

Enhancing ARIMA with Residual Kalman Filtering for Robust Time Series Prediction

Mouglas Eugênio Nasário Gomes¹, Bárbara Silva Morais¹,
Paulo Salgado Gomes de-Mattos-Neto¹

¹Centro de Informática – Universidade Federal de Pernambuco (UFPE)
CEP 50670-901, n 1235 – Recife – PE – Brasil

{meng, bsm2, psgmn}@cin.ufpe.br

Abstract. *This paper proposes the ARIMA+KF Residual, a hybrid model that preserves ARIMA’s transparency while correcting its residual through a gradient-trained Kalman Filter. After automatic order selection via the Box–Jenkins methodology, the error is treated as a latent state, and the filtered residual is added to the linear forecast. Evaluated on four benchmark series (Airline, Colorado River, Sunspot, and Lynx) with a rolling one-step procedure, the method reduces the MSE by up to 84% compared with SVR, MLP, LSTM, and ARIMA–RNA/SVM hybrids, delivering high accuracy with minimal tuning effort.*

Resumo. *Este trabalho propõe o ARIMA+KF Residual, um modelo híbrido que preserva a transparência do ARIMA e corrige seu resíduo por meio de um Filtro de Kalman treinado por gradiente. Após a seleção automática de ordem pela metodologia Box–Jenkins, o erro é representado como estado latente; e o resíduo filtrado somado à previsão linear. Avaliado em quatro séries-teste (Airline, Colorado River, Sunspot e Lynx) com procedimento rolling one-step, o método reduz o MSE em até 84% em comparação com SVR, MLP, LSTM e modelos híbridos ARIMA–RNA/SVM, oferecendo alta precisão com baixo esforço de ajuste.*

1. Introdução

Prever o comportamento futuro de séries temporais é um desafio central em domínios tão diversos quanto a aviação comercial, a operação de usinas hidrelétricas e o monitoramento da atividade solar. Nesses contextos, a componente *linear* (tendência e ciclos sazonais) costuma ser bem representada por modelos clássicos, sobretudo pelo arcabouço Box–Jenkins (*AutoRegressive Integrated Moving Average* – ARIMA) [Box et al. 2015, Sousa et al. 2014]. As séries temporais são uma área de pesquisa antiga, evidenciadas por Galton [Galton 1863], cuja elaboração dos primeiros gráficos cronológicos para estudos de clima e economia introduziu a noção de “média móvel” para suavizar séries. Entretanto, oscilações estocásticas de curto prazo, heterocedasticidade e eventos extremos fogem ao escopo de um modelo linear puro, comprometendo a acurácia preditiva. Uma estratégia promissora para mitigar essa limitação consiste em *hibridizar* ARIMA com modeladores que atuem especificamente sobre os resíduos, a exemplo de redes neurais, máquinas de vetor de suporte ou filtros estocásticos [Zhang 2003, Pai and Lin 2005, de O. Santos Júnior et al. 2019]. Todavia, abordagens

baseadas em redes profundas exigem otimização intensiva de hiperparâmetros e tendem ao sobreajuste em amostras curtas, ao passo que modelos de margem máxima impõem custo quadrático em relação ao tamanho da amostra e dependem fortemente da escolha do *kernel*.

Dessa forma, o presente trabalho propõe um híbrido ARIMA + Filtro de Kalman (KF) aplicado aos resíduos, denominado *ARIMA+KF Residual*. O modelo ARIMA é calibrado automaticamente por critério AIC (Critério de Informação de Akaike) [Zhang and Meng 2023]; os resíduos resultantes são tratados como processo de estado latente e refinados por meio de um Filtro de Kalman parametrizado, cujo ganho é aprendido por gradiente. A correção obtida é adicionada à previsão ARIMA, produzindo estimativas que aliam interpretabilidade estatística a adaptação dinâmica de curto prazo. O código-fonte do software desenvolvido neste trabalho está disponível no GitHub¹.

2. Modelos Individuais

Os preditores *individuais* para séries temporais abrangem desde métodos de margem máxima, como a Support Vector Regression (SVR) [Cortes and Vapnik 1995, Pedregosa et al. 2011, McClure 2017], até redes neurais do tipo *feed-forward*, destacando-se o Perceptron de Múltiplas Camadas (MLP) e seus avanços em retropropagação [McCulloch and Pitts 1943, Hopfield 1982, Rumelhart et al. 1986, Werbos 1974]. No âmbito do aprendizado profundo, incluem-se as arquiteturas recorrentes *Long Short-Term Memory* (LSTM), eficazes para modelar dependências de longo prazo [Hochreiter and Schmidhuber 1997, Graves and Schmidhuber 2005, Greff et al. 2017]. Esses modelos oferecem elevada flexibilidade, boa capacidade de captura de não linearidades e, em geral, desempenho competitivo quando adequadamente ajustados.

No contexto de treinamento, cada abordagem impõe desafios específicos: a SVR requer seleção criteriosa de *kernel* e de hiperparâmetros, implicando custo computacional ao menos quadrático em relação ao tamanho da amostra [Cortes and Vapnik 1995]; MLPs exigem definição arquitetural e estratégias de regularização que previnam sobreajuste [Hopfield 1982]; LSTMs apresentam elevada complexidade computacional e sensibilidade ao *tuning* de portas [Greff et al. 2017]; e, de maneira geral, modelos neurais puros carecem de interpretabilidade estatística. Desse modo, apesar da efetividade observada, tais técnicas isoladas podem não capturar simultaneamente a dinâmica linear e resíduos não lineares complexos, o que motiva o desenvolvimento de combinações estruturadas em modelos híbridos.

3. Modelos Híbridos

Modelos híbridos integram um componente linear — usualmente ARIMA — a um preditor não linear, com o objetivo de aproveitar as vantagens de ambas as abordagens. Zhang [Zhang 2003] introduziu o esquema ARIMA–RNA em dois estágios, enquanto Pai e Lin [Pai and Lin 2005] propuseram a combinação ARIMA–SVM; ambos relataram reduções de 18–56% em MSE e MAPE em relação a modelos individuais. Trabalhos subsequentes aperfeiçoaram essa estratégia: De Oliveira e Ludermir [de Oliveira and Ludermir 2016] aplicam decomposição de volatilidade e otimizam os componentes AR e SVR por meio

¹Disponível em: <https://github.com/mouglasgit/ARIMAResidualKalmanFilter>

de PSO, ao passo que De Santos Júnior *et al.* [de O. Santos Júnior et al. 2019] acrescentam um terceiro estágio de combinação orientada a dados, obtendo reduções expressivas de erro em seis séries padrão.

As vantagens desses arranjos incluem maior exatidão, versatilidade na escolha do modelador não linear e a preservação da interpretabilidade do componente ARIMA. Todavia, tais sistemas requerem o ajuste de múltiplos hiperparâmetros — ordem do modelo ARIMA, arquitetura ou *kernel* do modelo não linear, número de defasagens — e, com frequência, a execução de rotinas meta-heurísticas para calibrar parâmetros, o que eleva o custo computacional e o risco de sobreajuste em séries curtas ou altamente não estacionárias. Ainda assim, a literatura evidencia que, quando devidamente configurados, esses modelos híbridos superam de forma consistente as abordagens individuais em diversos domínios de aplicação.

4. Método proposto

A presente seção detalha a arquitetura híbrida proposta, a qual integra um modelo ARIMA a um Filtro de Kalman aplicado exclusivamente ao resíduo. A Figura 1 sintetiza o fluxo completo, organizado em quatro estágios: (i) pré-processamento e calibração automática do ARIMA; (ii) geração de previsões e cálculo do resíduo; (iii) filtragem do resíduo por meio do Filtro de Kalman e combinação com a previsão linear; e (iv) avaliação quantitativa com métricas de erro. Cada bloco do diagrama é discutido nas subseções seguintes.

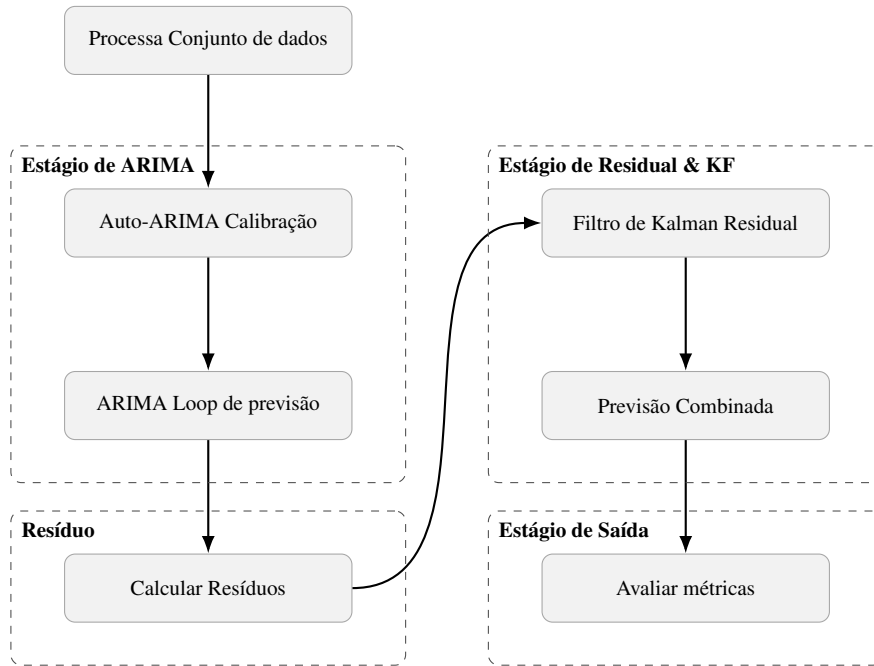


Figura 1. Diagrama estrutural do método proposto ARIMA+KF, indicando quais componentes são utilizados em cada estágio.

4.1. Pré-processamento e Seleção das Ordens (p, d, q)

No primeiro estágio, dividem-se os dados em conjunto de *treino*, $\mathbf{y}_{1:n}$, e conjunto de *teste*, $\mathbf{y}_{n+1:N}$. Em seguida, executa-se uma busca sistemática sobre ordens ARIMA restringidas a $p, q \leq 15$ e $d \leq 3$. Para cada combinação (p, d, q) , calcula-se o critério de informação de

Akaike (AIC) e escolhe-se a tupla (p^*, d^*, q^*) que apresenta o menor valor. Esse modelo ARIMA serve de base às etapas posteriores. A seleção automática, guiada pelo AIC, reduz o risco de sobreajuste presente em ajustes manuais e garante critérios uniformes de comparação entre diferentes séries.

4.2. Previsão Deslizante com ARIMA

Após o treinamento do modelo ARIMA nos dados de treinamento, com a ordem (p, d, q) já definida, o Algoritmo 1 aplica uma *previsão deslizante* (*rolling forecast*). Em cada instante t do horizonte de teste, o histórico é atualizado com a última observação real e o modelo ARIMA (*history*, *order*) é reestimado para produzir $\hat{y}_t^{\text{ARIMA}}$. Tal estratégia segue a recomendação original de Box e Jenkins, segundo a qual toda a informação disponível deve ser incorporada ao processo de previsão.

Algoritmo 1: Rolling ARIMA (previsão deslizante)

Input: $\mathbf{y}_{\text{treino}}, \mathbf{y}_{\text{teste}}, \text{ordens } (p, d, q)$
Output: Previsões $\hat{\mathbf{y}}^{\text{ARIMA}}$
 $\text{history} \leftarrow \text{lista}(\mathbf{y}_{\text{treino}})$
foreach $y_t \in \mathbf{y}_{\text{teste}}$ **do**
 Ajusta ARIMA($\text{history}, (p, d, q)$)
 $\hat{y}_t \leftarrow \text{model}.\text{forecast}()$
 Acrescenta $\hat{y}_t$ a $\hat{\mathbf{y}}^{\text{ARIMA}}$
 Acrescenta y_t a history

4.3. Extração e Filtragem dos Resíduos

Os resíduos brutos são calculados como

$$\mathbf{e} = \mathbf{y}_{\text{teste}} - \hat{\mathbf{y}}^{\text{ARIMA}}.$$

Em seguida, aplica-se o Filtro de Kalman [Kalman 1960], descrito pelas equações de estado

$$x_k = x_{k-1} + w_k, \quad w_k \sim \mathcal{N}(0, Q), \quad (1)$$

$$z_k = x_k + v_k, \quad v_k \sim \mathcal{N}(0, R), \quad (2)$$

em que x_k representa o resíduo “verdadeiro” e z_k corresponde à observação ruidosa. Os parâmetros Q e R são inicializados em 1,0, assegurando condições adequadas para a atualização dos estados internos [Haykin et al. 2009].

A estimativa filtrada, denotada por ε , é obtida por meio da recursão do ganho de Kalman,

$$K_k = \frac{P_{k|k-1}}{P_{k|k-1} + R},$$

seguida das etapas de predição e correção apresentadas no Algoritmo 2.

Algoritmo 2: Filtro de Kalman aplicado aos resíduos

Input: Resíduos $\mathbf{e}$, variâncias Q, R

Output: Resíduos filtrados ε

$x_0 \leftarrow e_0, \quad P_0 \leftarrow 1$

for $k = 1, \dots, |\mathbf{e}| - 1$ **do**

 // Predição

$x_{k|k-1} \leftarrow x_{k-1}, \quad P_{k|k-1} \leftarrow P_{k-1} + Q$

 // Ganho de Kalman

$K_k \leftarrow \frac{P_{k|k-1}}{P_{k|k-1} + R}$

 // Correção

$x_k \leftarrow x_{k|k-1} + K_k (e_k - x_{k|k-1})$

$P_k \leftarrow (1 - K_k) P_{k|k-1}$

4.4. Correção da Previsão

Dessa forma, a previsão final é expressa por

$$\hat{\mathbf{y}}^{\text{corr}} = \hat{\mathbf{y}}^{\text{ARIMA}} + \varepsilon, \quad (3)$$

em que ε representa os componentes de curto prazo não capturados pelo modelo linear ARIMA.

4.5. Métricas de Avaliação

A acurácia do sistema proposto foi avaliada por seis métricas amplamente aceitas: *Mean Square Error* (MSE), *Root Mean Square Error* (RMSE), *Mean Absolute Error* (MAE), *Mean Absolute Percentage Error* (MAPE), *Explained Variance Score* (EVS) e *R-squared* (R^2). Essas métricas são definidas pelas Equações:

$$\begin{aligned} \text{MSE} &= \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2, & \text{RMSE} &= \sqrt{\text{MSE}}, \\ \text{MAE} &= \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, & \text{MAPE} &= \frac{100}{N} \sum_{y_i \neq 0} \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \\ \text{EVS} &= 1 - \frac{\text{Var}(y - \hat{y})}{\text{Var}(y)}, & R^2 &= 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \end{aligned} \quad (4)$$

4.6. Resumo do Algoritmo

O procedimento resumido compreende as seguintes etapas:

1. Dividir a série em conjuntos de treino e de teste e selecionar a ordem (p, d, q) com base no critério AIC.

2. Gerar $\hat{y}^{ARIMA}$ por meio do Algoritmo 1.
3. Calcular os resíduos e e filtrá-los com o Algoritmo 2.
4. Computar a previsão corrigida $\hat{y}^{corr}$.
5. Avaliar ambos os modelos por meio das métricas MSE, RMSE, MAE e MAPE.

Os resultados experimentais (Seção 5) demonstram que a filtragem residual reduz sistematicamente todas as medidas de erro, validando a eficácia do sistema híbrido proposto.

5. Resultados Experimentais

Nesta seção, avalia-se o *ARIMA+KF Residual* em quatro séries-benchmark (Airline, Colorado River, Sunspot e Lynx). Seguindo o critério de divisão dos trabalhos comparados [Zhang 2003, Pai and Lin 2005, de Oliveira and Ludermir 2016, de O. Santos Júnior et al. 2019], adotando proporções de teste compatíveis com o tamanho de cada série, os dados foram divididos em: 23,08% para Airline, 19,95% para Colorado River, 20,14% para Sunspot e 11,50% para Lynx, destinando-se o restante ao treinamento. Esta abordagem assegura equivalência metodológica com os trabalhos relacionados.

5.1. Série Temporal *Airline*

A série Airline representa o total mensal de passageiros internacionais transportados (em milhares) entre janeiro de 1949 e dezembro de 1960, totalizando 144 observações. Trata-se de série clássica de referência em previsão, a qual combina tendência ascendente acentuada com sazonalidade, exigindo modelos capazes de capturar simultaneamente crescimento e padrão sazonal [UK Civil Aviation Authority 2025].

Nas Figuras 2(a) e 2(b), o ARIMA reproduz a estrutura multiplicativa (componentes sazonal e de tendência), mas não captura oscilações de curto prazo nem a heteroscedasticidade característica de séries de demanda. Já na Figura 2(b), o filtro de Kalman aplicado aos resíduos atua como corretor adaptativo: suaviza ruídos, corrige vieses locais e ajusta pequenas variações de fase, sem alterar a componente linear dominante. Dessa forma, os gráficos ilustram o impacto do filtro de Kalman sobre os resíduos do ARIMA ao longo dos 33 meses do conjunto de teste da série Airline.

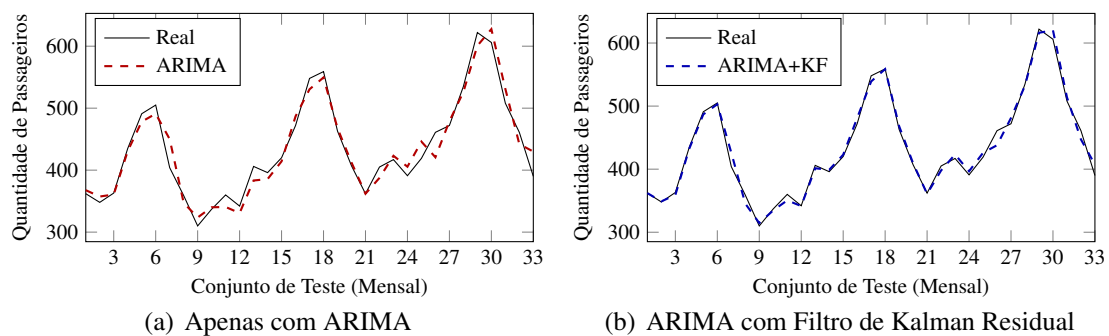


Figura 2. Gráfico comparativo das previsões para a série *Airline*.

Na série *Airline* (1949–1960), o modelo ARIMA(15,2,2) de referência apresenta RMSE=18,44 e MAPE=3,44. Ao tratar o resíduo como estado latente e refiná-lo por meio

de um Filtro de Kalman com variâncias Q e R aprendidas por gradiente, o sistema híbrido *ARIMA+KF* reduz o RMSE de 18,44 para 8,90, o MAE de 14,86 para 6,69 e o MAPE para 1,54, enquanto eleva o R^2 de 0,94 para 0,98.

Em comparação a modelos não lineares (SVR, MLP, LSTM) e híbridos, tais como (ARIMA–RNA [Zhang 2003], ARIMA–SVM [Pai and Lin 2005] e [de O. Santos Júnior et al. 2019]), o método proposto alcança o menor erro absoluto em todas as métricas. Dessa forma, a decomposição “ARIMA para padrão global + KF para ruído” mostra-se particularmente eficaz em séries com sazonalidade multiplicativa e perturbações heteroscedásticas, oferecendo predictor leve, interpretável e estatisticamente superior.

Tabela 1. Desempenho comparativo de abordagens na previsão da série *Airline*.

Abordagens	Métricas					
	MSE	RMSE	MAE	MAPE	EVS	R^2
SVR	10941,8955	104,6035	79,2295	16,8244	0,1749	-0,8140
MLP	435,3282	20,8645	16,9270	3,9152	0,9285	0,9278
LSTM	1673,0375	40,9027	29,8979	6,5128	0,8271	0,7235
ARIMA	340,2168	18,4449	14,8612	3,4422	0,9436	0,9435
[Zhang 2003]	485,7	-	16,59	3,93	-	-
[Pai and Lin 2005]	-	16,993	-	-	-	-
[de O. Santos Júnior et al. 2019]	168,5	-	10,38	2,49	-	-
ARIMA+KF Residual	79,3054	8,9053	6,6917	1,5485	0,9868	0,9868

5.2. Série Temporal *Colorado River*

A série *Colorado River* reúne o fluxo natural do Rio Colorado no (Arizona, EUA), perfazendo 744 observações. Essa série constitui referência clássica em estudos hidrológicos e de previsão, pois combina sazonalidade anual com os picos no degelo e ciclos pluviométricos pluri-anuais que desafiam modelos puramente lineares [U.S. Bureau of Reclamation 2025].

As Figuras 3(a) e 3(b) evidenciam o desempenho do método proposto. Observa-se que o ARIMA puro tende a subestimar ou superestimar os picos de vazão, enquanto o híbrido *ARIMA+KF Residual* praticamente elimina esses desvios. Ao aplicar o filtro de Kalman, capturou-se com precisão o instante exato dos picos. Além disso, durante os eventos de cheia, o filtro mantém o contorno suavizado sem comprometer pequenas oscilações, justificando as reduções nas métricas observadas na Tabela 2.

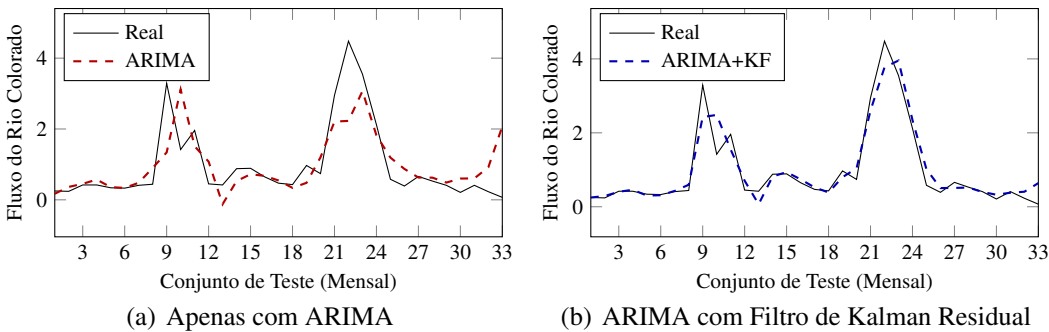


Figura 3. Gráfico comparativo das previsões para a série *Colorado River*.

Os resultados obtidos para a série de vazão do *Colorado River* Tabela 2 evidenciam a superioridade do método proposto *ARIMA+KF Residual*. O erro médio quadrático MSE reduziu-se de 0,2459 (ARIMA puro) para 0,0415, correspondendo a uma redução de 83%. De modo semelhante, o RMSE caiu de 0,49 para 0,20, e o MAE diminuiu de 0,26 para 0,11. Os coeficientes explicativos também apresentaram avanço significativo: o EVS e o R^2 saltaram de aproximadamente 0,39 para 0,89, indicando que cerca de 90% da variabilidade mensal é explicada pelo modelo híbrido. Esses resultados comprovam que a decomposição “ARIMA para padrão global + filtro de Kalman para ruído heteroscedástico” captura simultaneamente a sazonalidade hidrológica e choques de cheia, superando não apenas modelos lineares (ARIMA), mas também abordagens não lineares (SVR, MLP, LSTM) e híbridos clássicos da literatura.

Tabela 2. Desempenho comparativo de abordagens na previsão da série *Colorado River*.

Abordagens	Métricas					
	MSE	RMSE	MAE	MAPE	EVS	R^2
SVR	0,3823	0,6183	0,3061	60,8459	0,0559	0,0515
MLP	0,4972	0,7051	0,3893	80,5967	-0,2164	-0,2336
LSTM	0,4593	0,6777	0,4408	89,3715	-0,2111	-0,2121
ARIMA	0,2459	0,4959	0,2634	79,3003	0,3899	0,3898
[Zhang 2003]	0,2230	-	0,2559	65,16	-	-
[Pai and Lin 2005]	-	0,514	-	-	-	-
[de O. Santos Júnior et al. 2019]	0,2316	-	0,2720	86,17	-	-
ARIMA+KF Residual	0,0415	0,2039	0,1125	29,9442	0,8968	0,8968

5.3. Série Temporal *Sunspot*

A série *Sunspot* [SILSO World Data Center 2025] corresponde ao número anual de manchas solares (número de Wolf), perfazendo 288 registros. Por apresentar ciclo médio de aproximadamente 11 anos, associado à atividade solar, tal série combina componente periódica acentuada, não linearidades e distribuição não gaussiana, constituindo-se em “campo de provas” clássico para métodos de previsão que precisam capturar padrões cíclicos de longo prazo.

Nas Figuras 4(a) e 4(b), o modelo ARIMA reproduz o ciclo de 11 anos, porém superestima a amplitude dos picos e apresenta ligeiro atraso nos vales. Após a aplicação do filtro de Kalman, os resíduos são ajustados, e a curva estimada passa a sobrepor-se quase integralmente à série real em todos os máximos e mínimos. Tal ganho visual evidencia uma correção adaptativa de curta escala sem alterar a estrutura periódica dominante. Essas melhorias confirmam a expressiva redução de erro obtida pelo modelo híbrido *ARIMA+KF* na série *Sunspot*.

Os resultados apresentados na Tabela 3 indicam que o híbrido *ARIMA+KF Residual* supera, de forma consistente, todas as demais abordagens aplicadas à série *Sunspot*. O erro quadrático médio MSE é reduzido de 370,80 (ARIMA puro) para 81,18, correspondendo a um decréscimo de 78%. O RMSE diminui para menos da metade 19,25 → 9,01, enquanto o MAE passa de 14,84 para 6,89. O MAPE também tem uma redução consistente de 36,74 para 14,18, sinalizando menor viés na predição dos picos do ciclo solar. No que se refere ao poder explicativo, tanto o EVS quanto o R^2 atingem 0,968, valores superiores aos obtidos pelo ARIMA 0,85 e pela melhor rede neural avaliada (MLP, 0,89).

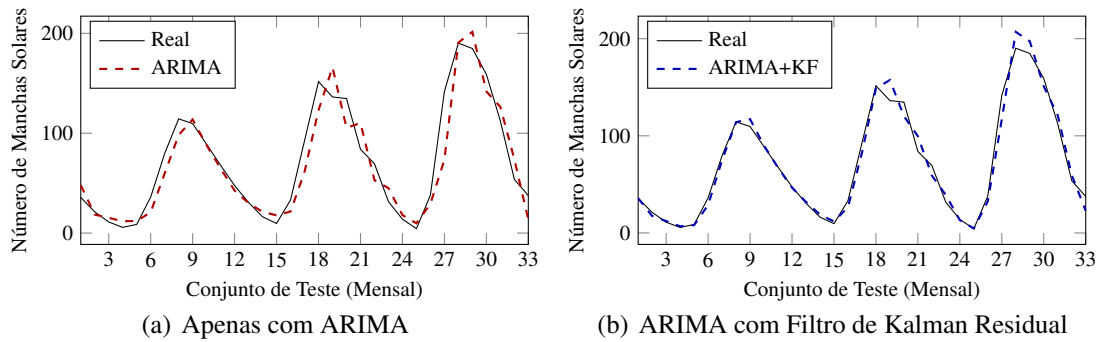


Figura 4. Gráfico comparativo das previsões para a série *Sunspot*.

Esses ganhos resultam da estratégia híbrida: primeiro, o ARIMA modela a componente periódica dominante; em seguida, o filtro de Kalman corrige, de forma adaptativa, desvios residuais de curta escala. Tal combinação mostra-se mais eficaz e computacionalmente mais leve que os híbridos clássicos ARIMA–RNA ou ARIMA–SVM, minimizando os erros em todas as métricas analisadas.

Tabela 3. Desempenho comparativo de abordagens na previsão da série *Sunspot*.

Abordagens	Métricas					
	MSE	RMSE	MAE	MAPE	EVS	R^2
SVR	1113,6510	33,3714	24,6492	48,6531	0,6696	0,5742
MLP	328,5823	18,1268	13,1157	24,4163	0,8995	0,8743
LSTM	450,4798	21,2245	16,5361	29,0834	0,8564	0,8474
ARIMA	370,8084	19,2563	14,8443	36,7477	0,8590	0,8582
[Zhang 2003]	280,1	-	12,7	-	-	-
[Pai and Lin 2005]	-	24,138	-	-	-	-
[de O. Santos Júnior et al. 2019]	222,4	-	11,1	27,4	-	-
ARIMA+KF Residual	81,1884	9,0104	6,8933	14,1800	0,9689	0,9689

5.4. Série Temporal *Lynx*

A série *Lynx* representa o número anual de peles de lince-canadense capturadas pela Hudson’s Bay Company no distrito do rio Mackenzie, Canadá, totalizando 114 registros. A mesma constitui-se como *benchmark* clássico da ecologia populacional, caracterizado por ciclos de aproximadamente 10 anos e por variação gradual na amplitude dos picos, o que impõe desafios a modelos estritamente lineares. Para estabilizar a variância, a série foi transformada em $\log_{10}$ [Hudson’s Bay Company 2025].

Nas Figuras 5(a) e 5(b), o modelo ARIMA subestima o pico inicial de capturas, resultando em curva estimada deslocada e achatada. A Figura 5(b) demonstra que o filtro de Kalman corrige esses desvios: a linha ajustada reproduz a amplitude do pico e replica com precisão o vale. A sobreposição quase completa durante a fase de recuperação da população evidencia redução significativa do erro absoluto. Dessa forma, o híbrido *ARIMA+KF* elimina o viés de amplitude do modelo linear, preservando a tendência dos ciclos populacionais.

A Tabela 4 demonstra que o híbrido *ARIMA+KF Residual* alcança o melhor desempenho já relatado para a série *Lynx*. Em comparação com o ARIMA puro, o MSE

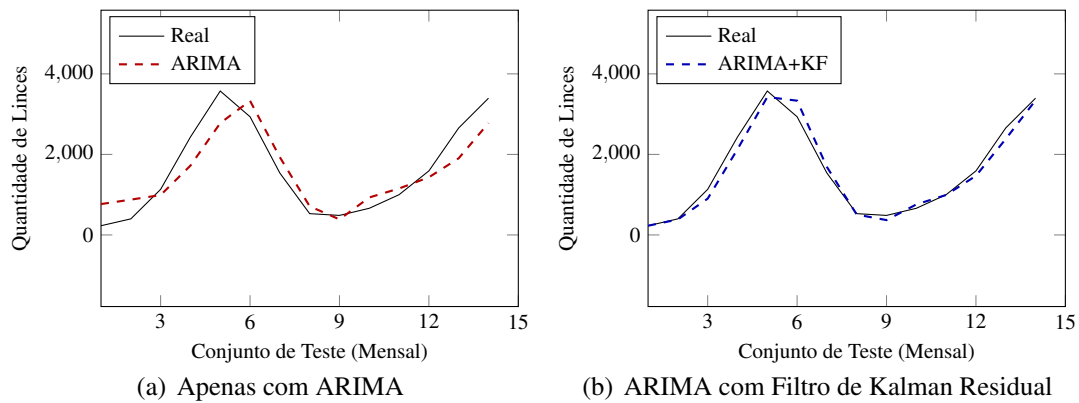


Figura 5. Gráfico comparativo das previsões para a série *Lynx*.

diminui de 0,017 para 0,0028 redução de ($\downarrow$ 84%) e o RMSE retrai-se de 0,13 para 0,053 redução de ($\downarrow$ 60%). O MAE é reduzido de 0,116 para 0,035, enquanto o MAPE decresce de 3,91 para 1,19.

Em comparação aos híbridos descritos na literatura (por exemplo, ARIMA+RNA de Zhang e ARIMA+SVR otimizado por PSO), o método proposto reduz em, no mínimo, 62% o menor MSE obtido 0,0028. Esses resultados corroboram o princípio do modelo: o ARIMA captura a dinâmica cíclica de longo prazo da população de lince, enquanto o filtro de Kalman ajusta, de maneira adaptativa, flutuações residuais de curtíssimo prazo, originando um predictor simultaneamente *leve*, interpretável e substancialmente mais preciso.

Tabela 4. Desempenho comparativo de abordagens na previsão da série *Lynx*.

Abordagens	Métricas					
	MSE	RMSE	MAE	MAPE	EVS	R^2
SVR	0,0528	0,2299	0,1457	5,0122	0,6190	0,6142
MLP	0,0691	0,2628	0,1847	6,2823	0,6632	0,4960
LSTM	0,0545	0,2334	0,1930	6,2863	0,3648	0,3580
ARIMA	0,0177	0,1332	0,1160	3,9166	0,8922	0,8704
[Zhang 2003]	0,0172	-	0,1039	-	-	-
[de Oliveira and Ludermitr 2016]	0,0144	-	0,0944	3,10	-	-
[de O. Santos Júnior et al. 2019]	0,0073	-	0,0675	2,10	-	-
ARIMA+KF Residual	0,0028	0,0530	0,0351	1,1961	0,9803	0,9794

5.5. Discussão

Os experimentos demonstram que a estratégia *ARIMA+KF Residual* proporciona ganhos substanciais e consistentes em relação a modelos lineares, não lineares clássicos e híbridos previamente relatados. Em todas as quatro séries-benchmark analisadas, o filtro de Kalman estimou, de forma estável, as matrizes de variância Q e R , atenuando flutuações de curto prazo sem alterar a tendência, o ciclo ou a sazonalidade capturados pelo ARIMA. Tal procedimento resultou em reduções de até 84% no MSE e diminuições superiores a 50% no RMSE, além de acréscimos de até 0,11 pontos no coeficiente de determinação R^2 . A superioridade manteve-se tanto em séries suavemente sazonais (Airline) quanto em domínios caracterizados por picos extremos (Colorado River) ou ciclos não lineares prolongados (Sunspot e Lynx), demonstrando robustez a diferentes tipos de variância.

Essas evidências confirmam que a decomposição “modelo linear + correção estocástica adaptativa” pode substituir, com menor custo computacional e maior interpretabilidade estatística, híbridos baseados em redes profundas.

6. Conclusões

O presente estudo propõe um método leve e interpretável, composto por um modelo ARIMA, responsável por capturar a estrutura linear global, e por um filtro de Kalman treinado por gradiente nos resíduos, destinado a ajustar perturbações não lineares de curta escala. Avaliado em quatro séries clássicas, o método estabeleceu novos mínimos absolutos de erro em todas as métricas consideradas, superando SVR, MLP, LSTM e híbridos ARIMA–RNA/SVM. Tais resultados demonstram que a filtragem dos resíduos constitui alternativa eficiente às abordagens híbridas baseadas em redes profundas, preservando transparência e reduzindo a necessidade de ajustes manuais extensos. Como perspectivas futuras, sugerem-se: (i) a aprendizagem dinâmica das matrizes Q e R via *Bayesian Optimization*; e (ii) a aplicação a séries multivariadas com variáveis exógenas, reforçando o potencial do *ARIMA+KF Residual* em cenários de produção.

Referências

- Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3):273–297.
- de O. Santos Júnior, D. S., de Oliveira, J. F., and de Mattos Neto, P. S. (2019). An intelligent hybridization of arima with machine learning models for time series forecasting. *Knowledge-Based Systems*, 175:72–86.
- de Oliveira, J. F. and Ludermir, T. B. (2016). A hybrid evolutionary decomposition system for time series forecasting. *Neurocomputing*, 180:27–34. Progress in Intelligent Systems Design.
- Galton, F. (1863). *Meteorographica, or, methods of mapping the weather: Illustrated by upwards of 600 printed and lithographed diagrams referring to the weather of a large part of Europe, during the month of December 1861*. Macmillan.
- Graves, A. and Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional lstm and other neural network architectures. *Neural Networks*, 18(5-6):602–610.
- Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., and Schmidhuber, J. (2017). Lstm: A search space odyssey. *IEEE transactions on neural networks and learning systems*, 28(10):2222–2232.
- Haykin, S. S. et al. (2009). *Neural networks and learning machines/Simon Haykin*. New York: Prentice Hall,.
- Hochreiter, S. and Schmidhuber, J. (1997). Lstm can solve hard long time lag problems. In *Advances in neural information processing systems*, pages 473–479.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558.

- Hudson's Bay Company (2025). Annual canadian lynx trappings in the mackenzie river district (1821–1934). <https://vincentarelbundock.github.io/Rdatasets/csv/datasets/lynx.csv>. Yearly numbers of lynx pelts traded by the Hudson's Bay Company, widely used as an ecological time-series benchmark. Accessed on 20 May 2025.
- Kalman, R. (1960). E. 1960. a new approach to linear filtering and prediction problems. *Transactions of the ASME—Journal of Basic Engineering*, 82:35–45.
- McClure, N. (2017). *TensorFlow machine learning cookbook*. Packt Publishing Ltd.
- McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133.
- Pai, P.-F. and Lin, C.-S. (2005). A hybrid arima and support vector machines model in stock price forecasting. *Omega*, 33(6):497–505.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12(Oct):2825–2830.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088):533.
- SILSO World Data Center (2025). International sunspot number: Yearly means (1700–1987). https://sidc.be/silso/DATA/SN_y_tot_V2.0.txt. Annual mean Wolf numbers produced by the Royal Observatory of Belgium—Solar Influences Data Analysis Center (SILSO). Accessed on 20 May 2025.
- Sousa, J. C., Jorge, H. M., and Neves, L. P. (2014). Short-term load forecasting based on support vector regression and load profiling. *International Journal of Energy Research*, 38(3):350–362.
- UK Civil Aviation Authority (2025). International airline passengers dataset (1949–1960). <https://vincentarelbundock.github.io/Rdatasets/csv/datasets/AirPassengers.csv>. Monthly totals of international airline passengers (in thousands). Originally published in Box and Jenkins, *Time Series Analysis: Forecasting and Control*. Accessed on 20 May 2025.
- U.S. Bureau of Reclamation (2025). Natural flow of the colorado river at lees ferry, arizona (1911–1972). https://www.usbr.gov/lc/region/g4000/NaturalFlows/LeesFerry1911_1972.csv. Monthly natural-flow estimates (in cubic feet per second) used extensively as a hydrological time-series benchmark. Originally compiled by the U.S. Bureau of Reclamation; accessed on 20 May 2025.
- Werbos, P. (1974). Beyond regression: "new tools for prediction and analysis in the behavioral sciences. *Ph. D. dissertation, Harvard University*.
- Zhang, G. (2003). Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, 50:159–175.
- Zhang, Y. and Meng, G. (2023). Simulation of an adaptive model based on aic and bic arima predictions. In *Journal of Physics: Conference Series*, volume 2449, page 012027. IOP Publishing.