

Atividade 02 - Modelos de Suavização Exponencial

Caio O. Azevedo e Caren G. C. Bermeo

Abril 2020

1 Introdução

A maioria dos métodos de previsão baseia-se na ideia de que observações passadas contêm informações sobre o padrão de comportamento da série temporal. O propósito dos métodos é distinguir o padrão de qualquer ruído que possa estar contido nas observações e então usar esse padrão para prever valores futuros da série (MORETTIN; TOLOI, 2006).

Uma grande classe de métodos de previsão, que tenta tratar ambas as causas de flutuações em séries de tempo, é das suavizações (MORETTIN; TOLOI, 2006; BUENO, 2011; HYNDMAN; ATHANASOPOULOS, 2018). Neste relatório apresentaremos três modelos de Suavização diferentes.

De início transcorremos a respeito da Suavização Exponencial Simples - SES -, um método que conforme o próprio nome diz, é simples e que tem boas vantagens, como a aplicação não dispendiosa, além da grande flexibilidade permitida pela variação da constante de suavização, dentre outras (MORETTIN; TOLOI, 2006).

Com este intuito utilizamos a série mensal do Consumo de Energia Elétrica (em Gwh) no Brasil (IPEADATA, 2020), para os períodos entre janeiro de 1990 e fevereiro de 2020. Para uma melhor aplicação da SES fora necessário fazer um recorte da série, no sentido de capturar um período sem tendência, assim, foram utilizados os períodos pós janeiro de 2012.

É importante destacar que a SES não é adequada para analisar séries temporais que apresentem tendência (MORETTIN; TOLOI, 2006) e neste sentido se apresenta a Suavização Exponencial de Holt - SEH.

Para a aplicação da SEH foi aproveitado a série anual do número de passageiros embarcados em vôos domésticos no Brasil, pós 1970 (IPEADATA, 2020). Esta série apresenta visualmente uma tendência positiva e aparentemente não contém sazonalidade.

A não presença de sazonalidade é importante devido a SEH não conseguir capturar esse elemento (HYNDMAN; ATHANASOPOULOS, 2018). Caso deseje-se investigar sazonalidade e tendência conjuntamente, deve-se usar a Suavização Exponencial de Holt-Winters - HW -, ou seja, o terceiro modelo explorado neste relatório.

Neste caso utilizou-se a série mensal do Produto Interno Bruto do Brasil (PIB) com dados disponíveis a partir de janeiro de 1990 (IPEADATA, 2020). Finalmente, modelo de

HW com sazonalidade aditiva foi aplicado as observações pós janeiro de 2010, por apresentarem presumivelmente uma maior tendência e sazonalidade. Para a aplicação e estimação dos parâmetros dos três modelos relacionados, foi utilizado o *software* R. Os dados utilizados, assim como toda a programação empregada estão disponíveis no link <<https://github.com/caio-azevedo/Series-Temporais>>.

2 Modelos de Suavização Exponencial

2.1 Modelos para séries localmente constantes

2.1.1 Suavização Exponencial Simples (SES)

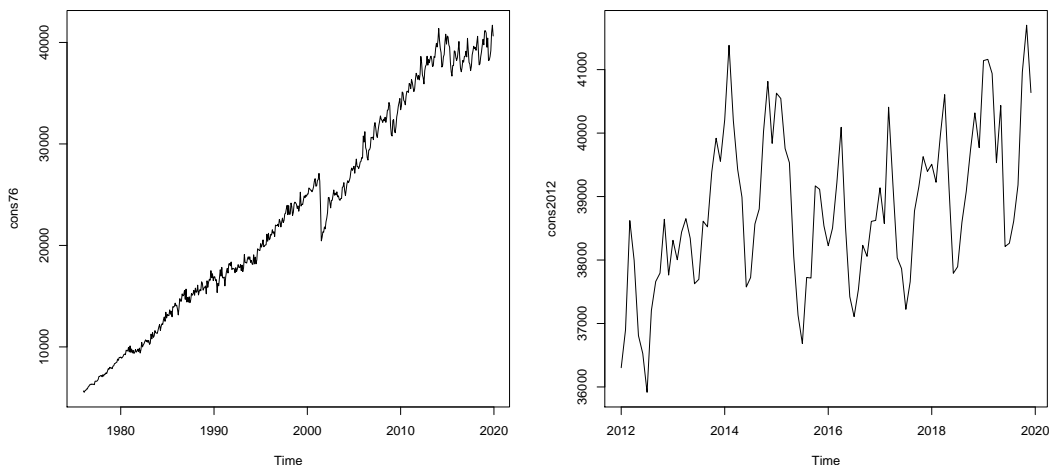
O mais simples dos métodos de suavização exponencial é chamado naturalmente de suavização exponencial simples. Este método é adequado para a previsão de dados sem tendência clara ou padrão sazonal.

Para esta finalidade foi usada a série mensal do Consumo de energia elétrica no Brasil para os períodos entre janeiro de 1990 e fevereiro de 2020. Entretanto, como é possível ver pela Figura 1a a série apresenta grande tendência até 2012, o que inviabilizaria uma boa adequação do método aos dados. E portanto, optou-se por fazer um recorte da série, no sentido de capturar um período sem tendência, e assim, foram utilizados os períodos pós janeiro de 2012.

Figura 1 – Consumo de Energia Elétrica - qde - Gwh.

(a) Jan/1990 - Fev/2020

(b) Jan/2010 - Fev/2020



Fonte: Elaboração própria a partir dos dados do Ipeadata.

Uma representação alternativa apresentada em Hyndman e Athanasopoulos (2018) mostra a esquematização do Método de Suavização Exponencial Simples. Para SES o único componente incluído é o nível, L_t . As representações da forma de componente dos métodos de suavização exponencial compreendem uma equação de previsão (1) e uma equação de suavização (Equação

2) para cada um dos componentes incluídos no método. Sendo assim, as equações da SES é dada por:

$$\hat{Y}_{t+h|t} = L_t, \quad \forall h > 0 \quad (1)$$

$$L_t = \alpha Y_t + (1 - \alpha)L_{t-1}, \quad 0 < \alpha < 1 \quad (2)$$

onde L_t é o nível (ou o valor suavizado) da série no momento t .

É importante salientar que substituindo L_t por $\hat{Y}_{t+1|t}$ da Equação 1 na Equação 2, assim como L_{t-1} por $\hat{Y}_{t|t-1}$, teremos a forma média ponderada de suavização exponencial simples mais conhecida - Equação 3.

Outro ponto de destaque é que como pode ser observado pela Equação 1 a previsão para qualquer período futuro ($h > 0$) será sempre a mesma, independente do valor de h .

$$\hat{Y}_{t+1|t} = \alpha Y_t + (1 - \alpha)\hat{Y}_{t|t-1} \quad (3)$$

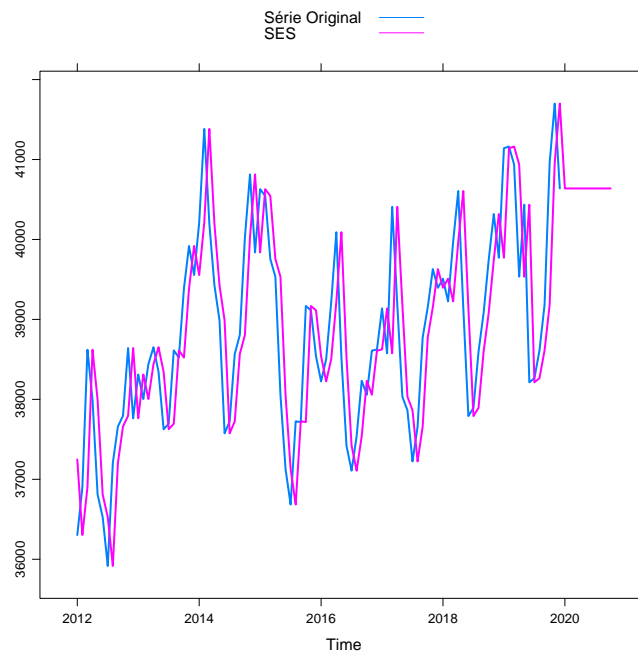
O valor estimado para α é de 0,999 e o estado inicial para L é de 37.248.

Tabela 1 – Previsão para o consumo de energia no Brasil (em Gwh) através da SES

	Valor Previsto	Lo.95	Hi.95
Jan 2020	40.638	38.971	42.306
Feb 2020	40.638	38.280	42.996

Fonte: Elaboração própria a partir dos dados do Ipeadata.

Figura 2 – Suavização Exponencial Simples (SES)



Fonte: Elaboração própria a partir dos dados do Ipeadata.

2.2 Modelo para séries que apresentam tendência

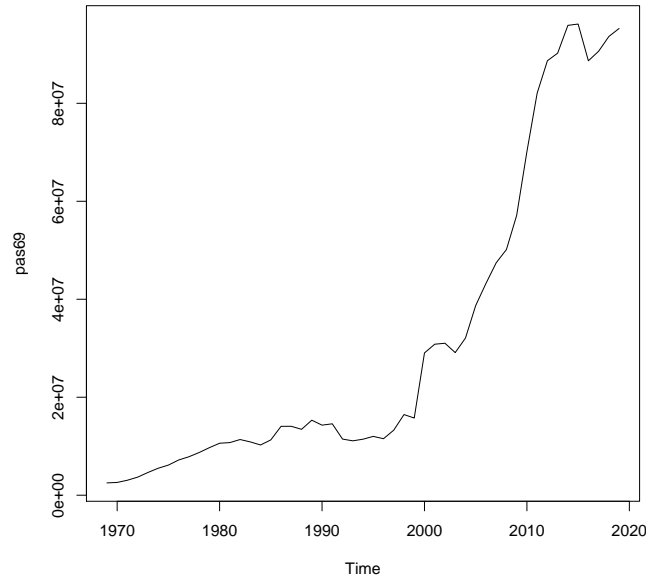
As técnicas vistas anteriormente não são adequadas para analisar séries temporais que apresentem tendência (MORETTIN; TOLOI, 2006). De tal modo que em seguida trataremos de um modelo para séries temporais não sazonais, com tendência e resíduo aleatório com média zero e variância constante (σ^2).

2.2.1 Suavização Exponencial de Holt (SEH)

O método de SES quando aplicado a uma série que apresenta tendência linear positiva(ou negativa), fornece previsões que subestimam (ou superestimam) continuamente os valores reais. Para evitar esse erro sistemático, um dos métodos aplicáveis é a SEH. Este método é similar, em princípio, à SES; a diferença é que em vez de suavizar só o nível, ele utiliza uma nova constante de suavização para "modelar" a tendência da série.(MORETTIN; TOLOI, 2006)

No intuito de apresentar uma série temporal com tendência, optamos pelo uso do Número de passageiros embarcados em vôos domésticos no Brasil. Esta série apresenta tendência positiva (ver Figura 3), tendo apenas alguns poucos períodos de queda.

Figura 3 – Números de passageiros embarcados em vôos domésticos no Brasil



Fonte: Elaboração própria a partir dos dados do Ipeadata.

Holt (1957) estendeu a suavização exponencial simples para permitir a previsão de dados com tendência. Esse método envolve uma equação de previsão (Equação 4) e duas equações de suavização - uma para o nível (Equação 5) e outra para a tendência (Equação 6):

$$\hat{Y}_{t+h|t} = L_t + hB_t, \quad \forall h > 0 \quad (4)$$

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + B_{t-1}), \quad 0 < \alpha < 1 \quad (5)$$

$$B_t = \beta^*(L_t - L_{t-1}) + (1 - \beta^*)B_{t-1}, \quad 0 < \beta^* < 1 \quad (6)$$

onde L_t denota uma estimativa do nível da série no momento t , B_t denota uma estimativa da tendência (inclinação) da série no momento t , α é o parâmetro de suavização para o nível, β^* é o parâmetro de suavização da tendência. Convém destacar que $\beta^* = \frac{\beta}{\alpha}$. (HYNDMAN; ATHANASOPOULOS, 2018)

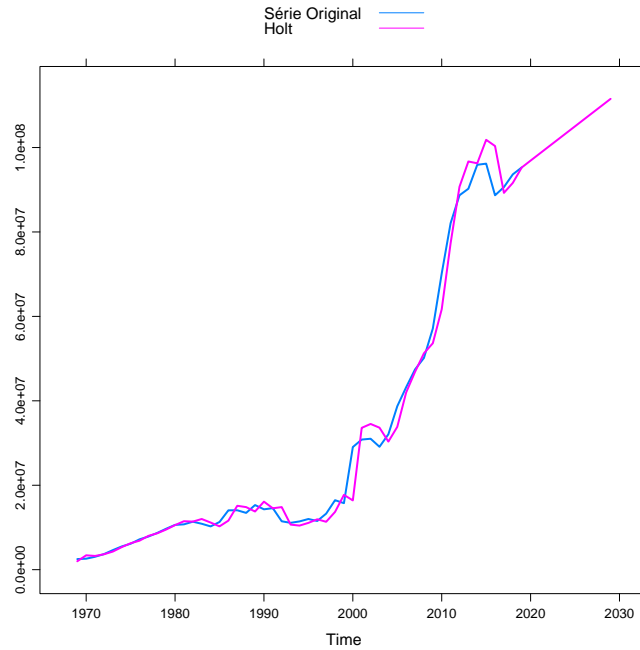
Os dados para a quantidade de passageiros embarcados em vôos domésticos no Brasil são disponibilizados de forma anual, onde a primeira observação é 1969 e a última de 2019.

Os valores dos parâmetros de suavização estimados podem ser observados abaixo. Para facilitar a compreensão e a disponibilização de cálculos introdutórios, está disponível no Apêndice uma tabela com os valores dos Estados Iniciais utilizados (Tabela 5), além disso também se encontra os critérios de informação do Modelo de Suavização Exponencial de *Holt* (Tabela 6).

$$\alpha = 0,9999 \quad \beta = 0,3082$$

Na Tabela 2 é possível visualizar os valores estimados para o número de passageiros embarcados em vôos domésticos no Brasil para os próximos 5 anos ($h = 5$). E a Figura 4 demonstra a adequação da SEH a série, assim como plota os valores estimados.

Figura 4 – Suavização Exponencial de Holt (SEH)



Fonte: Elaboração própria a partir dos dados do Ipeadata.

Tabela 2 – Previsão para a quantidade de passageiros embarcados através do Modelo de Holt

Ano	Valor Previsto	LI.95	LS.95
2020	96.905.230	89.872.733	103.937.727
2021	98.530.160	86.950.729	110.109.591
2022	100.155.091	83.928.947	116.381.234
2023	101.780.021	80.650.285	122.909.757
2024	103.404.951	77.079.802	129.730.100
2025	105.029.882	73.212.793	136.846.970

Fonte: Elaboração própria a partir dos dados do Ipeadata.

2.3 Modelo para séries sazonais

2.3.1 Suavização Exponencial de Holt-Winters (HW)

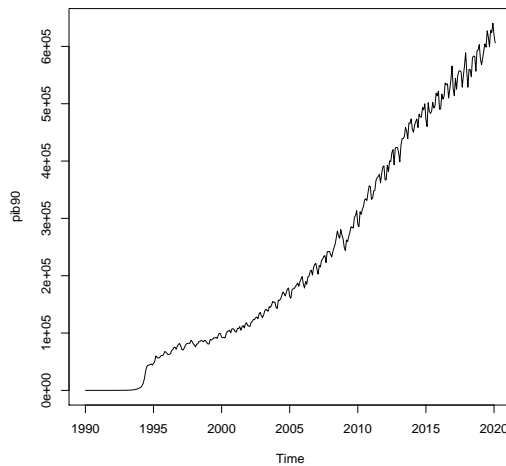
Holt (1957), Winters (1960) estenderam o método de Holt para capturar a sazonalidade. O método sazonal de *Holt-Winters* compreende a equação de previsão (Equação 7) e três equações de suavização - uma para o nível l_t (Equação 8), uma para a tendência b_t (Equação 9) e outra para o componente sazonal s_t (Equação 10), com os correspondentes parâmetros de suavização α , β^* e γ . Além disso, m serve para indicar a frequência da sazonalidade, ou seja, o número de estações do ano. Por exemplo, para dados trimestrais $m = 4$ e para dados mensais $m = 12$.

Existem dois tipos de procedimentos cuja utilização depende das características da série considerada. Basicamente, sendo a série sazonal e com tendência, esta sazonalidade poderá ser multiplicativa ou aditiva. (BUENO, 2011)

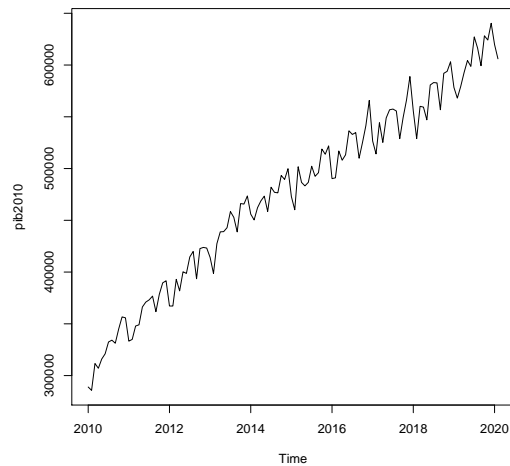
Optamos pela utilização da série temporal do Produto Interno Bruto, visando demonstrar a adequação do modelo HW com sazonalidade aditiva. Esta série apresenta tendência positiva por estar em seu formato nominal e além disso imagina-se possuir uma sazonalidade aditiva devido os ciclos econômicos (Ver Figura 5). Tais motivos justificam a escolha feita.

Figura 5 – Produto Interno Bruto do Brasil (PIB) - em R\$ milhões.

(a) Jan/1990 - Fev/2020



(b) Jan/2010 - Fev/2020



Fonte: Elaboração própria a partir dos dados do Ipeadata.

Os dados para o PIB brasileiro são disposto de forma mensal, sendo a primeira observação de janeiro de 1990, enquanto a última é de fevereiro de 2020, o que totaliza 362 observações. Por uma questão de demonstração da metodologia, preferiu-se utilizar os dados a partir de 2010, por apresentarem uma maior tendência, como pode ser visualizado na Figura 5b.

Pelo Método de *Holt-Winters* com sazonalidade aditiva as observações são estimadas conforme as equações abaixo:

$$\hat{Y}_{t+h|t} = L_t + hB_t + S_{t+h-m(k+1)}, \quad \forall h > 0 \quad (7)$$

$$L_t = \alpha(Y_t - S_{t-m}) + (1 - \alpha)(L_{t-1} + B_{t-1}), \quad 0 < \alpha < 1 \quad (8)$$

$$B_t = \beta^*(L_t - L_{t-1}) + (1 - \beta^*)B_{t-1}, \quad 0 < \beta^* < 1 \quad (9)$$

$$S_t = \gamma(Y_t - L_{t-1} - B_{t-1}) + (1 - \gamma)S_{t-m}, \quad 0 < \gamma < 1 \quad (10)$$

onde k é a parte inteira de $(h - 1)/m$, o que garante que as estimativas dos índices sazonais usados para previsão sejam provenientes do ano final da amostra. A equação do nível mostra uma média ponderada entre a observação com ajuste sazonal ($Y_t - S_{t-m}$) e a previsão não sazonal ($L_{t-1} + B_{t-1}$) para o tempo t . A equação da tendência é idêntica ao método linear de Holt. A equação sazonal mostra uma média ponderada entre o índice sazonal atual ($Y_t - L_{t-1} - B_{t-1}$) e o índice sazonal da mesma estação(ou mês) no ano passado (ou seja, m períodos atrás)¹, além disso destaca-se que $\beta^* = \frac{\beta}{\alpha}$. (HYNDMAN; ATHANASOPOULOS, 2018)

Fazendo uso de alguns pacotes amplamente utilizados no R, foram estimados $h = 10$ períodos para a série analisada.² Os parâmetros de suavização estimados, foram:

$$\alpha = 0,2761 \quad \beta = 0,0109 \quad \gamma = 0,0001$$

No apêndice é possível visualizar os estados iniciais utilizados na previsão do modelo HW (ver Tabela 7), assim como os critérios de informação obtidos (ver Tabela 8). Na Tabela 3 abaixo é possível ver os valores estimados para o PIB nos próximos períodos.

Tabela 3 – Previsão para o PIB através do Modelo de HW

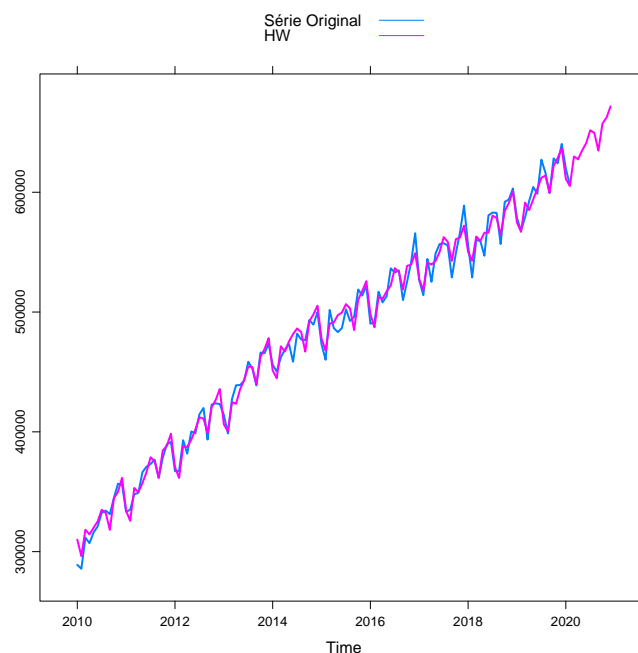
Data	Valor Previsto	LI(95%)	LS(95%)
Mar 2020	629.878,17	613.337,67	646.418,67
Apr 2020	627.469,26	610.261,08	644.677,44
May 2020	634.811,36	616.911,62	652.711,10
Jun 2020	640.960,39	622.346,12	659.574,67
Jul 2020	651.738,41	632.387,50	671.089,32

Fonte: Elaboração própria a partir dos dados do Ipeadata.

¹ Convém salientar que foi usado $m = 12$ para a previsão do modelo HW, que já é o default da função no R.

² Para o intuito principal dessa análise foi utilizado a biblioteca *Forecast*.

Figura 6 – Modelo de Suavização Holt-Winters



Fonte: Elaboração própria a partir dos dados do Ipeadata.

Referências

- BUENO, R. L. S. *Econometria de Séries Temporais*. 2^a. ed. São Paulo: Cengage Learning, 2011. 94-100 p.
- HOLT, C. Forecasting seasonals and trends by exponentially weighted averages. (O.N.R. MEMORANDUM NO.52). *Carnegie Institute of Technology*, 1957.
- HYNDMAN, R. J.; ATHANASOPOULOS, G. *Forecasting: principles and practice*. 2018. Disponível em: <<https://otexts.com/fpp2>>. Acesso em: 06 abr. 2020.
- IPEADATA. 2020. Disponível em: <<http://www.ipeadata.gov.br/Default.aspx>>. Acesso em: 06 abr. 2020.
- MORETTIN, P. A.; TOLOI, C. M. C. *Análise de Series Temporais*. 2^a edição ver. ampl. ed. São Paulo: Blucher, 2006. 89-100 p.
- WINTERS, P. R. Forecasting sales by exponentially weighted moving averages. *Management Science, INFORMS*, v. 6, n. 3, p. 324–342, 1960.

Apêndice

Tabela 4 – Critérios de Informação da SES

Critérios de Informação	
LogLik	-865,70
AIC	1.737,41
BIC	1.745,10
AICC	1.737,67
MSE	708.673,31
AMSE	1.438.242,33

Fonte: Elaboração própria a partir dos dados do Ipeadata.

Tabela 5 – Estados Iniciais utilizados na Previsão da SEH

Estados Iniciais	
l	1.256.274
b	735.688

Fonte: Elaboração própria a partir dos dados do Ipeadata.

Tabela 6 – Critérios de Informação da SEH

Critérios de Informação	
LogLik	-867,93
AIC	1.745,86
BIC	1.755,52
AICC	1.747,19
MSE	11.864.532.803.131,08
AMSE	41.588.376.339.972,26

Fonte: Elaboração própria a partir dos dados do Ipeadata.

Tabela 7 – Estados Iniciais utilizados na Previsão do Modelo HW

Estados Iniciais utilizados	
l	318.343
b	3.373
s1	17.244
s2	10.488
s3	8.076
s4	-11.775
s5	5.630
s6	10.440
s7	2.299
s8	-1.214
s9	-5.920
s10	-875
s11	-22.716
s12	-11.677

Fonte: Elaboração própria a partir dos dados do Ipeadata.

Tabela 8 – Critérios de Informação do Modelo de HW

Critérios de Informação	
LogLik	-1.387,43
AIC	2.808,86
BIC	2.856,52
AICC	2.814,74
MSE	61.879.554,20
AMSE	64.410.107,96

Fonte: Elaboração própria a partir dos dados do Ipeadata.