

Dados Solar

Author ^a

^a

Abstract

In this paper,

Keywords: text ...

1. Introdução

2. Material e Método

Considere que a potência gerada por placas fotovoltaicas limpas, Y , sejam monitoradas por n dias, $n \geq 1$. Suponha que para o j -ésimo dia sejam feitas k_j medições das potências geradas, para $j = 1, \dots, n$. Consideramos como instante inicial de medição do dia j , t_{0j} , o horário em que a primeira medida de potência é gerada, $Y > 0$. Então, a cada $h = 10$ minutos foram anotadas as medições das potências geradas até o instante $t_{k_j j}$, onde $t_{k_j j}$ é o último horário de anotação antes da placa deixar de gerar potência, $Y = 0$.

Para as análises descritas na próxima Seção, consideramos a transformação logarítmica para os valores das potências anotadas. Como ilustração, a Figura 1 mostra as potências geradas nos dias 1 e 2 de realização do experimento. No dia 1, a primeira medição foi feita às 5 horas e 29 minutos e a última medição foi feita às 17 horas e 39 minutos, totalizando $k_1 = 74$ medições. Para o dia 2, a primeira medição foi feita às 5 horas e 29 minutos e a última medição foi feita às 17 horas e 49 minutos, totalizando $k_1 = 75$ medições. No total, o experimento foi realizado durante $n = 20$ dias.

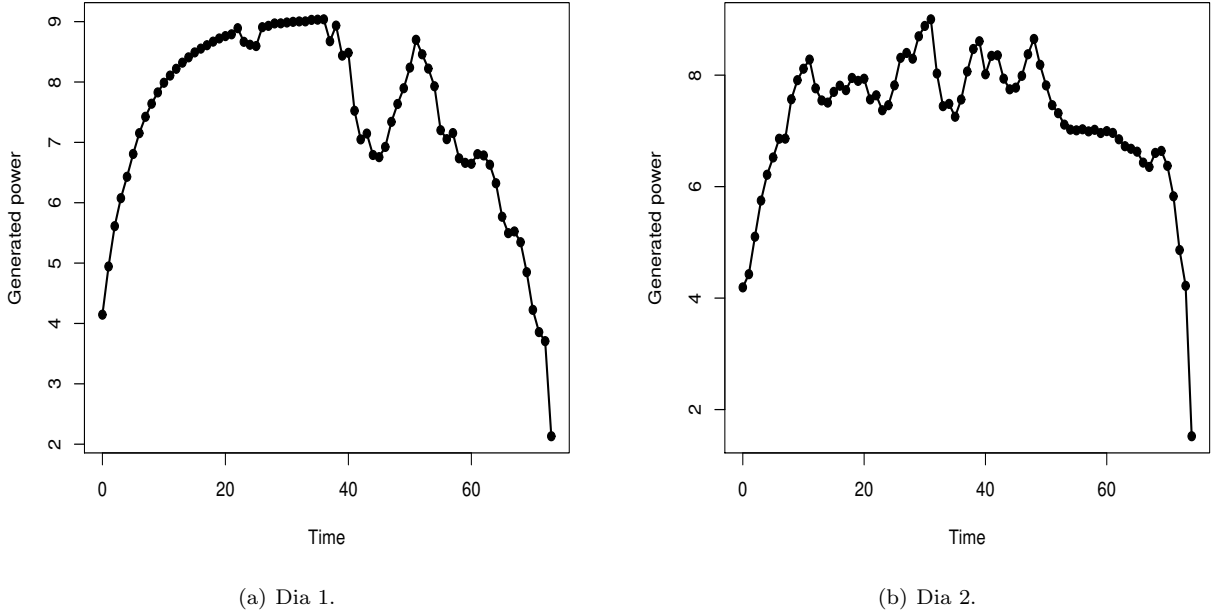


Figure 1: Potências geradas nos dias 1 e 2.

A Tabela 2 mostra as estatísticas resumo das potências geradas nos 20 dias de realização do experimento. O menor valor foi de 3.857. Na escala original, este valor equivale a 47.3232 kw/h. O primeiro quartil é igual a 6.904 (996.2518 kw.h), *i.e.*, 25% das potências geradas nos 20 dias são menores ou iguais a 996.2518 kw/h. O valor mediano foi de 8.065 (3181.156 kw/h), com média de 7.671 (2145.226 kw/h) e desvio-padrão de 1.2938. O terceiro quartil é igual a 8.775 (6470.444 kw.h), *i.e.*, 75% das potências geradas nos 20 dias são menores ou iguais a 2470.444 kw/h e maior valor gerado foi de 9.041 (8442.215 kw/h). A Figura 2 mostra o *boxplot* das potências geradas por dia.

Table 1: Estatísticas resumo.

Mínimo	3° quartil	mediana	Média	desvio-padrão	3° quartil	Máximo
3.857	6.904	8.065	7,671	1.2938	8.775	9.041

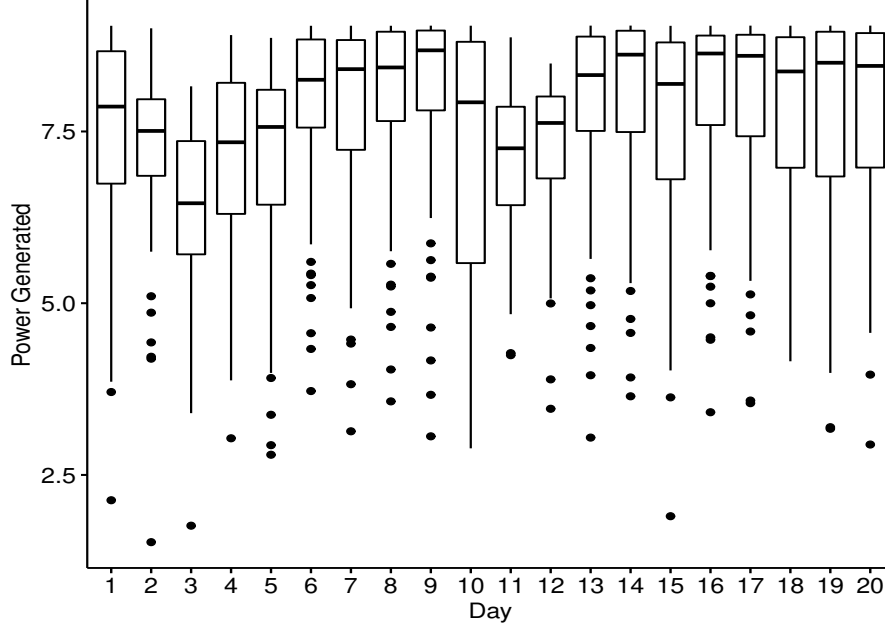


Figure 2: Potências geradas por dia.

2.1. Método

Seja Y_{ij} a potência gerada (log-transformada) no i -ésimo instante do dia j , para $i \in \{t_{0j}, \dots, t_{k_{jj}}\}$ e $j = 1, \dots, n$. Considere as seguintes variáveis explicativas

- T = Tempo;
- I = Irradiancia ;
- X = Temperatura;
- M = Massa acumulada de particulados.

Para mantermos um escala adequada para todas as variáveis, consideramos a transformação logarítmica para os valores observados das quatro variáveis explicativas.

De acordo com a Figura 1, considere a seguinte relação quadrática entre a variável resposta Y e o tempo T ,

$$\begin{aligned} Y_{ij} &= \gamma_{0j} + \beta_1 T_{ij} + \beta_2 T_{ij}^2 + \varepsilon_{ij}, \\ \gamma_{0j} &= \beta_0 + U_{0j}, \end{aligned} \quad (1)$$

onde $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$, $U_{0j} \sim \mathcal{N}(0, \sigma_0^2)$, $cov(\varepsilon_{ij}, \varepsilon_{ij'}) = 0$, $cov(U_{0j}, U_{0j'}) = 0$ e $\varepsilon_{i,j}$ e U_{0j} sendo independentes, para $i \in \{t_{0j}, \dots, t_{k_{jj}}\}$ e $j = 1, \dots, n$. Dessa forma, obtemos um modelo quadrático com intercepto aleatório. Denominamos este modelo de modelo M_0 .

Ajustamos o modelo M_0 aos dados observados utilizando o software R e o comando *lme* do pacote *lme4*. A Tabela 2 mostra as estimativas para os parâmetros dos efeitos fixos do modelo. Para um nível de significância $\alpha = \{1\%, 5\%, 10\%\}$, temos $p\text{-valor} < \alpha$, indicando que o intercepto e a variável explicativa T são significativas para explicação da variável resposta Y . As estimativas para os desvios-padrão são $\hat{\sigma} = 0.5717$ e $\hat{\sigma}_0 = 0.4931$.

Table 2: Estimativas para os efeitos fixos.

Parâmetros	Estimativa	Erro padrão	G.L.	Estatística t	p-valor
β_0	5.5118	0.119362	1.367	46.1767	0
β_1	0.1843	0.002840	1.367	64.8793	0
β_2	-0.0026	0.000037	1.367	-69.1586	0

Como as medidas de potência gerada dentro de cada dia são feitas longitudinalmente, é plausível assumir que os erros aleatórios ε_{ij} são correlacionados. Assim, considere que ε_{ij} e $\varepsilon_{(i+s)j}$ são os dois erros do j -ésimo dia separado por s unidades de tempo. Assim de acordo com um modelo auto-regressivo (AR) de primeira ordem, a correlação entre os dois erros é dada por $\text{corr}(\varepsilon_{ij}, \varepsilon_{(i+s)j}) = \rho^{|s|}$, para $i \in \{t_{0j}, \dots, t_{k_{jj}}\}$ e $j = 1, \dots, n$. Denominamos este modelo de M_1 .

Ajustamos o modelo M_1 utilizando o comando *lme* com a opção *autocorrelation = corAR1*. A Tabela 3 mostra as estimativas para os parâmetros dos efeitos fixos do modelo M_1 . As estimativas para os desvios-padrão são $\hat{\sigma} = 0.7233$ e $\hat{\sigma}_0 = 0.2891$. A estimativa do coeficiente de correlação é $\hat{\rho} = 0.9026$, indicando uma correlação forte.

Table 3: Estimativas para os efeitos fixos.

Parâmetros	Estimativa	Erro padrão	G.L.	Estatística t	p-valor
β_0	4.7481	0.169067	1.367	28.0839	0
β_1	0.2282	0.008926	1.367	25.5636	0
β_2	-0.0031	0.000111	1.367	-17.8462	0

Para verificar a significância estatística do coeficiente de correlação, consideramos o teste da razão de verossimilhanças. Neste teste, a hipótese nula assume que $\rho = 0$. Aplicamos o teste utilizando o comando *anova*(M_0, M_1). Como resultado do teste, temos $p\text{-valor} < 0.0001$, indicando a rejeição da hipótese nula.

Considere agora, o modelo M_1 acrescido da variável explicativa irradiância da seguinte maneira

$$\begin{aligned} Y_{ij} &= \gamma_{0j} + \beta_1 T_{ij} + \beta_2 T_{ij}^2 + \varepsilon_{ij}, \\ \gamma_{0j} &= \beta_0 I_{ij} + U_{0j}, \end{aligned}$$

para $i \in \{t_{0j}, \dots, t_{k_{jj}}\}$ e $j = 1, \dots, n$. Assim, temos que

$$Y_{ij} = U_{0j} + \beta_0 I_{ij} + \beta_1 T_{ij} + \beta_2 T_{ij}^2 + \varepsilon_{ij}, \quad (2)$$

onde, $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$, $U_{0j} \sim \mathcal{N}(0, \sigma_0^2)$, $\text{cov}(\varepsilon_{ij}, \varepsilon_{(i+1)j}) = \rho^{|s|}$, $\text{cov}(U_{0j}, U_{0j'}) = 0$ e $\varepsilon_{i,j}$ e U_{0j} sendo independentes, para $i \in \{t_{0j}, \dots, t_{k_{jj}}\}$ e $j = 1, \dots, n$. Ou seja, assumimos um modelo com intercepto aleatório e dependente da irradiância e do tempo. Denominamos este modelo de modelo M_2 .

A Tabela 4 mostra as estimativas para os parâmetros dos efeitos fixos do modelo. Como $p\text{-valor} < \alpha$, temos que, as variáveis explicativas I , T e T^2 são significativas para explicação da variável resposta Y , para $\alpha = \{1\%, 5\%, 10\%\}$. As estimativas para os desvios-padrão são $\hat{\sigma} = 0.1048$ e $\hat{\sigma}_0 = 2.3273$. A estimativa do coeficiente de correlação é $\hat{\rho} = 0.8832$, indicando uma correlação forte.

Table 4: Estimativas para os efeitos fixos.

Parâmetros	Estimativa	Erro padrão	G.L.	Estatística t	p-valor
β_0	0.9330	0.0040 35	1.367	231.2453	0
β_1	0.0169	0.001572	1.367	10.7648	0
β_2	-0.0002	0.000020	1.367	-10.9359	0

Considere agora o modelo M_3 . Este modelo é dado pelo modelo M_2 acrescido da variável temperatura (X) da seguinte maneira,

$$\begin{aligned} Y_{ij} &= \gamma_{0j} + \beta_1 T_{ij} + \beta_2 T_{ij}^2 + \varepsilon_{ij}, \\ \gamma_{0j} &= \beta_{01} I_{ij} + \beta_{02} X_{ij} + U_{0j}, \end{aligned}$$

para $i \in \{t_{0j}, \dots, t_{k_{jj}}\}$ e $j = 1, \dots, n$. Assim, temos que

$$Y_{ij} = U_{0j} + \beta_0 I_{ij} + \beta_1 T_{ij} + \beta_2 T_{ij}^2 + \beta_3 X_{ij} + \varepsilon_{ij}, \quad (3)$$

onde $\beta_3 = \beta_{02}$, para $i \in \{t_{0j}, \dots, t_{k_{jj}}\}$ e $j = 1, \dots, n$.

A Tabela 5 mostra as estimativas para os parâmetros dos efeitos fixos do modelo. Como p-valor associado ao parâmetro β_3 é maior do que o nível de significância α , temos que, a variável X não é significativa para explicação da variável resposta Y , para $\alpha = \{1\%, 5\%, 10\%\}$. As estimativas para os desvios-padrão são $\hat{\sigma} = 0.1052$ e $\hat{\sigma}_0 = 2.3766$. A estimativa do coeficiente de correlação é $\hat{\rho} = 0.8841$, indicando uma correlação forte.

Table 5: Estimativas para os efeitos fixos.

Parâmetros	Estimativa	Erro padrão	G.L.	Estatística t	p-valor
β_0	0.9336	0.004507	1.366	207.1206	0
β_1	0.0171	0.001621	1.366	10.5342	0
β_2	-0.0002	0.000021	1.366	-10.8353	0
β_3	-0.0151	0.047364	1.366	-0.3191	0.7497

Como modelo M_4 , considere o modelo M_2 acrescido da variável massa acumulada (M) da seguinte maneira,

$$\begin{aligned} Y_{ij} &= \gamma_{0j} + \beta_1 T_{ij} + \beta_2 T_{ij}^2 + \varepsilon_{ij}, \\ \gamma_{0j} &= \beta_{01} I_{ij} + \beta_{02} M_{ij} + U_{0j}, \end{aligned}$$

para $i \in \{t_{0j}, \dots, t_{k_{jj}}\}$ e $j = 1, \dots, n$. Assim, temos que

$$Y_{ij} = U_{0j} + \beta_0 I_{ij} + \beta_1 T_{ij} + \beta_2 T_{ij}^2 + \beta_3 M_{ij} + \varepsilon_{ij}, \quad (4)$$

onde $\beta_3 = \beta_{02}$, para $i \in \{t_{0j}, \dots, t_{k_{jj}}\}$ e $j = 1, \dots, n$.

A Tabela 6 mostra as estimativas para os parâmetros dos efeitos fixos do modelo. Como p-valor associado ao parâmetro β_3 é maior do que o nível de significância α , temos que, a variável m não é significativa para explicação da variável resposta Y , para $\alpha = \{1\%, 5\%, 10\%\}$. Este resultado já era esperado, pois a quantidade de massa acumulada de particulados ainda não foi suficiente para afetar a geração de energia. As estimativas para os desvios-padrão são $\hat{\sigma} = 0.1056$ e $\hat{\sigma}_0 = 2.3271$. A estimativa do coeficiente de correlação é $\hat{\rho} = 0.8849$, indicando uma correlação forte.

Table 6: Estimativas para os efeitos fixos.

Parâmetros	Estimativa	Erro padrão	G.L.	Estatística t	p-valor
β_0	0.9329	0.004038	1.366	231.0003	0
β_1	0.01700	0.001609	1.366	10.5652	0
β_2	-0.0002	0.000021	1.366	-10.8432	0
β_3	-0.00004	0.021231	1.366	0.0019	0.9985

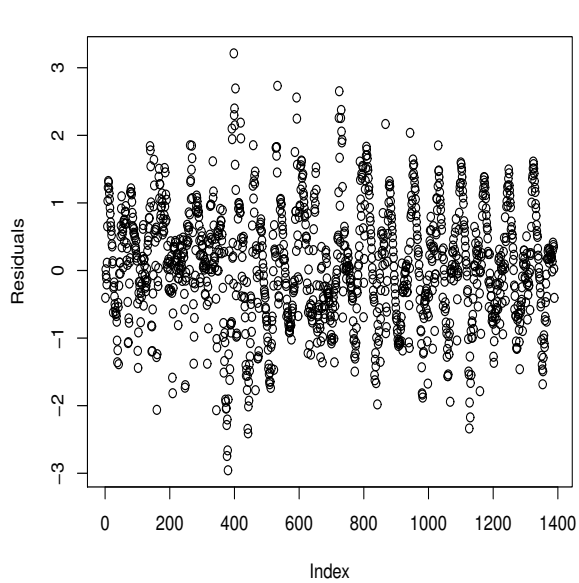
A Tabela 7 mostra os valores dos critérios de seleção de modelos AIC e BIC para os cinco modelos. O modelo com menor valor AIC ou BIC é o modelo que melhor explica os dados observados. De acordo com estes critérios, o modelo M_2 é o melhor modelo.

Table 7: Estimativas para os efeitos fixos.

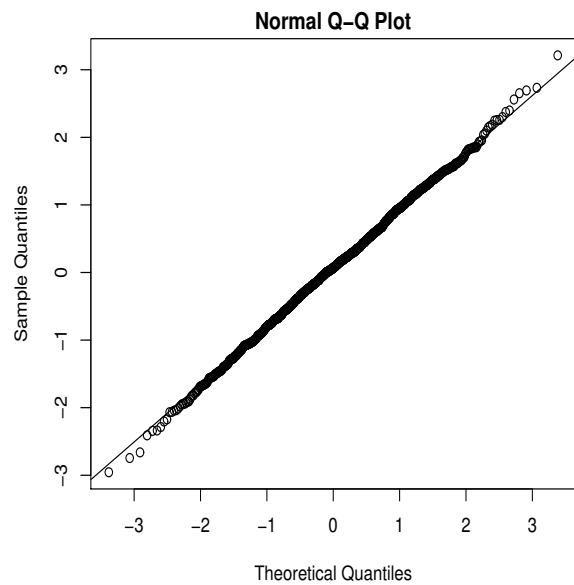
Modelo	M_0	M_1	M_2	M_3	M_4
AIC	2.508,67	783.56	-4.188,47	-4.182,29	-4.180,60
BIC	2.534,84	814.96	-4.157,07	-4.145,66	-4.143,96

A Figura 3, mostra o gráfico dos resíduos do modelo M_2 e o gráfico quantil-quantil. Note que, não temos evidências para rejeitar a suposição homogeneidade da variância e normalidade dos resíduos. O p-valor proveniente do teste de normalidade Shapiro-Wilk para os resíduos é 0.8771, indicando a não rejeição da suposição de normalidade dos resíduos.

Como ilustração da performance do modelo M_2 , a Figura 4 mostra os valores observados e os valores preditos pelo modelo para os dias 1 e 2. O erro quadrático médio para os valores destes dois dias são 0,005504 e 0.005461.

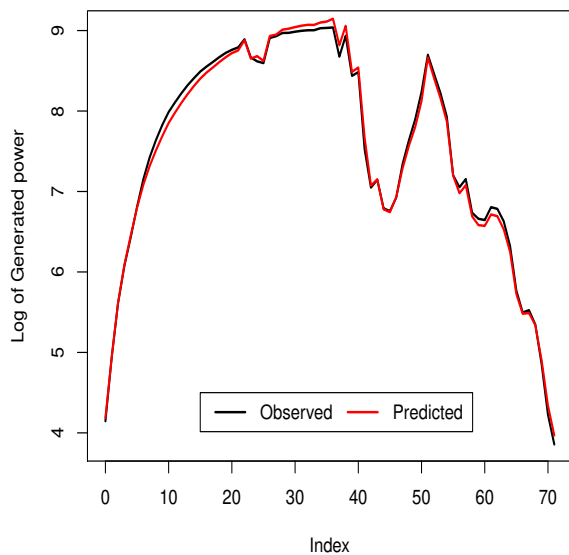


(a) Resíduos.

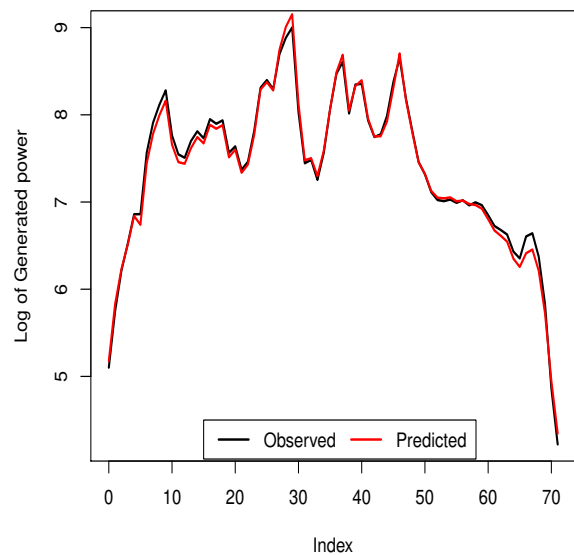


(b) qqplot.

Figure 3: Resíduos do modelo M_2 .



(a) Resíduos.



(b) qqplot.

Figure 4: Valores observados e valores preditos pelo modelo M_2 .

References

Antoniak, C. E. (1974). Mixture of processes dirichlet with applications to bayesian nonparametric problems.
The Annals of Statistics, **2**, 1142-1174.