



PRINCÍPIOS MATEMÁTICOS PARA O DESENVOLVIMENTO DE PESQUISAS CIENTÍFICAS

ANÁLISE DE DADOS NA CIÊNCIA | MODELAGEM
ESTATÍSTICA E OTIMIZAÇÃO

CIÊNCIA E TECNOLOGIA
REINARUCIA

| 4 a 7 de agosto, 2026



LOCAL:

IFPA Câmpus Ananindeua - Pa

CIÊNCIA E TECNOLOGIA
BACHARELADO

Prof. Dr. Denis Costa



ONDE A MATEMÁTICA ENCONTRA A CIÊNCIA, SURGEM NOVAS POSSIBILIDADES PARA TRANSFORMAR CONHECIMENTO EM INOVAÇÃO

ANÁLISE DE DADOS NA CIÊNCIA | MODELAGEM
ESTATÍSTICA E OTIMIZAÇÃO

CIÊNCIA & TECNOLOGIA
BACHARELADO

| 4 a 7 de agosto, 2026



LOCAL:

IFPA Câmpus Ananindeua - Pa

CIÊNCIA & TECNOLOGIA
BACHARELADO

Prof. Dr. Denis Costa

Modelagem

Processo de estruturação, Matemático-Estatístico, que compreende a análise visual dos dados até a validação do Modelo.

1. Coleta e Organização dos Dados → Padronização em pares ordenados (x_i, y_i) . Verifica-se se há dados ausentes, erros de digitação ou *outliers* (valores discrepantes) capazes de distorcer a análise.

2. Análise Exploratória (Diagrama de Dispersão) → A variável independente no eixo horizontal. O comportamento dos pontos: seguem uma linha reta? **Modelo Linear**. Formam uma parábola? **Modelo Quadrático**. Crescem ou decrescem de forma acelerada? **Modelo Exponencial ou Logarítmico**. Oscilam de forma periódica? **Modelo Senoidal**.

3. Proposição do Modelo Matemático → Escolha da função teórica $f(x, \beta)$ que melhor se adapta à tendência visual observada:

Tipo de Relação	Formato da Equação	Parâmetros a Ajustar
Linear	$y = ax + b$	a, b
Quadrática	$y = ax^2 + bx + c$	a, b, c
Exponencial	$y = a \cdot e^{bx}$	a, b
Potência	$y = a \cdot x^b$	a, b
Logarítmica	$y = a + b \cdot \ln x$	a, b
Senoidal	$y = A \cdot \text{sen}(B \cdot x + C) + D$	A, B, C, D

$$\forall \beta(a, b, c, d)$$

4. Estimaco dos Parmetros (Ajuste) → Mtodo dos Mnimos Quadrados (MMQ). O objetivo  minimizar a soma dos quadrados dos resduos (diferena entre o valor real y_i e o valor previsto $\hat{y}_i$):

Minimizar

$$S = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - f(x_i, \beta))^2$$

5. Avaliaco e Qualidade do Ajuste: a) Coeficiente de Determinaco (R^2) → Indica a proporo da varincia de y explicada por x ; melhor ajuste quanto mais prximo de 1. b) Anlise de Resduos → Grfico dos resduos ($e_i = y_i - \hat{y}_i$) em relao a x ; ajuste bom apresenta resduos aleatrios, distribudos em torno do zero. c) Erro Quadrtico Mdio (RMSE) → Mede a magnitude mdia do erro na varivel y .

6. Validação e Predição:

Validação → Se houver dados suficientes, divida o conjunto inicial em treino (para encontrar a equação) e teste (para verificar se a equação funciona bem com dados novos).

Predição → Use a equação final para estimar y para novos valores de x dentro do intervalo analisado (interpolação).

Aplicação Científica 1 → Considere um estudo de **Cinética de Adsorção** (muito comum na Química Ambiental e Ciência de Materiais), em que:

t : Tempo de contato em minutos (min).

$q(t)$: Capacidade de adsorção em miligramas por grama (mg/g).

- Em estudos práticos de laboratório, como nos ensaios de **cinética química, cinética de adsorção ou crescimento biológico**, frequentemente monitora-se a evolução de uma grandeza ao longo do tempo ou em resposta a uma condição controlada.
- Uma pesquisadora analisa como a **capacidade de retenção/acumulação de uma substância** varia à medida que o tempo passa. O objetivo da Modelagem Matemática-Estatística é identificar a lei matemática subjacente que rege essa dinâmica, permitindo prever comportamentos ou entender os mecanismos físicos/químicos do processo.

- Como todo ensaio de bancada está sujeito a pequenas variações (erros de amostragem, flutuações de temperatura ou limitações de precisão dos equipamentos), aplica-se uma parcela de **ruído experimental** sobre as medições para representar a incerteza real do processo.
- Variáveis envolvidas:

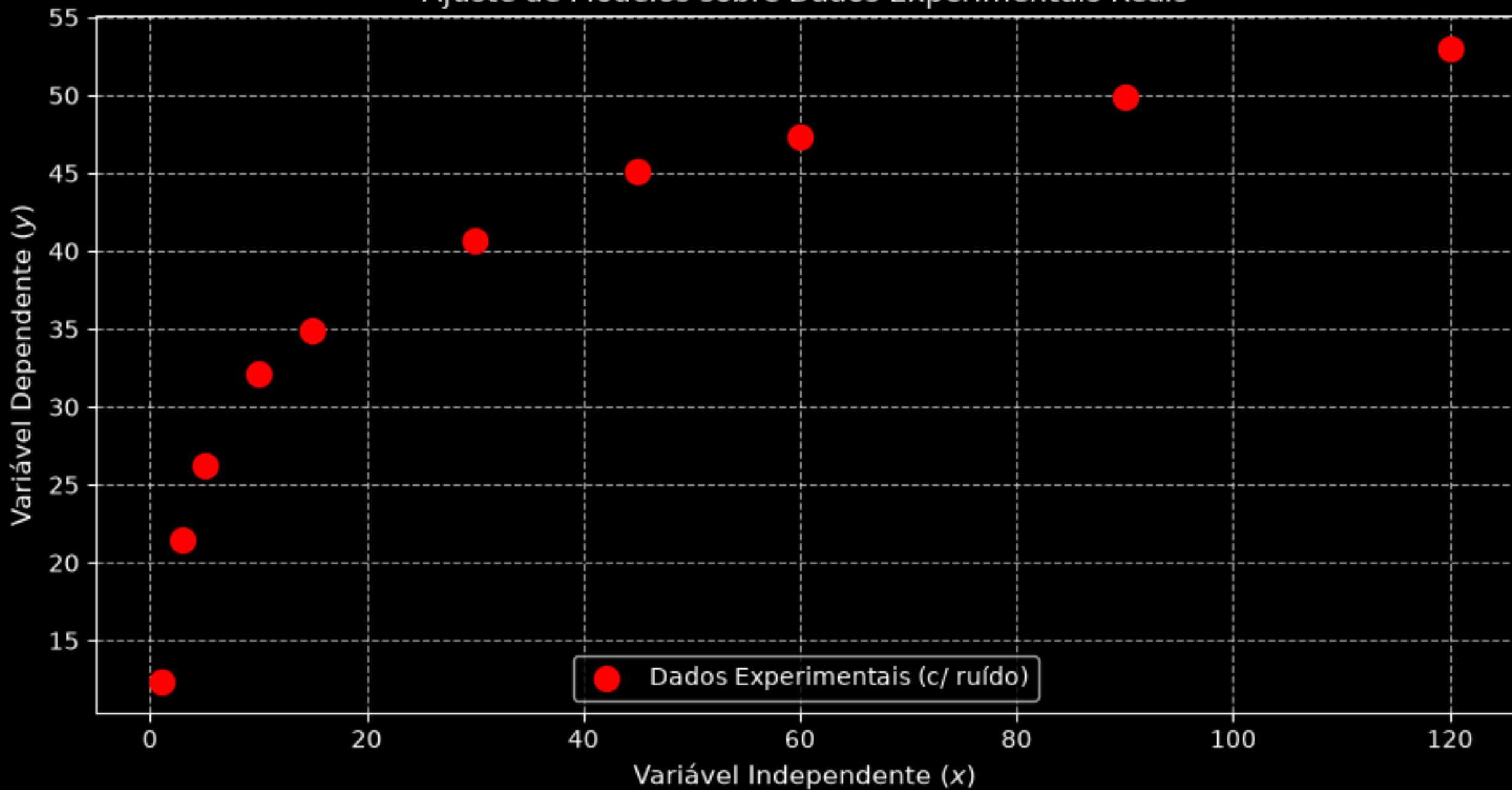
t (Variável Independente) → Representa o Tempo de Ensaio (em minutos). É a variável sob controle do pesquisadora.

q (Variável Dependente Teórica/Real) → Valor base de resposta de interesse (concentração ou capacidade de adsorção em mg/g).

Tabela 1: Dados Experimentais

Observação	Variável Independente (X) (Tempo - min)	Resposta Base (Ypuro) (Capacidade - mg/g)
1	1,0	12,10
2	3,0	21,50
3	5,0	25,80
4	10,0	31,20
5	15,0	35,00
6	30,0	40,80
7	45,0	44,10
8	60,0	46,90
9	90,0	50,20
10	120,0	52,70

Ajuste de Modelos sobre Dados Experimentais Reais



=====

RANKING DOS MODELOS TESTADOS

=====

Logarítmica		$R^2 = 0.9991$
Quadrática		$R^2 = 0.9022$
Linear		$R^2 = 0.7483$
Exponencial		$R^2 = 0.6725$
Senoidal		$R^2 = 0.4207$

=====

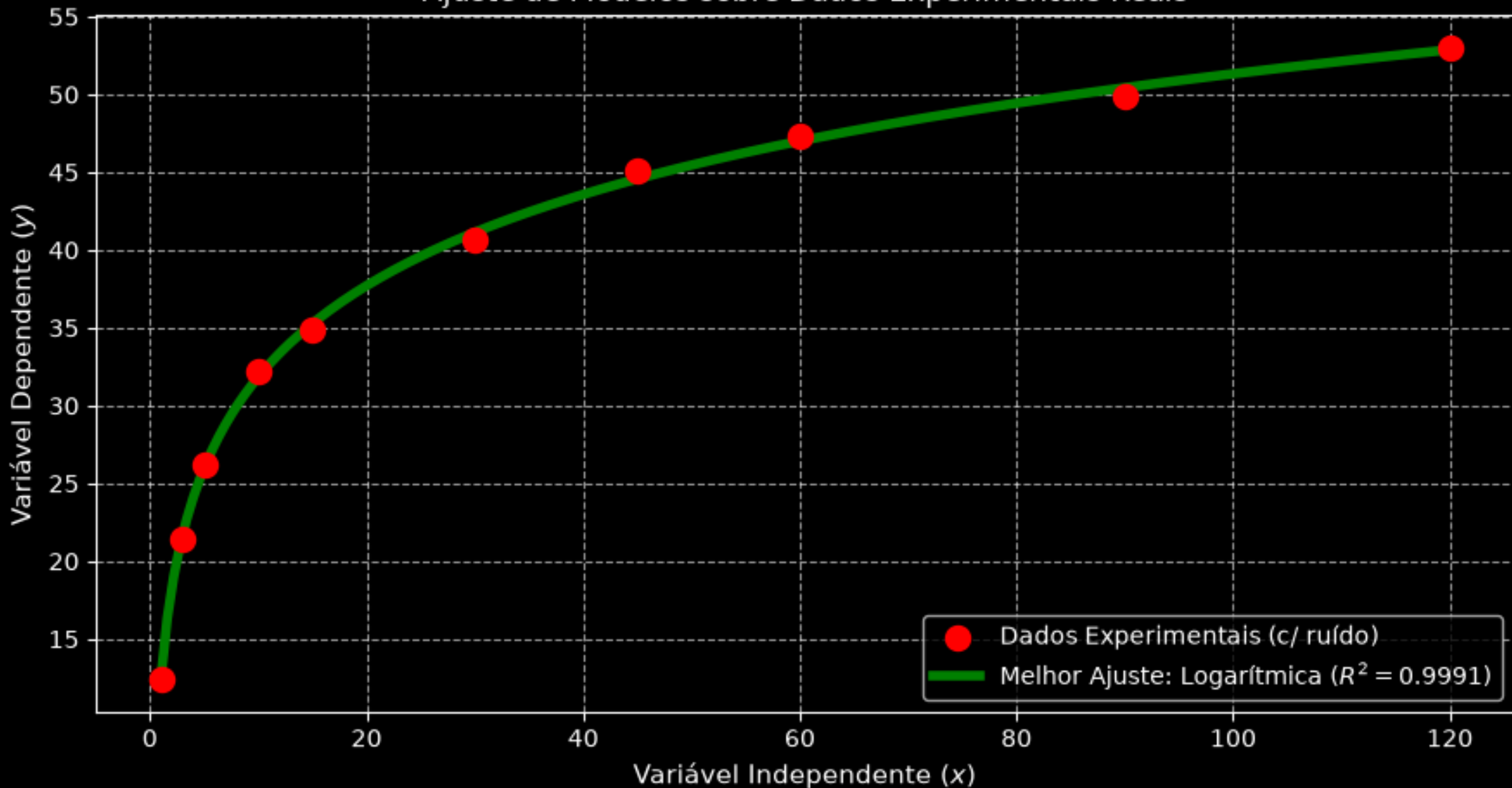
MELHOR MODELO SELECIONADO: LOGARÍTMICA

Estrutura Teórica: $y = a + b * \ln(x)$

Parâmetros Otimizados: [12.3551 8.4633]

Coeficiente de Determinação (R^2): 0.9991

Ajuste de Modelos sobre Dados Experimentais Reais



Aplicação Científica 2

O Bacharelado em Ciência e Tecnologia realiza estudos ambientais com o monitoramento da geração de energia solar fotovoltaica, essencial para a gestão da rede elétrica e dimensionamento de sistemas de armazenamento (baterias).

No estudo avalia-se a Potência Gerada por um painel solar desde o amanhecer até o anoitecer. O comportamento típico da radiação solar apresenta uma elevação progressiva durante a manhã, atinge um pico próximo ao meio-dia e decai suavemente até o final da tarde.

Esse comportamento oscilatório e simétrico com ciclo diário é ideal para ser modelado por Funções Senoidais (ou modelos gaussianos/parabólicos). No entanto, interrupções pontuais como passagem de nuvens, sombreamento transitório e variações de temperatura nos painéis introduzem instabilidades e ruídos nas medições reais.

A seguir estão os dados monitorados em um sistema fotovoltaico ao longo de um período de 12 horas (das 06:00 às 18:00):

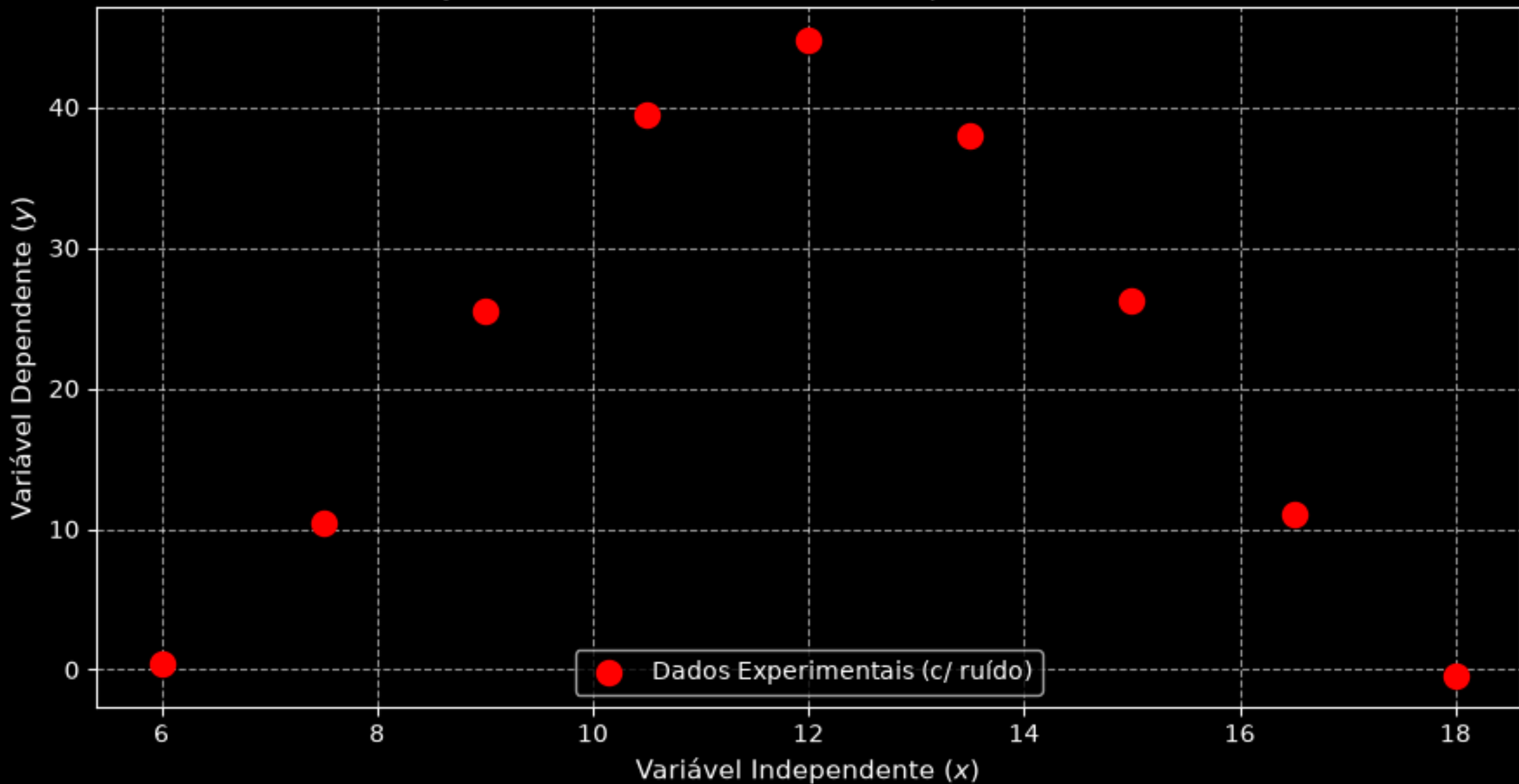
X (Variável Independente) $\rightarrow$ Representa o Horário do Dia (em horas decimais, de 6,0 h a 18,0 h).

Y_{puro} (Variável Dependente Teórica) $\rightarrow$ Geração ideal esperada em um dia de céu 100% limpo (em quilowatts - kW).

Tabela 2: Ydados (Com Ruído Experimental) → Potência efetivamente registrada pelos inversores, contendo as flutuações reais causadas por nuvens e variações térmicas.

Observações	Variável Independente (X) (Horário - horas)	Resposta Ideal (Ypuro) (Potência Teórica - kW)	Variável Dependente (Y) (Com Ruído de Nuvens/Medição)
1	6,0 (06:00)	0,00	0,82
2	7,5 (07:30)	10,50	11,98
3	9,0 (09:00)	25,00	23,12
4	10,5 (10:30)	38,20	40,15
5	12,0 (12:00)	45,00	42,91
6	13,5 (13:30)	38,20	39,84
7	15,0 (15:00)	25,00	23,45
8	16,5 (16:30)	10,50	11,73
9	18,0 (18:00)	0,00	-0,65

Ajuste de Modelos sobre Dados Experimentais Reais



=====

RANKING DOS MODELOS TESTADOS

=====

Senoidal		$R^2 = 0.9989$
Quadrática		$R^2 = 0.9486$
Logarítmica		$R^2 = 0.0205$
Linear		$R^2 = 0.0000$
Exponencial		$R^2 = 0.0000$

=====

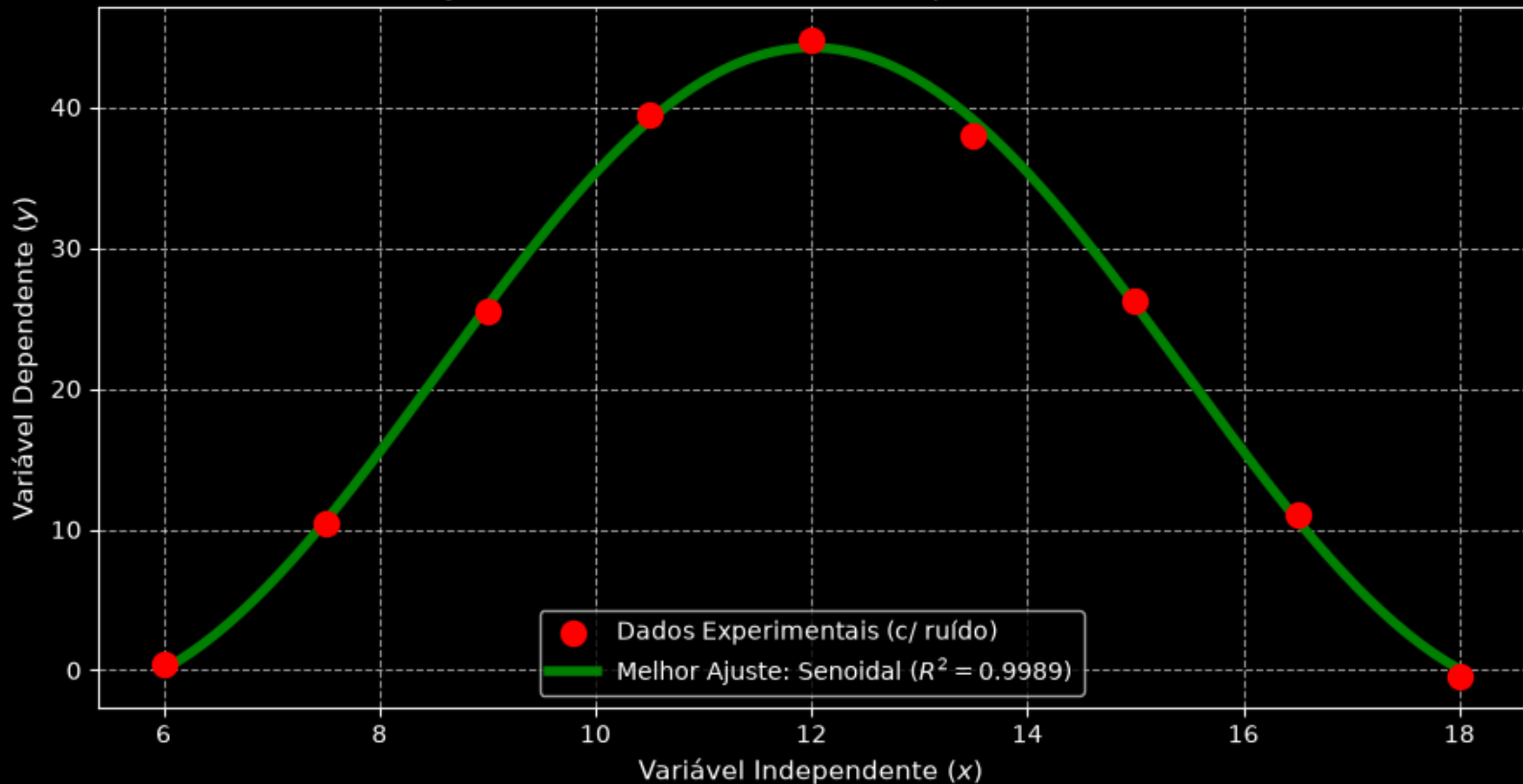
MELHOR MODELO SELECIONADO: SENOIDAL

Estrutura Teórica: $y = A \cdot \sin(B \cdot x + C) + D$

Parâmetros Otimizados: [23.0591 0.4546 -3.8856 21.1664]

Coeficiente de Determinação (R^2): 0.9989

Ajuste de Modelos sobre Dados Experimentais Reais

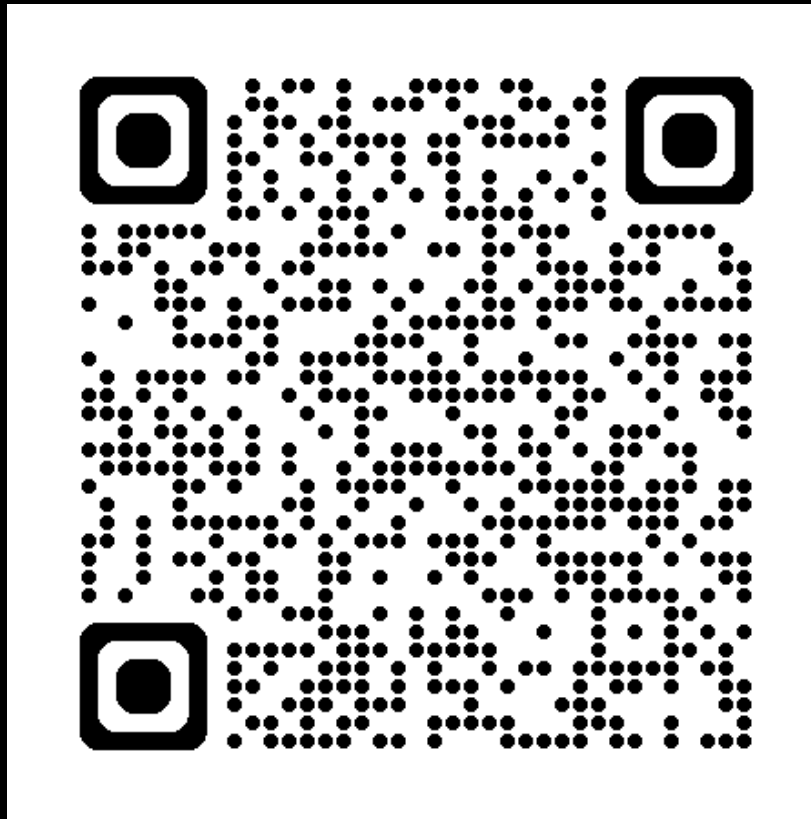


Princípios Matemáticos para o desenvolvimento de Pesquisas Científicas: Ajuste otimizado de curvas em 2D

@author: Denis C. L. Costa

Bacharelado em Ciência e Tecnologia

DOI → <https://doi.org/10.5281/zenodo.21793812>





Princípios Matemáticos para o desenvolvimento de Pesquisas Científicas:

Ajuste otimizado de curvas em 2D

Dr. Denis C. L. Costa, Professor do Bacharelado em Ciência e Tecnologia

Tabela 1: Dados Experimentais.

Observação	Variável Independente (X) (Tempo - min)	Resposta Base (Y _{puro}) (Capacidade - mg/g)
1	1,0	12,10
2	3,0	21,50
3	5,0	25,80
4	10,0	31,20
5	15,0	35,00
6	30,0	40,80
7	45,0	44,10
8	60,0	46,90
9	90,0	50,20
10	120,0	52,70

Tabela 2: Potência efetivamente registrada pelos inversores.

Observações	Variável Independente (X) (horas)	Resposta Ideal (Y _{puro})	Variável Dependente (Y)
1	6,0 (06:00)	0,00	0,82
2	7,5 (07:30)	10,50	11,98
3	9,0 (09:00)	25,00	23,12
4	10,5 (10:30)	38,20	40,15
5	12,0 (12:00)	45,00	42,91
6	13,5 (13:30)	38,20	39,84
7	15,0 (15:00)	25,00	23,45
8	16,5 (16:30)	10,50	11,73
9	18,0 (18:00)	0,00	-0,65