Time Series Analysis and Forecasting

Chapter 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

Outline

The Problem: Complex Seasonal Patterns

Real-World Examples

- **Hourly electricity demand**: Daily + Weekly + Annual patterns
- **Website traffic**: Daily + Weekly + Holiday effects
- **Retail sales**: Weekly + Monthly + Annual + Holiday effects
- **Call center volume**:
  - Hourly + Daily + Weekly patterns

SARIMA Limitation

Standard SARIMA$(p, d, q)(P, D, Q)_s$ handles only **one** seasonal period s.

For hourly data with daily AND weekly patterns, we need $s_1 = 24$ and $s_2 = 168$.

Solutions for Multiple Seasonalities

Traditional Approaches

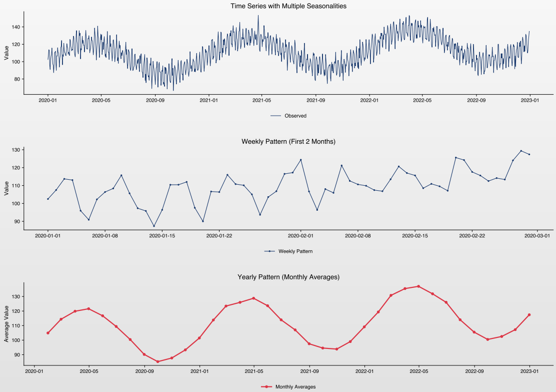
- **Fourier terms**: Add sin/cos regressors
- **Dummy variables**: Many parameters
- **Nested models**: Complex specification

Modern Approaches

- **TBATS**: Automatic, handles many periods
- **Prophet**: Flexible, interpretable
- **Neural methods**:
  - Deep learning

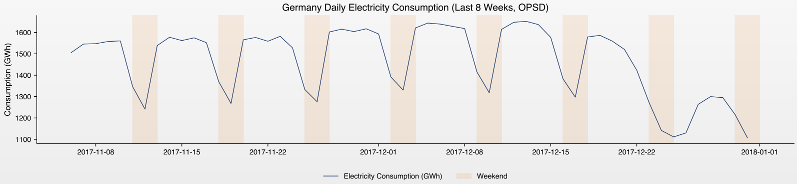
Method	Max Seasonalities	Interpretable
SARIMA	1	Yes
Fourier + ARIMA	Multiple	Moderate
TBATS	Multiple	Moderate
Prophet	Multiple	Yes

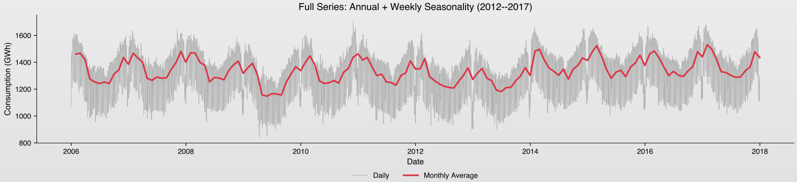
Example: Hourly Data with Multiple Seasonalities



TSA_ch9_multiple_seasonality

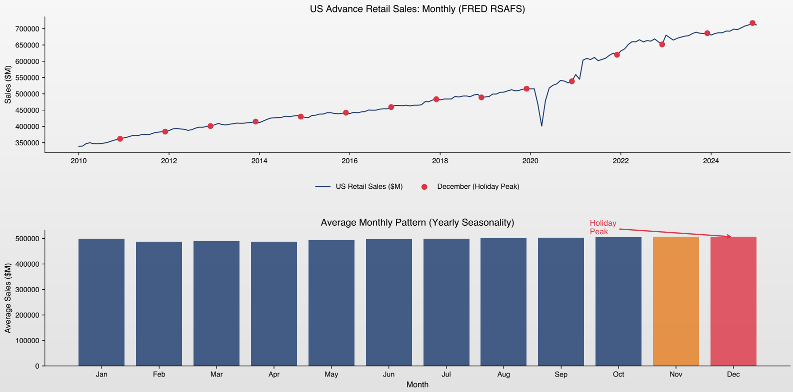
Real Example: Electricity Demand





TSA_ch9_electricity_demand

Real Example: Retail Sales with Holidays



TSA_ch9_retail_sales

TBATS: What Does It Stand For?

TBATS Components

- T **Trigonometric** seasonality using Fourier terms
- B **Box-Cox** transformation for variance stabilization
- A **ARMA** errors for remaining autocorrelation
- T **Trend** component (possibly damped)
- S **Seasonal** components (multiple allowed)

Key Innovation: Trigonometric Seasonality

$$s_t^{(i)} = \sum_{j=1}^{k_i} \left[s_j^{(i)} \cos\left(\frac{2\pi jt}{m_i}\right) + s_j^{*(i)} \sin\left(\frac{2\pi jt}{m_i}\right) \right]$$

m_i = seasonal period, k_i = number of harmonics

TBATS Model Structure

State Space Representation

$$y_t^{(\omega)} = \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^{T} s_{t-m_i}^{(i)} + d_t \tag{1}$$

$$\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha d_t, \quad b_t = \phi b_{t-1} + \beta d_t \tag{2}$$

- $y_t^{(\omega)}$: Box-Cox transformed
- ℓ_t: local level
- b_t: trend with damping ϕ
- $s_t^{(i)}$: seasonal components
- d_t: ARMA(p, q) errors
- Multiple periods: $m_1, \ldots, m_T$

TBATS: Trigonometric Seasonality

Why Fourier/Trigonometric Terms?

1. **Parsimonious**: Fewer parameters than dummy variables
2. **Smooth**: Captures smooth seasonal patterns naturally
3. **Flexible**: Number of harmonics k controls complexity
4. **Non-integer periods**: Can handle $s = 365.25$ for daily data

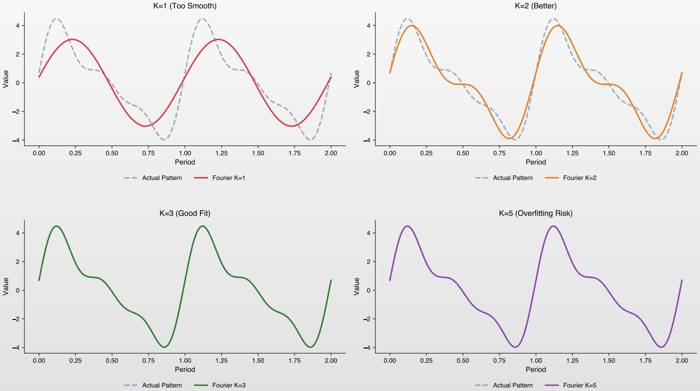
Low k (few harmonics)

- Smooth pattern
- Fewer parameters
- May miss sharp peaks

High k (many harmonics)

- Can capture any pattern
- More parameters
- Risk of overfitting

Fourier Approximation of Seasonality



TSA_ch9_fourier_approximation

TBATS: Key Features

Automatic Model Selection

TBATS automatically determines:

- Box-Cox parameter ω for variance stabilization
- Number of harmonics k_i for each seasonal period
- ARMA orders (p, q) for residual autocorrelation
- Damped vs non-damped trend specification

BATS vs TBATS

- **BATS**: Traditional seasonal states (dummy variables)
- **TBATS**: Trigonometric (Fourier) seasonal representation
- TBATS more parsimonious for long seasonal periods

TBATS: Advantages and Limitations

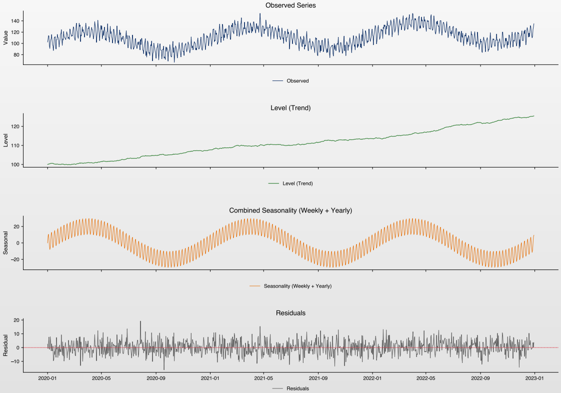
Advantages

- Handles **multiple** seasonal periods
- **Automatic** model selection
- Handles **non-integer** periods (365.25)
- **Box-Cox** for heteroskedasticity
- Good for **high-frequency** data

Limitations

- **Computationally intensive**
- No **external regressors**
- Less **interpretable** than Prophet
- Can be **slow** for very long series
- Requires **sufficient data** per season

TBATS Decomposition Example



TSA_ch9_tbats_decomposition

Prophet: Overview

What is Prophet?

Forecasting procedure developed by Facebook (Meta) in 2017 for **business time series**:

- Strong seasonal effects (daily, weekly, yearly)
- Holiday effects and trend changes (changepoints)
- Handles missing data and outliers

Key Philosophy: "Analyst-in-the-loop"

Designed for analysts with domain knowledge but without time series expertise.

Prophet Model Structure

Decomposition Approach

Prophet uses an **additive decomposition**:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t$$

$g(t)$: Trend

- Linear or logistic
- Automatic changepoints
- Growth saturation

$s(t)$: Seasonality

- Fourier series
- Multiple periods
- Custom seasonality

$h(t)$: Holidays

- Country holidays
- Custom events
- Window effects

Prophet: Trend Component

Linear Trend with Changepoints

$$g(t) = (k + \mathbf{a}(t)^T\delta) \cdot t + (m + \mathbf{a}(t)^T\gamma)$$

- k: base growth rate
- δ: rate adjustments
- $\mathbf{a}(t)$: active changepoints

Logistic Growth

For saturating trends:

$$g(t) = \frac{C(t)}{1 + e^{-k(t-m)}}$$

$C(t)$ = carrying capacity

Prophet: Seasonality Component

Fourier Series Representation

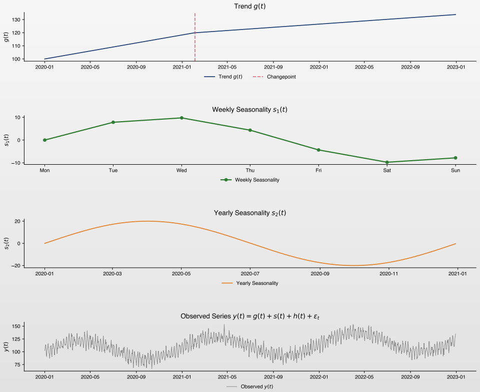
$$s(t) = \sum_{n=1}^{N} \left[a_n \cos\left(\frac{2\pi nt}{P}\right) + b_n \sin\left(\frac{2\pi nt}{P}\right) \right]$$

Default Settings

Seasonality	Period	Fourier Order
Yearly	365.25 days	10
Weekly	7 days	3
Daily	1 day	4

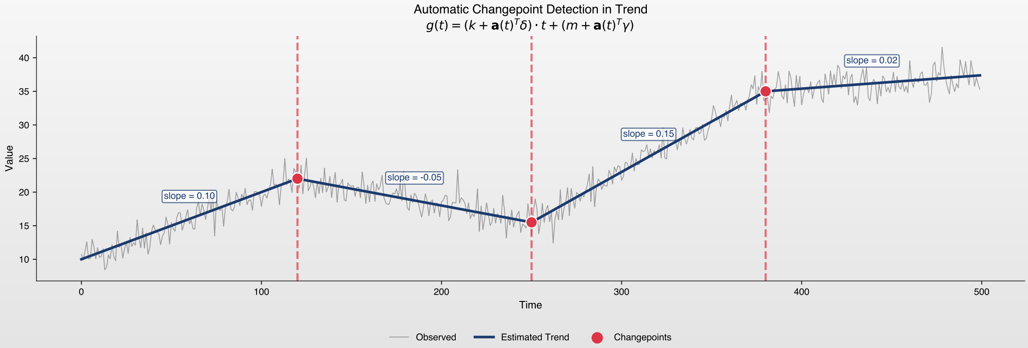
Higher N = more flexibility but risk of overfitting

Prophet Component Decomposition



TSA_ch9_prophet_components

Trend Changepoint Detection



TSA_ch9_changepoint_detection

Prophet: Holiday Effects

Holiday Model

$$h(t) = Z(t) \cdot \boldsymbol{\kappa}$$

where $Z(t)$ is an indicator matrix for holidays and $\boldsymbol{\kappa}$ are holiday effects.

Built-in Features

- 60+ countries supported
- Custom holiday definitions
- Window effects (before/after)

Holiday Types

- National holidays
- Religious observances
- Business events

Prophet: Customization Options

Seasonality Customization

- Add custom seasonal periods (monthly, quarterly)
- Control Fourier order for each seasonality
- Enable/disable default seasonalities

External Regressors

Prophet supports adding external variables:

- Weather data, promotions, special events
- Binary or continuous regressors
- Automatic regularization

Prophet: Uncertainty Quantification

Sources of Uncertainty

1. Trend uncertainty
2. Seasonality uncertainty
3. Observation noise

Prediction Intervals

- Default: 80% interval
- Configurable width
- Grows with horizon

Prophet: Tuning Parameters

Key Parameters

Parameter	Effect
`changepoint_prior_scale`	Trend flexibility (default: 0.05)
`seasonality_prior_scale`	Seasonality flexibility (default: 10)
`holidays_prior_scale`	Holiday effect size (default: 10)
`seasonality_mode`	'additive' or 'multiplicative'
`changepoint_range`	Portion of history for changepoints

Practical Tips

- **Overfitting trend?** Decrease `changepoint_prior_scale`
- **Underfitting seasonality?** Increase `seasonality_prior_scale`
- **Seasonal amplitude varies?** Use `seasonality_mode='multiplicative'`

Prophet: Advantages and Limitations

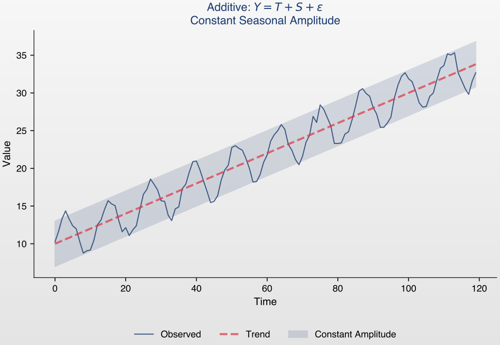
Advantages

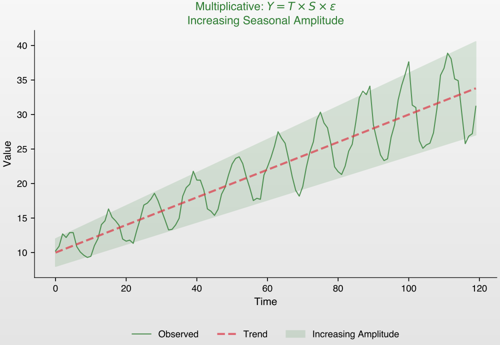
- **Easy to use**: Minimal tuning needed
- **Interpretable**: Clear decomposition
- **Handles missing data** well
- **Holiday effects** built-in
- **Multiple seasonalities**
- **External regressors** supported
- **Fast** fitting

Limitations

- **Not ARIMA-based**: No autocorrelation modeling
- **Daily data focus**: Less suited for very high frequency
- **Trend assumptions**: Linear/logistic may not fit
- **No built-in CV**: Must implement manually
- **Overfitting risk** with many seasonalities

Additive vs Multiplicative Seasonality





TSA_ch9_additive_vs_multiplicative

TBATS vs Prophet: Head-to-Head

Feature	**TBATS**	**Prophet**
Multiple seasonalities	Yes (automatic)	Yes (manual or auto)
Holiday effects	No	Yes (built-in)
External regressors	No	Yes
Trend changepoints	No (smooth)	Yes (automatic)
Missing data	Interpolation needed	Handles natively
Interpretability	Moderate	High
Computation speed	Slow	Fast
High-frequency data	Good	Moderate
Non-integer periods	Yes (e.g., 365.25)	Yes
Uncertainty intervals	Yes	Yes

When to Use Each Model

Use TBATS when:

- High-frequency data
- Multiple seasonal periods
- No external regressors
- Automatic selection preferred

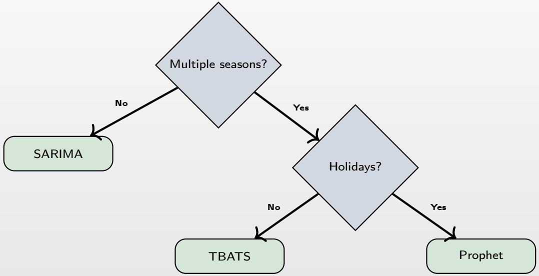
Use Prophet when:

- Business forecasting
- Holiday effects important
- Trend has changepoints
- External regressors available

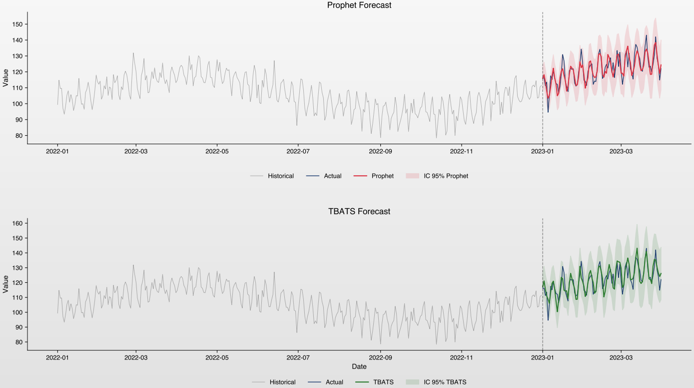
General Guideline

Prophet: business applications with daily data
TBATS: technical applications with high-frequency data

Decision Flowchart

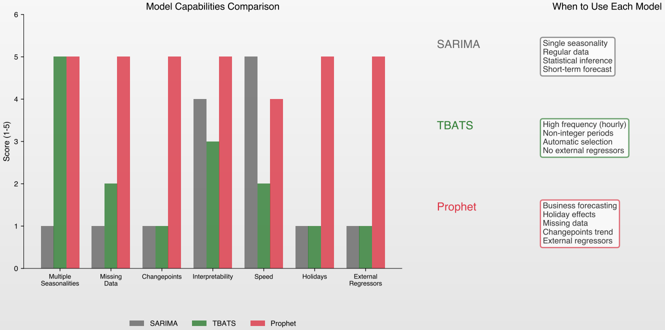


Prophet vs TBATS: Forecast Comparison



TSA_ch9_prophet_vs_tbats

Model Selection Guide



TSA_ch9_model_selection_guide

Case Study: Energy Demand Forecasting

Problem

Forecast hourly electricity demand with:

- **Daily pattern**: Peak at noon and evening
- **Weekly pattern**: Lower on weekends
- **Annual pattern**: Higher in summer (AC) and winter (heating)
- **Holiday effects**: Lower demand on holidays

Approach

1. Try TBATS with periods [24, 168, 8766]
2. Try Prophet with daily, weekly, yearly seasonality + holidays
3. Compare using cross-validation

Case Study: Results

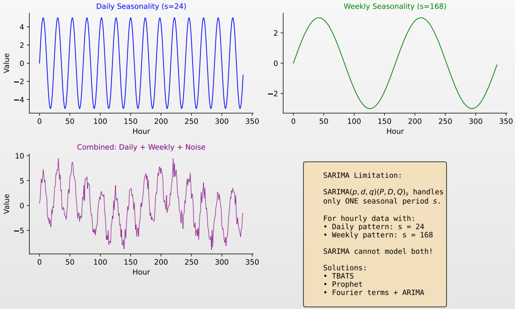
Model Comparison

Model	MAPE	RMSE	Coverage
SARIMA (daily only)	8.5%	450 MW	75%
TBATS	4.2%	220 MW	82%
Prophet	4.8%	250 MW	85%
Prophet + holidays	3.9%	200 MW	88%

Key Finding

Multiple seasonality models significantly outperform single-seasonality SARIMA.

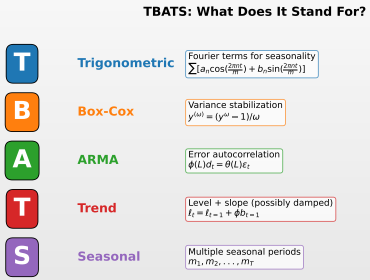
Quiz 1: Multiple Seasonality



Question: Why can't standard SARIMA$(p, d, q)(P, D, Q)_s$ model hourly electricity data with both daily and weekly patterns?

Answer: SARIMA handles only **one** seasonal period s. Cannot set $s = 24$ and $s = 168$ simultaneously.

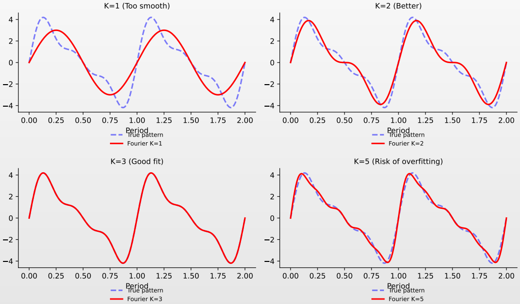
Quiz 2: TBATS Components



Question: What does each letter in TBATS represent?

Answer: **T**rigonometric seasonality, **B**ox-Cox transformation, **A**RMA errors, **T**rend, **S**easonal components.

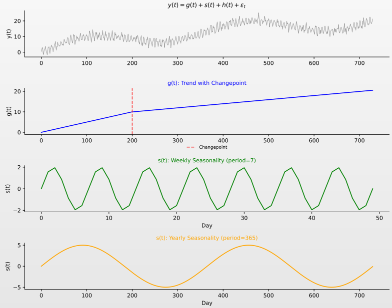
Quiz 3: Fourier Harmonics



Question: What happens when we increase the number of Fourier harmonics K?

Answer: Higher K captures more complex patterns but increases overfitting risk.

Quiz 4: Prophet Decomposition



Question: What are the main components in Prophet's model $y(t) = g(t) + s(t) + h(t) + \varepsilon_t$?

Answer: $g(t)$ = trend with changepoints, $s(t)$ = seasonality, $h(t)$ = holiday effects.

Quiz 5: Model Comparison

TBATS vs Prophet: Head-to-Head Comparison

Feature	TBATS	Prophet
Multiple seasonalities	Yes (automatic)	Yes (manual/auto)
Holiday effects	No	Yes (built-in)
External regressors	No	Yes
Trend changepoints	No (smooth)	Yes (automatic)
Missing data	Needs interpolation	Handles natively
Interpretability	Moderate	High
Computation speed	Slow	Fast
High-frequency data	Good	Moderate
Non-integer periods	Yes (e.g., 365.25)	Yes
Best for	Technical/high-freq	Business/daily

Question: What key features does Prophet have that TBATS lacks?

Answer: Holiday effects, external regressors, trend changepoints, native missing data handling.

Key Takeaways

Multiple Seasonalities

- Real-world data has multiple patterns
- SARIMA: one seasonal period
- TBATS/Prophet: multiple periods

Model Selection

- TBATS: high-frequency, automatic
- Prophet: interpretable, holidays
- Both use Fourier terms

Remember

Always validate with proper time series cross-validation!

Questions?

Questions?

Next Steps:

- Practice with the Jupyter notebook
- Try Prophet on your own data
- Explore NeuralProphet for deep learning extension