Forecasting seasonal data and nonparametric unit-root tests

Robert M. Kunst

Institute for Advanced Studies, Vienna, and University of Vienna

Abstract

Nonparametric unit-root tests are a useful addendum to the toolbox of time-series analysis. They tend to trade off power for enhanced robustness features. We consider a variant of the RURS (seasonal range unit roots) test statistic, a variant of the level-crossings count adapted to classes of seasonal patterns, and a new combined test. These tests exploit two main characteristics of seasonal unit-root models, the range expansion typical of integrated processes and the low frequency of changes among main seasonal shapes.

In standard designs, usual parametric tests based on the HEGY (Hylleberg, Engle, Granger, Yoo, 1990) dominate nonparametric rival tests in large samples. In small samples, however, surprising local power gains by range tests have been reported. It is of interest whether such power advantages transfer into enhanced predictive accuracy out of sample, particularly as decision rules based on HEGY tests have been shown to offer poor predictive performance.

This contribution explores the consequences of test-based decisions for predictions of seasonal time series. Apart from generating processes with seasonal unit roots and with deterministic seasonality, also processes with seasonal time deformation are considered.

The nature of seasonal cycles in time series

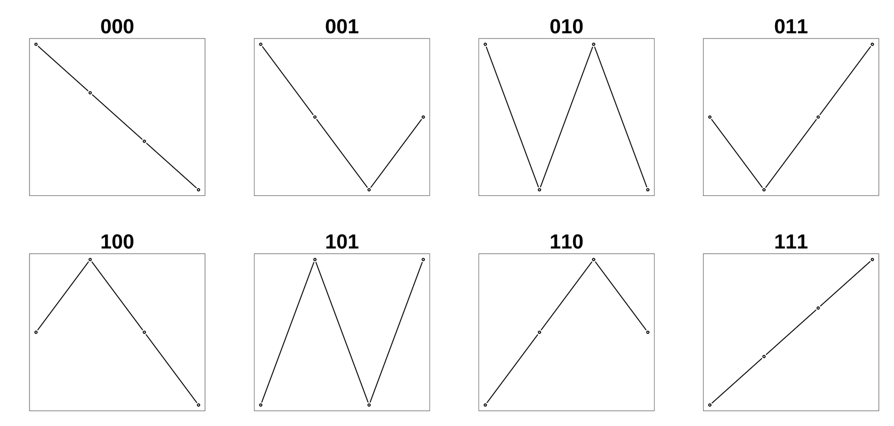
Seasonal cycles may be **repetitive** and well modeled by time-constant **deterministic** cycles. Changes of basic patterns remain episodic.

They may also be rarely but **persistently changing**. Such behavior is often seen as well modeled via **seasonal unit roots**. The simplest seasonal unit-root process is the **seasonal random walk** (SRW) $x_t = x_{t-4} + \varepsilon_t$ with white-noise (ε_t) for quarterly data.

There exist hypothesis tests for discriminating between these two concepts. These tests have comparatively low power in samples of typical size. We consider two types of tests: parametric tests of the HEGY type and non-parametric tests as inspired by jittered seasonal phase plots.

Eight qualitative seasonal patterns for quarterly data

In a first step, the intra-yearly seasonal patterns are classified into eight shape types. Deterministic seasonality should cause a return to the same shape class, unit roots persistent moves to different classes.

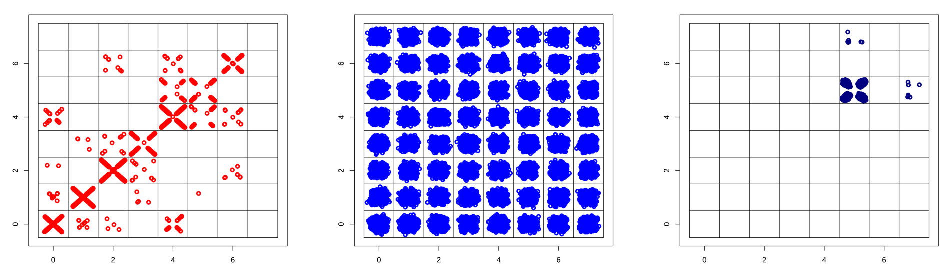


The shapes can be identified with binary number codes (0 for falling and 1 for rising movements), and thus with the binary representation of integers.

Jittered phase plots

Notation: we use x_t for the observed data at quarters $t = 1, \ldots, T$, index $\tau = 1, \ldots, N$ for years, $m(x_\tau)$ for the corresponding seasonal shape class, $(\tilde{x}_\tau, \tilde{x}_{\tau+1})$ for the coordinates within the bin.

The plots represent τ to $\tau+1$ transitions between shape classes. **Shallow** shapes close to constancy between quarters are in the center of the bin, **deep** shapes lie at the boundaries. Allocation to the left and right parts of bins is performed after binary randomization (jittering), such that X–type crosses (saltires) emerge for deep-to-deep and shallow-to-shallow transitions.



10,000 years of an SRW with Gaussian errors (left), of a random walk (center), and of the process with deterministic seasonality $x_t = 0.4x_{t-4} + \sum_{j=1}^{4} d_t + \varepsilon_t$, with $(d_1, \ldots, d_4) = (0, 8, 3, 10)$ (right).

Three non-parametric tests for seasonal unit roots

- The test statistic ζ_1 is based on **counting transitions** between seasonal shape classes. In detail, we adapt the zero-crossings count for random walks (Burridge, Guerre 1996) to shape classes:

$$\zeta_1 = \frac{\hat{\sigma}}{\widehat{\mathrm{MAD}}} N^{-0.5} \sum_{\tau=2}^{N} I(m(x_{\tau-1}) \neq m(x_\tau));$$

- The test statistic ζ_2 measures the **median distance from** the pure diagonal **saltire shape**

$$\zeta_2 = N^{0.5} \mathrm{med} ||\tilde{x}_\tau| - |\tilde{x}_{\tau-1}||/\sqrt{2},$$

  it is related to the RUR (range unit root) test by Aparicio, Escribano, Sipols (2006);
- The statistic ζ is the linear combination

$$\zeta = \frac{\zeta_1 + 17\,\zeta_2}{18},$$

  where the value 17 evolves from numerically maximizing local power.

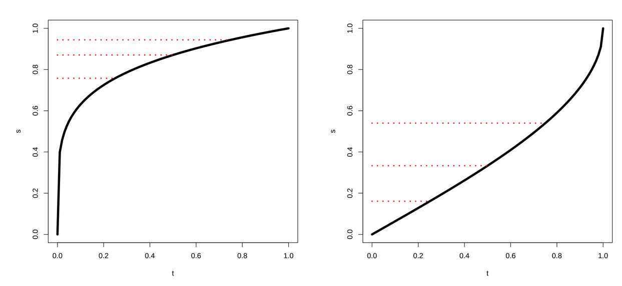
Distributions of ζ_1, ζ_2 are sensitive to deviations from the pure SRW under the null, a correction using long-run variances is called $\tilde{\zeta}_1$ and the evolving combined statistic $\tilde{\zeta}$. In small samples, a test based on $\tilde{\zeta}$ indeed turns out to be more powerful than tests based on ζ_2 or on $\tilde{\zeta}_1$.

Designs for forecasting experiments

1. In $x_t = \phi x_{t-4} + \varepsilon_t$, ϕ can be varied: $\phi = 1$ corresponds to the unit-root null (for all roots). $0.5 < \phi < 1$ to the alternative. We expect forecasts using the alternative model

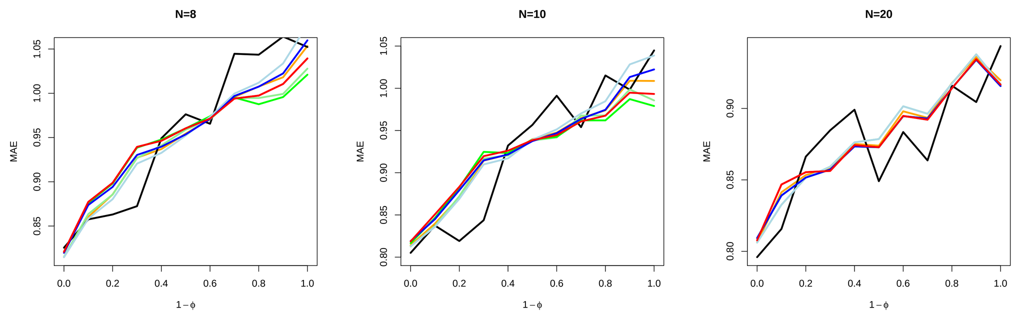
$$x_t = \sum_{j=1}^{4} \phi_j x_{t-j} + \sum_{j=1}^{4} d_j \delta_{j,t} + u_t$$

   to be more precise than those based on a SRW with drift if $\phi < 1$;
2. In models with **seasonal time deformation**, performance of all rival models is *a priori* uncertain. These models belong to neither the null nor the alternative.



Strongly persistent stable autoregressive processes are generated at high frequency, deformed by functions $s = t^\delta$ ('Box-Cox', left for $\delta = 0.2$) or $s = \frac{2}{\pi}\arcsin t$ (right), and aggregated into quarters.

Results of the prediction experiments



Graphs report the results for the first simulation design. Decisions based on HEGY–type tests (gray) are often better close to the null and at a large distance from the null. There are windows in an intermediate region, where the non-parametric test delivers better forecasts (color). 10% tests (more intense) preferable to 5% tests.

With seasonal time deformation, none of the prediction models is correctly specified. Seasonal unit-root tests tend to reject the null, even for weak deformation ($t^{0.9}$). Forecasts based on the dummy model are much better than those based on seasonal differences. Parametric tests reject more often, thus they deliver better forecasts on average.

Empirical applications

variable	% better by dummies
Austrian GDP	78
U.K. GDP	59
Heathrow precipitation	58
Heathrow temperature	55
industrial production	60
unemployment rate	70

Most real-life data examples support deterministic seasonal features rather than seasonal unit roots. Dummy-based models tend to predict better than those based on seasonal differencing, even if hypothesis tests support unit roots, as for some economic variables.

Summary and outlook

- Test power alone does not necessarily imply improved predictive accuracy for test-based selection of forecast models;
- Seasonal differences cannot be blindly recommended: forecast models with fixed seasonal dummies often deliver superior forecasts;
- In standard cases, the non-parametric test competes bravely with the parametric HEGY–type tests. In some non-standard cases, the parametric tests reject more often, which then benefits forecasting;
- Testing at 10% is usually better than at 5%: note that the asymptotically forecast-optimizing AIC operates at even more liberal rates;
- Identical power does not imply identical performance for test-based forecasts. What are the characteristics of trajectories from the same process that are classified differently by different tests?
- Forecasts based on the component tests ζ_1, ζ_2 do not perform much worse than those based on ζ. The optimum weight for prediction may differ from the power-optimizing weight;
- Breaks, outliers, non-normal distributions are often used to advertise non-parametric testing. Does the forecast performance for these cases based on ζ reflect this? Does this help with some real-life variables?

References

Aparicio F, Escribano A, Sipols AE (2006): Range unit-root (RUR) tests: robust against nonlinearities, error distributions, structural breaks and outliers. *J. of Time Series Analysis* **27**, 545–576.

Burridge P, Guerre E (1996) The Limit Distribution of Level Crossings of a Random Walk, and a Simple Unit Root Test. *Econometric Theory* **12**, 705–723.

Hylleberg S, Engle RF, Granger, CWJ, Yoo, BS (1990): Seasonal integration and cointegration. *J. of Econometrics* **44**, 215–238.