



PRΣTAP
PROGRAMA DE ESTATÍSTICA APLICADA

Comemoração do Dia do Estatístico

Turbinando o Excel com o R

José Francisco Moreira Pessanha

professorjfmp@hotmail.com

Universidade do Estado do Rio de Janeiro

Rio de Janeiro, 1 Junho de 2012

FreeFoto.com

Introdução



O R e o Excel são duas poderosas ferramentas computacionais para análise de dados.

Conhecer as possibilidades de integração entre estes dois programas contribui para agilizar a análise de dados, melhorar a apresentação de resultados e ainda possibilita desenvolver aplicativos.

Objetivo: descrever diferentes maneiras de estabelecer uma integração entre o R e o Excel

Para ilustrar a integração entre os programas será considerado um caso estudo sobre energia eólica.



Base de dados: registros anemométricos disponibilizados na página do Projeto Sonda (INPE/CPTEC).

Programa R

- Linguagem e ambiente computacional para fazer análises estatísticas e mineração de dados.
- É gratuito e open source.
- Disponibiliza uma variedade de funções para análise estatística (regressão linear e não linear, testes estatísticos, análise de séries temporais, estatística multivariada, planejamento de experimentos, etc.).
- Disponibiliza funções para a elaboração de diversos tipos de gráficos, úteis na análise exploratória de dados e visualização de resultados.
- É altamente extensível.
- Tem sido utilizado com sucesso em diversas aplicações da estatística nas mais diversas áreas do conhecimento.
- Rápida difusão e segundo algumas estimativas conta com cerca de 2 milhões de usuários.



MS Excel



- **Planilha eletrônica muito útil e fácil de aprender.**
- **Organiza e armazena as informações em forma tabular.**
- **Excelente interface com o usuário e um bom instrumento para comunicação e apresentação de informações.**
- **Fácil integração com outras bases de dados. Suporta programação em VBA.**
- **Contém funções estatísticas: distribuições de probabilidade, geração de números aleatórios, regressão linear, ANOVA, programação linear e não linear (Solver), etc.. Sem contar Add-Ins que podem incorporados.**
- **Forma conveniente de fazer What-If analyses.**
- **É um dos softwares mais utilizados no mundo.**

Como estabelecer uma integração entre o R e o Excel ?

Duas formas diferentes de integrar o R ao Excel

1) R como uma extensão do Excel (Excel as Host)

O usuário trabalha com uma versão da planilha Excel enriquecida pelas diversas funções estatísticas do R. (O R é embutido no Excel)

- Adição de novos menus e novos itens no menu
Produce resultados que não são atualizados automaticamente pelo Excel
- Novas funções que podem ser usadas como fórmulas do Excel
Tem a vantagem de permitir a atualização automática dos resultados.
- Funções embutidas em Macros VBA e que podem ser executadas quando solicitadas pelo usuário.

2) Excel acessado a partir do R (R as Host)

O usuário trabalha com o R e tem a possibilidade de acessar (controlar) o Excel e utilizá-lo como um local onde o R pode ler os dados de entrada e gravar os resultados das análises estatísticas. (O R controla o Excel)

Como fazer a integração entre o R e o Excel ?

Diferentes maneiras e níveis de integração

1) Pacotes **xlsReadWrite** e **xlsx**

Basicamente leitura/gravação de arquivos Excel .xls e .xlsx.

2) Pacote **Excel.Link**

Manuseia arquivos Excel a partir do R (R as host).

3) Pacote **XLConnect**

Manuseia arquivos Excel a partir do R (R as host).

3) Programas **ExcelDNA** e **R.NET**

Cria funções no Excel capazes de executar scripts no R (Excel as host).

5) Pacote **Rexcel**

Interface plena entre o R e o Excel

Estudo de caso sobre energia eólica

- A energia eólica é um recurso natural, limpo, abundante e inesgotável.
- O seu aproveitamento para produção de energia elétrica é uma das alternativas mais promissoras para reduzir as emissões dos gases do efeito estufa e mitigar os efeitos de crises no mercado de combustíveis fósseis.
- Rápida expansão da capacidade instalada da geração eólica no mundo, passando de 24.322 MW em 2001 para 159.213 MW em 2009, (*World Wind Energy Report 2009 - WWEA, 2010*)

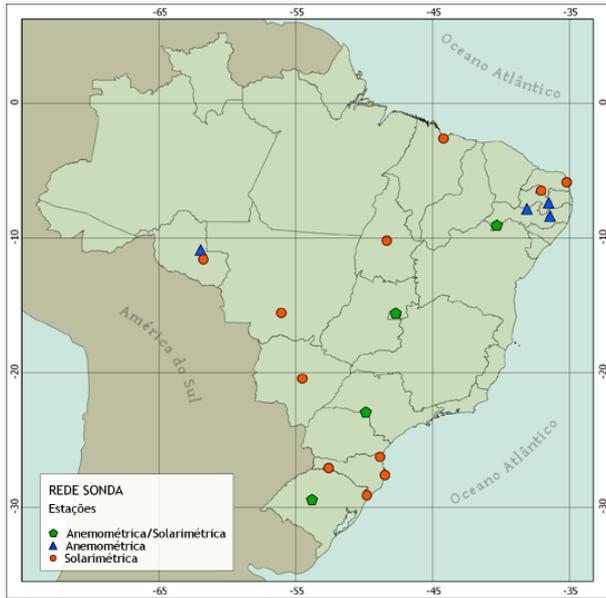


- Maior parte (48%) na Europa (76,218 MW),
Ásia 39.961 MW (26 GW estão na China)
América do Norte 38.748 MW (35.159 MW estão nos EUA).
Brasil, capacidade instalada de 600 MW em 2009.
- Potencial eólico brasileiro estimado em 143,5 GW a 50 metros de altura (Maior parte na região Nordeste – 52%).
- Importante complemento ao regime de geração hidrelétrica nos períodos de estiagem, contribuindo para assegurar a segurança energética do País e diversificar a matriz energética brasileira.

Projeto Sonda – Sistema de Organização Nacional de Dados Ambientais (INPE/CPTEC)



<http://sonda.ccst.inpe.br/>



ESTAÇÃO DE SÃO MARTINHO DA SERRA

(Incluída nas redes AERONET e BSRN)

Dados Ambientais

2012											
Jan	Feb	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
3.1	3.1	3.1	3.1								

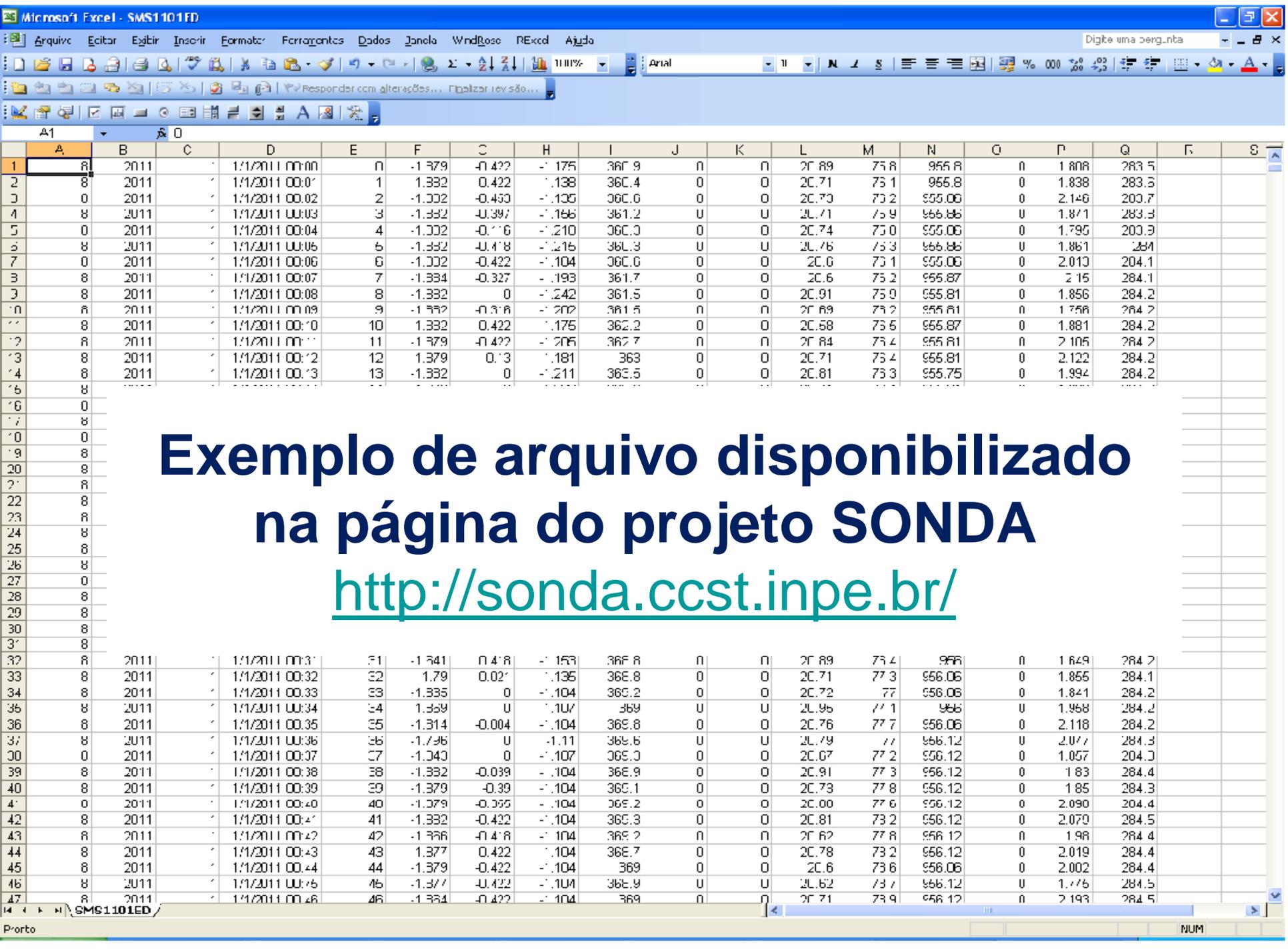
2011											
Jan	Feb	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1

2010				
ago	Set	Out	Nov	Dez
3.1	3.1	3.1	3.1	3.1

ago	Set	Out	Nov	Dez
3.0	3.0	3.0	3.0	3.0

ago	Set	Out	Nov	Dez
3.0	3.0	3.0	3.0	3.0

ago	Set	Out	Nov	Dez
3.0	3.0	3.0	3.0	3.0



Exemplo de arquivo disponibilizado na página do projeto SONDA

<http://sonda.ccst.inpe.br/>

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	8	2011	1/1/2011 00:00	0	-1.379	-0.422	-1.175	367.9	0	0	20.89	73.8	955.8	0	1.808	283.5			
2	8	2011	1/1/2011 00:01	1	1.332	0.422	-1.138	366.4	0	0	20.71	73.1	955.8	0	1.838	283.5			
3	0	2011	1/1/2011 00:02	2	-1.302	-0.450	-1.105	366.0	0	0	20.70	73.2	955.00	0	2.146	200.7			
4	8	2011	1/1/2011 00:03	3	-1.352	-0.397	-1.166	361.2	U	U	21.71	73.9	955.85	U	1.871	283.3			
5	0	2011	1/1/2011 00:04	4	-1.302	-0.416	-1.210	366.0	0	0	20.74	73.0	955.00	0	1.795	200.9			
6	8	2011	1/1/2011 00:05	5	-1.352	-0.418	-1.215	361.3	U	U	21.76	73.3	955.85	U	1.861	284			
7	0	2011	1/1/2011 00:06	6	-1.302	-0.422	-1.104	366.0	0	0	20.6	73.1	955.00	0	2.010	204.1			
8	8	2011	1/1/2011 00:07	7	-1.334	-0.327	-1.193	361.7	0	0	20.6	73.2	955.87	0	2.15	284.1			
9	8	2011	1/1/2011 00:08	8	-1.332	0	-1.242	361.5	0	0	20.91	73.0	955.81	0	1.856	284.2			
10	8	2011	1/1/2011 00:09	9	-1.352	-0.318	-1.202	361.5	0	0	20.89	73.2	955.81	0	1.758	284.2			
11	8	2011	1/1/2011 00:10	10	1.332	0.422	-1.175	362.2	0	0	20.58	73.5	955.87	0	1.881	284.2			
12	8	2011	1/1/2011 00:11	11	-1.379	-0.422	-1.205	362.7	0	0	20.84	73.4	955.81	0	2.105	284.2			
13	8	2011	1/1/2011 00:12	12	1.379	0.413	-1.181	363	0	0	20.71	73.4	955.81	0	2.122	284.2			
14	8	2011	1/1/2011 00:13	13	-1.352	0	-1.211	365.5	0	0	20.81	73.3	955.75	0	1.994	284.2			
15	8																		
16	0																		
17	8																		
18	0																		
19	8																		
20	8																		
21	8																		
22	8																		
23	8																		
24	8																		
25	8																		
26	8																		
27	0																		
28	8																		
29	8																		
30	8																		
31	8																		
32	8	2011	1/1/2011 00:31	31	-1.341	0.418	-1.153	366.8	0	0	20.89	73.4	955	0	1.849	284.2			
33	8	2011	1/1/2011 00:32	32	-1.379	0.021	-1.135	366.8	0	0	20.71	73.3	956.06	0	1.855	284.1			
34	8	2011	1/1/2011 00:33	33	-1.335	0	-1.104	365.2	0	0	20.72	77	956.06	0	1.841	284.2			
35	8	2011	1/1/2011 00:34	34	1.359	U	-1.107	369	U	U	21.95	77.1	956	U	1.958	284.2			
36	8	2011	1/1/2011 00:35	35	-1.314	-0.004	-1.104	365.8	0	0	20.76	77.7	956.06	0	2.118	284.2			
37	8	2011	1/1/2011 00:36	36	-1.395	U	-1.11	365.5	U	U	21.79	77	956.12	U	2.077	284.3			
38	0	2011	1/1/2011 00:37	37	-1.340	0	-1.107	365.0	0	0	20.67	77.2	956.12	0	1.057	204.0			
39	8	2011	1/1/2011 00:38	38	-1.332	-0.039	-1.104	365.9	0	0	20.91	77.3	956.12	0	1.83	284.4			
40	8	2011	1/1/2011 00:39	39	-1.379	-0.39	-1.104	365.1	0	0	20.73	77.8	956.12	0	1.85	284.3			
41	0	2011	1/1/2011 00:40	40	-1.379	-0.355	-1.104	365.2	0	0	20.00	77.6	956.12	0	2.090	204.4			
42	8	2011	1/1/2011 00:41	41	-1.332	-0.422	-1.104	365.3	0	0	20.81	73.2	956.12	0	2.070	284.5			
43	8	2011	1/1/2011 00:42	42	-1.356	-0.418	-1.104	365.2	0	0	20.62	77.8	956.12	0	1.98	284.4			
44	8	2011	1/1/2011 00:43	43	1.377	0.422	-1.104	365.7	0	0	20.78	73.2	956.12	0	2.019	284.4			
45	8	2011	1/1/2011 00:44	44	-1.379	-0.422	-1.104	369	0	0	20.6	73.6	956.06	0	2.002	284.4			
46	8	2011	1/1/2011 00:45	45	-1.377	-0.422	-1.104	365.9	U	U	21.52	73.7	956.12	U	1.775	284.5			
47	8	2011	1/1/2011 00:46	46	-1.354	-0.422	-1.104	369	0	0	20.71	73.9	956.12	0	2.193	284.5			

Pacote xlsReadWrite

Read and write Excel files (.xls)



Pacote xlsReadWrite: Lê e escreve arquivos Excel (.xls)

Autor: Hans-Peter Suter

Principais funções disponibilizadas:

`Objeto_R <- read.xls('nome do arquivo',nº da planilha,primeira linha da planilha)`

`dados <- read.xls('F:/SCR0601WD.xls', sheet=1, from=1)`

`dateTimeToStr(data, "formato da data")` ⇒ **converte datas do Excel para string**

`dateTimeToStr(dados[,4], 'dd/mm/yyyy hh:nn')`

`write.xls(variável, 'nome do arquivo',nº da planilha,primeira linha da planilha)`

`write.xls(saida, 'F:/saida.xls', sheet=1, from=1)`

Pacote xlsReadWrite

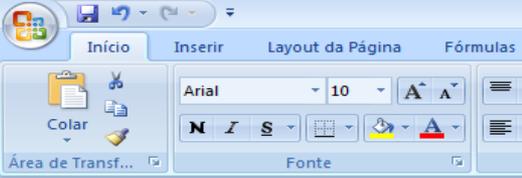
Código em R para extração dos registros de velocidade do vento a 25 e 50 metros de altura

```
#carrega pacote xlsReadWrite  
library(xlsReadWrite)
```

```
# leitura do arquivo de dados do SONDA  
dados<-data.frame(read.xls('F:/SCR0601WD.xls',sheet=1,from=1))  
dados[,4]<-dateTimeToStr(dados[1,4],'dd/mm/yyyy hh:nn')
```

```
# Seleciona velocidade do vento a 25m e 50 m  
# de altura ( colunas 6 e 9 de dados )  
velocidade <- dados[,c(6,9)]
```

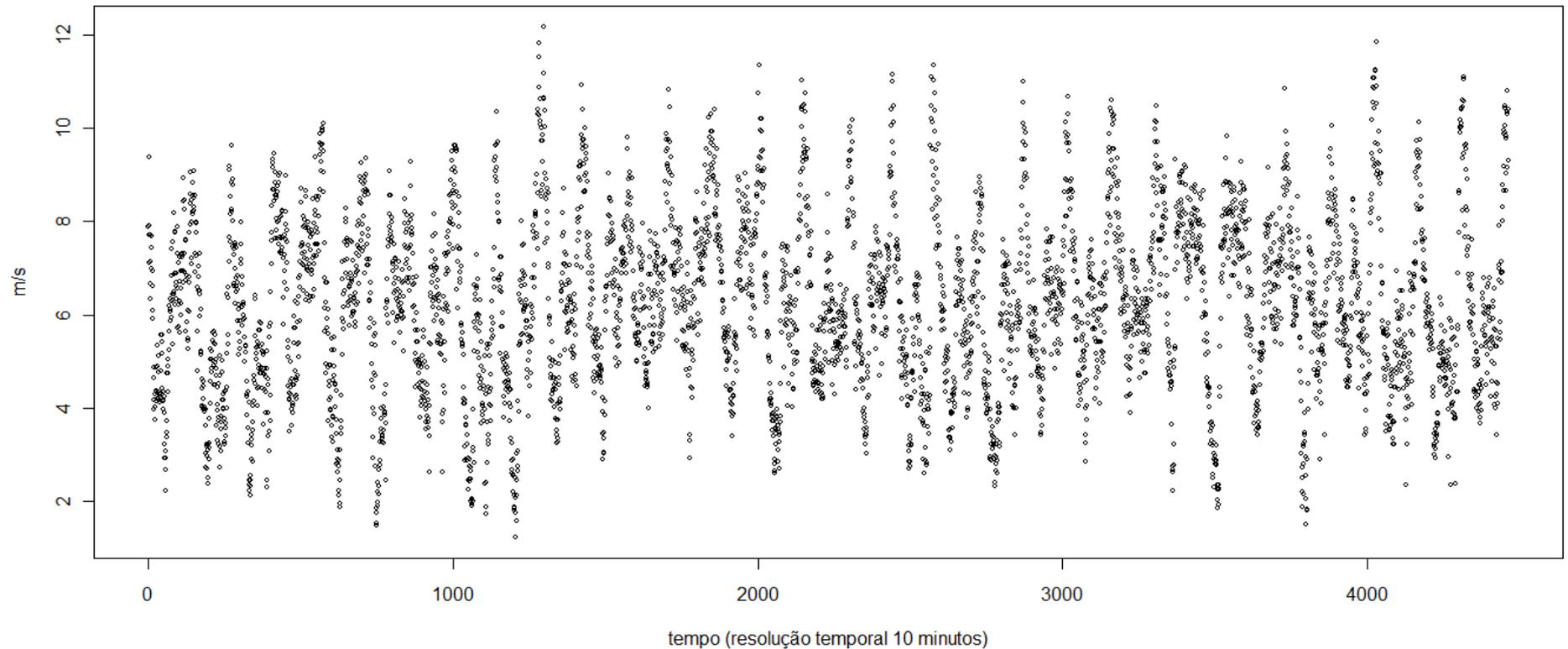
```
# grava em uma pasta separada as velocidades  
# a 25m e 50m  
saida<-data.frame(cbind(dados[,4],velocidade))  
write.xls(saida,'F:/saida.xls', sheet=1, from=1)
```



	A	B	C	D
1		ws_25	ws_50	
2	01/01/2006 00:00	6.78	7.88	
3	01/01/2006 00:00	6.6	7.9	
4	01/01/2006 00:00	8.38	9.39	
5	01/01/2006 00:00	6.2	7.12	
6	01/01/2006 00:00	6.6	7.7	
7	01/01/2006 00:00	6.62	7.7	
8	01/01/2006 00:00	5.26	6.33	
9	01/01/2006 00:00	5.81	6.67	
10	01/01/2006 00:00	6.06	7.15	
11	01/01/2006 00:00	6.56	7.69	
12	01/01/2006 00:00	6.46	7.44	
13	01/01/2006 00:00	6.18	7	
14	01/01/2006 00:00	6.06	6.95	
15	01/01/2006 00:00	5.22	5.93	
16	01/01/2006 00:00	5.55	6.61	
17	01/01/2006 00:00	4.97	6.1	
18	01/01/2006 00:00	3.94	4.87	
19	01/01/2006 00:00	3.76	4.65	
20	01/01/2006 00:00	3.46	4.18	
21	01/01/2006 00:00	2.93	3.97	

Gráfico gerado pelo R para a velocidade do vento a 50 metros de altura

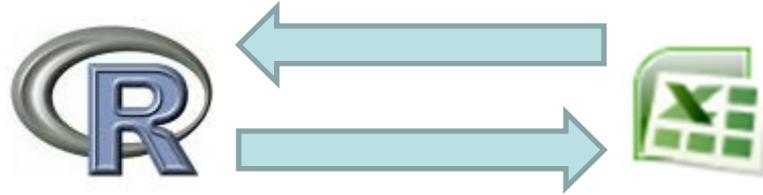
```
plot(velocidade[,2])
```



Pacote excel.link (R as Host)

Convenient way to work with data in Microsoft Excel

R é o programa
mestre



Pacote excel.link: Manuseia arquivos Excel a partir do R

Autor: Gregory Demin

Disponibiliza funções no R para manusear planilhas Excel, por exemplo:

`xl.workbook.add()` # abre um novo arquivo Excel a partir do R

`xl.workbook.open('F:/SCR0601WD.xls')` # abre um arquivo Excel

`xl.sheet.add('resultados')` # adiciona pasta no arquivo Excel

`xl.workbook.save("nome do arquivo")` grava um arquivo Excel

`xl.sheet.delete("nome da pasta")` # remove uma pasta do arquivo Excel

`xl[a1] <- objeto_R` # insere valor do objeto_R na célula a1 da planilha

Filtragem dos valores de velocidade do vento a 50 metros de altura

```
library(excel.link) # carrega pacote excel.link
```

```
# carrega script R com funções para filtragem de dados (LOESS)  
source('c:/turbinando/exemplo_excellink/rotinas_para_filtragem.txt')
```

```
# abre arquivo saida.xls a partir do R  
xl.workbook.open('h:/saida.xls')  
velocidade<-xl[c2:c4465] # velocidade a 50 metros de altura
```

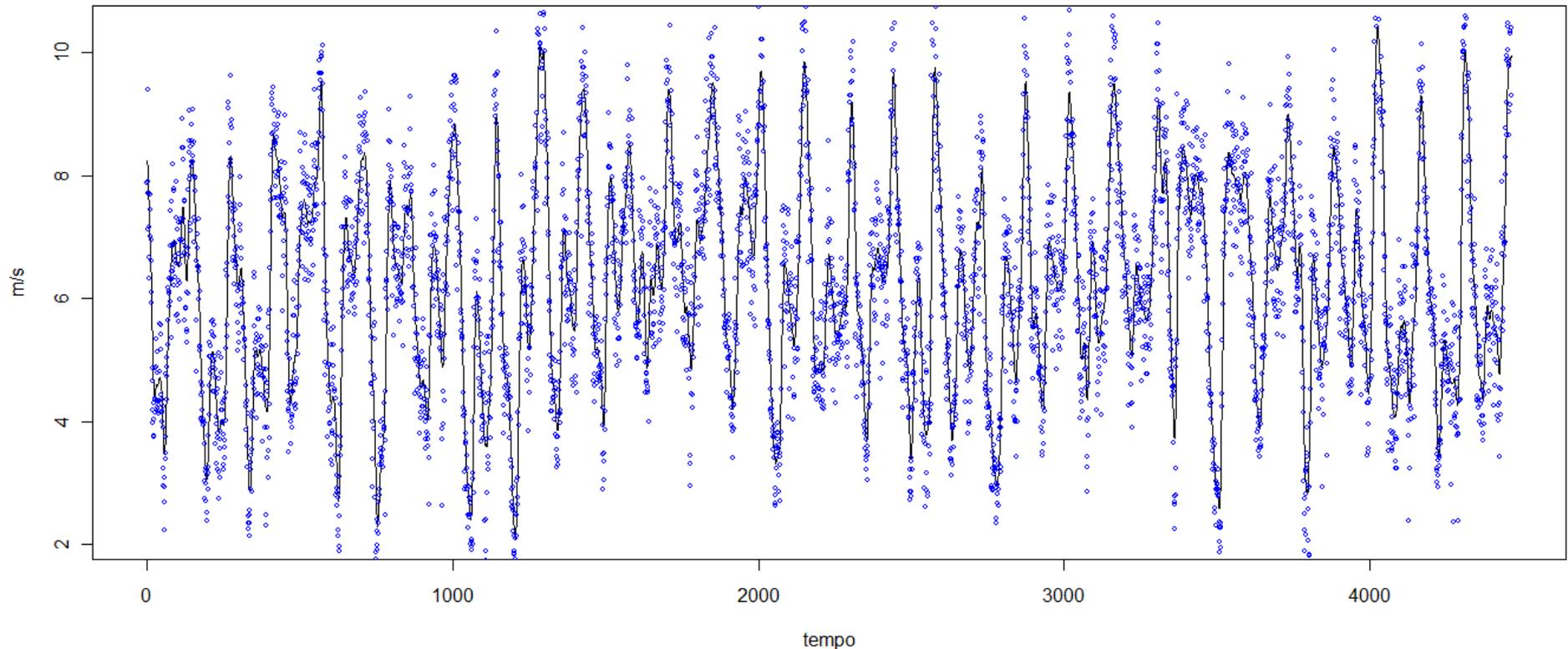
```
# executa funções para filtrar as medições  
# alfa constante de alisamento para o método LOESS  
resultado_filtragem<-filtragem(velocidade,alfa= xl[e1],3,0,'resultados')
```

```
xl.sheet.add('graficos') # adiciona uma worksheet chamada “graficos”  
xl[a1]=current.graphics() # insere o gráfico corrente na worksheet “graficos”
```

```
# inspeciona outliers  
xl.sheet.activate('resultados') # ativa worksheet “resultados”  
graficozoom(instanteselecionado=xl[h1],janela=20,velocidade,resultado_filtragem  
)
```

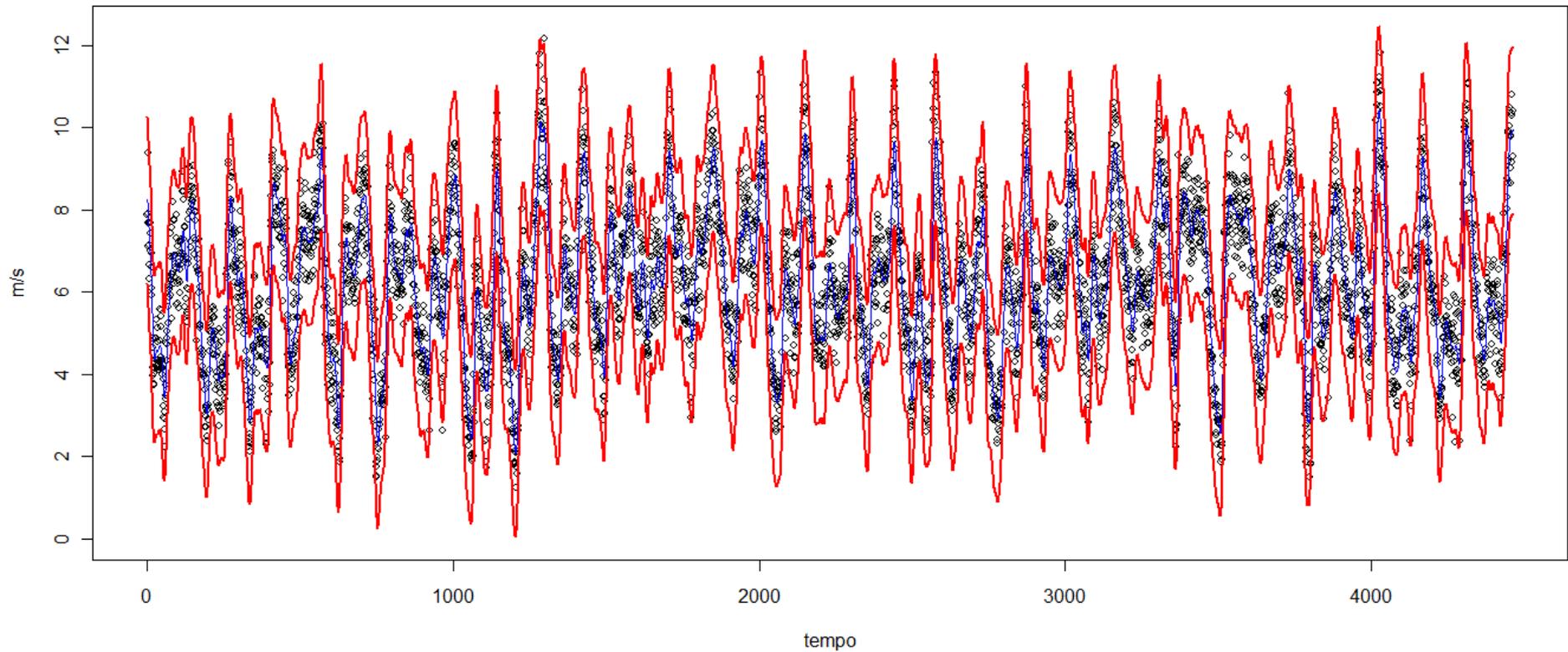
```
xl.workbook.save('dados_filtrados') # grava o arquivo Excel “dados_filtrados”  
xl.workbook.close('dados_filtrados') # fecha arquivo Excel
```

Medições de velocidade de vento e valores suavizados pelo procedimento LOESS

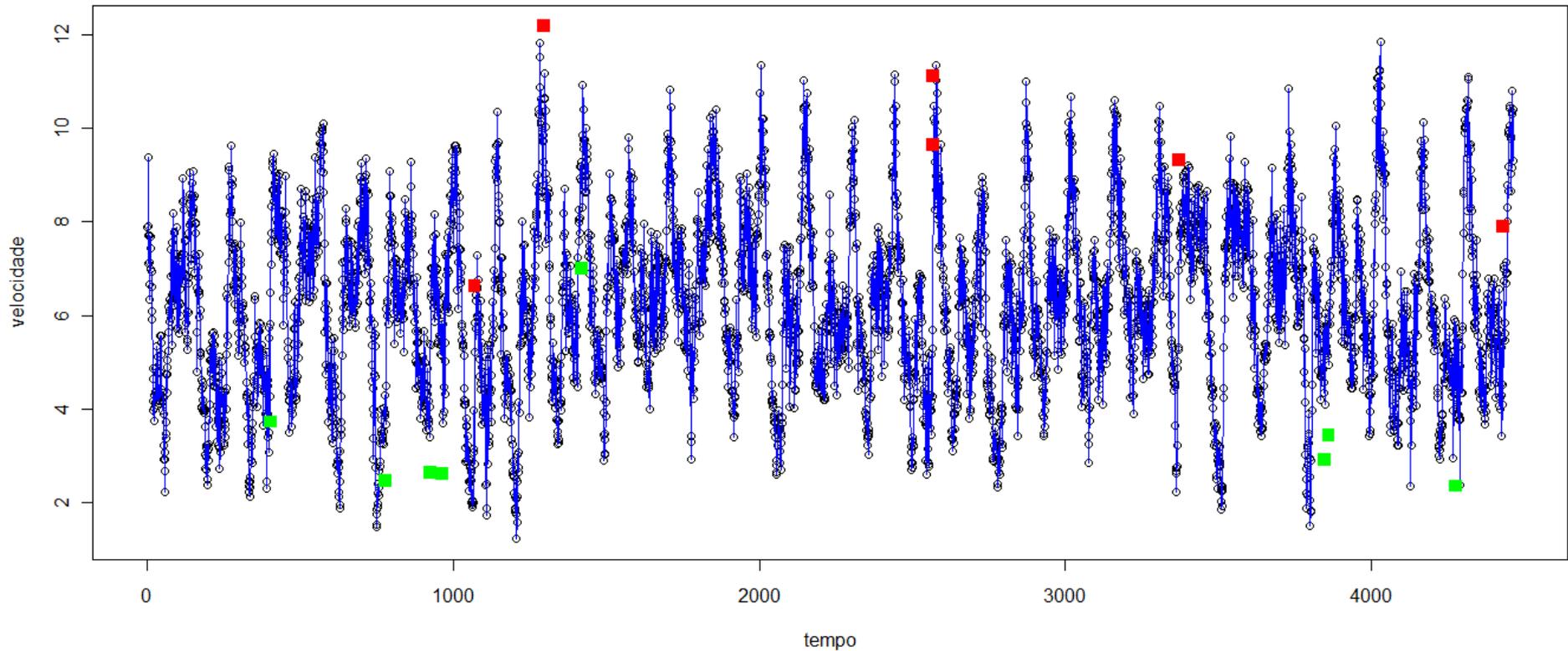


Loess significa Local Polynomial Regression Fitting
É uma função do pacote Stats

Medições de velocidade de vento, valores suavizados pelo procedimento LOESS e intervalos de confiança

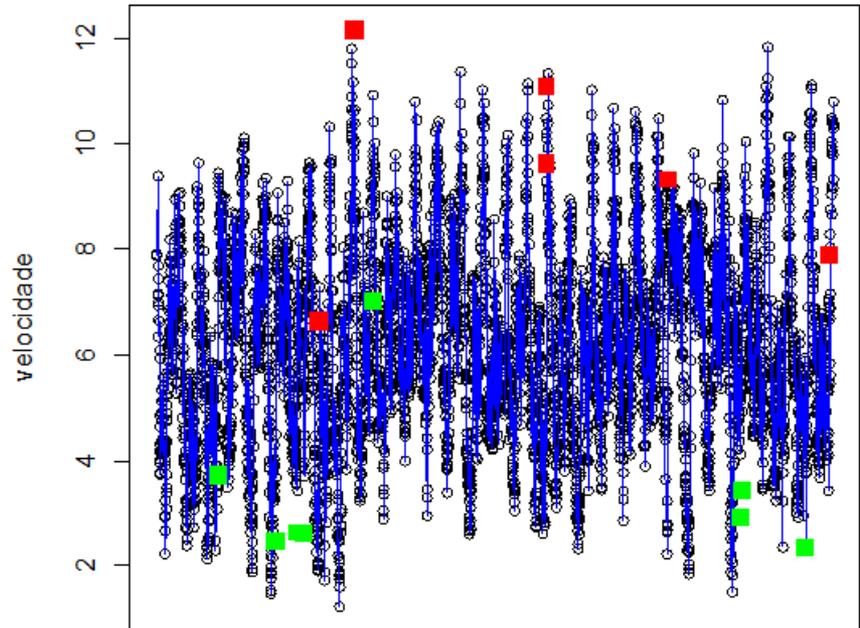


Medições de velocidade de vento, valores filtrados e outliers



Planilha com os resultados da filtragem realizada no R

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	valores medidos	valores filtrada	limite inferior	limite superior	alfa	14 outliers										
2	7.88	7.88	6.213743615	10.27641856	0.005	1070										
3	7.9	7.9	6.070533777	10.13320872		1297										
4	9.39	9.39	5.926093107	9.988768054		2568										
5	7.12	7.12	5.786108475	9.848783422		2569										
6	7.7	7.7	5.644644867	9.707319814		3372										
7	7.7	7.7	5.497682865	9.560357811		4432										
8	6.33	6.33	5.350772506	9.413447452		402										
9	6.67	6.67	5.200748483	9.26342343		779										
10	7.15	7.15	5.044445489	9.107120436		924										
11	7.69	7.69	4.876557357	8.939232304		964										
12	7.44	7.44	4.694353133	8.75702808		1420										
13	7	7	4.498530328	8.561205275		3846										
14	6.95	6.95	4.284758001	8.347432947		3861										
15	5.93	5.93	4.040045158	8.102720105		4275										
16	6.61	6.61	3.794058973	7.85673392												
17	6.1	6.1	3.570017919	7.632692866												
18	4.87	4.87	3.345956382	7.408631328												
19	4.65	4.65	3.099908747	7.162583694												
20	4.18	4.18	2.870398156	6.933073102												
21	3.97	3.97	2.697440464	6.76011541												
22	3.75	3.75	2.561309126	6.623984073												
23	3.76	3.76	2.441718015	6.504392962												
24	4.03	4.03	2.371672356	6.434347302												
25	4.3	4.3	2.349618956	6.412293902												
26	4.15	4.15	2.341063905	6.403738852												
27	4.38	4.38	2.356806612	6.419481558												
28	4.87	4.87	2.407646482	6.470321429												
29	4.36	4.36	2.46867143	6.531346376												
30	4.89	4.89	2.51018564	6.572860587												



Estudo de caso sobre Energia eólica

A distribuição de Weibull é o modelo probabilístico que melhor descreve a distribuição de frequência da velocidade do vento (CUSTÓDIO, 2009).

$$v \sim Weibull(k,c) \quad \Rightarrow \quad f(v) = \frac{k}{c} \left(\frac{v}{c}\right)^{k-1} e^{-\left(\frac{v}{c}\right)^k}$$

$v \geq 0$

v é a velocidade do vento em m/s ($v \geq 0$)

k é o fator de forma (adimensional e $k > 0$)

c é o fator de escala (em m/s e $c > 1$)

$$E(v) = c \Gamma\left(1 + \frac{1}{k}\right) \quad \sigma(v) = c \sqrt{\Gamma\left(1 + \frac{2}{k}\right) - \left[\Gamma\left(1 + \frac{1}{k}\right)\right]^2}$$

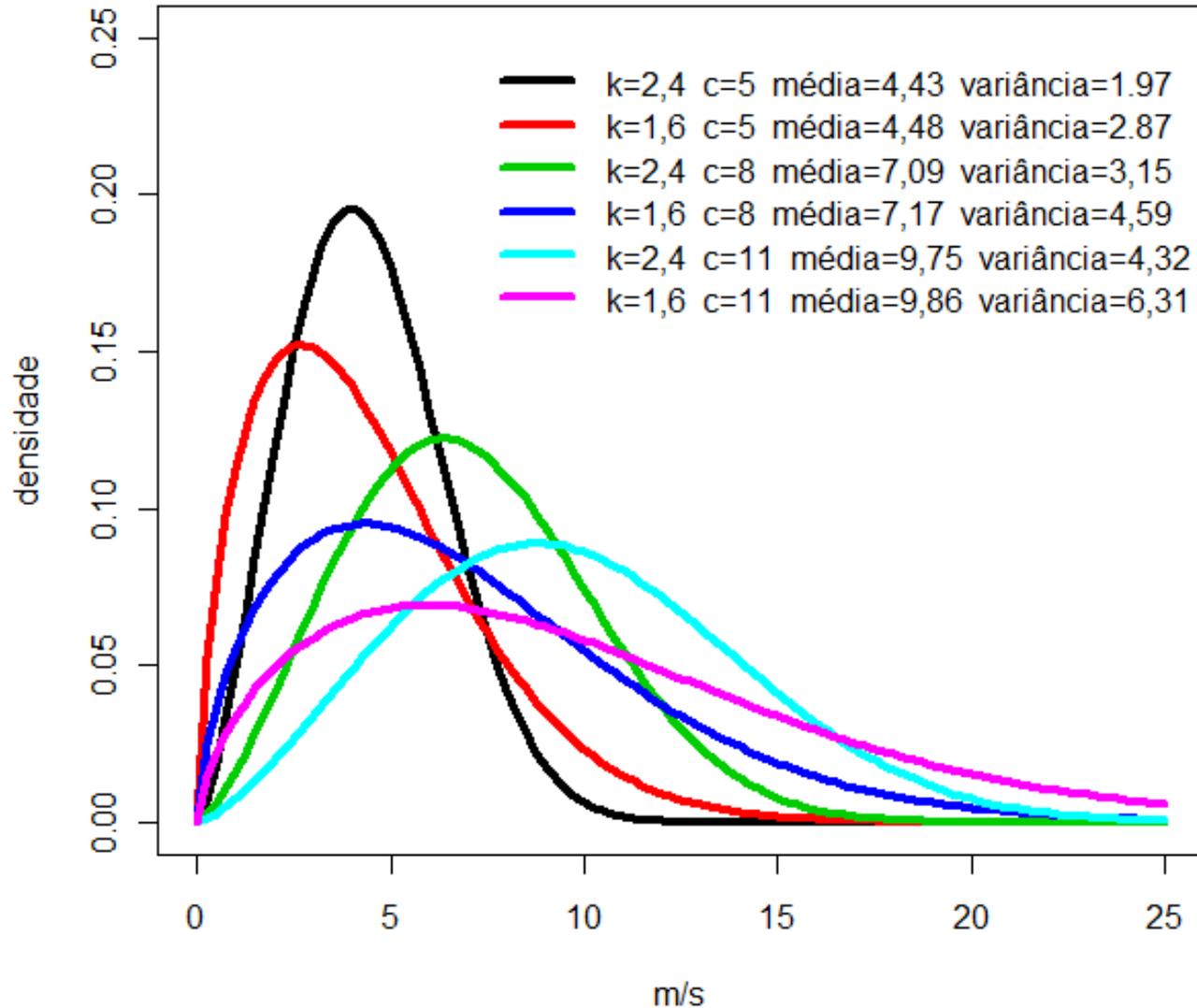
O fator de escala c tem relação direta com a velocidade média.

k é um indicador da constância do vento, quanto maior o valor de k menor a dispersão em torno da velocidade média. Tipicamente assume valores entre 2 e 3.

Em regiões de ventos alísios, como no Nordeste brasileiro, o parâmetro k pode atingir valores superiores a 6 (AMARANTE et al, 2001).

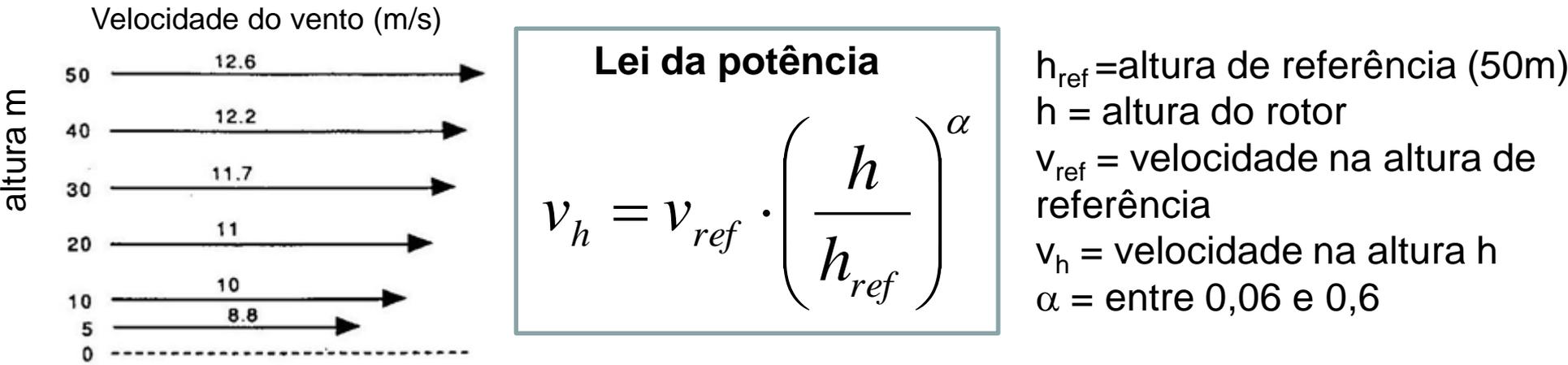
Estudo de caso sobre Energia eólica

Distribuição de Weibull para diferentes valores de c e k



Correção dos parâmetros c e k em função da altura do rotor da turbina

A velocidade do vento cresce com a altura



Correções dos fatores c e k em função da altura h do rotor

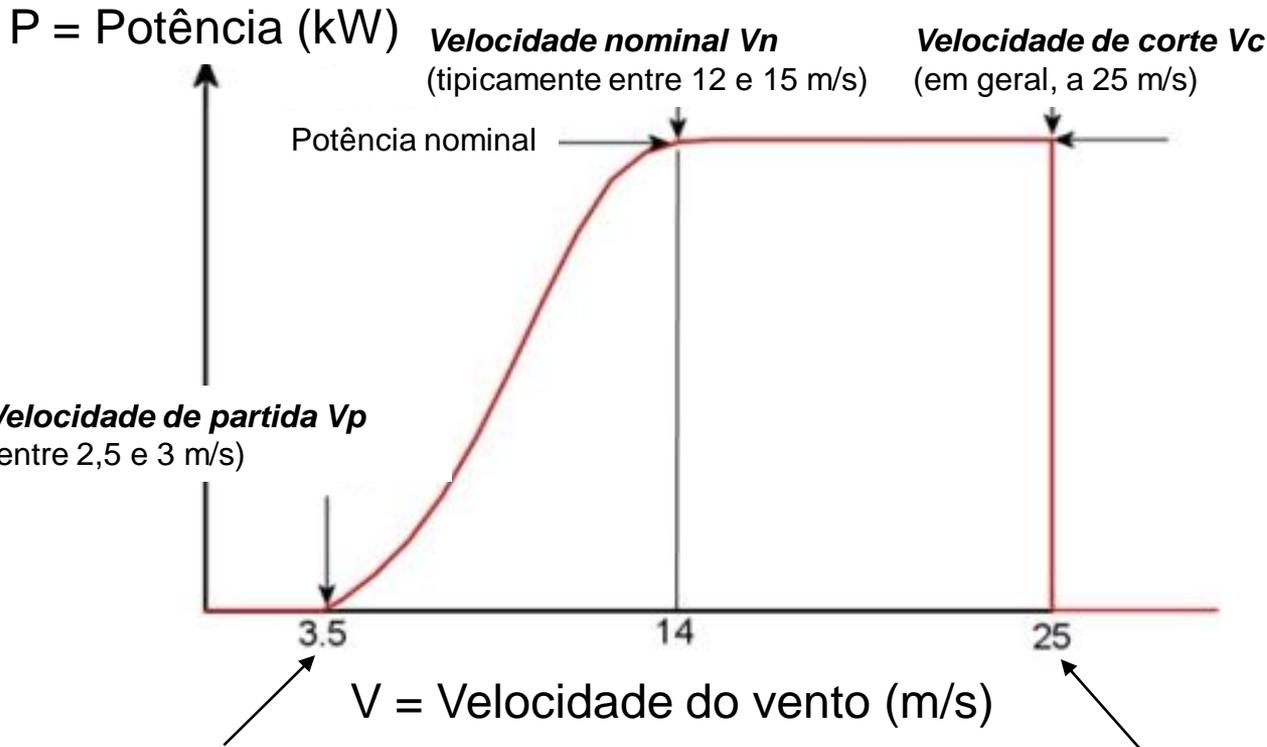
$$c_h = c_{ref} \cdot \left(\frac{h}{h_{ref}} \right)^n \quad \leftarrow n = \frac{0,37 - 0,088 \cdot \ln(c_{ref})}{1 - 0,088 \cdot \ln(h_{ref}/10)}$$

$$k_h = k_{ref} \cdot \frac{1 - 0,088 \cdot \ln(h_{ref}/10)}{1 - 0,088 \cdot \ln(h/10)}$$

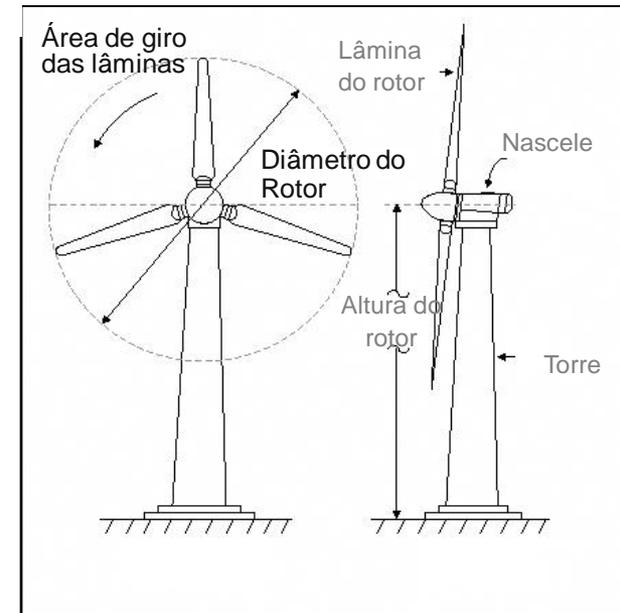
Aproximação que não considera a rugosidade do solo (CUSTÓDIO, 2009)

Estudo de caso sobre Energia eólica

A relação entre a velocidade do vento e a potência gerada é não linear



Turbina eólica



Conteúdo energético a partir 3,5 m/s

Potência gerada entre V_p e V_n

$$P = \frac{1}{2} C_p \cdot \eta_m \cdot \eta_g \cdot \rho \cdot A \cdot v^3$$

A turbina é desligada para protegê-la de esforços mecânicos

A = área de giro das lâminas m^2

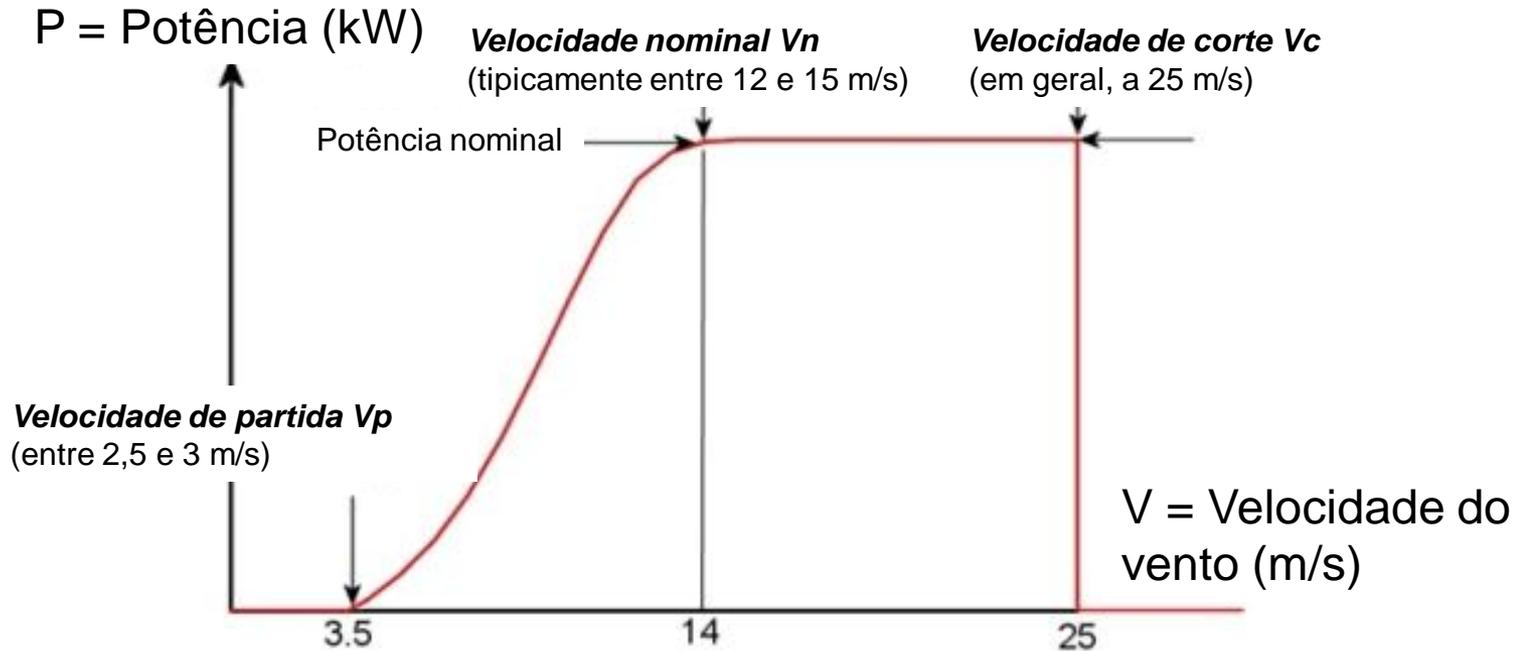
ρ = densidade do ar (kg/m^3)

C_p, η_m, η_g constantes que expressam os rendimentos nos estágios da conversão de energia



Estudo de caso sobre Energia eólica

Curva de potência



Curva de
potência
 $P(V)$

$$P = 0$$

velocidade $< v_p$ (velocidade de partida)

$$P = \frac{1}{2} C_p \cdot \eta_m \cdot \eta_g \cdot \rho \cdot A \cdot v^3$$

$v_p < \text{velocidade} < v_n$ (velocidade nominal)

$$P = \frac{1}{2} C_p \cdot \eta_m \cdot \eta_g \cdot \rho \cdot A \cdot v_n^3$$

$v_n < \text{velocidade} < v_c$ (velocidade de corte)

$$P = 0$$

velocidade $> v_c$

Estudo de caso sobre Energia eólica

Na implantação de um aproveitamento eólico é fundamental identificar a turbina cujas velocidades características v_p , v_n e v_c proporcionem a extração mais econômica da potência disponível no vento.

O fator de capacidade (FC) é um dos parâmetros considerados na escolha da turbina mais adequada.

O FC é definido pela razão entre a produção média da turbina e a sua produção potencial (máxima), se operada constantemente a plena capacidade.

$$FC = \frac{\text{potência média}}{\text{potência máxima}} = \frac{P_{\text{média}}}{P_{\text{máxima}}}$$

Estudo de caso sobre Energia eólica

Potência máxima $P_{m\acute{a}xima} = \frac{1}{2} C_p \cdot \eta_m \cdot \eta_g \cdot \rho \cdot A \cdot v_n^3$

←
Velocidade nominal

Potência média $P_{m\acute{e}dia} = \int_0^{\infty} P(v) \cdot f(v) \cdot dv$

←
Função densidade de probabilidade da velocidade do vento (Em geral, Weibull)

Curva de potência $P(v) = \begin{cases} 0 & v < v_p \\ \frac{1}{2} C_p \cdot \eta_m \cdot \eta_g \cdot \rho \cdot A \cdot v^3 & v_p \leq v < v_n \\ \frac{1}{2} C_p \cdot \eta_m \cdot \eta_g \cdot \rho \cdot A \cdot v_n^3 & v_n \leq v < v_c \\ 0 & v > v_c \end{cases}$

$$P_{m\acute{e}dia} = \int_{v_p}^{v_n} \frac{1}{2} C_p \eta_m \eta_g \rho A v^3 \cdot f(v) \cdot dv + \int_{v_n}^{v_c} \frac{1}{2} C_p \eta_m \eta_g \rho A v_n^3 \cdot f(v) \cdot dv$$

Estudo de caso sobre Energia eólica

JANGAMSHETTI & RAU (1999)

$$FC = \frac{P_{m\u00e9dia}}{P_{m\u00e1xima}} = \frac{\int_{v_p}^{v_n} \frac{1}{2} C_p \eta_m \eta_g \rho A v^3 \cdot f(v) \cdot dv + \int_{v_n}^{v_c} \frac{1}{2} C_p \eta_m \eta_g \rho A v_n^3 \cdot f(v) \cdot dv}{\frac{1}{2} C_p \cdot \eta_m \cdot \eta_g \cdot \rho \cdot A \cdot v_n^3}$$

$$FC = \frac{1}{v_n^3} \int_{v_p}^{v_n} v^3 f(v) dv + \int_{v_n}^{v_c} f(v) dv$$

O fator de capacidade depende das velocidades características da curva de pot\u00eancia da turbina e\u00f3lica (v_p , v_n e v_c) e da fun\u00e7\u00e3o densidade de probabilidade da velocidade do vento na altura do rotor da turbina.

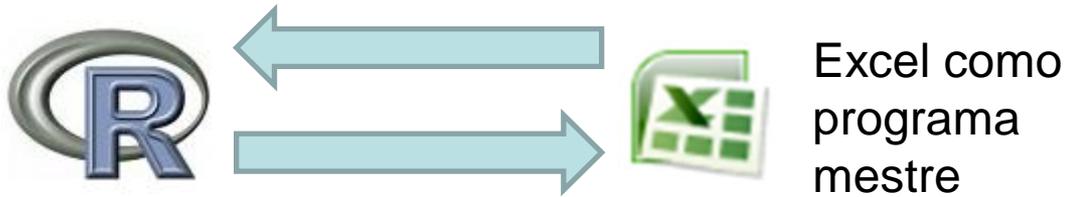
Estudo de caso sobre Energia eólica

A energia elétrica (Wh) gerada pela turbina eólica pode ser calculada pelo produto entre o fator de capacidade da turbina, a potência máxima da turbina e o período de tempo avaliado.

$$\text{energia} = FC \cdot \text{potência máxima} \cdot \text{período em horas}$$

Excel DNA + R.NET (Excel as Host)

Permite criar funções no Excel que são executadas no R



	A	B
1	=My	
2	<ul style="list-style-type: none">MyWeibullFcapMyWeibullFit	

Excel DNA (<http://exceldna.codeplex.com>) Integra a linguagem .NET no Excel
RdotNet (<http://rdotnet.codeplex.com>) Integra as linguages R e .NET

Excel DNA e RdoNet são open source.

Um bom lugar para visitar é o blog <http://mockquant.blogspot.com>

No blog podem ser obtidos os arquivos necessários e que devem ser armazenados em um mesmo diretório:

- ExcelDNA.Integration.dll
- R.NET.dll
- ExcelDNA.xll
- ExcelDNA.txt (arquivo com o programa)

Excel DNA + R.NET

```
<DnaLibrary RuntimeVersion="v4.0" Name="My First XLL" Language="CS">
<ExternalLibrary Path="R.NET.dll" />
<ExternalLibrary Path="ExcelDna.Integration.dll" />
<Reference Name="R.NET" />
<![CDATA[
using System;
using System.Collections.Generic;
using System.Linq;
using System.Text;
using ExcelDna.Integration;
using RDotNet;
namespace CSLib
{
    public class CSLib
    {
        static REngine engine = null;
        static CSLib()
        {
            // Local onde está instalado o arquivo R.dll
            REngine.SetDllDirectory(@"C:\Programas\R\R-2.13.2\bin\i386");
            engine = REngine.CreateInstance("RDotNet", new[] { "-q" });
        }
        [ExcelFunction(Description = "gera número aleatório normal padrão")]
        public static double [] MyRnorm(int number)
        {
            engine.EagerEvaluate("x<-rnorm("+ number +)").AsNumeric();
            NumericVector x = engine.GetSymbol("x").AsNumeric();
            return (x.ToArray());
        }
    }
}
]]>
</DnaLibrary>
```

Exemplo de um código para integrar o Excel ao R com o auxílio do Excel.DNA e R.NET

O código cria uma função Excel para gerar um número aleatório $N(0,1)$ usando a função `rnorm` do R

Cria função Excel
MyRnorm



Chama a função
`rnorm` do R



Excel DNA + R.NET



Função MyWeibullFit

```
[ExcelFunction(Description = "estima parametros Weibull")]  
public static double [] MyWeibullFit(int number)  
{  
    engine.EagerEvaluate("mes<-("+ number +")").AsNumeric();  
    NumericVector mes = engine.GetSymbol("mes").AsNumeric();  
    engine.EagerEvaluate(@"source('c:/ExcelDNA/ventos/fitweibull.txt')");  
    engine.EagerEvaluate("estimativas").AsNumeric();  
    NumericVector estimativas = engine.GetSymbol("estimativas").AsNumeric();  
    return (estimativas.ToArray());  
}
```



**Executa código R contido no arquivo fitweibull.txt responsável
peça estimação dos parâmetros da distribuição de Weibull da velocidade do
vento e apresentação de histograma.**

Estimação pelo método dos momentos

Foram utilizados os pacotes Stats, MASS e xlsReadWrite

Código R para estimação dos parâmetros da Weibull pelo método dos momentos (Blischke/Scheuer)

$$\bar{V} = c\Gamma\left(1 + \frac{1}{k}\right)$$
$$S^2 = c^2 \left[\Gamma\left(1 + \frac{2}{k}\right) - \left[\Gamma\left(1 + \frac{1}{k}\right) \right]^2 \right]$$

$\frac{S^2}{\bar{V}^2} = \frac{\Gamma\left(1 + \frac{2}{k}\right)}{\Gamma^2\left(1 + \frac{1}{k}\right)} - 1 \rightarrow \hat{k}$

$\hat{c} \leftarrow c = \frac{\bar{V}}{\Gamma\left(1 + \frac{1}{k}\right)}$

```
xbar <- mean(velocidade)
varx <- var(velocidade)
```

```
f <- function(k){
  return(gamma(1+2/k)/gamma(1+1/k)^2 - 1 - varx/xbar^2)
}
```

```
forma <- uniroot(f,c(0.02,50))$root
```

```
escala <- xbar/gamma(1+1/shape)
```

\hat{k}

\hat{c}

Excel DNA + R.NET

```
[ExcelFunction(Description = "calcula fator de capacidade")]  
public static double [] MyWeibullFcap(double value1,double value2,double value3,  
double value4,double value5)  
{  
    engine.EagerEvaluate("vp<-("+ value1 +)").AsNumeric();  
    NumericVector vp = engine.GetSymbol("vp").AsNumeric();  
    engine.EagerEvaluate("vn<-("+ value2 +)").AsNumeric();  
    NumericVector vn = engine.GetSymbol("vn").AsNumeric();  
    engine.EagerEvaluate("vc<-("+ value3 +)").AsNumeric();  
    NumericVector vc = engine.GetSymbol("vc").AsNumeric();  
    engine.EagerEvaluate("k<-("+ value4 +)").AsNumeric();  
    NumericVector k = engine.GetSymbol("k").AsNumeric();  
    engine.EagerEvaluate("s<-("+ value5 +)").AsNumeric();  
    NumericVector s = engine.GetSymbol("s").AsNumeric();  
    engine.EagerEvaluate(@"source('c:/ExcelDNA/ventos/calcfc.txt')");  
    engine.EagerEvaluate("fc").AsNumeric();  
    NumericVector fc = engine.GetSymbol("fc").AsNumeric();  
    return (fc.ToArray<double>());  
}
```



Função
MyWeibullFcap

**Insera os valores das
variáveis de entrada**



**Executa código R contido no arquivo calcfc.txt responsável pela
estimação do fator de capacidade da turbina eólica.
Foi utilizadoo pacote MASS.**

Código R para o cálculo do fator de capacidade (FC)

$$FC = \frac{1}{v_n^3} \int_{v_p}^{v_n} v^3 f(v) dv + \int_{v_n}^{v_c} f(v) dv$$

```
cte<-vn^3
```

```
aux1<-function(w) (w*w*w*dweibull(w,k,s))
```

```
aux2<-function(w) (dweibull(w,k,s))
```

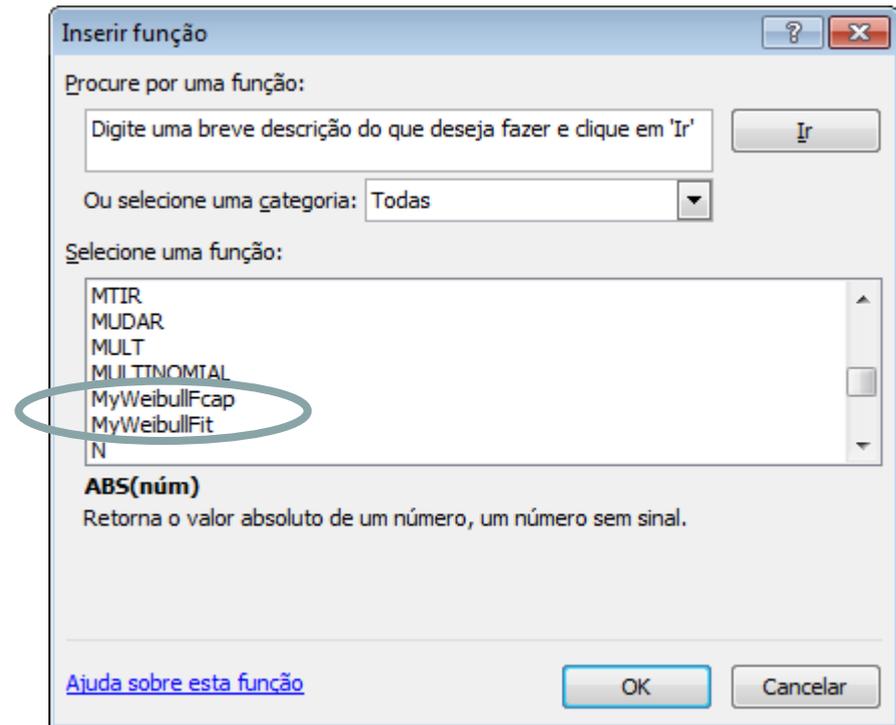
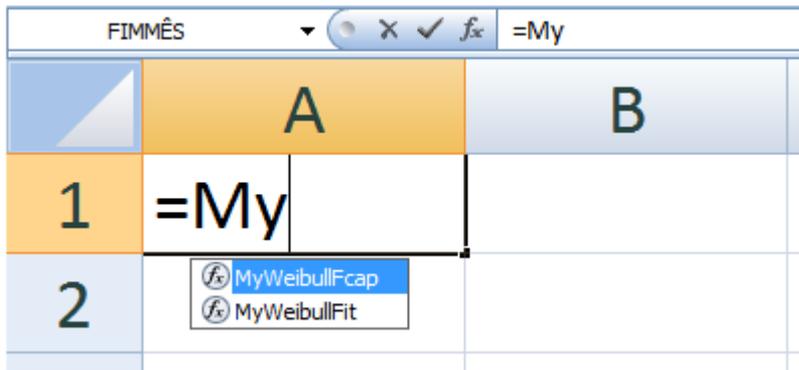
```
fc1<-integrate(aux1,vp,vn)
```

```
fc2<-integrate(aux2,vn,vc)
```

```
fc<-(1/cte)*fc1$value+fc2$value
```

Excel DNA + R.NET

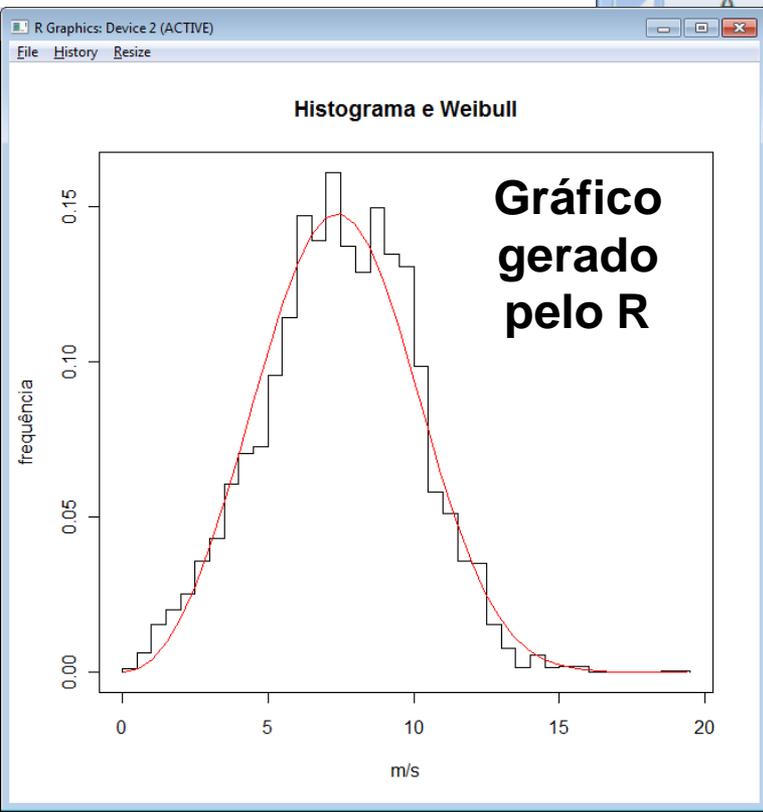
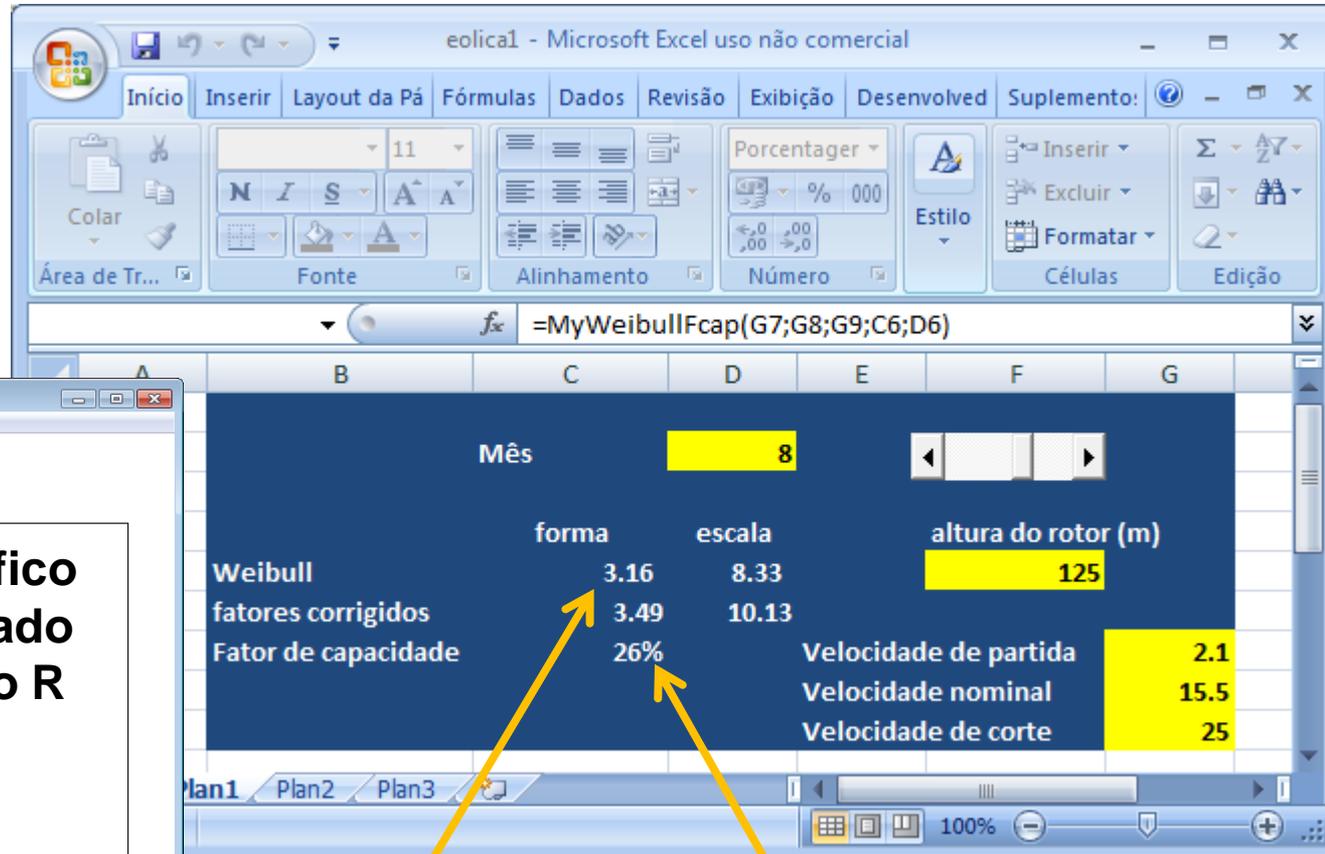
Para carregar as funções MyWeibullFit e MyWeibullFca basta abrir o arquivo ExcelDNA.xll no Excel.



Excel DNA + R.NET

Os registros anemométricos estão armazenados em outra planilha Excel denominada arquivo.xls, onde cada coluna guarda os registros de um mês

Planilha Excel de onde são chamadas as funções **MyWeibullFit** e **MyWeibullFcap**



MyWeibullFit(mês)

MyWeibullFcap(vp,vn,vc,k,c)

Pacote RExcel



Thomas Baier



Erich Neuwirth

Baier Thomas, & Neuwirth Erich (2007). Excel :: COM :: R. Computational Statistics, Volume 22, Number 1/April 2007. Physica Verlag. <http://rcom.univie.ac.at/>

Oferece diferentes formas de interface entre o R e o Excel.

Transferência de dados entre o R e o Excel

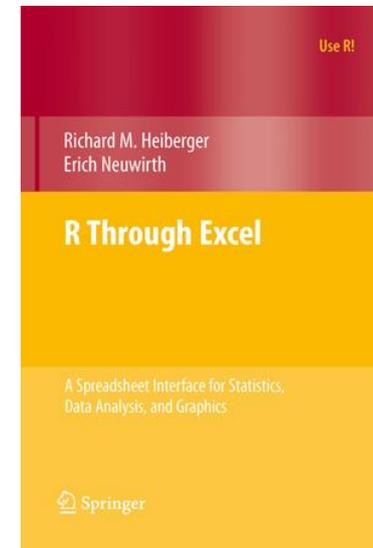
Excel como editor de dados para serem analisados no R

Excel como interface gráfica para o R

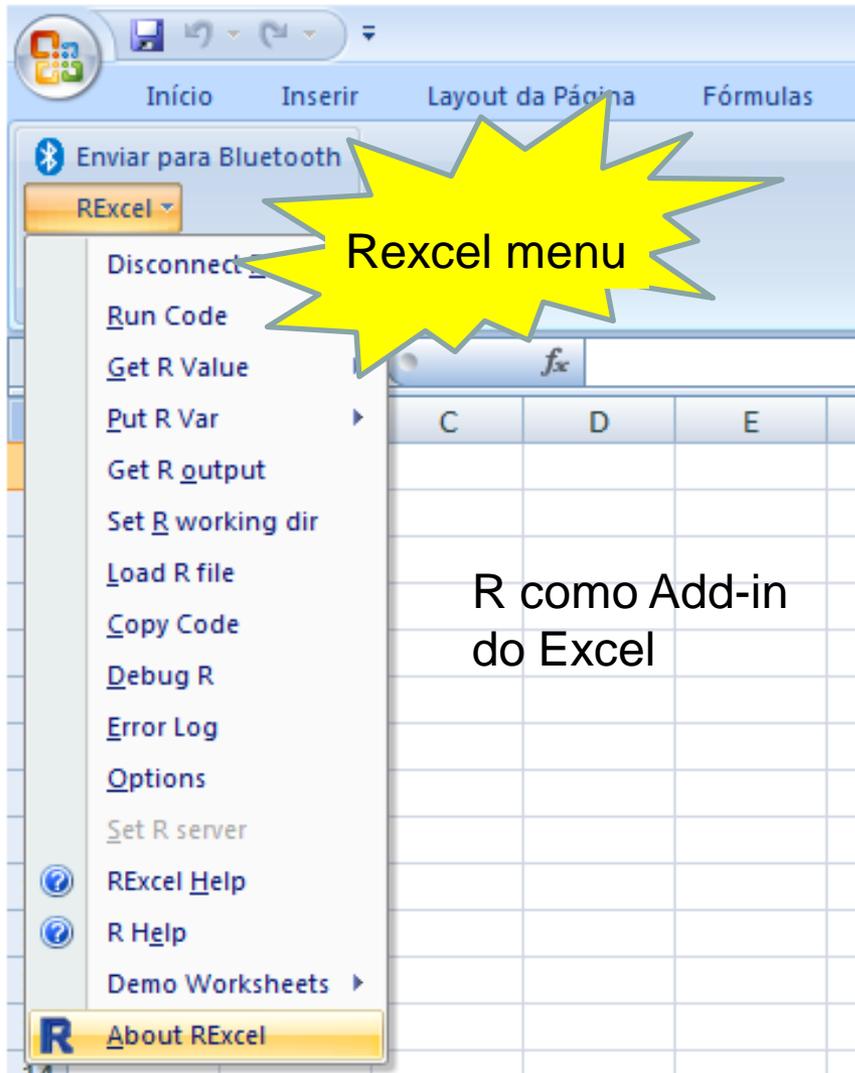
Excel é um local para armazenamento de dados e códigos

Executar comandos R a partir de macros em Excel VBA

Permite usar comandos R em funções do Excel.



Pacote RExcel



Três diferentes formas de usar o RExcel

1) Scratchpad mode: Executa código R escrito na planilha e transfere variáveis entre o R e o Excel (Usa o RExcel menu).

2) Macro mode: Chama funções do R a partir de macros no Excel (usa macros disponibilizados pelo RExcel.xla).

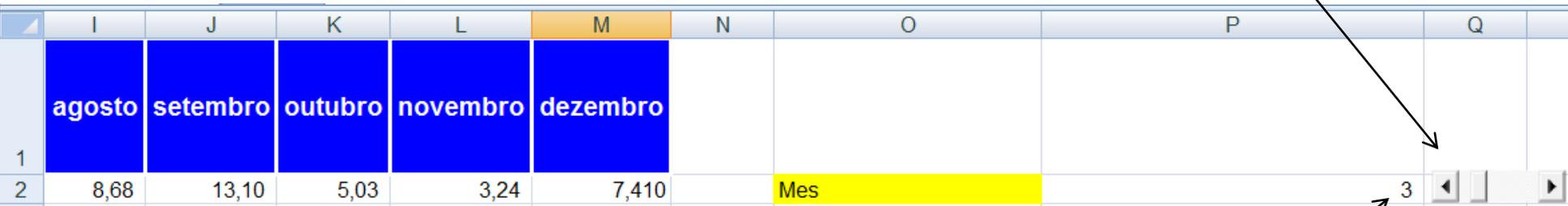
3) Worksheet functions: Chama funções do R diretamente das células de uma planilha.

Planilha do Estudo de caso

A matriz de dados tem 12 colunas, uma para cada mês

O primeiro passo consiste em escolher o mês que será analisado

Selecione o mês desejado com a Scrollbar



	I	J	K	L	M	N	O	P	Q
1	agosto	setembro	outubro	novembro	dezembro				
2	8,68	13,10	5,03	3,24	7,410	Mes			

A função **Rput("Mes";P2)** cria a variável Mes no R e que assume valor igual ao inserido na célula P2

O mês escolhido aparece na célula P3

Macro VBA para calcular estatísticas descritivas com funções do R

Rventos.xlsm - Plan1 (Código)

CommandButton1

Click

```
Private Sub CommandButton1_Click()  
Call RInterface.StartRServer  
Call RInterface.PutArray("velocidade", Range("Plan1:B2:M4465"))  
Call RInterface.RRun("library(moments)")  
Call RInterface.RRun("library(MASS)")  
Call RInterface.RRun("rm(x)")  
Call RInterface.RRun("x<-velocidade[,Mes]")  
Call RInterface.RRun("dados<-x[which(x[is.na(x)==FALSE]>=0)]")  
Call RInterface.RRun("media<-mean(dados)")  
Call RInterface.RRun("dp<-sd(dados)")  
Call RInterface.RRun("kurt<-kurtosis(dados)")  
Call RInterface.RRun("skew<-skewness(dados)")  
Call RInterface.RRun("mediana<-median(dados)")  
Call RInterface.RRun("maximo<-max(dados)")  
Call RInterface.RRun("minimo<-min(dados)")  
Call RInterface.RRun("nobs<-length(dados)")  
Call RInterface.GetArray("media", Range("P5:P5"))  
Call RInterface.GetArray("dp", Range("P6:P6"))  
Call RInterface.GetArray("skew", Range("P7:P7"))  
Call RInterface.GetArray("kurt", Range("P8:P8"))  
Call RInterface.GetArray("mediana", Range("P9:P9"))  
Call RInterface.GetArray("maximo", Range("P10:P10"))  
Call RInterface.GetArray("minimo", Range("P11:P11"))  
Call RInterface.GetArray("nobs", Range("P12:P12"))  
Call RInterface.StopRServer
```

StartRServer Inicia chamada do R

PutArray Transfere o conteúdo em B2:X4465 para a matriz velocidade R

RRun Executa um comando no R

GetArray Transfere variáveis do R para determinadas células na planilha Excel

StopRServer Encerra chamada do R

Planilha do Estudo de caso

Estima fatores de forma e escala (parâmetros da Weibull).

Modelos de turbinas

Modelo	vp m/s	vn m/s	vp m/s	altura m	potência kW
A	4	16	25	74	850
B	3.5	13	20	80	1650
C	4	12	20	95	1800
D	4	16	25	100	2000
E	3.5	15	25	105	3000
F	3	12	25	119	3000
G	4	12	25	125	1800

Modelo de turbina

G

Histograma, QQ Plot e KS

Seleciona do modelo da turbina

Faz histograma, QQ plot e Teste de Kolmogorov Smirnov

```
Rventos.xlsm - Módulo1 (Código)
(Geral) Botão5_Clique

Sub Botão5_Clique ()
Call RInterface.StartRServer
Call RInterface.RRun ("library (MASS) ")
Call RInterface.PutArray ("velocidade", Range ("Plan1!B2:M4465"))
Call RInterface.RRun ("x<-velocidade[,Mes]")
Call RInterface.RunRFile ("c:\turbinando\exemplo4\horario.txt")
Call RInterface.RRun ("y<-x[which (x[is.na (x)]==FALSE)>0] ")
Call RInterface.RunRCodeFromRange (Range ("Plan1!AA3:AA20"))
Call RInterface.GetArray ("s", Range ("P14:P14"))
Call RInterface.GetArray ("k", Range ("P15:P15"))
Call RInterface.StopRServer
End Sub
```

Macro VBA

RunFile Executa código em R armazenado no arquivo txt

RunCode executa código em R armazenado nas células W3:W20 da planilha

Macro VBA chama funções do R para estimar os parâmetros da distribuição de probabilidade da velocidade do vento

```
Sub Botão5_Clique()  
Call RInterface.StartRServer  
Call RInterface.RRun("library(MASS)")  
Call RInterface.PutArray("velocidade", Range("Plan1!B2:M4465"))  
Call RInterface.RRun("x<-velocidade[,Mes]")  
Call RInterface.RunRFile("c:\turbinando\exemplo4\horario.txt")  
Call RInterface.RRun("y<-x[which(x[is.na(x)]==FALSE)>0]")  
Call RInterface.RunRCodeFromRange(Range("Plan1!AA3:AA20"))  
Call RInterface.GetArray("s", Range("P14:P14"))  
Call RInterface.GetArray("k", Range("P15:P15"))  
Call RInterface.StopRServer  
End Sub
```

PutArray Transfere o conteúdo em B2:X4465 para a matriz velocidade R

RunCode executa código em R armazenado nas células W3:W20

GetArray Transfere resultados do R para o Excel de acordo com o modelo selecionado

RunCode executa código em R armazenado nas células W3:W20 da planilha

Código em R escrito na planilha Excel

Código R para estimação dos fatores de escala e de forma

```
par(mfrow=c(3,1))
```

```
h<-hist(y,breaks=30,ylab="frequência",xlab="m/s",main="Histograma")
```

```
xhist<-c(min(h$breaks),h$breaks)
```

```
yhist<-c(0,h$density,0)
```

```
xfit<-seq(min(y),max(y),length=40)
```

```
parweibull<-fitdistr(y,"weibull")
```

```
k<-parweibull[1]$estimate[1]
```

```
s<-parweibull[1]$estimate[2]
```

```
kstest<-ks.test(y,"pweibull",shape=k,scale=s)
```

```
yfit<-dweibull(xfit,shape=k,scale=s)
```

```
plot(xhist,yhist,type="s",ylim=c(0,max(yhist,yfit)),ylab="frequência",xlab="m/s",main='Histograma e Weibull')
```

```
lines(xfit,yfit,col='red')
```

```
# QQ plot
```

```
y.teo<-rweibull(nobs,shape=k,scale=s)
```

```
auxvar<-qqplot(y.teo,y,ylab="Distribuição empírica",xlab="Distribuição teórica",main="QQ-plot distr. Weibull")
```

```
abline(0,1)
```

```
modelo<-lm(auxvar$y~auxvar$x)
```

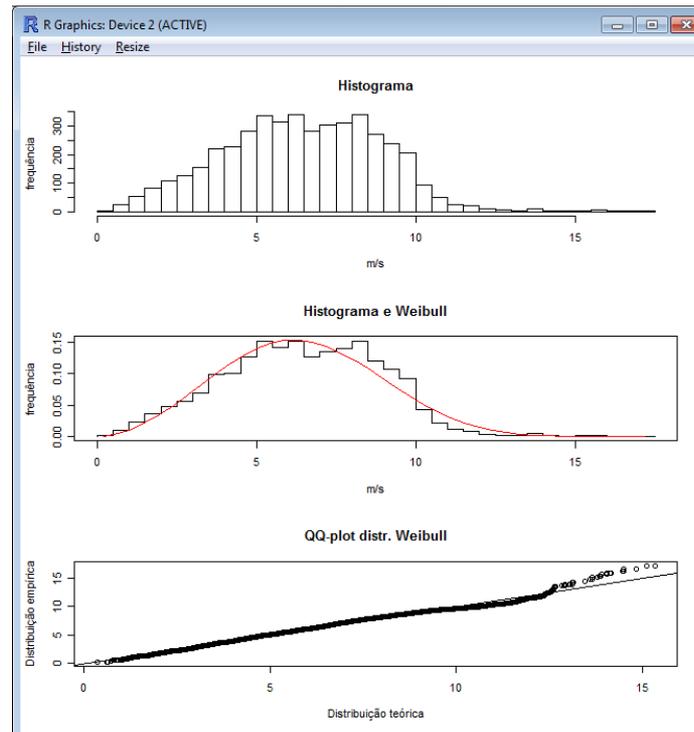
```
r2 <- modelo$coef[2]*(sqrt(var(auxvar$x))/sqrt(var(auxvar$y)))*cor(auxvar$y,auxvar$x)
```

Estima os parâmetros da função densidade de probabilidade selecionada pelo usuário, calcula a estatística de Kolmogorov-Smirnov e desenha histograma e QQ-Plot

Resultados gerados pelo R e exibidos na planilha Excel

fator de escala	8.30		
fator de forma	3.34		
Kstest	0.03	valor P	44.65%
r2 QQplot	0.9967		

Turbina	Fator de escala corrigido	10.10
G	Fator de forma corrigido	3.68
Altura	125	



Planilha do Estudo de caso

Cálculo do fator de capacidade

Parâmetros da turbina			
velocidade de partida (m/s)	4.0	vp	=RPut("vp";P20)
Velocidade nominal (m/s)	12.0	vn	=RPut("vn";P21)
Velocidade de corte (m/s)	25.0	vc	=RPut("vc";P22)
Potência nominal (kW)	1,800		
Fator de capacidade (%)	31%		
Energia MWh	217,632		

Funções que transferem valores do Excel para o R

Calcula Fator de Capacidade

Macro VBA para calcular fator de capacidade

```
Sub Botão6_Clique()  
Call RInterface.StartRServer  
Call RInterface.RunRFile("c:\RTGK\ventos\fc.txt")  
Call RInterface.GetArray("fc", Range("P24:P24"))  
Call RInterface.StopRServer  
End Sub
```

Código em R em arquivo txt

```
cte<-vn^3  
aux1<-function(w) (w*w*w*dweibull(w,k,s))  
aux2<-function(w) (dweibull(w,k,s))  
fc2<-integrate(aux2,vn,vc)  
fc1<-integrate(aux1,vp,vn)  
fc<<-(1/cte)*fc1$value+fc2$value
```

RunFile Executa código em R responsável pelo cálculo do fator de capacidade. Código em arquivo txt

Código em R para o cálculo do fator de capacidade (FC)

$$FC = \frac{1}{v_n^3} \int_{v_p}^{v_n} v^3 f(v) dv + \int_{v_n}^{v_c} f(v) dv$$

```
cte<-vn^3
```

```
aux1<-function(w) (w*w*w*dweibull(w,s,k))
```

```
aux2<-function(w) (dweibull(w,s,k))
```

```
fc2<-integrate(aux2, vn,vc)
```

```
fc1<-integrate(aux1, vp,vn)
```

```
fc<<- (1/cte)*fc1$value+fc2$value
```

Variáveis definidas no Excel

Conclusões

A integração entre o R e o Excel é fácil.

Oferece variadas formas novas e interessantes de usar ambos os programas.

O R ganha uma poderosa interface com o usuário.

O Excel é enriquecido com o poder do R.

Possibilita o desenvolvimento de aplicativos (com baixo custo) que podem ser disponibilizados aos clientes.

Referências bibliográficas

Amarante, O.A.C., Brower, M., Zack, J. e Sá, A.L., Atlas do Potencial Eólico Brasileiro, Centro de Pesquisas de Energia Elétrica, Brasília, 2001.

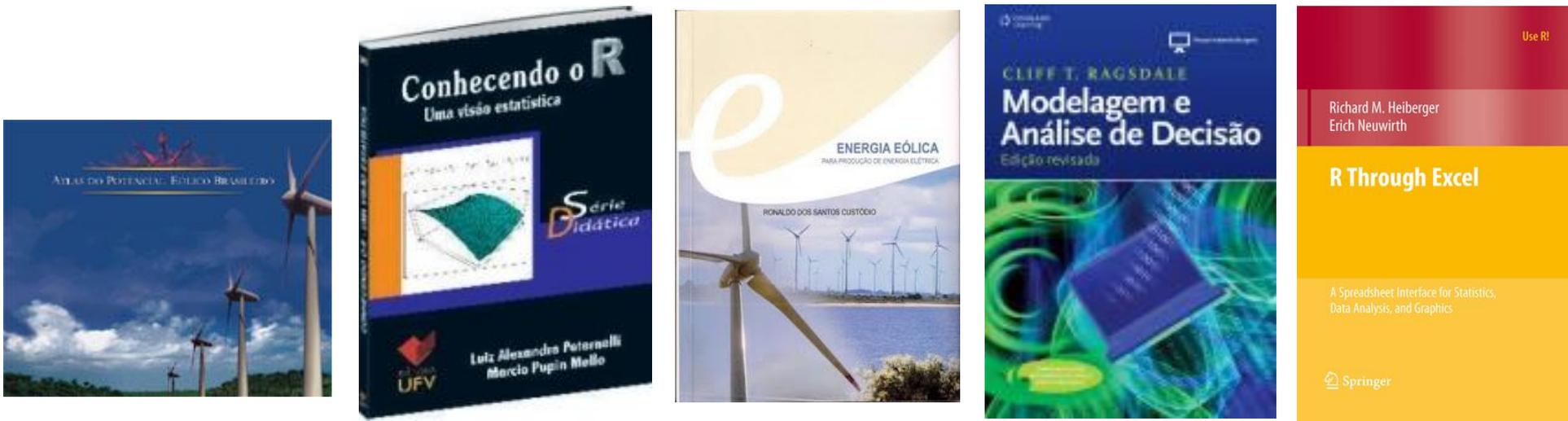
Baier Thomas.; Neuwirth Erich (2007). Excel :: COM :: R. Computational Statistics, Volume 22, Number 1/April 2007. Physica Verlag. <http://rcom.univie.ac.at/>

Custódio, R.S. Energia eólica para produção de energia elétrica, Eletrobras, Rio de Janeiro, 2009.

Jangamshetti, S.H.J.; Rau, V.G. (1999), Site matching of wind turbine generators: a case study, IEEE Transactions on Energy Conversion, v. 14, n. 4, December, 1537-1543.

Peternelli, L.A.; Mello, M.P. Conhecendo o R: uma visão estatística, Editora UFV, Viçosa, 2011.

Ragsdale, C.T. Modelagem e análise de dados, Cengage Learning, São Paulo, 2009.



Parabéns aos Estatísticos

José Francisco Moreira Pessanha (UERJ)
professorjfmp@hotmail.com

