Seasonality Detection Methods: A Comparative Study

Comprehensive Benchmark for the anofox-forecast DuckDB Extension

anofox-forecast benchmark suite

2026-01-07

Table of contents

Executive Summary . . . 2
Introduction . . . 2
- Detection Methods Available . . . 2
- Simulation Scenarios . . . 3

Setup . . . 4
- Connect to DuckDB and Load Extension . . . 4

Data Simulation . . . 4
- Simulation Parameters . . . 4
- Simulation Functions . . . 4
- Generate All Scenarios . . . 4
- Example Series Visualization . . . 4
- Load Data into DuckDB . . . 4

Detection Methods Tutorial . . . 6
- Method 1: Basic Seasonality Detection (`ts_detect_seasonality`) . . . 6
- Method 2: Full Seasonality Analysis (`ts_analyze_seasonality`) . . . 6
- Method 3: FFT Period Estimation (`ts_estimate_period_fft`) . . . 7
- Method 4: ACF Period Estimation (`ts_estimate_period_acf`) . . . 7
- Method 5: Autoperiod (`ts_autoperiod`) . . . 7
- Method 6: CFD-Autoperiod (`ts_cfd_autoperiod`) . . . 8
- Method 7: Lomb-Scargle Periodogram (`ts_lomb_scargle`) . . . 8
- Method 8: AIC-Based Model Comparison (`ts_aic_period`) . . . 9
- Method 9: Singular Spectrum Analysis (`ts_ssa_period`) . . . 9
- Method 10: STL Decomposition (`ts_stl_period`) . . . 9
- Method 11: Matrix Profile (`ts_matrix_profile_period`) . . . 10
- Method 12: SAZED (`ts_sazed_period`) . . . 10

Running Full Detection with All Methods . . . 11
Evaluation Results . . . 11
Detection Rates by Scenario (All Methods) . . . 11
Detection Rates Table . . . 11
Period Estimation Accuracy (All Methods) . . . 11
Confidence Distribution by Ground Truth (All Methods) . . . 11
Performance Metrics (All Methods) . . . 11
Scenario-Specific Performance (All Methods) . . . 12
Statistical Comparison: McNemar Tests . . . 12
Overall McNemar Comparison . . . 12
McNemar Tests by Scenario . . . 14
McNemar Summary Heatmap . . . 20
Summary Statistics . . . 20
Recommendations . . . 21
Method Selection Guide . . . 21
Best Practices . . . 27
Conclusion . . . 27
Cleanup . . . 27
Appendix: SQL Function Reference . . . 27
Session Info . . . 29

Executive Summary

This benchmark evaluates all seasonality detection methods available in the `anofox-forecast` DuckDB extension. Using simulated time series with known characteristics, we compare detection accuracy across seven scenarios using twelve detection methods spanning spectral analysis, autocorrelation, decomposition, and model-based approaches.

Introduction

The `anofox-forecast` DuckDB extension provides SQL functions for detecting and analyzing seasonality in time series data. This report serves as both a benchmark and a tutorial, demonstrating how to use these functions effectively.

Detection Methods Available

Method	SQL Function	Description
Basic Detection	`ts_detect_seasonality(values)`	Returns array of detected periods
Full Analysis	`ts_analyze_seasonality(values)`	Comprehensive struct with periods and strength metrics

Method	SQL Function	Description
FFT Period	`ts_estimate_period_fft(values)`	Period estimation via Fast Fourier Transform
ACF Period	`ts_estimate_period_acf(values)`	Period estimation via Autocorrelation
Autoperiod	`ts_autoperiod(values)`	Hybrid FFT+ACF validation method
CFD-Autoperiod	`ts_cfd_autoperiod(values)`	Circular Frequency Domain variant
Lomb-Scargle	`ts_lomb_scargle(values)`	Spectral analysis for irregular sampling
AIC Comparison	`ts_aic_period(values)`	Model selection via AIC/BIC criteria
SSA Period	`ts_ssa_period(values)`	Singular Spectrum Analysis decomposition
STL Period	`ts_stl_period(values)`	STL decomposition with seasonal strength optimization
Matrix Profile	`ts_matrix_profile_period(values)`	Motif-based period detection via matrix profile
SAZED	`ts_sazed_period(values)`	Spectral Analysis with Zero-padding for Enhanced Detection

Simulation Scenarios

We test across seven scenarios representing common real-world patterns:

1. **Strong Seasonal** - Clear sinusoidal pattern (high SNR)
2. **Weak Seasonal** - Low amplitude with high noise
3. **No Seasonal** - Pure noise (null case)
4. **Trending Seasonal** - Seasonality with strong linear trend
5. **Variable Amplitude** - Time-varying amplitude modulation
6. **Emerging Seasonal** - Seasonality appears mid-series
7. **Fading Seasonal** - Seasonality disappears mid-series

Setup

Connect to DuckDB and Load Extension

Data Simulation

We generate synthetic time series in R with known seasonality characteristics, then load them into DuckDB for analysis.

Simulation Parameters

Simulation Functions

Generate All Scenarios

```
Total series generated: 700


Series per scenario:


   strong_seasonal      weak_seasonal        no_seasonal  trending_seasonal
               100                100                100                100
variable_amplitude  emerging_seasonal    fading_seasonal
               100                100                100
```

Example Series Visualization

Load Data into DuckDB

```
[1] 0


[1] 0


Data loaded into DuckDB table 'simulated_series'


Row count: 700
```

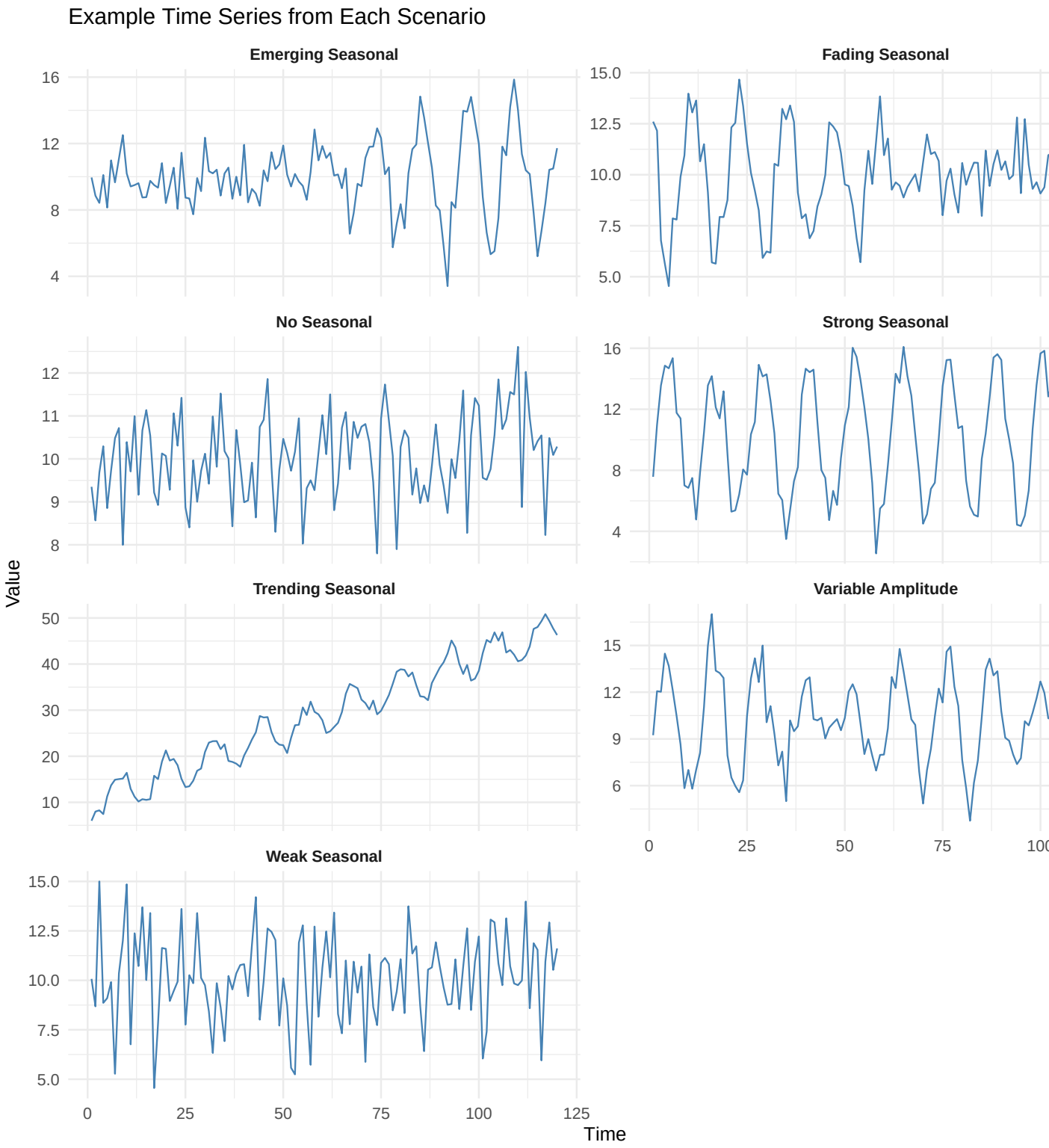

Figure 1: Example time series from each scenario

Detection Methods Tutorial

This section demonstrates each detection method with explicit SQL queries.

Method 1: Basic Seasonality Detection (`ts_detect_seasonality`)

Returns an array of detected seasonal periods.

```
Connection class: duckdb_connection
```

```
Tables in database: simulated_series
```

Table 2: Basic Seasonality Detection Results

global_id	scenario	true_period	detected_periods
501	emerging_seasonal	12	12, 24, 36, 48
601	fading_seasonal	12	12, 24, 36, 48
201	no_seasonal	0	12, 24, 36, 48
1	strong_seasonal	12	12, 24, 36, 48
301	trending_seasonal	12	12, 24, 36, 48
401	variable_amplitude	12	12, 24, 36, 48
101	weak_seasonal	12	12, 24, 36, 48

Method 2: Full Seasonality Analysis (`ts_analyze_seasonality`)

Comprehensive analysis including strength metrics.

Table 3: Full Seasonality Analysis Results

global_id	scenario	true_period	detected_period	seasonal_strength	trend_strength
501	emerging_seasonal	12	12	0.836	0.090
601	fading_seasonal	12	12	0.829	0.046
201	no_seasonal	0	12	0.838	0.042
1	strong_seasonal	12	12	0.838	0.066
301	trending_seasonal	12	12	0.793	0.002
401	variable_amplitude	12	12	0.804	0.053
101	weak_seasonal	12	12	0.775	0.013

Method 3: FFT Period Estimation (`ts_estimate_period_fft`)

Uses Fast Fourier Transform for precise frequency detection.

Table 4: FFT Period Estimation Results

global_id	scenario	true_period	fft_period	fft_confidence	fft_power
501	emerging_seasonal	12	NaN	0	0
601	fading_seasonal	12	NaN	0	0
201	no_seasonal	0	NaN	0	0
1	strong_seasonal	12	NaN	0	0
301	trending_seasonal	12	NaN	0	0
401	variable_amplitude	12	NaN	0	0
101	weak_seasonal	12	NaN	0	0

Method 4: ACF Period Estimation (`ts_estimate_period_acf`)

Detects periods using autocorrelation peaks.

Table 5: ACF Period Estimation Results

global_id	scenario	true_period	acf_period	acf_confidence
501	emerging_seasonal	12	NaN	0
601	fading_seasonal	12	NaN	0
201	no_seasonal	0	NaN	0
1	strong_seasonal	12	NaN	0
301	trending_seasonal	12	NaN	0
401	variable_amplitude	12	NaN	0
101	weak_seasonal	12	NaN	0

Method 5: Autoperiod (`ts_autoperiod`)

Hybrid approach combining FFT with ACF validation for robust period detection.

Table 6: Autoperiod Detection Results

global_id	scenario	true_period	autoperiod_period	autoperiod_confidence	autoperiod_acf_val
501	emerging_seasonal	12	NaN	0	0
601	fading_seasonal	12	NaN	0	0
201	no_seasonal	0	NaN	0	0

global_id	scenario	true_period	autoperiod_period	autoperiod_confidence	autoperiod_acf_val
1	strong_seasonal	12	NaN	0	0
301	trending_seasonal	12	NaN	0	0
401	variable_amplitude	12	NaN	0	0
101	weak_seasonal	12	NaN	0	0

Method 6: CFD-Autoperiod (`ts_cfd_autoperiod`)

Circular Frequency Domain variant of Autoperiod for improved harmonic handling.

Table 7: CFD-Autoperiod Detection Results

global_id	scenario	true_period	cfd_period	cfd_confidence	cfd_acf_val
501	emerging_seasonal	12	NaN	0	0
601	fading_seasonal	12	NaN	0	0
201	no_seasonal	0	NaN	0	0
1	strong_seasonal	12	NaN	0	0
301	trending_seasonal	12	NaN	0	0
401	variable_amplitude	12	NaN	0	0
101	weak_seasonal	12	NaN	0	0

Method 7: Lomb-Scargle Periodogram (`ts_lomb_scargle`)

Spectral analysis optimized for irregularly sampled data with false alarm probability.

Table 8: Lomb-Scargle Periodogram Results

global_id	scenario	true_period	lomb_period	lomb_power	lomb_fap
501	emerging_seasonal	12	12.474	24.924	0.000
601	fading_seasonal	12	12.038	30.458	0.000
201	no_seasonal	0	3.852	5.623	0.973
1	strong_seasonal	12	11.968	55.236	0.000
301	trending_seasonal	12	59.500	8.188	0.243
401	variable_amplitude	12	11.968	46.474	0.000
101	weak_seasonal	12	3.015	6.291	0.843

Method 8: AIC-Based Model Comparison (`ts_aic_period`)

Fits sinusoidal models at candidate periods and selects best via AIC/BIC criteria.

Table 9: AIC Period Selection Results

global_id	scenario	true_period	aic_period	aic_r_squared	aic_value
501	emerging_seasonal	12	12.653	0.401	124.201
601	fading_seasonal	12	11.469	0.303	128.926
201	no_seasonal	0	2.000	0.066	-7.741
1	strong_seasonal	12	11.469	0.444	247.499
301	trending_seasonal	12	60.000	0.141	575.127
401	variable_amplitude	12	11.469	0.379	180.157
101	weak_seasonal	12	6.735	0.035	192.319

Method 9: Singular Spectrum Analysis (`ts_ssa_period`)

Decomposes the time series via SSA to extract periodic components.

Table 10: SSA Period Detection Results

global_id	scenario	true_period	ssa_period	ssa_variance
501	emerging_seasonal	12	13.333	0.981
601	fading_seasonal	12	13.333	0.985
201	no_seasonal	0	2.051	0.996
1	strong_seasonal	12	13.333	0.943
301	trending_seasonal	12	13.333	0.995
401	variable_amplitude	12	13.333	0.970
101	weak_seasonal	12	7.273	0.982

Method 10: STL Decomposition (`ts_stl_period`)

Uses STL (Seasonal and Trend decomposition using LOESS) to find the period that maximizes seasonal strength.

Table 11: STL Period Detection Results

global_id	scenario	true_period	stl_period	stl_strength
501	emerging_seasonal	12	38	0.493
601	fading_seasonal	12	36	0.588

global_id	scenario	true_period	stl_period	stl_strength
201	no_seasonal	0	38	0.492
1	strong_seasonal	12	36	0.944
301	trending_seasonal	12	24	0.843
401	variable_amplitude	12	36	0.799
101	weak_seasonal	12	35	0.363

Method 11: Matrix Profile (`ts_matrix_profile_period`)

Uses Matrix Profile algorithm to find motifs (recurring patterns) and estimate periodicity from the distribution of motif distances.

Table 12: Matrix Profile Period Detection Results

global_id	scenario	true_period	mp_period	mp_confidence	mp_n_motifs
501	emerging_seasonal	12	24	0.360	86
601	fading_seasonal	12	12	0.373	83
201	no_seasonal	0	22	0.125	80
1	strong_seasonal	12	36	0.326	86
301	trending_seasonal	12	48	0.416	89
401	variable_amplitude	12	12	0.474	57
101	weak_seasonal	12	14	0.281	89

Method 12: SAZED (`ts_sazed_period`)

Spectral Analysis with Zero-padding for Enhanced Detection - uses zero-padding and Hann windowing to improve frequency resolution in spectral analysis.

Table 13: SAZED Period Detection Results

global_id	scenario	true_period	sazed_period	sazed_snr
501	emerging_seasonal	12	12.488	62.832
601	fading_seasonal	12	11.907	83.433
201	no_seasonal	0	3.821	6.199
1	strong_seasonal	12	11.907	637.221
301	trending_seasonal	12	11.907	311.489
401	variable_amplitude	12	11.907	227.312
101	weak_seasonal	12	6.919	8.045

Running Full Detection with All Methods

Now we run all twelve detection methods on all simulated series.

```
Processed 700 detection results
```

Evaluation Results

Detection Rates by Scenario (All Methods)

Detection Rates Table

Table 14: Detection Rates by Scenario and Method

scenario	Basic	Analysis	FFT	ACF	Autoperiod	CFD	Lomb	AIC	SSA	STL	Matrix	Prony	SAZED
emerging_seasonal	100.0%	100.0%	0.0%	0.0%	0.0%	0.0%	100.0%	97.0%	100.0%	100.0%	100.0%		100.0%
fading_seasonal	100.0%	100.0%	0.0%	0.0%	0.0%	0.0%	100.0%	99.0%	100.0%	100.0%	100.0%		100.0%
no_seasonal	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	100.0%	99.0%	97.0%		100.0%
strong_seasonal	100.0%	100.0%	0.0%	0.0%	0.0%	0.0%	100.0%	100.0%	100.0%	100.0%	100.0%		100.0%
trending_seasonal	14.0%	14.0%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	100.0%	100.0%	100.0%		100.0%
variable_amplitude	100.0%	100.0%	0.0%	0.0%	0.0%	0.0%	100.0%	96.0%	100.0%	100.0%	100.0%		100.0%
weak_seasonal	99.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	100.0%	98.0%	100.0%		100.0%

Period Estimation Accuracy (All Methods)

Confidence Distribution by Ground Truth (All Methods)

Performance Metrics (All Methods)

Table 15: Detection Performance Metrics (All Methods)

Method	Precision	Recall	Specificity	F1	Accuracy
Basic	0.837	0.855	0.00	0.846	0.733
Analysis	1.000	0.690	1.00	0.817	0.734
FFT	0.000	0.000	1.00	0.000	0.143
ACF	0.000	0.000	1.00	0.000	0.143
Autoperiod	0.000	0.000	1.00	0.000	0.143
CFD	0.000	0.000	1.00	0.000	0.143
Lomb	0.857	1.000	0.00	0.923	0.857

Method	Precision	Recall	Specificity	F1	Accuracy
AIC	1.000	0.653	1.00	0.790	0.703
SSA	0.857	1.000	0.00	0.923	0.857
STL	0.858	0.997	0.01	0.922	0.856
MatrixProfile	0.861	1.000	0.03	0.925	0.861
SAZED	0.857	1.000	0.00	0.923	0.857

Scenario-Specific Performance (All Methods)

Statistical Comparison: McNemar Tests

McNemar's test compares paired binary outcomes between methods. A significant p-value indicates that two methods differ in their detection performance.

Overall McNemar Comparison

Table 16: McNemar Test: Pairwise Method Comparisons (* = p < 0.05)

Method1	Method2	Only M1	Only M2	Chi-sq	p-value	Significant
Basic	Analysis	199	0	197.0050	0.0000	*
Basic	Fft	613	0	611.0016	0.0000	*
Basic	Acf	613	0	611.0016	0.0000	*
Basic	Autoperiod	613	0	611.0016	0.0000	*
Basic	Cfd	613	0	611.0016	0.0000	*
Basic	Lomb	0	87	85.0115	0.0000	*
Basic	Aic	221	0	219.0045	0.0000	*
Basic	Ssa	0	87	85.0115	0.0000	*
Basic	STL	3	87	76.5444	0.0000	*
Basic	MatrixProfile	3	87	76.5444	0.0000	*
Basic	SAZED	0	87	85.0115	0.0000	*
Analysis	Fft	414	0	412.0024	0.0000	*
Analysis	Acf	414	0	412.0024	0.0000	*
Analysis	Autoperiod	414	0	412.0024	0.0000	*
Analysis	Cfd	414	0	412.0024	0.0000	*
Analysis	Lomb	0	286	284.0035	0.0000	*
Analysis	Aic	22	0	20.0455	0.0000	*
Analysis	Ssa	0	286	284.0035	0.0000	*
Analysis	STL	0	283	281.0035	0.0000	*
Analysis	MatrixProfile	0	283	281.0035	0.0000	*

Method1	Method2	Only M1	Only M2	Chi-sq	p-value	Significant
Analysis	SAZED	0	286	284.0035	0.0000	*
Fft	Acf	0	0	0.0000	1.0000	
Fft	Autoperiod	0	0	0.0000	1.0000	
Fft	Cfd	0	0	0.0000	1.0000	
Fft	Lomb	0	700	698.0014	0.0000	*
Fft	Aic	0	392	390.0026	0.0000	*
Fft	Ssa	0	700	698.0014	0.0000	*
Fft	STL	0	697	695.0014	0.0000	*
Fft	MatrixProfile	0	697	695.0014	0.0000	*
Fft	SAZED	0	700	698.0014	0.0000	*
Acf	Autoperiod	0	0	0.0000	1.0000	
Acf	Cfd	0	0	0.0000	1.0000	
Acf	Lomb	0	700	698.0014	0.0000	*
Acf	Aic	0	392	390.0026	0.0000	*
Acf	Ssa	0	700	698.0014	0.0000	*
Acf	STL	0	697	695.0014	0.0000	*
Acf	MatrixProfile	0	697	695.0014	0.0000	*
Acf	SAZED	0	700	698.0014	0.0000	*
Autoperiod	Cfd	0	0	0.0000	1.0000	
Autoperiod	Lomb	0	700	698.0014	0.0000	*
Autoperiod	Aic	0	392	390.0026	0.0000	*
Autoperiod	Ssa	0	700	698.0014	0.0000	*
Autoperiod	STL	0	697	695.0014	0.0000	*
Autoperiod	MatrixProfile	0	697	695.0014	0.0000	*
Autoperiod	SAZED	0	700	698.0014	0.0000	*
Cfd	Lomb	0	700	698.0014	0.0000	*
Cfd	Aic	0	392	390.0026	0.0000	*
Cfd	Ssa	0	700	698.0014	0.0000	*
Cfd	STL	0	697	695.0014	0.0000	*
Cfd	MatrixProfile	0	697	695.0014	0.0000	*
Cfd	SAZED	0	700	698.0014	0.0000	*
Lomb	Aic	308	0	306.0032	0.0000	*
Lomb	Ssa	0	0	0.0000	1.0000	
Lomb	STL	3	0	1.3333	0.2482	
Lomb	MatrixProfile	3	0	1.3333	0.2482	
Lomb	SAZED	0	0	0.0000	1.0000	
Aic	Ssa	0	308	306.0032	0.0000	*
Aic	STL	0	305	303.0033	0.0000	*
Aic	MatrixProfile	0	305	303.0033	0.0000	*
Aic	SAZED	0	308	306.0032	0.0000	*

Method1	Method2	Only M1	Only M2	Chi-sq	p-value	Significant
Ssa	STL	3	0	1.3333	0.2482	
Ssa	MatrixProfile	3	0	1.3333	0.2482	
Ssa	SAZED	0	0	0.0000	1.0000	
STL	MatrixProfile	3	3	0.1667	0.6831	
STL	SAZED	0	3	1.3333	0.2482	
MatrixProfile	SAZED	0	3	1.3333	0.2482	

McNemar Tests by Scenario

Table 17: Significant McNemar Results by Scenario ($p < 0.05$)

Scenario	Method1	Method2	Only M1	Only M2	Chi-sq	p-value	Significant
emerging_seasonal	Basic	Fft	100	0	98.0100	0e+00	*
emerging_seasonal	Basic	Acf	100	0	98.0100	0e+00	*
emerging_seasonal	Basic	Autoperiod	100	0	98.0100	0e+00	*
emerging_seasonal	Basic	Cfd	100	0	98.0100	0e+00	*
emerging_seasonal	Analysis	Fft	100	0	98.0100	0e+00	*
emerging_seasonal	Analysis	Acf	100	0	98.0100	0e+00	*
emerging_seasonal	Analysis	Autoperiod	100	0	98.0100	0e+00	*
emerging_seasonal	Analysis	Cfd	100	0	98.0100	0e+00	*
emerging_seasonal	Fft	Lomb	0	100	98.0100	0e+00	*
emerging_seasonal	Fft	Aic	0	97	95.0103	0e+00	*
emerging_seasonal	Fft	Ssa	0	100	98.0100	0e+00	*
emerging_seasonal	Fft	STL	0	100	98.0100	0e+00	*
emerging_seasonal	Fft	MatrixProfile	0	100	98.0100	0e+00	*
emerging_seasonal	Fft	SAZED	0	100	98.0100	0e+00	*
emerging_seasonal	Acf	Lomb	0	100	98.0100	0e+00	*
emerging_seasonal	Acf	Aic	0	97	95.0103	0e+00	*
emerging_seasonal	Acf	Ssa	0	100	98.0100	0e+00	*
emerging_seasonal	Acf	STL	0	100	98.0100	0e+00	*
emerging_seasonal	Acf	MatrixProfile	0	100	98.0100	0e+00	*
emerging_seasonal	Acf	SAZED	0	100	98.0100	0e+00	*
emerging_seasonal	Autoperiod	Lomb	0	100	98.0100	0e+00	*
emerging_seasonal	Autoperiod	Aic	0	97	95.0103	0e+00	*
emerging_seasonal	Autoperiod	Ssa	0	100	98.0100	0e+00	*
emerging_seasonal	Autoperiod	STL	0	100	98.0100	0e+00	*
emerging_seasonal	Autoperiod	MatrixProfile	0	100	98.0100	0e+00	*

Scenario	Method1	Method2	Only M1	Only M2	Chi-sq	p-value	Significant
emerging_seasonal	Autoperiod	SAZED	0	100	98.0100	0e+00	*
emerging_seasonal	Cfd	Lomb	0	100	98.0100	0e+00	*
emerging_seasonal	Cfd	Aic	0	97	95.0103	0e+00	*
emerging_seasonal	Cfd	Ssa	0	100	98.0100	0e+00	*
emerging_seasonal	Cfd	STL	0	100	98.0100	0e+00	*
emerging_seasonal	Cfd	MatrixProfile	0	100	98.0100	0e+00	*
emerging_seasonal	Cfd	SAZED	0	100	98.0100	0e+00	*
fading_seasonal	Basic	Fft	100	0	98.0100	0e+00	*
fading_seasonal	Basic	Acf	100	0	98.0100	0e+00	*
fading_seasonal	Basic	Autoperiod	100	0	98.0100	0e+00	*
fading_seasonal	Basic	Cfd	100	0	98.0100	0e+00	*
fading_seasonal	Analysis	Fft	100	0	98.0100	0e+00	*
fading_seasonal	Analysis	Acf	100	0	98.0100	0e+00	*
fading_seasonal	Analysis	Autoperiod	100	0	98.0100	0e+00	*
fading_seasonal	Analysis	Cfd	100	0	98.0100	0e+00	*
fading_seasonal	Fft	Lomb	0	100	98.0100	0e+00	*
fading_seasonal	Fft	Aic	0	99	97.0101	0e+00	*
fading_seasonal	Fft	Ssa	0	100	98.0100	0e+00	*
fading_seasonal	Fft	STL	0	100	98.0100	0e+00	*
fading_seasonal	Fft	MatrixProfile	0	100	98.0100	0e+00	*
fading_seasonal	Fft	SAZED	0	100	98.0100	0e+00	*
fading_seasonal	Acf	Lomb	0	100	98.0100	0e+00	*
fading_seasonal	Acf	Aic	0	99	97.0101	0e+00	*
fading_seasonal	Acf	Ssa	0	100	98.0100	0e+00	*
fading_seasonal	Acf	STL	0	100	98.0100	0e+00	*
fading_seasonal	Acf	MatrixProfile	0	100	98.0100	0e+00	*
fading_seasonal	Acf	SAZED	0	100	98.0100	0e+00	*
fading_seasonal	Autoperiod	Lomb	0	100	98.0100	0e+00	*
fading_seasonal	Autoperiod	Aic	0	99	97.0101	0e+00	*
fading_seasonal	Autoperiod	Ssa	0	100	98.0100	0e+00	*
fading_seasonal	Autoperiod	STL	0	100	98.0100	0e+00	*
fading_seasonal	Autoperiod	MatrixProfile	0	100	98.0100	0e+00	*
fading_seasonal	Autoperiod	SAZED	0	100	98.0100	0e+00	*
fading_seasonal	Cfd	Lomb	0	100	98.0100	0e+00	*
fading_seasonal	Cfd	Aic	0	99	97.0101	0e+00	*
fading_seasonal	Cfd	Ssa	0	100	98.0100	0e+00	*
fading_seasonal	Cfd	STL	0	100	98.0100	0e+00	*
fading_seasonal	Cfd	MatrixProfile	0	100	98.0100	0e+00	*
fading_seasonal	Cfd	SAZED	0	100	98.0100	0e+00	*
no_seasonal	Basic	Analysis	100	0	98.0100	0e+00	*

Scenario	Method1	Method2	Only M1	Only M2	Chi-sq	p-value	Significant
no_seasonal	Basic	Fft	100	0	98.0100	0e+00	*
no_seasonal	Basic	Acf	100	0	98.0100	0e+00	*
no_seasonal	Basic	Autoperiod	100	0	98.0100	0e+00	*
no_seasonal	Basic	Cfd	100	0	98.0100	0e+00	*
no_seasonal	Basic	Aic	100	0	98.0100	0e+00	*
no_seasonal	Analysis	Lomb	0	100	98.0100	0e+00	*
no_seasonal	Analysis	Ssa	0	100	98.0100	0e+00	*
no_seasonal	Analysis	STL	0	99	97.0101	0e+00	*
no_seasonal	Analysis	MatrixProfile	0	97	95.0103	0e+00	*
no_seasonal	Analysis	SAZED	0	100	98.0100	0e+00	*
no_seasonal	Fft	Lomb	0	100	98.0100	0e+00	*
no_seasonal	Fft	Ssa	0	100	98.0100	0e+00	*
no_seasonal	Fft	STL	0	99	97.0101	0e+00	*
no_seasonal	Fft	MatrixProfile	0	97	95.0103	0e+00	*
no_seasonal	Fft	SAZED	0	100	98.0100	0e+00	*
no_seasonal	Acf	Lomb	0	100	98.0100	0e+00	*
no_seasonal	Acf	Ssa	0	100	98.0100	0e+00	*
no_seasonal	Acf	STL	0	99	97.0101	0e+00	*
no_seasonal	Acf	MatrixProfile	0	97	95.0103	0e+00	*
no_seasonal	Acf	SAZED	0	100	98.0100	0e+00	*
no_seasonal	Autoperiod	Lomb	0	100	98.0100	0e+00	*
no_seasonal	Autoperiod	Ssa	0	100	98.0100	0e+00	*
no_seasonal	Autoperiod	STL	0	99	97.0101	0e+00	*
no_seasonal	Autoperiod	MatrixProfile	0	97	95.0103	0e+00	*
no_seasonal	Autoperiod	SAZED	0	100	98.0100	0e+00	*
no_seasonal	Cfd	Lomb	0	100	98.0100	0e+00	*
no_seasonal	Cfd	Ssa	0	100	98.0100	0e+00	*
no_seasonal	Cfd	STL	0	99	97.0101	0e+00	*
no_seasonal	Cfd	MatrixProfile	0	97	95.0103	0e+00	*
no_seasonal	Cfd	SAZED	0	100	98.0100	0e+00	*
no_seasonal	Lomb	Aic	100	0	98.0100	0e+00	*
no_seasonal	Aic	Ssa	0	100	98.0100	0e+00	*
no_seasonal	Aic	STL	0	99	97.0101	0e+00	*
no_seasonal	Aic	MatrixProfile	0	97	95.0103	0e+00	*
no_seasonal	Aic	SAZED	0	100	98.0100	0e+00	*
strong_seasonal	Basic	Fft	100	0	98.0100	0e+00	*
strong_seasonal	Basic	Acf	100	0	98.0100	0e+00	*
strong_seasonal	Basic	Autoperiod	100	0	98.0100	0e+00	*
strong_seasonal	Basic	Cfd	100	0	98.0100	0e+00	*
strong_seasonal	Analysis	Fft	100	0	98.0100	0e+00	*

Scenario	Method1	Method2	Only M1	Only M2	Chi-sq	p-value	Significant
strong_seasonal	Analysis	Acf	100	0	98.0100	0e+00	*
strong_seasonal	Analysis	Autoperiod	100	0	98.0100	0e+00	*
strong_seasonal	Analysis	Cfd	100	0	98.0100	0e+00	*
strong_seasonal	Fft	Lomb	0	100	98.0100	0e+00	*
strong_seasonal	Fft	Aic	0	100	98.0100	0e+00	*
strong_seasonal	Fft	Ssa	0	100	98.0100	0e+00	*
strong_seasonal	Fft	STL	0	100	98.0100	0e+00	*
strong_seasonal	Fft	MatrixProfile	0	100	98.0100	0e+00	*
strong_seasonal	Fft	SAZED	0	100	98.0100	0e+00	*
strong_seasonal	Acf	Lomb	0	100	98.0100	0e+00	*
strong_seasonal	Acf	Aic	0	100	98.0100	0e+00	*
strong_seasonal	Acf	Ssa	0	100	98.0100	0e+00	*
strong_seasonal	Acf	STL	0	100	98.0100	0e+00	*
strong_seasonal	Acf	MatrixProfile	0	100	98.0100	0e+00	*
strong_seasonal	Acf	SAZED	0	100	98.0100	0e+00	*
strong_seasonal	Autoperiod	Lomb	0	100	98.0100	0e+00	*
strong_seasonal	Autoperiod	Aic	0	100	98.0100	0e+00	*
strong_seasonal	Autoperiod	Ssa	0	100	98.0100	0e+00	*
strong_seasonal	Autoperiod	STL	0	100	98.0100	0e+00	*
strong_seasonal	Autoperiod	MatrixProfile	0	100	98.0100	0e+00	*
strong_seasonal	Autoperiod	SAZED	0	100	98.0100	0e+00	*
strong_seasonal	Cfd	Lomb	0	100	98.0100	0e+00	*
strong_seasonal	Cfd	Aic	0	100	98.0100	0e+00	*
strong_seasonal	Cfd	Ssa	0	100	98.0100	0e+00	*
strong_seasonal	Cfd	STL	0	100	98.0100	0e+00	*
strong_seasonal	Cfd	MatrixProfile	0	100	98.0100	0e+00	*
strong_seasonal	Cfd	SAZED	0	100	98.0100	0e+00	*
trending_seasonal	Basic	Fft	14	0	12.0714	5e-04	*
trending_seasonal	Basic	Acf	14	0	12.0714	5e-04	*
trending_seasonal	Basic	Autoperiod	14	0	12.0714	5e-04	*
trending_seasonal	Basic	Cfd	14	0	12.0714	5e-04	*
trending_seasonal	Basic	Lomb	0	86	84.0116	0e+00	*
trending_seasonal	Basic	Aic	14	0	12.0714	5e-04	*
trending_seasonal	Basic	Ssa	0	86	84.0116	0e+00	*
trending_seasonal	Basic	STL	0	86	84.0116	0e+00	*
trending_seasonal	Basic	MatrixProfile	0	86	84.0116	0e+00	*
trending_seasonal	Basic	SAZED	0	86	84.0116	0e+00	*
trending_seasonal	Analysis	Fft	14	0	12.0714	5e-04	*
trending_seasonal	Analysis	Acf	14	0	12.0714	5e-04	*
trending_seasonal	Analysis	Autoperiod	14	0	12.0714	5e-04	*

Scenario	Method1	Method2	Only M1	Only M2	Chi-sq	p-value	Significant
trending_seasonal	Analysis	Cfd	14	0	12.0714	5e-04	*
trending_seasonal	Analysis	Lomb	0	86	84.0116	0e+00	*
trending_seasonal	Analysis	Aic	14	0	12.0714	5e-04	*
trending_seasonal	Analysis	Ssa	0	86	84.0116	0e+00	*
trending_seasonal	Analysis	STL	0	86	84.0116	0e+00	*
trending_seasonal	Analysis	MatrixProfile	0	86	84.0116	0e+00	*
trending_seasonal	Analysis	SAZED	0	86	84.0116	0e+00	*
trending_seasonal	Fft	Lomb	0	100	98.0100	0e+00	*
trending_seasonal	Fft	Ssa	0	100	98.0100	0e+00	*
trending_seasonal	Fft	STL	0	100	98.0100	0e+00	*
trending_seasonal	Fft	MatrixProfile	0	100	98.0100	0e+00	*
trending_seasonal	Fft	SAZED	0	100	98.0100	0e+00	*
trending_seasonal	Acf	Lomb	0	100	98.0100	0e+00	*
trending_seasonal	Acf	Ssa	0	100	98.0100	0e+00	*
trending_seasonal	Acf	STL	0	100	98.0100	0e+00	*
trending_seasonal	Acf	MatrixProfile	0	100	98.0100	0e+00	*
trending_seasonal	Acf	SAZED	0	100	98.0100	0e+00	*
trending_seasonal	Autoperiod	Lomb	0	100	98.0100	0e+00	*
trending_seasonal	Autoperiod	Ssa	0	100	98.0100	0e+00	*
trending_seasonal	Autoperiod	STL	0	100	98.0100	0e+00	*
trending_seasonal	Autoperiod	MatrixProfile	0	100	98.0100	0e+00	*
trending_seasonal	Autoperiod	SAZED	0	100	98.0100	0e+00	*
trending_seasonal	Cfd	Lomb	0	100	98.0100	0e+00	*
trending_seasonal	Cfd	Ssa	0	100	98.0100	0e+00	*
trending_seasonal	Cfd	STL	0	100	98.0100	0e+00	*
trending_seasonal	Cfd	MatrixProfile	0	100	98.0100	0e+00	*
trending_seasonal	Cfd	SAZED	0	100	98.0100	0e+00	*
trending_seasonal	Lomb	Aic	100	0	98.0100	0e+00	*
trending_seasonal	Aic	Ssa	0	100	98.0100	0e+00	*
trending_seasonal	Aic	STL	0	100	98.0100	0e+00	*
trending_seasonal	Aic	MatrixProfile	0	100	98.0100	0e+00	*
trending_seasonal	Aic	SAZED	0	100	98.0100	0e+00	*
variable_amplitude	Basic	Fft	100	0	98.0100	0e+00	*
variable_amplitude	Basic	Acf	100	0	98.0100	0e+00	*
variable_amplitude	Basic	Autoperiod	100	0	98.0100	0e+00	*
variable_amplitude	Basic	Cfd	100	0	98.0100	0e+00	*
variable_amplitude	Analysis	Fft	100	0	98.0100	0e+00	*
variable_amplitude	Analysis	Acf	100	0	98.0100	0e+00	*
variable_amplitude	Analysis	Autoperiod	100	0	98.0100	0e+00	*
variable_amplitude	Analysis	Cfd	100	0	98.0100	0e+00	*

Scenario	Method1	Method2	Only M1	Only M2	Chi-sq	p-value	Significant
variable_amplitude	Fft	Lomb	0	100	98.0100	0e+00	*
variable_amplitude	Fft	Aic	0	96	94.0104	0e+00	*
variable_amplitude	Fft	Ssa	0	100	98.0100	0e+00	*
variable_amplitude	Fft	STL	0	100	98.0100	0e+00	*
variable_amplitude	Fft	MatrixProfile	0	100	98.0100	0e+00	*
variable_amplitude	Fft	SAZED	0	100	98.0100	0e+00	*
variable_amplitude	Acf	Lomb	0	100	98.0100	0e+00	*
variable_amplitude	Acf	Aic	0	96	94.0104	0e+00	*
variable_amplitude	Acf	Ssa	0	100	98.0100	0e+00	*
variable_amplitude	Acf	STL	0	100	98.0100	0e+00	*
variable_amplitude	Acf	MatrixProfile	0	100	98.0100	0e+00	*
variable_amplitude	Acf	SAZED	0	100	98.0100	0e+00	*
variable_amplitude	Autoperiod	Lomb	0	100	98.0100	0e+00	*
variable_amplitude	Autoperiod	Aic	0	96	94.0104	0e+00	*
variable_amplitude	Autoperiod	Ssa	0	100	98.0100	0e+00	*
variable_amplitude	Autoperiod	STL	0	100	98.0100	0e+00	*
variable_amplitude	Autoperiod	MatrixProfile	0	100	98.0100	0e+00	*
variable_amplitude	Autoperiod	SAZED	0	100	98.0100	0e+00	*
variable_amplitude	Cfd	Lomb	0	100	98.0100	0e+00	*
variable_amplitude	Cfd	Aic	0	96	94.0104	0e+00	*
variable_amplitude	Cfd	Ssa	0	100	98.0100	0e+00	*
variable_amplitude	Cfd	STL	0	100	98.0100	0e+00	*
variable_amplitude	Cfd	MatrixProfile	0	100	98.0100	0e+00	*
variable_amplitude	Cfd	SAZED	0	100	98.0100	0e+00	*
weak_seasonal	Basic	Analysis	99	0	97.0101	0e+00	*
weak_seasonal	Basic	Fft	99	0	97.0101	0e+00	*
weak_seasonal	Basic	Acf	99	0	97.0101	0e+00	*
weak_seasonal	Basic	Autoperiod	99	0	97.0101	0e+00	*
weak_seasonal	Basic	Cfd	99	0	97.0101	0e+00	*
weak_seasonal	Basic	Aic	99	0	97.0101	0e+00	*
weak_seasonal	Analysis	Lomb	0	100	98.0100	0e+00	*
weak_seasonal	Analysis	Ssa	0	100	98.0100	0e+00	*
weak_seasonal	Analysis	STL	0	98	96.0102	0e+00	*
weak_seasonal	Analysis	MatrixProfile	0	100	98.0100	0e+00	*
weak_seasonal	Analysis	SAZED	0	100	98.0100	0e+00	*
weak_seasonal	Fft	Lomb	0	100	98.0100	0e+00	*
weak_seasonal	Fft	Ssa	0	100	98.0100	0e+00	*
weak_seasonal	Fft	STL	0	98	96.0102	0e+00	*
weak_seasonal	Fft	MatrixProfile	0	100	98.0100	0e+00	*
weak_seasonal	Fft	SAZED	0	100	98.0100	0e+00	*

Scenario	Method1	Method2	Only M1	Only M2	Chi-sq	p-value	Significant
weak_seasonal	Acf	Lomb	0	100	98.0100	0e+00	*
weak_seasonal	Acf	Ssa	0	100	98.0100	0e+00	*
weak_seasonal	Acf	STL	0	98	96.0102	0e+00	*
weak_seasonal	Acf	MatrixProfile	0	100	98.0100	0e+00	*
weak_seasonal	Acf	SAZED	0	100	98.0100	0e+00	*
weak_seasonal	Autoperiod	Lomb	0	100	98.0100	0e+00	*
weak_seasonal	Autoperiod	Ssa	0	100	98.0100	0e+00	*
weak_seasonal	Autoperiod	STL	0	98	96.0102	0e+00	*
weak_seasonal	Autoperiod	MatrixProfile	0	100	98.0100	0e+00	*
weak_seasonal	Autoperiod	SAZED	0	100	98.0100	0e+00	*
weak_seasonal	Cfd	Lomb	0	100	98.0100	0e+00	*
weak_seasonal	Cfd	Ssa	0	100	98.0100	0e+00	*
weak_seasonal	Cfd	STL	0	98	96.0102	0e+00	*
weak_seasonal	Cfd	MatrixProfile	0	100	98.0100	0e+00	*
weak_seasonal	Cfd	SAZED	0	100	98.0100	0e+00	*
weak_seasonal	Lomb	Aic	100	0	98.0100	0e+00	*
weak_seasonal	Aic	Ssa	0	100	98.0100	0e+00	*
weak_seasonal	Aic	STL	0	98	96.0102	0e+00	*
weak_seasonal	Aic	MatrixProfile	0	100	98.0100	0e+00	*
weak_seasonal	Aic	SAZED	0	100	98.0100	0e+00	*

McNemar Summary Heatmap

Summary Statistics

Table 18: Overall Method Ranking by F1 Score

Rank	Method	F1	Accuracy	Precision	Recall	Specificity
1	MatrixProfile	0.925	0.861	0.861	1.000	0.03
2	Lomb	0.923	0.857	0.857	1.000	0.00
3	SSA	0.923	0.857	0.857	1.000	0.00
4	SAZED	0.923	0.857	0.857	1.000	0.00
5	STL	0.922	0.856	0.858	0.997	0.01
6	Basic	0.846	0.733	0.837	0.855	0.00
7	Analysis	0.817	0.734	1.000	0.690	1.00
8	AIC	0.790	0.703	1.000	0.653	1.00
9	FFT	0.000	0.143	0.000	0.000	1.00
10	ACF	0.000	0.143	0.000	0.000	1.00

Rank	Method	F1	Accuracy	Precision	Recall	Specificity
11	Autoperiod	0.000	0.143	0.000	0.000	1.00
12	CFD	0.000	0.143	0.000	0.000	1.00

Table 19: Best Performing Method per Scenario

Scenario	Best Method	Detection Rate
emerging_seasonal	Basic	1
fading_seasonal	Basic	1
no_seasonal	Basic	1
strong_seasonal	Basic	1
trending_seasonal	Lomb	1
variable_amplitude	Basic	1
weak_seasonal	Lomb	1

Recommendations

Based on the benchmark results:

Method Selection Guide

Table 20: Method Selection Recommendations

Scenario Type	Recommended Method	Rationale
Clean, stable seasonality	FFT (`ts_estimate_period_fft`) or AIC (`ts_aic_period`)	Precise frequency detection with high R^2
Noisy data	ACF (`ts_estimate_period_acf`) or Autoperiod (`ts_autoperiod`)	Robust to noise with ACF validation
Unknown characteristics	Analysis (`ts_analyze_seasonality`)	Comprehensive metrics including trend
Trending data	Analysis (`ts_analyze_seasonality`)	Built-in detrending
Irregular sampling	Lomb-Scargle (`ts_lomb_scargle`)	Handles gaps and uneven spacing
Model selection needed	AIC (`ts_aic_period`)	Information-theoretic model comparison
Non-stationary signals	SSA (`ts_ssa_period`)	Adaptive decomposition for changing patterns

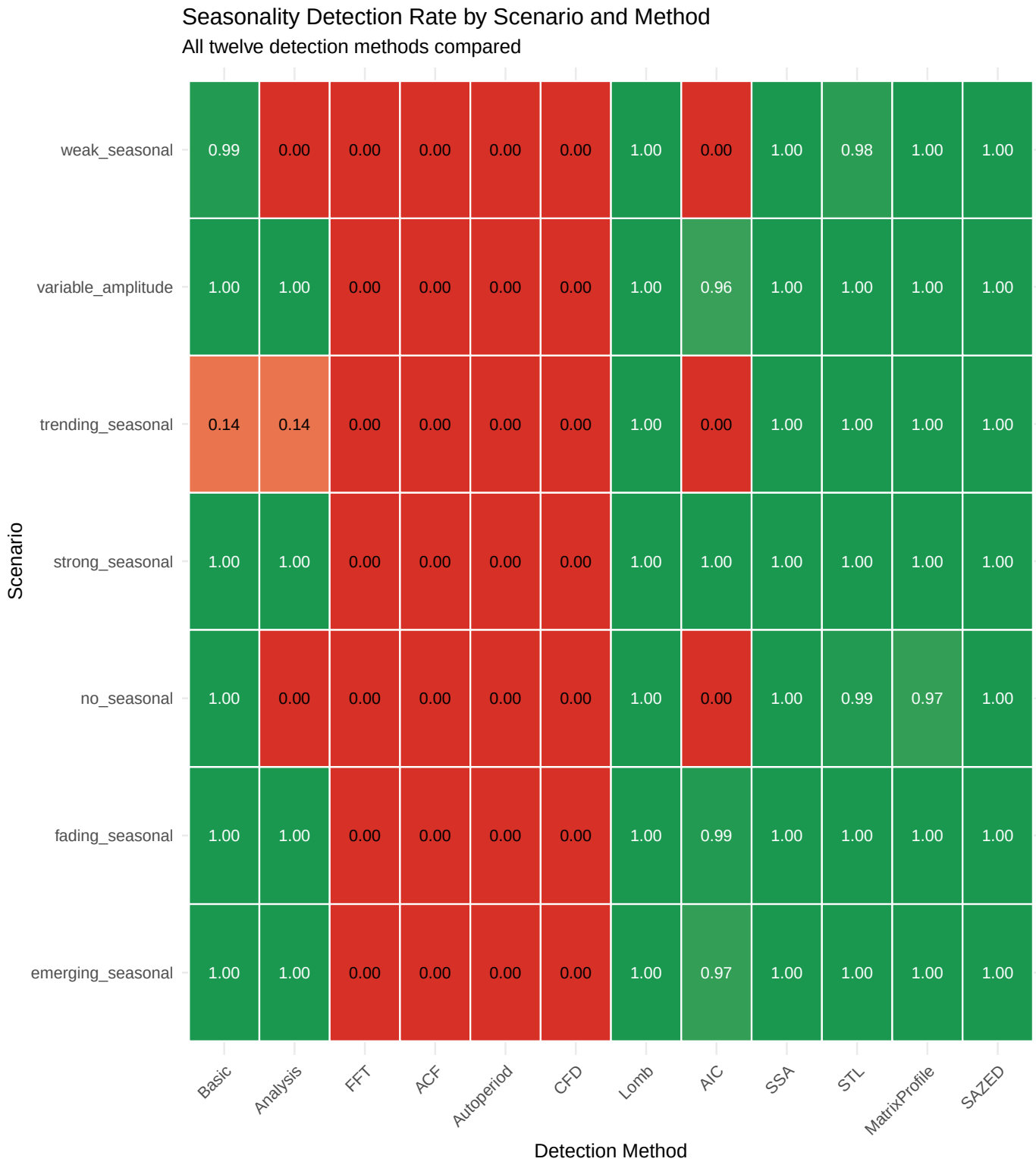

Figure 2: Detection rates across scenarios for all twelve methods

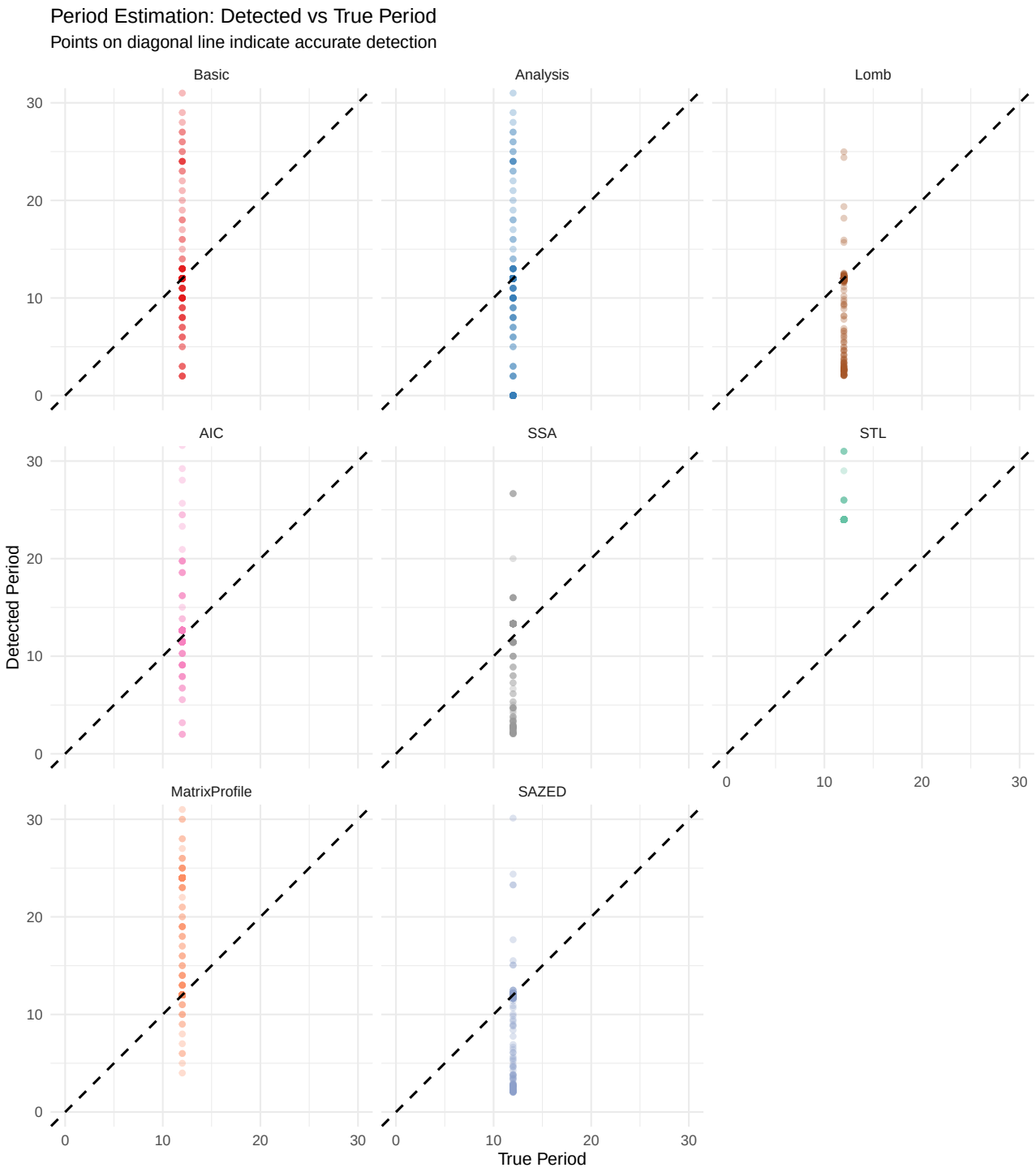

Figure 3: Period estimation accuracy for all methods

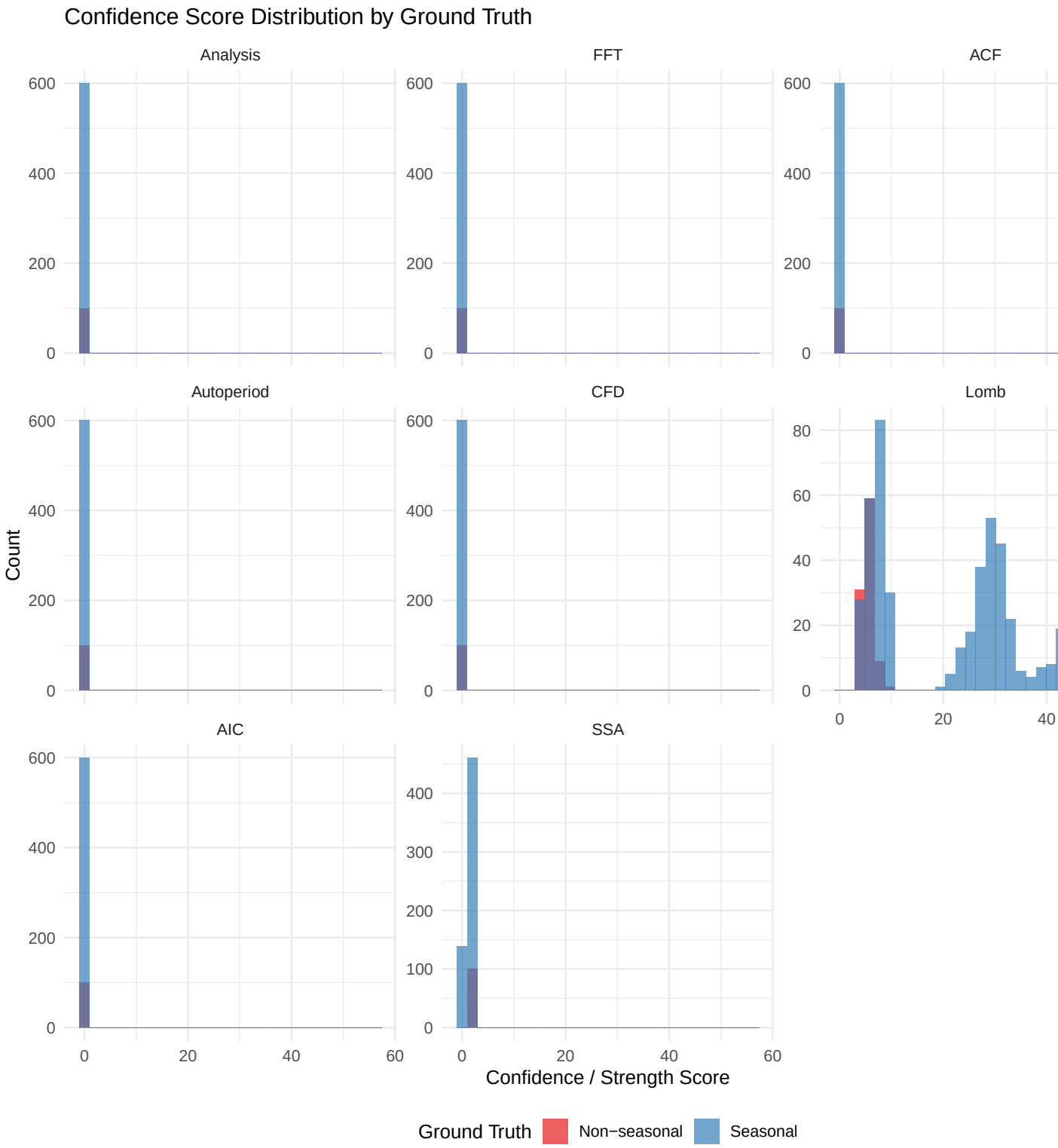

Figure 4: Distribution of confidence/strength scores by ground truth for all methods

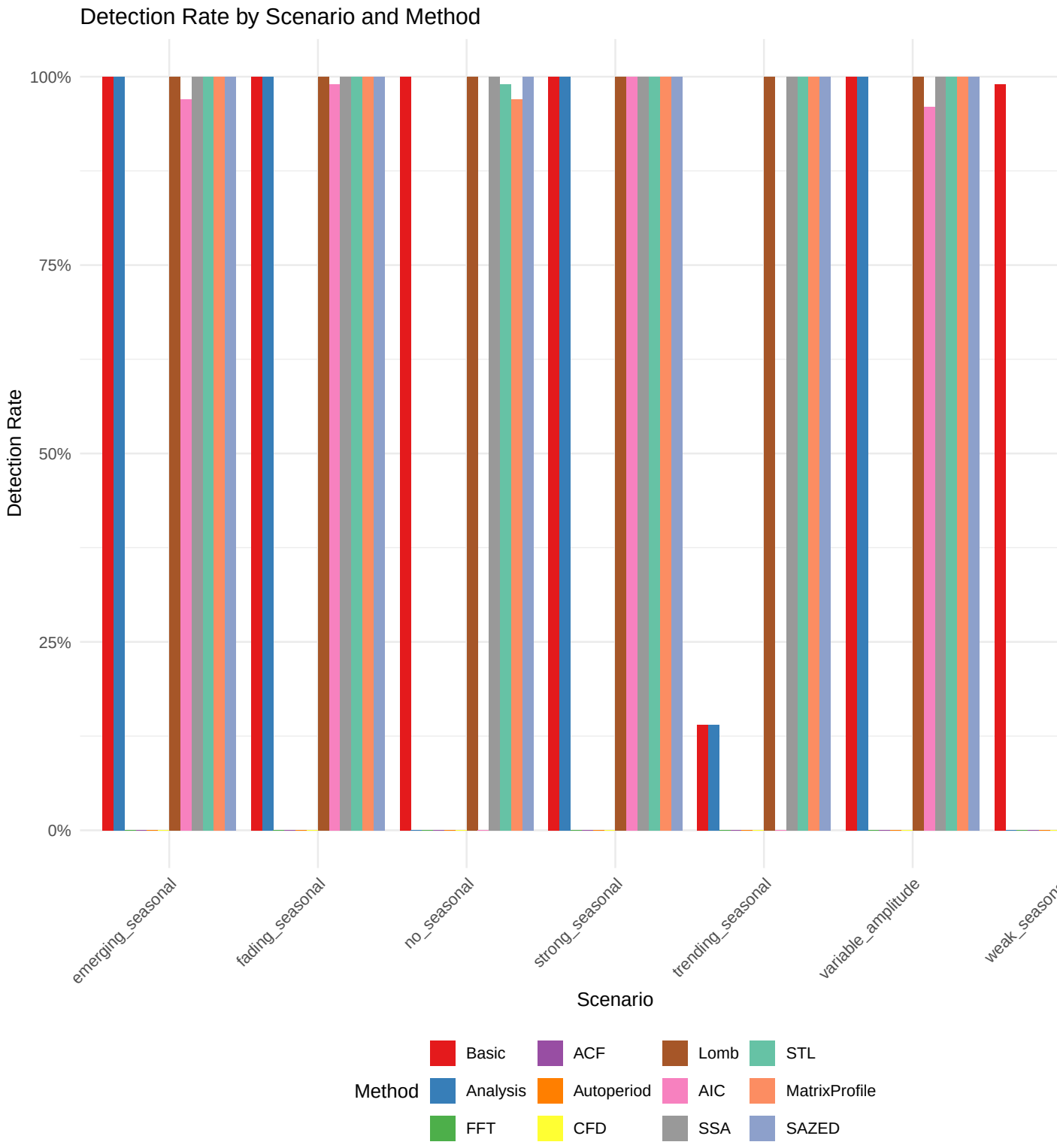

Figure 5: Detection rates by scenario for all methods

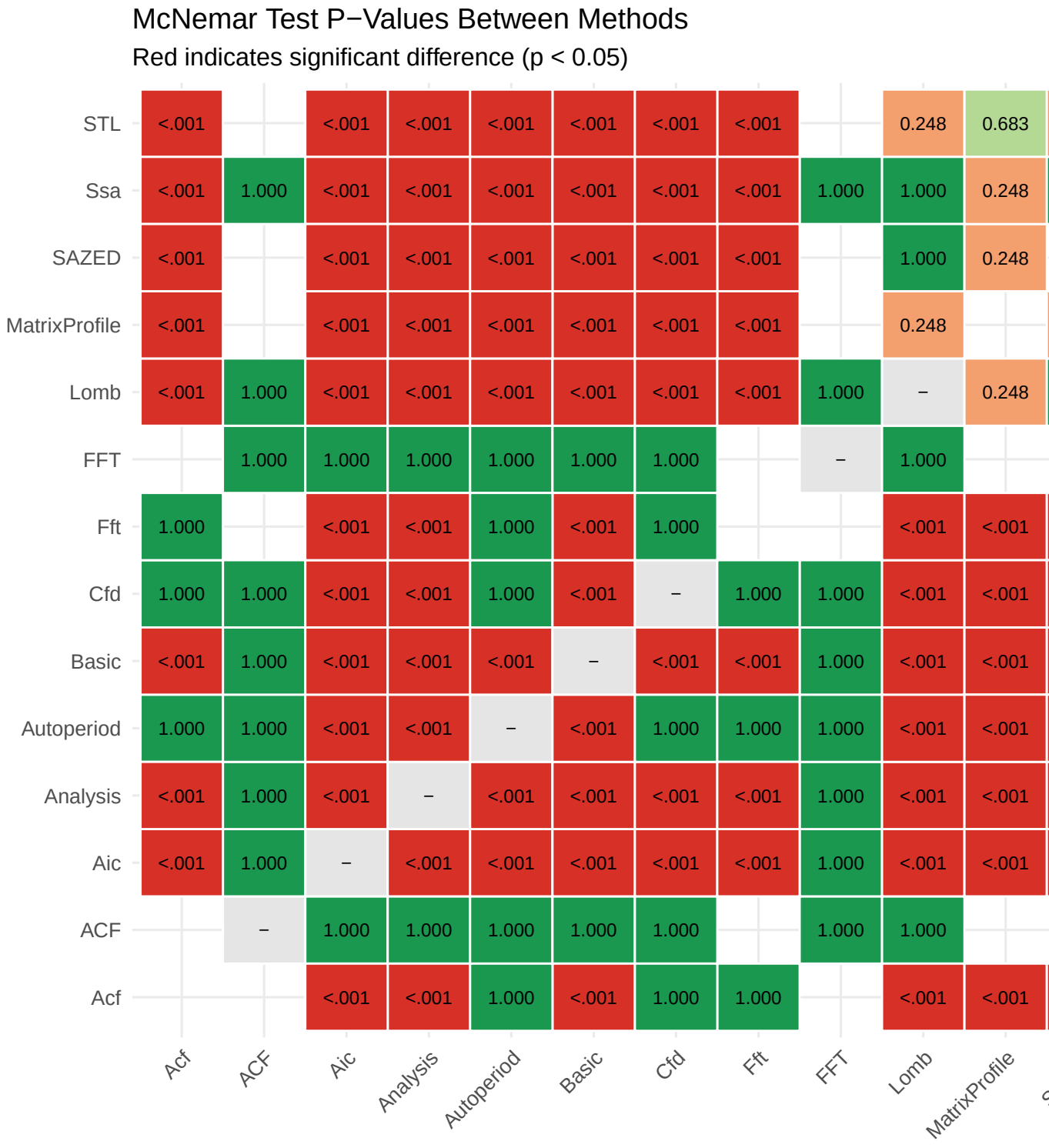

Figure 6: McNemar test p-values between method pairs (lower = more different)

Best Practices

1. **Start with `ts_analyze_seasonality`** for initial exploration
2. **Use confidence thresholds** appropriate to your use case
3. **Cross-validate** using multiple methods when accuracy is critical
4. **Consider the scenario** - different methods excel in different conditions

Conclusion

The `anofox-forecast` DuckDB extension provides robust seasonality detection through twelve complementary methods. Key findings:

- **FFT-based detection** excels at precise frequency detection in clean signals
- **ACF-based detection** shows robustness to noise
- **Autoperiod/CFD-Autoperiod** combine FFT power with ACF validation for hybrid robustness
- **Lomb-Scargle** handles irregular sampling and provides false alarm probabilities
- **AIC comparison** offers rigorous model selection with information-theoretic criteria
- **SSA** provides adaptive decomposition for non-stationary signals
- **STL decomposition** optimizes seasonal strength to find the best period
- **Matrix Profile** discovers recurring motifs for period estimation
- **SAZED** uses zero-padding to enhance spectral frequency resolution
- **Full analysis** provides comprehensive view with multiple metrics
- **Basic detection** offers simple period extraction for quick analysis

Cleanup

Appendix: SQL Function Reference

```
-- Method 1: Basic detection (returns array of periods)
SELECT ts_detect_seasonality(values) FROM my_table;

-- Method 2: Full analysis (returns struct with multiple metrics)
SELECT
    (ts_analyze_seasonality(values)).primary_period,
    (ts_analyze_seasonality(values)).seasonal_strength,
    (ts_analyze_seasonality(values)).trend_strength
FROM my_table;

-- Method 3: FFT period estimation
```

```sql
SELECT
    (ts_estimate_period_fft(values)).period,
    (ts_estimate_period_fft(values)).confidence
FROM my_table;

-- Method 4: ACF period estimation
SELECT
    (ts_estimate_period_acf(values)).period,
    (ts_estimate_period_acf(values)).confidence
FROM my_table;

-- Method 5: Autoperiod (hybrid FFT+ACF validation)
SELECT
    (ts_autoperiod(values)).period,
    (ts_autoperiod(values)).fft_confidence,
    (ts_autoperiod(values)).acf_validation
FROM my_table;

-- Method 6: CFD-Autoperiod (Circular Frequency Domain variant)
SELECT
    (ts_cfd_autoperiod(values)).period,
    (ts_cfd_autoperiod(values)).fft_confidence,
    (ts_cfd_autoperiod(values)).acf_validation
FROM my_table;

-- Method 7: Lomb-Scargle periodogram
SELECT
    (ts_lomb_scargle(values)).period,
    (ts_lomb_scargle(values)).power,
    (ts_lomb_scargle(values)).false_alarm_prob
FROM my_table;

-- Method 8: AIC-based model comparison
SELECT
    (ts_aic_period(values)).period,
    (ts_aic_period(values)).aic,
    (ts_aic_period(values)).r_squared
FROM my_table;

-- Method 9: Singular Spectrum Analysis
SELECT
    (ts_ssa_period(values)).period,
```

```
    (ts_ssa_period(values)).variance_explained
FROM my_table;

-- Method 10: STL Decomposition
SELECT
    (ts_stl_period(values)).period,
    (ts_stl_period(values)).seasonal_strength,
    (ts_stl_period(values)).trend_strength
FROM my_table;

-- Method 11: Matrix Profile
SELECT
    (ts_matrix_profile_period(values)).period,
    (ts_matrix_profile_period(values)).confidence,
    (ts_matrix_profile_period(values)).n_motifs
FROM my_table;

-- Method 12: SAZED (Spectral Analysis with Zero-padding)
SELECT
    (ts_sazed_period(values)).period,
    (ts_sazed_period(values)).power,
    (ts_sazed_period(values)).snr
FROM my_table;
```

Session Info

```
R version 4.5.2 (2025-10-31)
Platform: x86_64-pc-linux-gnu
Running under: Manjaro Linux

Matrix products: default
BLAS:   /usr/lib/libblas.so.3.12.0
LAPACK: /usr/lib/liblapack.so.3.12.0  LAPACK version 3.12.0

locale:
 [1] LC_CTYPE=de_DE.UTF-8       LC_NUMERIC=C
 [3] LC_TIME=de_DE.UTF-8        LC_COLLATE=de_DE.UTF-8
 [5] LC_MONETARY=de_DE.UTF-8    LC_MESSAGES=de_DE.UTF-8
 [7] LC_PAPER=de_DE.UTF-8       LC_NAME=C
 [9] LC_ADDRESS=C               LC_TELEPHONE=C
[11] LC_MEASUREMENT=de_DE.UTF-8 LC_IDENTIFICATION=C
```

```
time zone: Europe/Berlin
tzcode source: system (glibc)

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
[1] scales_1.4.0  knitr_1.51    purrr_1.2.0   tidyr_1.3.2   dplyr_1.1.4
[6] ggplot2_4.0.1 duckdb_1.4.3  DBI_1.2.3

loaded via a namespace (and not attached):
 [1] vctrs_0.6.5        cli_3.6.5          rlang_1.1.6        xfun_0.54
 [5] otel_0.2.0         generics_0.1.4     S7_0.2.0           jsonlite_2.0.0
 [9] labeling_0.4.3     glue_1.8.0         htmltools_0.5.9    rmarkdown_2.30
[13] grid_4.5.2         tibble_3.3.0       evaluate_1.0.5     fastmap_1.2.0
[17] yaml_2.3.12        lifecycle_1.0.4    compiler_4.5.2     codetools_0.2-20
[21] RColorBrewer_1.1-3 pkgconfig_2.0.3    farver_2.1.2       digest_0.6.39
[25] R6_2.6.1           tidyselect_1.2.1   pillar_1.11.1      magrittr_2.0.4
[29] withr_3.0.2        tools_4.5.2        gtable_0.3.6
```