WARMA: A Window-Based Autoregressive Moving Average Model for Non-Stationary Time Series[*†]

Arthur Ronald Garcia[1], Jorge Soares[1], Laura Assis[1], Fabio Porto[2], Dayse Pastore[1], Eduardo Ogasawara[1]

[1]CEFET/RJ - Centro Federal de Educação Tecnológica Celso Suckow da Fonseca

[2]LNCC - Laboratório Nacional de Computação Científica

{arthur.several, jorge.soares, laura.assis}@cefet-rj.br

fporto@lncc.br, dayse.pastore@cefet-rj.br, eogasawara@ieee.org

Abstract. *Many time-series models assume stationarity. Analysts therefore often apply global transformations such as differencing. Although these transformations can be effective, they also change the mean, variance, and covariance simultaneously, making the original series harder to interpret. We present WARMA, a model that combines window-based preprocessing with ARMA modeling. The model operates locally over sliding windows. WARMA first adjusts the mean and then the variance, so the preprocessing remains interpretable in terms of steps. Experiments on synthetic and real economic-financial series show that WARMA reaches stationarity in at most two steps and delivers forecasting accuracy comparable to that of ARIMA while providing a more localized view of the preprocessing effect.*

1. Introduction

Most classical linear time-series models require stationarity, that is, constant mean, variance, and covariance over time, a condition that real series rarely satisfy without prior transformation [Box et al., 2015; Tsay, 2010; Gujarati, 2021]. In practice, the most common strategy for addressing this issue is to apply global transformations, such as the successive differencing used by ARIMA [Box et al., 2015]. The problem is that differencing acts on the series globally, even when non-stationarity is driven mainly by local changes in level or scale. As a consequence, part of the original structure becomes less interpretable, which hinders analysis and diagnosis [Salles et al., 2019]. This point is particularly relevant in economic-financial series, where regime changes, shocks, and shifts in volatility often occur locally [Tsay, 2010].

This paper starts from the premise that, although the full series may be non-stationary, local subsequences often exhibit approximately stable behavior. When this occurs, modeling these subsequences and making explicit which properties change over time becomes more meaningful than applying a global transformation that distorts the original signal. From this perspective, we ask whether a window-based local preprocessing scheme can produce representations with integration order zero while preserving

*The authors thank CAPES, FAPERJ, and CNPq for their partial support of this research.

†Generative AI tools were used only for linguistic revision, with a focus on textual clarity. The authors reviewed all suggestions and remain fully responsible for the paper's content.

interpretability and maintaining predictive performance competitive with ARIMA. To answer this question, we propose WARMA (*Window-based AutoRegressive Moving Average*), which combines: (*i*) sliding-window preprocessing that acts sequentially on local statistical properties of the series; and (*ii*) an ARMA model fitted to the resulting stationary representation.

WARMA is organized around the concept of *steps*. Rather than applying a single global transformation, the model estimates how many steps are required to make the series stationary. In this interpretation, the step structure is treated as a conceptual hypothesis: one step suggests the predominance of a non-constant mean, whereas two steps suggest the need to stabilize local variance.

This paper makes three main contributions. First, we formalize the WARMA model, which integrates local window-based preprocessing with ARMA modeling. Second, we make explicit a stepwise interpretation as a conceptual hypothesis for handling local non-stationarity in both mean and variance. Third, we empirically evaluate the proposed model on a synthetic dataset and monthly economic-financial time series, demonstrating its viability in scenarios where global transformations may hinder interpretability.

2. Background and Related Work

A covariance-stationary process has constant mean and variance, and covariance that depends only on the lag. This is the classical setting for AR (AutoRegressive), MA (Moving Average), and ARMA models. In contrast, ARIMA adds an integration component to model non-stationary stochastic processes through the difference operator [Box et al., 2015; Yang and Zurbenko, 2010]. In this setting, diagnosing non-stationarity is also essential, since unit-root tests are traditionally used to distinguish unit-root processes from stationary or trend-stationary series [Herranz, 2017].

The literature also describes transformations that target specific properties of the series, from deterministic or stochastic trend models to variance-stabilizing transformations such as Box-Cox [Yang and Zurbenko, 2010; Box and Cox, 1964]. In an experimental review of transformation methods for forecasting non-stationary series, Salles et al. [2019] show that no procedure performs best in every case. Local approaches such as the Haar-Fisz technique [Fryzlewicz et al., 2006] also highlight the importance of preserving local variation in the series.

Against this background, WARMA is closely related to the Adaptive Normalization technique proposed by Ogasawara et al. [2010]. As a direct antecedent, Ronald et al. [2019] explore an autoregressive model based on adaptive integration over sliding windows, showing that local preprocessing can effectively support forecasting. Building on this line of work, WARMA integrates this preprocessing with an ARMA model and makes stationarization explicit in terms of steps. Thus, the contribution of WARMA is not to introduce a wholly new principle of local preprocessing, but to combine this preprocessing with ARMA modeling and make its interpretation explicit in terms of steps. Unlike ARIMA, whose integration order indicates how many differencing operations are applied, the steps aim to reflect which statistical property must be handled locally [Box et al., 2015; Yang and Zurbenko, 2010].

3. The WARMA Model

Let X $= \langle x_1, \ldots, x_n \rangle$ be a time series. WARMA analyzes X by moving a window of size w along the series with shift τ, where $3 \le w \le n$ and $1 < \tau \le w$. This process produces $h = \left\lfloor \frac{(n-w)}{\tau} \right\rfloor + 1$ subsequences. The residual offset is defined as $o = (n - w) \bmod \tau$, and with this offset, the i-th subsequence is $seq_{i,w,\tau}(\mathrm{X}) = \langle x_{1+o+(i-1)\tau}, \ldots, x_{w+o+(i-1)\tau} \rangle$. The offset aligns the windowing with the end of the series; as a consequence, up to o initial observations may be excluded, reflecting WARMA's priority of maintaining a consistent local representation of the series dynamics rather than preserving every individual observation.

These subsequences form a matrix $\mathbf{A} \in \mathbb{R}^{h \times w}$. Each row corresponds to one window of the series. This representation captures subsequences whose local behavior is approximately stable. The two parameters affect preprocessing in different ways. The value of w defines the span over which the mean and variance are computed. The value of τ controls how much neighboring subsequences overlap. With a small τ, adjacent windows show greater continuity. With $\tau = w$, they become non-overlapping. This flexibility allows the model to adapt to different kinds of local behavior.

This distinction is important for interpretation. When w is small, WARMA reacts to short-lived local changes and tends to produce a more fine-grained representation of the series. When w is larger, the model evaluates a broader local context and favors smoother descriptions of the same process. The shift parameter τ plays a complementary role: it controls how densely these local views are sampled along time. In this sense, w determines the scale at which local properties are measured, whereas τ determines the continuity between neighboring local summaries.

In step 1, WARMA normalizes the mean of each row, that is, $b_{ij} = a_{ij} - \mu(a_i)$, where $\mu(v) = \frac{1}{|v|} \sum_{j=1}^{|v|} v_j$. After this phase, discrepant windows may optionally be discarded using an interquartile-range criterion, as in the Adaptive Normalization technique [Ogasawara et al., 2010]. If the series remains non-stationary, step 2 is applied to stabilize the variance, that is, $c_{ij} = \frac{b_{ij}}{\sigma(b_i)}$, where $\sigma(b_i)$ denotes the standard deviation of the elements in the i-th row. When $\sigma(b_i) = 0$, the resulting row is mapped to zero. At the end of preprocessing, denoting the resulting matrix by $\mathbf{D}$, the series used for modeling is $\mathrm{Vec}\big(\mathbf{D}_{1:h,((w-\tau)+1):w}\big)$, that is, the vectorization of the submatrix formed by the last τ columns.

This final vectorization step is not merely a notational convenience. Because adjacent windows may overlap, using the whole matrix would replicate local information more than necessary. By retaining only the last τ columns of each row before vectorization, WARMA preserves the progression induced by the sliding windows while avoiding excessive reuse of overlapping observations. Therefore, the resulting series remains aligned with the local preprocessing logic without inflating the representation.

The previous steps can also be written in compact form. Let $s \in \{0, 1, 2\}$ be the number of steps required to stationarize the series. For each observation x_t, the normalization uses the local parameters of its corresponding window. This leads to the stationary representation $y_t = \frac{x_t - f_\mu(s)\mu(a^{(g(t))})}{\left(\sigma(a^{(g(t))})\right)^{f_\sigma(s)}}$, where $g(t)$ indicates the window associated with the observation, $f_\mu(s) = 0$ if $s = 0$ and 1 otherwise, and $f_\sigma(s) = 1$ if $s = 2$ and 0 otherwise. The equation describes three cases: *(i)* an already stationary series, *(ii)* a series that requires only mean normalization, and *(iii)* a series that requires normalization of both mean and variance.

In the fitting process, s is defined as the smallest number of steps for which the transformed series has integration order zero according to the *ndiffs* function from the *forecast* package [Hyndman and Khandakar, 2008; Herranz, 2017], and the search over (w, τ) exhaustively evaluates the combinations considered in the experiment, with w starting at 3 and τ ranging from 2 to w.

Applying the transformation to ARMA yields the final model, $\phi(B)y_t = \theta(B)\varepsilon_t$, where $\phi(B)$ and $\theta(B)$ are, respectively, the autoregressive and moving-average polynomials in the lag operator. To highlight the preprocessing hyperparameters as well, we adopt the notation WARMA$(p, (w, s, \tau), q)$. When only the ARMA component is of interest, it is sufficient to report (p, q).

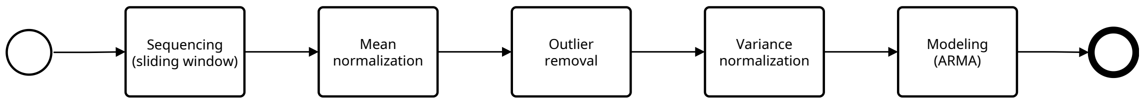

Figure 1. Overview of WARMA

4. Experimental Evaluation

This section is organized into three parts. Subsection 4.1 describes the experimental design, Subsection 4.2 discusses the evidence supporting the conceptual hypothesis of steps, and Subsection 4.3 compares the predictive performance of WARMA and ARIMA.

4.1. Experimental Design

The evaluation proceeds along two fronts. First, it examines the behavior of WARMA under different non-stationary patterns using synthetic time series, including level shifts, deterministic trends, and random walks. Second, it compares the performance of WARMA and ARIMA on a set of 80 non-stationary monthly economic-financial series spanning seven domains, after excluding those that are stationary. These series were selected from a broader collection of 92 monthly economic-financial series assembled from public data portals maintained by B3, Investing.com, the Central Bank of Brazil, IBGE, and Ipeadata [B3 - Brasil, Bolsa, Balcão, 2026; Investing.com, 2026; Banco Central do Brasil, 2026; Instituto Brasileiro de Geografia e Estatística, 2026; Instituto de Pesquisa Econômica Aplicada, 2026].

Model selection follows the automated Box-Jenkins logic via *auto.arima*, using AICc to select the ARMA order for the preprocessed series [Hyndman and Khandakar, 2008; Hurvich and Tsai, 1989]. In WARMA, the procedure exhaustively scans candidate combinations of (w, τ), with w starting at 3 and τ ranging from 2 to w. It defines s as the smallest number of steps for which the transformed series reaches integration order zero according to *ndiffs* and then fits an ARMA model to the resulting series. In this study, the optional step of discarding discrepant windows after mean normalization remains inactive. This choice keeps the comparison with ARIMA methodologically tighter, so that the difference between the models arises primarily from local preprocessing rather than from an additional filtering stage. The overall procedure follows the stages of identification, estimation, and diagnosis.

4.2. Evidence on the Conceptual Hypothesis of Steps

In the synthetic series, WARMA is consistent with the conceptual hypothesis underlying the interpretation-by-steps approach. Level shifts and deterministic trends reach integration order zero in a single step, that is, with only local mean normalization. The same occurs for the analyzed random walk, showing that the window-based representation can locally capture patterns that global differencing tends to obscure.

These examples offer a qualitative reading of the model. More specifically, the fact that WARMA responds with steps compatible with these structures provides favorable evidence for the adopted conceptual hypothesis. Extending the analysis from synthetic to real data, the main structural finding in the same direction is that all 80 non-stationary series reach integration order zero in at most two steps according to the *ndiffs* function used in the study. Of these, 82.5% require only one step and 17.5% require two steps. For this dataset, these results are consistent with the adopted conceptual interpretation of steps. This suggests that, for this set, a large share of non-stationarity can be explained by local-level variations and, to a lesser extent, by local-scale variations. Two-step cases are concentrated in the commodities, economic activity, and stock market domains. This indicates that stabilizing local variance is less common and depends on the type of series under analysis.

The data also clarify the relationship between WARMA and ARIMA. Integration orders 1 and 2 do not correspond directly to 1 or 2 WARMA steps. A series that is differenced twice does not necessarily require separate treatment of two statistical properties, just as a series that reaches stationarity in two WARMA steps is not simply the local counterpart of second-order differencing. The two quantities summarize different things. The integration order counts how many global differencing operations are applied, whereas the number of steps indicates which local statistical properties need to be stabilized. In many cases, WARMA indicates that stabilizing the local mean is enough. Series with integration order two appear only in net public sector debt, whereas two-step cases also appear in other domains. This reinforces the interpretation that integration and steps capture complementary information about non-stationarity rather than equivalent measures of the same phenomenon.

A related empirical point concerns the preprocessing hyperparameters selected by the search procedure. In the analyzed set, the chosen WARMA models are concentrated at very short windows, especially $w = 3$, with corresponding values of τ concentrated at small admissible values, particularly 2 and 3. A second, smaller concentration appears around w between 12 and 14. This pattern suggests that, for many of the monthly economic-financial series in the study, the local behavior relevant for stationarization is captured at a short scale, although some series benefit from a broader local context. Even when forecasting performance remains close to that of ARIMA, these selected hyperparameters provide an additional descriptive view of how local the non-stationary behavior appears to be.

4.3. Predictive Comparison with ARIMA

At the aggregate level, WARMA and ARIMA show very similar predictive performance, as measured by MASE. At horizon $(1–3)$, the average values are practically identical $(2.40$ for ARIMA and 2.39 for WARMA); at $(1–12)$, they remain close $(4.08$ and $4.12)$.

In the cumulative results up to 14 steps, the difference again becomes negligible (4.43 and 4.42). Even so, this overall proximity hides systematic differences across domains.

Table 1. Avg MASE by domain (representative horizons and cumulative periods).

Method	Period	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
	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
	1–12	**2.27**	**4.57**	**1.77**	5.87	3.50	1.64	6.22

As Table 1 shows, the domain-level reading is more informative. For interest rates and net public sector debt, WARMA outperforms ARIMA at both the short and annual horizons, with a particularly pronounced advantage in net public sector debt at the longer horizon. In economic activity, WARMA maintains an advantage across both summary horizons, although this difference diminishes as the horizon increases. In the stock market, WARMA performs better in the short term but worse at the annual horizon. In exchange rates, CPI, and commodities, ARIMA yields a lower mean MASE at both horizons, with small differences in exchange rates and larger ones in commodities.

WARMA nevertheless remains predictively competitive. Its interpretive value is therefore not only conceptual. The model handles non-stationarity locally through windows instead of relying solely on a global transformation. This gives a more explicit view of which statistical properties need stabilization, making WARMA a plausible alternative when the analyst's interest is not limited to forecasting but also includes understanding whether the series requires local correction of level, scale, or both.

5. Conclusion

The results for the analyzed dataset provide a positive answer to the research question. The proposed local window-based preprocessing scheme produces representations with integration order zero according to the operational criterion based on *ndiffs*, while offering a more transparent interpretation of the underlying statistical properties. Moreover, the predictive performance of WARMA remains competitive with ARIMA. The proposed formalization describes local preprocessing and its integration with the ARMA component in a unified manner, and the experimental results support the conceptual hypothesis based on stepwise transformations, with integration order zero achieved in at most two steps for the evaluated series.

From a practical perspective, this makes WARMA a plausible alternative when the objective extends beyond forecasting to include interpretability, particularly in identifying whether local corrections in level, scale, or both are required. Nevertheless, the current formulation has a well-defined scope: it explicitly addresses local non-stationarity in mean and variance, but does not address non-stationarity in covariance or directly model seasonality. Future work includes extending the model to incorporate covariance-related dynamics, analyzing the sensitivity of the hyperparameters and computational cost, and developing a seasonal extension of the framework.

References

B3 - Brasil, Bolsa, Balcão (2026). B3 Data and Information Services.

Banco Central do Brasil (2026). Sistema Gerenciador de Séries Temporais.

Box, G. E. P. and Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26:211–243.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, USA, 5th edition.

Fryzlewicz, P., Sapatinas, T., and Rao, S. S. (2006). A Haar-Fisz technique for locally stationary volatility estimation. *Biometrika*, 93:687 – 704.

Gujarati, D. N. (2021). *Essentials of Econometrics*. SAGE, Los Angeles.

Herranz, E. (2017). Unit root tests. *Wiley Interdisciplinary Reviews: Computational Statistics*, 9.

Hurvich, C. M. and Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76:297 – 307.

Hyndman, R. J. and Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 27:1 – 22.

Instituto Brasileiro de Geografia e Estatística (2026). SIDRA: Banco de Tabelas Estatísticas.

Instituto de Pesquisa Econômica Aplicada (2026). Ipeadata.

Investing.com (2026). Investing.com Market Data.

Ogasawara, E., Martinez, L. C., De Oliveira, D., Zimbrão, G., Pappa, G. L., and Mattoso, M. (2010). Adaptive Normalization: A novel data normalization approach for non-stationary time series. In *Proceedings of the International Joint Conference on Neural Networks*.

Ronald, A., Salles, R., Belloze, K., Pastore, D., and Ogasawara, E. (2019). Modelo Autorregressivo de Integração Adaptativa. In *Anais do Simpósio Brasileiro de Banco de Dados (SBBD)*, pages 175–180. SBC.

Salles, R., Belloze, K., Porto, F., Gonzalez, P. H., and Ogasawara, E. (2019). Nonstationary time series transformation methods: An experimental review. *Knowledge-Based Systems*, 164:274 – 291.

Tsay, R. S. (2010). *Analysis of Financial Time Series*. Wiley.

Yang, W. and Zurbenko, I. (2010). Nonstationarity. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2:107 – 115.