Eidgenössische Technische Hochschule Zürich

Seminar für
Statistik

CO2 impact on plantlouse-system

Zhou,Xiaobei
Professor: C. B. Schwierz , M. Kalisch and W. A. Stahel

19th January 2011

1 Introduction and Problem Description

This report is based on the reaserch of my client Jeannine Klaiber, who is the sudent of Institute of Plant, Animal and Agroecosystem Sciences of the ETH Zurich.It is to study selected effects of carbon dioxide (CO2) on the Brussels sprouts - cabbage aphid system. The effects of different levels of CO2, which is ambient CO2 and elevated CO2, on selected physical and chemical plant parameters of Brussels sprouts plants grown for an extended period under climatic regimes would be tested in this report. I also take interest on the acceptance of such plants by winged cabbage aphids and subsequent aphid performance.

There are two factors in this report: One is treatment and the other is time. The factor of treatment contains ambient CO2 and elevated CO2, which is the climatic regime the selected plants grow. The time factor includes 7, 11 and 15 weeks, which is the period selected plant parameters were measured and bioassays were conducted.

To quantify whether different level of CO2 influences selected physical and chemical plant traits, 10 plants per treatment were randomly chosen for destructive sampling at each of the three sampling dates. The following physical and chemical parameters were measured on these plants: Physical plant parameters: Stomatal conductance, leaf flourescence, plant height, leaf area and number and total wet/dry weight Chemical plant parameters: Foliar concentration of primary metabolites (C/N content)

Further, in order to test whether Cabbage aphid winged adults show any differential response in terms of host acceptance towards host plants exposed to increased CO2 levels compared to plants grown under ambient CO2 conditions, a "neutral" control plant, which is considered as a non-host plant coming from different species from selected plants , is introduced in the experiment. In the dual-choice wind tunnel bioassay, groups of 30 winged aphids were presented with the choice between a host plant, which is selected plant grown under different CO2 level and a non-host control plant. The number of aphids landed on either plant and the number of nymphs laid by those aphids were recorded after 24hours. Plants containing winged aphids and nymphs were kept for 16 additional days to allow further counting of aphid numbers on those plants. Aphid countings were conducted from day 2 to day 8 (after the day of the bioassay) and at day 16. At each sampling date, bioassays were replicated 8 times for each treatment.

2 Data

Here is a short explanation of the sampling data and parameters in the report. The sample contains totally 54 observations, which is diveded into two datasets:30 observations(a) are to test the significance of effects of different level of CO2 on physiochemical parameters of the plants, and 24 observations(b) are to test the preference and performance of the aphids and their offspring developing.

For the dataset(a), the parameter names used in R are as follow:

- **N** is nitrogen content of the selected plant.
- **C** is carbon content.
- **C.N** is C:N ratio (=nirtogen content / carbon content).
- **height** is plant height in mm.
- **LeafNr** is number of green leaves still attached to the stem.
- **Stomatal cond** is stomatal conductance (indicator of a plants gas exchange capacity).
- **Leaf area** is surface area of a plants leaves in cm2.
- **wet weight** is wet weight in g from all aboveground plantparts.
- **dry weight** is dry weight in g from all aboveground plant parts.

For the dataset(b),there are two paramers:

•**on Host** is percentage of aphid (out of 30) settled on host plant 24hrs after bioassay starts.**on Host** is the parameter to test preference of aphid on selected plant within different level CO2.

•**Day16** is number of nymphs present on a plant on day 16 after Bioassay, produced by the previousely landed aphids (on host plant). **Day16** is to be employed to test the performance of the offspring of aphids.

Addtionally,as discussed on previously introduction part, this report contains two factors: The factor of treatment is defined as **Treatment** and the factor of time is defined as **Weeks** in R. Notation: Both **Treatment** and **Weeks** are dummy variables. The commond **as.factor** is employed to define a factor in R.

3 Analysis

For dataset(a), the scatterplot matrix **Figure1** and **Figure2** present an general overview for the parameters **N**, **C**, **C.N**, **height**, **LeafNr**, **Stomatal cond**, **Leaf area**, **wet weight** and **dry weight**.These parameters are transformed to their logarithms (indicated by first letter l). The collinearity between each parameter is not discussed here, since the purpose of this report is not to sudy the relationship between physiochemical parameters of selected plants.

Moreover, the scatterplot matrix **Figure3** is employeded to plot dataset(b).

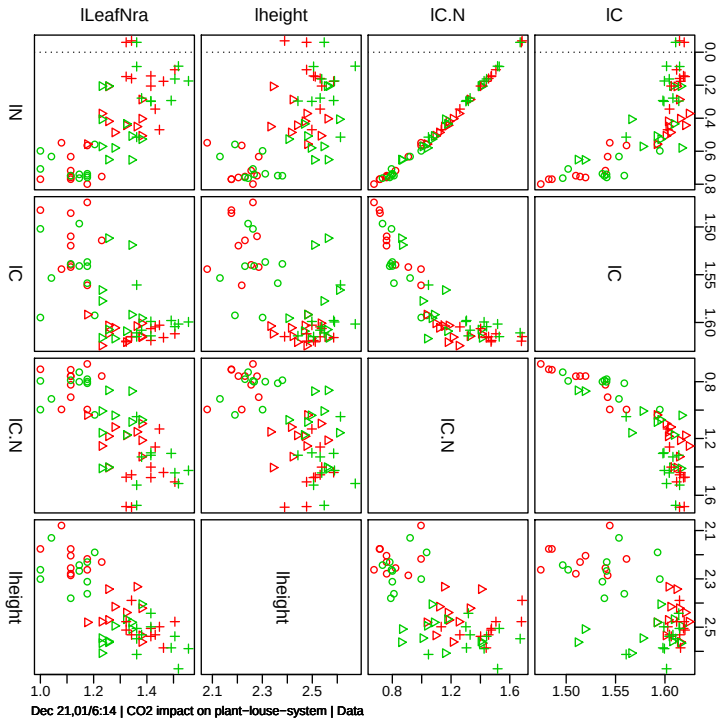
lLeafNra
lheight
lC.N
lC
lN
lC
lC.N
lheight
Dec 21,01/6:14 | CO2 impact on plant−louse−system | Data

Figure 1:

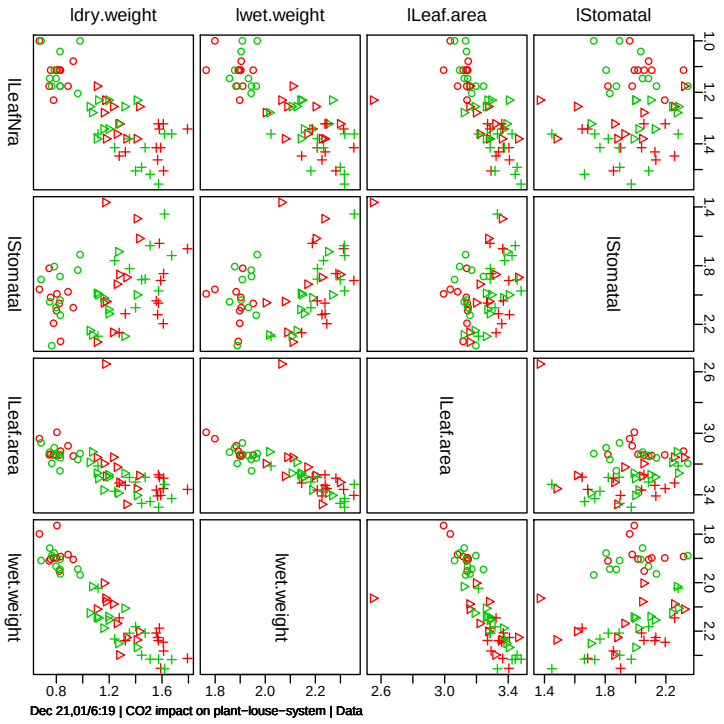

Figure 2: There is an extreme value(the 23th observation) on **lLeaf area**

.

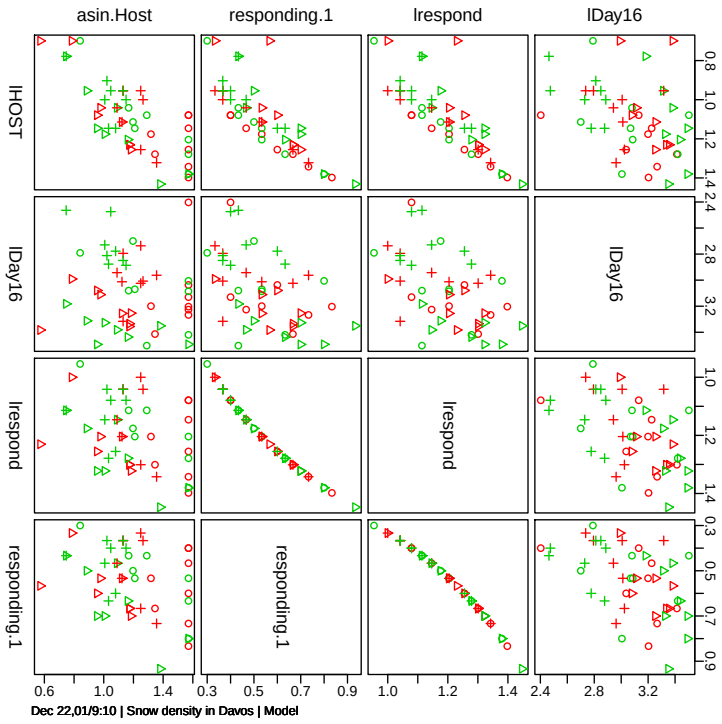

Figure 3:

.

3.1 Introduction to linear regression

The full model(1.0) contains the main effects of **Treatment** and **Weeks** and their interaction effect. The model(1.1) contains the main effects. Additonally, the model(1.11) and (1.12) includes **Treatment** and **Weeks** separately.

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \gamma_1 X_1 X_2 + \epsilon \qquad (1.0)$$

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \epsilon \qquad (1.1)$$

$$Y = \alpha + \beta_1 X_1 + \epsilon \qquad (1.11)$$

$$Y = \alpha + \beta_2 X_2 + \epsilon \qquad (1.12)$$

- β_1=Treatment effect, β_2=Weeks effect, γ=interaction effect, α=intercept X_1=factor(Treatment), X_2=factor(Week), Y=parameter

Incremental F-test is used to test the significance of each effect. The level of significance is 5%(0.05). The interaction effect between **Treatment** and **Weeks** is firstly tested. If the interaction effect is significant, then the main effect is also significant. If the interaction effect is not significant. each main effect has to be tested and interpreted separately.

3.2 Linear regression of physiochemical parameters

For dataset(a), each parameter **N**, **C**, **C.N**, **height**, **LeafNr**, **Stomatal cond**, **Leaf area**, **wet weight** and **dry weight** is pluged in the linear model(1), (1.1), (1.11) and(1.12) in turn until finding a suitable model .For parameter **log10(N)**, which is log_{10}- transformationin of **N** in dataset(a), the full model is:

```
m0.1<-regr(log10(N)∼Treatment+Weeks+Treatment:Weeks,...)
```

```
Terms:
                      coef    stcoef    signif  R2.x df p.value
(Intercept)      0.7064003        NA  8.524892    NA  1   0.000
Treatment       -0.0071923 -0.014663 -0.061375 0.667  1   0.903
Weeks                   NA        NA  3.701483 0.500  2   0.000
Treatment:Weeks         NA        NA  0.352821 0.646  2   0.676
Coefficients for factors:
$Weeks
       7       11       15
 0.00000 -0.29342 -0.54405
$‘Treatment:Weeks‘
     0:7      1:7     0:11     1:11     0:15     1:15
0.000000 0.000000 0.000000 0.068204 0.000000 0.057627
St.dev.error:  0.131   on 54 degrees of freedom
Multiple R^2:  0.744    Adjusted R-squared: 0.721
F-statistic:   31.5   on 5 and 54 d.f.,  p.value: 7.44e-15
```

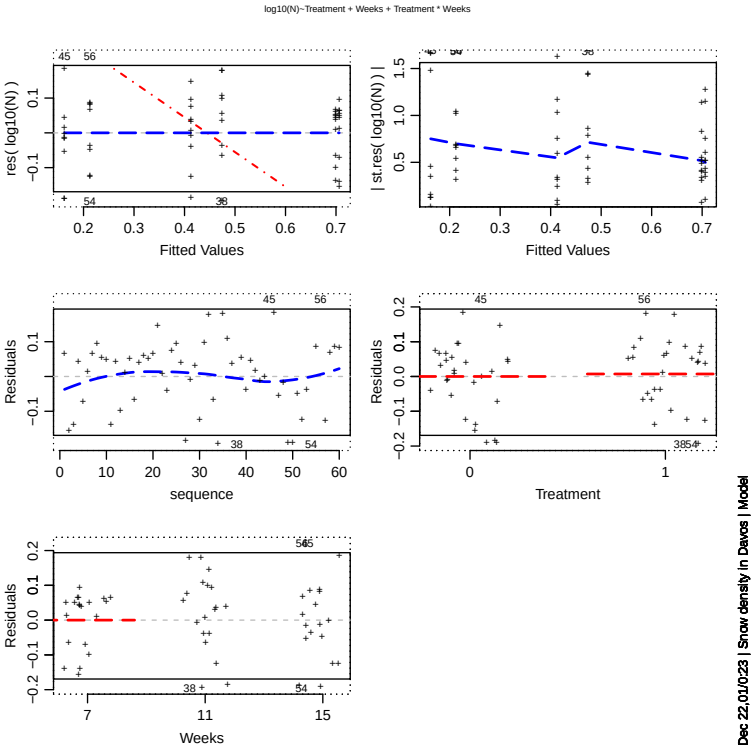

Figure 4: Residual plots show that there is a chance that the interaction effect occurs.

.

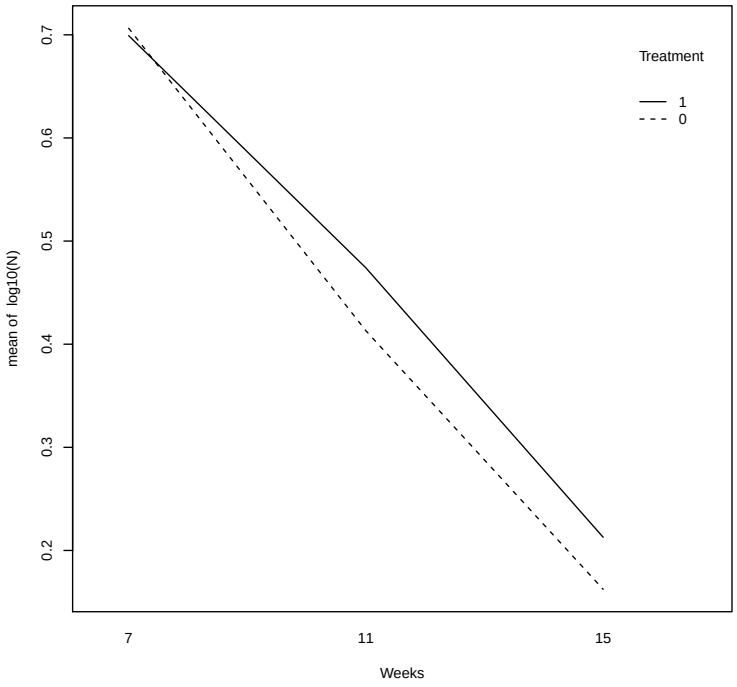

Figure 5: The interaction plot shows no clear interaction effect.

.

The command **drop1(m0.1, test="F")** is used to test interaction effect automatically in R program:

```
Model:
log10(N) ~ Treatment + Weeks + Treatment : Weeks
                Df Sum of Sq   RSS  AIC F value Pr(F)
<none>                       0.922 -238
Treatment:Weeks  2    0.0135 0.936 -242    0.39  0.68
```

Obviously, the interaction effect on parameter **log10(N)** is not significant. Then the reduced model is:

```
m0.2<-regr(log10(N)∼Treatment+Weeks,...)

Terms:
```

```
              coef stcoef signif R2.x df p.value
(Intercept) 0.6854     NA  10.25   NA  1   0.000
Treatment   0.0348 0.0708   0.52    0  1   0.302
Weeks           NA     NA   5.01    0  2   0.000
Coefficients for factors:
$Weeks
     7     11     15
 0.000 -0.259 -0.515
St.dev.error:  0.129   on 56 degrees of freedom
Multiple R^2:  0.741    Adjusted R-squared: 0.727
F-statistic:   53.3   on 3 and 56 d.f.,  p.value: 2.22e-16
```

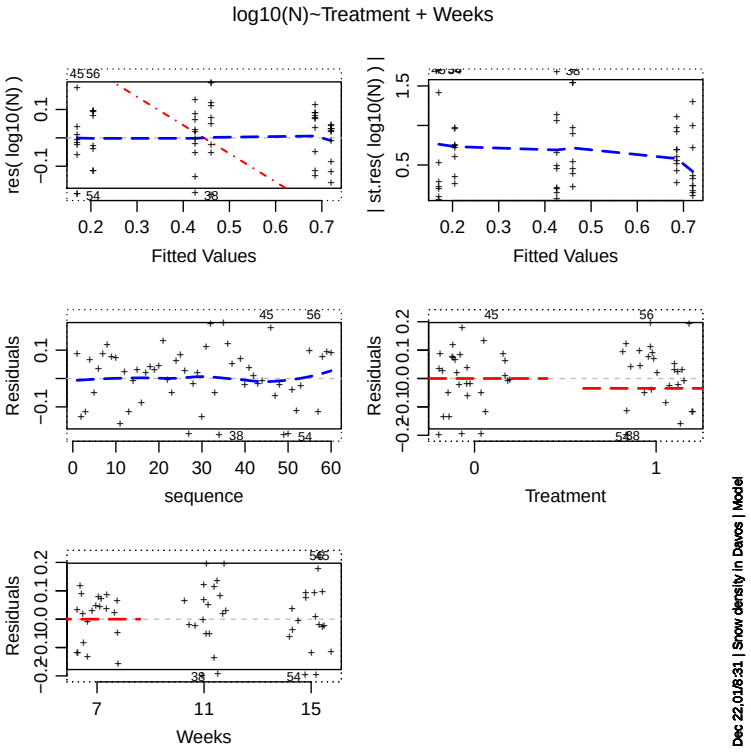

Figure 6: There is an improvement throughout the residual plot.

.

The command **drop1(m0.2, test="F")** is also used to test main effect in R

program:

```
Model:
log10(N) ~ Treatment + Weeks
          Df Sum of Sq  RSS  AIC F value  Pr(F)
<none>                 0.94 -242
Treatment  1     0.018 0.95 -242    1.08    0.3
Weeks      2     2.655 3.59 -165   79.42 <2e-16 ***
```

For parameter **log10(N)**, the effect of **Treatment** is not significant, the effect of **Weeks** is the only factor which is significant. The final reduced model for **log10(N)** is:

```
m0.3<-regr(log10(N)~Weeks,...)

Terms:
             coef stcoef signif R2.x df p.value
(Intercept) 0.703     NA  12.13   NA  1       0
Weeks          NA     NA   5.01    0  2       0
Coefficients for factors:
$Weeks
     7     11     15
 0.000 -0.259 -0.515
St.dev.error:  0.129   on 57 degrees of freedom
Multiple R^2:  0.736    Adjusted R-squared: 0.726
F-statistic:   79.3   on 2 and 57 d.f.,  p.value:    0
```

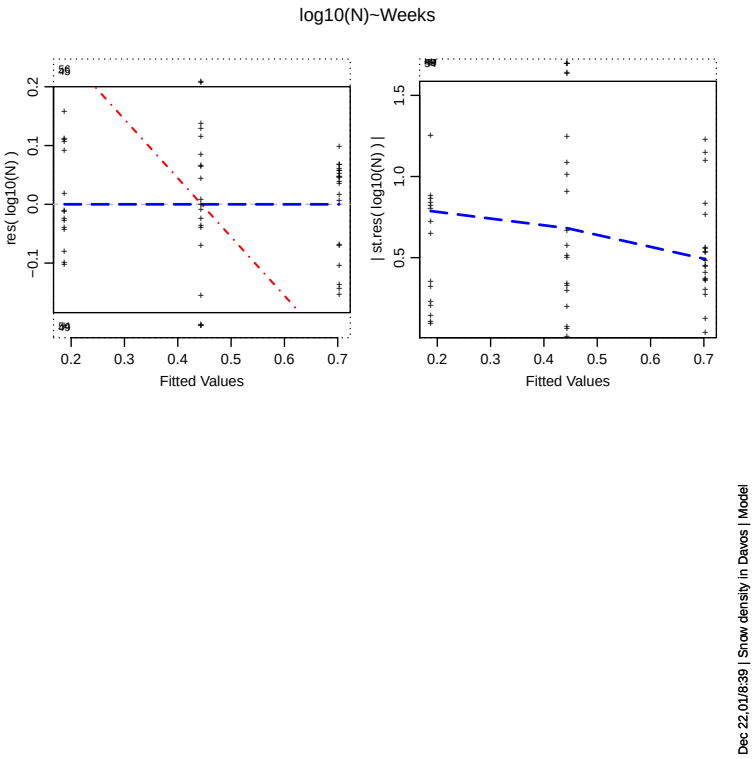

Figure 7:

Notation: The analysis and interpretation of other parameters including **C**, **C.N**, **height**, **LeafNr**, **Stomatal cond**, **Leaf area**, **wet weight** and **dry weight** are presentated in the part of appendix

3.3 Linear regression of aphid preference and performance

For parameter **asin(sqrt(on.Host))**, which is $arc - sine$- transformation of square root of **N** in dataset(b), the final reduced model is:

`m9.7<-regr(asin(sqrt(on.Host))~Weeks,...)`

```
Terms:
            coef stcoef signif R2.x df p.value
(Intercept) 1.41     NA  13.58   NA  1       0
Weeks         NA     NA   2.08    0  2       0
```

```
Coefficients for factors:
$Weeks
     7     11     15
 0.000 -0.360 -0.294
St.dev.error:  0.206   on 45 degrees of freedom
Multiple R^2:  0.382    Adjusted R-squared: 0.354
F-statistic:   13.9   on 2 and 45 d.f.,  p.value: 2e-05
```

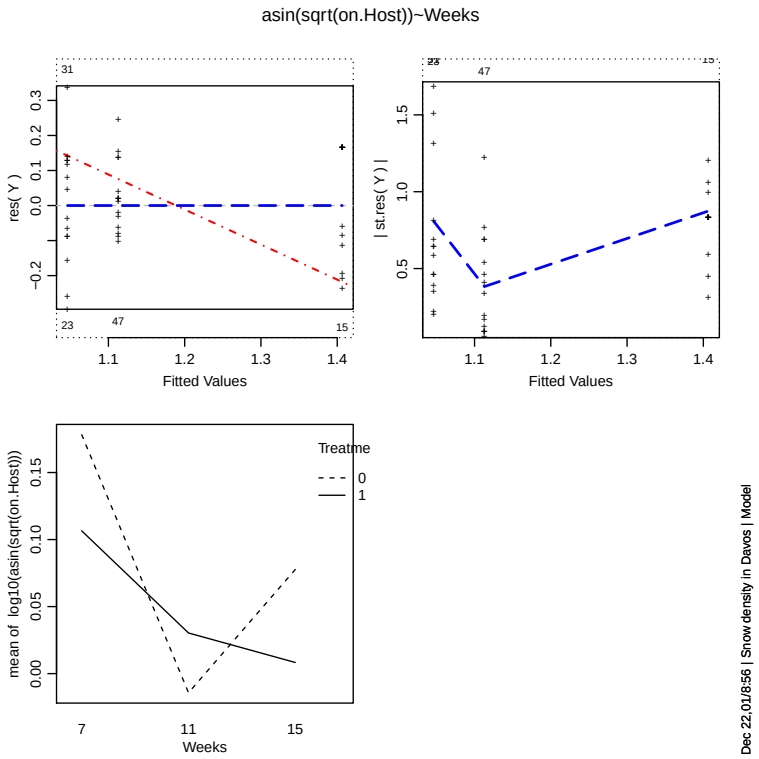

Figure 8:

For parameter **log10(Day16/Host)**, which is log_{10}- transformation of the ratio of **Day16/Host** in dataset(b), the final reduced model is:

```
m11.3<-regr(log10(Day16/Host)~Weeks,...)

Terms:
            coef stcoef signif R2.x df p.value
```

```
(Intercept) 1.93     NA  17.54   NA  1       0
Weeks         NA     NA   2.02    0  2       0
Coefficients for factors:
$Weeks
     7     11     15
 0.000  0.278 -0.103
St.dev.error:  0.218   on 45 degrees of freedom
Multiple R^2:  0.367    Adjusted R-squared: 0.338
F-statistic:     13   on 2 and 45 d.f.,  p.value: 3.45e-05
```

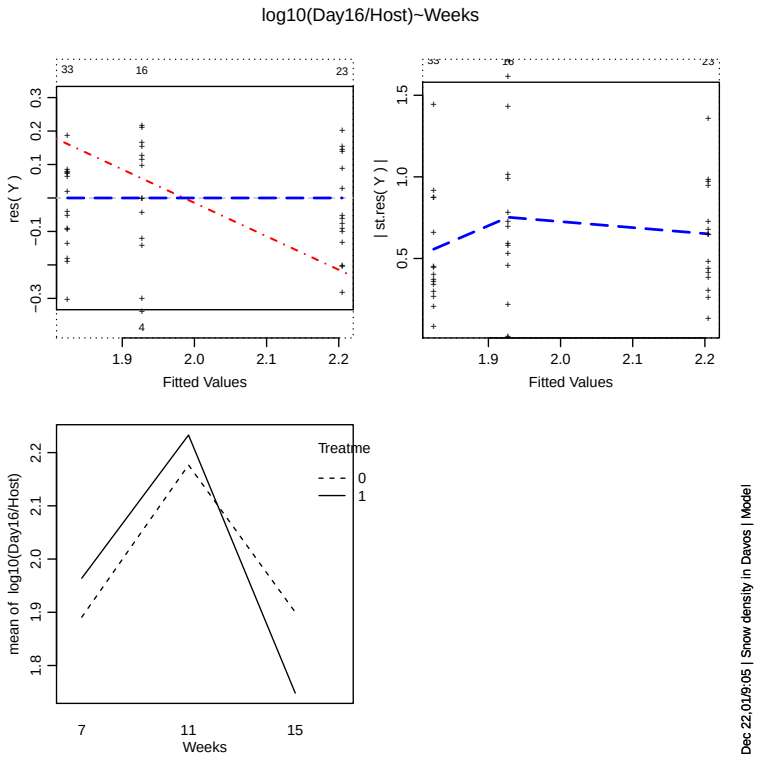

Figure 9:

4 Conclusions

Table1 shows the result of effects of different levels of CO2, which is ambient CO2 and elevated CO2, on selected physical and chemical plant parameters.

Throughout Table1, the factor **treatment** is significant on parameter **log10(C)** and **log10(height)**. And the factor of **Weeks** is working with all the parameters except parameter **log10(Stomatal.cond)**. The interaction effect of both factor **treatment** and **Weeks** is only significant on parameter **log10(C)**.

	Treatment	Week	Interaction
log10(N)	0	+	0
log10(C)	+	+	+
log10(C.N)	0	+	0
log10(height)	+	+	0
log10(LeafNra)	0	+	0
log10(Stomatal)	0	0	0
log10(Leaf.area)	0	+	0
log10(wet.weight)	0	+	0
log10(dry.weight)	0	+	0

Table 1: 0 means not significant and + means significant

The result of effects of different levels of CO2 on the preference and performance of the aphids and their offspring developing is presented on Table2. Here the factor of **treatment** is not significant on **asin(sqrt(on.Host))** and **log10(Day16/Host)**.

	Treatment	Week	Interaction
asin(sqrt(on.Host))	0	+	0
log10(Day16/Host)	0	+	0

Table 2:

5 Reference

"Modern Applied Statistics with S", 4th edition, W.N. Venables, B.D. Ripley, Springer

"Extending the Linear Model with R", J.J. Faraway, Chapman & Hall

Appendix

For dataset(a), the final reduced model of the parameter of **log10(C)** is: `m1.1<-regr(log10(C)~Weeks+Treatment:Weeks,...)`

```
Terms:
                  coef stcoef signif  R2.x df p.value
(Intercept)     1.5173     NA  95.57    NA  1   0.000
Treatment       0.0283  0.332   1.26 0.667  1   0.014
Weeks              NA     NA   3.84 0.500  2   0.000
Treatment:Weeks    NA     NA   1.46 0.646  2   0.002

Coefficients for factors:
$Weeks
     7     11     15
0.0000 0.0919 0.0957

$`Treatment:Weeks`
    0:7     1:7    0:11    1:11    0:15    1:15
 0.0000  0.0000  0.0000 -0.0563  0.0000 -0.0402

St.dev.error:  0.025   on 54 degrees of freedom
Multiple R^2:  0.688    Adjusted R-squared: 0.659
F-statistic:   23.8   on 5 and 54 d.f.,  p.value: 1.41e-12
```

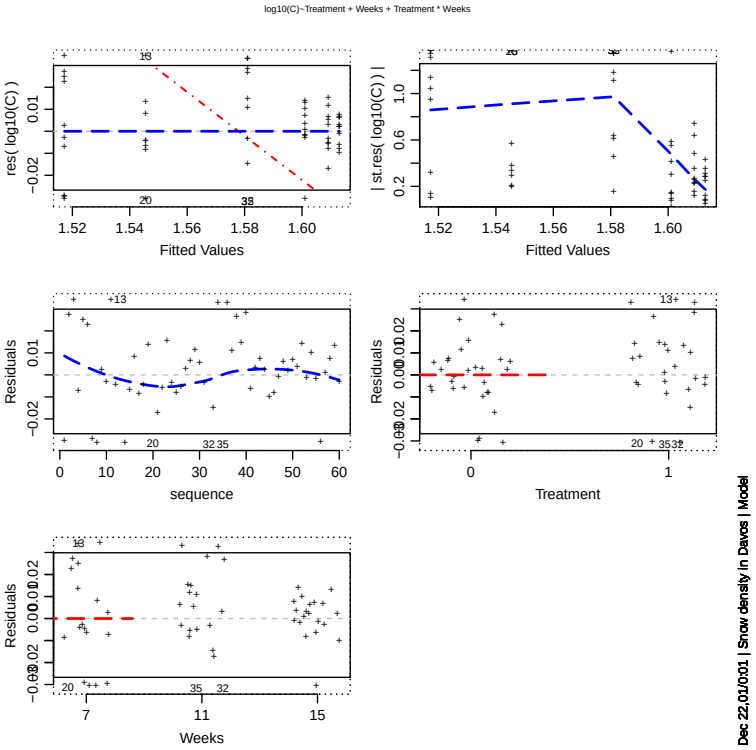
log10(C)~Treatment + Weeks + Treatment * Weeks
res(log10(C))
Fitted Values
| st.res(log10(C)) |
Fitted Values
Residuals
sequence
Residuals
Treatment
Residuals
Weeks
Dec 22,01/0:01 | Snow density in Davos | Model

Figure 10:

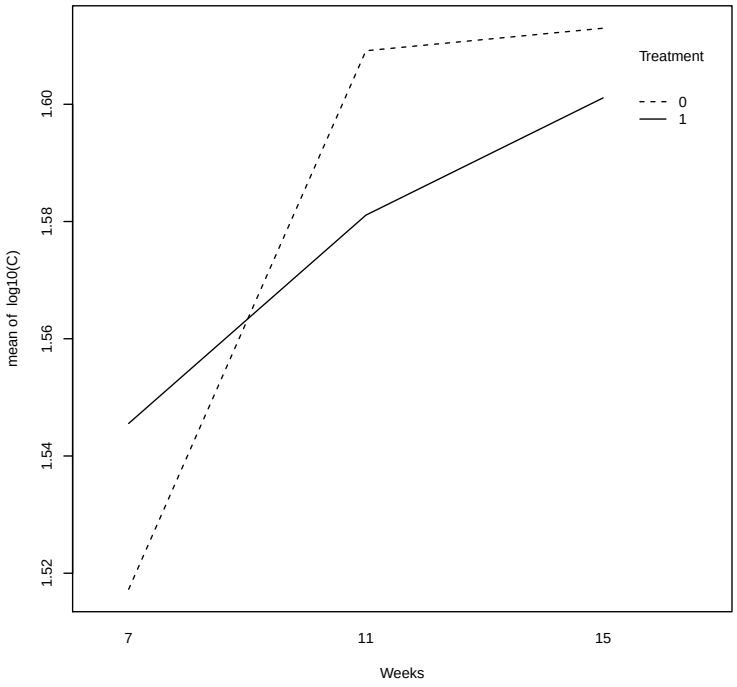

Figure 11:

The final reduced model of the parameter of **log10(C.N)** is:

```
m2.3<-regr(log10(C.N)~Weeks,...)

Terms:
             coef stcoef signif R2.x df p.value
(Intercept) 0.829     NA  12.64   NA  1       0
Weeks          NA     NA   5.09    0  2       0
Coefficients for factors:
$Weeks
    7    11    15
0.000 0.323 0.591
St.dev.error:  0.146   on 57 degrees of freedom
Multiple R^2:  0.741    Adjusted R-squared: 0.732
F-statistic:   81.7   on 2 and 57 d.f.,  p.value:    0
```

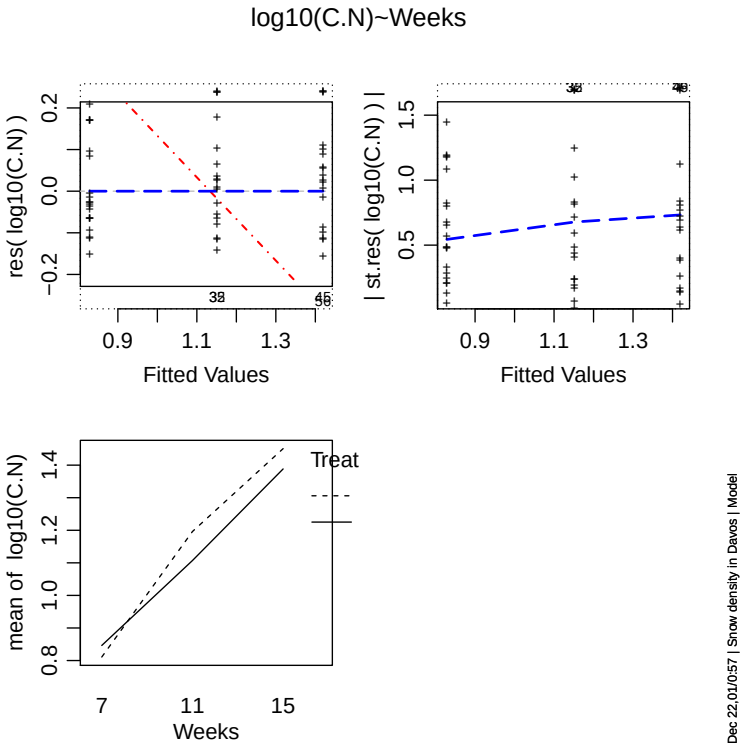

Figure 12:

The final reduced model of the parameter of **log10(height)** is:

```
m3.2<-regr(log10(C)~Weeks+Treatment:Weeks,...)

Terms:
             coef stcoef signif R2.x df p.value
(Intercept) 2.214     NA  66.60   NA  1   0.000
Treatment   0.057  0.199   1.72    0  1   0.001
Weeks          NA     NA   6.02    0  2   0.000
Coefficients for factors:
$Weeks
    7    11    15
0.000 0.241 0.286
St.dev.error:  0.0643   on 56 degrees of freedom
Multiple R^2:  0.812    Adjusted R-squared: 0.801
F-statistic:   80.4   on 3 and 56 d.f.,  p.value:    0
```

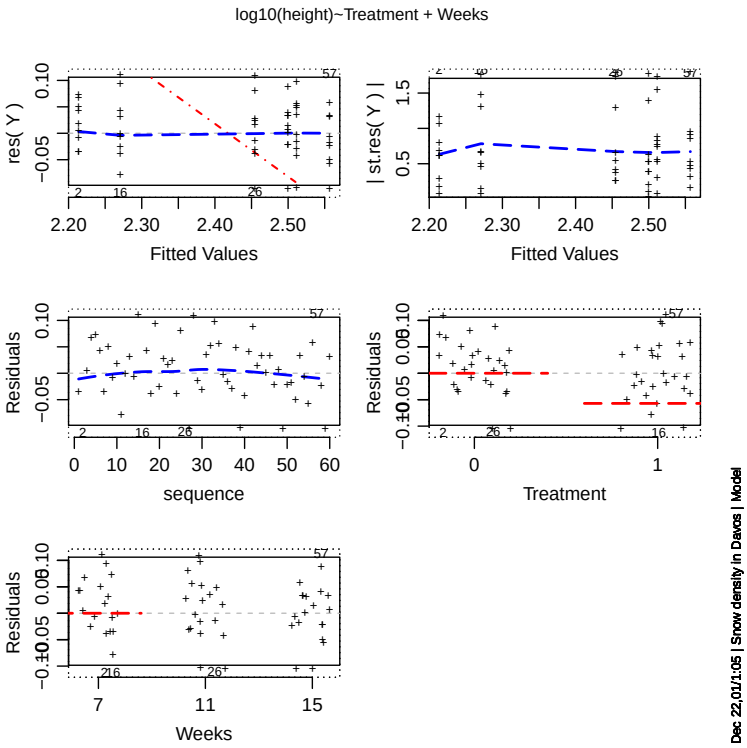
log10(height)~Treatment + Weeks
res(Y)
Fitted Values
| st.res(Y) |
Residuals
sequence
Treatment
Weeks
Dec 22,01/1:05 | Snow density in Davos | Model

Figure 13:

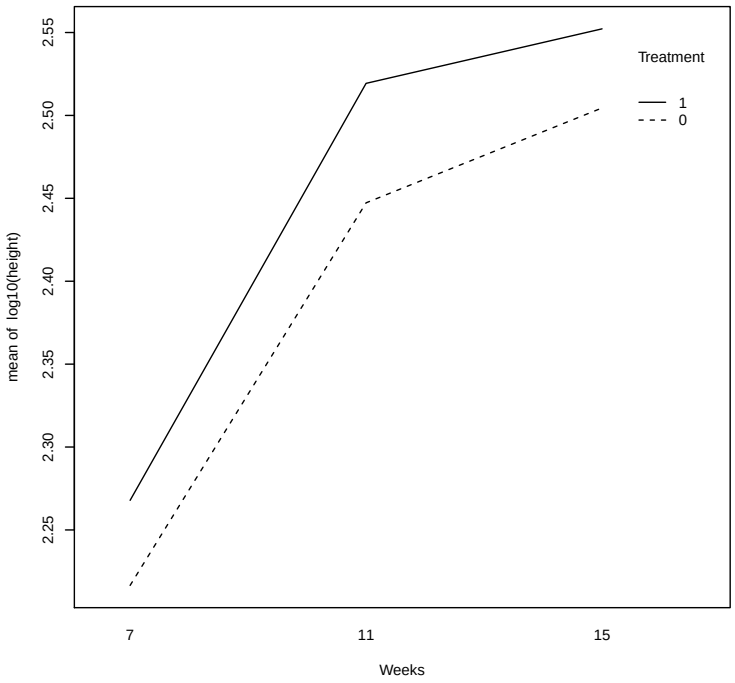

Figure 14:

The final reduced model of the parameter of **log10(LeafNra)** is:

```
m4.3<-regr(log10(LeafNra)~Weeks,...)

Terms:
            coef stcoef signif R2.x df p.value
(Intercept) 1.12     NA   37.2   NA  1       0
Weeks         NA     NA    5.6    0  2       0
Coefficients for factors:
$Weeks
    7    11    15
0.000 0.174 0.298
St.dev.error:  0.0673   on 57 degrees of freedom
Multiple R^2:  0.776    Adjusted R-squared: 0.768
F-statistic:   98.9   on 2 and 57 d.f.,  p.value:    0
```

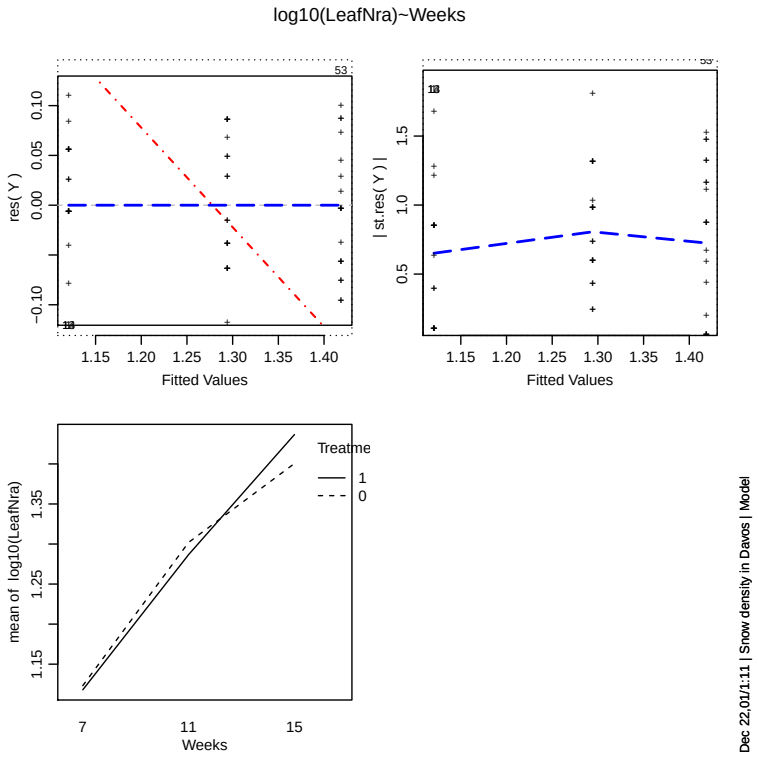

Figure 15:

The final reduced model of the parameter of **log10(Leaf.area)** is:

```
m6.3<-regr(log10(Leaf.area)~Weeks,...)

Terms:
            coef stcoef signif R2.x df p.value
(Intercept) 3.13     NA  58.75   NA  1       0
Weeks         NA     NA   2.38    0  2       0
Coefficients for factors:
$Weeks
    7    11    15
0.000 0.113 0.224
St.dev.error:  0.119   on 57 degrees of freedom
Multiple R^2:  0.385    Adjusted R-squared: 0.363
F-statistic:   17.8   on 2 and 57 d.f.,  p.value: 9.66e-07
```

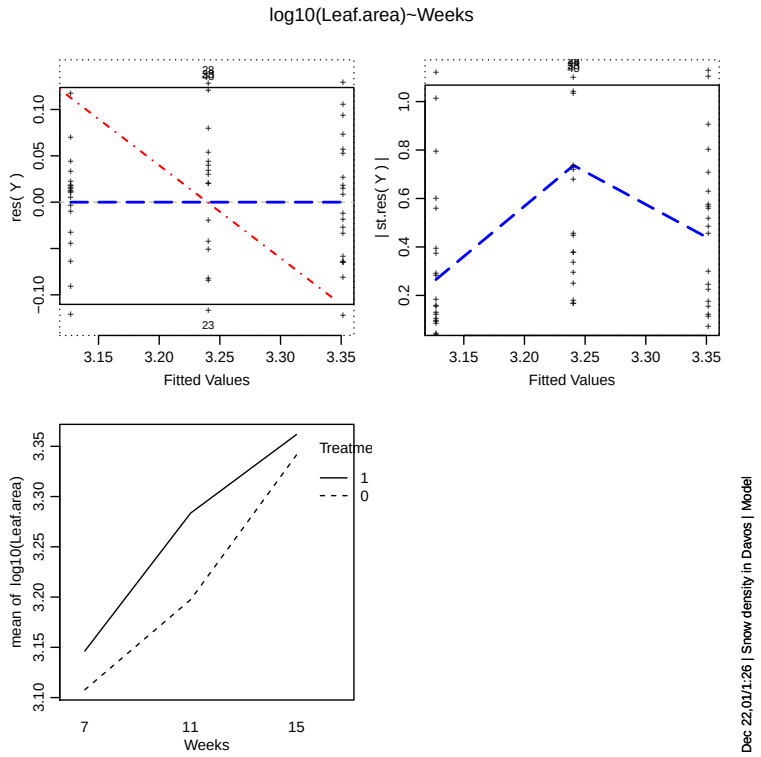

Figure 16:

The final reduced model of the parameter of **log10(wet.weight)** is:

```
m7.3<-regr(log10(wet.weight)~Weeks,...)

Terms:
            coef stcoef signif R2.x df p.value
(Intercept)  1.9     NA  59.35   NA  1       0
Weeks         NA     NA   6.28    0  2       0
Coefficients for factors:
$Weeks
    7    11    15
0.000 0.251 0.345
St.dev.error:  0.0715   on 57 degrees of freedom
Multiple R^2:  0.814    Adjusted R-squared: 0.807
F-statistic:    124   on 2 and 57 d.f.,  p.value:    0
```

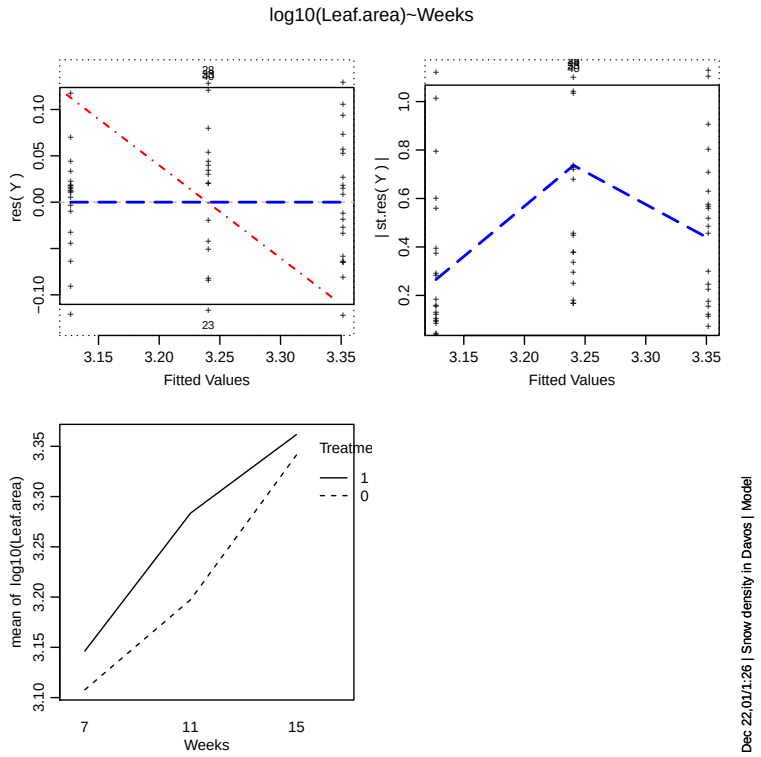

Figure 17:

The final reduced model of the parameter of **log10(dry.weight)** is:

```
m8.3<-regr(log10(dry.weight)~Weeks,...)

Terms:
             coef stcoef signif R2.x df p.value
(Intercept) 0.812     NA  14.56   NA  1       0
Weeks          NA     NA   6.94    0  2       0
Coefficients for factors:
$Weeks
    7    11    15
0.000 0.409 0.682
St.dev.error:  0.125   on 57 degrees of freedom
Multiple R^2:  0.842    Adjusted R-squared: 0.837
F-statistic:    152   on 2 and 57 d.f.,  p.value:    0
```

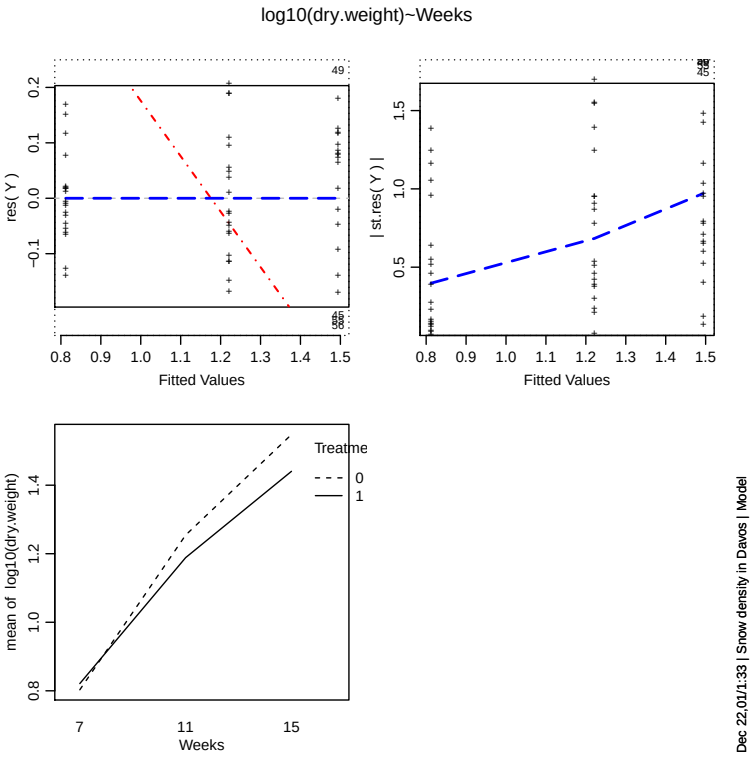
log10(dry.weight)~Weeks
res(Y)
Fitted Values
| st.res(Y) |
Fitted Values
mean of log10(dry.weight)
Weeks
Treatme
0
1
Dec 22,01/1:33 | Snow density in Davos | Model

Figure 18: