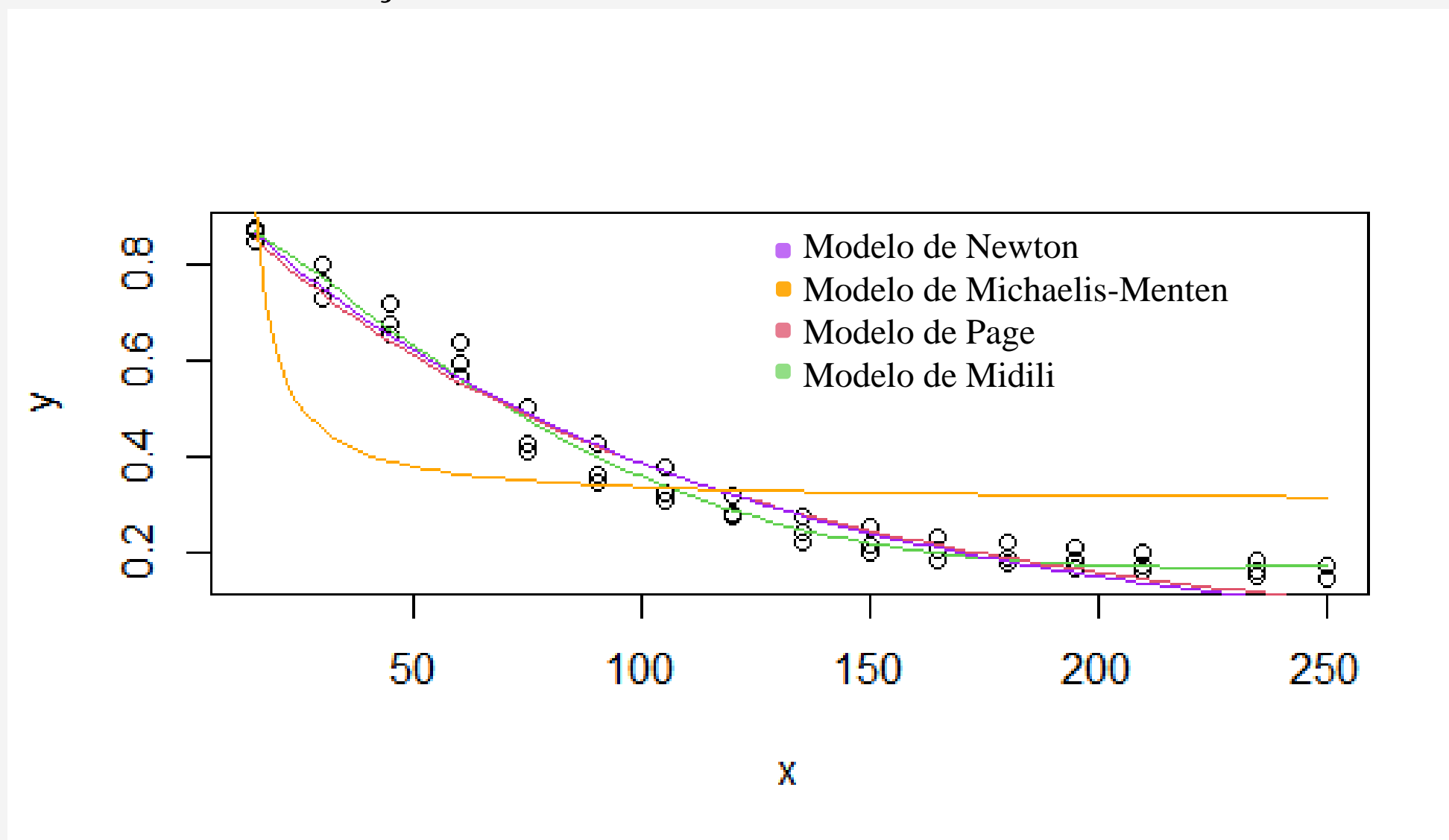


# Seleção de Modelos

**Figura 7:** Gráficos dos modelos ajustados anteriormente aos dados de desidratação das flores de *Viola x wittrockiana*.



Observando os ajustes é difícil identificar qual é o modelo mais adequado. Portanto, para conseguirmos fazer isso, precisamos utilizar alguns indicadores.



# Indicadores


- Para esse curso serão utilizados quatro indicadores:
  - R-quadrado ( $R^2$ )
  - R-quadrado ajustado ( $R_a^2$ )
  - Erro Médio Relativo
  - Erro Médio Estimado



# Coeficiente de determinação ( $R^2$ )

- Fornece a proporção da variação de  $Y$  explicada pela variável  $X$  e mostra o quão “próximo” as medidas reais estão do nosso modelo.
- Essa medida varia de 0 a 1, e quanto mais próxima de 1, maior a precisão do modelo.
- É dado por:

$$R^2 = \frac{S_{XY}^2}{S_{XX} \times S_{YY}}$$



# Coeficiente de determinação ajustado ( $R_a^2$ )

- O  $R_a^2$ , diferente do  $R^2$ , faz uma ponderação pelo número de parâmetros do modelo. Ele é dado por:

$$R_a^2 = 1 - \left[ \frac{(1 - R^2)(n - i)}{n - p} \right]$$

- $R^2$  é o coeficiente de determinação
- $n$  é o número de medições
- $p$  é o número de parâmetros
- $i$  está relacionado com o ajuste do intercepto da curva



# Erro Médio Estimado (EME)

- É utilizado para indicar o desvio dos valores observados em relação a curva estimada pelo modelo.
- É dado pela seguinte equação:

$$EME = \sqrt{\frac{\sum (y - \hat{y})^2}{GLR}}$$

*GLR*: Graus de liberdade do resíduo do modelo ajustado



# Erro Médio Relativo (EMR)

- Da mesma forma que o EME, o EMR também é utilizado para indicar o desvio dos valores observados em relação a curva estimada pelo modelo.

- É dado pela seguinte equação:

$$EMR = \frac{100}{n} \sum_{i=1}^n \left( \frac{|y - \hat{y}|}{y} \right)$$

- Para interpretação dos resultados, de acordo com Radünz et al. (2010), os ajustes médios relativos abaixo de 10% indicam ajuste adequado dos valores observados em relação ao modelo ajustado.



# Cálculo dos indicadores

- Para calcular os indicadores vamos utilizar novamente o RStudio. Como exemplo para realizar os cálculos no R, vamos utilizar o modelo de Newton.





# Cálculo dos indicadores

- Primeiramente, é preciso criar o objeto da regressão no R. No caso do modelo de Newton, isso é feito com seguinte comando:

```
reg1=nls(y ~ exp(-k*x), start=list(k=0.003))
```





# $R^2$ e $R_a^2$

- Para utilizar os próximos comandos, deve-se baixar o pacote qpcR na aba Packages, e chamar o pacote com o comando *library(qpcR)*.
- Com esse pacote é possível calcular tanto o  $R^2$  quanto o  $R_a^2$ , com os comandos:

*Rsq(reg1) #para  $R^2$*

*Rsq.ad(reg1) #para  $R^2$  ajustado*



# $R^2$ e $R_a^2$

- Para o modelo de Newton os resultados obtidos utilizando os comandos foram:

```
> Rsq(reg1)
```

```
[1] 0.9698351
```

```
> Rsq.ad(reg1)
```

```
[1] 0.9698351
```



# Erro Médio Estimado (EME)

- Para o cálculo do EME no R para o modelo de Newton, os comando utilizados são:

```
n<-length(x)
```

```
eme1=(100/n)*sum(abs(pf-fitted(reg1))/y); eme1
```

O resultado obtido foi:

```
[1] 13.59066
```



# Erro Médio Relativo (EMR)

- Para o cálculo do EMR no R para o Modelo de Newton, é necessário primeiro calcular a soma de quadrados do resíduo, com os comandos:

*$yest1=fitted(reg1)$*

*$SQR1=sum((y-fitted(reg1))^2); SQR1$*



# Erro Médio Relativo (EMR)

- Em seguida, para calcular o EMR, os comandos utilizados são os seguintes:

```
n<-length(x)
```

```
p1<-1 #número de parâmetros
```

```
ep1=sqrt(SQR1/(n-p1))
```

```
ep1
```

- O resultado obtido foi:

```
[1] 0.0428095
```

# Seleção de Modelos

**Tabela 2:** Indicadores dos modelos ajustados aos dados de desidratação de folhas de *Viola × wittrockiana*.

Modelo	$R^2$	$R^2_a$	EME	EMR
Newton	0,9698	0,9698	13,59	0,0428
Page	0,9668	0,9661	12,71	0,0425
Page modificado	0,9668	0,9661	12,72	0,0425
Thompson	0,9498	0,9487	14,77	17,56
Henderson e Pabis	0,9695	0,9689	13,57	0,0433
Logarítmico	0,9781	0,9771	8,47	0,0349
Dois termos	0,9805	0,9791	8,12	0,0332
Exponencial de dois termos	0,9678	0,9671	11,66	0,0405
Wang e Sing	0,9790	0,9785	10,14	0,0353
Midilli et al.	0,9843	0,9832	7,5281	0,0299
Michaelis-Menten	0,3750	0,3614	46,4043	0,1608



# Seleção de Modelos

- Na tabela apresentada vemos os resultados dos cálculos de todos os indicadores. A partir deles podemos identificar quais são os melhores ajustes. Entre os modelos apresentados, o melhor ajustado foi o Midili et al.





# Seleção de Modelos: Midili et al.

- Pode-se observar que para este modelo o coeficiente de determinação ajustado foi de 0,9832 e, conforme Silva (2015), valores superiores a 98% indicam um bom ajuste dos modelos aplicados para representação do fenômeno de desidratação.
- O erro médio estimado foi de 7,5281 e o erro médio relativo foi de 0,0299 e, conforme Radünz et al. (2010), os ajustes médios relativos abaixo de 10% indicam ajuste adequado dos valores observados em relação ao modelo ajustado.

# Estimativas

- Para o modelo com a melhor qualidade de ajuste pode ser fornecido um **Intervalo de Confiança** para a perda média de massa da referida curva de desidratação.





# Intervalo de Confiança

- Ele é obtido a partir do cálculo da variância das estimativas dos parâmetros do modelo, dada por:

$$\hat{V}(\hat{\beta}) = (X'X)^{-1}\hat{\sigma}^2$$

- $X$  é a matriz de primeiras derivadas parciais do modelo;
- $\hat{\beta}$  é o vetor de parâmetros do modelo;
- $\hat{\sigma}^2$  é o quadrado médio do resíduo.



# Intervalo de Confiança

- $\hat{\sigma}^2$  é dado por:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (y - \hat{y}_i)^2}{n - p}$$

Em que:

- $n$  é o número de observações;
- $p$  é o número de parâmetros do modelo ajustado.



# Intervalo de Confiança

- O erro padrão da estimativa é dado por:

$$ep(\hat{\beta}_i) = \sqrt{\hat{V}(\hat{\beta}_i)}$$

Em que:

- $\hat{V}(\hat{\beta}_i)$  é um elemento da diagonal principal da matriz de variância-covariância.



# Intervalo de Confiança

- Portanto, o intervalo de 95% de confiança para o parâmetro  $\beta_i$  do modelo é dado por:

$$IC(\beta_i) = \hat{\beta}_i \pm t_{(v;0,025)} \times ep(\hat{\beta}_i)$$

Em que:

- $t_{(v;0,025)}$  é o quantil superior da distribuição  $t$  de Student, com  $\alpha = 5\%$  e graus de liberdade  $v = n - p$ .



# Cálculo do Intervalo de Confiança

- No RStudio, vamos calcular o IC para o modelo de Midilli et al. Primeiro vamos criar a regressão do modelo de Midilli.

```
m=nls(y~a*exp(-k*x^n)+b*x,start=list(a=0.95,    k=0.0021,  
b=-0.00001, n=1.07))  
summary(m)
```



# Cálculo do Intervalo de Confiança

- Em seguida vamos utilizar os seguintes comandos:

*yest=fitted(m)*

*SQR=sum((y-fitted(m))^2); SQR*

*n<-length(x)*

*p<-4* → Em que 4 é o número de parâmetros do modelo

*ep=sqrt(SQR/(n-p))*

*ep*

- *p<-4* pode ser substituído por *p<-length(coef(m));p*



# Cálculo do Intervalo de Confiança

- Com esses comandos obtemos o erro padrão (ep). A partir dele é possível obter o intervalo de confiança. O erro padrão obtido para o modelo de Midilli foi de 0,02993909.



# Cálculo do Intervalo de Confiança

- Em seguida, utilizam-se os seguintes comandos:

```
xnew <- seq(0, 250, length = 10000)
```

```
ynewm <- predict(m, list(x = xnew))
```

- *xnew* corresponde a 10.000 valores entre 0 e 250, e *ynewm* corresponde aos valores do peso final em relação aos 10.000 valores de *xnew* atribuídos ao tempo de secagem.



# Cálculo do Intervalo de Confiança

- Para calcular os valores máximo e mínimo do intervalo de confiança, usamos:

$$LSm = ynewm + qt(0.975, n-p) * ep$$

$$LIm = ynewm - qt(0.975, n-p) * ep$$



# Cálculo do Intervalo de Confiança

- Agora vamos construir o gráfico com o intervalo de confiança usando os comandos:

```
plot(x,y)
```

```
lines(xnew,ynewm, lwd=1,col="green")
```

```
lines(xnew, LIm, lwd=1, lty=2, col="green")
```

```
lines(xnew, LSm, lwd=1, lty=2, col="green")
```

```
legend("topright",
```

```
legend=c("Estimated", "95% CI"),
```

```
lwd=c(1,1), lty=c(1,2),
```

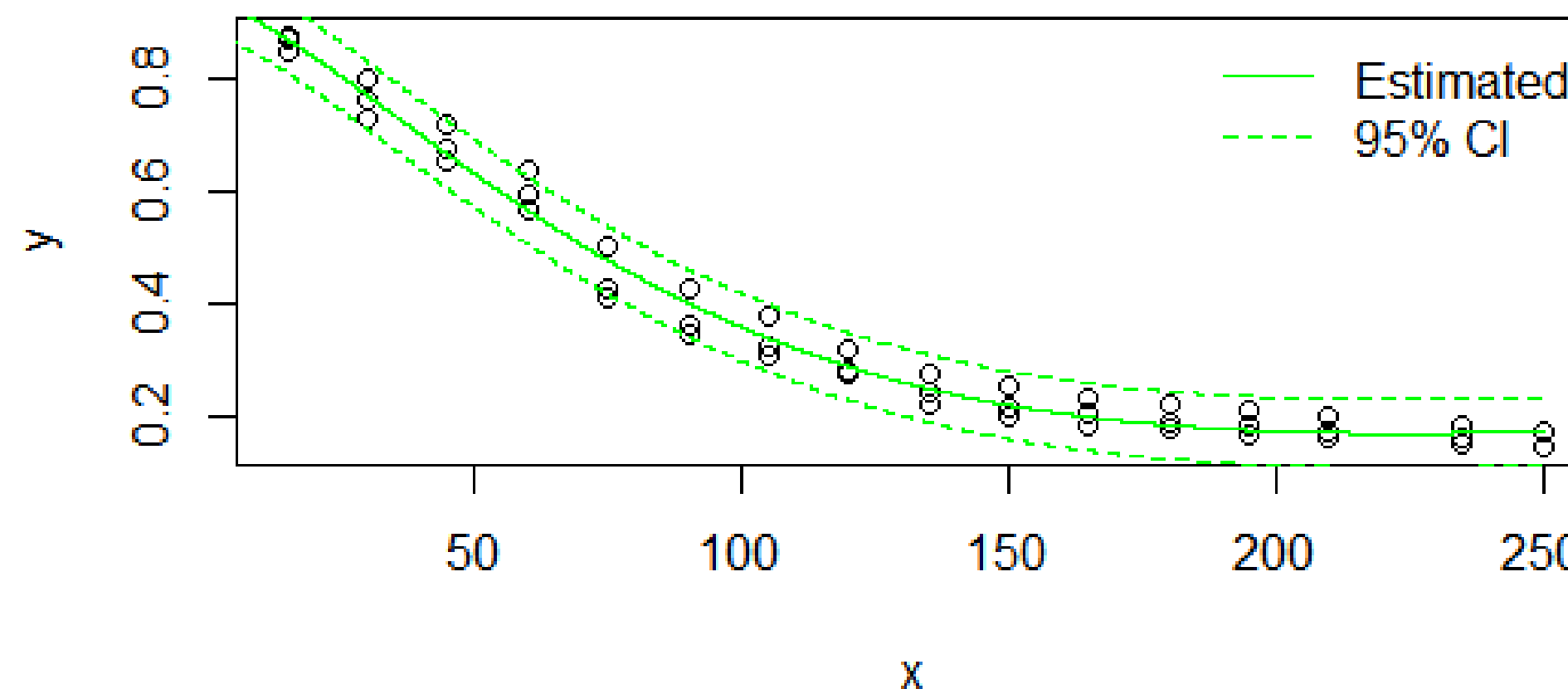
```
col=c("green","green"),
```

```
bty="n", cex=1)
```

# Cálculo do Intervalo de Confiança

- Assim, obtemos a curva com o Intervalo de Confiança, representado pelas linhas pontilhadas:

**Figura 8:** Modelo de Midilli et al. ajustado aos dados de secagem de *Viola x wittrockiana* com o intervalo de confiança de 95% .





# Previsão de um valor particular

- Para prever um valor particular para os dados de secagem, podemos utilizar o seguinte comando:

```
predict(m,list(x=c( )))
```

Ao adicionar um valor no parênteses, o R nos entrega o peso final estimado para determinado tempo.





# Previsão de um valor particular

- Por exemplo, caso se queira saber o peso final para um tempo de secagem de 150 minutos, devemos usar o comando da seguinte maneira:

```
predict(m,list(x=c(150)))
```

Assim encontramos que para 150 minutos de secagem, o peso final é de 0,2185733 gramas.



# Referências

AVEN, M., FOLDES, F. The Chemical Kinetics of Procaine and Chlorprocaine Hydrolysis. Science, 24 Aug 1951. <https://doi.org/10.1126/science.114.2956.206.b>

CHARNET, et al. Análise de modelos de regressão linear: com aplicações. 2. ed. Campinas: Editora da UNICAMP, 2008. 368p

FERNANDES, T. Curva de crescimento do fruto do cafeeiro em diferentes alinhamentos de plantio utilizando modelos não lineares. Dissertação (Mestrado) - Universidade Federal de Lavras. Lavras, 2012.

FONSECA, J. Composição centesimal e modelagem da curva de Desidratação para perda de massa de *Viola × wittrockiana*. Trabalho de Conclusão de curso (Curso de Ciência e Tecnologia de Alimentos) - Universidade Federal do Pampa. Itaquí, 2018.

GASPARIN, P. P. (2012). Secagem da *Mentha piperita* em leito fixo utilizando diferentes temperaturas e velocidades de ar.



# Referências

RADÜNZ, L. L. et al. Avaliação da cinética de secagem de carqueja. Revista engenharia na agricultura - REVENG, v. 19, n. 1, p. 19–27, 26 jan. 2011. doi10.13083/reveng.v19i1.147

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

RITZ, C.; STREIBIG, J. C. Nonlinear Regression with R. 2009. 151p. <https://doi.org/10.1007/978-0-387-09616-2>

SILVA, L A.; ZAMBIAZI, R. C.; CHAVES, F. C.; FISCHER, S. Z. (2015). Avaliação das condições de extração sobre o teor de compostos fenólicos totais de *Viola x wittrockiana*. In: Congresso Brasileiro de Processamento mínimo e Pós-colheita de frutas, flores e hortaliças, 001. Aracaju-SE.