9). As pointed out, this is the effect seen when the three widths are averaged. It could

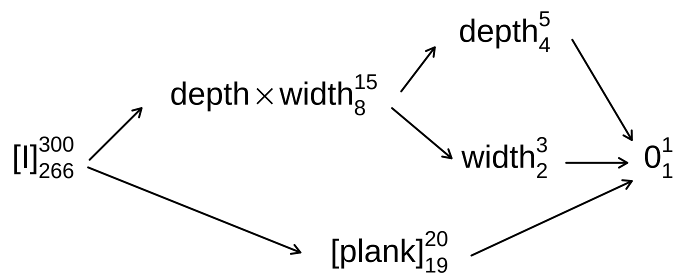

Figure 3.1: The factor structure diagram

be that the depth effect is different for widths close to the side of the plank (width=1) than for widths in the center (width=3). In other words, there could be a plank*width interaction effect, that we wouldn't find in the plots above. Instead similar plots are given in the bottom diagrams of figure 3.2 for the widths and depths by averaging over the planks (that is, plotting the 15 average values).

The depth structure already seen is recognized. Also, it is seen that there is a clear shift in humidity level from width to width and that the depth humidity pattern seems to be roughly the same for the three widths. However, there are some deviations from parallel patterns and the uncertainties in the deviations from parallel patterns are not visible. A similar increasing-decreasing width pattern, that was not clearly visible from the top diagram is now seen. This pattern seems to be roughly the same for all depths (with the same precautions as before) and the low humidity levels for the top and bottom depths are clearly seen. Note again that the two bottom plots contain the same information: had there been clearly non-parallel patterns in one figure (an interaction effect) this would also appear in the other figure. The next step is to start the actual statistical analysis of the data.

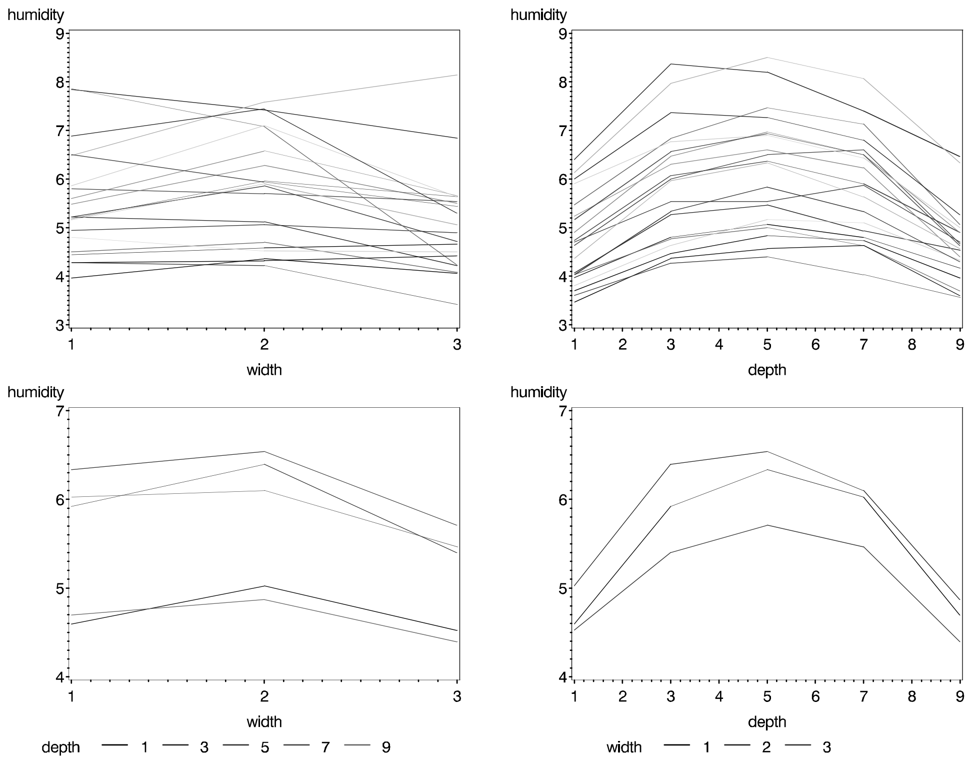

Figure 3.2: Four average humidity profiles

3.3 Test of overall effects/model reduction

A statistical analysis of this kind is commonly carried out in several steps, starting with the basic model found from the factor structure considerations. This model usually contains every possible effect there may be in the data. However, it is of interest to simplify things into easily interpretable results, if possible! So, the idea is to remove non-signifcant ”complex stuff” from the model before summarizing the results.

Carrying out the mixed model analysis corresponding to the model given by (3.2) gives the following ANOVA table of fixed effects:

Source of variation	Numerator degrees of freedom	Denominator degrees of freedom	F-statistics	P-values
depth	4	266	78.26	<0.0001
width	2	266	29.65	<0.0001
depth*width	8	266	1.08	0.3745

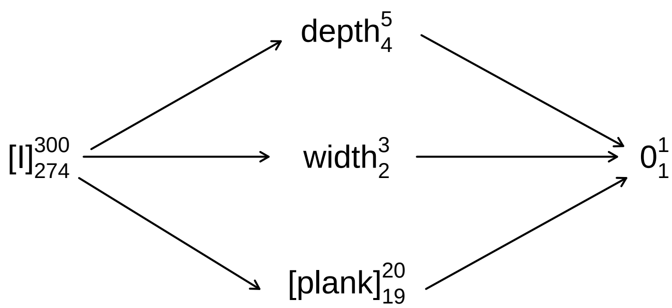

Figure 3.3: The factor structure diagram

We see, that the depth*width interaction effect is non-significant. Hence, we remove the interaction term and do the analysis based on the model:

$$Y_i = \mu + \alpha(\texttt{width}_i) + \beta(\texttt{depth}_i) + d(\texttt{plank}_i) + \epsilon_i, \tag{3.3}$$

where $d(\texttt{plank}_i) \sim N(0, \sigma^2_{Plank})$ and $\epsilon_i \sim N(0, \sigma^2)$. This model is illustrated by the factor structure diagram in figure 3.3.

Note how the 8 degrees of freedom from the interaction effect has now been added to the error degrees of freedom. The table of fixed effects then becomes:

Source of variation	Numerator degrees of freedom	Denominator degrees of freedom	F-statistics	P-values
depth	4	274	78.07	<0.0001
width	2	274	29.57	<0.0001

Note that the removal of the non-significant interaction effect only has minor effects on the conclusions regarding the depth and width effects: They are both extremely significant, confirming what we "explored" above. Since there are no more non-significant fixed effects, the model given by 3.3 is the final model to use for summarizing the results.

3.4 Post hoc analysis and summarizing the results

3.4.1 Estimates of the variance parameters

The final model is given by (3.3), since main effects of as well width as depth are clearly significant. Estimates of the two variance parameters are:

$$\widehat{\sigma}^2_{Planks} = 0.9797, \quad \widehat{\sigma}^2 = 0.4047$$

Uncertainties of these estimates are not automatically provided by SAS. In a later module we will see how confidence intervals for variance parameters can be constructed.

The remaining part of this subsection on post-hoc analysis and presentation of results illustrates how the information in factors can be summarized whenever the factor does not interact with any other factor.

3.4.2 Estimates of the fixed parameters

Estimates of the expected values (LSMEANS) for each level of depth, together with their uncertainties and 95% confidence intervals are:

	Parameter	Estimate	SE	Lower	Upper
Depth 1	$\mu + \beta(1)$	4.7150	0.2361	4.2270	5.2030
Depth 3	$\mu + \beta(2)$	5.9050	0.2361	5.4170	6.3930
Depth 5	$\mu + \beta(3)$	6.1950	0.2361	5.7070	6.6830
Depth 7	$\mu + \beta(4)$	5.8633	0.2361	5.3753	6.3514
Depth 9	$\mu + \beta(5)$	4.6533	0.2361	4.1653	5.1414

and correspondingly for each level of width:

	Parameter	Estimate	SE	Lower	Upper
Width 1	$\mu + \alpha(1)$	5.5140	0.2303	5.0352	5.9928
Width 2	$\mu + \alpha(2)$	5.7860	0.2303	5.3072	6.2648
Width 3	$\mu + \alpha(3)$	5.0990	0.2303	4.6202	5.5778

3.4.3 Comparisons of the fixed parameters

A commonly used post hoc analysis is to compare either specific pairs of depths (resp. widths) or compare all combinations within each factor. For the former, a standard t-tests can be used, e.g.

$$t = \frac{\widehat{\beta}(1) - \widehat{\beta}(2)}{SE\left(\widehat{\beta}(1) - \widehat{\beta}(2)\right)}$$

using the error degrees of freedom (274). Or equivalently expressed by a 95% confidence interval:

$$\widehat{\beta}(1) - \widehat{\beta}(2) \pm t_{.975,274} SE\left(\widehat{\beta}(1) - \widehat{\beta}(2)\right)$$

In this case, the estimates of the fixed effects are raw averages of the data based on the same number of observations for each level, so the standard error of the difference between two depth levels is given by

$$SE\left(\widehat{\beta}(1) - \widehat{\beta}(2)\right) = \sqrt{2}\sqrt{\widehat{\sigma}^2/60}$$

This means that two depth levels are claimed signifcantly different if they differ by more than

$$t_{.975,274}\sqrt{2}\sqrt{\widehat{\sigma}^2/60}$$

from each other. This is also called the *95% Least Significant Difference (LSD)* value.

It would be tempting to do such tests for all combinations of levels within each factor. This is generally NOT an acceptable approach, since the probability of "significance-by-chance" becomes too large when many tests are performed simultaneously. This is called the "multiplicity problem". With five depth levels there are $5 \times 4/2 = 10$ possible depth pairs to compare. Comparing two specific (decided before seeing the data) levels is not the same as comparing the smallest among five with the largest among five. In a case with no effects one would always expect the latter two to be more different by chance than the former.

There are numerous solutions to properly handle this problem, if all comparisons indeed are made. All of them amounts to requiring differences to be larger than required by the usual t-test to be claimed significant. One general idea, that can be used whenever numerous tests are performed simultaneously, is the *Bonferroni* correction: If k tests are performed simultaneously, then use level α/k in each test rather than α. For instance, if all depth levels are compared, standard pair-wise t-test output (using the PDIFF option of the LSMEANS statement in SAS) can be used, but employing level 0.5% in each test rather than 5%: So only claiming those differences significant for which the usual P-value is less than 0.005. This method is known to be somewhat conservative, meaning that it may be too critical, or in other words again: it may miss some actual differences.

Another solution is to use another distribution than the t-distribution, when comparisons are made. With the so-called *Tukey-Kramer* method two depth levels would be claimed signifcantly different if they differ by more than

$$\nu_{.975,J,274}\sqrt{\widehat{\sigma}^2/60}$$

from each other, where J is the number of groups to be compared and $\nu_{0.975,J,274}$ is the 97.5%-quantile of the so-called “studentized range” distribution with J groups. This distribution takes into account that the two levels that we compare in a single test is coming from J groups all together. This distribution is, just like the t-distribution, tabulated or ”available” in the computer. Note that if $J = 2$, then the studentized range distribution corresponds to the t-distribution,

$$\nu_{.975,2,274} = t_{.975,274}\sqrt{2}$$

We just use the option ADJUST=TUKEY in the LSMEANS statement of SAS to get the corrected confidence bands and corrected p-values:

Depth difference	Parameter	Estimate	SE	Lower	Upper	P-value
1-3	$\beta(1) - \beta(2)$	-1.1900	0.1162	-1.5090	-0.8710	<0.0001
1-5	$\beta(1) - \beta(3)$	-1.4800	0.1162	-1.7990	-1.1610	<0.0001
1-7	$\beta(1) - \beta(4)$	-1.1483	0.1162	-1.4673	-0.8294	<0.0001
1-9	$\beta(1) - \beta(5)$	0.06167	0.1162	-0.2573	0.3806	0.9841
3-5	$\beta(2) - \beta(3)$	-0.2900	0.1162	-0.6090	0.02896	0.0943
3-7	$\beta(2) - \beta(4)$	0.04167	0.1162	-0.2773	0.3606	0.9964
3-9	$\beta(2) - \beta(5)$	1.2517	0.1162	0.9327	1.5706	<0.0001
5-7	$\beta(3) - \beta(4)$	0.3317	0.1162	0.01271	0.6506	0.0370
5-9	$\beta(3) - \beta(5)$	1.5417	0.1162	1.2227	1.8606	<0.0001
7-9	$\beta(4) - \beta(5)$	1.2100	0.1162	0.8910	1.5290	<0.0001

Note that since the P-values are ”corrected”, that is, based on the more proper studentized range distribution, they can be used directly without any additional Bonferroni correction. Similarly for the width effect:

Width difference	Parameter	Estimate	SE	Lower	Upper	P-value
1-2	$\alpha(1) - \alpha(2)$	-0.2720	0.08997	-0.4840	-0.05998	0.0077
1-3	$\alpha(1) - \alpha(3)$	0.4150	0.08997	0.2030	0.6270	<0.0001
2-3	$\alpha(2) - \alpha(3)$	0.6870	0.08997	0.4750	0.8990	<0.0001

Frequently, the key information of the two tables for each effect is summarized into a single table in which the lsmeans are ordered by size:

	Estimate
Depth 9	4.6533_a
Depth 1	4.7150_a
Depth 7	5.8633_b
Depth 3	5.9050_{bc}
Depth 5	6.1950_c

The letter subscripts express the 5% significance results of the 10 pair-wise comparisons:

- Two depths sharing a subscript are NOT significantly different
- Two depths NOT sharing a subscript are significantly different

So the pattern already observed in Figure 3.2 can now be statistically confirmed: there is a clear lower humidity close to the top and the bottom (and no difference between top and bottom). Also there is an indication that the center position has significantly higher humidity than the in between positions (between which no difference is seen).

For the width effect, the summary table becomes particularly simple, since all three differences are significant:

	Estimate
Width 3	5.0990_a
Width 1	5.5140_b
Width 2	5.7860_c

For these data, a figure of the raw data, like one of the bottom plots of figure 3.2 together with a statement of the lack of significant width*depth interaction and the two summary tables would probably suffice for most purposes. In later modules we will see how additional plots of the model expectations/details will provide informative figures for interpretation.

Other types (than the multiple comparison approach) of post hoc analysis may be employed, especially when quantitative information about the factor levels are available. In this case we know exactly the positions that corresponds to the different widths and depths and this could be used in the analysis. For instance, it could be investigated whether a quadratic function of the depths could be used to describe the humidity pattern. Apart from the nice direct functional interpretation of the dependence of humidity on depth, it could possibly provide more powerful tests for interaction effects. In fact this would still be a ”linear” model, and could be handled by PROC MIXED. We will return to such analyzes in a later module. Non-linear models (using e.g. exponentials etc) could also be an option in some cases, but then the model will no longer be a linear model, and additional theory and software packages would be needed.

The summary approach above was based on the assumption of no interaction between width and depth, that is, the conclusions regarding widths hold for all the depths, and vice versa. Had there been a significant interaction, we would have to present, say, the depth effects for each of the three widths (and/or vice versa), since the significance tells us that these three conclusions will NOT be the same. In practice, we proceed as above, BUT for the combined width*depth factor with 15 levels rather than for each of them separately. We will see examples of this later.

One important step in the analysis given is missing: An investigation of the validity of the model assumptions! We return to this issue in Module 6, where we then finish the analysis of this data set on the humidity of beech wood planks.