

WARMA: Window-based Autoregressive Moving Average

Arthur Ronald Garcia · Jorge Soares · Laura Assis · Fabio Porto · Dayse Pastore · Eduardo Ogasawara

Introdução

Previsão de séries temporais é uma tarefa central em diversas áreas, e modelos lineares clássicos como AR, MA, ARMA e ARIMA são amplamente utilizados para isso. Porém, esses modelos dependem de premissas que frequentemente não se sustentam na prática.

Premissa Violada

Esses modelos assumem covariância-estacionariedade — média, variância e covariância constantes ao longo do tempo. Séries econômico-financeiras reais raramente satisfazem essa condição sem transformação prévia.

Transformações Globais

A abordagem mais comum é a diferenciação global (ARIMA). O problema: ela age sobre toda a série de uma vez, mesmo quando a não-estacionariedade é apenas local.

Estrutura Local Oculta

Embora a série completa seja não-estacionária, subsequências locais frequentemente se comportam de forma estável. Modelar essas subsequências preserva a interpretabilidade e evita distorções desnecessárias do sinal original.

Problema e Questão de Pesquisa

Transformações globais como a diferenciação atuam sobre toda a série de uma vez, mesmo quando a não-estacionariedade é impulsionada principalmente por mudanças locais de nível ou escala. Com isso, parte da estrutura original torna-se menos interpretável, dificultando a análise e o diagnóstico.

> **Questão de Pesquisa**
>
> Um esquema de pré-processamento local baseado em janelas consegue produzir representações com ordem de integração zero, preservando a interpretabilidade e mantendo desempenho preditivo competitivo com o ARIMA?

O Modelo WARMA

O Problema Global

A série completa pode ser não estacionária — mas isso não significa que cada subsequência local também o seja. Transformações globais tratam toda a série de forma uniforme, ignorando variações locais.

A Solução Local

Subsequências locais podem ser aproximadamente estáveis. O WARMA usa janelas deslizantes para estimar propriedades locais e, em seguida, ajusta um modelo ARMA após a estacionarização local — uma mudança fundamental na unidade de pré-processamento.

Pipeline WARMA

Sequenciamento

Janela deslizante extrai subsequências locais da série temporal original.

Normalização da Média

Remove a tendência local subtraindo a média de cada janela.

Remoção de Outliers

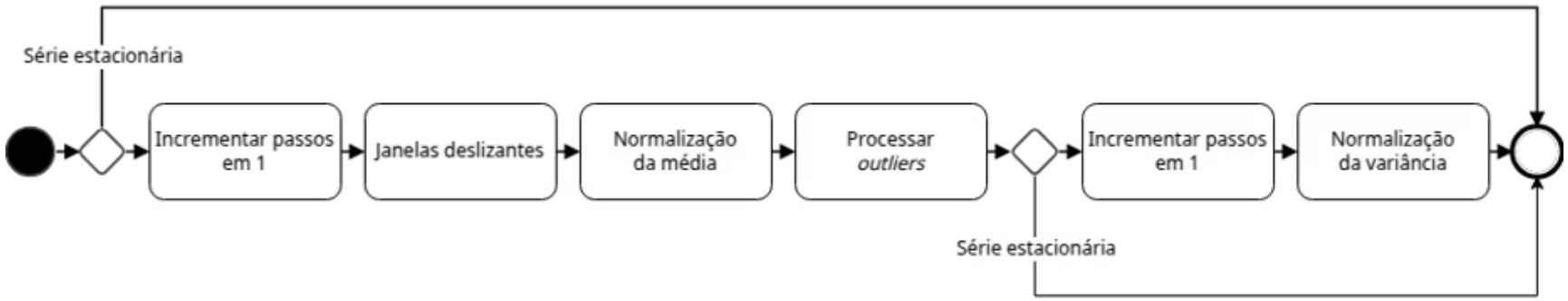
Identifica e trata valores atípicos dentro de cada subsequência local.

Normalização da Variância

Estabiliza a escala dividindo pelo desvio padrão local de cada janela.

Modelagem (ARMA)

Ajuste do modelo ARMA sobre as subsequências localmente estacionarizadas.



Sequenciamento por Janelas Deslizantes

Definição Formal

Dada a série $\mathbf{X = \langle x_1, \ldots, x_n \rangle}$, com restrições $\mathbf{3 \leq w \leq n}$ e $\mathbf{1 < \tau \leq w}$, o tamanho de janela **w** e o deslocamento **τ** definem as subsequências. São geradas **h** janelas alinhadas ao final da série, armazenadas na matriz $\mathbf{A \in \mathbb{R}^{(h \times w)}}$ (uma subsequência por linha).

- $\mathbf{h} = \lfloor (n - w) / \tau \rfloor + 1$
- $\mathbf{o} = (n - w) \bmod \tau$
- $\mathbf{seq}_{i,w,\tau}(X) = \langle x_{1+o+(i-1)\tau}, \ldots, x_{w+o+(i-1)\tau} \rangle$

> O offset **o** alinha o janelamento com o final da série — até **o** observações iniciais podem ser excluídas.

Interpretação dos Parâmetros

w

Define a escala local da janela — a granularidade em que propriedades locais (média e variância) são estimadas.

τ

Controla a densidade de amostragem ao longo do tempo. Valores pequenos de τ aumentam a continuidade entre janelas adjacentes; τ = w produz janelas sem sobreposição.

Passo 1: Normalização da Média

Centralização por Janela

Cada janela é centralizada pela sua média local. Uma etapa sugere predominância de média não constante.

Fórmula: $b_{ij} = a_{ij} - \mu(a_i)$, onde $\mu(v) = (1/|v|) \Sigma_j v_j$

Após essa etapa, janelas discrepantes podem ser opcionalmente descartadas por um critério de **amplitude interquartílica (IQR)**, conforme a técnica de Normalização Adaptativa (Ogasawara et al., 2010). Nos experimentos, essa filtragem permanece inativa para manter a comparação com o ARIMA metodologicamente mais rigorosa.

Conexão Interpretativa

Corrige **mudanças de nível** locais

Trata **tendências determinísticas** locais

Captura também padrões de **passeio aleatório** localmente

Preserva a **interpretabilidade** da correção

Passo 2: Normalização da Variância

Estabilização de Escala

Se a estacionariedade não é atingida após a Etapa 1, a variância é estabilizada. Duas etapas indicam necessidade de tratar a escala local. Linhas com variância zero são mapeadas para zero.

Fórmula: $c_{ij} = b_{ij} / \sigma(b_i)$

Valor Interpretativo

O número de etapas indica qual propriedade local precisou de tratamento:

- **s = 1** → correção de nível
- **s = 2** → correção de nível e escala

O modelo distingue explicitamente correções de nível e escala.

Representação Compacta e Modelagem ARMA

01

Retenção das Últimas Colunas

A série usada para modelagem é **Vec(D1:h, (w−τ+1):w)**, a vetorização da submatriz formada pelas últimas τ colunas de cada linha. Isso evita o reuso excessivo de observações sobrepostas.

02

Modelagem ARMA

A série estacionária resultante y_t é modelada com ARMA: $\boldsymbol{\varphi(B)y_t = \theta(B)\varepsilon_t}$, onde φ(B) e θ(B) são os polinômios autorregressivo e de médias móveis no operador de defasagem. Notação completa: **WARMA(p,(w,s,τ),q)**.

03

Fórmula Compacta

Seja s ∈ {0,1,2} o número de etapas. Para cada observação x_t, a representação estacionária é:

$$\boldsymbol{y_t = (x_t - f\mu(s)\cdot\mu(ag(t))) / (\sigma(ag(t)))f\sigma(s)}$$

onde g(t) indica a janela associada a x_t, fμ(s) = 0 se s=0 e 1 caso contrário, fσ(s) = 1 se s=2 e 0 caso contrário. Isso cobre três casos: **(i)** já estacionária (s=0); **(ii)** normalização de média apenas (s=1); **(iii)** normalização de média e variância (s=2).

s é definido como o menor número de etapas para o qual a série transformada possui ordem de integração zero, segundo a função `ndiffs` (pacote *forecast*, Hyndman & Khandakar, 2008).

Algoritmo de Seleção do Modelo WARMA

A seleção do modelo WARMA segue uma busca exaustiva sobre os hiperparâmetros (w, τ), avaliando para cada combinação a série pré-processada e selecionando o modelo com menor AICc.

```
Input: série temporal X, window.length máximo
Output: modelo ARMA otimizado

AICc.model ← ∞
ARMA.model ← NULL

para w ∈ {3, ..., window.length}, τ ∈ {1, ..., w} faça
   s ← menor valor em {0,1,2} tal que ndiffs(warma.preprocessing(X,w,τ,s)) = 0
   Y ← warma.preprocessing(X, w, τ, s)
   obj ← auto.arima(Y)
   se AICc(obj) < AICc.model então
      AICc.model ← AICc(obj)
      ARMA.model ← ARMA(obj)
   fim
fim
retorne ARMA.model
```

Avaliação Experimental

Séries Sintéticas

Mudanças de nível, tendências determinísticas e passeios aleatórios — para validação controlada da hipótese de etapas. A busca avalia exaustivamente combinações de (w, τ), com w iniciando em 3 e τ variando de 1 até w.

Dados Reais

80 séries mensais econômico-financeiras não estacionárias, cobrindo 7 domínios: juros, dívida líquida do setor público (DLSP), atividade econômica, ações, câmbio, IPCA e commodities. Selecionadas de um conjunto amplo de 92 séries provenientes de B3, Investing.com, Banco Central do Brasil, IBGE e Ipeadata. Séries estacionárias foram excluídas.

Hipótese Conceitual dos Passos: Evidências

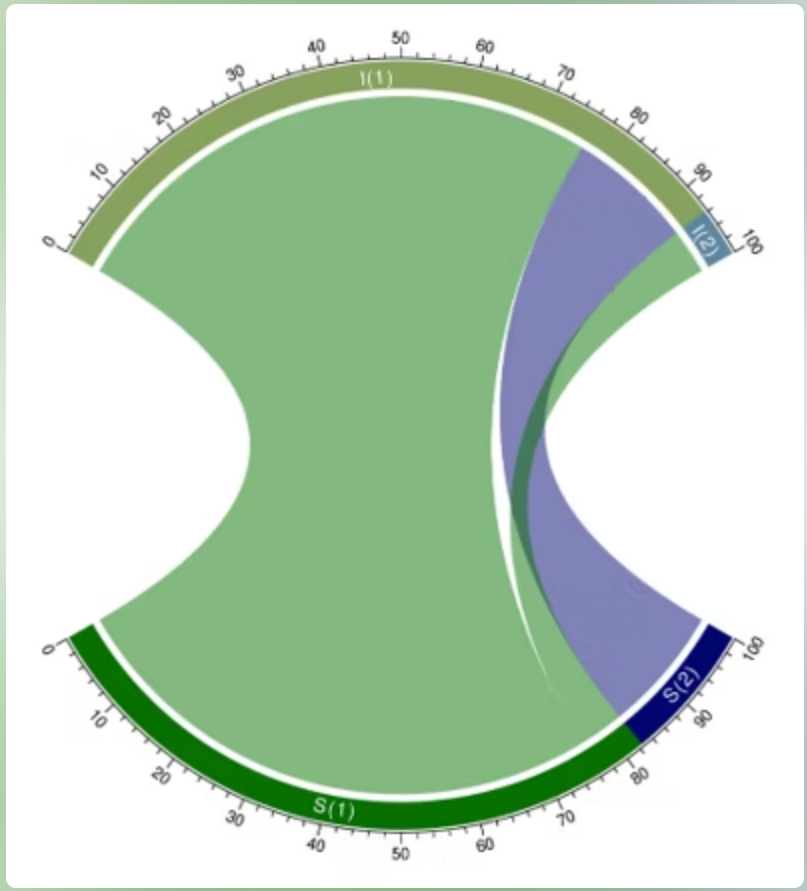
Séries Sintéticas

- Mudanças de nível, tendências determinísticas e passeio aleatório analisado atingem **ordem de integração zero em uma única etapa** (s=1)
- A representação baseada em janelas captura localmente padrões que o operador de diferença global tende a obscurecer

Séries Reais (80 séries)

- Todas as 80 séries atingem estacionariedade em **no máximo duas etapas**
- **82,5%** das séries atingem estacionariedade em **uma etapa** (s=1) — correção de nível suficiente
- **17,5%** requerem **duas etapas** (s=2) — correção de nível e escala; concentrados em **commodities**, **atividade econômica** e **mercado acionário**

⚠ As ordens de integração 1 e 2 **não correspondem diretamente** a 1 ou 2 etapas WARMA — elas capturam informações complementares sobre não-estacionariedade. Séries com ordem de integração 2 aparecem apenas na dívida líquida do setor público (DLSP), ao passo que casos de duas etapas WARMA ocorrem também em outros domínios.



Análise dos Hiperparâmetros

Os hiperparâmetros selecionados revelam como a não estacionariedade se manifesta localmente nas séries analisadas.

Janelas Curtas

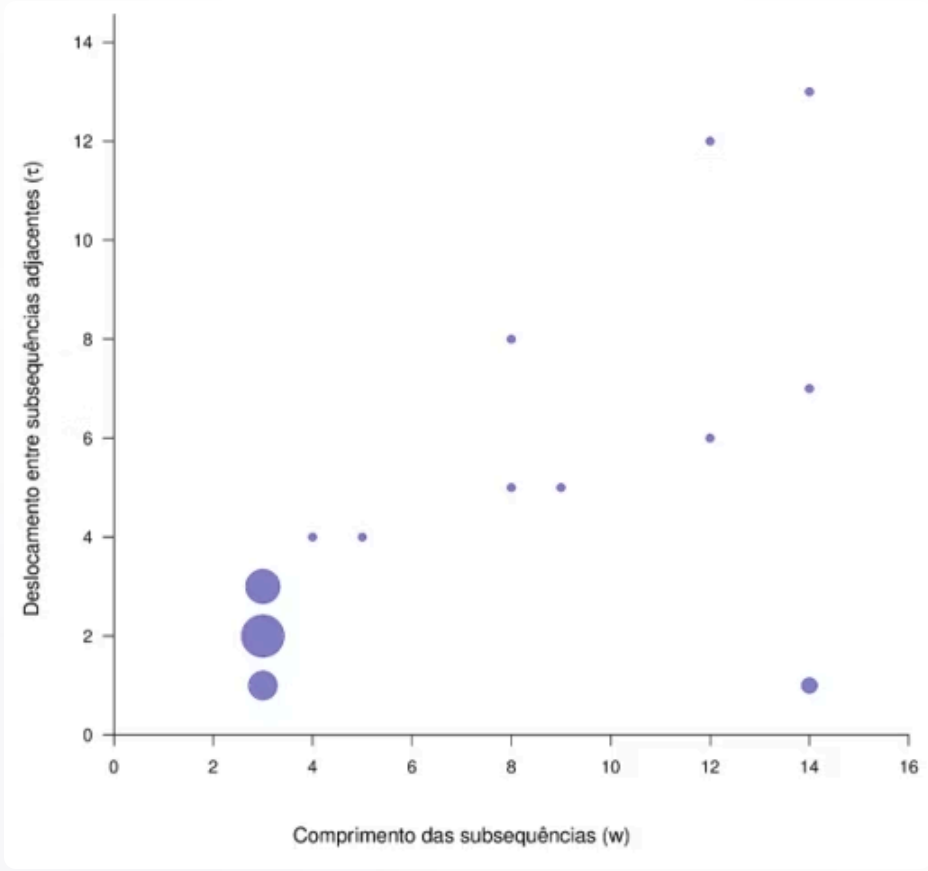
Modelos concentram-se em janelas curtas, especialmente **w = 3**, indicando não estacionariedade de escala local muito pequena.

Deslocamento τ Pequeno

τ aparece frequentemente em valores pequenos admissíveis, especialmente **2 e 3**, controlando sobreposição entre janelas.

Concentração Secundária

Uma concentração menor ocorre em torno de **w = 12 a 14**, sugerindo sazonalidade mensal em algumas séries.

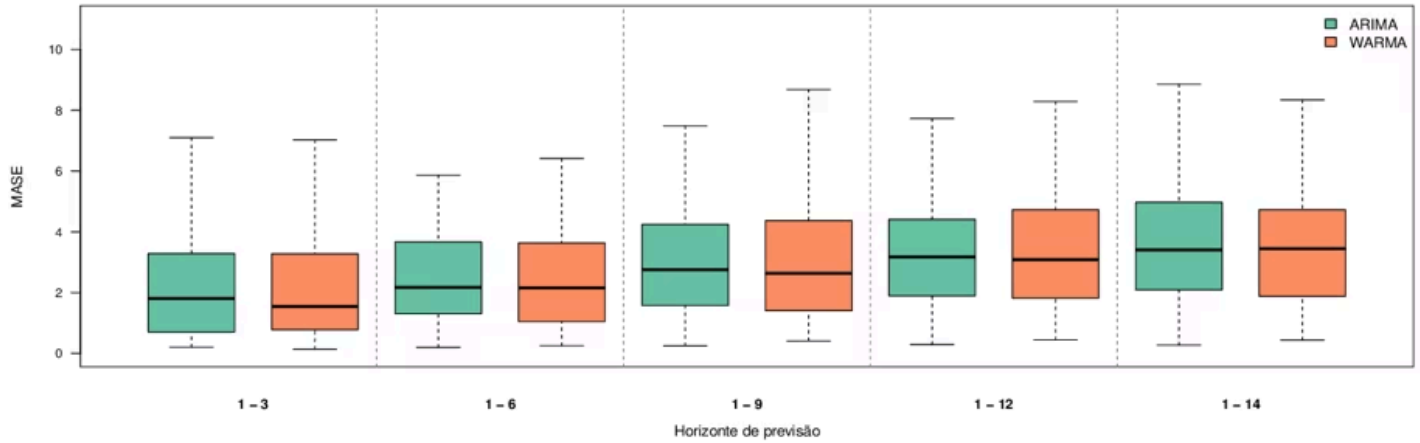


Desempenho Preditivo: WARMA vs. ARIMA

O MASE agregado é próximo entre WARMA e ARIMA, demonstrando competitividade do modelo proposto com diferenças por domínio.

Método	Período	Interest	NPSD	Econ. act.	Stocks	Exchange	CPI	Commod.
ARIMA	1–3	0.99	2.85	1.93	3.92	3.37	1.30	1.95
ARIMA	1–12	2.84	5.73	1.79	5.10	3.35	1.52	5.86
WARMA	1–3	**0.63**	**2.66**	**1.37**	**3.26**	3.45	1.50	2.39
WARMA	1–12	**2.27**	**4.57**	1.77	5.87	3.50	1.64	6.22

O WARMA supera o ARIMA para taxas de juros e NPSD em ambos os horizontes, e para atividade econômica e ações no horizonte curto. O ARIMA é superior para câmbio, CPI e commodities. MASE agregado: horizonte 1–3 (ARIMA 2.40, WARMA 2.39), horizonte 1–12 (ARIMA 4.08, WARMA 4.12), acumulado até 14 passos (ARIMA 4.43, WARMA 4.42).



Conclusões e Trabalhos Futuros

Estacionarização Local

O WARMA oferece estacionarização local com interpretação explícita das etapas de correção.

Desempenho Competitivo

O desempenho permanece competitivo com o ARIMA em todos os horizontes avaliados.

Interpretabilidade

As etapas identificam explicitamente a correção local necessária: nível, escala ou ambos.

Trabalhos Futuros

Covariância, sazonalidade, sensibilidade de hiperparâmetros e custo computacional.

Use o WARMA quando a interpretabilidade da não estacionariedade local for tão importante quanto a acurácia de previsão.



CEFET/RJ

Agradecimentos

FAPERJ

CAPES

CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico

Arthur Ronald Garcia · Jorge Soares · Laura Assis · Fabio Porto · Dayse Pastore · Eduardo Ogasawara

https://eic.cefet-rj.br/dal