**APPENDIX 2: DETAILED METHOD DESCRIPTIONS**

# METHODS DESCRIPTION

## Edwards’ test

The Edwards' test was originally based on the approach outlined in [9] whereby a sine curve is fit and from the fitted curve, estimates of the "seasonal intensity" (peak-to-low ratio of the fitted curve) can be obtained. This approach was modified by Brookhart and Rothman in [10] due to the bias found in the Edwards estimator when dealing with rare events or events exhibiting very large or very small amounts of seasonality. The approach was further modified by Weinstein et al. [11], by using a Poisson generalized linear model to model event counts as a sinusoidal pattern along with a linear component for long-term trends. The model for the expected count of events E[Ni] in time period i is as follows:

where τ is the number of months since the study start date,

and

The ratio of peak-to-low occurrence rates of the process is given by

A value of R > 1 indicates that peaks are higher than troughs in the fitted model, meaning there is seasonality. The null hypothesis, R = 1, is tested against the alternative hypothesis that R > 1. The implementation and notation follow [11].

## Friedman’s test

The Friedman test is a non-parametric method for determining whether samples are drawn from the same population or from populations with equal medians. Friedman’s test was originally introduced in [12], where the procedure is described as the “method of ranks” because it relies on ranks of the observations rather than distributional assumptions. Let be the rank of the ith observation in the jth year and . The test statistic is expressed as

Where n is the number of observations in each period i, and . The null hypothesis that there is no seasonality follows a distribution with k-1 degrees of freedom. The implementation and notation follow [21,22].

## Auto ARIMA test

The auto ARIMA test uses a variation of the Hyndman-Khandakar algorithm [23], which combines unit root tests, minimization of the AICc and MLE to obtain an ARIMA model. The test considers a time series seasonal if the model chosen contains a seasonal component. The implementation follows [13,14,19,24].

## ARIMA hypothesis test

The ARIMA hypothesis test compares an optimally chosen ARIMA model with a seasonal component (chosen by using the auto ARIMA test) with an optimally chosen ARIMA model that does not contain a seasonal component (again, chosen by using the auto ARIMA test). Let N1 and N2 be the number of parameters estimated by models and respectively. The series is considered seasonal if the seasonal component is statistically significant by Wilks’ Theorem [25] and the test statistic is given by

where the difference will be approximately with Δ = N1 - N2 degrees of freedom. The implementation follows [13,14,19,24] .

## QS test

The QS test is based on [15]. The test statistic QS is computed by examining the autocorrelation function at appropriate seasonal lags. Let n be the number of elements in the series and m the seasonal frequency (m=12 for monthly data). The test statistic is given by

where the autocovariance is defined as and the autocorrelation is defined as , for the time series . The probability distribution is approximately with 2 degrees of freedom. The implementation and notation follow [21,22].

## ETS hypothesis test

The ETS test is an exponential smoother that implements a state space approach [14]. Each model consists of a measurement equation that describes the observed data, and some state equations that describe how the unobserved components or states (level, trend, seasonal) change over time [14] . Hence, these are referred to as state space models. Each state space model is labeled as ETS (Error, Trend, Seasonal) [14] . This label can also be thought of as ExponenTial Smoothing [14] . Model selection is implemented by utilizing likelihood maximization and information criterion (AIC, AICc, BIC) depending on the nature of the errors (additive or multiplicative) [14] . If a model with a seasonal component is chosen, it is tested for statistical significance using Wilks’ theorem as described in the Arima Hypothesis test above. The implementation follows [13,14,18,19,24] .

## Kruskal-Wallis test

The Kruskal-Wallis test [16] is similar to the Friedman test (same null hypothesis), but the test statistic is computed as

where ranks are produced over the entire sample N rather than each period year. The test follows a distribution with k-1 degrees of freedom under the null hypothesis. The implementation and notation follow [21,22].

## Welch’s test

As explained in [21], Welch’s test [17] is essentially a one-way ANOVA without repeated measures for heteroskedastic data. Let . The test statistic is

where . WE follows an F distribution under the null hypothesis, with k-1 and degrees of freedom. The implementation and notation follow [21,22].