Time Series Analysis and Forecasting

Seminar 9: Prophet and TBATS

Daniel Traian PELE

Bucharest University of Economic Studies
IDA Institute Digital Assets
Blockchain Research Center
AI4EFin Artificial Intelligence for Energy Finance
Romanian Academy, Institute for Economic Forecasting
MSCA Digital Finance

Seminar Outline

- **Multiple Choice Quiz** – Knowledge check
- **True/False** – Conceptual checks
- **Calculation Exercises** – Applied practice
- **Worked Examples** – Detailed solutions
- **AI-Assisted Exercise** – Critical thinking
- **Summary** – Key takeaways

Quiz 1: Multiple Seasonality Challenge

Question

Why can't standard SARIMA handle hourly electricity demand data?

Answer choices

(A) SARIMA can only handle monthly data

(B) SARIMA allows only one seasonal period (m parameter)

(C) SARIMA doesn't support trend components

(D) SARIMA requires normally distributed data

Answer on next slide...

Quiz 1: Answer

Answer: B – SARIMA allows only one seasonal period

Question: Why can't standard SARIMA handle hourly electricity demand data?

Answer choices

(A) SARIMA can only handle monthly data ×

(B) **SARIMA allows only one seasonal period (*m* parameter)** ✓

(C) SARIMA doesn't support trend components ×

(D) SARIMA requires normally distributed data ×

- Hourly data has daily (24h), weekly (168h), and annual (8760h) patterns
- SARIMA's single m parameter cannot capture all these simultaneously
- Use TBATS or Prophet for **multiple seasonality**

TSA_ch9_quiz1_multiple_seasonality

Quiz 2: TBATS Acronym

Question

What does TBATS stand for?

Answer choices

(A) Trend, Baseline, ARMA, Transform, Seasonal

(B) Trigonometric, Box-Cox, ARMA, Trend, Seasonal

(C) Time-Based Automatic Time Series

(D) Temporal Bayesian Adaptive Trend System

Answer on next slide...

Quiz 2: Answer

Answer: B – Trigonometric, Box-Cox, ARMA, Trend, Seasonal

Question: What does TBATS stand for?

Answer choices

(A) Trend, Baseline, ARMA, Transform, Seasonal ×

(B) **Trigonometric, Box-Cox, ARMA, Trend, Seasonal** ✓

(C) Time-Based Automatic Time Series ×

(D) Temporal Bayesian Adaptive Trend System ×

- **T**rigonometric (Fourier for seasonality), **B**ox-Cox (variance stabilization)
- **A**RMA (error autocorrelation), **T**rend (damped local), **S**easonal (multiple periods)

TSA_ch9_quiz2_tbats_components

Quiz 3: Fourier Terms

Question

In TBATS, increasing the number of Fourier harmonics (K) for a seasonal pattern:

Answer choices

(A) Always improves forecast accuracy

(B) Allows more flexible (complex) seasonal shapes

(C) Reduces the model complexity

(D) Eliminates the need for Box-Cox transformation

Answer on next slide...

Quiz 3: Answer

Answer: B – Allows more flexible seasonal shapes

Question: In TBATS, increasing the number of Fourier harmonics (K) for a seasonal pattern:

Answer choices

(A) Always improves forecast accuracy ×

(B) **Allows more flexible (complex) seasonal shapes ✓**

(C) Reduces the model complexity ×

(D) Eliminates the need for Box-Cox transformation ×

- **Trade-off**: More harmonics = more flexibility but also more parameters
- $s_t^{(i)} = \sum_{j=1}^{K_i} \left[a_j^{(i)} \cos\left(\frac{2\pi jt}{m_i}\right) + b_j^{(i)} \sin\left(\frac{2\pi jt}{m_i}\right) \right]$

TSA_ch9_quiz3_fourier_harmonics

Quiz 4: Prophet Decomposition

Question

Prophet decomposes a time series into which components?

Answer choices

(A) AR, MA, and seasonal components

(B) Trend, seasonality, holidays, and error

(C) Mean, variance, and autocorrelation

(D) Level, slope, and curvature

Answer on next slide...

Quiz 4: Answer

Answer: B – Trend, seasonality, holidays, and error

Question: Prophet decomposes a time series into which components?

Answer choices

(A) AR, MA, and seasonal components ×

(B) **Trend, seasonality, holidays, and error** ✓

(C) Mean, variance, and autocorrelation ×

(D) Level, slope, and curvature ×

- $y(t) = g(t) + s(t) + h(t) + \varepsilon_t$
- $g(t)$ = trend, $s(t)$ = seasonality (Fourier), $h(t)$ = holidays, ε_t = error

TSA_ch9_quiz4_prophet_decomposition

Quiz 5: Prophet vs TBATS

Question

When would you choose Prophet over TBATS?

Answer choices

(A) When you need automatic model selection

(B) When you have known holidays and changepoints to incorporate

(C) When you need the most parsimonious model

(D) When your data has no trend

Answer on next slide...

Quiz 5: Answer

Answer: B – Known holidays and changepoints

Question: When would you choose Prophet over TBATS?

Answer choices

(A) When you need automatic model selection ×

(B) **When you have known holidays and changepoints to incorporate ✓**

(C) When you need the most parsimonious model ×

(D) When your data has no trend ×

- **Prophet**: Easy holiday integration, analyst-in-the-loop, handles missing data, interpretable
- **TBATS**: Automatic model selection, handles complex seasonality without domain expertise

TSA_ch9_quiz5_prophet_vs_tbats

Quiz 6: Seasonality Mode

Question

For retail sales data where December sales are 3x the monthly average, which seasonality mode is more appropriate in Prophet?

Answer choices

(A) Additive seasonality

(B) Multiplicative seasonality

(C) Both work equally well

(D) Neither—use ARIMA instead

Answer on next slide...

Quiz 6: Answer

Answer: B – Multiplicative seasonality

Answer choices

(A) Additive seasonality ×

(B) **Multiplicative seasonality** ✓

(C) Both work equally well ×

(D) Neither—use ARIMA instead ×

- When seasonal amplitude scales with the level, use multiplicative
- **Additive**: $y = g(t) + s(t)$ (constant seasonal effect)
- **Multiplicative**: $y = g(t) \cdot (1 + s(t))$ (proportional seasonal effect)

TSA_ch9_additive_vs_multiplicative

Quiz 7: Prophet Changepoints

Question

In Prophet, changepoints allow the model to:

Answer choices

(A) Change the seasonal period automatically

(B) Adjust the trend slope at specific points in time

(C) Switch between additive and multiplicative modes

(D) Detect and remove outliers

Answer on next slide...

Quiz 7: Answer

Answer: B – Adjust trend slope at specific points

Question: In Prophet, changepoints allow the model to:

Answer choices

(A) Change the seasonal period automatically ×

(B) **Adjust the trend slope at specific points in time** ✓

(C) Switch between additive and multiplicative modes ×

(D) Detect and remove outliers ×

- Changepoints allow piecewise linear trend with different slopes
- $g(t) = (k + \mathrm{a}(t)^\top \delta) \cdot t + (m + \mathrm{a}(t)^\top \gamma)$
- Prophet automatically detects changepoints or you can specify them manually

TSA_ch9_changepoint_detection

Quiz 8: Model Selection

Question

You have daily call center data with weekly seasonality only. Which model is most appropriate?

Answer choices

(A) TBATS (designed for multiple seasonality)

(B) Prophet (handles any seasonality well)

(C) Standard SARIMA (simpler and sufficient)

(D) LSTM neural network (most flexible)

Answer on next slide...

Quiz 8: Answer

Answer: C – Standard SARIMA is sufficient

Question: You have daily call center data with weekly seasonality only. Which model is most appropriate?

Answer choices

(A) TBATS (designed for multiple seasonality) ×

(B) Prophet (handles any seasonality well) ×

(C) **Standard SARIMA (simpler and sufficient)** ✓

(D) LSTM neural network (most flexible) ×

- **Principle of parsimony**: Use the simplest model that fits the data
- With only weekly seasonality ($m = 7$), SARIMA works fine
- Use TBATS/Prophet when you *need* multiple seasonalities or special features

Quiz 9: Prophet Uncertainty

Question

Prophet generates prediction intervals by:

Answer choices

(A) Assuming normally distributed residuals

(B) Sampling from the posterior distribution of parameters

(C) Using bootstrap resampling of historical errors

(D) Applying a fixed multiplier to point forecasts

Answer on next slide...

Quiz 9: Answer

Answer: B – Sampling from posterior distribution

Question: Prophet generates prediction intervals by:

Answer choices

(A) Assuming normally distributed residuals ×

(B) **Sampling from the posterior distribution of parameters** ✓

(C) Using bootstrap resampling of historical errors ×

(D) Applying a fixed multiplier to point forecasts ×

- Prophet uses **Bayesian estimation**: MAP for point forecasts, MCMC/simulation for intervals
- Uncertainty from both trend (changepoints) and observation noise

Note: by default, Prophet uses MAP estimation. Prediction intervals come from simulating future trend changepoints and adding historical residual variability, not from full posterior sampling (MCMC is optional via `mcmc_samples > 0`).

Quiz 10: Practical Application

Question

For forecasting hourly energy demand with daily, weekly, and annual patterns plus holiday effects, which approach is best?

Answer choices

(A) SARIMA with $m = 24$

(B) TBATS with three seasonal periods

(C) Prophet with custom holidays

(D) Either TBATS or Prophet, depending on whether holidays are important

Answer on next slide...

Quiz 10: Answer

Answer: D – TBATS or Prophet depending on needs

Answer choices

(A) SARIMA with $m = 24$ ×

(B) TBATS with three seasonal periods ×

(C) Prophet with custom holidays ×

(D) **Either TBATS or Prophet, depending on whether holidays are important** ✓

- Both handle multiple seasonality
- Holiday effects crucial $\Rightarrow$ Prophet; automatic selection $\Rightarrow$ TBATS
- Often try both and compare via cross-validation

TSA_ch9_electricity_demand

True or False? — Questions

Statement	T/F?
1. Prophet was developed by Facebook (Meta) for business forecasting.	?
2. TBATS can only handle two seasonal periods at most.	?
3. In Prophet, the default trend is logistic growth.	?
4. Fourier terms approximate seasonality using sine and cosine functions.	?
5. Prophet requires equally spaced time series data.	?
6. The Box-Cox transformation in TBATS stabilizes variance.	?

True or False? — Answers

Statement	T/F	Explanation
1. Prophet was developed by Facebook (Meta) for business forecasting.	T	**Released 2017, analyst-in-the-loop**
2. TBATS can only handle two seasonal periods at most.	F	**Any number of periods**
3. In Prophet, the default trend is logistic growth.	F	**Default is piecewise linear**
4. Fourier terms approximate seasonality using sine and cosine functions.	T	$\sum[a_k \cos + b_k \sin]$
5. Prophet requires equally spaced time series data.	F	**Handles missing/irregular**
6. The Box-Cox transformation in TBATS stabilizes variance.	T	$y^{(\lambda)}$ **transformation**

Exercise 1: Fourier Terms Calculation

Problem

- **Data**: Daily data with weekly seasonality ($m = 7$), using $K = 3$ Fourier harmonics
- **Calculate**: How many parameters does this add to the model?
- **Formula**: $s(t) = \sum_{k=1}^{K} \left[a_k \cos\left(\frac{2\pi kt}{m}\right) + b_k \sin\left(\frac{2\pi kt}{m}\right)\right]$

Solution

- Each harmonic requires 2 parameters (sine and cosine coefficients)
- $k = 1$: a_1, b_1 (fundamental frequency 1/7 cycles per day)
- $k = 2$: a_2, b_2 (first overtone, 2/7 cycles per day)
- $k = 3$: a_3, b_3 (second overtone, 3/7 cycles per day)
- **Total**: $2 \times K = 2 \times 3 = \mathbf{6}$ parameters

TSA_ch9_fourier_approximation

Exercise 2: Choosing Seasonality Mode

Problem

- **Data**: Monthly hotel bookings — July 2020: 1000, Jan 2020: 400, July 2023: 2000, Jan 2023: 800
- **Question**: Should you use additive or multiplicative seasonality? Why?

Solution

Year	July	January	Ratio (Jul/Jan)
2020	1000	400	2.5
2023	2000	800	2.5

- The *ratio* stays constant (2.5), not the difference!
- Additive: $1000 - 400 = 600$ vs $2000 - 800 = 1200$ (not constant)
- **Conclusion**: Use multiplicative: `seasonality_mode='multiplicative'`

TSA_ch9_additive_vs_multiplicative

Exercise 3: TBATS Model Interpretation

Problem

- **TBATS model**: Box-Cox $\lambda = 0.5$, seasonal periods $m_1 = 24$, $m_2 = 168$, Fourier terms $K_1 = 5$, $K_2 = 3$
- **Calculate**: What does each component tell you? Total seasonal parameters?

Solution

- **Box-Cox** $\lambda = 0.5$: Square root transformation ($y^{(0.5)} = \sqrt{y}$), data had increasing variance
- $m_1 = 24$: Daily pattern (24 hours); $m_2 = 168$: Weekly pattern (7×24)
- $K_1 = 5$: Complex intraday pattern (5 harmonics for peaks, valleys)
- $K_2 = 3$: Simpler weekly pattern (weekday vs weekend)
- **Total seasonal parameters**: $2(K_1 + K_2) = 2(5 + 3) = \mathbf{16}$

TSA_ch9_tbats_decomposition

Exercise 4: Prophet Holiday Effects

Problem

- **Task**: Forecast daily restaurant revenue with holiday effects
- **Holidays**: Valentine's Day (Feb 14, boost), Easter (variable, closed), Christmas (Dec 25, closed)
- **Write**: Python code to create the holidays dataframe for 2024–2025

Solution

```
holidays = pd.DataFrame({
'holiday': ['valentines', 'valentines', 'easter', 'easter',
'christmas', 'christmas'],
'ds': pd.to_datetime(['2024-02-14', '2025-02-14',
'2024-03-31', '2025-04-20', '2024-12-25', '2025-12-25']),
'lower_window': [-1, -1, 0, 0, -1, -1],
'upper_window': [0, 0, 0, 0, 0, 0]})
model = Prophet(holidays=holidays)
```

TSA_ch9_prophet_components

Visual: Retail Sales Forecasting with Prophet

Scenario

Monthly retail sales (2018-2023): December peaks, COVID-19 break in 2020, growing trend.

Prophet Configuration

```
model = Prophet(seasonality_mode='multiplicative',
changepoint_prior_scale=0.5, yearly_seasonality=True)
model.add_country_holidays(country_name='US')
```

Key Decision

Multiplicative seasonality: December effect proportional to baseline level.

TSA_ch9_retail_sales

Visual: Energy Demand with TBATS

Scenario

Hourly electricity: intraday (24h), weekly (168h), annual (8760h) patterns.

TBATS in R

```
library(forecast)
energy_msts <- msts(energy_data, seasonal.periods = c(24, 168, 8760))
fit <- tbats(energy_msts); fc <- forecast(fit, h = 168)
```

Note

TBATS automatically selects K for each seasonal period via AIC.

TSA_ch9_electricity_demand

Visual: Cross-Validation Comparison

Objective

Compare Prophet, TBATS, and SARIMA on 2 years of daily sales data.

Prophet Cross-Validation

```
from prophet.diagnostics import cross_validation, performance_metrics
df_cv = cross_validation(model, initial='365 days',
period='90 days', horizon='30 days')
metrics = performance_metrics(df_cv)
print(f"MAPE: {metrics['mape'].mean():.2%}")
```

Typical Results

Model	MAPE	Computation Time
SARIMA (weekly only)	8.5%	Fast
TBATS (weekly + yearly)	6.2%	Moderate
Prophet (weekly + yearly + holidays)	5.8%	Fast

TSA_ch9_prophet_vs_tbats

Discussion: When to Use Which Model?

Key Question

You have a new forecasting task. How do you choose between SARIMA, TBATS, and Prophet?

Decision Framework

1. **How many seasonal periods?**
   - One $\Rightarrow$ SARIMA may suffice
   - Multiple $\Rightarrow$ TBATS or Prophet
2. **Do you have domain knowledge to encode?**
   - Holidays, events, changepoints $\Rightarrow$ Prophet
   - Let the data speak $\Rightarrow$ TBATS
3. **Interpretability requirements?**
   - Need to explain components $\Rightarrow$ Prophet
   - Just need forecasts $\Rightarrow$ Either

TSA_ch9_model_selection_guide

Discussion: Overfitting with Fourier Terms

Key Question

Can you have too many Fourier terms? What are the symptoms?

Answer: Yes!

Symptoms of overfitting:

- In-sample fit is excellent, but out-of-sample is poor
- Seasonality looks "jagged" or unrealistic
- Forecasts oscillate wildly

Guidelines

- Maximum $K \leq m/2$ (Nyquist limit)
- Start with $K = 3$–5 for most applications
- Use cross-validation to select K
- Prophet default: $K = 10$ for yearly, $K = 3$ for weekly

TSA_ch9_fourier_approximation

Discussion: Handling Structural Breaks

Scenario

Your historical data includes COVID-19 period (2020-2021). How do you handle this when forecasting 2024?

Options

1. **Exclude COVID period**: Train only on pre-COVID and post-COVID data
2. **Use changepoints**: Let Prophet detect/specify breaks
3. **Add regressors**: Include COVID indicator variable
4. **Adjustment**: Manually adjust 2020-2021 values to "normal"

Prophet Approach

```
model = Prophet(changepoints=['2020-03-15', '2021-06-01'])
df['covid'] = (df['ds'] >= '2020-03-15') & (df['ds'] < '2021-06-01')
model.add_regressor('covid')
```

TSA_ch9_changepoint_detection

AI Exercise: Critical Thinking

Prompt to test in ChatGPT / Claude / Copilot

"Use Prophet to forecast daily Wikipedia page views for the 'Bitcoin' article. Include holiday effects, identify changepoints, and compare with a TBATS model. Evaluate using time series cross-validation."

Exercise:

1. Did the AI correctly specify the seasonality periods (weekly, yearly)?
2. Are the changepoints reasonable? Do they correspond to real events?
3. Is the cross-validation properly implemented (expanding window)?
4. How does Prophet handle the extreme spikes in Bitcoin-related traffic?
5. Does the TBATS comparison use the same evaluation metrics?

Warning: AI-generated code may run without errors and look professional. *That does not mean it is correct.*

TSA_ch9_prophet_vs_tbats

Take-Home Exercises

1. **Theoretical**: Prove that $K = m/2$ Fourier terms can represent any periodic function with period m (for even m).
2. **Computation**: For the seasonal pattern below (daily data, weekly cycle), determine the minimum number of Fourier harmonics needed:

   Mon: 100, Tue: 110, Wed: 115, Thu: 110, Fri: 120, Sat: 80, Sun: 65
3. **Applied**: Download hourly electricity demand data from a public source:
   - Fit both TBATS (in R) and Prophet (in Python)
   - Compare forecast accuracy using RMSE and MAPE
   - Visualize the component decompositions
4. **Critical Thinking**: Why might Prophet perform poorly on high-frequency financial data (e.g., minute-by-minute stock prices)?

TSA_ch9_tbats_decomposition

Exercise Solutions Hints

Hints

1. By Fourier theorem, any periodic function can be represented as sum of sines and cosines. With period m, frequencies are k/m for $k = 1, \ldots, m/2$.
2. The pattern has:
   - One peak (Friday) and one trough (Sunday)
   - Fairly smooth transitions
   - $K = 2$ or $K = 3$ likely sufficient (try and compare)
3. For electricity data:
   - Include daily (24h) and weekly (168h) patterns
   - Add holidays for your region in Prophet
   - Expect MAPE around 3-5% for hourly forecasts
4. Financial data issues:
   - No clear seasonality (market efficiency)
   - High noise-to-signal ratio
   - Prophet designed for "business" data with trends and seasons

TSA_ch9_fourier_approximation

Summary: Chapter 9

Key Concepts

1. **TBATS**: Automatic, Fourier-based, handles any number of seasonal periods
2. **Prophet**: Analyst-friendly, explicit holiday/event handling, interpretable
3. **Seasonality mode**: Additive (constant amplitude) vs Multiplicative (proportional)
4. **Fourier terms**: More = flexible but risk overfitting; use CV to select
5. **Changepoints**: Allow trend to adapt to structural breaks

Questions?

Bibliography I

Time Series Fundamentals

- Hyndman, R.J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice*, 3rd ed., OTexts.
- Shumway, R.H., & Stoffer, D.S. (2017). *Time Series Analysis and Its Applications*, 4th ed., Springer.
- Brockwell, P.J., & Davis, R.A. (2016). *Introduction to Time Series and Forecasting*, 3rd ed., Springer.

Financial Time Series

- Tsay, R.S. (2010). *Analysis of Financial Time Series*, 3rd ed., Wiley.
- Franke, J., Härdle, W.K., & Hafner, C.M. (2019). *Statistics of Financial Markets*, 4th ed., Springer.

Bibliography II

Modern Approaches and Machine Learning

- Nielsen, A. (2019). *Practical Time Series Analysis*, O'Reilly Media.
- Petropoulos, F., et al. (2022). *Forecasting: Theory and Practice*, International Journal of Forecasting.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). The M4 Competition, International Journal of Forecasting.

Online Resources and Code

- **Quantlet**: `https://quantlet.com` — Code repository for statistics
- **Quantinar**: `https://quantinar.com` — Quantitative methods learning platform
- **GitHub TSA**: `https://github.com/QuantLet/TSA/tree/main/TSA_ch9` — Python code for this seminar

Thank You!

Questions?

Seminar materials are available at: https://danpele.github.io/Time-Series-Analysis/

Quantlet Quantinar