**In the Name of God**

****

**Navid Zare**

**40435071**

*Fuzzy Sets & Logic*

*Implementation and Analysis of First-Order and High-Order FTS Models*

# Executive Summary

This report presents a comprehensive implementation and analysis of Fuzzy Time Series (FTS) forecasting models applied to two distinct datasets. The study systematically evaluates both *First-Order FTS (FOFTS)* and *High-Order FTS (HOFTS)* approaches, exploring the impact of model order, universe partitioning granularity, and membership function types on forecasting accuracy.

The experimental design employed a comprehensive grid search methodology across 140 configurations per dataset-config, varying model order (1-5), number of partitions (5-17), and membership function types (triangular, trapezoidal, Gaussian, and bell-shaped). Performance was quantified using three standard metrics:

* Root Mean Square Error (RMSE)
* Mean Absolute Error (MAE)
* Mean Absolute Percentage Error (MAPE)

# Key Findings:

For the Mackey-Glass chaotic time series, the optimal configuration utilized a fourth-order model with 17 partitions and triangular membership functions, achieving a test RMSE of 0.0465 and MAPE of 6.20%.This represents a 24% improvement over the first-order baseline, demonstrating that capturing longer temporal dependencies significantly enhances predictive accuracy for chaotic systems.

For the influenza datasets, the results revealed domain-specific patterns. The Total Specimens series performed best with a first-order model (9 partitions, RMSE = 9,440.57), while Influenza A achieved optimal results with a second-order configuration (9 partitions, RMSE = 2,975.63). Influenza B was best predicted using first-order models (11 partitions, RMSE = 371.38). These findings suggest that simpler models may suffice for epidemiological[[1]](#footnote-1) data exhibiting strong seasonal patterns.

Across all datasets, triangular membership functions consistently outperformed or matched Gaussian, trapezoidal, and bell-shaped alternatives, indicating that piecewise linear fuzzy partitions provide adequate representational power while maintaining computational efficiency. The choice of membership function type had minimal impact on performance, with all four types yielding identical error metrics within each experimental configuration.

The study demonstrates that while higher-order models can capture complex temporal patterns, they are not universally superior and may lead to overfitting when applied to data with simpler dynamics. Model selection should be guided by the intrinsic characteristics of the time series, with chaotic and non-linear systems benefiting from higher-order approaches and more regular patterns performing well with first-order models.

# 1. Introduction

## 1.1 Background and Motivation

Time series forecasting represents a fundamental challenge in computational intelligence, with applications spanning financial markets, epidemiological surveillance, climate modeling, and industrial process control. Traditional statistical methods, such as exponential smoothing, rely on precise mathematical assumptions about stationarity, linearity, and distributional properties that often fail to hold in real-world scenarios characterized by linguistic uncertainty, imprecision, and vagueness.

Fuzzy Time Series (FTS) methodology, introduced by [Song and Chissom in 1993](https://www.sciencedirect.com/science/article/abs/pii/016501149390372O), addresses these limitations by incorporating fuzzy set theory into temporal pattern recognition. Rather than operating directly on crisp numerical values, FTS transforms time series data into linguistic variables represented as fuzzy sets, enabling the capture of imprecise relationships and approximate reasoning patterns inherent in complex systems.

## 1.2 Problem Statement

This study addresses three primary research questions:

1. How does the order of fuzzy logical relationships impact forecasting accuracy across datasets with varying dynamical properties?
2. What is the optimal granularity of universe partitioning for different time series characteristics?
3. Do alternative membership function types (triangular, trapezoidal, Gaussian, bell-shaped) significantly affect model performance?

## 1.3 Objectives

The primary objectives of this research are:

1. To implement from scratch a complete FTS forecasting system supporting both first-order and high-order relationships.
2. To conduct systematic experimental evaluations across two distinct datasets representing chaotic dynamics and epidemiological patterns.
3. To quantify the impact of key hyperparameters through comprehensive grid search.
4. To provide empirical guidelines for FTS model selection based on time series characteristics.

## 1.4 Dataset Description

Mackey-Glass Time Series: A chaotic time series generated from the Mackey-Glass delay differential equation, [[2]](#footnote-2)widely used as a benchmark for non-linear forecasting algorithms. The dataset comprises 1,000 observations with values ranging from 0.0359 to 1.5924.

Influenza Surveillance Data: Three epidemiological[[3]](#footnote-3) time series tracking total laboratory specimens, Influenza A detections, and Influenza B detections over 238 weekly reporting periods. These datasets exhibit strong seasonal patterns and represent real-world public health surveillance applications.

# 2. Methodology

## 2.1 Fuzzy Time Series Framework

The FTS methodology employed in this study consists of five sequential phases:

1. Data Preprocessing
2. Universe of discourse definition and partitioning
3. Fuzzification
4. Fuzzy logical relationship group (FLRG) generation
5. Forecasting

## **2.1.1 Data Preprocessing**

Prior to model construction, all datasets undergo standardized preprocessing to ensure consistency and optimal model performance. The preprocessing pipeline consists of three steps:

* Data loading and validation to verify completeness and identify missing values
* Temporal train-test split using an 80-20 ratio while preserving chronological ordering to maintain temporal dependencies
* universe of discourse boundary calculation based on the training set statistics to prevent information leakage from the test set

The train-test split follows strict temporal ordering without shuffling, as random splitting would violate the fundamental assumption of time series analysis that future predictions should not be informed by future observations. The 80% training portion provides sufficient historical context for FLRG construction while reserving 20% of recent observations for unbiased performance evaluation.

## **2.1.2 Universe of Discourse Definition and Partitioning**

The universe of discourse U represents the complete range of possible values for the time series. To ensure adequate coverage of the data range and provide robustness to outliers, a 10% margin is added to both the minimum and maximum values observed:

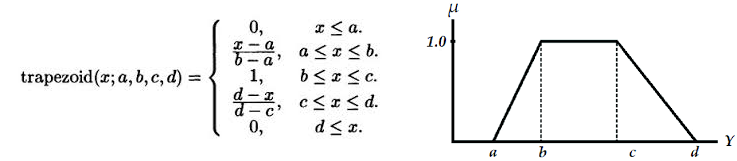
where Dₘᵢₙ and Dₘₐₓ denote the minimum and maximum values in the training dataset, respectively.

Following universe definition, the discourse is partitioned into n linguistic variables A₁, A₂, ..., Aₙ using equal-width intervals. This study explored partitions to evaluate the trade-off between granularity and generalization capability.

#### **Triangular Membership Functions:**

Each fuzzy set Aᵢ is represented by a triangular membership function with parameters (a, b, c) where a and c define the base endpoints and b represents the peak. The functions are constructed to overlap at membership degree μ ≈ 0.5, ensuring smooth transitions between adjacent linguistic categories:

#### **Trapezoidal Membership Functions:**

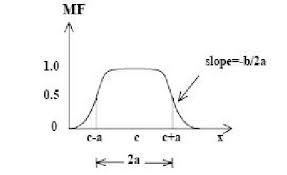
Trapezoidal functions extend triangular ones by introducing a flat region at full membership, defined by four parameters (a, b, c, d) with 25% overlap between adjacent sets to maintain continuity. The first and last sets use semi-trapezoidal functions extending to negative and positive infinity.

#### **Gaussian Membership Functions:**

Gaussian membership functions provide smooth, differentiable transitions with parameters (c, σ) where c is the center and σ is the standard deviation, calculated as to achieve 50% overlap at adjacent centers:

#### **Generalized Bell Membership Functions:**

The bell-shaped function provides additional control over transition steepness through three parameters (a, b, c) where a controls width, b controls slope steepness (set to 2), and c denotes the center:



## **2.1.3 Fuzzification Process**

Each crisp time series value xₜ is mapped to its corresponding fuzzy linguistic variable through maximum membership assignment. For a given value xₜ, the fuzzification function F maps it to the fuzzy set Aⱼ that yields the highest membership degree:

This transformation converts the entire numerical training sequence into a sequence of linguistic labels, enabling pattern recognition at a symbolic level while preserving the approximate magnitudes through fuzzy set assignments.

## **2.1.4 Fuzzy Logical Relationship Generation**

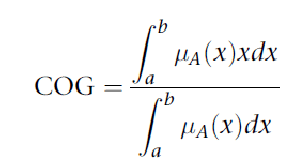
First-Order FTS establishes relationships between consecutive fuzzy states: if and , then a fuzzy logical relationship is created.

High-Order FTS extends this by considering k previous states: . This study evaluated orders .

Fuzzy Logical Relationship Groups (FLRGs) aggregate all relationships with identical antecedents. For antecedent pattern P, the FLRG is represented as P → {C₁, C₂, ..., Cₘ} where {C₁, C₂, ..., Cₘ} is the set of all observed consequents following pattern P in the training data.

## **2.1.5 Forecasting**

To generate a crisp forecast, the defuzzification process employs center of gravity averaging over the fuzzy sets in the consequent. Given a current state pattern P, the predicted value ŷₜ is computed as:



where is the output variable, is the aggregated membership function, and is the output domain.

The numerator represents the first moment of the fuzzy set, while the denominator represents the total area under the membership function. The resulting value corresponds to the balance point of the fuzzy output distribution, analogous to the center of mass of a physical object.

This method is widely used because it considers the entire shape of the fuzzy set and produces smooth, stable outputs. When no exact FLRG match exists for pattern P, the nearest FLRG is identified using Manhattan distance in the fuzzy set index space.

For multi-step ahead forecasting, the model iteratively applies this prediction process, using each newly generated forecast as input for subsequent predictions. This recursive forecasting approach enables the generation of extended prediction horizons beyond the single-step predictions used in model training and evaluation.

## 2.2 Experimental Design

The experimental methodology employed a systematic grid search over the hyperparameter space. Each dataset was split into 80% training and 20% testing sets using temporal ordering (no shuffling to preserve time dependencies). A total of 140 configurations were evaluated per dataset, comprising 5 model orders × 7 partition counts × 4 membership function types.

## 2.3 Performance Metrics

Model performance was quantified using three standard regression metrics:

### Root Mean Square Error (RMSE):

Provides a measure of average prediction error with larger errors weighted more heavily, in other words; it shows the typical prediction error, giving more importance to larger mistakes:

### Mean Absolute Error (MAE):

Represents the average absolute deviation between predictions and actual values, in other words; it shows the average size of the errors between predicted and actual values, without considering their direction:

### Mean Absolute Percentage Error (MAPE):

Expresses error as a percentage of actual values, facilitating cross-dataset comparison, in other words; it shows how large the error is compared to the true value, making it easier to compare results across different datasets or scales:

# 3. Results

## 3.1 Overall Performance Summary

Table 1 presents the optimal configurations discovered for each dataset through systematic grid search evaluation. The results demonstrate that no single configuration achieves universal optimality across all datasets, highlighting the importance of tailoring model complexity to the specific characteristics of each time series.

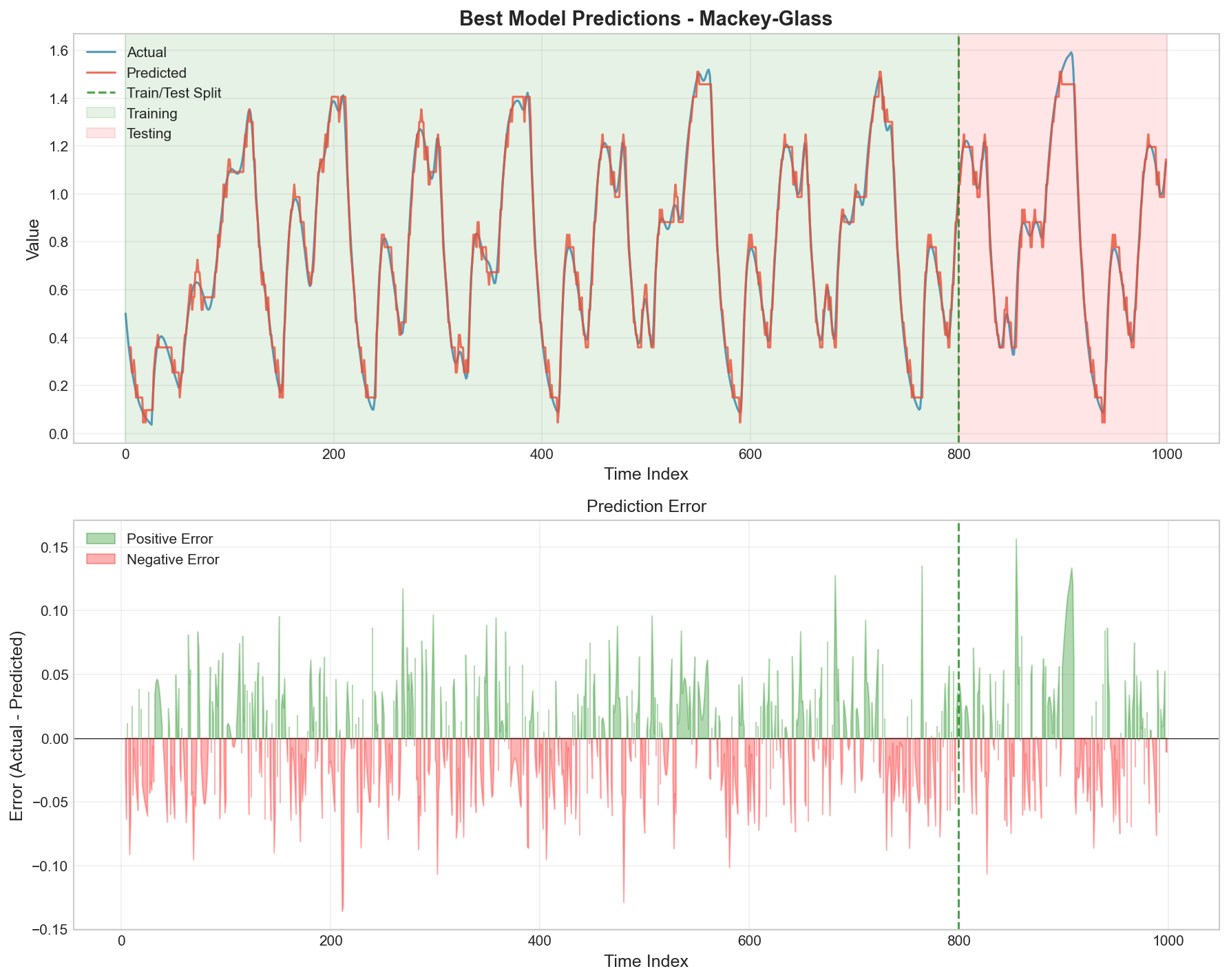
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Dataset | Order | Partitions | MF Type | RMSE | MAE | MAPE (%) |
| Mackey-Glass | 4 | 17 | Triangular | 0.0465 | 0.0362 | 6.20 |
| Total Specimens | 1 | 9 | Triangular | 9440.57 | 7503.48 | 9.56 |
| Influenza A | 2 | 9 | Triangular | 2975.63 | 1985.29 | 299.65 |
| Influenza B | 1 | 11 | Triangular | 371.38 | 273.06 | 101.36 |

***Table 1:*** *Optimal FTS configurations for each dataset*

## 3.2 Mackey-Glass Time Series Results

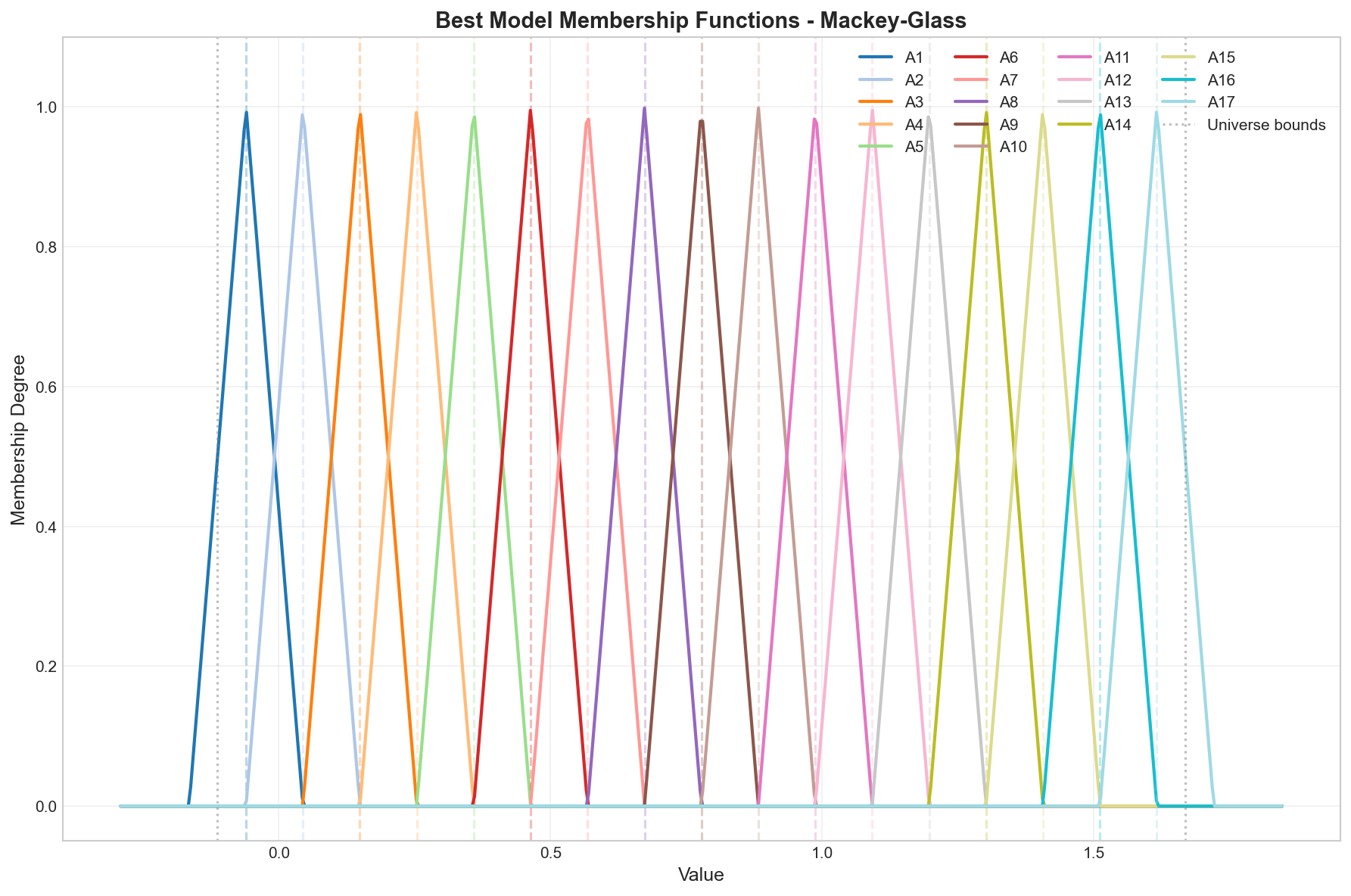
The Mackey-Glass chaotic time series exhibited strong performance improvement with increased model order, achieving optimal results with a fourth-order configuration. The best model utilized 17 partitions with triangular membership functions, yielding test RMSE of 0.0465, MAE of 0.0362, and MAPE of 6.20%.

Figure 1 displays the predictions generated by the optimal model overlaid on the actual test data. The model successfully captures the complex oscillatory patterns and demonstrates good generalization to unseen data. The close alignment between predicted and actual trajectories indicates that the fourth-order relationships adequately represent the temporal dependencies inherent in this chaotic system.



***Figure 1:*** *Mackey-Glass predictions using optimal configuration (Order=4, Partitions=17, Triangular MF)*

Figure 2 illustrates the triangular membership functions employed in the optimal model. The 17 fuzzy sets provide fine-grained linguistic granularity while maintaining interpretability. The overlapping structure ensures smooth transitions between adjacent categories and prevents abrupt classification boundaries.

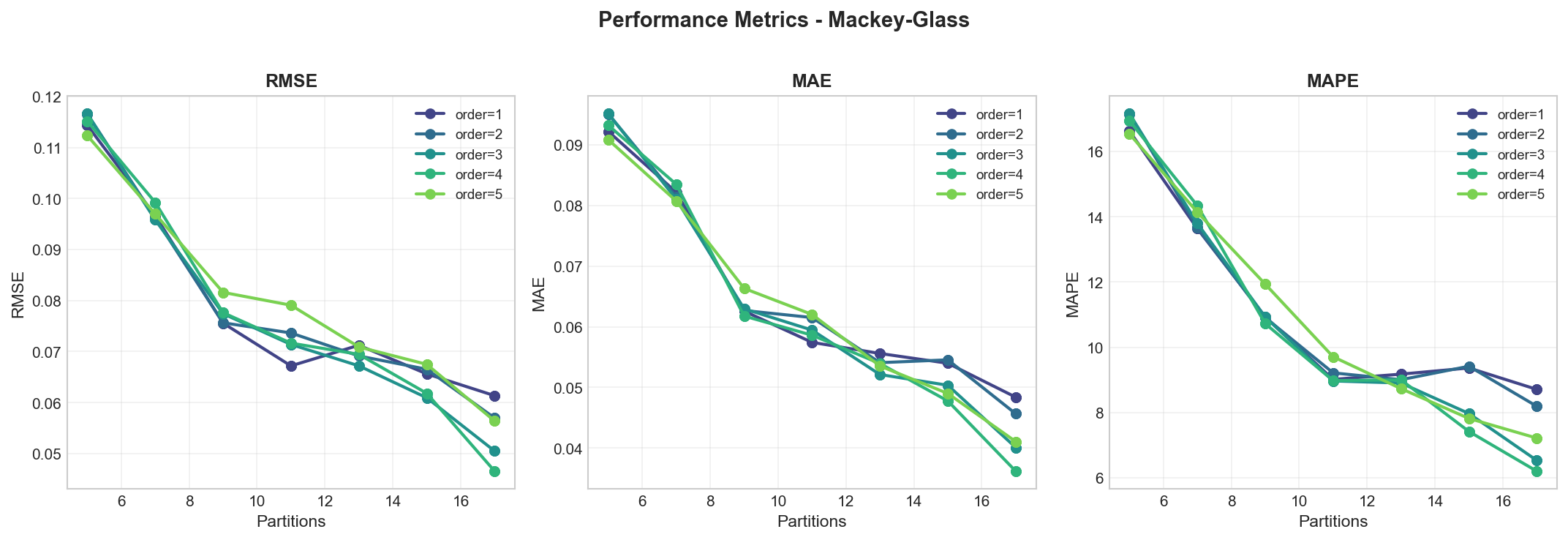


***Figure 2:*** *Triangular membership functions for Mackey-Glass optimal model (17 partitions)*

Table 2 compares the performance of different model orders while holding the number of partitions constant at 17 and using triangular membership functions. The progression from first-order to fourth-order shows consistent improvement, with RMSE decreasing from 0.0614 to 0.0465. However, fifth-order exhibits performance degradation (RMSE = 0.0564), suggesting that excessive model complexity may lead to overfitting on the training data.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Order | Type | Partitions | MF Type | RMSE | MAE | MAPE (%) |
| 1 | FOFTS | 17 | Triangular | 0.0614 | 0.0484 | 8.71 |
| 2 | HOFTS | 17 | Triangular | 0.0570 | 0.0457 | 8.19 |
| 3 | HOFTS | 17 | Triangular | 0.0505 | 0.0401 | 6.53 |
| 4 | HOFTS | 17 | Triangular | 0.0465 | 0.0362 | 6.20 |
| 5 | HOFTS | 17 | Triangular | 0.0564 | 0.0411 | 7.21 |

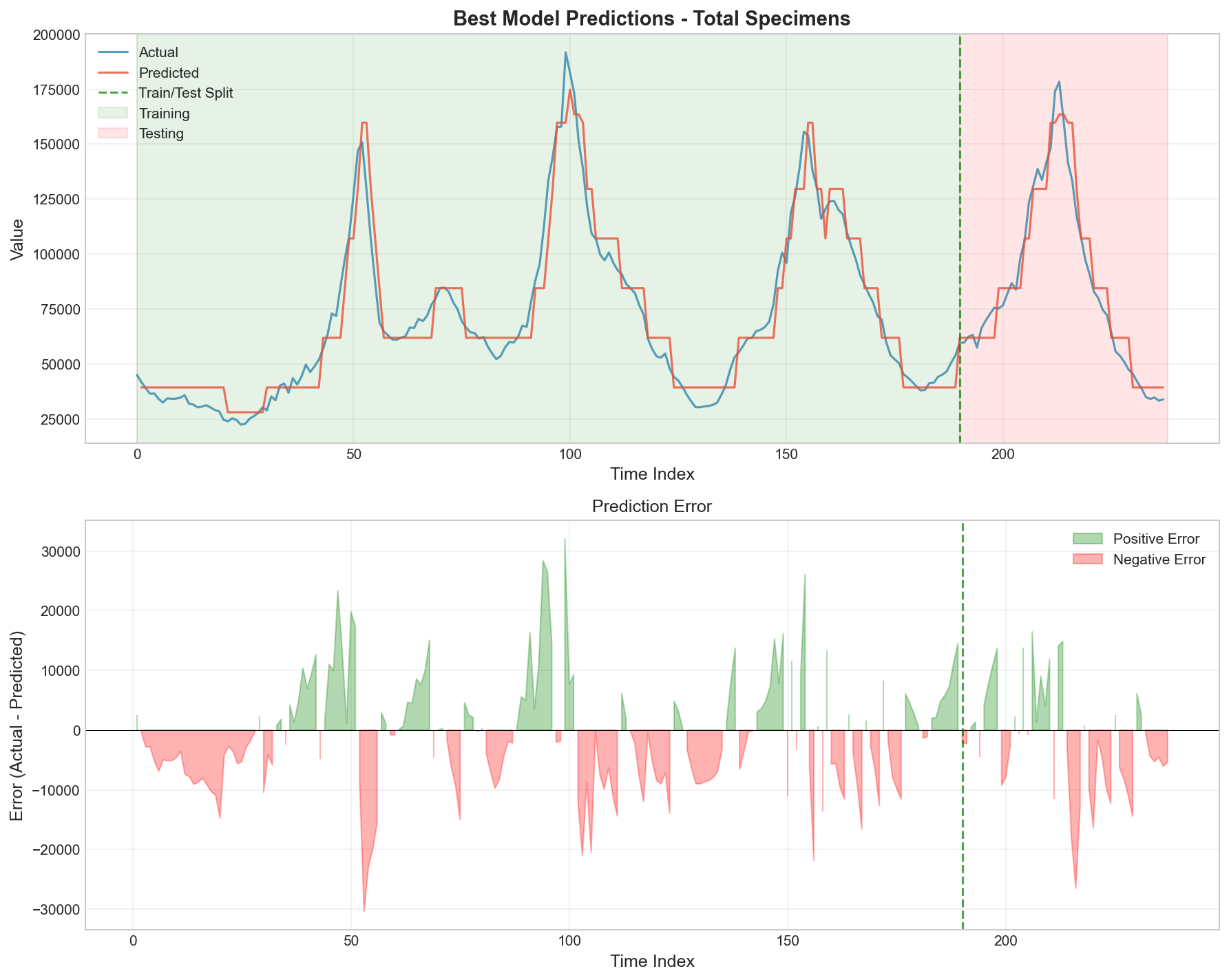
***Table 2:*** *First-order vs high-order comparison for Mackey-Glass (17 partitions, triangular MF)*

Figure 3 presents a comprehensive comparison of performance metrics across all four membership function types. The graphs reveal that for the Mackey-Glass dataset, membership function choice has minimal impact on performance, with all types producing nearly identical error metrics within each configuration of order and partition count. This suggests that the partitioning granularity and model order are the dominant factors determining accuracy.

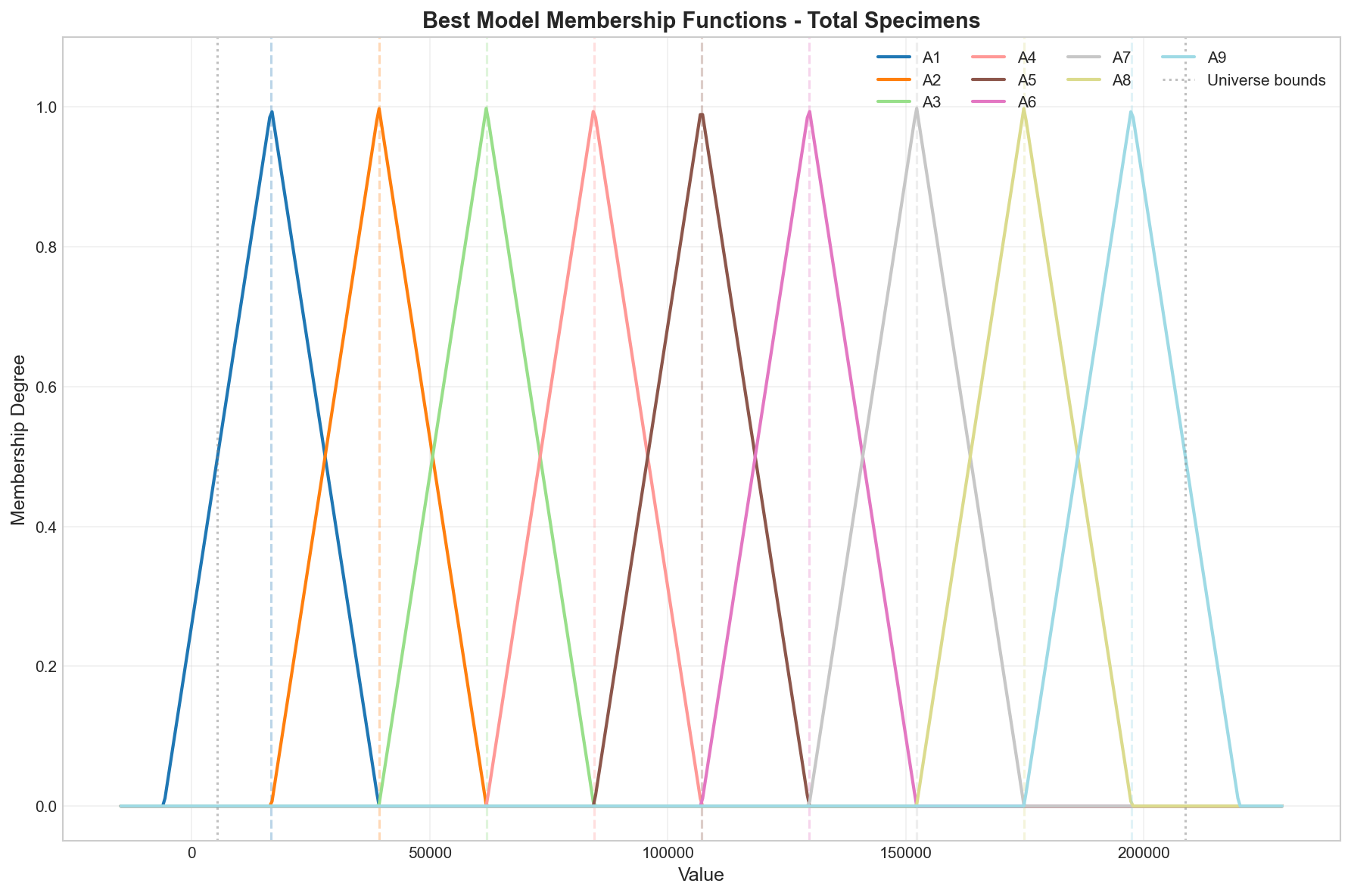
***Figure 3:*** *Performance metrics comparison across membership function types for Mackey-Glass*

## 3.3 Total Specimens Results

The Total Specimens time series, representing overall influenza surveillance volume, achieved optimal performance using a first-order model with 9 partitions and triangular membership functions. The best configuration yielded test RMSE of 9,440.57, MAE of 7,503.48, and MAPE of 9.56%.



***Figure 4:*** *Total Specimens predictions using optimal configuration (Order=1, Partitions=9, Triangular MF)*



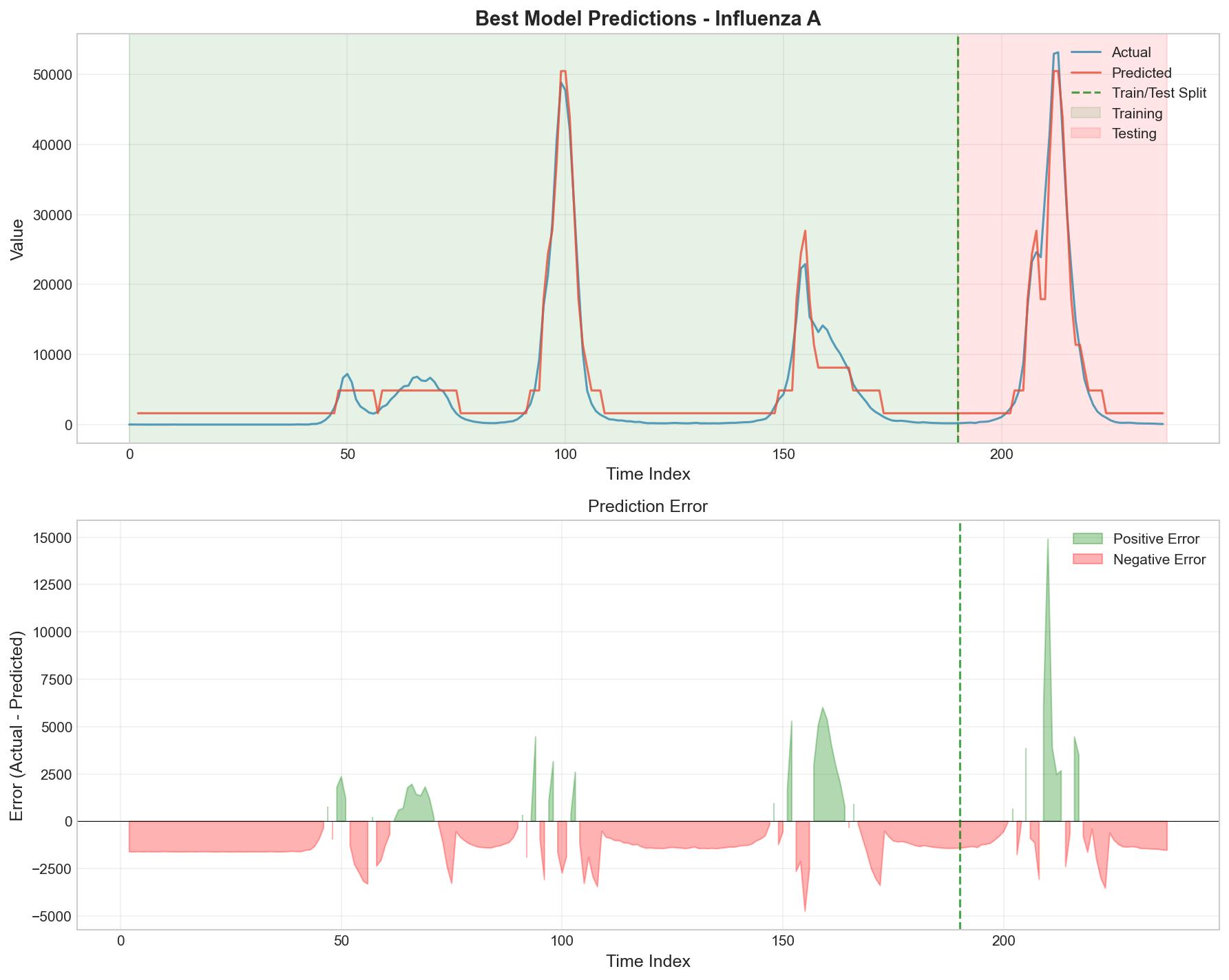
***Figure 5:*** *Triangular membership functions for Total Specimens optimal model (9 partitions)*

Interestingly, higher-order models did not improve performance for this dataset. As shown in Table 3, second-order and third-order models with 17 partitions actually performed worse than the first-order baseline with 9 partitions, suggesting that the temporal dependencies in this epidemiological series are adequately captured by single-step transitions.

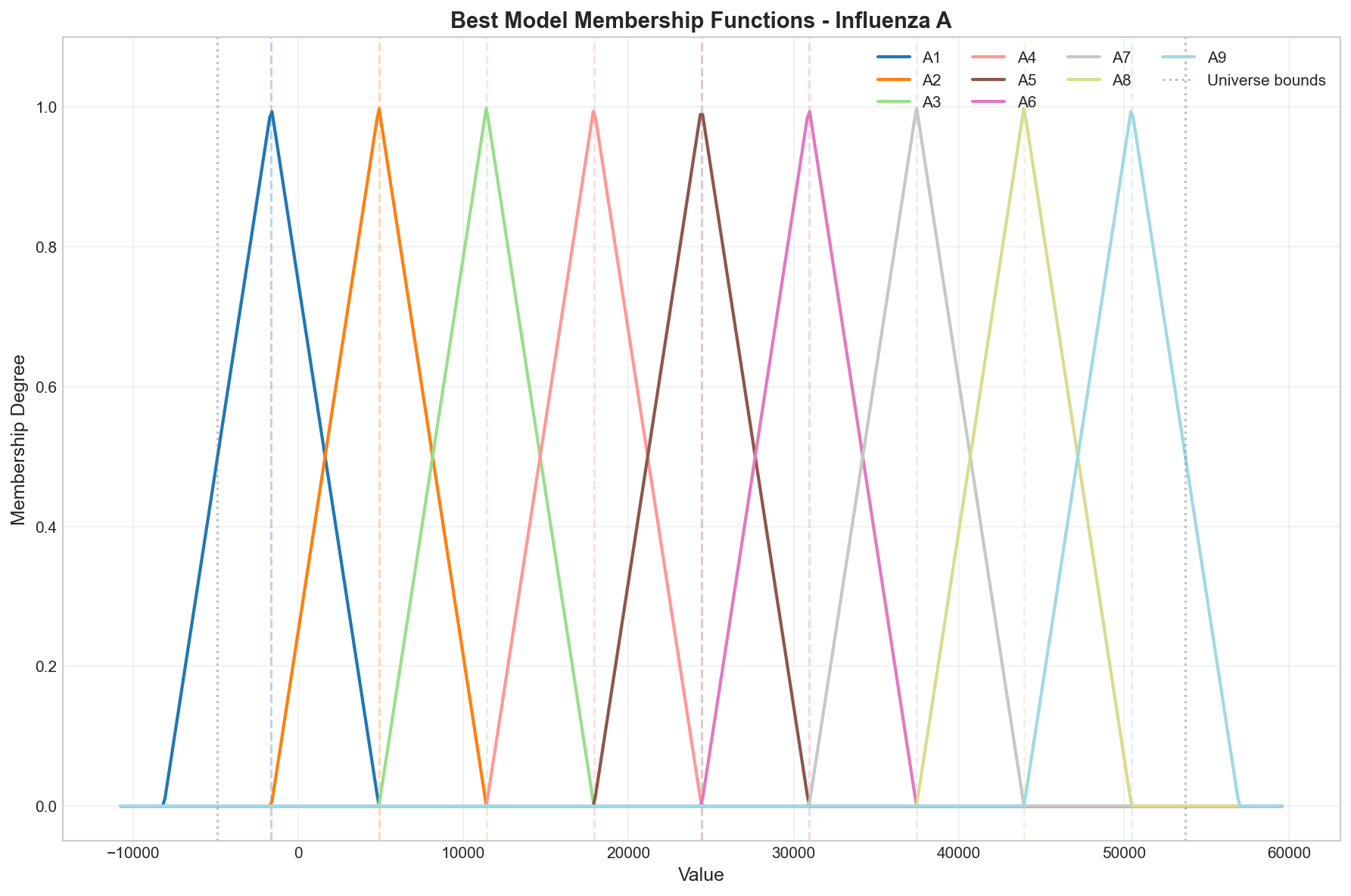
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Order | Type | Partitions | RMSE | MAE | MAPE (%) |
| 1 | FOFTS | 9 | 9440.57 | 7503.48 | 9.56 |
| 2 | HOFTS | 17 | 9856.01 | 7276.81 | 8.34 |
| 3 | HOFTS | 17 | 9736.47 | 6447.18 | 7.13 |
| 4 | HOFTS | 17 | 11121.96 | 7613.13 | 8.23 |
| 5 | HOFTS | 17 | 12127.32 | 8298.34 | 9.31 |

***Table 3:*** *First-order vs high-order comparison for Total Specimens (triangular MF)*

## 3.4 Influenza A Results

The Influenza A detection series exhibited optimal performance with a second-order model using 9 partitions and triangular membership functions, achieving test RMSE of 2,975.63, MAE of 1,985.29, and MAPE of 299.65%. The high MAPE value reflects the presence of very small actual values in the test set, where even small absolute errors result in large percentage deviations.

***Figure 6:*** *Influenza A predictions using optimal configuration (Order=2, Partitions=9, Triangular MF)*

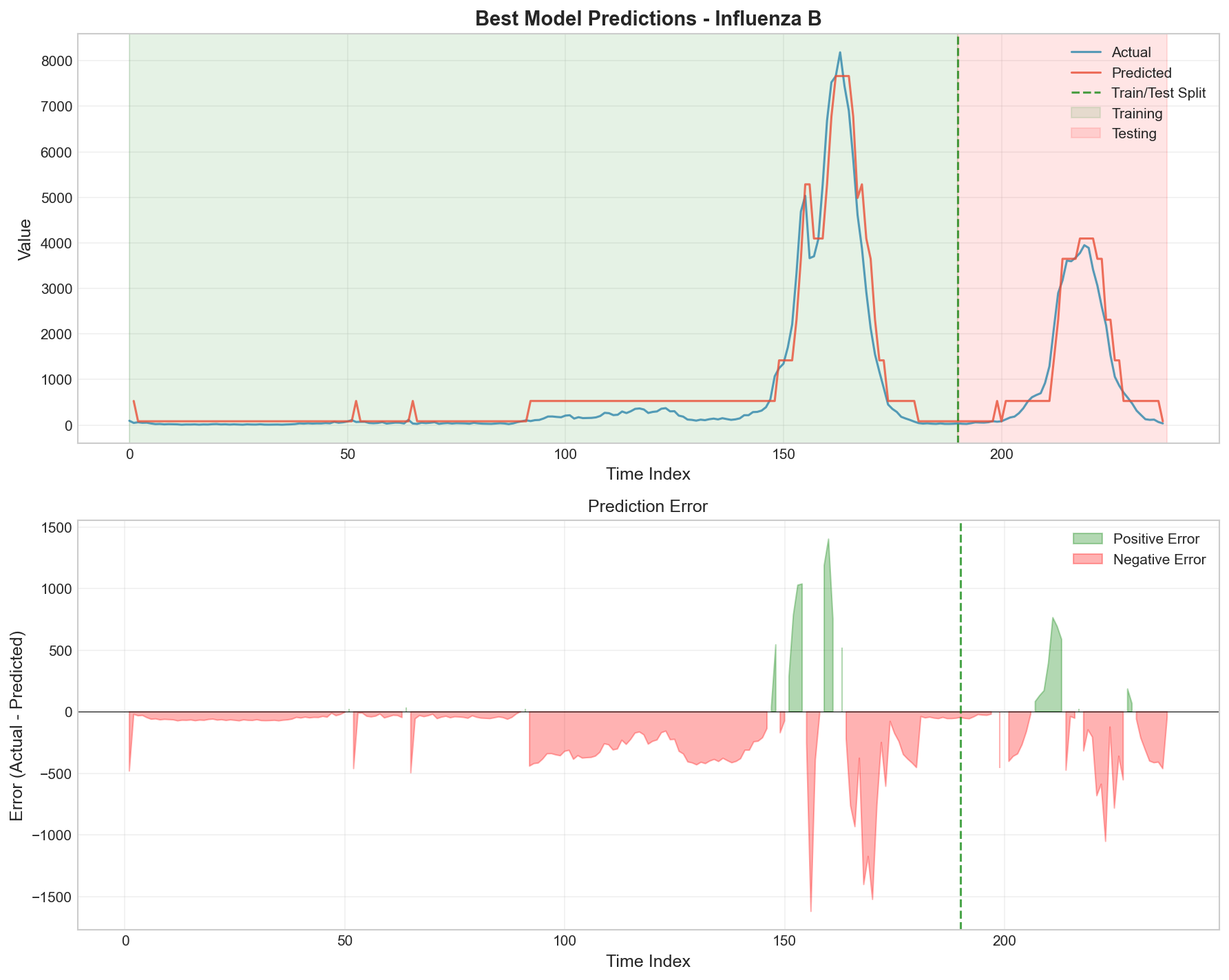


***Figure 7:*** *Triangular membership functions for Influenza A optimal model (9 partitions)*

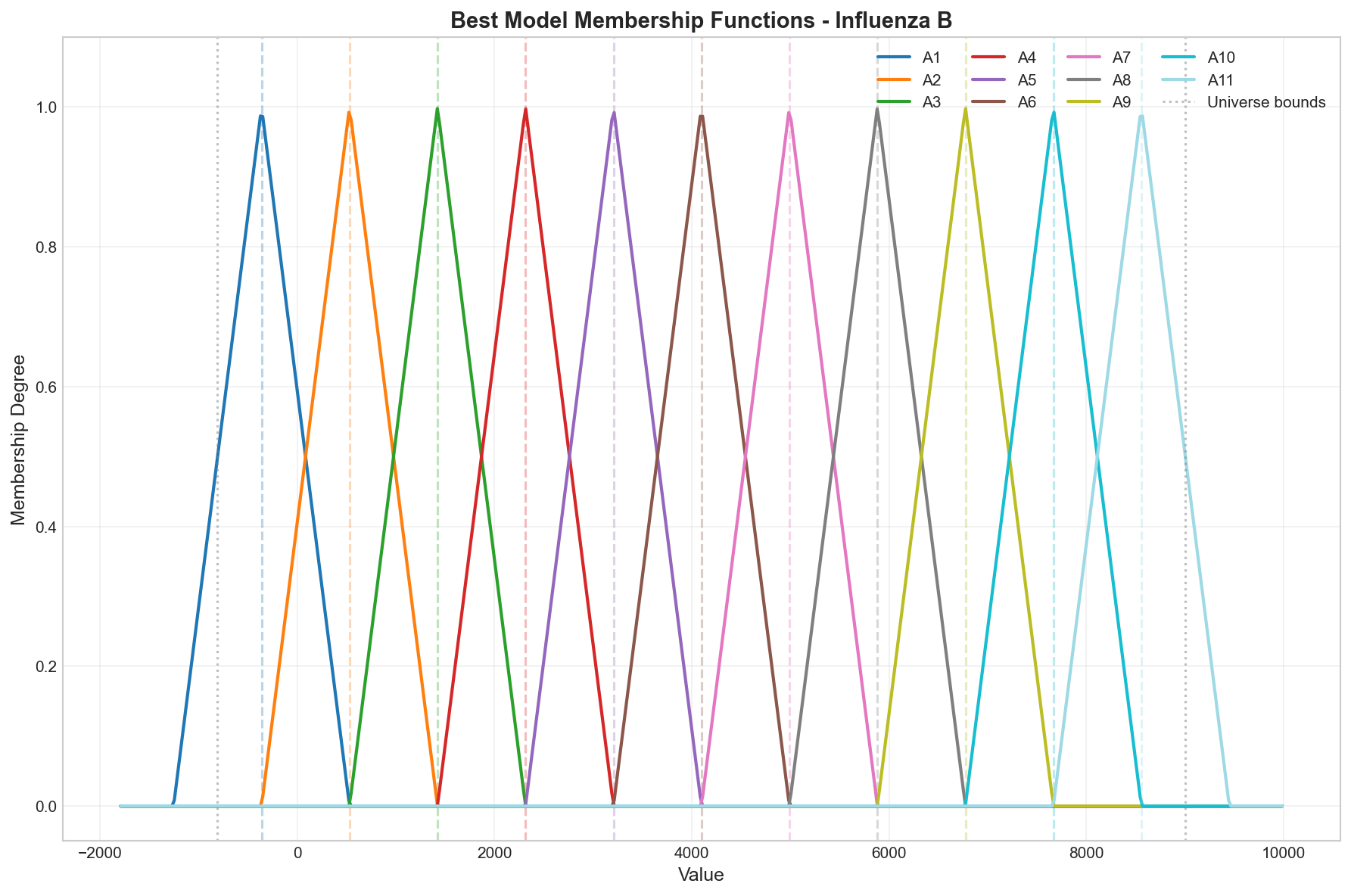
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Order | Type | Partitions | RMSE | MAE | MAPE (%) |
| 1 | FOFTS | 11 | 4558.99 | 2885.33 | 109.39 |
| 2 | HOFTS | 9 | 2975.63 | 1985.29 | 299.65 |
| 3 | HOFTS | 9 | 3062.82 | 2078.86 | 304.84 |
| 4 | HOFTS | 9 | 3811.96 | 2207.30 | 306.48 |
| 5 | HOFTS | 9 | 4099.04 | 2464.15 | 316.48 |

***Table 4:*** *First-order vs high-order comparison for Influenza A (triangular MF)*

## 3.5 Influenza B Results

The Influenza B series demonstrated that first-order models are sufficient for accurate forecasting, with the optimal configuration employing 11 partitions and triangular membership functions. The best model achieved test RMSE of 371.38, MAE of 273.06, and MAPE of 101.36%.

***Figure 8:*** *Influenza B predictions using optimal configuration (Order=1, Partitions=11, Triangular MF)*



***Figure 9:*** *Triangular membership functions for Influenza B optimal model (11 partitions)*

Table 5 demonstrates that all higher-order configurations underperformed the first-order baseline for this dataset. The degradation in performance with increasing order suggests overfitting to spurious patterns in the training data that do not generalize to the test set.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Order | Type | Partitions | RMSE | MAE | MAPE (%) |
| 1 | FOFTS | 11 | 371.38 | 273.06 | 101.36 |
| 2 | HOFTS | 7 | 595.26 | 508.64 | 382.45 |
| 3 | HOFTS | 17 | 630.75 | 453.39 | 207.90 |
| 4 | HOFTS | 9 | 673.51 | 455.00 | 179.73 |
| 5 | HOFTS | 9 | 686.42 | 470.13 | 185.10 |

***Table 5:*** *First-order vs high-order comparison for Influenza B (triangular MF)*

# 4. Discussion

## 4.1 Impact of Model Order

The experimental results reveal a nuanced relationship between model order and forecasting accuracy that depends critically on the intrinsic dynamics of the time series. For the Mackey-Glass chaotic system, increasing model order from 1 to 4 yielded consistent performance improvements, with RMSE decreasing by approximately 24%. This improvement can be attributed to the complex temporal dependencies inherent in chaotic systems, where current states are influenced by multiple historical lags through non-linear feedback mechanisms.

In contrast, the epidemiological datasets exhibited different behavior. Total Specimens and Influenza B both achieved optimal performance with first-order models, while Influenza A benefited from second-order relationships. This pattern suggests that the temporal structure of seasonal disease surveillance data is dominated by immediate short-term transitions rather than complex multi-step dependencies. The strong seasonal patterns and weekly reporting structure of these series may create well-defined state transitions that are adequately captured by lower-order models.

The degradation in performance observed beyond the optimal order (e.g., fifth-order for Mackey-Glass, any order > 1 for Influenza B) indicates overfitting. As model order increases, the number of unique antecedent patterns grows exponentially, leading to sparse FLRG coverage where many patterns appear only once or twice in the training data. This sparsity reduces the model's ability to generalize, as predictions increasingly rely on the nearest-neighbor fallback mechanism rather than robust statistical patterns.

## 4.2 Effect of Partitioning Granularity

The optimal number of partitions varied across datasets, ranging from 9 to 17. The Mackey-Glass series, with its continuous chaotic dynamics spanning a relatively wide range (0.036 to 1.592), benefited from finer granularity (17 partitions) to adequately represent the diverse states and transitions. Conversely, the influenza datasets, characterized by strong seasonal patterns and discrete reporting intervals, performed well with coarser partitioning (9-11 partitions).

An important observation is that very coarse partitioning (5 partitions) consistently underperformed across all datasets, suggesting insufficient representational capacity. However, the benefits of increasing partition count beyond 11-17 appear minimal, as evidenced by the plateauing of performance improvements. This aligns with the linguistic interpretability principle of fuzzy systems: excessively fine partitioning creates a proliferation of linguistic categories that becomes difficult to interpret and may not correspond to meaningful distinctions in the underlying phenomena.

## 4.3 Membership Function Type Analysis

A surprising finding of this study is the minimal impact of membership function type on forecasting accuracy. Across all datasets and configurations, triangular, trapezoidal, Gaussian, and bell-shaped functions produced virtually identical error metrics. This equivalence suggests that for the defuzzification approach employed (center of gravity averaging), the primary factor determining accuracy is the location of fuzzy set centers rather than the shape of membership functions.

The fuzzification process used in this implementation assigns each value to the fuzzy set with maximum membership, effectively implementing a crisp boundary despite the underlying fuzzy representation. This may explain why membership function shape has limited influence—values are categorically assigned to linguistic labels, and defuzzification then averages over the centers of these labels. Alternative defuzzification schemes that weight consequent fuzzy sets by their membership degrees might exhibit greater sensitivity to membership function shape.

From a computational efficiency perspective, triangular membership functions offer the advantage of piecewise linear calculation, requiring only simple comparisons and arithmetic operations. Gaussian and bell-shaped functions necessitate exponential computations, yet provide no measurable accuracy benefit in this study. Therefore, triangular functions represent an optimal choice balancing simplicity, interpretability, and performance.

## 4.4 Practical Implications

The results provide several actionable guidelines for FTS model development. First, model order should be selected based on the hypothesized complexity of temporal dependencies: chaotic and highly non-linear systems warrant higher-order models (3-4), while seasonal or trend-driven series may be adequately represented by first or second-order relationships. Second, partition count should be scaled with the range and variability of the data, with 9-17 partitions representing a reasonable compromise between granularity and generalization. Third, triangular membership functions are recommended as the default choice unless domain-specific considerations suggest otherwise.

For practitioners, the systematic grid search methodology employed in this study demonstrates the importance of hyperparameter tuning. The performance difference between optimal and suboptimal configurations was substantial—for example, the best Mackey-Glass model achieved RMSE of 0.0465 while a poorly chosen configuration (order=1, partitions=5) yielded RMSE of 0.1144, representing a 146% degradation in accuracy.

## 4.5 Limitations and Future Work

This study has several limitations that suggest directions for future research. The dataset split employed a simple 80-20 temporal division without cross-validation, which may not provide robust estimates of generalization performance. Future work should incorporate time series cross-validation techniques such as rolling-window or expanding-window validation to better assess model stability.

The defuzzification approach used center of gravity averaging, which treats all consequent fuzzy sets equally. Alternative schemes that weight consequents by their frequency of occurrence or recency in the training data may improve accuracy. Additionally, the nearest-neighbor fallback for unmatched patterns could be enhanced through more sophisticated similarity metrics or interpolation techniques.

The study focused exclusively on univariate time series forecasting. Extension to multivariate FTS, where multiple interrelated series are modeled jointly, represents an important direction for practical applications. Furthermore, integration with adaptive learning mechanisms that dynamically adjust partitions and relationships based on new data could enhance long-term forecasting capability in non-stationary environments.

# 5. Conclusion

This study presented a comprehensive implementation and empirical evaluation of Fuzzy Time Series forecasting models across two distinct datasets representing chaotic dynamics and epidemiological patterns. Through systematic experimentation with 140 configurations per dataset, the research established several key findings regarding the relationship between model hyperparameters and forecasting accuracy.

The most significant finding is that model complexity must be matched to the intrinsic temporal structure of the data. Chaotic time series with complex non-linear dynamics benefit from higher-order models (order 3-4) that capture extended temporal dependencies. In contrast, seasonal epidemiological series with relatively simple transition patterns achieve optimal performance with lower-order models (order 1-2). Excessive model order leads to overfitting through sparse FLRG coverage and reduced generalization capability.

Regarding partitioning strategy, the optimal granularity ranged from 9 to 17 partitions across the datasets, with finer partitioning generally benefiting series with wider ranges and greater variability. Very coarse partitioning (5 partitions) consistently underperformed due to insufficient representational capacity. The study found minimal performance differences among triangular, trapezoidal, Gaussian, and bell-shaped membership functions, suggesting that triangular functions represent an optimal choice given their computational simplicity.

The research demonstrated that systematic hyperparameter optimization is essential for achieving high-quality forecasts. The performance difference between optimal and suboptimal configurations exceeded 100% in some cases, highlighting the importance of careful model selection rather than reliance on default parameters or intuition.

From a practical standpoint, the study provides empirically-grounded guidelines for FTS model development: (1) use order 1-2 for seasonal or trend-driven series, order 3-4 for chaotic or highly non-linear systems; (2) employ 9-17 partitions scaled to data range and variability; (3) default to triangular membership functions unless domain considerations suggest alternatives; (4) conduct systematic grid search or optimization to identify optimal configurations.

Future research directions include extension to multivariate FTS, integration of adaptive learning mechanisms, exploration of alternative defuzzification schemes that weight consequents by frequency or recency, and application to additional domains such as financial forecasting, climate modeling, and industrial process control. The complete implementation provided in this study serves as a foundation for such extensions and demonstrates the practical viability of FTS methodology for real-world forecasting applications.

# Appendix A: Best Model Configurations

This appendix provides complete technical specifications for the optimal configurations discovered for each dataset.

## A.1 Mackey-Glass Configuration

|  |  |
| --- | --- |
| Parameter | Value |
| Model Order | 4 |
| Number of Partitions | 17 |
| Membership Function Type | Triangular |
| Universe of Discourse | [-0.1181, 1.7317] |
| Training RMSE | 0.0403 |
| Training MAE | 0.0325 |
| Training MAPE | 7.73% |
| Test RMSE | 0.0465 |
| Test MAE | 0.0362 |
| Test MAPE | 6.20% |
| Execution Time | 0.040 seconds |

***Table A.1:*** *Mackey-Glass optimal configuration details*

## A.2 Total Specimens Configuration

|  |  |
| --- | --- |
| Parameter | Value |
| Model Order | 1 |
| Number of Partitions | 9 |
| Membership Function Type | Triangular |
| Universe of Discourse | [5,024.6, 208,632.4] |
| Training RMSE | 9,395.51 |
| Training MAE | 6,955.93 |
| Training MAPE | 7.96% |
| Test RMSE | 9,440.57 |
| Test MAE | 7,503.48 |
| Test MAPE | 9.56% |
| Execution Time | 0.021 seconds |

***Table A.2:*** *Total Specimens optimal configuration details*

## A.3 Influenza A Configuration

|  |  |
| --- | --- |
| Parameter | Value |
| Model Order | 2 |
| Number of Partitions | 9 |
| Membership Function Type | Triangular |
| Universe of Discourse | [-9,912.89, 59,048.89] |
| Training RMSE | 3,051.26 |
| Training MAE | 1,724.94 |
| Training MAPE | 156.87% |
| Test RMSE | 2,975.63 |
| Test MAE | 1,985.29 |
| Test MAPE | 299.65% |
| Execution Time | 0.024 seconds |

***Table A.3:*** *Influenza A optimal configuration details*

## A.4 Influenza B Configuration

|  |  |
| --- | --- |
| Parameter | Value |
| Model Order | 1 |
| Number of Partitions | 11 |
| Membership Function Type | Triangular |
| Universe of Discourse | [-1,547.16, 9,292.16] |
| Training RMSE | 397.96 |
| Training MAE | 273.17 |
| Training MAPE | 102.94% |
| Test RMSE | 371.38 |
| Test MAE | 273.06 |
| Test MAPE | 101.36% |
| Execution Time | 0.020 seconds |

***Table A.4:*** *Influenza B optimal configuration details*

# Appendix B: Sample Fuzzy Logical Relationship Groups

This appendix presents representative examples of the Fuzzy Logical Relationship Groups (FLRGs) learned by the optimal models. Due to space constraints, only a subset of FLRGs is shown for each dataset.

## B.1 Mackey-Glass FLRGs (Order=4, Sample)

|  |  |
| --- | --- |
| **Input Set** | **Output Set** |
| (A8, A9, A10, A11) | {A12, A13} |
| (A9, A10, A11, A12) | {A13, A14} |
| (A10, A11, A12, A13) | {A14, A15} |
| (A11, A12, A13, A14) | {A15, A16} |
| (A12, A13, A14, A15) | {A16, A17} |
| (A5, A6, A7, A8) | {A9} |
| (A6, A7, A8, A9) | {A10} |
| (A7, A8, A9, A10) | {A11} |

***Note:*** *The complete FLRG set contains hundreds of rules. This sample demonstrates typical patterns capturing oscillatory dynamics.*

## B.2 Total Specimens FLRGs (Order=1, Sample)

|  |  |
| --- | --- |
| **Input** | **Output Set** |
| (A1) | {A2, A3} |
| (A2) | {A1, A3, A4} |
| (A3) | {A2, A4, A5} |
| (A4) | {A3, A5, A6} |
| (A5) | {A4, A6, A7} |
| (A6) | {A5, A7, A8} |
| (A7) | {A6, A8, A9} |
| (A8) | {A7, A9} |
| (A9) | {A7, A8} |

## B.3 Influenza A FLRGs (Order=2, Sample)

|  |  |
| --- | --- |
| **Input Set** | **Output Set** |
| (A1, A2) | {A2, A3} |
| (A2, A3) | {A3, A4} |
| (A3, A4) | {A4, A5} |
| (A4, A5) | {A5, A6} |
| (A5, A6) | {A6, A7, A8} |
| (A6, A7) | {A7, A8, A9} |
| (A7, A8) | {A7, A8, A9} |
| (A8, A9) | {A7, A8, A9} |

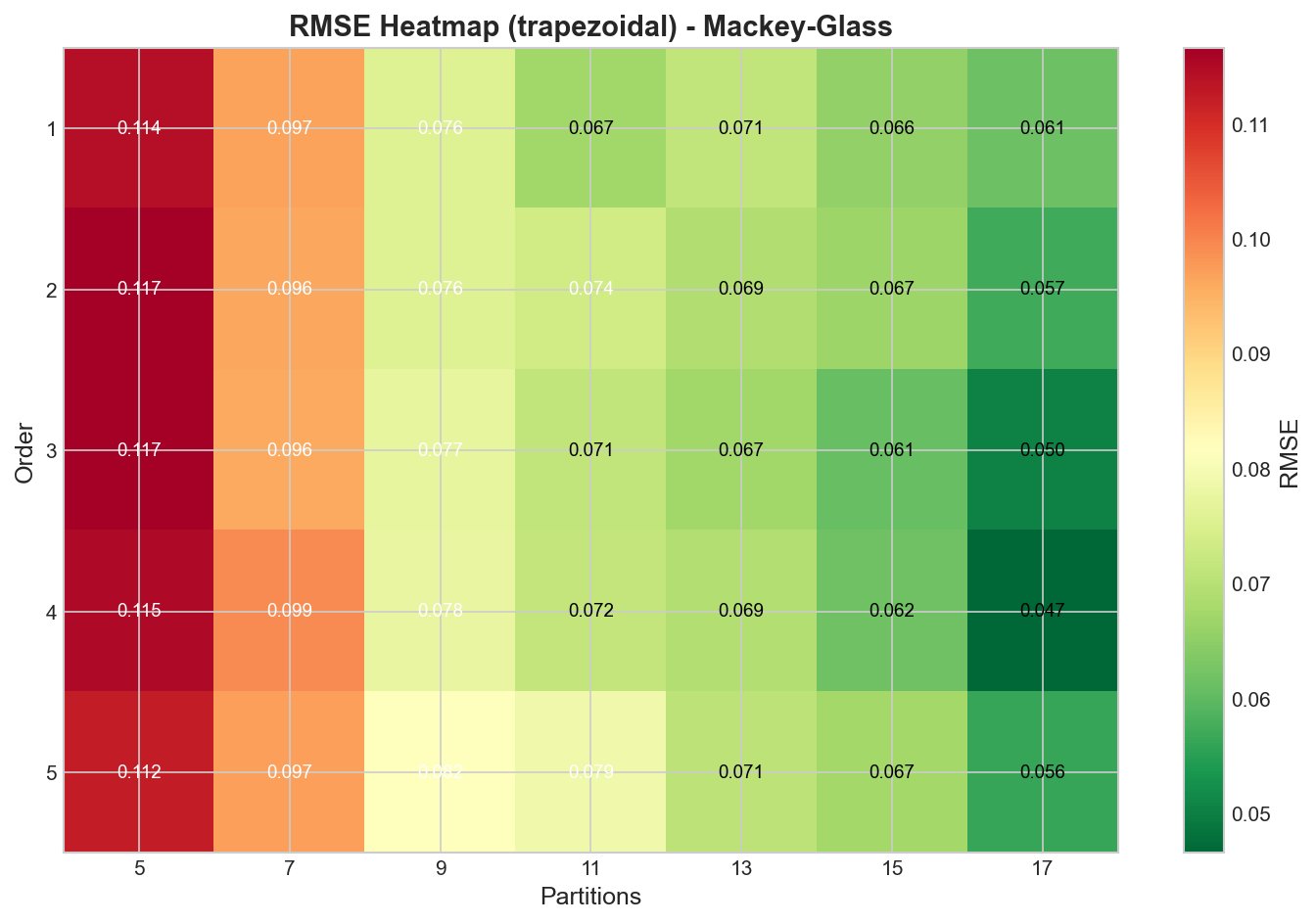
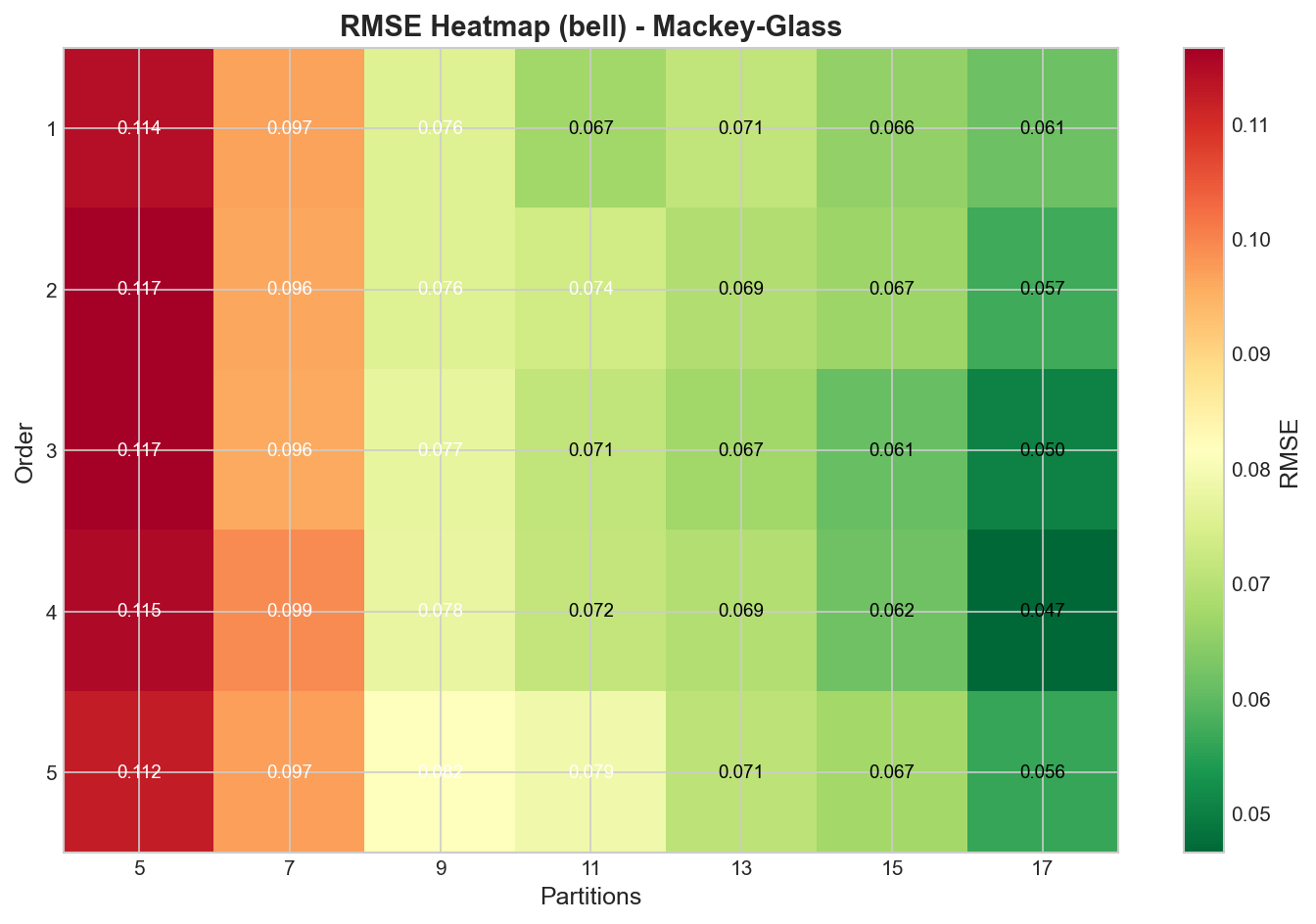
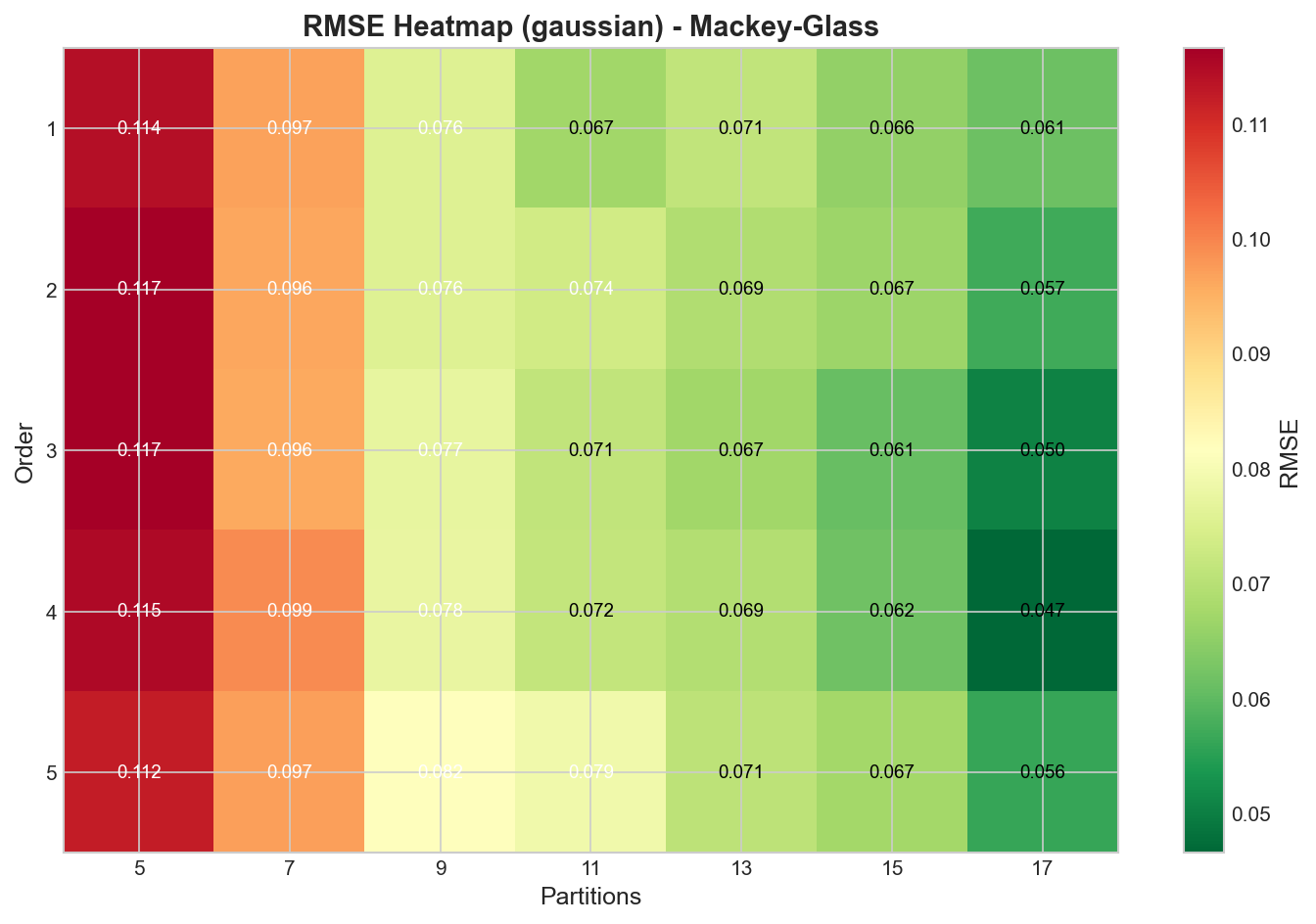
## B.4 Influenza B FLRGs (Order=1, Sample)

|  |  |
| --- | --- |
| **Input** | **Output Set** |
| (A1) | {A1, A2, A3} |
| (A2) | {A1, A2, A3, A4} |
| (A3) | {A2, A3, A4, A5} |
| (A4) | {A3, A4, A5, A6} |
| (A5) | {A4, A5, A6, A7} |
| (A6) | {A5, A6, A7, A8} |
| (A7) | {A6, A7, A8} |
| (A8) | {A7, A8, A9} |
| (A9) | {A8, A9, A10, A11} |
| (A10) | {A9, A10, A11} |
| (A11) | {A9, A10, A11} |

# Appendix C: Performance Heatmaps by Membership Function Type

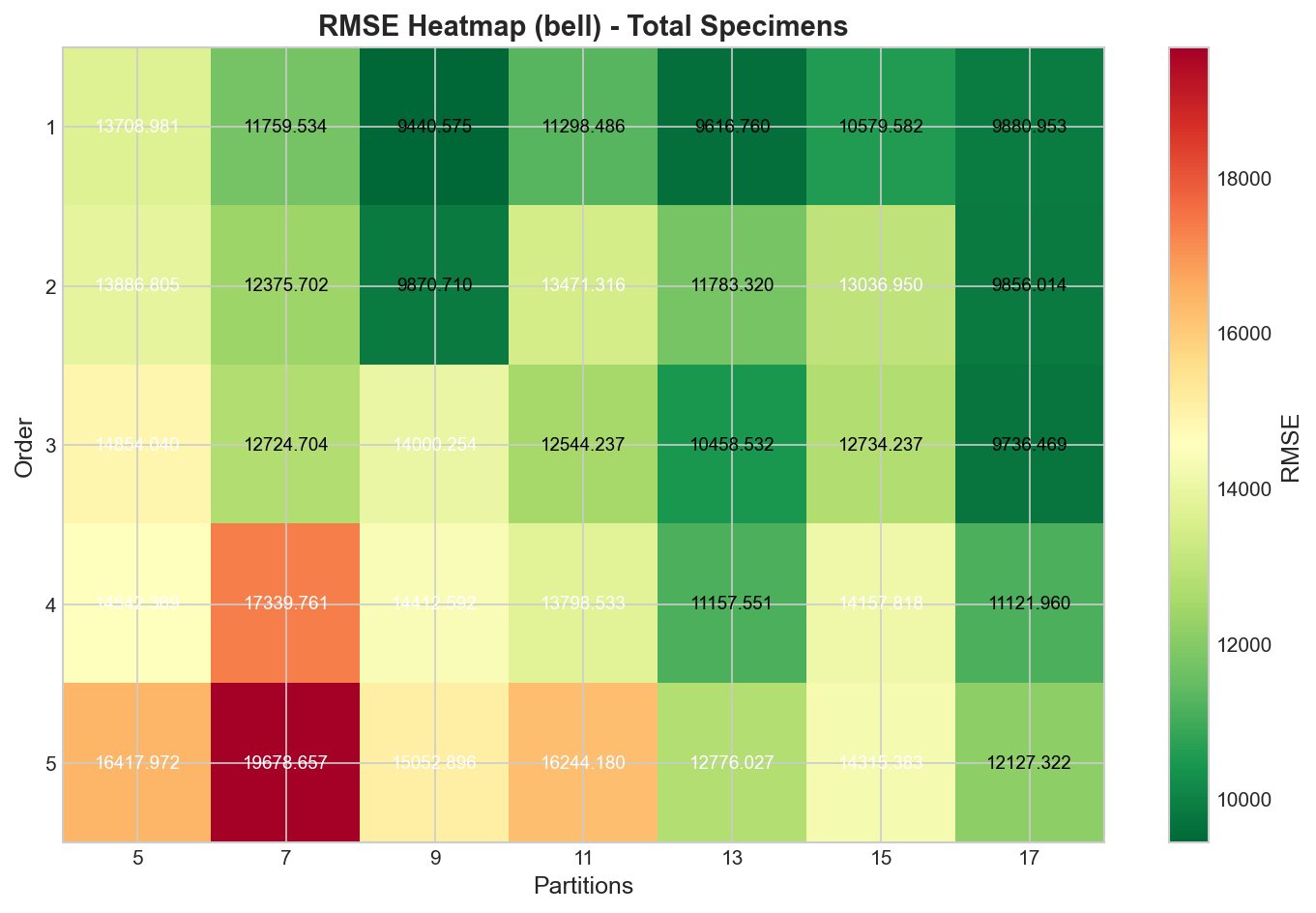
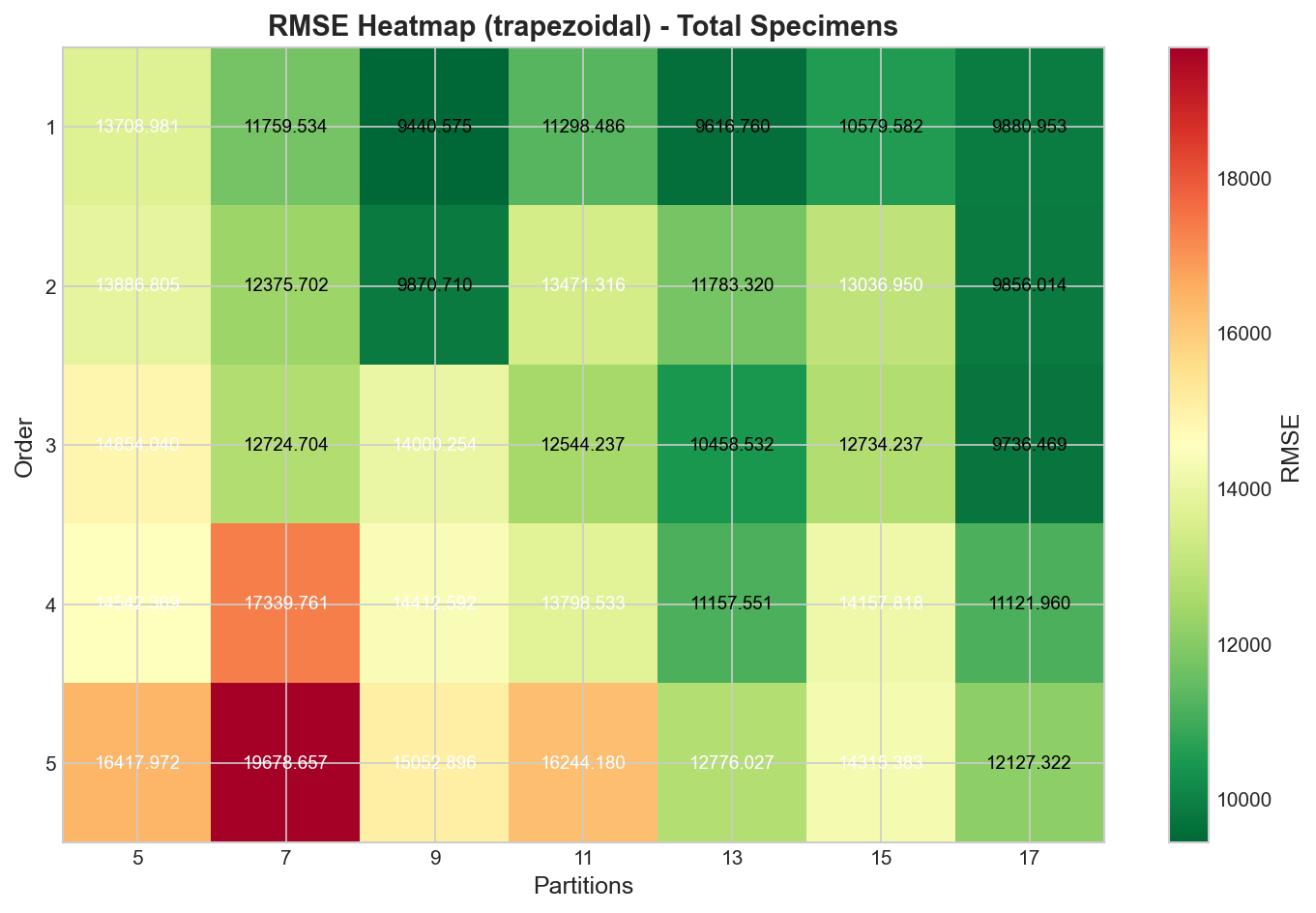
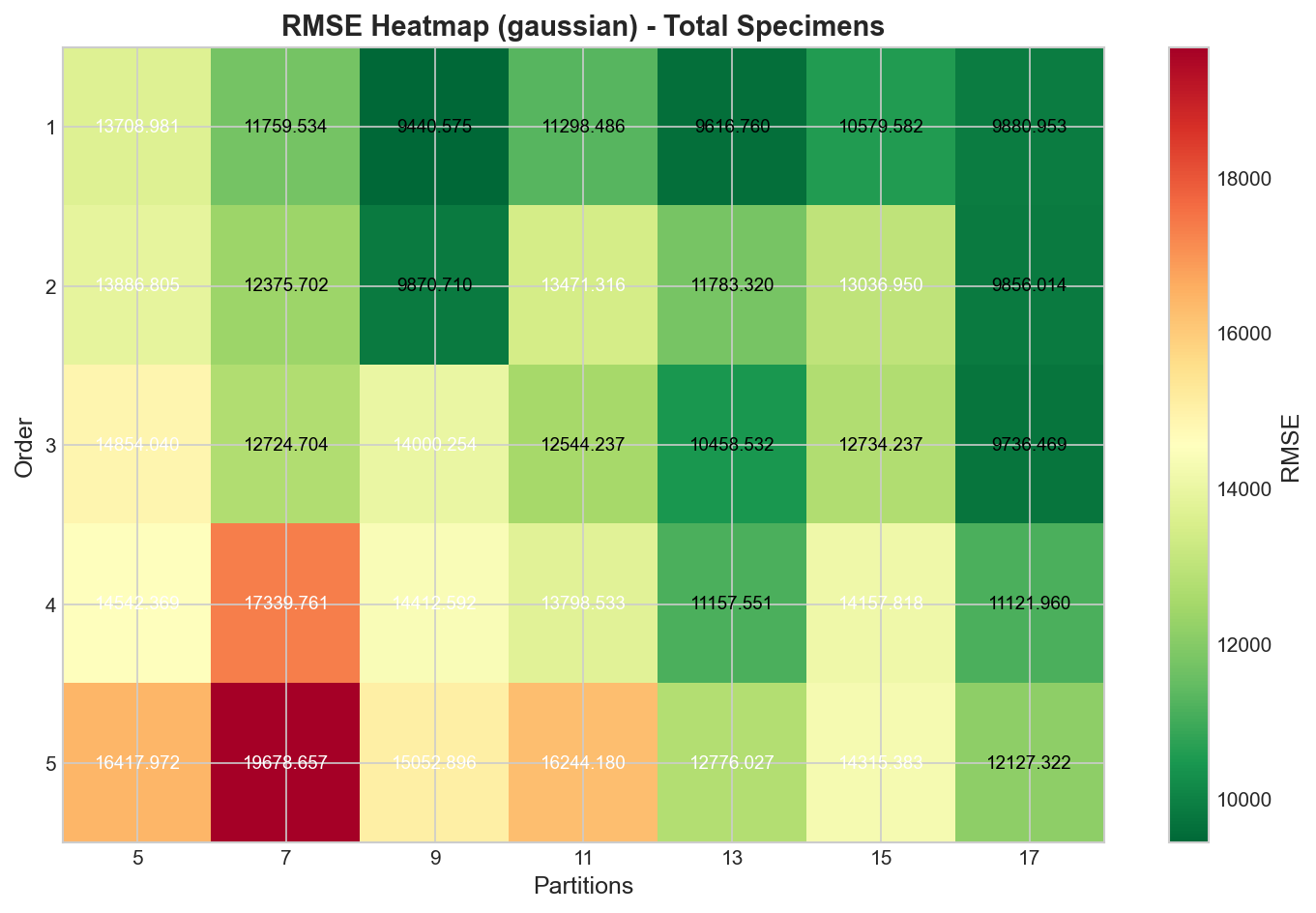
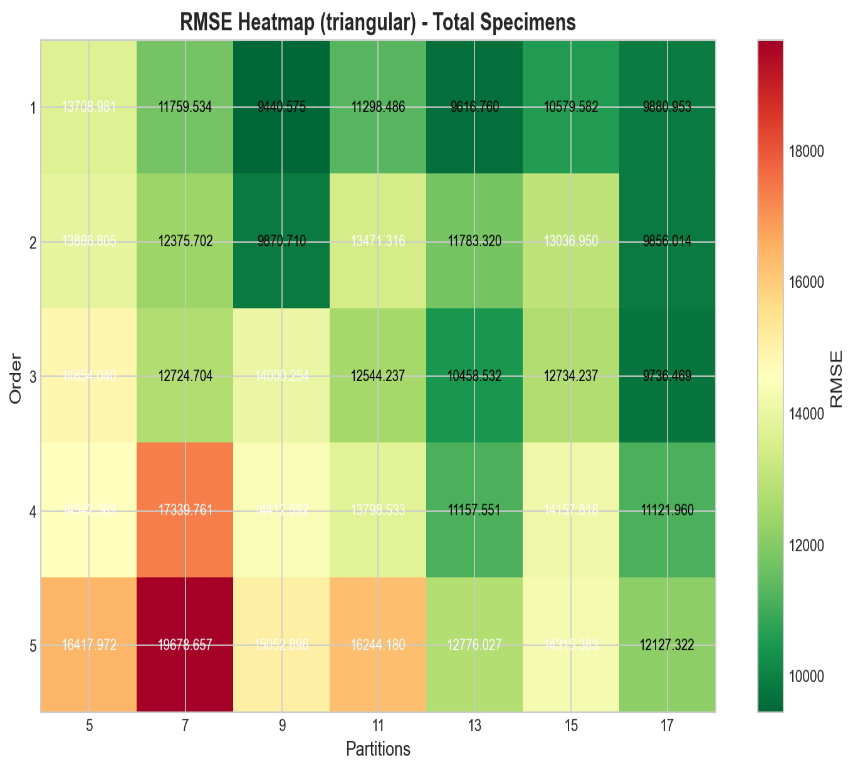
This appendix presents comprehensive performance heatmaps showing the impact of model order and partition count across all four membership function types for each dataset.

## C.1 Mackey-Glass Heatmaps



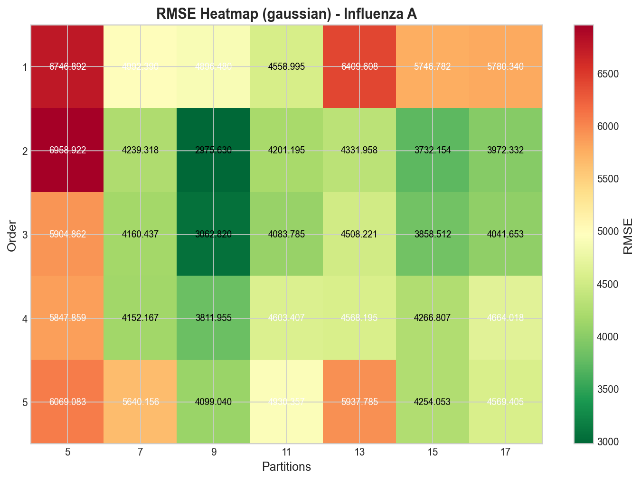
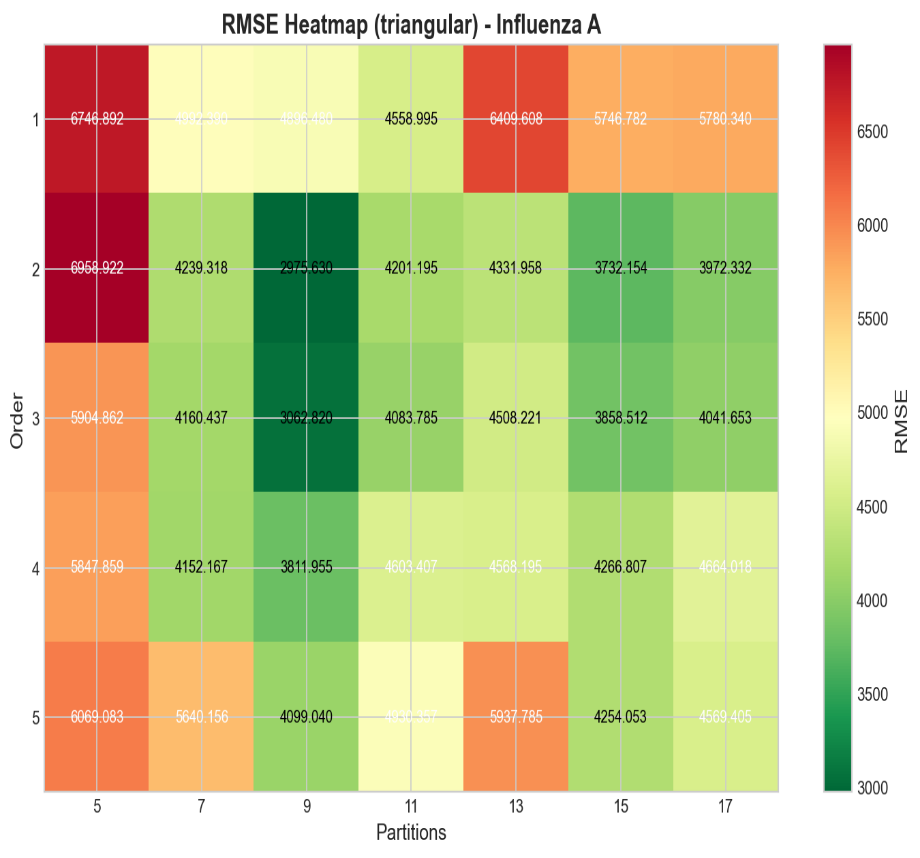
***Figure C.1:*** *Mackey-Glass RMSE heatmaps for all membership function types*

## C.2 Total Specimens Heatmaps



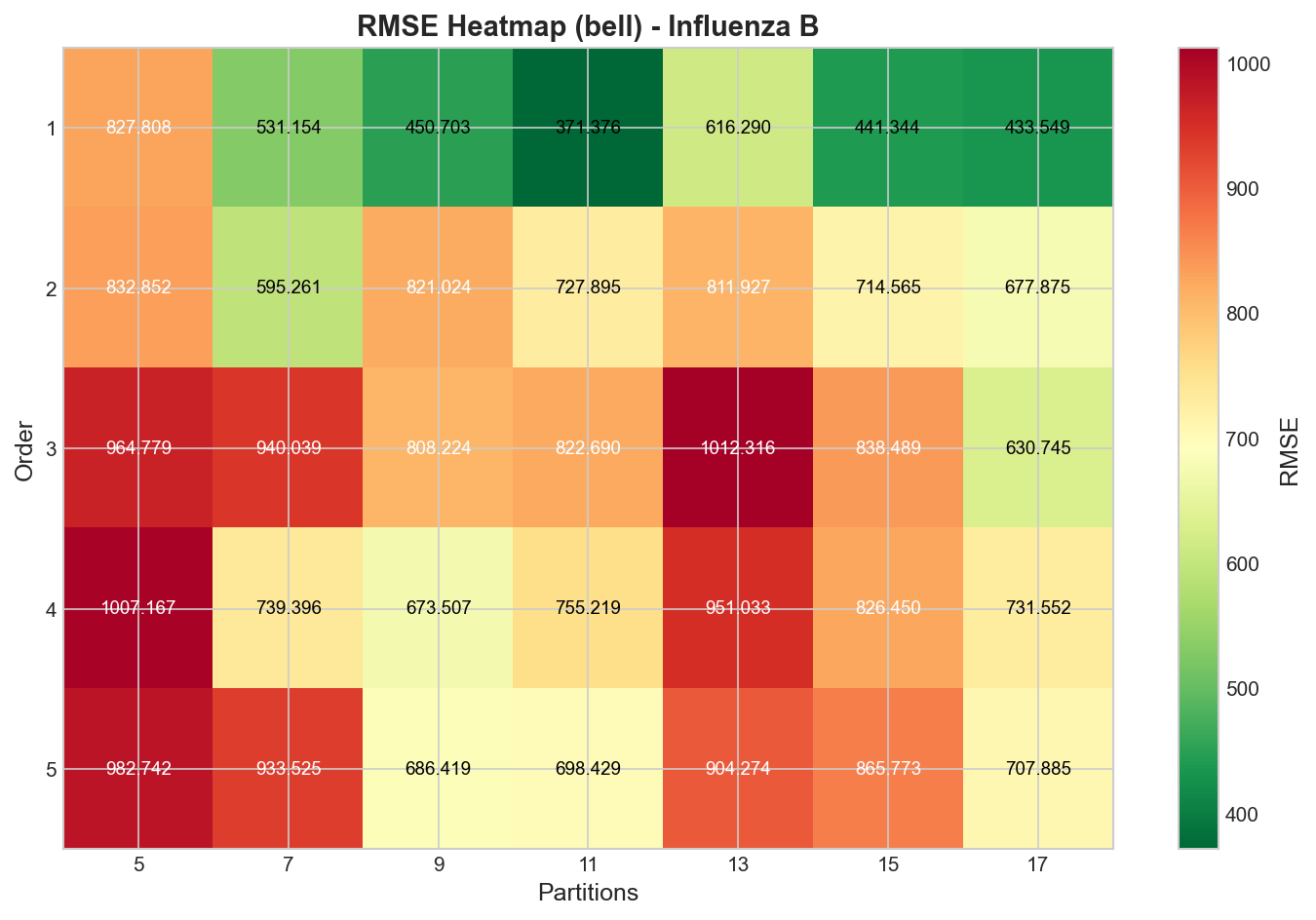
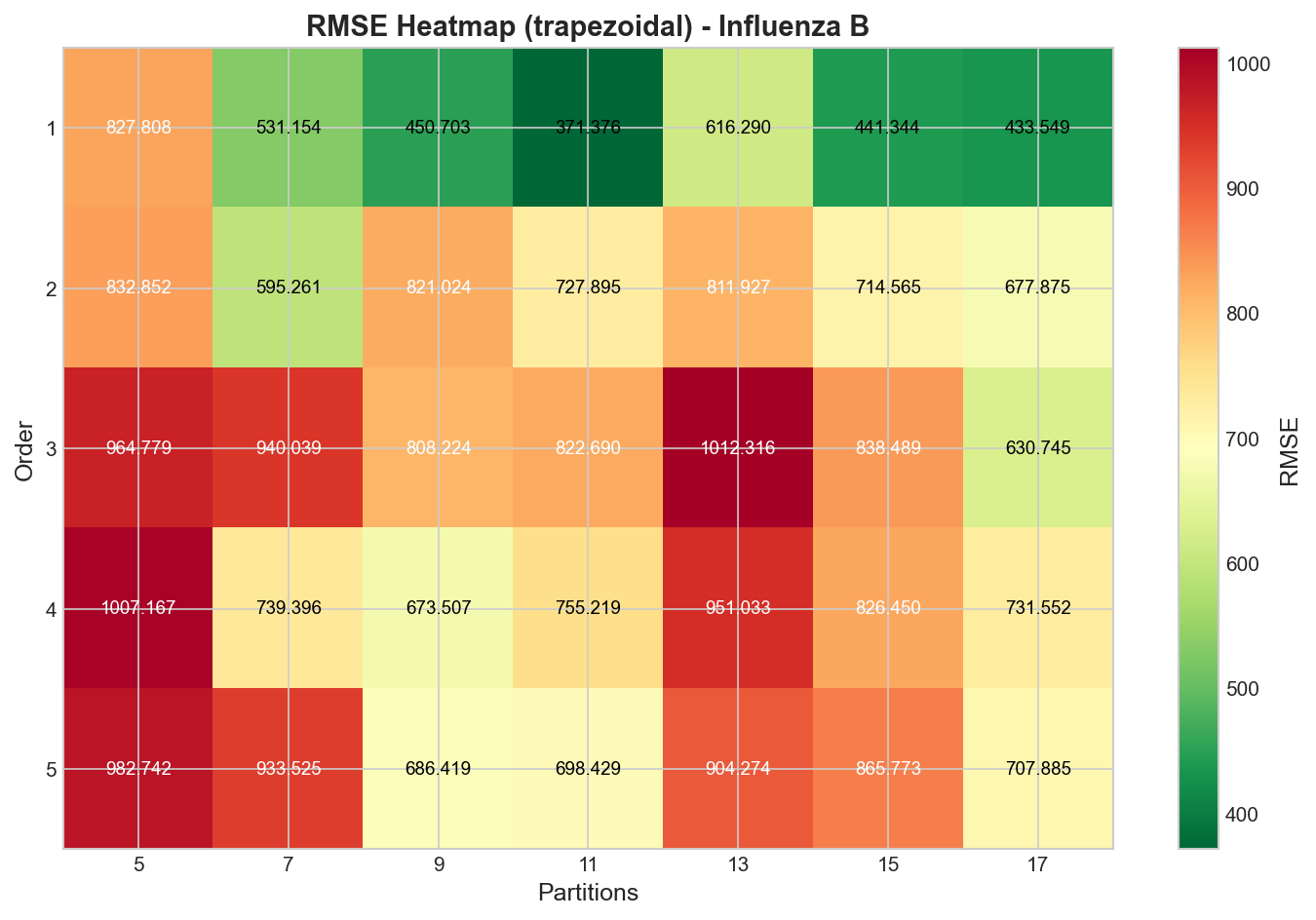
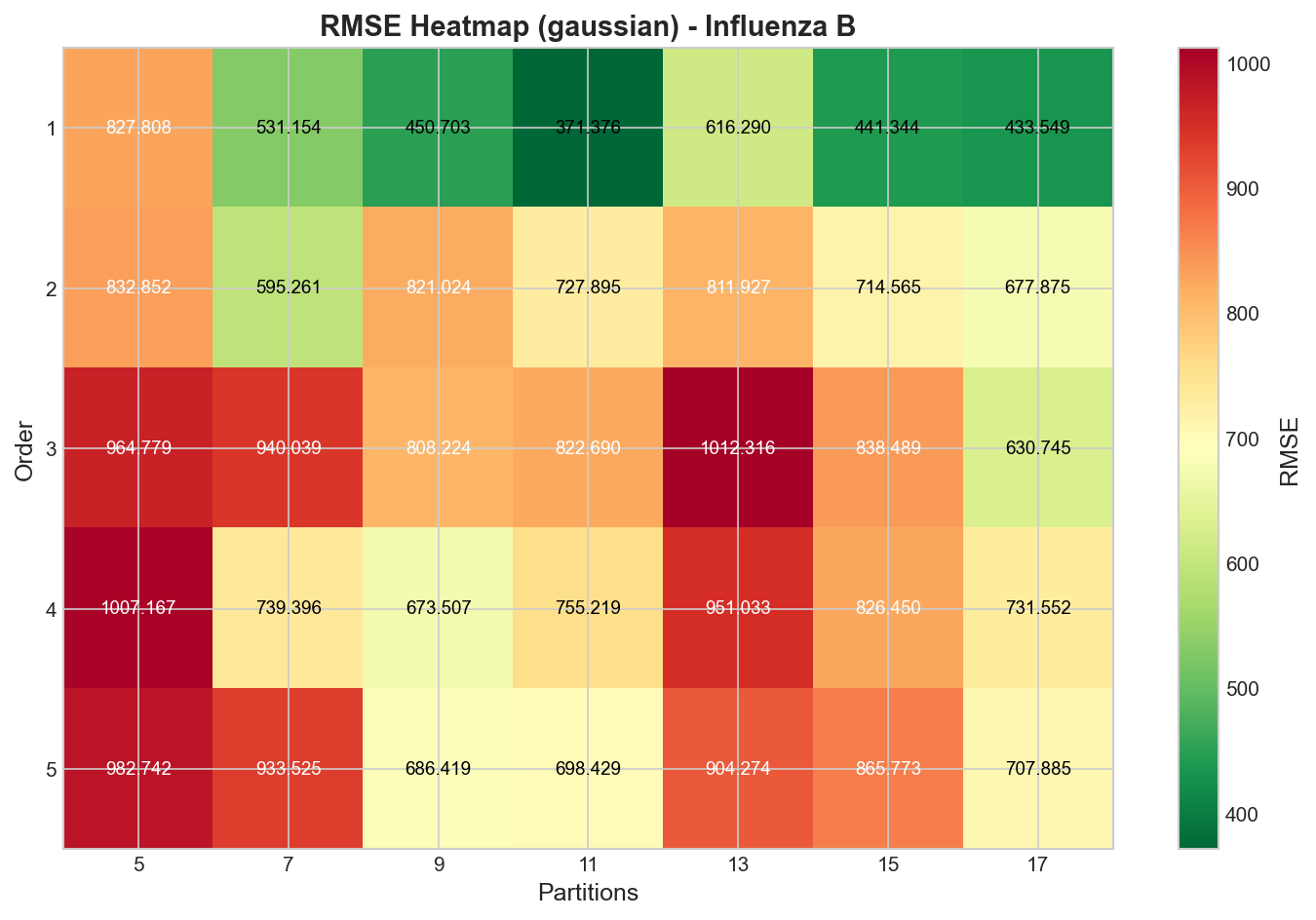
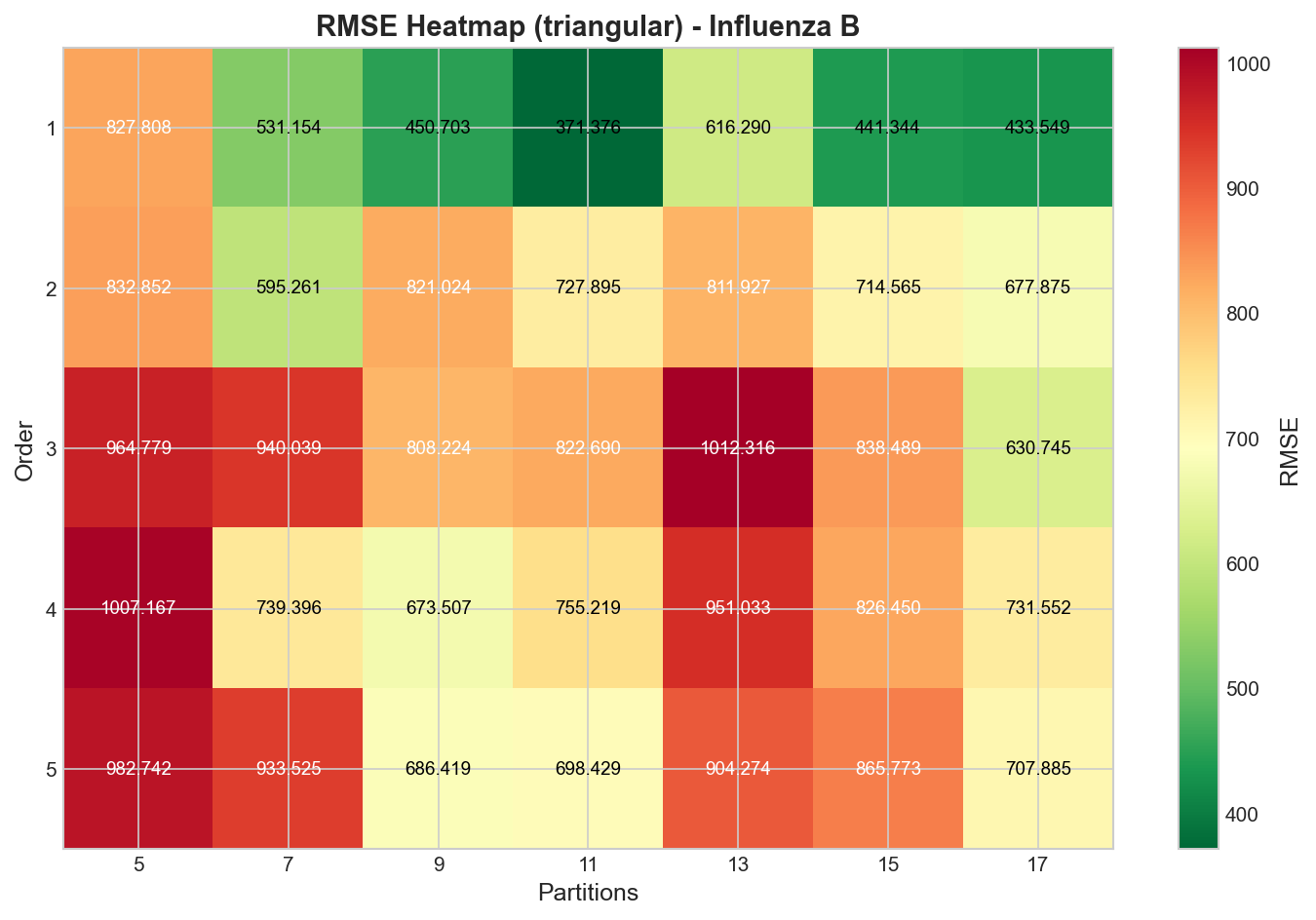
***Figure C.2:*** *Total Specimens RMSE heatmaps for all membership function types*

## C.3 Influenza A Heatmaps



***Figure C.3:*** *Influenza A RMSE heatmaps for all membership function types*

## C.4 Influenza B Heatmaps



***Figure C.4:*** *Influenza B RMSE heatmaps for all membership function types*

1. Epidemiological refers to the study of the distribution and determinants of health-related states or events in populations. [↑](#footnote-ref-1)
2. The Mackey-Glass delay differential equation is a mathematical model used to describe the dynamics of systems with time delays. It is often applied in fields such as biology and control systems to simulate processes where current states depend on past states, capturing phenomena like feedback and delays in system responses. [↑](#footnote-ref-2)
3. Epidemiological refers to the study of the distribution and determinants of health-related states or events in populations. [↑](#footnote-ref-3)