

# ***ANÁLISE DE REGRESSÃO***

# Análise de Regressão

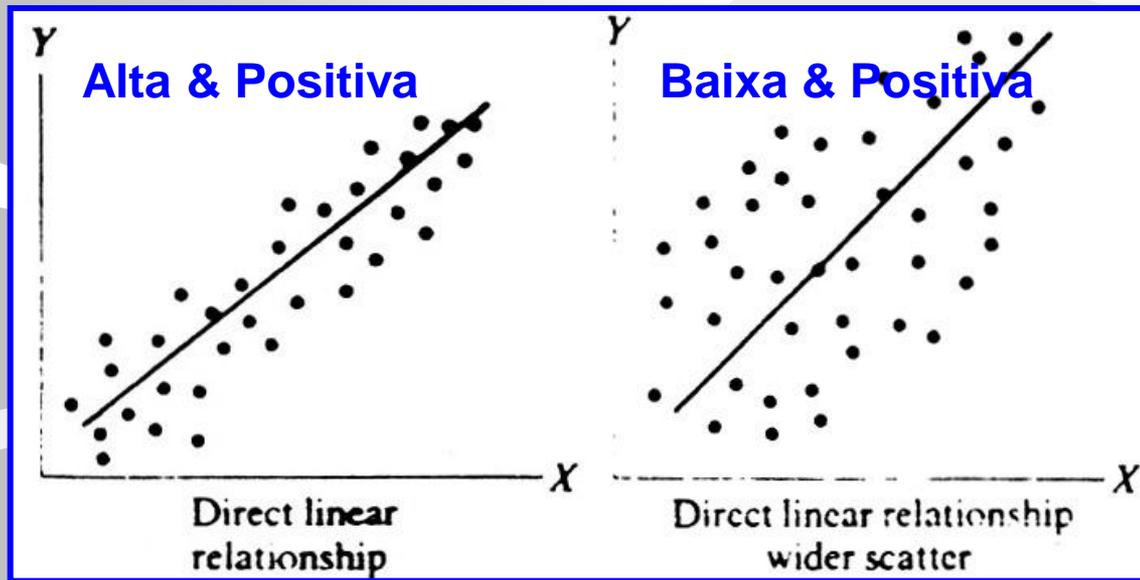
- Procedimento formal de desenvolver modelos ou equações a partir de dados históricos de navios semelhantes.

## Projeto do navio

Ferramenta de alto valor p/ gerar informação de navios existentes

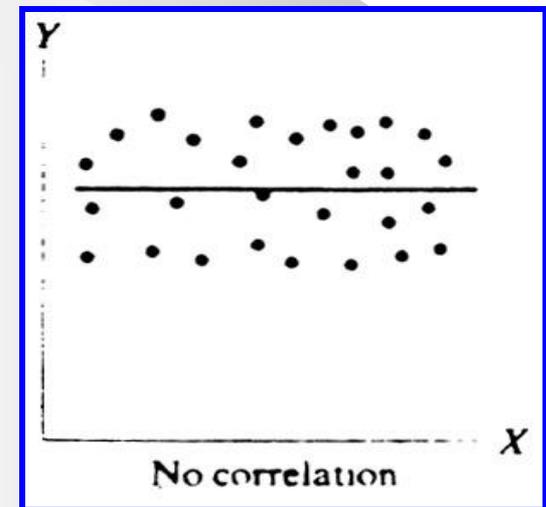
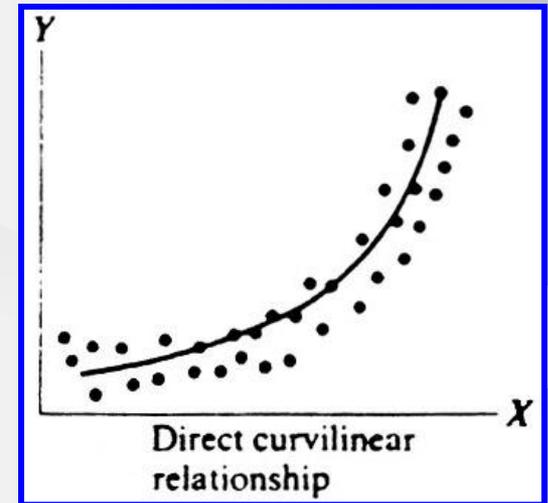
- Técnica de adaptar (“fitting”) curva qdo não existe relação óbvia que explique variação da variável(s) dependente(s) .
- Consiste, simplificada, em:
  - Plotar dados (diagrama de dispersão) e observar distribuição;
  - Assumir uma forma de equação;
  - Transformá-la, caso necessário, em expressão linear;
  - Efetuar procedimento de mínimos quadrados;
  - Testar a aderência;
  - Testar, caso necessária, outra equação que melhor se adapte.

# Análise de Regressão



Correlação linear

Correlação não-linear



# Curvas Típicas

■ Formas típicas:

- Linear

$$y = a_0 + a_1 x$$

- Hiperbólica

$$y = 1 / (a_0 + a_1 x)$$

- Exponencial

$$y = a b^x$$



Linear semi-log

$$\log y = \log a + x \log b$$

- Linear múltipla

$$y = a_0 + a_1 x_1 + a_2 x_2$$

- Polinomial

$$y = a_0 + a_1 x + a_2 x^2 + \dots$$

- Geométrica

$$y = a x^b$$



Linear log-log

$$\log y = \log a + b \log x$$

# Parâmetros de Aderência

- **Estimativa erro padrão**

Hipótese: Resíduos distribuídos - Curva normal

**Erro padrão = Desvio padrão dos resíduos**

Medida de dispersão dos dados em relação à curva de regressão

- **Coeficiente de determinação ( $R^2$  ou  $r^2$ ):** Percentual ou fração de variação de **Y** em torno da média explicada pela curva de regressão (adimensional)

$$0.0 \leq R^2 \leq 1.0$$

- **Coeficiente de correlação:** Mede quão próximo conjunto dados se distribuem em torno da curva de regressão

$$-1.0 \leq R \leq 1.0$$

# Observações Gerais

- Técnica de adaptação de curva: checar dados (equações de ordem superior), evitar extrapolação ou uso fora intervalo dos dados;
- Precisão coerente com dados;
- Curvas consistentes com teoria e julgamento prático de engenharia;
- Impossível remover uma variável independente sem perder precisão;
- Manter equação razoavelmente reduzida ( $x_i$ ,  $i \leq 10$ );
- Impossível melhorar pela adição de 1 ou 2 variáveis independentes;
- Objetivo:  $R \geq 0,95$ ,  $R^2 \geq 0,90$
- Evitar problemas com variáveis independentes altamente dependentes;
- Qualquer variável com valor grande do coeficiente deve ser observada cuidadosamente, especialmente dois termos com coeficientes de sinal oposto.

# Exemplo Análise Navios Ro-Ro

- 38 navios – entre décadas 70 e 80
- Estimativas iniciais  $L_{pp}$ , B e T  
Novo projeto  
 $10.000 \text{ ton} \leq \text{DWT} \leq 20.000 \text{ ton}$
- Observação:  
Espaços em branco ignorar

	NOME	DWT [tons]	LPP [m]	B [m]	T [m]	V [nós]	POT. [kW]
1	Mahano	6439.0	121.90	19.20	5.94	19.0	9709.0
2	Seaway Prince	4216.0	121.90	19.70	5.58	18.2	9023.0
3	Lys. Endeavor	8098.0	131.90	22.60	7.32	18.0	11931.0
4	Iron Monarch	14941.0	167.90	25.00	8.53	20.5	14166.0
5	Aust. Enterprise	14308.0	167.90	25.00	8.99	25.0	19440.0
6	Union Rotorua	14783.0	188.20	26.10	7.92	20.0	20474.0
7	Aust. Emblem	23857.0	205.00	30.00	10.52	26.7	34302.0
8	Saudi Abha	38500.0	232.00	32.30	10.80	19.5	20300.0
9	Mercand. Pres.	3500.0	96.00	18.80	4.95	15.3	3310.0
10	Mercand. Duke	7200.0	122.90	19.40	6.20	15.7	3600.0
11	Roll-Vigo	5500.0	112.80	18.40	6.20	17.6	10000.0
12	Baltic Eagle	9450.0	126.50	23.00	8.20	18.0	9694.0
13	Dana America	8002.0	124.00	24.00		15.3	5995.0
14	Arcturus	13000.0	146.00	25.00	8.00	18.5	13200.0
15	Esja	1100.0	64.00	13.00	3.55	13.8	2400.0
16	Monthlery	2200.0	102.00	18.00	5.25	16.0	
17	Mario	3250.0	97.50	19.70	4.70		
18	Skaugran	42400.0	173.20	32.30	12.00	14.8	11186.0
19	Ville Dunkg.	10000.0	160.60	25.00	7.98		
20	Linne	8300.0	152.50	18.00		17.2	
21	Dana Optima	3450.0	96.10	18.80	4.95	15.3	
22	Tajura	2900.0	98.50	17.50	5.17	16.0	
23	Barber Tampa	44000.0	246.40	32.30	11.70	20.5	26920.0
24	Jacqueline	7250.0	131.00	22.10	6.80	16.0	
25	Karisma	6750.0	136.00	22.00	6.90		
26	Con Ros NKK	17930.0	170.00	32.30	9.15	17.0	
27	Atl. Comp.	37000.0	233.60	32.30	10.80	17.5	17475.0
28	Ro-Ro Sea Link	6700.0	111.20	21.00	6.20	15.0	
29	GD Quincy	22500.0	199.00	32.30	9.00	18.0	
30	Nopal Barbro	17863.0	180.00	32.30	9.72	19.7	13524.0
31	Transfinland	11645.0	146.00	24.60	6.24		14500.0
32	Andra Mazario	19800.0	174.00	32.30	9.05	19.3	16550.0
33	Boogabila	31500.0	210.00	32.30	10.08	22.0	20238.0
34	Helenic Explorer	31123.0	175.00	28.50	11.90	20.6	18880.0
35	Atl. Prosper	8670.0	151.00	19.90	7.20		
36	Atl. Prelude	9511.0	162.30	19.40	7.20		
37	Boxer	9000.0	133.50	25.00	6.50	17.0	
38	Nordana	8002.0	135.00	24.00	6.80		

# Exemplo: Navios Ro-Ro

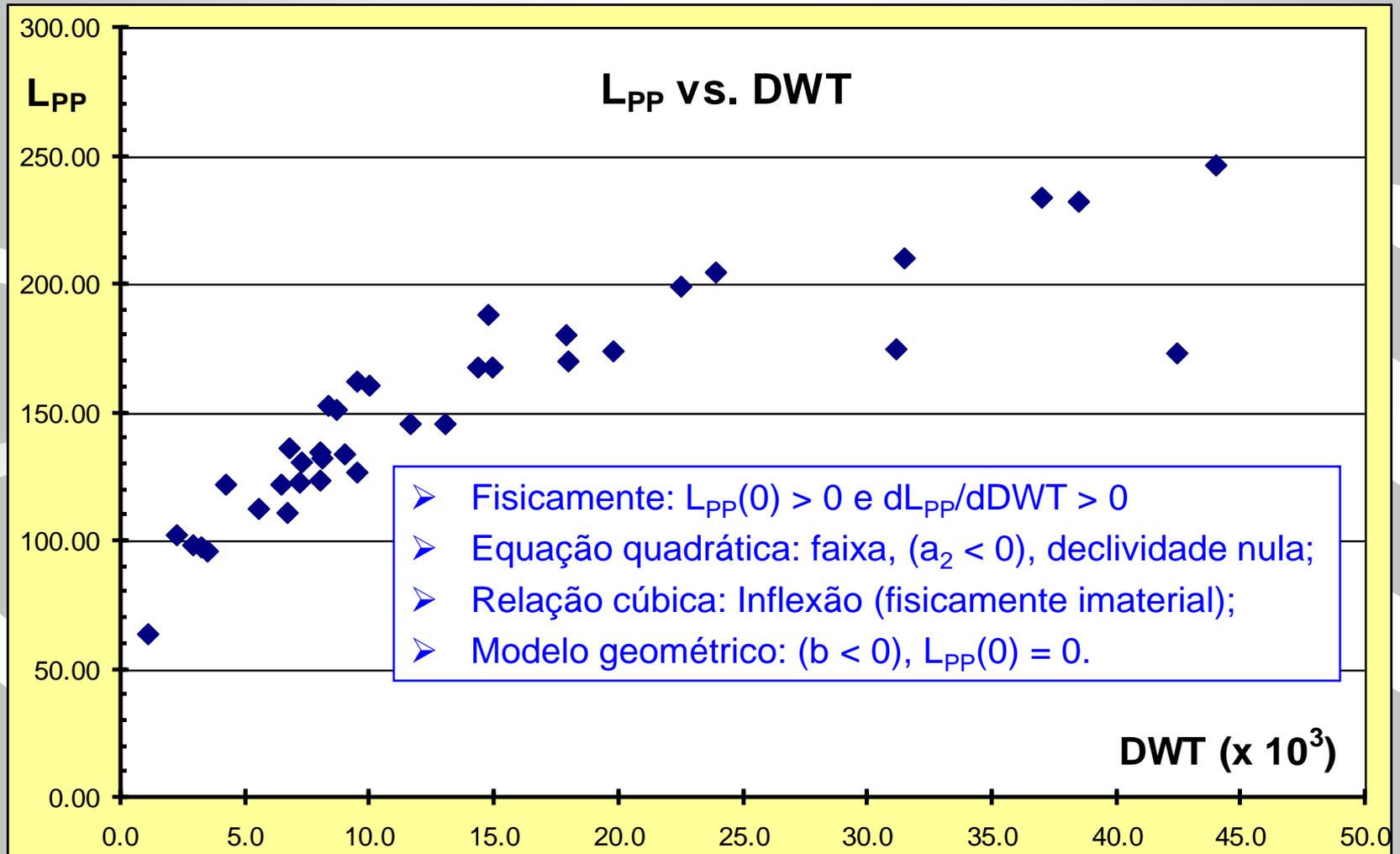
	NOME	LPP [m]
1	Mahano	121.90
2	Seaway Prince	121.90
3	Lys. Endeavor	131.90
4	Iron Monarch	167.90
5	Aust. Enterprise	167.90
6	Union Rotorua	188.20
7	Aust. Emblem	205.00
8	Saudi Abha	232.00
9	Mercand. Pres.	96.00
10	Mercand. Duke	122.90
11	Roll-Vigo	112.80
12	Baltic Eagle	126.50
13	Dana America	124.00
14	Arcturus	146.00
15	Esja	64.00
16	Monthlery	102.00
17	Mario	97.50
18	Skaugran	173.20
19	Ville Dunkg.	160.60
20	Linne	152.50
21	Dana Optima	96.10
22	Tajura	98.50
23	Barber Tampa	246.40
24	Jacqueline	131.00
25	Karisma	136.00
26	Con Ros NKK	170.00
27	Atl. Comp.	233.60
28	Ro-Ro Sea Link	111.20
29	GD Quincy	199.00
30	Nopal Barbro	180.00
31	Transfinland	146.00
32	Andra Mazario	174.00
33	Boogabila	210.00
34	Helenic Explorer	175.00
35	Atl. Prosper	151.00
36	Atl. Prelude	162.30
37	Boxer	133.50
38	Nordana	135.00

- Valor máximo e mínimo ⇒ Verificação rápida dados de entrada.
- Outros parâmetros: Análise dados em relação distribuição normal;
- Assimetria (“Skewness”) ⇒ assimetria em relação à média;
- Curtose ⇒ Achatamento (“peakness”) dados em relação à distribuição normal.

Estatísticas	
Variável	L <sub>PP</sub>
Observações	38
Faixa	182.40
Mínimo	64.00
Máximo	246.40
Mediana	146.00
Valor médio	150.09
Erro padrão	6.8626
Variância	1.7896E+03
Desvio padrão	4.2304E+01
Coef. Variação	2.8186E+01
Skewness	3.7930E-01
Kurtosis	-4.4486E-01

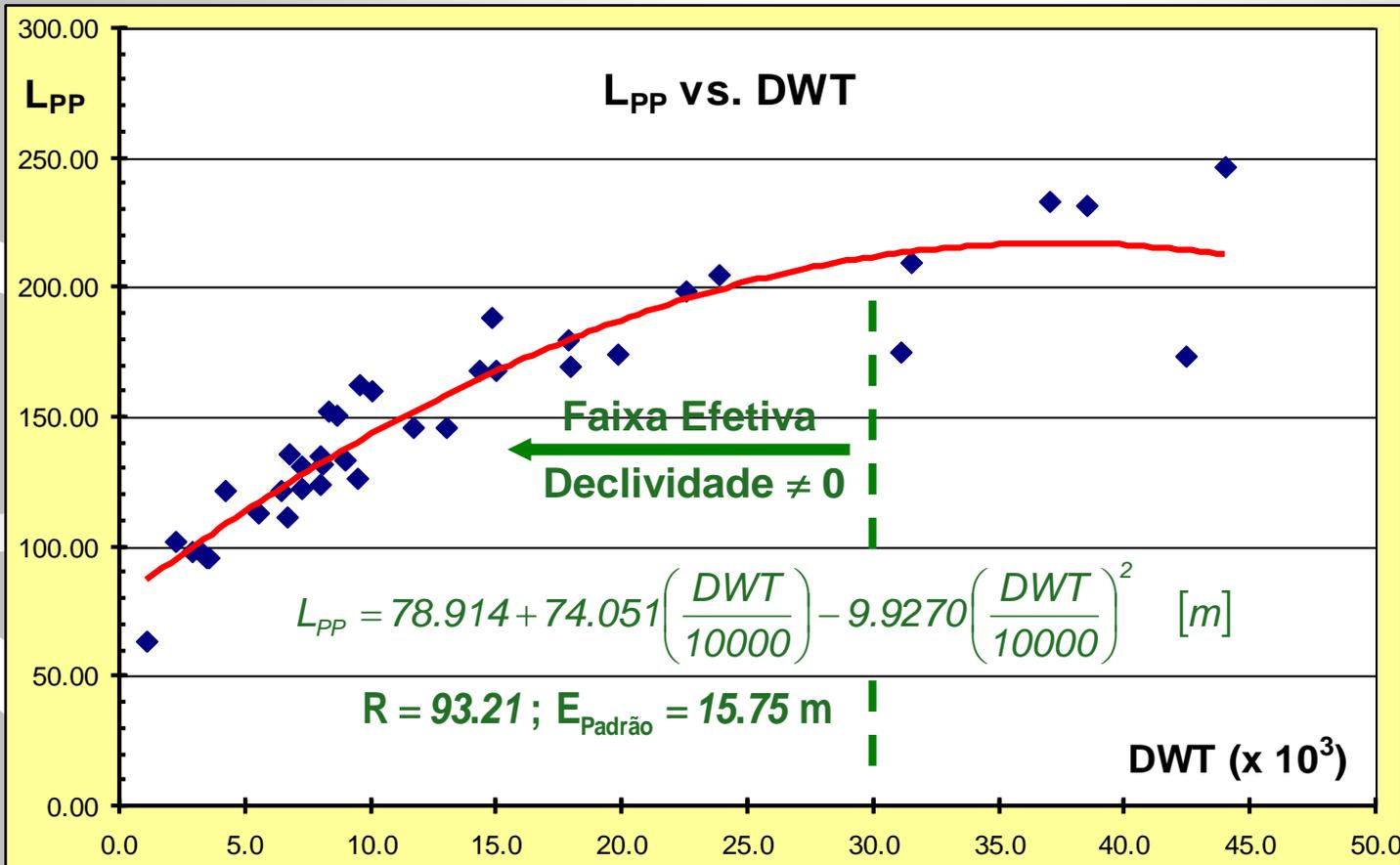
# Exemplo: Navios Ro-Ro

## ■ Diagrama de dispersão - $L_{pp}$ vs DWT



# Exemplo: Navios Ro-Ro

■  $L_{PP}$  vs DWT ; Modelo quadrático

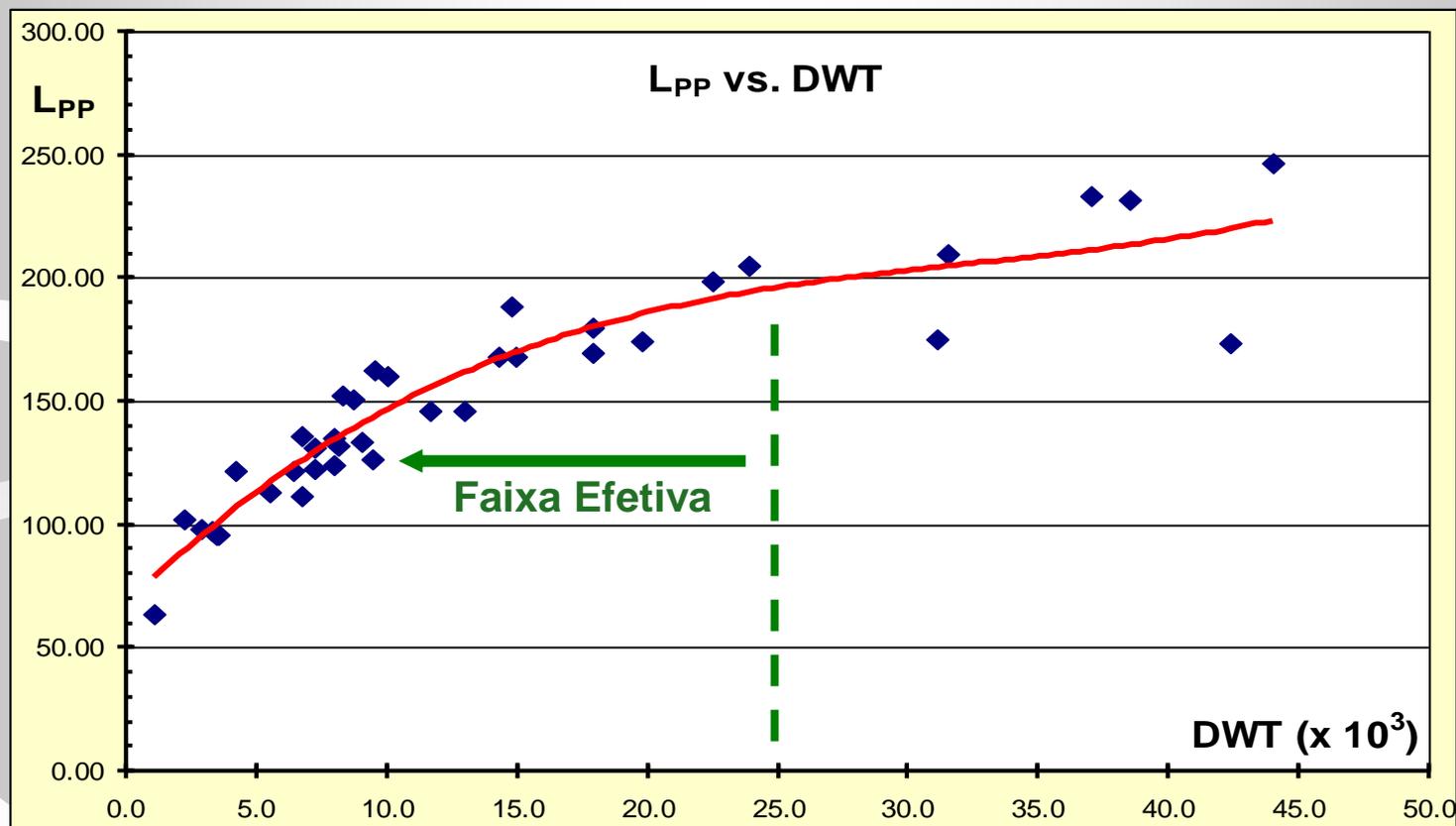


Variável dependente: $L_{PP}$	
Coeficiente	
K	7.8914E+01
DWT	7.4051E-03
DWT <sup>2</sup>	-9.9270E-08
Estimativa Erro padrão	
K	6.6460E+00
DWT	6.7115E-04
DWT <sup>2</sup>	1.9736E-08
Estatística t	
K	1.1874E+01
DWT	8.5003E+00
DWT <sup>2</sup>	-5.0299E+00
Prob > 1	
K	5.2200E-14
DWT	3.9500E-10
DWT <sup>2</sup>	1.3700E-05

➤ Boa aproximação: Estatística t alta e prob > t baixa

# Exemplo: Navios Ro-Ro

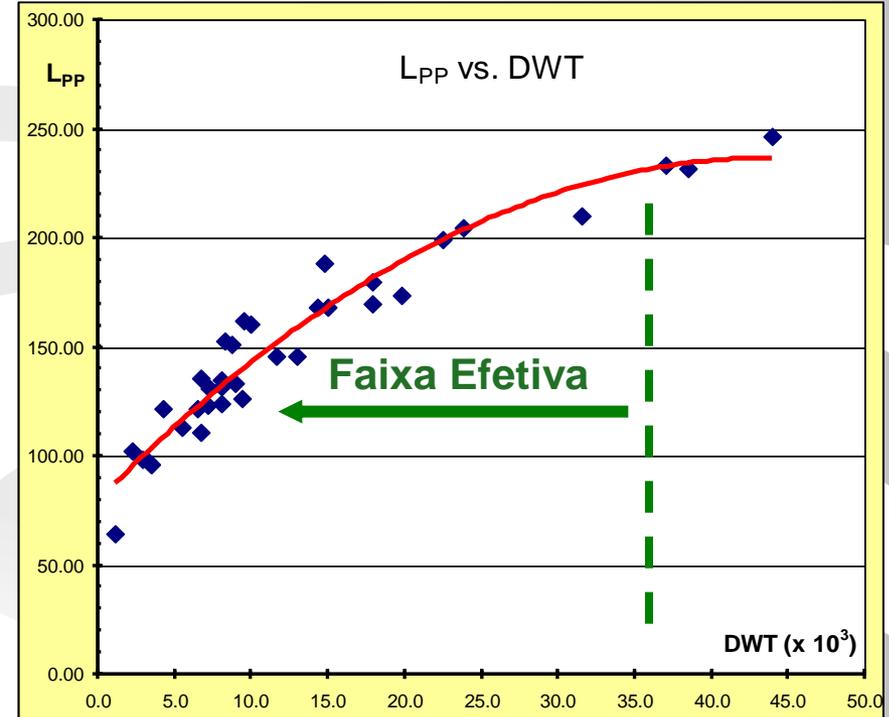
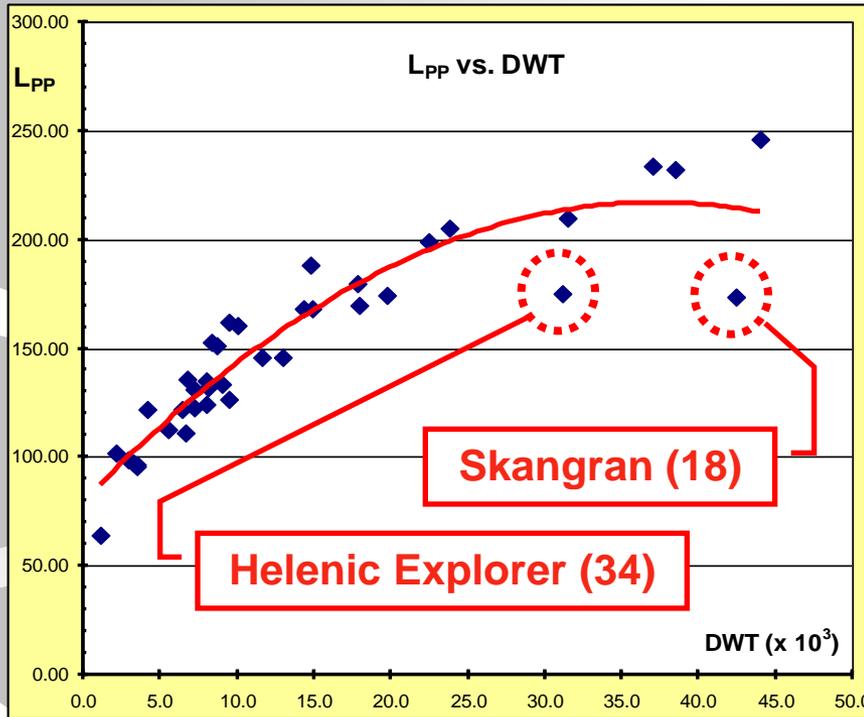
■  $L_{PP}$  vs DWT ; Modelo cúbico



	Coeficiente Determinação	Coeficiente Correlação	Erro Padrão Estimativa	Estatística Durbin-Watson	Razão F	Prob > F
<b>Cúbica</b>	<b>0.879</b>	<b>0.938</b>	1.5332E+01	2.1827	8.2559E+01	1.1000E-15
<b>Quadrática</b>	0.869	0.932	1.5751E+01	2.0656	<b>1.1595E+02</b>	<b>3.6300E-16</b>

# Exemplo: Navios Ro-Ro

## ■ $L_{PP}$ vs DWT ; Modelo quadrático



- Helenic Explorer: Primeiros Ro-Ro construídos;
- Skangran: Produtos florestais 1 pernada rota.

$$L_{PP} = 79.594 + 71.300 \left( \frac{DWT}{10000} \right) - 8.085 \left( \frac{DWT}{10000} \right)^2 \quad [m]$$

$$\begin{cases} R = 96.69 \\ E_{\text{Padrão}} = 12.32m \end{cases}$$

$$\begin{cases} DWT = 20000 \text{ ton} \Rightarrow L_{PP} = 189.9m \\ E_{\text{Padrão}} = \sigma = 12.32m \\ P(\bar{x} - \sigma \leq x \leq \bar{x} + \sigma) = 92\% \end{cases}$$

# Exemplo: Navios Ro-Ro

- Velocidade → Volume praça de máquinas ⇒  $L_{pp}$  para DWT

$$L_{pp} = a + bDWT + cDWT^2 + dV_s$$

- Regressão não-linear (multi-linear).

$$L_{pp} = 56.589 + 62.433 \left( \frac{DWT}{10000} \right) - 6.2355 \left( \frac{DWT}{10000} \right)^2 + 1.4973V_s \quad [m]$$

$$\left\{ \begin{array}{l} R = 97.94 \\ E_{\text{Padrão}} = 10.00m \end{array} \right.$$



**FIM**