Assessing Strength of Seasonality Through Sample Entropy: A Simulation Study

John Carlo P. Daquis, Maria Lizeth M. Laus and Nikki E. Supnet
University of the Philippines Diliman

This paper investigates the behaviour of sample entropy when used as a measure of seasonality of time series. Sample entropy decreases when the series becomes less complex or when regular patterns emerge. The more regular patterns in seasonal data compared to those of non-seasonal data is used in providing evidence that sample entropy is inversely related to the likelihood that seasonality exists in the data. A simulation study was conducted to assess the behaviour of the sample entropy in relation to seasonality. Sample entropy yields large values for time series without seasonality, and as the extent of seasonality becomes dominant, the value decreases. The sample entropy becomes a more reliable measure of seasonality as the length of the time series increases.

Keywords: entropy, sample entropy, seasonality, time series

1. Introduction

A time series is an ordered sequence of observations taken over a period such as month, quarter or year. We find time series data in fields such as agriculture, business and economics, engineering, medical studies and many other fields. Many time series contain a seasonal phenomenon that repeats itself after a regular period of time.

Seasonality often arises from factors such as technology and calendar related influences such as natural conditions like weather fluctuations that are representative of the season; business and administrative procedures like start and end of school terms; and social and cultural behaviour like Christmas. Moreover, seasonality can also be caused by calendar related systematic effects that are not stable in their annual timing such as trading day effect or the recurring effects associated with individual days of the week; and the moving holiday effect or the effect of holidays which occur each year but whose exact timing shifts.

Along with other features of data like trend and cycle, seasonality is very important in time series analysis. There are existing tests and procedures for detecting and assessing seasonality in a time series.

The easiest way to detect seasonality is to plot the time series data. If the graph shows cyclic patterns, then seasonality may indeed be present. Once the periodicity of the seasonality is known, the strength of seasonality can further be inspected using seasonal subseries plots (Cleveland, 1993) wherein the observed values are ordered by seasonality period. Boxplots are used as an alternative to seasonal subseries plots especially for long time series. If the period of seasonality is not known, then autocorrelation plots (Box and Jenkins, 1976) or the spectral plots (Jenkins and Watts, 1968) can be used to determine it.

Aside from the graphical methods, several tests and procedures are designed for the objective of detecting seasonality. Moineddin et al. (2003) proposed that the coefficient of determination of an autoregressive model fitted to the time series can quantify the strength of seasonality. Test statistics based on spectral peaks were developed by McElroy and Holan (2009) in detecting for seasonality. Several other tests were developed to test for seasonality, some of which are found and discussed in Edwards (1961), Hewitt et al. (1971), Walter and Elwood (1975), Freedman (1979), and Franses (1992).

Seen graphically, seasonal data have more regular cyclic patterns than non-seasonal data. Thus, under the assumption that other identifiable patterns (e.g. trend) are not present in a time series other than seasonality, a series with no seasonality obviously exhibits a more random movement than a seasonal one. This fact gave rise to the idea of using entropy as a measure of seasonality strength. In physics, entropy is defined as the amount of disorder in a system. That is, a more chaotic system has higher entropy. In information theory,entropy is defined as measure of uncertainty associated with a random variable. Shannon (1948) in finding fundamental limits on signal processing operations, defined the Shannon entropyH(X) of a random variable X represented in expectation form as $H(X) = -E\left(\log_b p(x)\right)$ where $p(x)$ is the probability that X takes on a particular value x.. The value of b is usually 2 or e. The expression $-log_b\, p(x)$ is interpreted as the amount of information contained in the outcome X = x. Thus Shannon entropy is a measure of the amount of information contained in the random variable X.

The value of entropy is large for complex data or for data with irregularities. When regular patterns emerge, the data becomes simpler and entropy decreases. Hence a completely deterministic system has zero entropy. One can view seasonality as patterns in a time series. We thus hypothesized entropy to be small in seasonal data and decreases as seasonality strength increases.

We propose sample entropy as a descriptive measure for seasonality strength. Simulated data of varying seasonality strength and length are considered in investigating the behaviour of the sample entropy estimates. The procedure is demonstrated in actual data to complement the results of the simulation study.

2. Information Measures

Perhaps one of the most important measure based on the concept of information theory is the Akaike Information Criterion (AIC), Akaike (1974). The AIC is defined as AIC = $2k–2\ln(L)$, where k is the number of parameters in the statistical model while $ln(L)$ is the log-likelihood function evaluated at the maximum likelihood estimate of the parameters in the model. The AIC is a useful tool in model selection from a list of candidate models, where it selects the model which best fits the data while penalizing for the number of parameters being estimated. Consider for example, an unknown true model f being estimated by candidate models $h_{i,}$ $i = 1,2,\ldots,n$. Then AIC_i is the amount of information lost from using h_i in representing f. Thus, one chooses the model with minimum AIC.

Entropy is also used as a basis of several goodness-of-fit tests. Vasicek (1976) proposed a normality test based on sample entropy, with results indicating that the test is comparable with other normality tests. On the other hand, Dudewicz and Van der Meulen (1981) investigated the power properties of an entropy-based test for uniformity. Gokhale (1983) proposed a generalized goodness-of-fit test. A recent goodness-of-fit test based on maximum entropy was proposed by Lee et al. (2011).

In measuring system complexity, Pincus (1991) developed a family of approximate entropy (ApEn). Establishing that a measure of system complexity, however, requires a large number of data points; that is why Pincus (1991) proposed the family of system parameters ApEn(m,r), where m is the pattern length and r is the criterion similarity; and related statistics ApEn(m,r,N) where N is the number of observations, as means of measuring the regularity or the rate of generation of new information that can be applied to typically short and noisy time series clinical data. Pincus (1991) developed the approximate entropy based on the works of Grassberger and Procaccia (1983), Takens (1983) and Eckmann and Ruelle (1985).

The mean rate of new information creation, known as Kolmogorov-Sinai (KS) entropy, is a useful parameter to characterize the system dynamics. Grassberger and Procaccia (1983) used the KS entropy as a motivation to calculate the rate of information generation in a time series data. Takens (1983) on the other hand, generalized the measure by introducing a distance metric; and Eckmann and Ruelle (1985) then modified the Takens formula to calculate the KS entropy for the physical invariant measure presumed to underlie the data distribution. Eckmann and Ruelle (1985) defined entropy (ER Entropy) as follows. For a given time series data $\{y(N)\} = \{y_1, y_2,\ldots, y_N\}$, form N-m+1 sequence of vectors x(i)'s defined by x(i) = [y(i),y(i+1),...,y(i+m-1)]. Then define for each i, $1\leq i \leq$ N-m+1,

$$C_i^m(r) = \frac{\text{number of j such that d[x(i), x(j)]} \leq r}{N-m+1} \quad (1)$$

where $d[x(i), x(j)] = \max_{k=1,2,\ldots,m} \{|y_{i+k-1} - y_{j+k-1}|\}$

$$\text{Define } \Phi^m(r) = (N-m+1)^{-1} \sum_{i=1}^{N-m+1} \log C_i^m(r) \quad (2)$$

Then,

$$\text{ER entropy} = \lim_{N\to\infty} \lim_{N\to\infty} \lim_{N\to\infty} [\Phi^m(r) - \Phi^{m+1}(r)] \quad (3)$$

Note that $\Phi^m(r) - \Phi^{m+1}(r)$ = average over i of the logarithm of the conditional probability that $|y_{j+m} - y_{i+m}| \leq r$, given that $|y_{j+k} - y_{i+k}| \leq r$ for $k = 0,1,\ldots,m-1$.

However, Pincus (1991) noted that the ER entropy is not suitable for experimental data since its value is infinity for a process of superimposed noise of any magnitude. Thus, the approximate entropy was introduced for analysis of short and noisy time series data by

$$ApEn\ (m(r) = \lim_{N\to\infty} [\Phi^m(r) - \Phi^{m+1}] \quad (4)$$

For N data points ApEn(m,r) is estimated by

$$ApEn\ (m,r,N) = \Phi^m(r) - \Phi^{m+1} \quad (5)$$

Low values of ApEn indicate high degree of regularity.

Pincus (1991) also demonstrated that for $m = 2$ and $N = 1000$, choices of r ranging from 0.1 to 0.2 times the standard deviation of the series would produce reasonable statistical validity of ApEn(m,r,N). However, Richman and Moorman (2000) pointed out that ApEn lead to inconsistent results and developed a new complexity measure called sample entropy (SampEn). ApEn algorithm includes "self-matches" which results to a biased estimate. This bias causes ApEn to be uniformly lower than expected for short records and to absence of relative consistency. Lack of relative consistency means that if ApEn of one data set is higher than that of another, it should, but does not remain higher for all conditions tested. Moreover, ApEn(m,r,N) suggests more similarity than is present. Since ApEn includes "self-matches" as stated earlier, ApEn(m,r,N) is certain to be defined under all circumstances. Also, the expected value of ApEn(m,r,N) is less than the parameter ApEn(m,r). The differences between SampEn and ApEn are that SampEn does not count templates as matching themselves and does not employ a template-wise strategy for calculating probabilities.

ApEn statistic is expected to increase when r decreases and independent of record length. Richman and Moorman (2000) tested ApEn and SampEn on uniform independent and identically distributed random numbers and found out that SampEn agrees with theory more closely than ApEn. Moreover, ApEn is expected to be relatively consistent. That is, when ApEn of a record S is lower than the ApEn of record T for one pair of parameter m and r, it is expected to do so for all other pairs. However, ApEn showed lack of relative consistency in some cases. Richman and Moorman (2000) found out that SampEn shows relative consistency in cases where ApEn does not. Nevertheless, though SampEn is more consonant with theory than ApEn, SampEn statistics deviate from predictions for very short data sets because of correlation of templates. Overall, SampEn statistics provide an improved evaluation of time series regularity.

Sample entropy has been a widely used tool in detecting physiological patterns. Lake et al. (2002) used the sample entropy as a regularity measure of heart rates of infants in the course of neonatal sepsis. Botting et al. (2009) used the approximate sample entropy in the prediction of cervical cancer and got very promising results. It was also used by Radomski (2009) as a tool for predicting upcoming labour.

3. Methodology

Richman and Moorman (2000) viewed *SampEn(m, r, N)* calculation as a process of sampling information about regularity in the time series where *N* is the length of the time series, *m* is the length of the sequences to be compared, and *r* is the tolerance for accepting matches.

Given a time series with n time points $\{y(n)\} = \{y_1, y_2, \ldots, y_n\}$, SampEn can be computed with the following steps.

1. Form *N-m+1* vector sequences $x_m(i)$ of length m where $x_m(i)=\{y_i, y_{i+1},\ldots,y_{i+m-1}\}$ and $1 \leq i \leq N-m+1$.
2. Compute the distance between vectors $x_m(i)$ and $x_m(j)$, i.e., $d[x(i), x(j)] = \max\{|yi+k - yj+k|\}, 0 \leq k \leq m-1$.
3. Let Bi be the number of vectors $x_m(i)$ that is within the r of xm(j), $1 \leq i \leq N-m$. That is, Bi counts the distance of $x_m(i)$ between another vector which is less than r. In calculating Bi, we call $x_m(i)$ a template and the instance that $x_m(i)$ is within r of $x_m(j)$ a template match. Unlike in ApEn, SampeEn does not compare $x_m(i)$ with itself and this results to lower bias for SampEn.

4. Compute the following:

$$B_i^m(r) = \frac{B_i}{N-m+1} \quad (6)$$

$$B^m(r) = \frac{1}{N-m}\sum_{i=1}^{N-m} B_m^i(r) \quad (7)$$

where $B^m(r)$ is the probability that any vector $x_m(j)$ is within *r* of $x_m(i)$.

5. Increase the dimension to m+1and calculate the number of $x_{m+1}(i)$ within r of $x_{m+1}(j)$ as A_i. Then, compute:

$$A_i^m(r) = \frac{A_i}{N-m+1} \quad (8)$$

$$A^m(r) = \frac{1}{N-m}\sum_{i=1}^{N-m} A_m^i(r) \quad (9)$$

6. *SampEn* is then defined as

$$SampEn(m,r) = \lim_{n\to\infty}\{-\ln\left[\frac{A^m(r)}{B^m(r)}\right]\} \quad (10)$$

and estimated by

$$SampEn(m,r,N) = -\ln\left[\frac{A^m(r)}{B^m(r)}\right] \quad (11)$$

We use (11) in estimating the sample entropy of the generated time series. Even though they are very crucial in calculating the value of SampEn, there are no existing guidelines in choosing the parameters *r* and *m*, though Pincus (1991) suggested to use m = 2 or 3 and r = 0.25 times the standard deviation of the time series.

4. Simulation Study

To assess the behaviour of sample entropy, the *AR(1)* model is generated from

$$y_t = \varphi y_{t-1} + \theta e_t \quad (12)$$

A non-seasonal stationary model is chosen so the series generated can be easily modified for seasonal variation:

$$y_t = 1000 + 0.5y_{t-1} + 10e_t \quad (13)$$

To make the time series seasonal, observations generated by (13) are multiplied correspondingly by a seasonal index. There are ten sets of temporal indices of varying seasonality strength. The standard deviations of the sets of seasonal indices are manipulated. Standard deviations of the indices indicate the strength of the seasonality that would be produced. The larger the standard deviation of the indices the stronger the seasonality would be and vice versa. Table 1 below summarizes these ten sets of seasonal indices.

Table 1. Temporal Indices Used in Simulation

TEMPORAL INDICES										
	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Set 10
1	1	0.956	0.912	0.867	0.822	0.776	0.729	0.681	0.634	0.585
2	1	0.936	0.871	0.806	0.740	0.672	0.604	0.535	0.465	0.394
3	1	1.107	1.216	1.327	1.438	1.552	1.667	1.783	1.901	2.021
4	1	0.956	0.912	1.867	0.822	0.776	0.729	0.681	0.634	0.585
5	1	0.936	0.871	0.806	0.740	0.672	0.604	0.535	0.465	0.394
6	1	1.107	1.216	1.327	1.438	1.552	1.667	1.783	1.901	2.021
7	1	0.956	0.912	0.867	0.822	0.776	0.729	0.681	0.634	0.585
8	1	0.936	0.871	0.806	0.740	0.672	0.604	0.535	0.465	0.394
9	1	1.107	1.216	1.327	1.438	1.552	1.667	1.783	1.901	2.021
10	1	0.956	0.912	0.867	0.822	0.776	0.729	0.681	0.634	0.585
11	1	0.936	0.871	0.806	0.740	0.672	0.604	0.535	0.465	0.394
12	1	1.107	1.216	1.327	1.438	1.552	1.667	1.783	1.901	2.021
STDEV	0	0.080	0.161	0.243	0.326	0.410	0.495	0.582	0.670	0.759

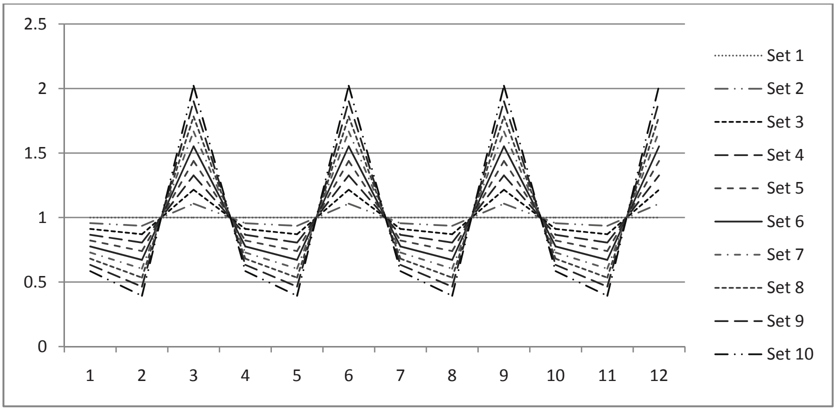
Temporal indices emulate that of a quarterly seasonal data. It should also be emphasized that theses sets are not fixed. To summarize, a time series following equation (13) is generated. It is then multiplied to each of the ten temporal series to produce ten new sets of time series of varying seasonality levels. These resemble quarterly seasonal data. It is also important to note that each set of temporal indices have an average of one so as not to affect the mean of the original stationary time series. Obviously, set 1 time series is the original non-seasonal stationary data. Sets 5 to 10 are considered to be highly seasonal.The original series generated are of varying length (short, average, and long) and seasonality strength (no seasonality to strong seasonality). Since there are three series lengths and 10 different seasonality strengths, there are a total of 30 scenarios. For every scenario, the SampEn was computed. To produce reliable estimates, each scenario has 100 replicates. The table below summarizes the scenarios:

Table 2. Summary of scenarios

Time Series Length	Short	36 Obs
	Medium	72 Obs
	Long	180 Obs
Seasonality	10 Sets of Varying Seasonality	
No. of Replicates	100	

After the data were simulated and the seasonality injected, the SampEn for every simulation per scenario was computed. Figure 1 shows a multiple line chart for the mean SampEn of the 100 simulations per length of time series.

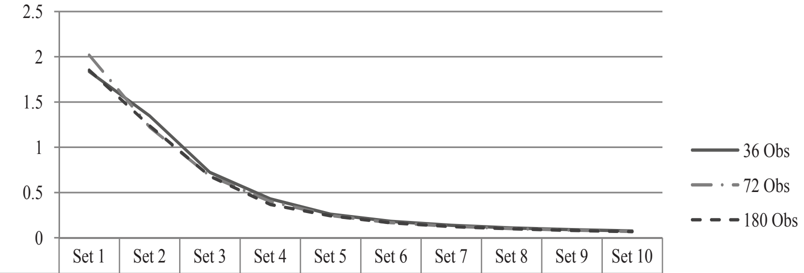
Figure 1. Line graphs of the sets of temporal indices



As seasonality becomes more apparent, the mean sample entropy estimate decreases, signifying an inverse relation. For highly seasonal time series starting at Set 5, the average of the estimates indeed go to zero. The stationary Set 1 mean sample entropy estimates are close to two, this is however not an upper bound for the statistic. It is also important to note that the three lines almost coincide, indicating that the sample entropy is robust to the number of time series points.

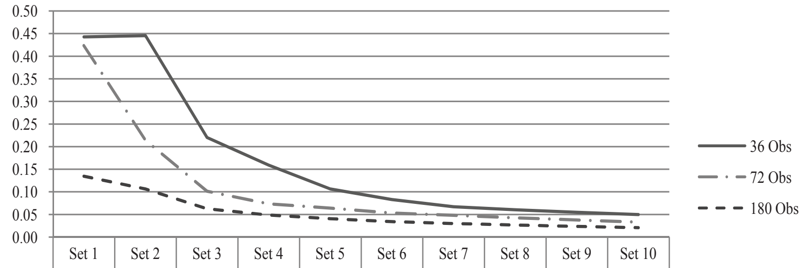
Figure 3 shows the standard errors of the mean SampEn mean for every time series length. It exhibits the behavior of the average SampEn estimates across the time series lengths. Notice that the estimates have larger standard error for fewer time points. Among the three series lengths, the standard error of the longest series, which is 180 time points, is the smallest and it approaches zero faster. This shows that SampEn will be more reliable when used in a longer time series. Although for highly seasonal data, the standard error of the resulting sample entropy estimate are reliable enough even for short time series.

Figure 2. Average sample entropy estimates



	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Set 10
36 Obs	1.8375	1.3476	0.7231	0.4302	0.2588	0.1805	0.1373	0.1074	0.0885	0.0744
72 Obs	2.0190	1.2241	0.6841	0.4006	0.2455	0.1670	0.1266	0.1010	0.0833	0.0689
180 Obs	1.8531	1.2385	0.6804	0.3669	0.2427	0.1643	0.1233	0.0980	0.0807	0.0674

Figure 3. Standard errors of the average sample entropy estimates



	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Set 10
36 Obs	0.44269	0.4455	0.22014	0.15904	0.10631	0.08285	0.06706	0.06029	0.05482	0.04975
72 Obs	0.42354	0.21403	0.10149	0.07347	0.06395	0.05324	0.04819	0.04266	0.0377	0.03305
180 Obs	0.13467	0.10642	0.06254	0.04879	0.04057	0.03425	0.02995	0.02649	0.02347	0.02075

5. Application to Actual Data

To verify if the results from the simulated data also apply to actual data, the value of sample entropy was computed for Philippine Gross Domestic Product (GDP) and Consumer Price Index (CPI). Implications of the sample entropy estimate is then compared against the existing seasonality tests in the X11 seasonal adjustment procedure.

Figures 4 and 5 shows the original and first difference GDP, respectively, covering the period from first quarter of 1981 to the fourth quarter of 2010. This is quite a long time series with 120 observations. Hence based from the simulation study, we expect a more reliable estimate of the sample entropy.

Figure 4. Plot of original GDP data: Q1 1981 to Q4 2010 (Source: NSCB)

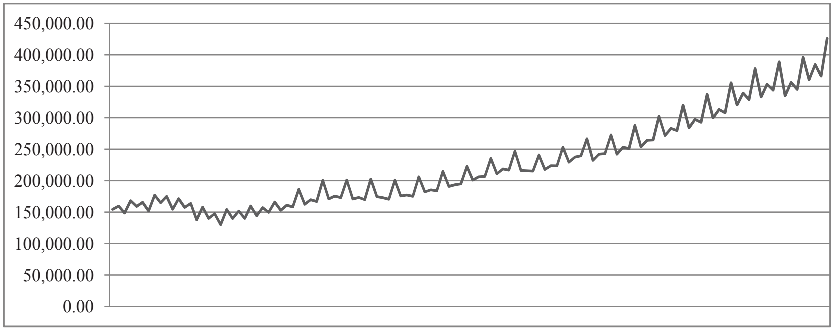
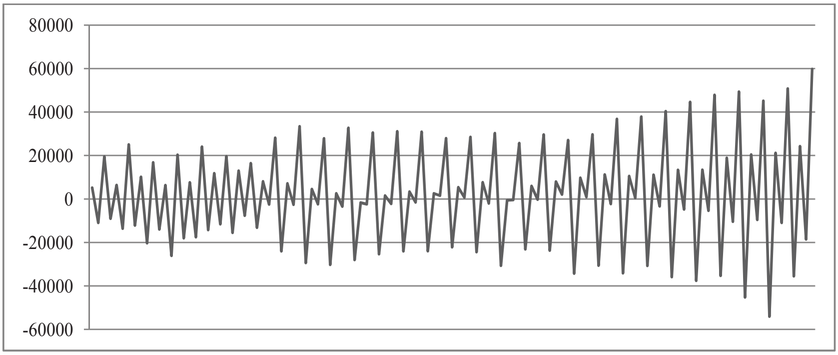


Figure 5. Plot of GDP data, first difference: Q1 1981 to Q4 2010



Based on both graphs, GDP have a high seasonality. It peaks every fourth quarter and it subsides every second quarter. To verify this, the sample entropy for the first difference of GDP was computed. The value of the computed SampEn is 0.400, which indicates high seasonality and confirms the information seen from the graph. This sample entropy estimate is consistent with the test for seasonality in the X11 procedure in Table 3.

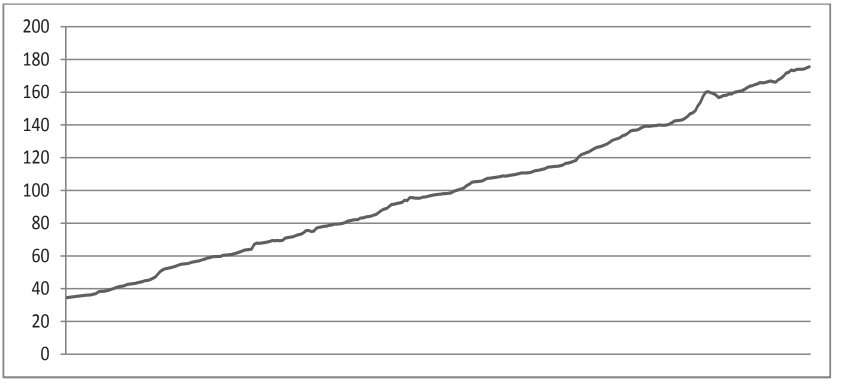
Table 3 shows the test (based on X11 decomposition) for the presence of stable, moving and identifiable seasonality in the GDP data, see Lothian and Morry (1978) for details of the test. The moving and stable seasonality tests are F-test types while for identifiable seasonality, a nonparametric Kruskal-Wallis test is used. Indeed, seasonality is present in GDP. The seasonality is constant and does not evolve in time, and it is identifiable.

The second data, CPI, covers the period January 1988 to November 2011 and it was also plotted in Figure 6 to inspect its seasonality visually.

Table 3. X11 Results of testing for seasonality for the variable GDP

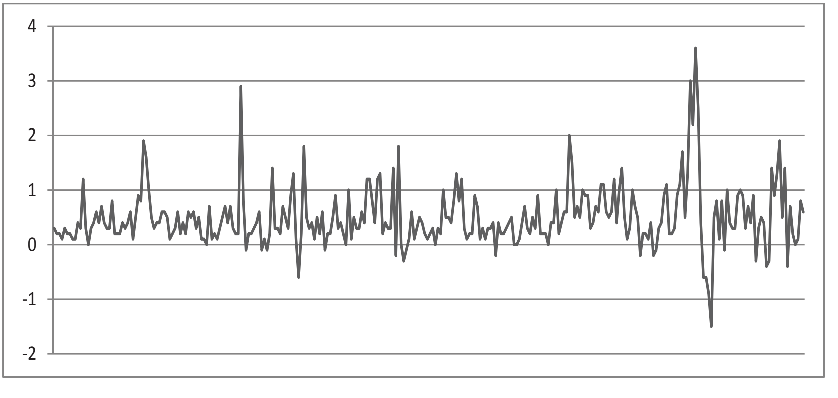
Test	F-Statistic	p-value	Decision
Stable Seasonality	354.31	<0.0001	Present
Moving Seasonality	0.75	0.8056	Not Present
Combined Identifiable Seasonality			Present

Figure 6. Plot of original CPI data: Jan1988 to Nov2011 (Source: NSO)



The graph of the first difference of the CPI in Figure 7 does not exhibit seasonality. To measure the strength of seasonality, sample entropy was computed. The computed entropy was 1.223, which is an indication of weak seasonality. From the computed sample entropy values, we can say that GDP is more seasonal than CPI since the former has a smaller SampEn value than the latter.

Figure 7. Plot of CPI data, first difference: January 1988 to November 2011



Test based on X11 decomposition show that stable seasonality is present in CPI, although a 4.51 F-statistic indicates that it is not as strong as the seasonality in GDP (F-stat = 354.31). Note that the seasonality of CPI evolves through time (moving seasonality is present), but still makes it identifiable.

Table 4. X11 results of testing for seasonality for the variable CPI

Test	F-Statistic	p-value	Decision
Stable Seasonality	4.51	<0.0001	Present
Moving Seasonality	2.96	<0.0001	Present
Combined Identifiable Seasonality			Present

6. Concluding Notes

The simulation study as well as the economic data illustrates that sample entropy decreases as seasonality of data becomes stronger and vice versa. The decrease in the standard error of the sample entropy means that as the number of time points increases, the entropy becomes a more reliable measure of seasonality. Furthermore, the method agrees with the usual tests for seasonality in seasonal decomposition methods.

Whereas procedures like the X11 gives a formal test on the presence of seasonality, sample entropy can be a descriptive tool in quantifying the extent of seasonality present in a time series data without the rigid assumptions in more formal tests.

References

AKAIKE, H., 1974, A new look at the statistical model identification, *EEE Transactions on Automatic Control*, Vol. 19: 716–723.

BOTTING, S., J. TRZECIAKOWSKI, M. BENOIT, S. SALAMA and C. DIAZ-ARRASTIA, 2009, Sample Entropy Analysis of Cervical Neoplasia Gene-expression Signatures, *BioMed Central Bioinformatics*. Vol. 10: 66.

BOX, G.E.P., and G. JENKINS, 1976, *Time Series Analysis: Forecasting and Control*, Holden-Day.

CLEVELAND, W.S., 1993, Visualizing Data, Hobart Press, Summit, New Jersey, USA.

DUDEWICZ, E., and E. VAN DER MEULEN, 1981, Entropy-based Tests of Uniformity, *Journal of the American Statistical Association,* Vol. 76: 967-974.

ECKMANN, J., and D. RUELLE, 1985, Ergodic theory of chaos and strange attractors, *Reviews of Modern Physics*, Vol. 57: 617-656.

EDWARDS, J., 1961, The Recognition and Estimation of Cyclic Trends, *Annals of Human Genetics*, Vol. 25: 83-87.

FRANSES, P., 1992, Testing for Seasonality, *Economics Letters*, Vol. 38: 259-262.

FREEDMAN, L., 1979, The Use of a Kolmogorov-type Statistic in Testing Hypothesis About Seasonal Variation, *Journal of Epidimiology and Community Health*, Vol. 33: 223-228.

GOKHALE, D., 1983, On entropy-based goodness of fit test, *Journal of Computational Statistics and Data Analysis*, Vol. 1: 157–165.

GRASSBERGER, P., and I. PROCACCIA, 1983, Estimation of the Kolmogorov Entropy from a Chaotic Signal, *Physical Review A*, Vol. 28: 2592-2593.

HEWITT, D., J. MILNER, A. CSIMA, and A. PAKULA, 1971, On Edwards' Criterion of Seasonality and a Nonparametric Alternative, *British Journal of Preventive and Social Medicine*, Vol. 25: 174-176.

JENKINS, G., and D. WATTS, 1968, Spectral Analysis and Its Applications, Holden-Day.

LAKE, D., J. RICHMAN, M. GRIFFIN and J. MOORMAN, 2002, Sample Entropy Analysis of Neonatal Heart Rate Variability, *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, Vol. 283: 789-797.

LEE, S., I. VONTA and A. KARAGRIGORIOU, 2011, A Maximum Entropy Type Test of Fit, *Journal of Computational Statistics and Data Analysis*, Vol. 55: 2635-2643.

LOTHIAN, J., and M. MORRY, 1978, A Test for the Presence of Identifiable Seasonality When Using the X-11-ARIMA Program, StatCan Staff Paper STC2118, Seasonal Adjustment and Time Series Analysis Staff, Statistics Canada, Ottawa.

MCELROY, T., and S. HOLAN, 2009, Using Spectral Peaks to Detect Seasonality, Proceedings of the Federal Committee on Statistical Methodology.

MOINEDDIN, R., R. UPSHUR, E. CRIGHTON, and M. MAMDANI, 2003, Autoregression as a means of assessing the strength of seasonality in a time series, *Population Health Metrics*, 1:10.

PINCUS, S.M., 1991, Approximate entropy as a measure of system complexity, Proceedings of the National Academy of Sciences, USA. 1991;88:2297-2301.

RICHMAN, J.S., and J.R. MOORMAN, 2000, Physiological time-series analysis using approximate entropy and sample entropy, *American Journal of Physiology - Heart & Circulatory Physiology*, *278(6):H2039-H2049.*

RADOMSKI, D., 2009, Estimation of A Sample Entropy Value Useful For Prediction Of An Upcoming Labour Based On Electrohysterographical Signal, Proceedings of the 30th Annual Conference of the International Society for Clinical Biostatistics, August 23-27, 2009. Prague, Czech Republic.

SHANNON, C., 1948, A Mathematical Theory of Communication, *Bell System Technical Journal*, Vol. 27: 379-423, 623-656.

TAKENS, F., 1983, Invariants related to dimension and entropy, Proceedings of the 13th Colloquium, Brasileiro de Matematica, Rio de Janeiro, Brazil, 1983.

WALTER, S.D., and J.M. ELWOOD, 1975, A Test for Seasonality of Events with a Variable Population at Risk, *British Journal of Preventive and Social Medicine*, Vol. 29: 18-21.

VASICEK, O., 1976, A Test for Normality based on Sample Entropy, *Journal of the Royal Statistical Society*, Vol. 38: 54-59.