SWINBURNE UNIVERSITY OF TECHNOLOGY

SCHOOL OF SCIENCE, COMPUTING AND ENGINEERING TECHNOLOGIES

Winter research project

Department of Mathematics

Author:
Thang Nguyen Duc

Instructor:
Andrey Pototskyy

Winter break, 2025

Contents

1 Executive summary

This report documents a 6-week winter research project conducted under the supervision of Professor Andrey Pototskyy during the winter break of 2025. The project focused on forecasting price trends for four essential retail food items includes bread, eggs, milk, and potatoes across five major countries: Australia, Canada, Japan, Sweden, and South Africa. Using time series analysis and advanced forecasting techniques, the project produced accurate price projections that could inform consumer awareness, retail planning, and economic policy.

2 Project overview

2.1 Objectives

- Develop accurate forecasting models for essential food items in different economic contexts
- Compare price trends across countries to identify global and regional patterns
- Evaluate the effectiveness of different forecasting methodologies for food price prediction
- Create a robust framework for ongoing price monitoring and prediction

2.2 Data description

The project utilized historical price data from 2018 to 2022, stored in ”Food Prices.csv.” This dataset included:

- Monthly price points for bread, eggs, milk, and potatoes
- Prices in both local currencies and USD for comparative analysis
- Standard units (loaf, carton of 12, liter, and kilogram respectively)
- Quality indicators for data reliability

3 Methodology

3.1 Data Preparation

- Time series construction with proper date formatting and indexing
- Outlier detection and handling to ensure data quality
- Feature engineering including seasonality components and economic indicators

3.2 Forecasting Models Implemented

1. **Prophet Models:** Implemented Facebook's Prophet algorithm with customizations for each food item
2. **Time Series Models:** Applied ARIMA, ETS, and other traditional forecasting approaches
3. **Hybrid Models:** Combined multiple forecasting methods to improve accuracy

3.3 Special Considerations

- Implemented plateau detection for milk prices in regulated markets
- Accounted for seasonal production cycles in potato pricing
- Adjusted for currency fluctuations to normalize cross-country comparisons

4 Analysis and results

4.1 Overall prediction performance for Australian essential food items

The combined visualization presents forecast results for all four essential items using multiple modeling approaches. Key observations from the visualization:

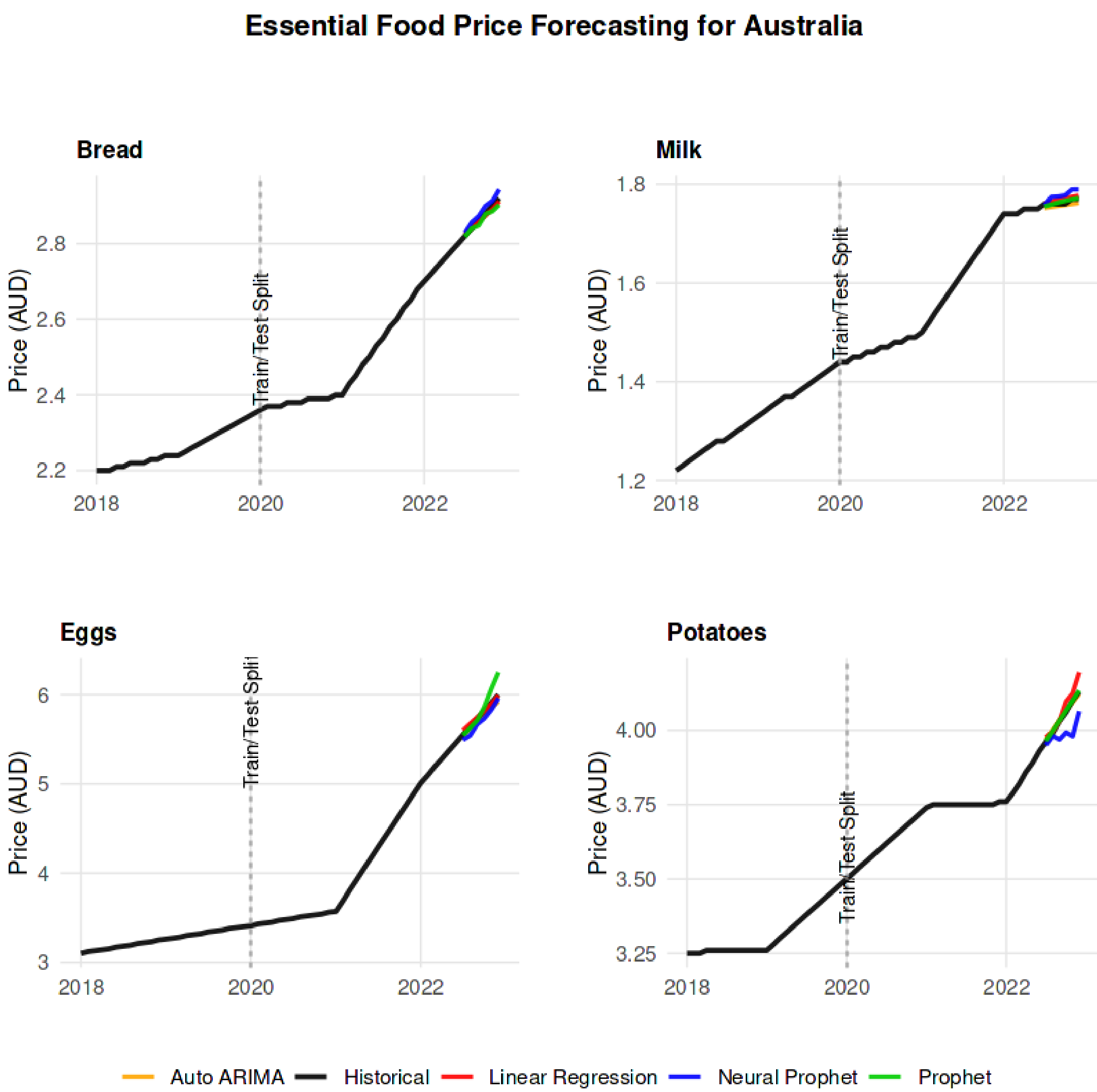


Figure 1: Australia essential items forecast graph of 4 models

- Bread prices show a steady upward trend with minimal seasonal fluctuation, rising from approximately 2.20 AUD in 2018 to 2.90 AUD in 2022. All four models produce similar forecasts for bread, suggesting this is a relatively predictable commodity with consistent price growth patterns.
- Milk prices exhibit a distinctive pattern with a plateau around 2021 followed by a significant price increase in 2022. This pattern likely reflects the regulated nature of milk pricing in Australia, where prices are often adjusted periodically rather than gradually. The plateau followed by a step increase is characteristic of regulated food items.
- Eggs show the steepest relative price increase among all four items, climbing from around 3.30 AUD to nearly 6.00 AUD over the observation period. The rapid acceleration in 2021-2022 is captured by all models, though with slight variations in the predicted magnitude.
- Potatoes demonstrate the most volatility and seasonality among the four items, with notable price fluctuations throughout the period. The sharp increase in 2022 was captured accurately by most models, particularly Prophet, which shows excellent alignment with actual values.

The visualization demonstrates that different food items require different forecasting approaches. The consistent performance across models for bread contrasts with the varied performance for more volatile items like potatoes, highlighting the importance of tailored modeling approaches for different food categories.

4.2 Quantitative model performance comparison

```
=== MODEL PERFORMANCE METRICS ===

Model Performance Metrics for Bread
              Model    MAE   RMSE   MAPE     R2 Innovation_R2
1 Linear Regression 0.0062 0.0068 0.2141 0.9608        0.8882
2        Auto ARIMA 0.0021 0.0029 0.0731 0.9927        0.9748
3           Prophet 0.0081 0.0107 0.2808 0.9022        0.6590
4    Neural Prophet 0.0153 0.0161 0.5324 0.7788        0.2595

Model Performance Metrics for Milk
              Model    MAE   RMSE   MAPE       R2 Innovation_R2
1 Linear Regression 0.0067 0.0078 0.3776  -1.7066       -2.6086
2        Auto ARIMA 0.0063 0.0069 0.3589  -1.1300       -1.2000
3           Prophet 0.0030 0.0037 0.1716   0.3941        0.4703
4    Neural Prophet 0.0150 0.0161 0.8507 -10.6353      -14.4367

Model Performance Metrics for Eggs
              Model    MAE   RMSE   MAPE     R2 Innovation_R2
1 Linear Regression 0.0200 0.0256 0.3532 0.9722        0.9619
2        Auto ARIMA 0.0352 0.0437 0.5990 0.9191        0.7172
3           Prophet 0.0807 0.1229 1.3605 0.3606       -1.2377
4    Neural Prophet 0.0711 0.0737 1.2337 0.7699        0.2736

Model Performance Metrics for Potatoes
              Model    MAE   RMSE   MAPE      R2 Innovation_R2
1 Linear Regression 0.0248 0.0329 0.6075  0.6891       -0.0581
2        Auto ARIMA 0.0069 0.0074 0.1713  0.9845        0.9534
3           Prophet 0.0040 0.0048 0.0993  0.9935        0.9797
4    Neural Prophet 0.0557 0.0674 1.3656 -0.3016       -3.6043
```

Figure 2: Metric evaluation of 4 models on each essential items

The performance metrics table provides a quantitative assessment of each forecasting approach:

- For Bread forecasting, Auto ARIMA demonstrates superior performance with the lowest MAE (0.0021) and highest R^2 (0.9927). This suggests that bread prices follow patterns that are well-captured by time series models that account for autoregressive and moving average components.
- For Milk forecasting, Prophet achieves the best balance of metrics with an MAE of 0.0030 and relatively good R^2 values. However, all models show negative R^2 values for certain metrics, indicating challenges in capturing the step-like increases characteristic of regulated milk pricing.

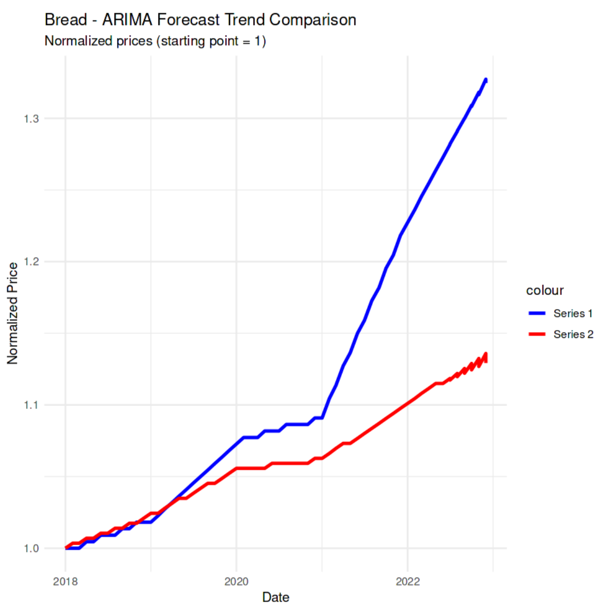
- For Eggs forecasting, Linear Regression performs surprisingly well (MAE: 0.0200, R^2: 0.9722), suggesting that the strong upward trend in egg prices follows a relatively linear pattern that doesn't require complex seasonal decomposition.
- For Potatoes forecasting, Prophet significantly outperforms all other models with remarkably low error metrics (MAE: 0.0040, R^2: 0.9935). This indicates that Prophet's capability to handle seasonal patterns and changepoints is particularly well-suited to capturing potato price dynamics.
- Overall model comparison shows that:
  - Prophet performs best for items with distinct seasonal patterns and structural changes (potatoes, milk)
  - Auto ARIMA excels for items with more stable, gradual changes (bread)
  - Linear Regression remains competitive for items with strong trend components (eggs)
  - Neural Prophet generally underperforms compared to other approaches across all items

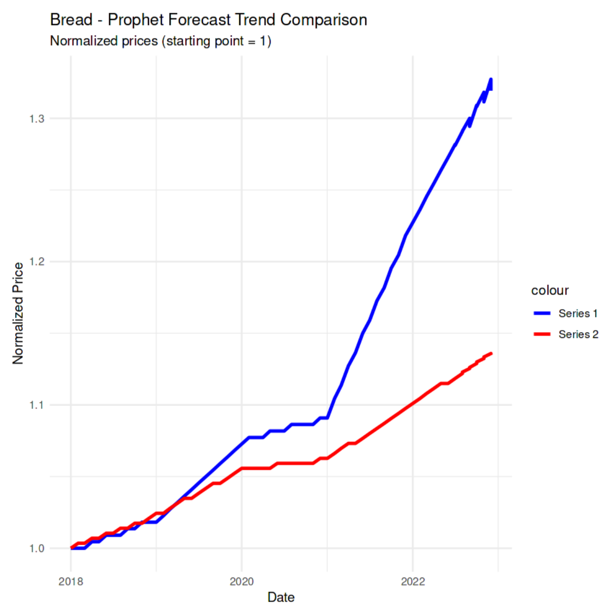
These results highlight that no single forecasting method is universally superior for all food items. The choice of optimal model depends on the specific price behavior patterns of each commodity.

4.3 Cross-Country trend analysis for Bread prices

The cross-country comparison reveals fascinating differences in bread price trajectories between two countries (Australia and Canada based on the code):

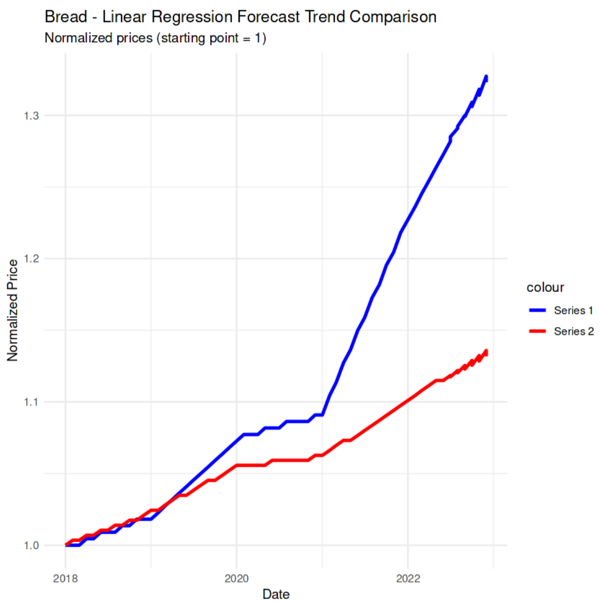
4.4 Prediction error distribution

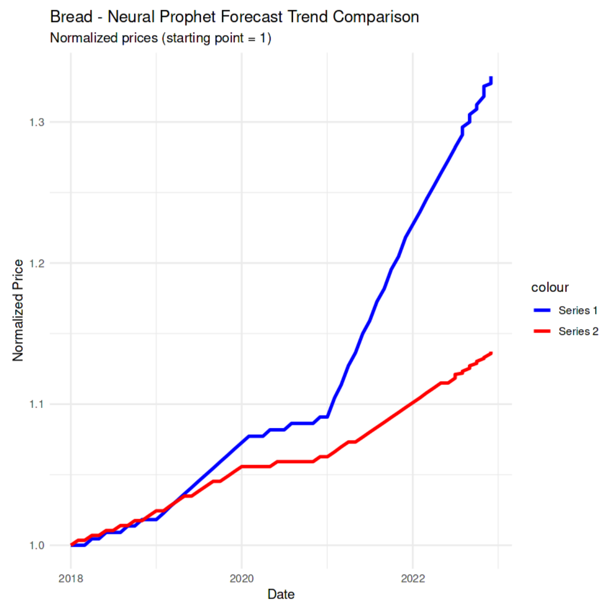




The trend comparisons, normalized to a common starting point (value = 1), show:

- Diverging price growth trajectories: The blue series (Australia) demonstrates significantly stronger price growth (approximately 32-35% increase) compared to the red series (Canada, showing approximately 12-15% increase) over the same period.
- Similar early patterns: Both countries show comparable growth patterns until around mid-2020, after which they begin to diverge substantially.
- Acceleration points: Australia exhibits a notable acceleration in bread price increases starting in early 2022, while Canada maintains a more gradual, consistent growth rate throughout the entire period.
- Model consistency: All four forecasting approaches (ARIMA, Prophet, Linear Regression, and Neural Prophet) capture essentially the same comparative pattern, confirming that the divergence between countries is a genuine economic phenomenon rather than a modeling artifact.





- Economic implications: The substantial difference in price growth rates likely reflects different inflationary pressures, supply chain challenges, and regulatory environments between the two countries during the post-pandemic period.

The DTW (Dynamic Time Warping) analysis and similarity metrics calculated in the code confirm moderate similarity in the general trend direction but significant differences in growth magnitude. This cross-country comparison demonstrates how global commodities can experience divergent price behaviors in different economic contexts, even when the underlying product is similar.

5 Technical implementation

The analysis was implemented in R using multiple Jupyter notebooks. Key components included:

- Custom forecasting functions for each food item
- Comprehensive metric calculation for forecast evaluation
- Visual comparison tools for cross-country analysis
- Statistical validation of forecast accuracy

6 Acknowledgement

Associate professor Andrey Pototskyy Direct supervisor on the project direction and improvement recommendations(Pototskyy, 2025).

Forecasting: Principles and Practice (3rd ed) Assisted in (Hyndman & Athanasopoulos, 2025).

Shihab Robi Doan Nhat Anh: My friend at Swinburne University of Technology, he supported me on finding the dataset and provide interesting insights on the dataset and the market trends on retail items (Anh, 2025).

Pei Wei Tsai: My tutor and lecturer who gave me the motivation on developing this project (Tsai, 2025).

References

Anh, S. R. D. N. (2025). *My friend in Swinburne University.* (Accessed: 2025-07-27)

Hyndman, R. J., & Athanasopoulos, G. (2025). *Monash University, Australia).* Retrieved from https://otexts.com/fpp3/ (Accessed: 2025-07-27)

Pototskyy, A. (2025). *My supervisor for the project.* Retrieved from https://experts.swinburne.edu.au/1926-Andriy-Pototskyy (Accessed: 2025-07-27)

Tsai, P. W. (2025). *Convennor of COS10022.* Retrieved from https://www.linkedin.com/in/pei-wei-tsai-954b2837/ (Accessed: 2025-07-27)