

Article

Demographic Seasonality

Johan Fellman

Department of Finance and Statistics, Hanken School of Economics, Helsinki, Finland

Abstract

The seasonality of demographic data has been of great interest. It depends mainly on the climatic conditions, and the findings may vary from study to study. Commonly, the studies are based on monthly data. The *population at risk* plays a central role. For births or deaths over short periods, the population at risk is proportional to the lengths of the months. Hence, one must analyze the number of births (and deaths) per day. If one studies the seasonality of multiple maternities, the population at risk is the total monthly number of confinements and the number of multiple maternities in a given month must be compared with the monthly number of all maternities. Consequently, when one considers the monthly rates of multiple maternities, the monthly number of births is eliminated and one obtains an unaffected seasonality measure of the rates. In general, comparisons between the seasonality of different data sets presuppose standardization of the data to indices with common means, mainly 100. If one assumes seasonality as ‘non-flatness’ throughout a year, a chi-squared test would be an option, but this test calculates only the heterogeneity and the same test statistic can be obtained for data sets with extreme values occurring in consecutive months or in separate months. Hence, chi-squared tests for seasonality are weak because of this arbitrariness and cannot be considered a model test. When seasonal models are applied, one must pay special attention to how well the applied model fits the data. If the goodness of fit is poor, nonsignificant models obtained can erroneously lead to statements that the seasonality is slight, although the observed seasonal fluctuations are marked. In this study, we investigate how the application of seasonal models can be applied to different demographic variables.

Keywords: population at risk; births; deaths; twins; triplets; trigonometric regression; indices; chi-squared test; *t* test; *F* test; circadian rhythm; adjusted coefficient of determination; degrees of freedom

(Received 19 March 2019; accepted 9 April 2019)

The seasonality of population data has been of great interest in demographic studies. Seasonality depends mainly on the climatic conditions, and hence the findings may vary from study to study. Commonly, the studies are based on monthly data. The *population at risk* plays a central role. In a study of seasonal variation in the number of births or deaths, the population at risk is the product of the size of the population and the length of the month. For short periods, the population can be assumed to be constant, and therefore, the population at risk is proportional to the lengths of the months. Hence, for studies of monthly birth and death data, one must analyze the number of cases per day.

If one studies the seasonal variation in multiple maternities or in the occurrence of an innate disease, the population at risk is the total number of confinements. Hence, the number of multiple maternities in a given month must be compared with the monthly number of all maternities. Similar attempts should also be applied when one studies stillbirths and the births of males (sex ratio). Hence, one must consider the monthly rates of multiple maternities of stillbirths and of males. In the last case, one often considers secondary sex ratio as the rate of males (in %) among all births. If one considers the monthly rates, the monthly number of birth data is eliminated and one obtains unaffected seasonality measures of the rates.

The test for seasonality is not simple because the definition of seasonality is somewhat arbitrary. If one assumes seasonality as ‘non-flatness’ throughout a year, a chi-squared test (11 degrees of freedom for monthly data) would be an option. But this test is weak because it does not consider the order of the months, it calculates the same test statistic for the data set with high (or low) values occurring in consecutive months and for that in separate months, and consequently, chi-squared tests are not appropriate for model testing (Fellman & Eriksson, 2009).

Model fit is important, but which model is appropriate cannot be fixed *a priori*. Ultimately, the test for seasonality might be integrated into the test for clustering, but importantly, the clustering must be cyclic, with a one-year period. In demographic data, the sine curve has been a common feature of studies on seasonal variation (intraannual fluctuations). However, such attempts are often unsatisfactory.

Multiple trigonometric regression models have proved useful for studying all kinds of periodic data for they are flexible when the data differ from a simple sine curve. Diggle (1996) discusses these models under the name of *periodograms* and states that they are good summary descriptions of time series showing a periodic pattern. In its simplest form, the trigonometric regression model is a sine curve and is comparable to the original and generalized St Leger models (Fellman & Eriksson, 1999, 2000, 2002). When seasonal models are applied, one must pay special attention to how well the applied model fits the data. A poor fit can erroneously result in a statement that the seasonality is slight, although the

Author for correspondence: Johan Fellman, Email: fellman@hanken.fi

Cite this article: Fellman J. (2019) Demographic Seasonality. *Twin Research and Human Genetics* 22: 187–194 <https://doi.org/10.1017/thg.2019.18>

© The Author(s) 2019.

observed seasonal fluctuations are marked (Fellman, 2017; Fellman & Eriksson, 2009).

In general, comparisons between the seasonality of different data sets presuppose standardization of the data to indices with common means, mainly 100. A good knowledge of the seasonal variation during normal years is of fundamental importance for studies of the effects of wars, famines, epidemics or similar privations (Fellman & Eriksson, 2002; Land & Cantor, 1983). In this study, we investigate how seasonal models can be applied to different demographic variables.

Methods and Materials

Methods

In earlier papers, Fellman and Eriksson (2000, 2002) gave extensive presentations of seasonal models. They presented alternative models, but their studies were mainly based on multiple trigonometric regression models. It will be seen in the applications that in many situations, although there are marked seasonal variations, the simple sine curve model does not fit the data. For example, this is the case when the annual data show more than one marked peak, but data showing one peak and one trough may also differ markedly from the sine curve. Such cases could be when the peak or the trough is too long or when the distance between them differs from six months. Therefore, multiple trigonometric regression models are noteworthy alternatives.

Following Fellman and Eriksson (2002), we consider the multiple trigonometric regression model.

$$\begin{aligned} R(t_i) &= K + \sum_{m=1}^M R_m \sin(mt_i + \alpha_m) + \varepsilon_i \\ &= K + \sum_{m=1}^M (A_m \cos(mt_i) + B_m \sin(mt_i)) + \varepsilon_i. \end{aligned}$$

Hence, the applicable multiple trigonometric regression model is

$$R(t_i) = K + \sum_{m=1}^M (A_m \sin(mt_i) + B_m \cos(mt_i)) + \varepsilon_i, \quad (1)$$

where M is the number of pairs of trigonometric terms, $A_m = R_m \sin \alpha_m$ and $B_m = R_m \cos \alpha_m$. The error terms ε_i are assumed to be independent and homoscedastic. The parameters A_m and B_m ($m = 1, \dots, M$) and K in model (1) are estimated by ordinary least squares for the monthly data. With monthly data, one must introduce the restriction $M \leq 5$. The basic parameters α_m and the amplitudes R_m are estimated by the equations

$$\tan(\hat{\alpha}_m) = \frac{\hat{A}_m}{\hat{B}_m} \quad (2)$$

and

$$\hat{R}_m = \sqrt{\hat{A}_m^2 + \hat{B}_m^2}. \quad (3)$$

The angle should be chosen so that $\sin(\alpha_m)$ and A_m (or alternatively $\cos(\alpha_m)$ and B_m) have the same sign. These restrictions give unique angles α_m . Because the angle α_m and the amplitudes R_m have to be estimated from formulae (2) and (3), Fellman and

Eriksson (2002) stressed that if the target is parameter estimates ($\hat{\alpha}_m$ and \hat{R}_m), then the statistically nonsignificant estimates cannot be ignored, and therefore, irrespective of the significance of the parameter estimates (\hat{A}_m and \hat{B}_m), they recommended full pairs of trigonometric terms. Another argument for this is that \hat{R}_m may differ significantly from zero, but the angle α_m may be such that \hat{A}_m or \hat{B}_m is close to zero and is consequently considered nonsignificant. From this, it follows that the tests of parameter estimate significance should be applied to \hat{R}_m , but not to \hat{A}_m or \hat{B}_m .

The simple regression model $M = 1$ indicates the sinusoidal model. Good agreement with the sinusoidal model demands that the data pattern has one peak and one trough within a year.

In this study, we pay more attention to the model fit and less to the estimates of the individual parameters. Hence, we accept only significant parameter estimates. We assume that the model is compact and optimal when all the nonsignificant parameter estimates are eliminated and all the remaining parameter estimates are significant.

However, Fellman and Eriksson (2002) observed that the regressor vectors are almost orthogonal, resulting in parameter estimates that are approximately uncorrelated and relatively independent of the number of trigonometric terms in the models. Consequently, the model chosen has little influence on the parameter estimates. Exact orthogonality should be obtained if the data are equidistant, that is, if the months are of equal length (e.g., 30 days), hence corresponding exactly to 30° . Exactly equidistant data can be observed when circadian (daily) rhythms are studied (24 h). On the other hand, the error variance is based on the unexplained sum of squares, and accordingly, parameter tests are strongly affected by the model chosen. Furthermore, we assume in this study that the error terms are independent and homoscedastic. If this is not the case, the estimates obtained, although unbiased and consistent, will not be efficient. When the data are based on a large number of observations, the monthly rates can be assumed to be asymptotically normal. The whole model can then be tested by the F test and the individual parameters by t tests (Fellman & Eriksson, 2000, 2002). The goodness of fit of the regression model can also be analyzed with the coefficients of determination (R^2 or \bar{R}^2). Only if there is good agreement between the data and the model, will the model-based tests detect the seasonal variation.

The sinusoidal model can easily describe the seasonality of the demographic variable and locate the peak and the trough. If the optimal model is a multiple regression model, the seasonality is difficult to interpret. Therefore, tests assuming a specific seasonal pattern should be combined with other tests. Regression models are included in Figures 2, 5, 6 and 14, and sinusoidal models in Figures 11, 12, 16 and 17.

Materials

In this study, we apply the methods on various demographic seasonal data used in earlier papers. In Table 1, we present data concerning all and multiple maternities from Switzerland in 1976–1890 initially published by Weinberg (1901). Obviously, the majority of multiple maternities are twin maternities.

During his work on his thesis (Eriksson, 1973) Aldur W. Eriksson collected a large data set concerning twinning from the Åland Islands (Finland). Later, our group (Eriksson and Fellman) continued the data collection and the obtained data sets have been a good source for our continuing twinning studies.

Table 1. Seasonality of all and multiple births in Switzerland (1876–1890) according to Weinberg (1901). Obviously, the majority of multiple maternities are twin maternities

Month	Days	Switzerland		MUR ^a
		Multiple maternities	All maternities	
J	31	92	7455	12.341
F	28.25	88	6906	12.743
M	31	101	7645	13.211
A	30	98	7298	13.428
M	31	97	7388	13.129
J	30	92	7208	12.764
J	31	89	7417	11.999
A	31	90	7367	12.217
S	30	85	7166	11.862
O	31	83	7123	11.652
N	30	82	6871	11.934
D	31	81	7082	11.437
Total	365.25	1078	86,926	12.401

^aMUR = multiple maternity rate.

Table 2. Seasonal twinning data for the Åland Islands (1750–1949; Eriksson, 1973)

Month	Days	Åland		TWR ^a
		Twin maternities	Maternities	
J	31	158	7419	21.297
F	28.25	132	6860	19.242
M	31	173	7479	23.131
A	30	147	7328	20.060
M	31	148	6859	21.577
J	30	131	5693	23.011
J	31	157	5536	28.360
A	31	165	6109	27.009
S	30	156	5861	26.617
O	31	157	6165	25.466
N	30	158	6482	24.375
D	31	126	6953	18.122
Total	365.25	1808	78,744	22.960

^aTWR = twin maternity rate.

Seasonality twinning data initially published by Eriksson (1973) are presented here in Table 2. These data, as well as the Switzerland 1876–1890 data in Table 1 from Weinberg (1901), are republished in this study because the original data may be difficult to find. Other data analyzed in this study are published in journals readily available and are therefore not republished. The seasonal data in Table 2 are used and were published in several studies by Eriksson and Fellman. These published data tables are identical and the obtained results match, but in some studies the periods presented are mistakenly different. The period 1750–1949 initially given by Eriksson (1973) is the correct one. In this study, we refer to this period.

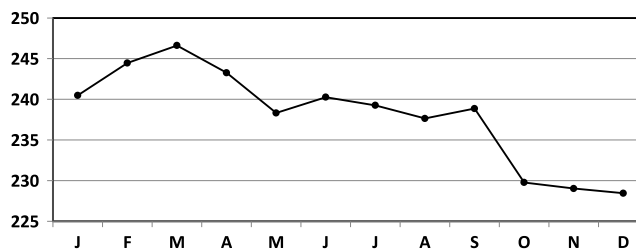


Fig. 1. Seasonality of births in Switzerland (1876–1890).

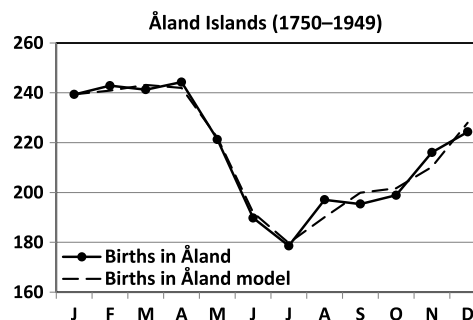


Fig. 2. Seasonality of births in Åland (Finland).

Results

In the following, we apply the trigonometric regression model on data concerning the demographic variables births (both live and still), deaths, twins, triplets and sex ratios.

Births

Fellman and Eriksson (1999) have built sinusoidal and Walter–Elwood (Walter & Elwood, 1975) models for twinning data for Switzerland (1876–1890) and for Åland (Finland) (1750–1949; Tables 1 and 2, respectively). In this study, we build trigonometric regression models for these data and compare the results with earlier results. The seasonality of births in Switzerland and in Åland (Finland) are presented in Figures 1 and 2, respectively. In Figure 1, one can locate a peak in March and a trough in December. The long distance between the peak and the trough speaks against a sinusoidal pattern, and a trigonometric regression model should be applied. The trigonometric regression model gives a better fit, but the interpretation of the model is difficult to perform. In Figure 2, one observes the peak in April and the trough in July. Now, the distance is too short for a sinusoidal model.

The optimal regression model for births in Åland (1750–1949) is

$$Y_i = 215.6064 + 21.5316 \sin(t_i) - 20.45052 \cos(t_i) - 6.976 \cos(2t_i) + 6.251 \cos(3t_i).$$

The goodness of fit is acceptable ($\bar{R}^2 = 0.962$ and $F = 70.70$), but the model is difficult to interpret.

In Figure 3, we compare the seasonality of births in Sweden for 1841–1850 and 1941–1950. One observes that for the decade in the 19th century, there are two almost identical peaks in January and September. For the decade 1941–1950, the spring peak is increased and moved to April and the September peak is markedly decreased.

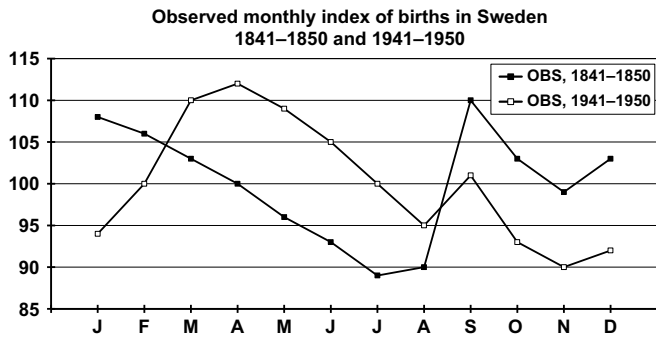


Fig. 3. Seasonality of births in Sweden for the decades 1841–1850 and 1941–1950. There are marked seasonality discrepancies between the two decades.

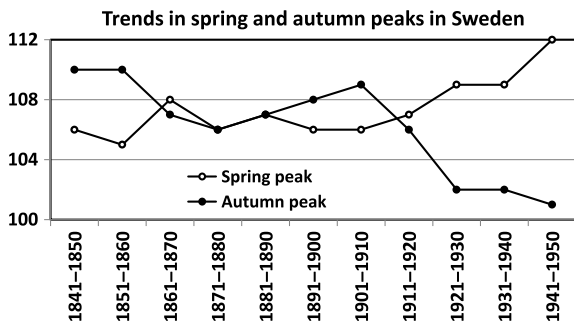


Fig. 4. Trends in levels of peaks of births in Sweden (1841–1950).

If sinusoidal models are assumed, it is the best attempt to introduce the model angle and follow its temporal trend. Alternatively, one can follow temporal trends of the troughs and peaks. For both periods in this study of the Swedish data, the two peaks indicate that a sinusoidal model is not acceptable. In Figure 4, one can follow the trends of the spring and autumn peaks. Both trends show rather constant levels of the peaks during the 19th century, but thereafter the spring peak shows a marked increase and the autumn peak shows a marked decrease.

A comparison between James' (1980) data for England and Wales (1952–1975) and Elster and Bleyl's (1991) data for US births (1985–1988) demands standardization of data to indices with means of 100. The regression model of the seasonality of birth data for England and Wales is

$$Y_i = 99.977 - 1.985 \cos(t_i) + 5.163 \sin(t_i) - 1.607 \cos(2t_i) + 1.6 \sin(2t_i),$$

and the data are given in Figure 5. The model is significant ($\bar{R}^2 = 0.885$ and $F = 22.0$). The corresponding model for US birth data is

$$Y = 100.012 - 3,205 \cos(t_i) - 3.038 \sin(t_i) - 0.919 \cos(2t_i) + 2.148 \sin(2t_i) - 0.749 \sin(4t_i) - 0.696 \sin(5t_i),$$

and the graphs are presented in Figure 6. The model is significant ($F = 67.48$ and $\bar{R}^2 = 0.973$).

The comparison between the seasonality patterns is presented in Figure 7. Note that the two data sets are obtained for comparable periods: 1952–1975 for James and 1985–1988 for Elster and Bleyl, and

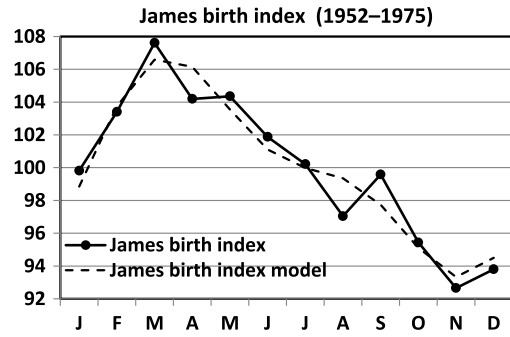


Fig. 5. Seasonality of birth data for England and Wales (1952–1975; James, 1980).

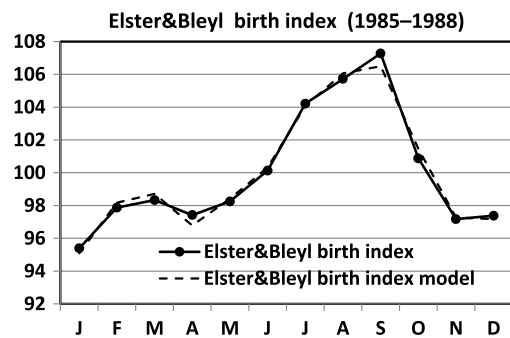


Fig. 6. Seasonality of births for the USA (1985–1988; Elster & Bleyl, 1991).

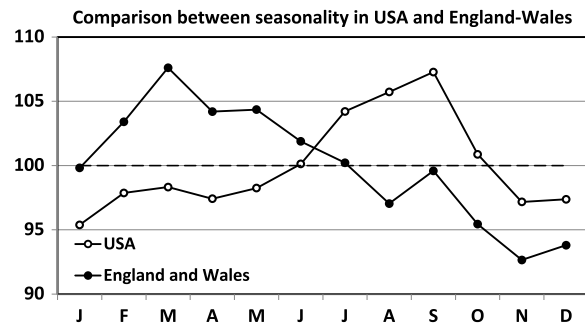


Fig. 7. Comparison between the James' data (England and Wales) and the Elster and Bleyl data (USA). Note the strong differences in seasonality.

that the data are from countries with a similar temperate climate in the northern hemisphere (England and Wales vs. USA). Despite this, marked differences in seasonality can be observed. For the England and Wales data, we observe a maximum in the spring, in contrast to the maximum for the US data being in the autumn.

Stillbirths. A good knowledge of the seasonal variation during normal years is of fundamental importance for studies of the effects of wars, famines, epidemics or similar privations (Land & Cantor, 1983). For the period 1881–1890 in Finland, the seasonality of live births is presented in Figure 8(a) and that of stillbirths in Figure 8(b). The comparison between the graphs demands standardized indices. Note the marked differences in the seasonality expressed in Figure 8(c).

Mortality. During the 1860s, there was a severe famine in Finland. Fellman and Eriksson (2002) gave a thorough and detailed presentation of the demographic effects of this deprivation. In this study,

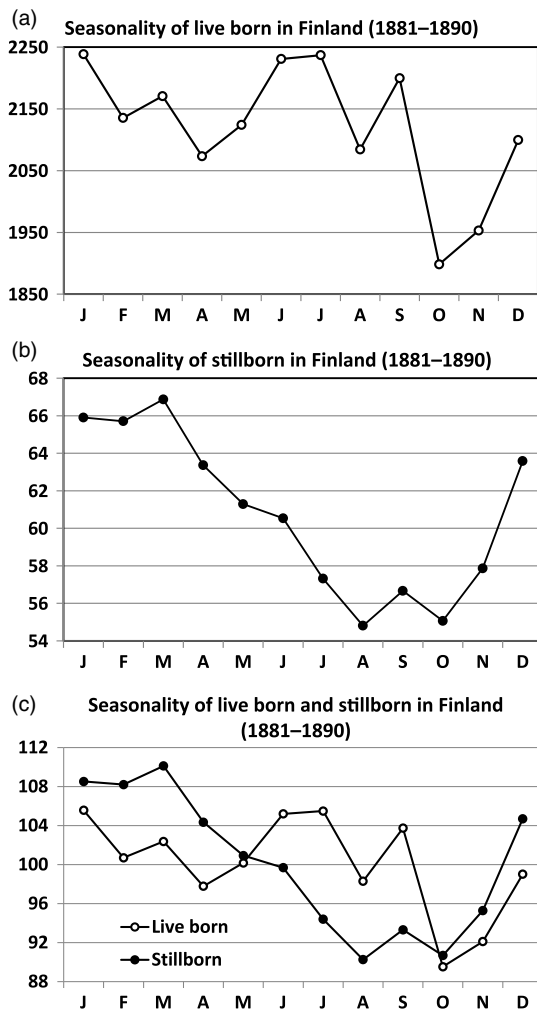


Fig. 8. Seasonality of live and stillbirths in Finland (1881–1890). The seasonality of live births is given in (a) and of stillbirths in (b). Note the marked differences in seasonality expressed in (c).

we return to these austere days in Finland. We consider the seasonal variations in the mortality. The famine started in 1866, with the most severe period occurring in 1868, and the deprivation terminated in 1870. The birth rate in Finland during 1868 was only about 70% of that in normal years and showed an aberrant seasonality, with a strong trough from October 1868 to February 1869, that is, conceptions were fewer between January and May 1868, when the food shortage was extremely severe (Fellman & Eriksson, 2002).

During famines and other extreme conditions, the seasonality of mortality often shows a markedly different pattern. According to Fellman and Eriksson (2002), as a consequence of the severe famine in Finland in 1867–1868, the mortality for the whole year in 1868 was almost four times as high as during normal years, and the seasonality of the mortality was even more accentuated. In Figure 9, we present the seasonality of the mortality in Finland during 1866, 1868 and 1870. The years 1866 and 1870 were almost normal with respect to famine, but during 1868 the famine was exceptionally hard and resulted in extreme mortality values during the spring (April and May), when the food showed the greatest deficiency before the new harvest.

Icelandic Historical Statistics (IHS, 1997) has published a long-term series concerning demographic data. Fellman and Eriksson (2002) and Fellman (2016, 2019) presented a detailed historic

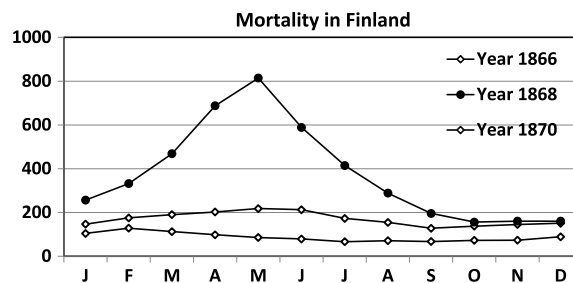


Fig. 9. Seasonality of mortality in Finland during 1866, 1868 and 1870. In 1868, the famine was exceptionally severe resulting in seasonality of mortality, with an extreme peak in the spring (April and May). For details, see the text.

analysis of the mortality in Iceland based on the Icelandic data. They especially noted that the seasonal variation in the mortality data from Iceland, 1856–1990, shows a strong secular decrease, and the connection between this temporal trend and the epidemiological transition was considered. For this temporal finding, they analyzed and presented socioeconomic explanations and obtained interesting results (Fellman & Eriksson, 2002). They considered the mortality data from Iceland for the period 1856–1990. To obtain temporal comparisons, Fellman and Eriksson transformed the rates measuring deaths per day to indices. Figure 4 in Fellman and Eriksson (2002) shows the time series of the seasonal indices and the annual rates of deaths per 1000. After 1856, the death rate shows a steadily decreasing trend from over 30 per thousand in the 1850s and 1860s to about 7 during the second half of the 20th century. The extreme peak for the period 1861–1870 is largely caused by an increased number of drownings and accidents of other types. In 1861–1865, the death rate in males caused by drowning was about 3 per thousand. This was followed by a clear continuous decreasing trend. At the beginning of the 20th century, the corresponding death rate was between 1 and 2 per thousand. After World War II, the decreasing trend is even more noticeable (IHS, 1997).

Omran (1971) proposed the epidemiological transition, consisting of three stages, and later Olshansky and Ault (1986) introduced a fourth stage. Rogers and Hackenberg (1987) named the fourth stage the hybriatic stage and suggested that the respective turning points of the stages were 1875, 1930 and 1970. The accepted stages are: (a) age of pestilence and famine (up to 1875); (b) age of receding pandemics (1876–1930); (c) age of degenerative and man-made diseases (1931–1970); and (d) hybriatic stage (1971–).

Following these stages, Fellman and Eriksson (2002) performed a pooling of the data. They pooled the data into four subperiods: 1856–1875, 1876–1930, 1931–1970 and 1971–1990, and compared the seasonal patterns for these subperiods. A diminishing trend in the seasonality is clearly discernible. The Icelandic data for the period 1856–1875 show strong seasonal variation, with one marked mortality peak in June. During the period 1876–1930, the summer peak declined and the mortality in the winter and spring increased. A small peak is discernible in late autumn. For the period 1931–1970, the mortality values are high during the spring and early summer (peaks in February, April and June). The autumn peak, although diminished, is still discernible. This indicates further reduced seasonal variation. For the last period, 1971–1990, the seasonal variation has almost disappeared. As Fellman and Eriksson (2002) already stressed, a specific aspect of the mortality pattern in Iceland in the 19th century is the high level of mortality caused by drowning. Being a cause of death that almost entirely affected males (over 95% of the victims were males), it is an important factor

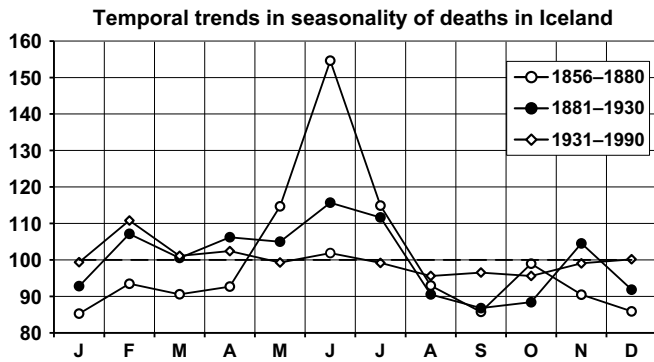


Fig. 10. Seasonality of three periods.

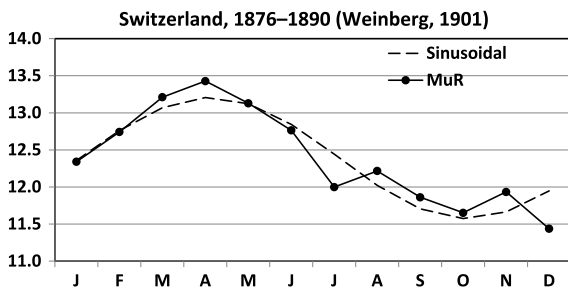


Fig. 11. Seasonality of MUR in Switzerland in 1876-1890. The obtained sinusoidal model is included.

when the mortality in Iceland is studied. Figure 7 in Fellman and Eriksson (2002) presents the mortality per 1000 for drowning and other accidents.

Although the number of drownings has decreased markedly since the 19th century, drowning was still a notable cause of death for males in the first half of the 20th century (Figure 6; Fellman & Eriksson, 2002).

In Figure 10, we give a short presentation of the central result of Fellman and Eriksson (2002). They grouped the period 1856-1990 into three subperiods: the 19th century period, 1856-1880; the turn of the century, 1881-1930; and the recent period, 1931-1990. The temporal changes in the seasonality discussed in detail by Fellman and Eriksson (2002) can easily be seen in Figure 10. One identifies the strong seasonality for the oldest period with the marked peak in the summer, slight seasonality during the turn of the century and the almost eliminated seasonality of today. The slight seasonality of today can be explained by the Icelandic climate, which is virtually identical for all four seasons.

Twinning. For Switzerland, the rates are according to Weinberg multiple maternity rates (MUR), but obviously twin confinements dominate these rates. Fellman (2017) and Fellman and Eriksson (1999) built sinusoidal models for the twinning rates (TWRs) for Switzerland (1876-1890) given in Table 1 and for Åland (Finland; 1750-1949) given in Table 2. In this study, we build trigonometric regression models for these data and compare the results with earlier findings. For MUR in Switzerland (1876-1890), the obtained seasonal model is a sinusoidal model $Y_i = 12.391 + 0.779 \sin(t_i) - 0.247 \cos(t_i)$ (Fellman, 2017; Fellman & Eriksson, 1999).

The model is significant ($F = 26.72, p < .001, \bar{R}^2 = 0.824$), and in Figure 11, one can locate a peak in April and a trough in October. The distance between the peak and the trough seems to be about six months, indicating that a sinusoidal pattern is acceptable.

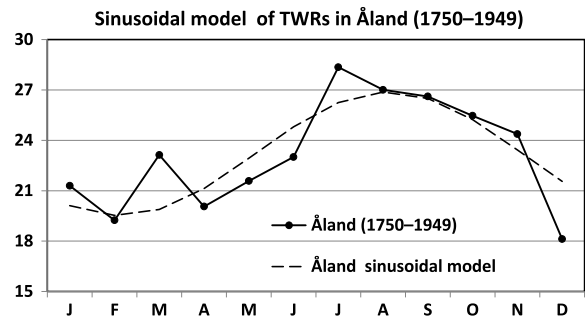


Fig. 12. Seasonality of the TWR in Åland, 1750-1949. An acceptable sinusoidal model is included.

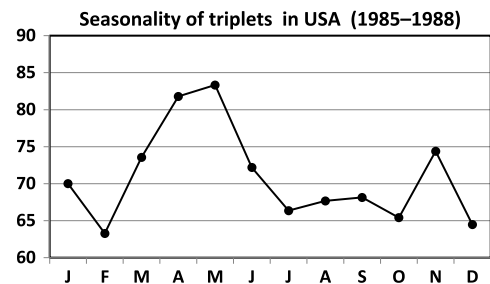


Fig. 13. Seasonality of triplets in the USA according to Elster and Bleyl (1991).

Also for Åland, the obtained seasonal model is a sinusoidal model $Y_i = 23.208 - 2.458 \cos(t_i) - 2.738 \sin(t_i)$ (Fellman, 2017; Fellman & Eriksson, 1999).

The model is significant ($F = 10.35, P < 0.01, \bar{R}^2 = 0.630$). In Figure 12, one observes the trough in February and a peak in July. Now the distance is acceptable for a sinusoidal pattern.

Triplets. Seasonality of triplets in the USA is according to Elster and Bleyl (1991) given in Figure 13. The only possible model is a multiple trigonometric regression, but it is difficult to interpret.

The seasonality of triplets in England and Wales is presented in James (1980). The only possible model for the triplet rate (TRR) index is a multiple trigonometric regression

$$Y_i = 109.621 + 9.875 \cos(t_i) + 4.489 \sin(t_i) - 4.683 \sin(2t_i) - 4.567 \sin(3t_i) - 4.242 \sin(4t_i) + 5.693 \cos(5t_i),$$

but it is difficult to interpret. The model is significant, that is, $F = 13.41$ with $p < .01$, and $\bar{R}^2 = 0.871$.

Sex ratios. In this study, we also analyse the seasonal variation in the secondary sex ratio in Åland. In order to make it possible to investigate temporal effects of the seasonality, we divide the temporal data set into the subperiods 1653-1750, 1751-1850 and 1851-1950. We measured the secondary sex ratio as the proportion (%) of males. The seasonal variation shows large differences and decreasing seasonalities between the three centuries. The standard deviation is 1.382 for 1653-1750, 1.088 for 1751-1850 and 0.564 for 1851-1950. For 1653-1750, the sex ratio shows a slight increasing trend during the year. A slight similar trend can be seen during 1851-1950. For 1751-1850, such an increase is not observed, but for July, an extreme value can be noted. For 1653-1750, the overall sex ratio is 51.29, for 1751-1850 it is 51.05 and for 1851-1950 it is 51.40. No temporal trends are apparent. Figure 15 shows that no statistically significant monthly heterogeneity emerges.

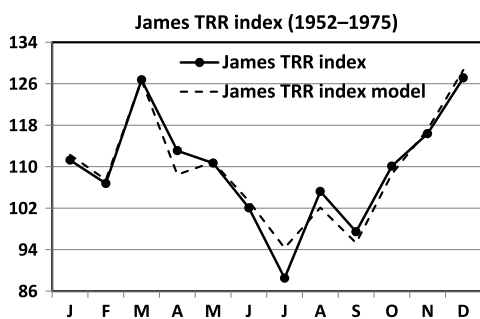


Fig. 14. Seasonality of TRR in England and Wales according to James (1980). The trigonometric regression model is included.

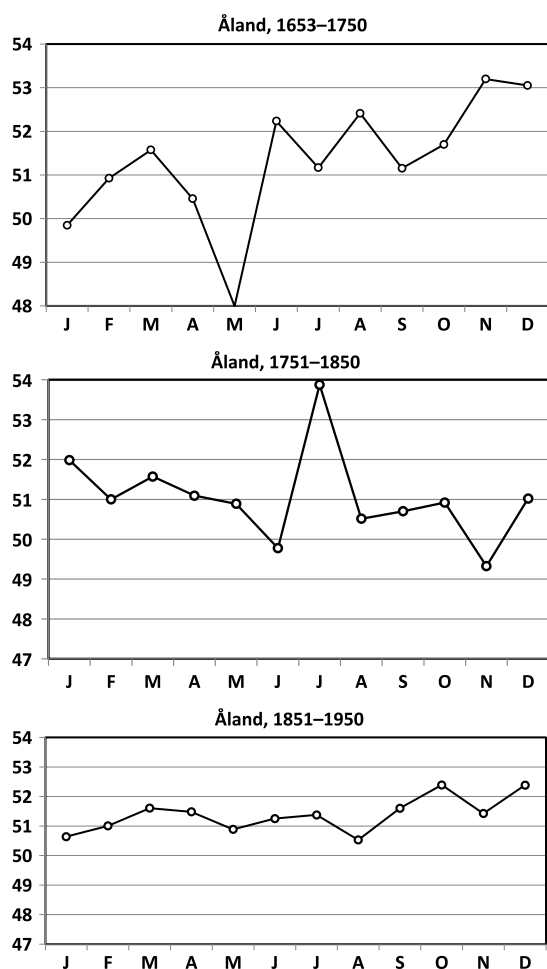


Fig. 15. Seasonal variation in the sex ratio observed in Åland during the centuries 1653–1750, 1751–1850 and 1851–1950. For detailed discussion, see the text.

Using chi-squared tests with 11 degrees of freedom, we obtained nonsignificant monthly test values for all periods: 8.67 for 1653–1750, 18.61 for 1751–1850, 6.90 for 1851–1950 and 11.06 for the whole data set (the 5% critical value is 19.67).

Discussion

Fellman and Eriksson (2000) considered monthly data and paid special attention to seasonality problems. They noted that several authors have proposed that the monthly data should be pooled to

quarterly data (e.g., Nonaka & Miura, 1987a, 1987b). The opinion of Fellman and Eriksson (2000) was that this should be done with great caution and tested empirically. Pooling must be based entirely on assumptions and the proposed model, but under no circumstances on the data. Furthermore, it reduces the number of observations, so that it is difficult to model the seasonal pattern statistically. In addition, pooling may mask individual months with exceptional data, for example, birth data in September (Fellman & Eriksson, 1999). They also stressed their opinion that monthly data are well suited for the study of seasonal variation. The analyses can be based on a sufficient number of time points, which may not be the case with quarterly data. If the year is split into smaller parts than months (e.g., into weeks), the seasonal pattern, especially when rare events are considered, is easily masked by random fluctuations, and the data obtained are not large enough for consistent estimates. Also, official seasonal statistics of demographic items is often based on monthly data (Fellman & Eriksson, 2000).

Bobak and Gjonca (2001) stressed that the number of births varies markedly by season, but the causes of this variation are not fully understood. Proposed explanations include temperature or photoperiod, seasonal variation in pregnancy loss or cultural factors. In their paper, they examined whether birth seasonality is influenced by sociodemographic factors. Bobak and Gjonca used data on all live births registered in the Czech Republic in 1989–1991. Differences in the degree of seasonality between sociodemographic groups were examined by inspection of curves, by comparing coefficients of variations of monthly numbers of births and by calculating the ratios of the number of births in the three peak months (March–May) to the number of births in the three lowest months (October–December). They found large differences in the size of the seasonal variation in births by sociodemographic factors. Especially, the seasonal variation was highly pronounced in mothers who were 25–34 years old, had higher education, were married and were pregnant with their second or third child. By contrast, birth seasonality was weak in mothers who were aged below 19 years or above 35 years, were unmarried, had low education and were expecting their first or fourth or higher order birth. Furthermore, they found that in a multivariate model, all four sociodemographic variables contributed significantly to seasonal variation. Their results suggested that the seasonality of births is, in addition to climatic factors, strongly influenced by sociodemographic factors.

Gebremedhin (2014) compared the multiple birth rates across 12 months using a bivariate logistic regression model. However, no significant seasonal variation was observed. For Icelandic males, Fellman and Eriksson (2002) had to include drownings as one important indicator of the harsh living conditions. In the first half of the 19th century, the fishermen had a hard life practicing their profession and braving the Atlantic storms in the small vessels of that time. Until the 20th century, the sex ratio for the Icelandic population was low relative to the sex ratio for both Norway and Sweden, indicating the greater effect of mortality from drownings and malnutrition and its sequelae on males than on females (Tomasson, 1977).

Above, we have applied trigonometric regression models to different demographic variables. The results show that the models usually provided a good fit. However, when the obtained model is a multiple one, it shows complicated fluctuations. Hence, it is often difficult to interpret. The interpretation is easier when the obtained model is simpler, particularly when it is a sinusoidal one.

In Figure 16, we present a three-year sketch based on the Switzerland twinning data in Figure 11, and in Figure 17, we present a three-year sketch based on the Åland twinning data in Figure 12.

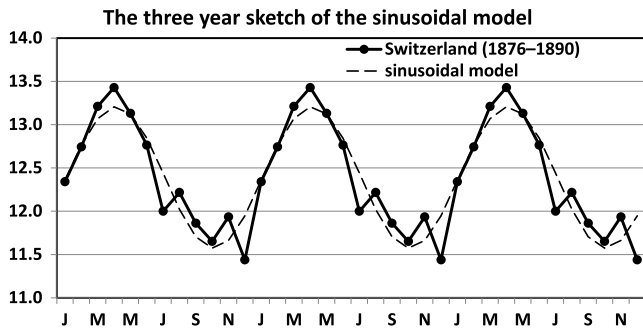


Fig. 16. Three-year sketch is based on Figure 11.

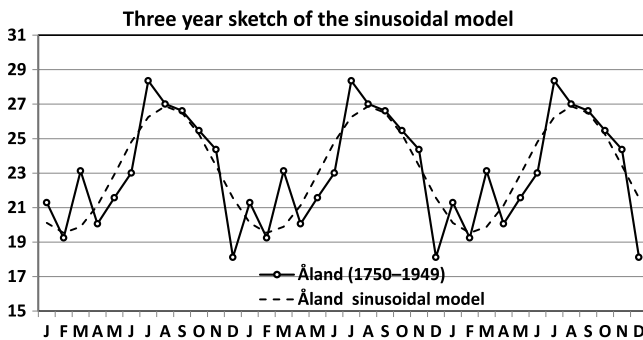


Fig. 17. Three-year sketch is based on Figure 12.

The figures show the periodic fluctuations of the sinusoidal model. A sinusoidal model shows annual variation with a high and low value. Consequently, when the sinusoidal model yields a rather good fit and potential improvement is minute, the sinusoidal model is easily explained and should be used.

Financial support. This study was in part supported by a grant from the Magnus Ehrnrooths Stiftelse Foundation.

References

- Bobak, M., & Gjonca, A. (2001). The seasonality of live birth is strongly influenced by sociodemographic factors. *Human Reproduction*, 16, 1512–1517.
- Diggle, P. J. (1996). *Time series: A biostatistical introduction*. Oxford, UK: Oxford Science Publications Clarendon Press.
- Elster, A. D., & Bleyl, J. (1991). Seasonality of triplet births in the United States. *Human Biology*, 63, 711–718.

- Eriksson, A. W. (1973). Human twinning in and around the Åland Islands. *Commentationes Biologicae*, 64, 1–159.
- Fellman, J. (2016). Historic demography of Iceland. *Journal of Advances in Medicine and Medical Research*, 2, 1–13.
- Fellman, J. (2017). Studies of the seasonal pattern of multiple maternities. *Twin Research and Human Genetics*, 20, 250–256.
- Fellman, J. (2019). *Mortality history of Iceland*. Manuscript submitted for publication.
- Fellman, J., & Eriksson, A. W. (1999). Statistical analysis of the seasonal variation in twinning rate. *Twin Research*, 2, 22–29.
- Fellman, J., & Eriksson, A. W. (2000). Statistical analysis of the seasonal variation in demographic data. *Human Biology*, 72, 851–876.
- Fellman, J., & Eriksson, A. W. (2002). Regional, temporal and seasonal variations in births and deaths: The effects of famines. *Social Biology*, 48, 86–104.
- Fellman, J., & Eriksson, A. W. (2009). Temporal and regional variations in the seasonality of births in Åland (Finland), 1653–1950. *Biodemography and Social Biology*, 55, 103–112.
- Gebremedhin, S. (2014). Multiple births in Sub-Saharan Africa: Epidemiology, postnatal survival, and growth pattern. *Twin Research and Human Genetics*, 18, 100–107.
- Icelandic Historical Statistics (IHS). (1997). *Hagskinna Sögulegar hagtölur um Ísland* ed. Gudmundur Jonson and Magnus S. Magnusson. Reykjavik. 957 pp. Retrieved from <https://www.statice.is>
- James, W. H. (1980). Seasonality in twin and triplet births. *Annals of Human Biology*, 7, 163–175.
- Land, K. C., & Cantor, D. (1983). Arima models of seasonal variation in U.S. birth and death rates. *Demography*, 20, 541–568.
- Nonaka, K., & Miura, T. (1987a). Incidence of twinning in London from 1581 to 1760. *Acta Geneticae Medicae et Gemellologiae (Roma)*, 36, 365–372.
- Nonaka, K., & Miura, T. (1987b). Principles in methods of epidemiological studies on birth seasonality. In T. Miura (Ed.), *Progress in biometeorology* (vol. 6, pp. 13–24). The Hague, the Netherlands: SPB Academic.
- Olshansky, J., & Ault, A. B. (1986). The fourth stage of the epidemiologic transition: The age of delayed degenerative diseases. *Milbank Memorial Fund Quarterly*, 64, 355–391.
- Omran, A. R. (1971). The epidemiologic transition. A theory of the epidemiology of population change. *Milbank Memorial Fund Quarterly*, 49, 509–538.
- Rogers, R. G., & Hackenberg, R. (1987). Extending epidemiologic transition theory: A new stage. *Social Biology*, 34, 234–243.
- Tomasson, R. F. (1977). A millennium of misery: The demography of the Icelanders. *Population Studies*, 31, 405–427.
- Walter, S. D., & Elwood, J. M. (1975). A test for seasonality with a variable population at risk. *British Journal of Preventive & Social Medicine*, 29, 18–21.
- Weinberg, W. (1901). Beiträge zur Physiologie und Pathologie der Mehrlingsgeburten beim Menschen [Contributions in physiology and pathology to the human multiple maternities]. *Pflügers Archiv für die gesamte Physiologie des Menschen und der Tiere*, 88, 346–430.