02424 - Advanced Dataanalysis and Statistical Modelling
Assingment 3

Tymoteusz Barcinski - s221937
Soren Skjernaa - s223316

May 10, 2023

Contents

1 Introduction **2**

2 Mixed effects models **2**

2.1 Description of experiment and an initial look at the data . . . 2
2.2 Previous findings from fitting a generalized linear gamma model . . . 4
2.3 Mixed linear model using subjects as random effects . . . 4
2.3.1 Model description . . . 4
2.3.2 Model diagnostics . . . 5
2.3.3 Model reduction . . . 6
2.3.4 Model presentation . . . 7
2.4 Mixed linear model using the nested random effects subjects and subDay . . . 8
2.4.1 Model description . . . 8
2.4.2 Model diagnostics . . . 8
2.4.3 Model reduction . . . 9
2.4.4 Model presentation . . . 10
2.5 Repeated measurement model . . . 11
2.5.1 Model description . . . 11
2.5.2 Model diagnostics . . . 12
2.5.3 Model reduction . . . 12
2.5.4 Model presentation . . . 13
2.6 Interpretation of the final model . . . 14
2.7 Conclusion . . . 15

3 Hierarchical models: Random varaiance **16**

3.1 First model . . . 16
3.2 Second model . . . 17
3.3 Third model . . . 18
3.4 Proof of the reformulated Theorem 6.7 . . . 19
3.5 Hierarchical model . . . 20
3.5.1 Derivation of $\gamma_i|Y_i$ . . . 23
3.6 Final Normal model . . . 23
3.7 Comparison between models . . . 25

4 References **28**

1 Introduction

The purpose of the following report is to establish a model for the variation in clothing insulation level (clo), measured at an experiment conducted at the Laboratory of Occupant Behavior, Satisfaction Thermal comfort and Environmental Research (LOBSTER).

Having previously fitted generalized linear gamma models to the data set, we will in this report investigate whether a linear mixed effects model, or a hierachical model explains the variance in the data better than the fixed effects model.

All tests in the following two sections are carried out using a significance level of $\alpha = 0.05$.

2 Mixed effects models

2.1 Description of experiment and an initial look at the data

The experiment was carried out at the LOBSTER, where the clo was measured for 47 different subjects over multiple days. Each subject participated for 1-4 days (with most subjects participating for 3 days), and had 2-6 measurements recorded each day (with most having 6 recordings per day). All in all the data set consists of 803 observations with no missing data. The following variables were recorded for each observation:

Variable	**Type**	**Explanation**
clo	Clothing insulation level	Positive variable, with higher values implying higher insulation.
tOut	Outdoors air temperature	Measured in C^o
tInOp	Indoor operating temperature	Measured in C^o
sex	Sex	Female/male.
subjId	Subject ID	Uniqiue ID for each subject.
time	Time	Time difference since last observation for the subject (continues variable, but unit not given).
day	Day	Number of experimentation day for the subject.
subDay	Subject $\times$ Day ID	Unique ID for each combination of subject and day.

Table 1: Overview of the variables recorded.

The main variables of interest are the outcome clo, and how it is affected by the temperatures and the sex of the subject. We first present plots of the data. Note that some of these plots and the points made are a repeat of the previous assignment.

Figure 1 shows how the outcome, clo is distributed for different levels of the independent variables. The top-left plot shows that there is a large subject-to-subject variation in the distribution of clo. We also see that for some subjects, there is no within-subject variation, while for others there is a large within-subject variation. From all four plots, we clearly see that there is a larger variation for the clo measurements for female subjects than for male subjects. Furthermore, it seems like the clo level is lower for the male subjects compared to the female subjects. From the two bottom plots, it is not clear if clo depends on the two temperature measurements, although there might be a tendency of clo decreasing when the temperatures increase.

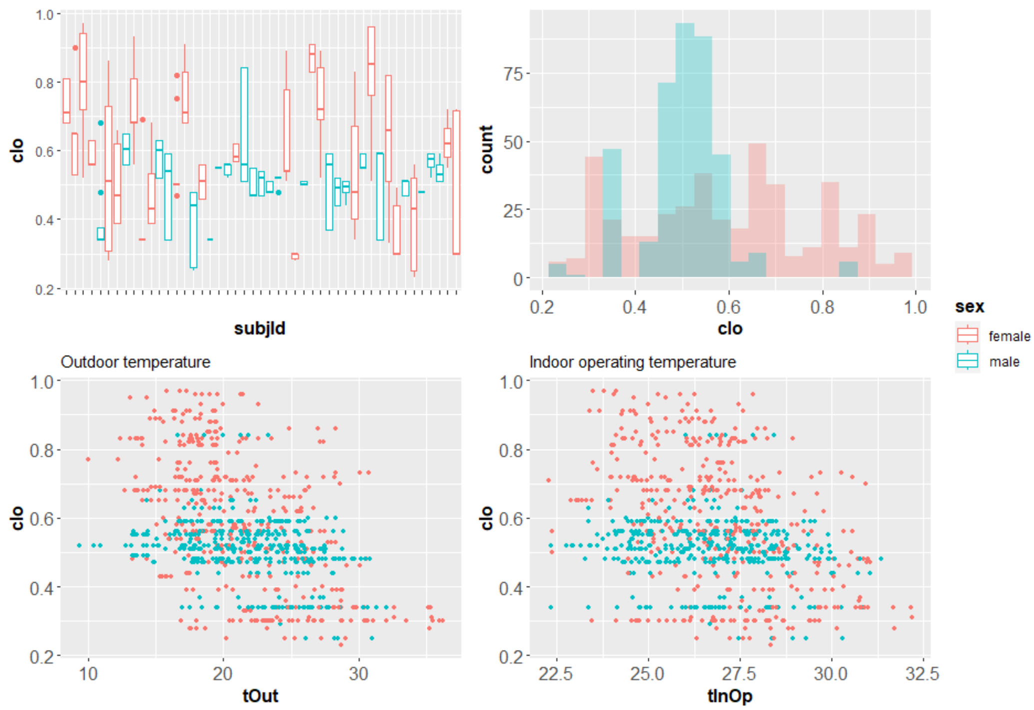

Figure 1: Plots of clo against the different independent variables.

Figure 2 shows the variation in clo for the different combinations of subjects and days. Adjacent boxes colored with the same color represents observations from the same subject on different experimental days. We see that for most combinations of subjects and days, the variation within the day is very small. Furthermore, we see that some subjects may have a small variance in their clo level for one day, but a large variance in their clo level for other days.

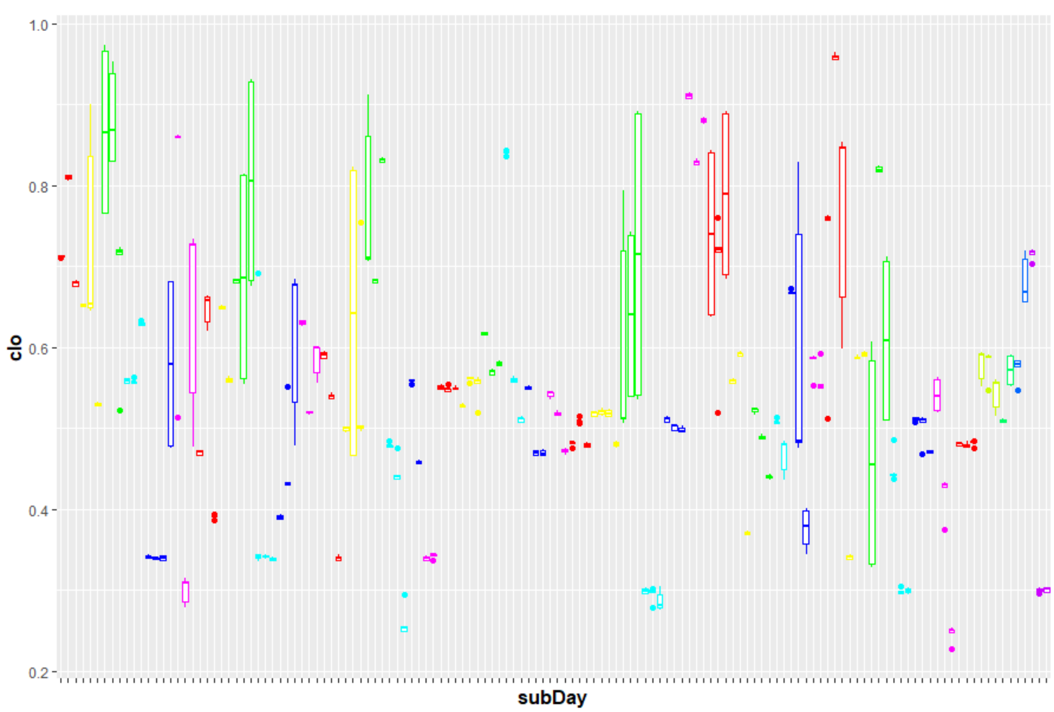

Figure 2: Plots of clo against subDay.

2.2 Previous findings from fitting a generalized linear gamma model

We have previously fitted a generalized linear gamma model to the data of the form $Y_i \sim \text{Gamma}\left(\phi, \frac{\mu_i}{\phi}\right)$ with the inverse link $\frac{1}{\mu_i} = \eta_i$ and a constant precision parameter $\phi > 0$. Here we used a linear predictor of the form

$$
\begin{aligned}
\eta_i = \beta_0 &+ \beta_1(\text{sex}_i) \\
&+ \beta_2(\text{sex}_i)\cdot \text{tOut}_i + \beta_3(\text{sex}_i)\cdot \text{tOut}_i^2 \\
&+ \beta_4(\text{sex}_i)\cdot \text{tInOp}_i + \beta_5(\text{sex}_i)\cdot \text{tInOp}_i^2 \\
&+ \beta_6(\text{sex}_i)\cdot \text{tOut}_i \cdot \text{tInOp}_i \\
&+ \beta_7(\text{sex}_i)\cdot \text{tOut}_i \cdot \text{tInOp}_i^2 \\
&+ \beta_8(\text{sex}_i)\cdot \text{tOut}_i^2 \cdot \text{tInOp}_i \\
&+ \beta_9(\text{sex}_i)\cdot \text{tOut}_i^2 \cdot \text{tInOp}_i^2.
\end{aligned} \tag{1}
$$

We found that the we could not fully reduce the three-way interactions between sex and the two temperatures, and that the polynomials were also significant. Furthermore we fitted the corresponding model with a similar linear predictor, but based on the subjects instead of sex. In this model we found significant auto-correlation between observations on the same subject within each day. This finding is not surprising as the experiment deals with repeated measurements within subjects. This correlation within combinations of subjects and days is also observed in figure 2.

2.3 Mixed linear model using subjects as random effects

2.3.1 Model description

Based on the previous findings we chose to start our modelling with a linear predictor based on (1). As we are not interested in the results of modelling specific subjects but more in taking the variation between subjects into account, we choose to view the subjects as being randomly drawn from a larger sample population. This leads us to use a mixed effect model. In this section we will only consider random effects for each subject, while we in later sections consider the additional correlation of observations within each subject and day combination.

As each subject may have a different reaction to temperature increases, we choose to include random effects structure based on subjId for both the intercept and the temperatures. This gives us a random effects structure of the following form

$$
\phi_i = u_1(\text{subjId}_i) + u_2(\text{subjId}_i)\cdot \text{tOut}_i + u_3(\text{subjId}_i)\cdot \text{tOut}_i^2 + u_4(\text{subjId}_i)\cdot \text{tInOp}_i + u_5(\text{subjId}_i)\cdot \text{tInOp}_i^2 \tag{2}
$$

where the random effects are assumed independent between subjects, but we allow for a covariance of the random effects within a subject. Thus we have the following distributional assumption for the random effects.

$$
\begin{pmatrix} u_1(\text{subjId}_i) \\ u_2(\text{subjId}_i) \\ u_3(\text{subjId}_i) \\ u_4(\text{subjId}_i) \\ u_5(\text{subjId}_i) \end{pmatrix} \sim N\left(0, \begin{pmatrix} \sigma_{u_1}^2 & \sigma_{u_1u_2} & \sigma_{u_1u_3} & \sigma_{u_1u_4} & \sigma_{u_1u_5} \\ \sigma_{u_1u_2} & \sigma_{u_2}^2 & \sigma_{u_2u_3} & \sigma_{u_2u_4} & \sigma_{u_2u_5} \\ \sigma_{u_1u_3} & \sigma_{u_2u_3} & \sigma_{u_3}^2 & \sigma_{u_3u_4} & \sigma_{u_3u_5} \\ \sigma_{u_1u_4} & \sigma_{u_2u_4} & \sigma_{u_3u_4} & \sigma_{u_4}^2 & \sigma_{u_4u_5} \\ \sigma_{u_1u_5} & \sigma_{u_2u_5} & \sigma_{u_3u_5} & \sigma_{u_4u_5} & \sigma_{u_5}^2 \end{pmatrix}\right)
$$

As clo is a positive variable, we choose to log transform the variable, to move its support to the entire real line. Thus our initial model has the form

$$
\log(\text{clo}_i) = \eta_i + \phi_i + \epsilon_i
$$

where $i = 1, 2, \ldots, 803$ and $\epsilon_i \sim N(0, \sigma^2)$ are assumed independent from each other and from all the random effects.

2.3.2 Model diagnostics

We first analyse the distribution of the conditional residuals

$$\hat{\epsilon}_i = \log(\text{clo}_i) - \eta_i - \phi_i$$

The model diagnostic plots based on these are seen in figure 3. For all six plots, the observations belonging to males and female are clearly marked, with black observations being female and red observations being male. Looking at the five plots we draw the following conclusions (plots are discussed row by row):

1. Looking at plot 1 we see a clear shape to the distribution of the residuals, with residuals falling on diagonal lines. Looking closer into the residuals we have found that residuals coming from the same subject an day combination falls on the same diagonal lines. Thus we clearly need to model a structure depending on subDay. We return to that in the following section. Furthermore it is seen that an increase in the fitted value, leads to a decrease in the residuals.

2. For the QQ-plot we see no evidence against assuming a normal distribution of the conditional residuals.

3. From plot 3, it is clear that there is a difference in size of the residual variance for males and females. Therefore, in order to improve the model we could include a sex specific variance. We look into this when we fit the hierachical models.

4. The model used for clo dependence on tInOp, seems to contain no systematic errors, as the residuals are evenly distributed.

5. The same conclusion holds for clo dependence on tOut.

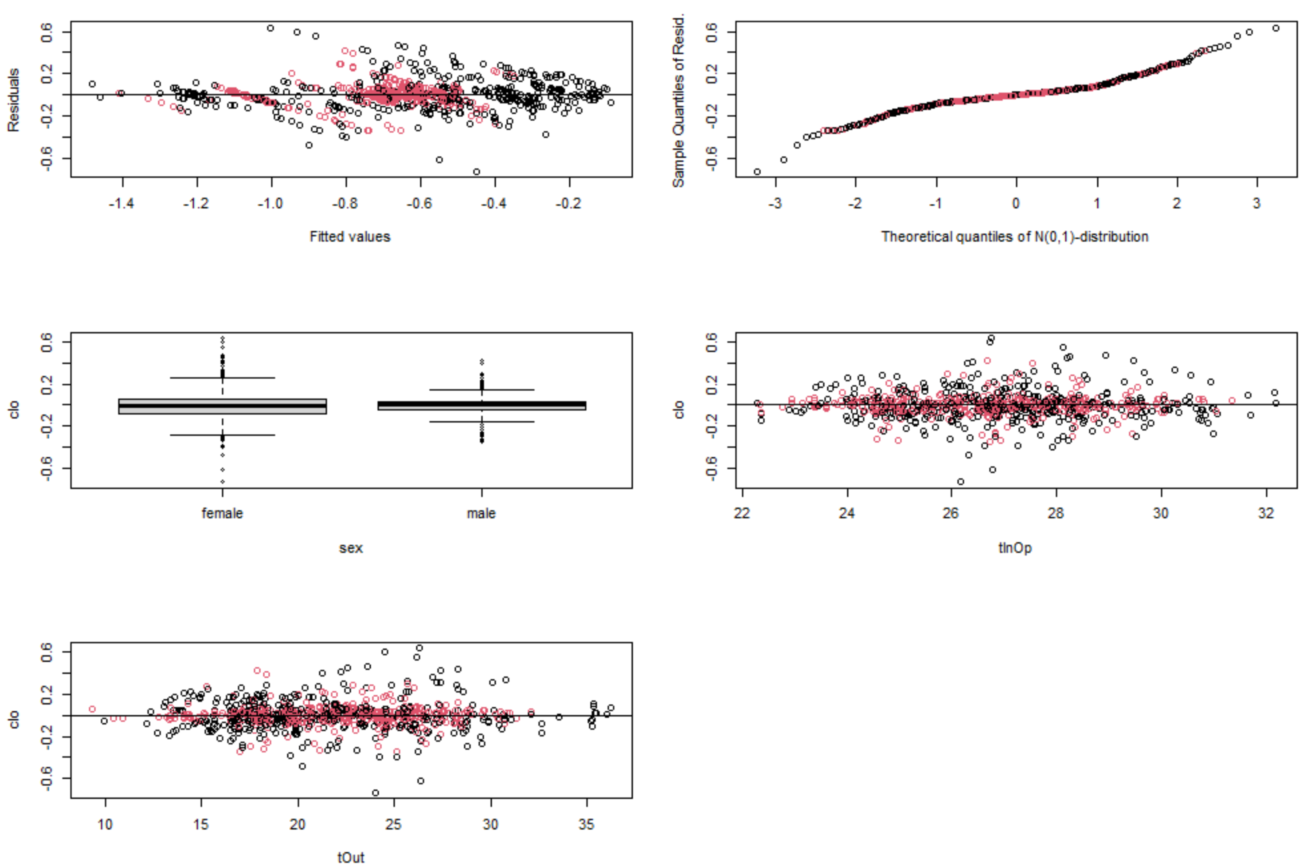

Figure 3: Diagnostics plots for model 1.

Figure 4 shows a QQ-plot of the random effects. We do not see any clear indications against assuming a normal distribution for the random effects.

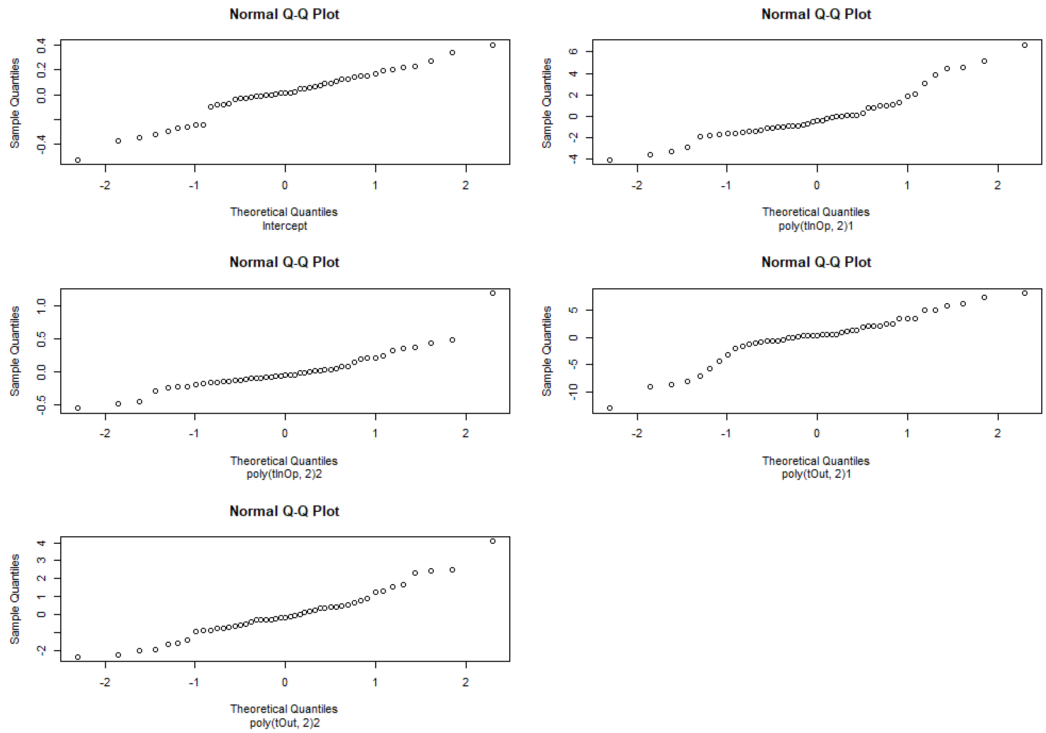

Figure 4: QQ-plots of the estimated random effects.

2.3.3 Model reduction

We first investigate if it is possible to reduce the variance structure. We do this by fitting the initial model, and its corresponding sub-models using the restricted-maximum-likelihood (REML) method. Using the found restricted-log-likelihoods l_{re} we compare the models using the likelihood ratio test (LRT). Trying to reduce the initial model we get the following ANOVA-table

Removed term	N Parameters	AIC_{re}	BIC_{re}	l_{re}	Df	p-value
Initial model	34	-583.94	-424.54	325.97		
$tOut^2 \mid subjId$	29	-515.24	-379.28	286.62	5	< 0.001
$tInOp^2 \mid subjId$	29	-591.21	-455.25	324.61	5	0.742

Table 2: ANOVA table for reducing random effects.

Based on this we remove the random effect $u_5(\text{subjId}_i) \cdot \text{tInOp}_i^2$ and the covariances corresponding to it. Continuing with the model reduction we get the following ANOVA-table

Removed term	N Parameters	AIC_{re}	BIC_{re}	l_{re}	Df	p-value
Reduced model	29	-591.21	-455.25	324.61		
$tInOp \mid subjId$	25	-537.87	-420.67	293.94	4	< 0.001
$tOut^2 \mid subjId$	25	-581.49	-464.85	315.75	4	0.001

Table 3: ANOVA table for reducing random effects.

so no further reduction of the model can be done.

Having established our random effects structure, we continue on to reduce the fixed effects structure. Reducing the fixed effect structure we use models fitted with maximum-likelihood estimates and based on the standard-log-likelihoods l we use the LRT with a type II partitioning of the model deviance. The following ANOVA-table shows the reductions we have done. The effects are listed in the order we removed the terms, starting from the initial model based on the found random effects structure down to the final model.

Model	N Parameters	AIC	BIC	l	Df	p-value
Initial model	29	-548.64	-412.68	303.32		
All three-way-interactions removed	25	-553.69	-436.48	301.85	4	0.567
All sex $\times$ tInOp interactions removed	23	-557.47	-449.64	301.74	2	0.897
All sex $\times$ tOut interactions removed	21	-556.35	-457.89	299.17	2	0.077
All tInOp $\times$ tOut interactions removed	17	-555.99	-476.29	295.00	4	0.079
$\beta_5\cdot$ tInOp ^2_i removed	16	-557.84	-482.82	294.92	1	0.694

Table 4: ANOVA table for reducing fixed effects.

Following the convention that we do not remove fixed effects if their corresponding random effects are still in the model, we deduce that no further model reductions are possible. The reasoning behind the convention is that random effect such as $u_1(\mathrm{subjId}_i)$ provides a random intercept for each subject. As we assume it is distributed with mean zero, removing the fixed intercept would only work to shift the mean of u_1.

2.3.4 Model presentation

Our final model is given by the equation

$$\begin{aligned}\log(\mathrm{clo}_i) =&\beta_0 + \beta_1(\mathrm{sex}_i) + \beta_2(\mathrm{sex}_i)\cdot \mathrm{tOut}_i + \beta_3(\mathrm{sex}_i)\cdot \mathrm{tOut}_i^2 + \beta_4(\mathrm{sex}_i)\cdot \mathrm{tInOp}_i + \\ &u_1(\mathrm{subjId}_i) + u_2(\mathrm{subjId}_i)\cdot \mathrm{tOut}_i + u_3(\mathrm{subjId}_i)\cdot \mathrm{tOut}_i^2 + u_4(\mathrm{subjId}_i)\cdot \mathrm{tInOp}_i + \epsilon_i\end{aligned}$$

with

$$\begin{pmatrix} u_1(\mathrm{subjId}_i) \\ u_2(\mathrm{subjId}_i) \\ u_3(\mathrm{subjId}_i) \\ u_4(\mathrm{subjId}_i)\end{pmatrix} \sim N\left(0, \begin{pmatrix} \sigma^2_{u_1} & \sigma_{u_1u_2} & \sigma_{u_1u_3} & \sigma_{u_1u_4} \\ \sigma_{u_1u_2} & \sigma^2_{u_2} & \sigma_{u_2u_3} & \sigma_{u_2u_4} \\ \sigma_{u_1u_3} & \sigma_{u_2u_3} & \sigma^2_{u_3} & \sigma_{u_3u_4} \\ \sigma_{u_1u_4} & \sigma_{u_2u_4} & \sigma_{u_3u_4} & \sigma^2_{u_4}\end{pmatrix}\right) \quad , \quad \epsilon_i \sim N(0,\sigma^2).$$

The errors and all random effects are assumed pairwise independent. Furthermore the random effects are assumed independent between different subjects. Table 5 and 6 presents the model parameters estimated using REML, together with 95% profile confidence intervals.

Parameter	Estimate	Std. Error	Lower CI (2.5 %)	Upper CI (97.5 %)
(Intercept)	-0.58	0.05	-0.67	-0.49
sexmale	-0.11	0.07	-0.25	0.02
poly(tInOp, 1)	0.98	0.50	-0.00	1.96
poly(tOut, 2)1	-3.18	0.74	-4.64	-1.72
poly(tOut, 2)2	-0.63	0.42	-1.46	0.20

Table 5: Model parameters and their uncertainties for the final model.

Std. deviations and correlations	Lower CI (2.5 %)	Estimate	Upper CI (97.5 %)
$\tilde{\sigma}$	0.13	0.14	0.15
sd((Intercept))	0.18	0.23	0.29
sd(poly(tInOp, 1))	2.32	3.02	3.94
sd(poly(tOut, 2)1)	3.52	4.57	5.93
sd(poly(tOut, 2)2)	1.07	2.05	3.91
cor((Intercept),poly(tInOp, 1))	-0.56	-0.26	0.09
cor((Intercept),poly(tOut, 2)1)	-0.06	0.29	0.58
cor((Intercept),poly(tOut, 2)2)	-0.64	-0.15	0.42
cor(poly(tInOp, 1),poly(tOut, 2)1)	-1.00	-0.98	0.44
cor(poly(tInOp, 1),poly(tOut, 2)2)	-0.38	-0.02	0.34
cor(poly(tOut, 2)1,poly(tOut, 2)2)	-0.51	-0.16	0.24

Table 6: Estimate of the variance parameters in the final model.

2.4 Mixed linear model using the nested random effects subjects and subDay

2.4.1 Model description

As previously mentioned, Figure 2 and the model diagnostics from the mixed effects model with a random subjId effect, shows a correlation of observations within the same subject×day combination. Thus we fit a linear mixed effects model to the data using the random effects subjId and subDay. Note that subDay is nested within subjId, i.e. subjId is a coarser factor than subDay.

Our starting point for the fixed effects structure is again given by

$$\begin{aligned}\eta_i = \beta_0 &+ \beta_1(\mathrm{sex}_i)\\ &+ \beta_2(\mathrm{sex}_i)\cdot \mathrm{tOut}_i + \beta_3(\mathrm{sex}_i)\cdot \mathrm{tOut}_i^2\\ &+ \beta_4(\mathrm{sex}_i)\cdot \mathrm{tInOp}_i + \beta_5(\mathrm{sex}_i)\cdot \mathrm{tInOp}_i^2\\ &+ \beta_6(\mathrm{sex}_i)\cdot \mathrm{tOut}_i\cdot \mathrm{tInOp}_i\\ &+ \beta_7(\mathrm{sex}_i)\cdot \mathrm{tOut}_i\cdot \mathrm{tInOp}_i^2\\ &+ \beta_8(\mathrm{sex}_i)\cdot \mathrm{tOut}_i^2\cdot \mathrm{tInOp}_i\\ &+ \beta_9(\mathrm{sex}_i)\cdot \mathrm{tOut}_i^2\cdot \mathrm{tInOp}_i^2.\end{aligned}$$

As the model contains random effects based on two groupings, we limit the random effects structure to only including random intercepts and first order terms of the temperature variables. This is done in order to keep the model simple and thus easier to estimate. The random effects structure is given by the equation

$$\begin{gathered}\phi_i = u_1(\mathrm{subjId}_i) + u_2(\mathrm{subjId}_i)\cdot \mathrm{tOut}_i + u_3(\mathrm{subjId}_i)\cdot \mathrm{tInOp}_i +\\ v_1(\mathrm{subDay}_i) + v_2(\mathrm{subDay}_i)\cdot \mathrm{tOut}_i + v_3(\mathrm{subDay}_i)\cdot \mathrm{tInOp}_i.\end{gathered}$$

We again allow for a covariance of the random effects within a grouping. Thus the distributional assumptions are

$$U_i = \begin{pmatrix} u_1(\mathrm{subjId}_i)\\ u_2(\mathrm{subjId}_i)\\ u_3(\mathrm{subjId}_i)\end{pmatrix} \sim N\left(0, \begin{pmatrix} \sigma_{u_1}^2 & \sigma_{u_1u_2} & \sigma_{u_1u_3}\\ \sigma_{u_1u_2} & \sigma_{u_2}^2 & \sigma_{u_2u_3}\\ \sigma_{u_1u_3} & \sigma_{u_2u_3} & \sigma_{u_3}^2\end{pmatrix}\right)$$

$$V_i = \begin{pmatrix} v_1(\mathrm{subDay}_i)\\ v_2(\mathrm{subDay}_i)\\ v_3(\mathrm{subDay}_i)\end{pmatrix} \sim N\left(0, \begin{pmatrix} \sigma_{v_1}^2 & \sigma_{v_1v_2} & \sigma_{v_1v_3}\\ \sigma_{v_1v_2} & \sigma_{v_2}^2 & \sigma_{v_2v_3}\\ \sigma_{v_1v_3} & \sigma_{v_2v_3} & \sigma_{v_3}^2\end{pmatrix}\right).$$

Our initial model can therefore be stated as

$$\log(\mathrm{clo}_i) = \eta_i + \phi_i + \epsilon_i$$

where $i = 1, 2, \ldots, 803$, $\epsilon_i \sim N(0, \sigma^2)$ and the errors terms, the bundled vectors of random effects U_i and V_i are all pairwise independent.

2.4.2 Model diagnostics

Analysing the model diagnostic plots based on the conditional residuals (seen in figure 5) we draw the following conclusions

1. Looking at the plot of residuals against fitted values, we see that the clear diagonal lines are gone from the plot. Apart from this the variation in the residuals seems to equal the corresponding plot from the previous model.
2. For the QQ-plot we note that the tails are a bit heavier than we would expect from a normal distribution.
3. We again see that the residual variance for males is lower than the residual variance for females.
4. The plot of residuals against the temperature tInOp, shows no sign of a missing temperature structure.
5. The same conclusion holds for clo dependence on tOut.

Figure 6 shows a QQ-plot of the random effects. We do not see any clear indications against assuming a normal distribution for the random effects.

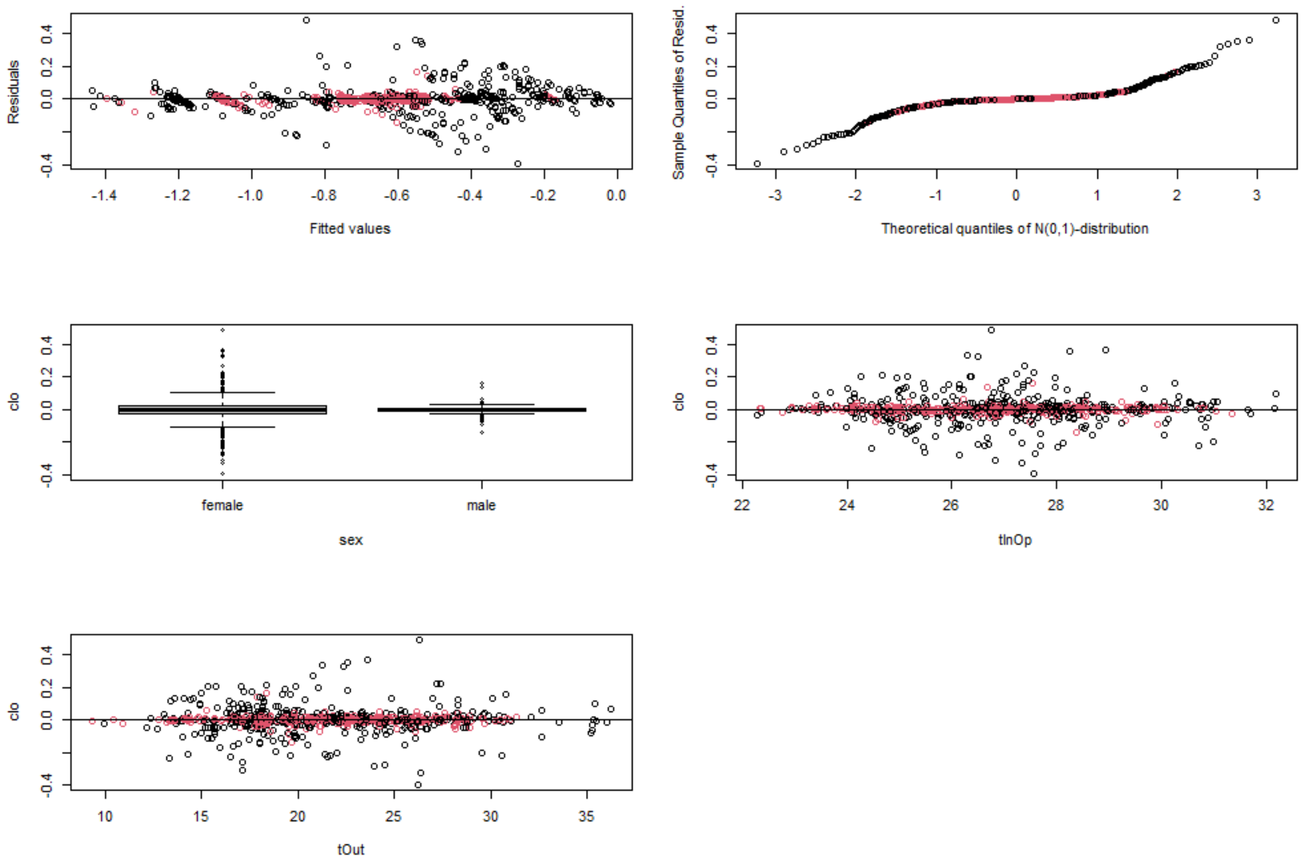

Figure 5: Diagnostics plots for the mixed effects model.

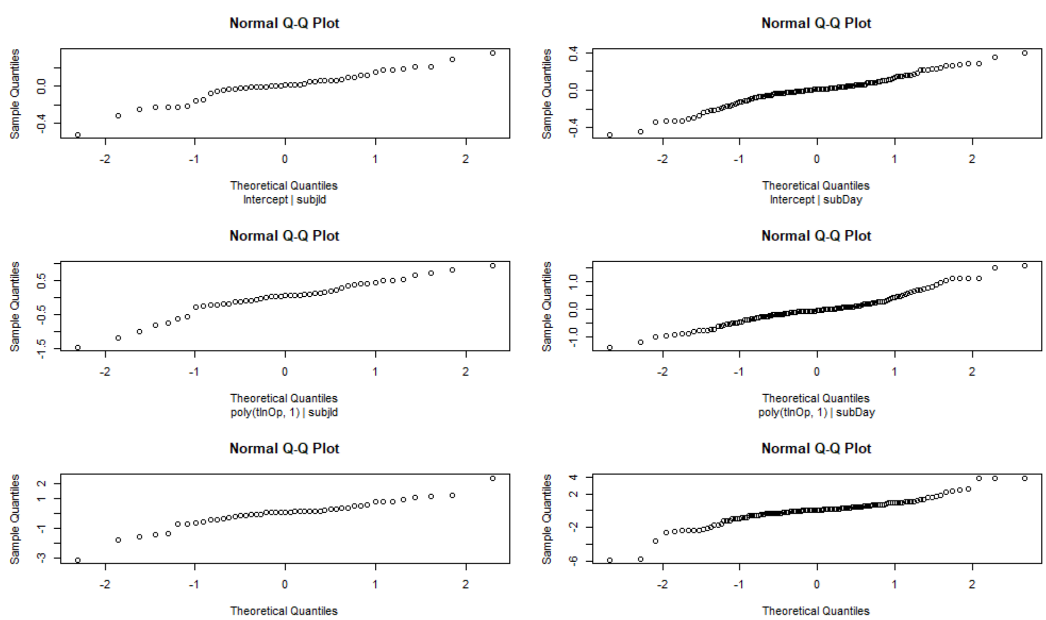

Figure 6: QQ-plots of the estimated random effects.

2.4.3 Model reduction

Our strategy for reducing the model is equal to our strategy for reducing the previous model. First we try to reduce the random effects structure. Here we follow the convention that we reduce the nested random effect due to subDay, before removing random effects due to subjId. The idea is that if we cannot remove the finer random effect due to subDay, then we should also keep the coarser effect due to subjId, as the random effect due to subDay gives a deviation from the subjId effect.

Fitting the initial model with REML-estimates and finding the restricted-log-likelihood l_{re} for it and its sub-models we reduce the random effects structure using LRT. The following ANOVA-table shows the reductions we have done in the order we removed the terms.

Removed term	N Parameters	AIC_{re}	BIC_{re}	l_{re}	Df	p-value
Initial model	31	-1137.4	-992.09	599.72		
tInOp | subDay	28	-1137.8	-1006.53	596.90	3	0.131
tOut | subDay	26	-1137.43	-1016.13	594.72	2	0.112

Table 7: ANOVA table for reducing random effects.

The LRT for removing the remaining random effects are all significant. The results are seen in the following ANOVA table

Removed term	N Parameters	AIC_{re}	BIC_{re}	l_{re}	Df	p-value
Intercept | subDay	25	-581.49	-464.85	315.75	1	< 0.0001
tInOp | subjId	28	-1135.05	-1027.22	590.53	3	0.039
tOut | subjId	23	-1130.88	-1023.57	588.44	3	0.006

Table 8: ANOVA table for reducing random effects.

Thus the final random effects structure is given by

$$\phi_i = u_1(\text{subjId}_i) + u_2(\text{subjId}_i) \cdot \text{tOut}_i + u_3(\text{subjId}_i) \cdot \text{tInOp}_i + v_1(\text{subDay}_i).$$

Having established the random effects structure, we reduce the fixed effects structure. We again reduce the fixed effects structure using LRT based on models fitted with maximum-likelihood-estimates, and use a TYPE II partitioning of the model deviance. The following ANOVA-table shows the reductions we have done in the order we removed the terms.

Model	N Parameters	AIC	BIC	l	Df	p-value
Initial model	26	-1110.39	-988.50	581.20		
All three-way-interactions removed	22	-1114.77	-1011.65	579.40	4	0.463
All tInOp $\times$ tOut interactions removed	18	-1122.18	-1037.79	579.09	4	0.961
All sex $\times$ tOut interactions removed	16	-1125.86	-1050.84	578.93	2	0.851
$\beta_3 \cdot \text{tOut}_i^2$ removed	15	-1127.73	-1057.41	578.87	1	0.723

Table 9: ANOVA table for reducing fixed effects.

Test for further model reduction are all rejected.

2.4.4 Model presentation

Our final model is given by the equation

$$\begin{aligned}\log(\text{clo}_i) =& \beta_0 + \beta_1(\text{sex}_i) + \beta_2 \cdot \text{tOut}_i + \beta_4(\text{sex}_i) \cdot \text{tInOp}_i + \beta_5(\text{sex}_i) \cdot \text{tInOp}_i^2 + \\ & u_1(\text{subjId}_i) + u_2(\text{subjId}_i) \cdot \text{tOut}_i + u_3(\text{subjId}_i) \cdot \text{tInOp}_i + v_1(\text{subDay}_i) + \epsilon_i\end{aligned}$$

with

$$U_i = \begin{pmatrix} u_1(\text{subjId}_i) \\ u_2(\text{subjId}_i) \\ u_3(\text{subjId}_i) \end{pmatrix} \sim N\left(0, \begin{pmatrix} \sigma_{u_1}^2 & \sigma_{u_1 u_2} & \sigma_{u_1 u_3} \\ \sigma_{u_1 u_2} & \sigma_{u_2}^2 & \sigma_{u_2 u_3} \\ \sigma_{u_1 u_3} & \sigma_{u_2 u_3} & \sigma_{u_3}^2 \end{pmatrix}\right)$$

$$v_1(\text{subDay}_i) \sim N\left(0, \sigma_{v_1}^2\right)$$

$$\epsilon_i \sim N\left(0, \sigma^2\right).$$

The errors are assumed independent from each other and from all random effects, and the random effects are assumed independent between different subjects and days. Table 10 and 11 presents the model parameters estimated using REML, together with 95% profile confidence intervals.

Parameter	Estimate	Std. Error	Lower CI (2.5 %)	Upper CI (97.5 %)
(Intercept)	-0.589	0.044	-0.674	-0.503
sexmale	-0.100	0.061	-0.223	0.022
poly(tInOp, 2)1	-1.022	0.257	-1.527	-0.518
poly(tInOp, 2)2	-0.094	0.159	-0.407	0.218
poly(tOut, 1)	-0.541	0.351	-1.229	0.148
sexmale:poly(tInOp, 2)1	1.048	0.346	0.367	1.728
sexmale:poly(tInOp, 2)2	0.003	0.260	-0.509	0.514

Table 10: Model parameters and their uncertainties for the final model.

Std. deviations and correlations	Estimate
$\tilde{\sigma}$	0.083
sd(Intercept) | subjId	0.189
sd(poly(tInOp, 1)) | subjId	0.657
sd(poly(tOut, 1)) | subjId	1.493
cor(Intercept, poly(tInOp, 1)) | subjId	-0.853
cor(Intercept, poly(tOut, 1)) | subjId	0.576
cor(poly(tInOp, 1), poly(tOut, 1)) | subjId	-0.065
sd(Intercept) | subDay	0.184

Table 11: Estimate of the variance parameters in the final model.

2.5 Repeated measurement model

The previous model shows that we should take correlation between observations on the same subjects within the same day into account. However, using the previous model we have the same correlation between all observations in the same subject on the same day, thus ignoring any time factor. In this section we will model the correlation in time, such that observations on the same subject that are closer together in time are more correlated than observations farther apart in time.

2.5.1 Model description

We again consider the linear predictor

$$\begin{aligned}
\eta_i = \beta_0 &+ \beta_1(\text{sex}_i) \\
&+ \beta_2(\text{sex}_i)\cdot \text{tOut}_i + \beta_3(\text{sex}_i)\cdot \text{tOut}_i^2 \\
&+ \beta_4(\text{sex}_i)\cdot \text{tInOp}_i + \beta_5(\text{sex}_i)\cdot \text{tInOp}_i^2 \\
&+ \beta_6(\text{sex}_i)\cdot \text{tOut}_i\cdot \text{tInOp}_i \\
&+ \beta_7(\text{sex}_i)\cdot \text{tOut}_i\cdot \text{tInOp}_i^2 \\
&+ \beta_8(\text{sex}_i)\cdot \text{tOut}_i^2\cdot \text{tInOp}_i \\
&+ \beta_9(\text{sex}_i)\cdot \text{tOut}_i^2\cdot \text{tInOp}_i^2.
\end{aligned}$$

for our initial fixed effects structure. Regarding the variance structure, we only consider models with random intercepts in this section. The marginal distribution of $\log(\text{clo}_i)$ therefore has the following form

$$\log(\text{clo}_i) \sim N(\eta_i, V), \qquad \text{where}$$
$$V_{i_1,i_2} = \begin{cases} 0 & \text{, if subDay}_{i_1} \neq \text{subDay}_{i_2} \text{ and } i_1 \neq i_2 \\ \sigma^2_{\text{subDay}} + \sigma^2\cdot\lambda(t_{i_1}-t_{i_2}) & \text{, if subDay}_{i_1} = \text{subDay}_{i_2} \text{ and } i_1 \neq i_2 \\ \sigma^2_{\text{subDay}} + \sigma^2 & \text{, if } i_1 = i_2 \end{cases}$$

for $i, i_1, i_2 = 1, 2, ..., 803$. Here $\lambda(t_{i_1} - t_{i_2})$ is a function on the distance between the observations, modelling the reduction in the correlation between observations that are further away from each other in time. We consider three possible choices for λ: a Gaussian covariance structure, an Exponential covariance structure and an Autoregressive(1) covariance structure. These are seen in table 12 together with the AIC_{re} and BIC_{re} from fitting the models using REML.

Covariance structure	Correlation term ($\lambda(t_{i_1} - t_{i_2})$)	AIC$_{re}$	BIC$_{re}$
Gaussian	$\exp\left(\frac{-(t_{i_1} - t_{i_2})^2}{\rho^2}\right)$	-1194.65	-1092.00
Exponential	$\exp\left(\frac{-\lvert t_{i_1} - t_{i_2}\rvert}{\rho}\right)$	-1196.23	-1093.59
Autoregressive(1)	$\rho^{\lvert t_{i_1} - t_{i_2}\rvert}$	-1204.14	-1106.16

Table 12: AIC and BIC for the three choices of covariance structure.

As the models are not nested, we cannot test them using LRT so we instead rely on the model AIC$_{re}$ and BIC$_{re}$ for chosing a model structure. We see that the AIC$_{re}$ and BIC$_{re}$ are quite close for the three covariance structures. As the AR(1) structure has the lowest AIC$_{re}$ and BIC$_{re}$, we use this for further modelling.

2.5.2 Model diagnostics

The model diagnostic plots based on the conditional residuals are seen in figure 7. For all six plots, the observations belonging to males and female are clearly marked, with black observations being female and red observations being male. Looking at the six plots we draw the following conclusions (plots are discussed row by row):

1. Looking at the plot of residuals against the fitted values, we see that the diagonal lines from previous models are gone. Overall the plot of the residuals against fitted values does not raise concerns about the model structure.
2. For the QQ-plot we note that the tails are a bit heavier than we would expect from a normal distribution.
3. We again see that the residual variance for males is lower than the residual variance for females.
4. The plot of residuals against the temperature tInOp, shows no sign of a missing temperature structure.
5. The same conclusion holds for clo dependence on tOut.
6. The random intercept effect seems to follow a normal distribution.

2.5.3 Model reduction

We first test for reducing the covariance structure to a compound symmetry structure (no time dependent correlation). As the compound symmetry model is a sub-model of the AR(1) model we can test this using LRT. Finding the restricted-log-likelihoods and calculating the test statistic $\lambda = -2\left(l_{re}(\text{Compound-symmetry-model}) - l_{re}(\text{AR(1)-model})\right)$, we get a test statistic of $\lambda = 113.21$ and comparing it to the $\chi^2(1)$-distribution we get a p-value of $p < 0.0001$. Thus as suspected we reject the reduction of the variance structure.

Next we focus on reducing the fixed effects structure. We again reduce the fixed effects structure using LRT based on models fitted with maximum-likelihood-estimates, and use a TYPE II partitioning of the model deviance. The following ANOVA-table shows the reductions we have done in the order we removed the terms.

Model	N Parameters	AIC	BIC	l	Df	p-value
Initial model	21	-1178.49	-1080.04	610.25		
All three-way-interactions removed	17	-1185.45	-1105.74	609.72	4	0.903
All tInOp $\times$ tOut interactions removed	13	-1193.26	-1132.31	609.63	4	0.996
All sex $\times$ tOut interactions removed	11	-1196.12	-1144.55	609.06	2	0.566
All tOut terms removed	9	-1194.95	-1152.75	606.47	2	0.075
$\beta_5(\text{sex}_i)\cdot\text{tInOp}_i^2$ removed	7	-1197.93	-1165.11	605.97	2	0.602

Table 13: ANOVA table for reducing fixed effects.

Finally, trying to reduce the sex$\times$tInOp interaction yields a test statistic of $\lambda = 7.6$ and comparing this with the $\chi^2(1)$-distribution we get $p = 0.006$ thus rejecting further model reduction.

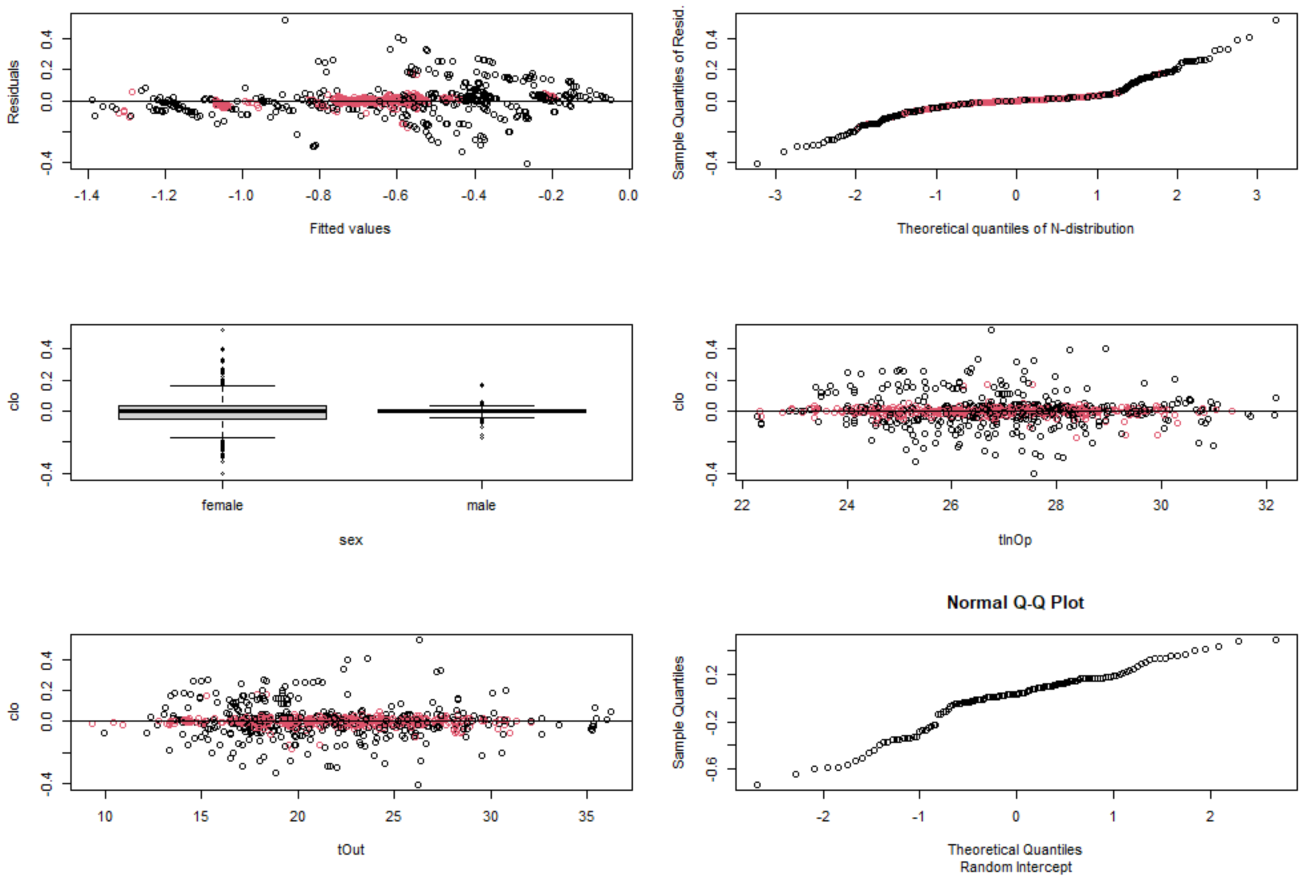

Figure 7: Diagnostics plots for the model.

2.5.4 Model presentation

Our final model is given by

$$\log(\text{clo}_i) \sim N(\eta_i, V), \qquad \text{where}$$

$$\eta_i = \beta_1(\text{sex}_i) + \beta_4(\text{sex}_i) \cdot \text{tInOp}_i \tag{4}$$

$$V_{i_1,i_2} = \begin{cases} 0 & \text{, if subDay}_{i_1} \neq \text{subDay}_{i_2} \\ \sigma^2_{\text{subDay}} + \sigma^2 \cdot \rho^{|t_{i_1} - t_{i_2}|} & \text{, if subDay}_{i_1} = \text{subDay}_{i_2} \end{cases}$$

for $i, i_1, i_2 = 1, 2, ..., 803$.

Table 14 and 15 presents the model parameters estimated using REML together with 95% profile confidence intervals

Parameter	Estimate	Std. Error	Lower CI (2.5 %)	Upper CI (97.5 %)
Intercept β_1(female)	-0.070	0.109	-0.286	0.146
Intercept β_1(male)	-0.596	0.113	-0.819	-0.372
Slope β_4(female	-0.019	0.004	-0.027	-0.012
Slope β_4(male)	-0.004	0.004	-0.012	0.004

Table 14: Model parameters and their uncertainties for the final model.

Std. deviations and correlations	Lower CI (2.5 %)	Estimate	Upper CI (97.5 %)
$\tilde{\sigma}$	0.091	0.101	0.112
$\tilde{\sigma}_{\text{subDay}}$	0.235	0.267	0.303
$\tilde{\rho}$	0.403	0.512	0.606

Table 15: Estimate of the variance parameters in the final model.

2.6 Interpretation of the final model

Having fitted three different models to the data, we found an AIC of -557.84 for the final mixed effects model using random effects based on subjId, an AIC of -1127.73 for the final mixed effects model using both subjId and subDay as random effects, and finally an AIC of -1197.93 for the final repeated measurements model. This, coupled with the fact that the repeated measurements model gives a model that is easier to interpret, means that we will go with the repeated measurements model as specified in equation (4) as our final model.

As we treat the combination of subjects and days as random effects, we are not interested in the specific values for the subjects. Instead we are interested in the variance between them. In table 15, we see that the subDay-to-subDay standard deviation is around twice as large as the within subject standard deviation. Thus a large part of the variation in the model comes from the difference between the sampled subjects.

As we have log-transformed the observations, we need to back transform the fixed effects estimate for us to interpret them. We find that the intercept for the indoors operating temperature is $0.93(0.75, 1.16)$ for females and 0.55 (0.44, 0.69) for males. Furthermore, we found that a 1-degree increase in the indoor operating temperature yields a 1.9% (1.2%, 2.7%) decrease in the clo level for female and a 0.38% decrease (1.17% decrease, 0.42% increase) in the clo level for males. We note that as a 0% increase is in the confidence interval for males, the temperature effect is not significant for males.

Thus, males are less susceptible to change their clothing insulation level depending on the indoor operating temperature. This tendency is also visualised in figure 8, where we have plotted the fitted lines based on the fixed effects.

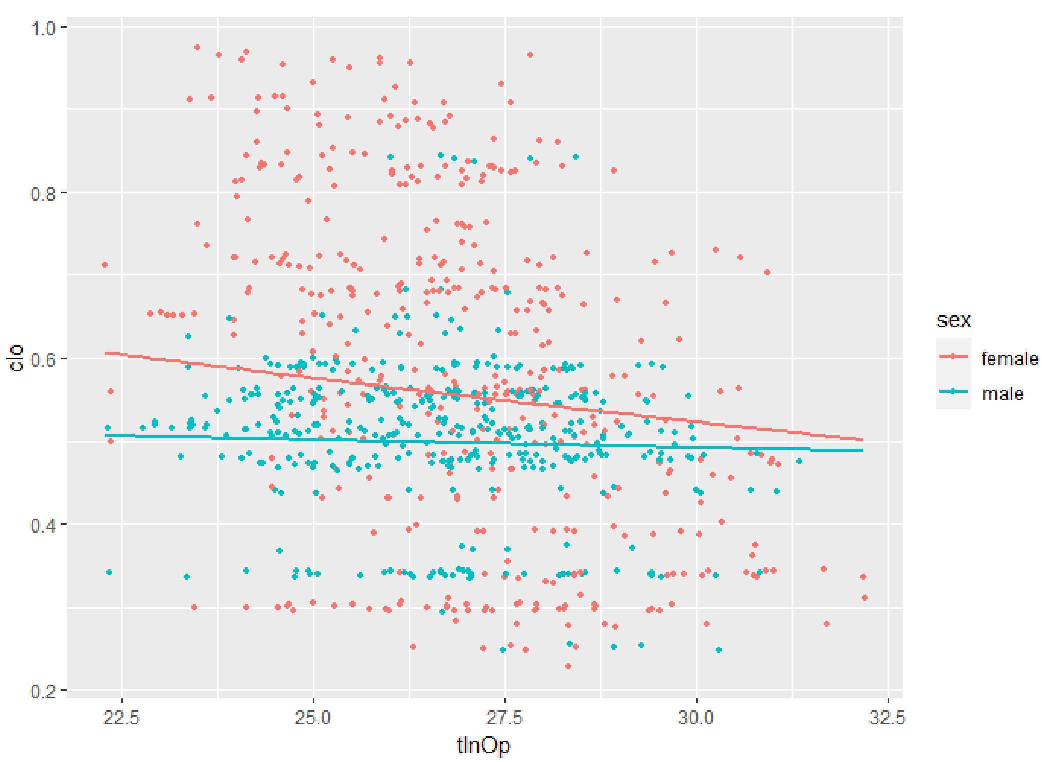

Figure 8: Regression lines for each sex.

In figure 9 we have plotted the fitted values for six random selected females and males, to visualise the found correlation between observations in time. The vertical dotted lines denotes the changing to a new experimental day.

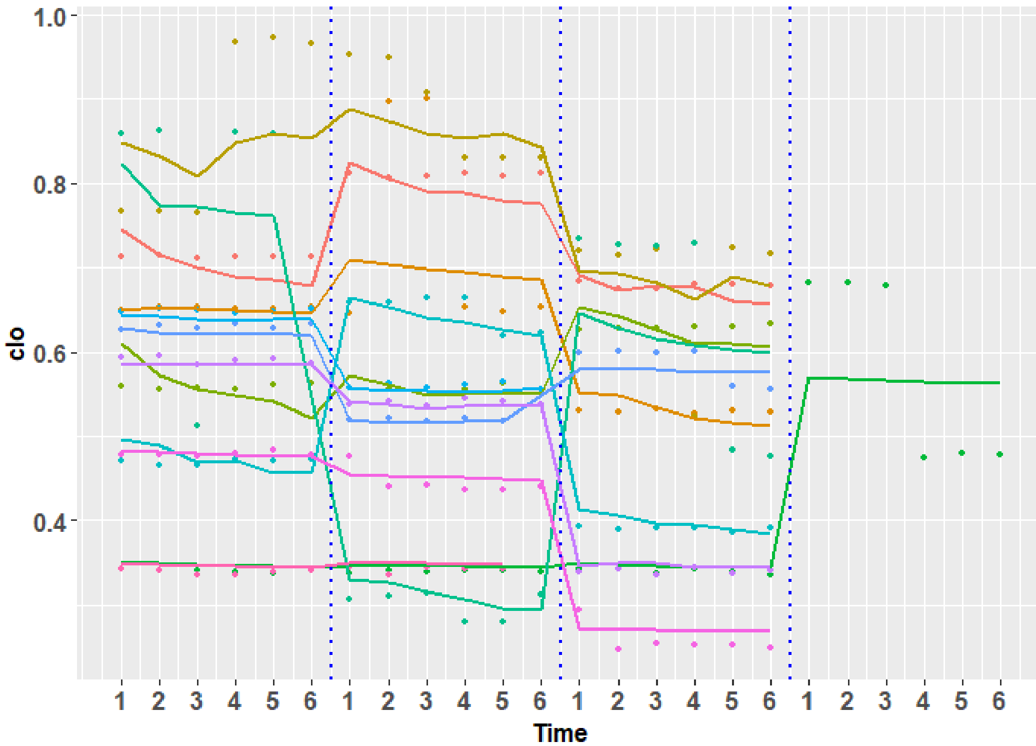

Figure 9: Fitted values for six random selected females and males.

2.7 Conclusion

As the aim of the experiment is to generalize the relation between sex, temperature and clothing insulation to a larger population, it is natural to treat the experimental subjects as random effects. Having fitted three different mixed effects models, we found that it was not only beneficial to treat observations on the same day for each subjects as correlated, but also to take into account the effect of observations being closer in time, as being more correlated.

Fitting a mixed effects model with an AR(1)-covariance structure for each combination of subject and day, we found that the clo level is only affected by the sex of the subject and the indoors temperature. Here we saw that males have a lower clo level at room temperature (22 degrees) than females, and that the clo level decreases as the indoor temperature increase. We furthermore found a significant difference between the indoor temperatures effect on the clo level in males and females, with the temperature effect being non-significant for males.

From the model diagnostics of all three mixed effect models, we observed that there is a large difference in the residual variation between males and females, with females having a larger variance. As we have not modelled this effect in part 1, we will look into modelling a sex specific variance in part 2.

In the previous assignment, we found that a fixed effects generalized linear model, with a Gamma distribution was a better fit to the data than the normal distribution. Given more time, it could be interesting to examine the fit of a generalized mixed effects Gamma model to the data.

3 Hierarchical models: Random varaiance

3.1 First model

We consider the model

$$Y_{i,j,k} = \log(\text{clo}_{i,j,k}) = \mu + \beta\left(\text{sex}_i\right) + u_i + \varepsilon_{ijk}; \quad u_i \sim N\left(0, \sigma_u^2\right) \quad \varepsilon_{ijk} \sim \left(0, \sigma^2\right)$$

where $Y_{i,j,k}$ is the logarithm of the clothing insulation level for subject i on day j and k refer to the observation number within the day. We have that

$$\mathrm{E}\left[Y_{ijk}\right] = \mu + \beta(\text{sex}_i); \quad \mathrm{Cov}\left[Y_{ijk}, Y_{hlm}\right] = \begin{cases} \sigma_u^2 + \sigma^2 & \text{for } (i,j) = (h,l) \quad \text{(same subject, same observation)} \\ \sigma_u^2 & \text{for } i = h, \quad j \neq l \quad \text{(same subject)} \\ 0 & \text{for } i \neq h \end{cases} \tag{2}$$

We can further write the model in the following form for subject i:

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{U} + \varepsilon_i; \quad \mathbf{U} \sim N(0, \sigma_u^2), \quad \varepsilon_i \sim N_{n_i}(0, \sigma^2 I) \tag{3}$$

Therefore we can write the model as a multivariate normal distribution

$$\mathbf{Y}_i \sim N_{n_i}\left(\mathbf{X}_i\boldsymbol{\beta}, \sigma_u^2\mathbf{Z}\mathbf{Z}^T + \sigma^2 I\right) \tag{4}$$

The model is parametrized by $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_u^2, \sigma^2)$. We can write the log likelihood for the model

$$l(\boldsymbol{\theta}; \boldsymbol{y}) = \log L(\boldsymbol{\theta}; \boldsymbol{y}) = \sum_{i=1}^{n} l_i(\boldsymbol{\theta}; \boldsymbol{y}_i) \tag{5}$$

where n is the number of subjects. We implement the likelihood function and perform the optimisation with respect to the parameters. The detailed implementation is presented:

```
n = dim(df)[1]
X = cbind(rep(1, n), df$sex_optimization)
p = dim(X)[2]
Z = cbind(rep(1, n))
u_number = dim(Z)[2]
y = log(clo)

obj_1 = function(beta){
  result = 0
  for(subject_i in unique(df$subjId)){
    X_i = X[df$subjId == subject_i, , drop = F]
    n_i = dim(X_i)[1]
    y_i = y[df$subjId == subject_i]
    y_hat_i = X_i %*% beta[1:p]
    Psi_i = exp(beta[p+1])*diag(u_number)
    Z_i = Z[df$subjId == subject_i, , drop = F]
    Sigma_full_i = exp(beta[p+2])*diag(n_i) + Z_i %*% Psi_i %*% t(Z_i)
    result = result +
      sum(dmvnorm(y_i, mean = y_hat_i, sigma = Sigma_full_i, log = TRUE))
  }
  return(-result)
}
```

We verify our results with the model estimated by the library lme4. The estimated parameters are presented in the table 16

	own implementation		lme4	
Parameter	Estimate	Std. Error	Estimate	Std. Error
(Intercept)	-0.584	0.047	-0.584	0.047
sexmale	-0.109	0.067	-0.109	0.067
σ_u^2	0.050	-	0.050	0.224
σ^2	0.035	-	0.035	0.186

Table 16: Model parameters and their uncertainties for the final model.

Since the variance parameters were estimated in the log domain constructing the Wald confidence interval for them would require taking the transformation into account as the Wald confidence intervals are not invariant to parameter transformations. We chose to not report the uncertainty for those parameters for our optimization method.

3.2 Second model

We consider the model

$$Y_{i,j,k} = \log(\text{clo}_{i,j,k}) = \mu + \beta\left(\text{sex}_i\right) + u_i + v_{ij} + \varepsilon_{ijk}; \quad u_i \sim N\left(0, \sigma_u^2\right) \quad v_{ij} \sim N\left(0, \sigma_v^2\right) \quad \varepsilon_{ijk} \sim \left(0, \sigma^2\right)$$

where $Y_{i,j,k}$ is the logarithm of the clothing insulation level for subject i on day j and k refer to the observation number within the day. We have that

$$\mathrm{E}\left[Y_{ijk}\right] = \mu + \beta(\text{sex}_i); \tag{6}$$

$$\mathrm{Cov}\left[Y_{ijk}, Y_{hlm}\right] = \begin{cases} \sigma_u^2 + \sigma_v^2 + \sigma^2 & \text{for } (i,j,k) = (h,l,m) \quad \text{(same subject, same day, same observation)} \\ \sigma_u^2 + \sigma_v^2 & \text{for } (i,j) = (h,l), \quad k \neq m \quad \text{(same subject, same day)} \\ \sigma_u^2 & \text{for } i = h, \quad j \neq l \text{(same subject)} \\ 0 & \text{for } i \neq h \end{cases} \tag{7}$$

We can further write the model in the following form for subject i and day j:

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_{1,i}\mathbf{U} + \mathbf{Z}_{2,i}\mathbf{V} + \varepsilon_i; \quad \mathbf{U} \sim N(0, \sigma_u^2), \quad \mathbf{V} \sim N_{days_i}(0, \sigma_v^2 I), \quad \varepsilon_i \sim N_{n_i}(0, \sigma^2 I) \tag{8}$$

Therefore we can write the model as a multivariate normal distribution

$$\mathbf{Y}_i \sim N_{n_i}\left(\mathbf{X}_i\boldsymbol{\beta}, \sigma_u^2 \mathbf{Z_{1,i}}\mathbf{Z_{1,i}}^T + \sigma_v^2 \mathbf{Z_{2,i}}\mathbf{Z_{2,i}}^T + \sigma^2 I\right) \tag{9}$$

The model is parametrized by $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_u^2, \sigma_v^2, \sigma^2)$. We can write the log likelihood for the model

$$l(\boldsymbol{\theta}; \boldsymbol{y}) = \log L(\boldsymbol{\theta}; \boldsymbol{y}) = \sum_{i=1}^{n} l_i(\boldsymbol{\theta}; \boldsymbol{y}_i) \tag{10}$$

where n is the number of subjects. We implement the likelihood function and perform the optimisation with respect to the parameters. The detailed implementation is presented:

```
n = dim(df)[1]
X = cbind(rep(1, n), df$sex_optimization)
p = dim(X)[2]
Z1 = cbind(rep(1, n))
u_number1 = dim(Z1)[2]
Z2 = cbind(rep(1, n))
u_number2 = dim(Z2)[2]
y = log(clo)

obj_2 = function(beta){
  result = 0
  for(subject_i in unique(df$subjId)){
    X_i = X[df$subjId == subject_i, , drop = F]
    n_i = dim(X_i)[1]
    y_i = y[df$subjId == subject_i]
    y_hat_i = X_i %*% beta[1:p]

    Psi1_i = exp(beta[p+1])*diag(u_number)
    Z1_i = Z1[df$subjId == subject_i, , drop = F]
    Z2_i = Z2[df$subjId == subject_i, , drop = F]

    subject_i_days_length = length(unique(df$day[df$subjId == subject_i]))
    day_matrix_list = vector("list", subject_i_days_length)
    for(index_day in seq(subject_i_days_length)){
      subject_i_days = df$day[df$subjId == subject_i]
      day_j = unique(subject_i_days)[index_day]
      Psi2_ij_tmp = exp(beta[p+3])*diag(u_number2)
      Z2_ij = Z2_i[subject_i_days == day_j, , drop = F]
      matrix_tmp = Z2_ij %*% Psi2_ij_tmp %*% t(Z2_ij)
      day_matrix_list[[index_day]] = matrix_tmp
    }
    Psi2_ij = as.matrix(bdiag(day_matrix_list))
    Sigma_full_i = exp(beta[p+2])*diag(n_i) +
      Z1_i %*% Psi1_i %*% t(Z1_i) + Psi2_ij

    result = result +
      sum(dmvnorm(y_i, mean = y_hat_i, sigma = Sigma_full_i, log = TRUE))
  }
  return(-result)
}
```

We verify our results with the model estimated by the library lme4. The estimated parameters are presented in the table 17. The comment from the section about model 1 about the Wald confidence intervals for variance parameters still applies.

	own implementation		lme4	
Parameter	Estimate	Std. Error	Estimate	Std. Error
(Intercept)	-0.583	0.047	-0.583	0.047
sexmale	-0.111	0.067	-0.111	0.067
σ_u^2	0.039	-	0.039	0.197
σ_v^2	0.038	-	0.038	0.195
σ^2	0.0079	-	0.0079	0.089

Table 17: Model parameters and their uncertainties

3.3 Third model

We consider the model

$$Y_{i,j,k} = \log(\text{clo}_{i,j,k}) = \mu + \beta\,(\text{sex}_i) + u_i + v_{ij} + \varepsilon_{ijk}; \tag{11}$$

$$u_i \sim N\left(0, \sigma_u^2\alpha(\text{sex}_i)\right) \quad v_{ij} \sim N\left(0, \sigma_v^2(\alpha\,\text{sex}_i)\right) \quad \varepsilon_{ijk} \sim \left(0, \sigma^2\alpha(\text{sex}_i)\right) \tag{12}$$

where $Y_{i,j,k}$ is the logarithm of the clothing insulation level for subject i on day j and k refer to the observation number within the day. We have that

$$\text{E}\left[Y_{ijk}\right] = \mu + \beta(\text{sex}_i); \tag{13}$$

$$\text{Cov}\left[Y_{ijk}, Y_{hlm}\right] = \begin{cases} (\sigma_u^2 + \sigma_v^2 + \sigma^2)\alpha(\text{sex}_i) & \text{for } (i,j,k) = (h,l,m) \quad \text{(same subject, same day, same observation)} \\ (\sigma_u^2 + \sigma_v^2)\alpha(\text{sex}_i) & \text{for } (i,j) = (h,l), \quad k \neq m \quad \text{(same subject, same day)} \\ \sigma_u^2\alpha(\text{sex}_i) & \text{for } i = h, \quad j \neq l \text{(same subject)} \\ 0 & \text{for } i \neq h \end{cases} \tag{14}$$

We can further write the model in the following form for subject i and day j:

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_{1,i}\mathbf{U} + \mathbf{Z}_{2,i}\mathbf{V} + \varepsilon_i;$$
$$\mathbf{U} \sim N(0, \sigma_u^2\alpha(\text{sex}_i)), \quad \mathbf{V} \sim N_{days_i}(0, \sigma_v^2\alpha(\text{sex}_i)I), \quad \varepsilon_i \sim N_{n_i}(0, \sigma^2\alpha(\text{sex}_i)I)$$

Therefore we can write the model as a multivariate normal distribution

$$\mathbf{Y}_i \sim N_{n_i}\left(\mathbf{X}_i\boldsymbol{\beta}, \left(\sigma_u^2\mathbf{Z_{1,i}}\mathbf{Z_{1,i}}^T + \sigma_v^2\mathbf{Z_{2,i}}\mathbf{Z_{2,i}}^T + \sigma^2 I\right)\alpha(\text{sex}_i)\right) \tag{15}$$

The model is parametrized by $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_u^2, \sigma_v^2, \sigma^2, \alpha)$. Here α gives the weight for the variance for males, i.e. if $\tilde{\sigma^2}_{female}$ is the estimated variance parameter for females then the corresponding variance parameter for males is given by $\tilde{\sigma^2}_{male} = \alpha\tilde{\sigma^2}_{female}$.

We can write the log-likelihood for the model

$$l(\boldsymbol{\theta}; \boldsymbol{y}) = \log L(\boldsymbol{\theta}; \boldsymbol{y}) = \sum_{i=1}^{n} l_i(\boldsymbol{\theta}; \boldsymbol{y}_i) \tag{16}$$

where n is the number of subjects. We implement the likelihood function and perform the optimisation with respect to the parameters. The detailed implementation is presented:

```
n = dim(df)[1]
X = cbind(rep(1, n), df$sex_optimization)
p = dim(X)[2]
Z1 = cbind(rep(1, n))
u_number1 = dim(Z1)[2]
Z2 = cbind(rep(1, n))
u_number2 = dim(Z2)[2]
X_alphas = X # note that it's the case in our model
y = log(clo)

obj_3 = function(beta){

  result = 0
  for(subject_i in unique(df$subjId)){
    X_i = X[df$subjId == subject_i, , drop = F]
    n_i = dim(X_i)[1]
```

```r
    y_i = y[df$subjId == subject_i]
    y_hat_i = X_i %*% beta[1:p]

    X_alphas_i = unique(X_alphas[df$subjId == subject_i, , drop = F])
    Psi1_i = as.numeric(
      exp(X_alphas_i %*% beta[c("sigma_u^2_log", "alpha")])
    )*diag(u_number1)

    Z1_i = Z1[df$subjId == subject_i, , drop = F]
    Z2_i = Z2[df$subjId == subject_i, , drop = F]

    subject_i_days_length = length(unique(df$day[df$subjId == subject_i]))
    day_matrix_list = vector("list", subject_i_days_length)
    for(index_day in seq(subject_i_days_length)){
      subject_i_days = df$day[df$subjId == subject_i]
      day_j = unique(subject_i_days)[index_day]
      X_alphas_ij = X_alphas_i
      Psi2_ij_tmp = as.numeric(
        exp(X_alphas_ij %*% beta[c("sigma_v^2_log", "alpha")])
      )*diag(u_number2)
      Z2_ij = Z2_i[subject_i_days == day_j, , drop = F]

      matrix_tmp = Z2_ij %*% Psi2_ij_tmp %*% t(Z2_ij)
      day_matrix_list[[index_day]] = matrix_tmp
    }
    Psi2_ij = as.matrix(bdiag(day_matrix_list))

    Sigma_i = as.numeric(
      exp(X_alphas_i %*% beta[c("sigma^2_log", "alpha")])
    )*diag(n_i)

    Sigma_full_i = Sigma_i + Z1_i %*% Psi1_i %*% t(Z1_i) + Psi2_ij

    result = result +
      sum(dmvnorm(y_i, mean = y_hat_i, sigma = Sigma_full_i, log = TRUE))
  }
  return(-result)
}
```

The estimated parameters are presented in the table 18. The variance parameters and the α parameter were estimated in the log domain, since they are positive numbers. The standard errors in the table correspond to the uncertainty in the domain where the optimization was performed: the regular domain for fixed effects regression parameters and log domain for the rest. The estimates in the linear domain for all parameters were also presented

Parameter	Estimate log domain	Std. Error log domain	Estimate	Std. Error
(Intercept)	-	-	-0.5833	0.0797
sexmale	-	-	-0.1117	0.0830
σ_u^2	-2.4366	0.3886	0.0874	-
σ_v^2	-1.680	0.1775	0.186	-
σ^2	-4.365	0.0696	0.0127	-
α	-2.499	0.1049	0.0821	-

Table 18: Model parameters and their uncertainties

3.4 Proof of the reformulated Theorem 6.7

Given

$$Y_i \mid \gamma_i \sim N\left(\mu, \frac{\sigma^2}{\gamma_i}\right); \quad \gamma_i \sim \text{Gamma}(1, \lambda); \quad E[\gamma_i] = 1 \quad Var[\gamma_i] = \frac{1}{\lambda} \tag{17}$$

We want to show that the marginal distribution of Y_i is

$$f_{Y_i} \sim \frac{1}{\sigma} f_0\left(\frac{y-\mu}{\sigma}; 2\lambda\right)$$

where f_0 is the pdf of a student t-distributed random variable with 2λ degrees of freedom.
Proof: We note that the canonical parametrization of the Gamma distribution is

$$Y \sim \text{G}(\alpha, \beta); \quad E[Y] = \alpha\beta \quad Var[Y] = \alpha\beta^2 \tag{18}$$

where α is the shape parameter and β is the scale parameter. We choose to reparametrize the distribution in way such that the expected value is the mean value parameter μ_{G} and the variation is characterized by the precision parameter λ as it is done on page 96.

$$\alpha = \lambda \quad \beta = \frac{\mu_{\mathrm{G}}}{\lambda} \tag{19}$$

We obtain the following density function after reparametrization.

$$f_{\gamma_i}(\gamma_i, \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \gamma_i^{\alpha-1} \exp\left(-\frac{\gamma_i}{\beta}\right) = \frac{\lambda^\lambda}{\mu_{\mathrm{G}}^\lambda \Gamma(\lambda)} \gamma_i^{\lambda-1} \exp\left(-\frac{\gamma_i \lambda}{\mu_{\mathrm{G}}}\right) \tag{20}$$

In the Theorem we have $\mu_{\mathrm{G}} = 1$, hence we obtain

$$f_{\gamma_i}(\gamma_i, \lambda) = f_{\gamma_i}(\gamma_i, \mu_{\mathrm{G}} = 1, \lambda) = \frac{\lambda^\lambda}{\Gamma(\lambda)} \gamma_i^{\lambda-1} \exp\left(-\gamma_i \lambda\right) \tag{21}$$

The conditional density of $Y_i|\gamma_i$ is the normal distribution

$$f_{Y|\gamma_i}(y; \gamma_i) = \frac{\sqrt{\gamma_i}}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\gamma_i}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right) \tag{22}$$

The density of the student t distribution with 2λ degrees of freedom is

$$f_0(y) = \frac{\Gamma\left(\frac{2\lambda+1}{2}\right)}{\sqrt{2\lambda\pi}\Gamma\left(\lambda\right)} \left(1 + \frac{y^2}{2\lambda}\right)^{-(2\lambda+1)/2} \tag{23}$$

Consider

$$\begin{aligned} f_{Y_i}(y) &= \int_0^\infty f_{Y_i|\gamma_i}(y_i; \gamma_i) f_{\gamma_i}(\gamma_i, \lambda) d\gamma_i = \int_0^\infty \frac{\sqrt{\gamma_i}}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\gamma_i}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right) \frac{\lambda^\lambda}{\Gamma(\lambda)} \gamma_i^{\lambda-1} \exp\left(-\gamma_i\lambda\right) d\gamma_i \\ &= \frac{\lambda^\lambda}{\sigma\sqrt{2\pi}\Gamma(\lambda)} \int_0^\infty \gamma_i^{(\lambda+\frac{1}{2})-1} \exp\left(-\gamma_i\left(\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2 + \lambda\right)\right) d\gamma_i \end{aligned}$$

The integrand is seen as the kernel of a Gamma distribution: $\mathrm{G}\left(\alpha = \lambda + \frac{1}{2}, \beta = 1/\left(\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2 + \lambda\right)\right)$. Therefore we need to adjust the integrating constant to obtain the true distribution which will integrate to 1. After doing so we obtain the following:

$$f_{Y_i}(y) = \frac{1}{\sqrt{2\pi}\sigma} \frac{\lambda^\lambda}{\Gamma(\lambda)} \frac{\Gamma\left(\lambda + \frac{1}{2}\right)}{\left(\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 + \lambda\right)^{\lambda+\frac{1}{2}}}$$

Thus substituting $t = \frac{y-\mu}{\sigma}$ and $v = 2\lambda$ we get

$$f_{Y_i}(y) = \frac{1}{\sigma} \frac{\sigma\left(\frac{v+1}{2}\right)}{\sqrt{2\pi}\Gamma(v/2)} \lambda^{-\frac{1}{2}} \lambda^{\lambda+\frac{1}{2}} \left(\frac{t^2+v}{2}\right)^{-\frac{v+1}{2}} = \frac{1}{\sigma} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\Gamma(v/2)} \left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}} = \frac{1}{\sigma} f_0\left(\frac{y-\mu}{\sigma}; 2\lambda\right)$$

which was to be proven.

3.5 Hierarchical model

We now consider the model

$$\begin{aligned} Y_{i,j,k}|u_i, v_{ij}, \gamma_i = \log(\mathrm{clo}_{i,j,k})|u_i, v_{ij}, \gamma_i &\sim N\left(\mu + \beta\left(\mathrm{sex}_i\right) + u_i + v_{ij},\, \sigma^2\alpha(\mathrm{sex}_i)/\gamma_i\right) \\ u_i|\gamma_i &\sim N\left(0, \sigma_u^2\alpha(\mathrm{sex}_i)/\gamma_i\right) \\ v_{ij}|\gamma_i &\sim N\left(0, \sigma_v^2\alpha(\mathrm{sex}_i)/\gamma_i\right) \\ \gamma_i &\sim G\left(1, \lambda\right) \end{aligned}$$

where $Y_{i,j,k}$ is the logarithm of the clothing insulation level for subject i on day j and k refer to the observation number within the day. We have that

$$\mathrm{E}\left[Y_{ijk}|\gamma_i\right] = \mu + \beta(\mathrm{sex}_i); \tag{24}$$

$$\mathrm{Cov}\left[Y_{ijk}|\gamma_i, Y_{hlm}|\gamma_h\right] = \begin{cases} (\sigma_u^2 + \sigma_v^2 + \sigma^2)\alpha(\mathrm{sex}_i)\gamma_i^{-1} & \text{for } (i,j,k) = (h,l,m) \quad \text{(same subject, same day, same observation)} \\ (\sigma_u^2 + \sigma_v^2)\alpha(\mathrm{sex}_i)\gamma_i^{-1} & \text{for } (i,j) = (h,l), \quad k \neq m \quad \text{(same subject, same day)} \\ \sigma_u^2\alpha(\mathrm{sex}_i)\gamma_i^{-1} & \text{for } i = h, \quad j \neq l \text{(same subject)} \\ 0 & \text{for } i \neq h \end{cases} \tag{25}$$

We can further write the model in the following form for subject i and day j:

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_{1,i}\mathbf{U} + \mathbf{Z}_{2,i}\mathbf{V} + \varepsilon_i;$$
$$\mathbf{U}|\gamma_i \sim N(0, \sigma_u^2\alpha(\mathrm{sex}_i)\gamma_i^{-1}), \quad \mathbf{V}|\gamma_i \sim N_{days_i}(0, \sigma_v^2\alpha(\mathrm{sex}_i)\gamma_i^{-1}I), \quad \varepsilon_i|\gamma_i \sim N_{n_i}(0, \sigma^2\alpha(\mathrm{sex}_i)\gamma_i^{-1}I)$$

Therefore we can write the model as a multivariate normal distribution

$$\mathbf{Y}_i|\gamma_i \sim N_{n_i}\left(\mathbf{X}_i\boldsymbol{\beta}, \left(\sigma_u^2\mathbf{Z_{1,i}}\mathbf{Z_{1,i}}^T + \sigma_v^2\mathbf{Z_{2,i}}\mathbf{Z_{2,i}}^T + \sigma^2 I\right)\alpha(\mathrm{sex}_i)\gamma_i^{-1}\right) \tag{26}$$
$$\gamma_i \sim G(1, \lambda)$$

The model is parametrized by $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_u^2, \sigma_v^2, \sigma^2, \alpha, \lambda)$. We can write the log-likelihood for the model

$$l(\boldsymbol{\theta}; \boldsymbol{y}) = \log L(\boldsymbol{\theta}; \boldsymbol{y}) = \sum_{i=1}^{n} l_i(\boldsymbol{\theta}; \boldsymbol{y}_i) \tag{27}$$

where n is the number of subjects.

We note that equation (27) is the multivariate equivalent of the conditional distribution from Theorem 6.7 mentioned in section 3.4. Theorem 6.7 generalizes to the multivariate setting as changing the proof to account for the dispersion matrix Σ in the multivariate setting is straightforward. As such we will not go further into detail with the proof here. The theorem tells us that the marginal distribution for $Y_{i,j,k}$ is

$$Y_i \sim t_{n_i}\left(\mathbf{X}_i\boldsymbol{\beta}, \left(\sigma_u^2\mathbf{Z_{1,i}}\mathbf{Z_{1,i}}^T + \sigma_v^2\mathbf{Z_{2,i}}\mathbf{Z_{2,i}}^T + \sigma^2 I\right)\alpha(\mathrm{sex}_i), 2\lambda\right)$$

where t_{n_i} is the multivariate t-distribution with mean, scale and degrees of freedom as specified.

The model is parametrized by $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_u^2, \sigma_v^2, \sigma^2, \alpha, \lambda)$. We can write the log-likelihood for the model

$$l(\boldsymbol{\theta}; \boldsymbol{y}) = \log L(\boldsymbol{\theta}; \boldsymbol{y}) = \sum_{i=1}^{n} l_i(\boldsymbol{\theta}; \boldsymbol{y}_i) \tag{28}$$

We implement the likelihood function and perform the optimisation with respect to the parameters. The detailed implementation in R is presented:

```
n = dim(df)[1]
X = cbind(rep(1, n), df$sex_optimization)
p = dim(X)[2]
Z1 = cbind(rep(1, n))
u_number1 = dim(Z1)[2]
Z2 = cbind(rep(1, n))
u_number2 = dim(Z2)[2]
X_alphas = X # note that it's the case in our model
y = log(clo)

obj_5 = function(beta){
  result = 0
  for(subject_i in unique(df$subjId)){
    X_i = X[df$subjId == subject_i, , drop = F]
    n_i = dim(X_i)[1]
    y_i = y[df$subjId == subject_i]
    y_hat_i = X_i %*% beta[1:p]

    X_alphas_i = unique(X_alphas[df$subjId == subject_i, , drop = F])
    Psi1_i = as.numeric(
```

```
        exp(X_alphas_i %*% beta[c("sigma_u^2_log", "alpha")])
      )*diag(u_number1)

      Z1_i = Z1[df$subjId == subject_i, , drop = F]
      Z2_i = Z2[df$subjId == subject_i, , drop = F]

      subject_i_days_length = length(unique(df$day[df$subjId == subject_i]))
      day_matrix_list = vector("list", subject_i_days_length)
      for(index_day in seq(subject_i_days_length)){
        subject_i_days = df$day[df$subjId == subject_i]
        day_j = unique(subject_i_days)[index_day]
        X_alphas_ij = X_alphas_i
        Psi2_ij_tmp = as.numeric(
          exp(X_alphas_ij %*% beta[c("sigma_v^2_log", "alpha")])
        )*diag(u_number2)
        Z2_ij = Z2_i[subject_i_days == day_j, , drop = F]

        matrix_tmp = Z2_ij %*% Psi2_ij_tmp %*% t(Z2_ij)
        day_matrix_list[[index_day]] = matrix_tmp
      }
      Psi2_ij = as.matrix(bdiag(day_matrix_list))

      Sigma_i = as.numeric(
        exp(X_alphas_i %*% beta[c("sigma^2_log", "alpha")])
      )*diag(n_i)
      Sigma_full_i = Sigma_i + Z1_i %*% Psi1_i %*% t(Z1_i) + Psi2_ij

      result = result +
        sum(dmvt(y_i, delta = y_hat_i, type = "shifted",
                 sigma = Sigma_full_i,
                 df = 2*exp(beta["lambda_log"]),
                 log = TRUE))
    }
    return(-result)
  }
```

We also choose to estimate the model with the TMB package with the hierarchical model formulation of equation (26), because then the random effects would be estimated and later compared in section 3.7. The TMB packages estimates the parameters using the Laplace approximation to the marginal log-likelihood for Y_i. The Laplace-approximation gives an approximation to the marginal log-likelihood by integrating a second order Taylor approximation of the joint log-likelihood around the optimum of the random effects, over the random effects. TMB speeds up the optimization by finding the derivatives of the model based on automatic differentiation of the Laplace approximation.

The detailed implementation of the main parts of the likelihood function in C++ is presented:

```
for(int i=0; i < nsubjects; i++){
  f -= dnorm(u[i], mean_random_subject,
             sqrt(exp(sigma2_u_log) *exp(alpha * sex[index]) / gamma[i]),
             true);
}

for(int j=0; j < ndays; j++){
  i = subjectId_day_factor_gamma[j];
  f -= dnorm(v[j], mean_random_day,
             sqrt(exp(sigma2_v_log) *exp(alpha * sex[index]) / gamma[i]),
             true);
}

for(int i=0; i < nsubjects; i++){
  f -= dgamma(gamma[i], exp(lambda), (1/exp(lambda)), true);
}

for(int index = 0; index < nobs; index++){
  i = subjectId_factor[index];
  j = subjectId_day_factor[index];
  f -= dnorm(y[index], (beta[0] + beta[1]*sex[index] + u[i] + v[j]),
             sqrt(exp(sigma2_log) *exp(alpha * sex[index]) / gamma[i]),
             true);
}
```

The estimated parameters for both optimization methods are presented in the table 19. We note that for the regular optimization, we were unable to estimate the parameters uncertainty because the hessian matrix at the optimum was singular, and the inversion of it was impossible. We see that the parameter estimates for the TMB are slightly different than for the regular optimization which is explained by the fact that TMB uses the Laplace approximation which introduces some level of approximation, wheres the estimation through the marginal distribution for Y doesn't introduce those inaccuracies.

	marginal t-student		TMB		
Parameter	Estimate log	Estimate	Estimate log	Std. Error log	Estimate
(Intercept)	-	-0.475	-0.484	0.043	-
sexmale	-	0.199	-0.189	0.053	-
σ_u^2	-3.846	0.0214	-4.035	0.409	0.018
σ_v^2	-3.845	0.0214	-4.246	0.304	0.014
σ^2	-8.8047	0.0002	-8.741	0.333	0.0002
α	-0.7178	0.4878	-0.762	0.335	0.467
λ	-0.984	0.3738	-0.959	0.172	0.383

Table 19: Model parameters and their uncertainties

3.5.1 Derivation of $\gamma_i|Y_i$

Consider

$$f_{\gamma_i|Y_i=y_i} = \frac{f_{Y_i=y_i|\gamma_i} f_{\gamma_i}}{f_{Y_i=y_i}} = \frac{\text{(Multivariate normal)} \cdot \text{(gamma)}}{\text{multivarite t-student}}$$

$$Y_i|\gamma_i \sim N\left(\mu, \frac{\sigma^2}{\gamma_i} I\right); \quad \gamma_i \sim G(1, \lambda)$$

Consider

$$\begin{aligned} f_{Y_i=y_i|\gamma_i} f_{\gamma_i} = \exp\left(-\frac{1}{2}(y_i-\mu)^T \Sigma^{-1} (y_i-\mu)\right) \frac{1}{\sqrt{(2\pi)^k \det(\Sigma)}} \\ \frac{\lambda^\lambda}{\Gamma(\lambda)} \gamma_i^{\lambda-1} \exp\left(-\gamma_i \lambda\right) \end{aligned}$$

Note that

$$\Sigma^{-1} = \left(\frac{\sigma^2}{\gamma} I\right)^{-1} = \frac{\gamma}{\sigma^2} I; \quad \det(\Sigma) = \det\left(\frac{\sigma^2}{\gamma} I\right) = \left(\frac{\sigma^2}{\gamma}\right)^k \det(I) = \frac{\sigma^{2k}}{\gamma^k}$$

Hence

$$\begin{aligned} f_{Y_i=y_i|\gamma_i} f_{\gamma_i} = &\exp\left(-\frac{1}{2}||y_i-\mu||_2^2 \frac{\gamma_i}{\sigma^2}\right) (2\pi)^{\frac{-k}{2}} \frac{\gamma_i^{k/2}}{\sigma^k} \frac{\lambda^\lambda}{\Gamma(\lambda)} \gamma_i^{\lambda-1} \exp\left(-\gamma_i \lambda\right) = \\ &(2\pi)^{\frac{-k}{2}} \frac{\lambda^\lambda}{\Gamma(\lambda)} \gamma_i^{(\lambda+\frac{k}{2})-1} \exp\left(-\gamma_i \left(\frac{1}{2\sigma^2}||y_i-\mu||_2^2 + \lambda\right)\right) \end{aligned}$$

We see the kernel of a Gamma distribution. Thus

$$\gamma_i|Y_i \sim \mathrm{G}\left(\alpha = \lambda + \frac{k}{2}, \beta = 1/\left(\frac{||y_i-\mu||_2^2}{2\sigma^2} + \lambda\right)\right)$$

where the constants in front cancel out with the multivariate t-student distribution in the denominator.

3.6 Final Normal model

Finally, we consider the model

$$\begin{aligned} Y_{i,j,k}|u_i, v_{ij}, \gamma_i = \log(\text{clo}_{i,j,k})|u_i, v_{ij}, \gamma_i &\sim N\left(\mu + \beta\left(\text{sex}_i\right) + u_i + v_{ij}, \sigma^2 \alpha\left(\text{sex}_i\right) e^{-\gamma_i}\right) \\ u_i \mid \gamma_i &\sim N\left(0, \sigma_u^2 \alpha\left(\text{sex}_i\right) e^{-\gamma_i}\right) \\ v_{ij} \mid \gamma_i &\sim N\left(0, \sigma_v^2 \alpha\left(\text{sex}_i\right) e^{-\gamma_i}\right) \\ \gamma_i &\sim N\left(0, \sigma_G^2\right) \end{aligned}$$

where $Y_{i,j,k}$ is the logarithm of the clothing insulation level for subject i on day j and k refer to the observation number within the day. We have that

$$\mathrm{E}\left[Y_{ijk}|\gamma_i\right] = \mu + \beta(\mathrm{sex}_i); \tag{29}$$

$$\mathrm{Cov}\left[Y_{ijk}|\gamma_i, Y_{hlm}|\gamma_h\right] = \begin{cases} (\sigma_u^2 + \sigma_v^2 + \sigma^2)\alpha(\mathrm{sex}_i)e^{-\gamma_i} & \text{for } (i,j,k) = (h,l,m) \quad \text{(same subject, same day, same observation)} \\ (\sigma_u^2 + \sigma_v^2)\alpha(\mathrm{sex}_i)e^{-\gamma_i} & \text{for } (i,j) = (h,l), \quad k \neq m \quad \text{(same subject, same day)} \\ \sigma_u^2\alpha(\mathrm{sex}_i)e^{-\gamma_i} & \text{for } i = h, \quad j \neq l \text{(same subject)} \\ 0 & \text{for } i \neq h \end{cases} \tag{30}$$

We can further write the model in the following form for subject i and day j:

$$\begin{aligned} &\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_{1,i}\mathbf{U} + \mathbf{Z}_{2,i}\mathbf{V} + \varepsilon_i; \\ &\mathbf{U}|\gamma_i \sim N(0, \sigma_u^2\alpha(\mathrm{sex}_i)e^{-\gamma_i}), \quad \mathbf{V}|\gamma_i \sim N_{days_i}(0, \sigma_v^2\alpha(\mathrm{sex}_i)e^{-\gamma_i}I), \quad \varepsilon_i|\gamma_i \sim N_{n_i}(0, \sigma^2\alpha(\mathrm{sex}_i)e^{-\gamma_i}I) \end{aligned}$$

Therefore we can write the model as a multivariate normal distribution

$$\begin{aligned} \mathbf{Y}_i|\gamma_i &\sim N_{n_i}\left(\mathbf{X}_i\boldsymbol{\beta}, \left(\sigma_u^2\mathbf{Z_{1,i}}\mathbf{Z_{1,i}}^T + \sigma_v^2\mathbf{Z_{2,i}}\mathbf{Z_{2,i}}^T + \sigma^2 I\right)\alpha(\mathrm{sex}_i)e^{-\gamma_i}\right) \\ \gamma_i &\sim N(0, \sigma_G^2) \end{aligned} \tag{31}$$

The model is parametrized by $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_u^2, \sigma_v^2, \sigma^2, \alpha, \sigma_G^2)$. We can write the log-likelihood for the model

$$l(\boldsymbol{\theta}; \boldsymbol{y}) = \log L(\boldsymbol{\theta}; \boldsymbol{y}) = \sum_{i=1}^{n} l_i(\boldsymbol{\theta}; \boldsymbol{y}_i) \tag{32}$$

where n is the number of subjects.

The detailed implementation of the main parts of the likelihood function in C++ is presented:

```
for(int i=0; i < nsubjects; i++){
  f -= dnorm(u[i], mean_random_subject,
             sqrt(exp(sigma2_u_log) *exp(alpha * sex[index]) * exp(-gamma[i])),
            true);
}

for(int j=0; j < ndays; j++){
  i = subjectId_day_factor_gamma[j];
  f -= dnorm(v[j], mean_random_day,
             sqrt(exp(sigma2_v_log) *exp(alpha * sex[index]) * exp(-gamma[i])),
             true);
}

for(int i=0; i < nsubjects; i++){
  f -= dnorm(gamma[i], mean_random_gamma, sqrt(exp(sigma2_G_log)), true);
}

for(int index = 0; index < nobs; index++){
  i = subjectId_factor[index];
  j = subjectId_day_factor[index];
  f -= dnorm(y[index], (beta[0] + beta[1]*sex[index] + u[i] + v[j]),
             sqrt(exp(sigma2_log) *exp(alpha * sex[index]) * exp(-gamma[i])),
             true);
}
```

The estimated parameters are presented in the table 20.

Parameter	Estimate log	Estimate	Std. Error log
(Intercept)	-0.4906	0.0407	-
sexmale	-0.1810	0.0533	-
α	-1.1344	0.3545	0.3216
σ_u^2	-2.2760	0.4717	0.1027
σ_v^2	-2.5828	0.3761	0.0756
σ_G^2	1.5996	0.2225	4.9512
σ^2	-6.8391	0.3755	0.0011

Table 20: Model parameters and their uncertainties

3.7 Comparison between models

Figure 13 presents the posterior distribution of $\gamma_i|Y_i$ in the original domain (top row) and the linear domain (bottom row). Based on the plot in the bottom right corner we see that in the linear domain the random effects have a clear linear relationship indicating that the models hierarchical and final normal are not that different.

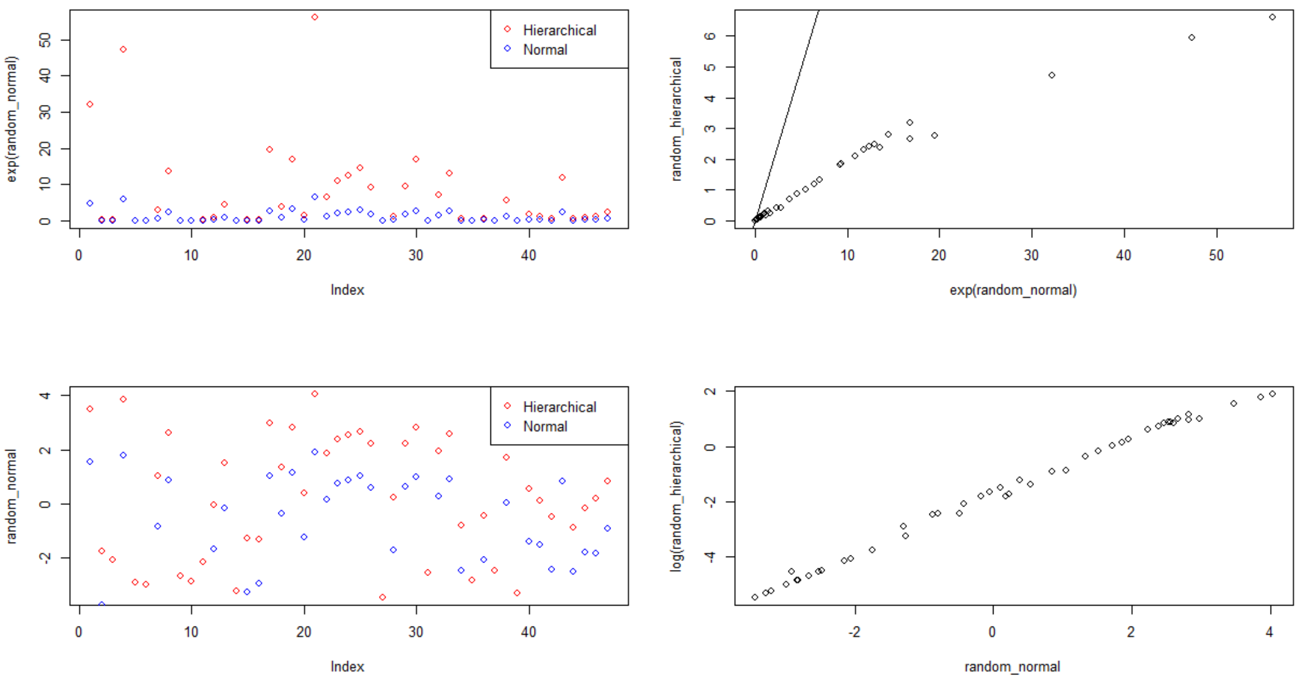

Figure 10: Plots of clo against the different independent variables.

Table 21 presents the likelihoods and the AIC for the models considered in Part 2. We see that the models with random effect in the dispersion for distribution of the dependent variable are significantly better than other models. 21.

Model	log-likelihood	number of parameters	AIC
First	134.7031	4	-261.4063
Second	541.9209	5	-1073.842
Third	769.2759	6	-1526.552
Hierarchical marginal	1216.219	7	-2418.438
Hierarchical TMB	1212.806	7	-2411.612
Final Normal model	1204.501	7	-2395.003

Table 21: Comparison of likelihood and AIC between estimated models

The following QQ plots indicate that the standardised residuals for models with random effects in the variance regression do not have heavy tails. In the following plots the Hierarchical TMB model is indicated with the red color, the Final normal model is indicated with the red color, and the Third TMB model is indicated with the green color.

The table 22 presents the parameter estimates of the hierarchial and the final normal model. We have $a \in \{\lambda, \sigma_G^2\}$. We note that $\gamma_i \sim \mathrm{G}(1, \lambda)$; $\quad Var(\gamma_i) = \frac{1}{\lambda} = 2.611$.

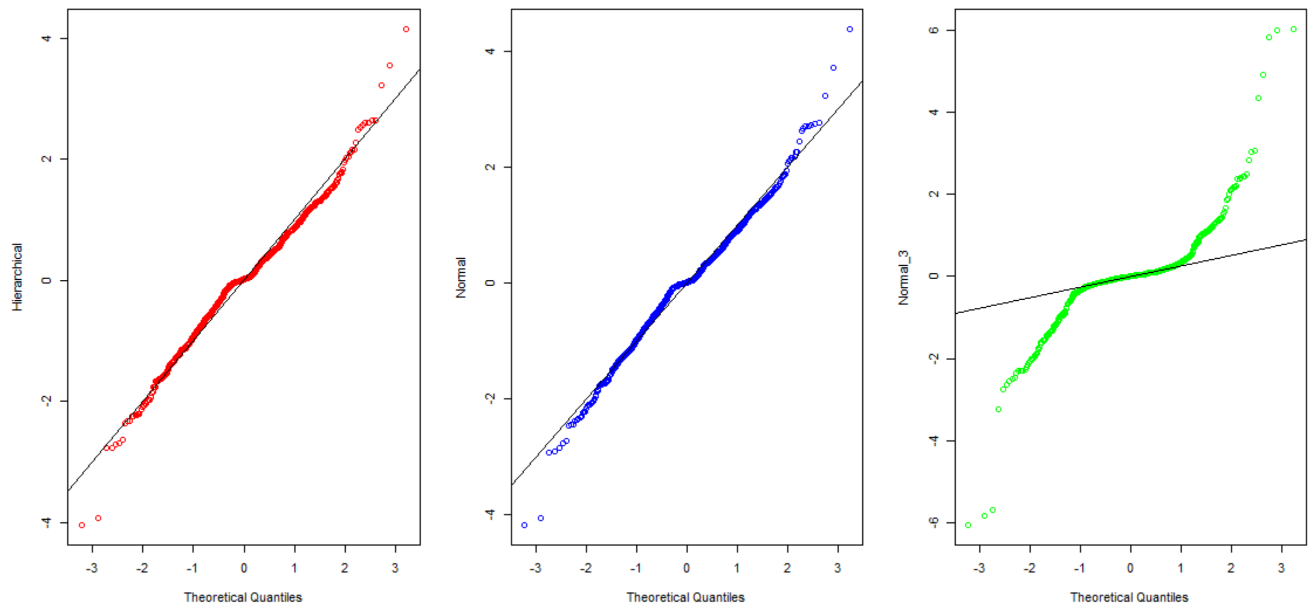

Figure 11: Plots of clo against the different independent variables.

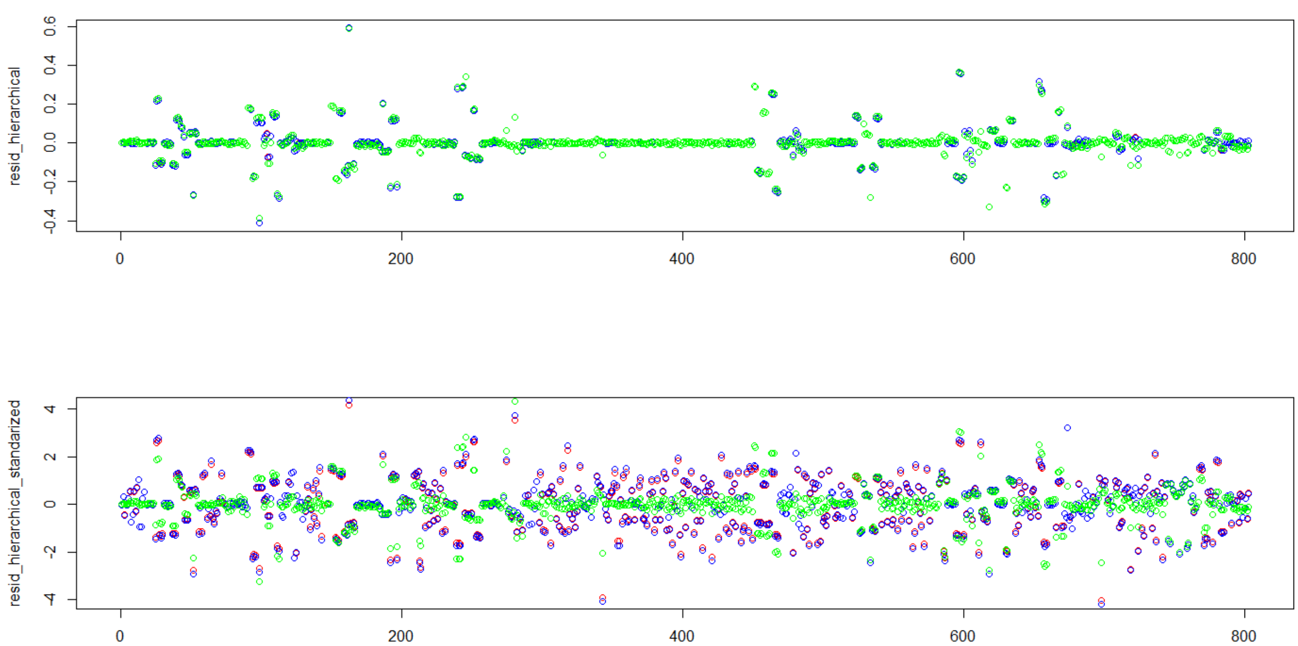

Figure 12: Plots of residuals and standardized residuals

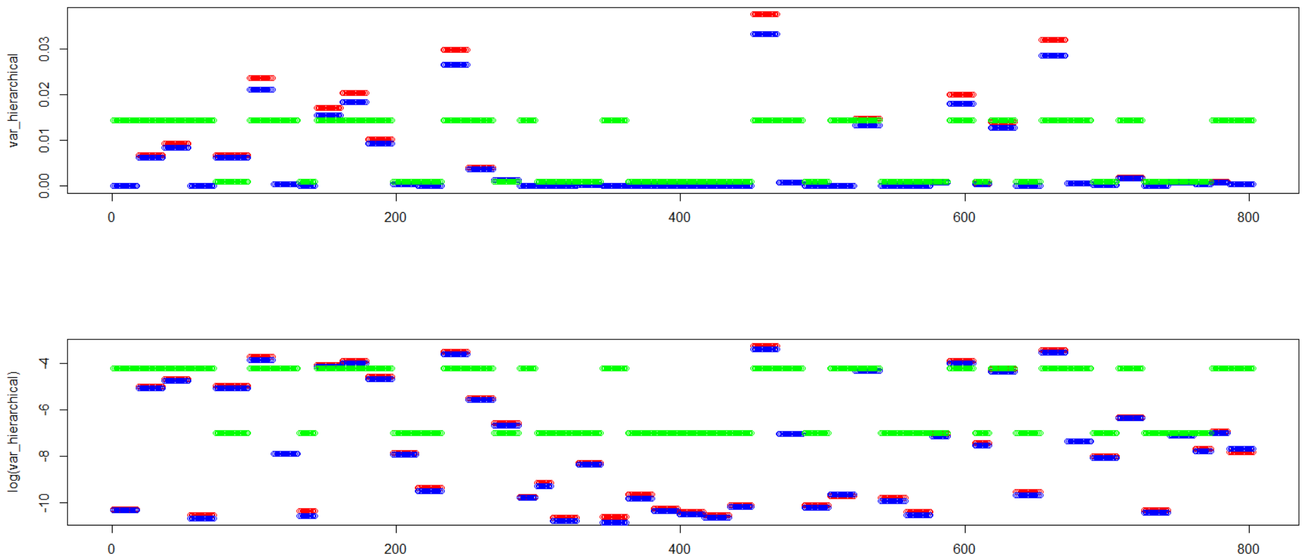

Figure 13: Plots of variance and log variance

	Hierarchical TMB				Final normal TMB			
	log	sd log	regular	sd	log	sd log	regular	sd
μ	-	-	-0.484	0.04	-	-	-0.490	0.04
β	-	-	-0.189	0.05	-	-	-0.181	0.05
α	-0.762	0.335	0.467	-	-1.134	0.355	0.322	-
σ_u^2	-4.035	0.409	0.018	-	-2.276	0.472	0.103	-
σ_v^2	-4.246	0.304	0.014	-	-2.583	0.376	0.076	-
a	-0.959	0.172	0.383	-	1.5996	0.223	4.951	-
σ^2	-8.741	0.333	**0.00015**	-	-6.839	0.376	**0.0011**	-

Table 22: Comparison of parameters between models

4 References

1. Conradsen, K., Christensen, A.M., Nielsen, A.A., Ersbøll, B.K. (2019, v.0.94). *Multivariate Statistics: For the Technical Sciences*. DTU Compute Lyngby.

2. Madsen H., Thyregod P. (2011). *Introduction to General and Generalized Linear Models*. CRC Press.