Chapter 9
Seasonality of Causes of Death

9.1 Decomposing Seasonal Data

The majority of deaths in most countries can be attributed to causes that feature a distinct seasonal pattern. Figure 9.1 depicts the relative monthly frequencies of nine selected causes of death in the United States for women and men combined for the years 1959–2014. The reported number of counts in parentheses in the title of each panel is the actual number of deaths. To control for varying lengths of months, the monthly columns in each histogram have been adjusted for a uniform length (30 days). The horizontal reference lines denote the expected value of a uniform distribution (=1/12).

The typical distribution follows a sinusoidal pattern with highest mortality in winter and relatively few cases in the summer. Primarily, those are circulatory diseases (e.g., heart diseases, cerebrovascular diseases)—as shown in the first row of Fig. 9.1—and respiratory diseases such as chronic obstructive pulmonary disease ("COPD"), pneumonia or influenza (Eurowinter Group 1997, 2000; Mackenbach et al. 1992; Kunst et al. 1990; Rau 2007; Yen et al. 2000; Seretakis et al. 1997), which are displayed in three horizontal panels in the middle of Fig. 9.1.

If diseases, and ultimately, mortality occur seasonally, it has been argued that "an environmental factor has to be considered in the etiology of that disease" (Marrero 1983, p. 275).[1] The main environmental factor to trigger higher mortality during winter for circulatory diseases and respiratory diseases—the rows on top of

---

[1] It should be noted that the impact of environmental factors on diseases and deaths is as not a finding of the latter part of the twentieth century but is well known for more than 2000 years. In about 400BC Hippocrates started his treatise "On Airs, Waters, and Places" with the following words: "Whoever wishes to investigate medicine properly, should proceed thus: in the first place to consider the seasons of the year, and what effects each of them produces for they are not at all alike, but differ much from themselves in regard to their changes."
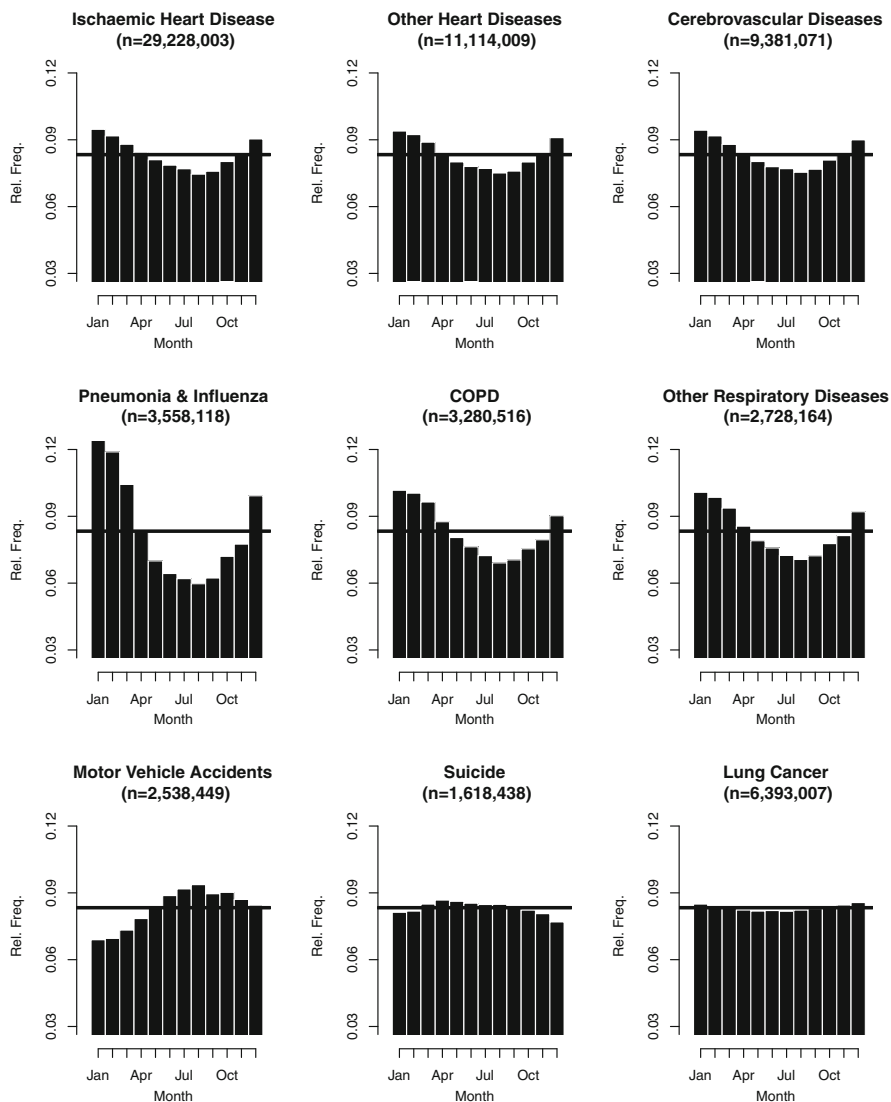
Ischaemic Heart Disease
(n=29,228,003)

Other Heart Diseases
(n=11,114,009)

Cerebrovascular Diseases
(n=9,381,071)

Pneumonia & Influenza
(n=3,558,118)

COPD
(n=3,280,516)

Other Respiratory Diseases
(n=2,728,164)

Motor Vehicle Accidents
(n=2,538,449)

Suicide
(n=1,618,438)

Lung Cancer
(n=6,393,007)

Fig. 9.1 Seasonality of selected causes of death in the United States for both sexes combined for the years 1959–2014. The counts reported in each panel denote the actual numbers of death. The relative frequencies in each histogram are adjusted for a uniform length of 30 days per month (Data source: National Center for Health Statistics and National Bureau of Economic Research)

Fig. 9.1—is well understood: temperature. Cold temperatures constrict the blood vessels and change the composition of the blood; furthermore, low temperatures facilitate the survival of bacteria in droplets and increase the risk for pulmonary infections (Eurowinter Group 1997, 2000; Huynen et al. 2001; Rau 2007).

The patterns observed in the three panels at the bottom of Fig. 9.1 deviate from the ones for circulatory and respiratory diseases above. Motor vehicle accidents do not peak in winter but around July and August. Many people assume that the reason for the peak in all-cause mortality is due to suicides in winter. The middle panel at the bottom of Fig. 9.1 illustrates why this assumption is wrong for three reasons: (1) The seasonal pattern is less pronounced for suicide than for other causes. (2) If one can speak of a seasonal pattern at all, the peak occurs definitely not during winter. (3) The 30,000 observed deaths are less than 1.5% of all deaths; not enough to shape the pattern for all causes. Lung cancer, whose impact on mortality in the United States was discussed in previous chapters, is—like many malignant neoplasms—an example of no or only negligible seasonality.

Figure 9.1 displays an aggregated picture of monthly deaths. In our analysis we want to investigate, however, whether the seasonal pattern for selected causes of death differs by age as well as whether the seasonal pattern changed over calendar time. The multiplicative model[2] suggested by Eilers et al. (2008) to decompose seasonal data allows such an analysis. The model is, at its core, another application of smoothing data via P-splines (Eilers and Marx 1996) as in Chap. 5. It is rather flexible since it allows the estimation not only of counts but also of rates. Exposures are then included as log offsets if the latter is desired, similar to Camarda's approach (2012, 2015) employed in Chaps. 5, 6, and 7. We use the model in its most simple form: The model is estimating counts assuming an annual unimodal pattern in the data. Not allowing for bimodal patterns or even higher frequencies should not induce any problems in our analysis since the causes in which we are interested in feature clear patterns with one peak and one trough (see Fig. 9.1).

We model the expected value of death counts y over age a and time t, $\mu_{ta} = E(y_{ta})$, to be Poisson distributed using a log-link function

$$\log (\mu_{ta}) = v_{ta} + f_{ta} \cos (\omega t) + g_{ta} \sin (\omega t)$$

with $\omega = 2\pi/p$, where p is the period. In our case of monthly values $p = 12$. Further technical details are given in Eilers et al. (2008).

The estimation yields three smooth matrices/surfaces, v_{ta} for the trend as well as the smooth cosine and sine surfaces f_{ta} and g_{ta}. The trend surface captures any major changes in the overall pattern that could be caused by varying population sizes, survival improvements, competing risks .... We are mainly not interested in this trend surface nor in the the actual sine and cosine surfaces. The two latter surfaces allow us, however, to obtain an estimate for the amplitude and the phase over age and time via simple trigonometric functions. The latter denotes the location of the annual peak of the death counts and is expressed in the difference in days from the 1st of January; i.e., a value of 30 corresponds to late January whereas -30 indicates that mortality is highest in the beginning of December.

---

[2]Since the logarithm of death counts is modeled, it actually becomes an additive model.

9.2 Results

Our data and results are displayed in five panels for each selected cause. On the first page for each cause, we show the observed ("raw") monthly numbers of death by calendar time and single age (adjusted for a duration of 30 days) in the upper panel. The panel below plots the fit of the model, i.e., the combined pattern of the trend and the sine and the cosine surfaces, which is equivalent to the observed counts minus the (raw) residuals; see, for example, Fig. 9.2 on page 103 for mortality from all causes combined for women. Our main interest is displayed on the second page for each cause. The top panel shows the estimated trend surface v_{at}. In the case of seasonality of all-cause mortality among US women (Fig. 9.3), we can see that the number of deaths from that category increases with age and reaches its "hotspot" for octogenarians before the numbers of death decline again. As the trend surface plots the seasonally-adjusted *density* of deaths, the lower number of deaths for nonagenarians are the consequence of less people being alive rather than a decline in the risk of dying. Even without the additional seasonal component, up to 3,500 women died at a single age during a single month. The height of "excess mortality" is depicted by the amplitude in the middle panel. Higher ages correspond not only to higher mortality; the colors and the contour lines suggest that mortality differences between winter and summer also become larger at higher ages. Increasing seasonality with age has already been described by Adolphe Quetelet in 1838 and is typically also found in more contemporary populations (Feinstein 2002; McDowall 1981; Rau and Doblhammer 2003; Rau 2007). Over time we can not really discern a clear trend. It seems rather that deaths for 70-year-old women in the US are about 10% higher during the peak season and about 15% higher for 90-year-old women than on average during a year. If we multiply the seasonal estimate of a given age and calendar time (e.g., 1.1) with the corresponding square of the trend surface (e.g., 1,500 deaths), we obtain the fitted value (e.g., 1,650 deaths) shown in the lower panel on the previous page. When the peak season occurs in a year is illustrated in the lower panel. The colors indicate a value slightly below 30. Hence, deaths occur most often in the end of January, regardless of age or calendar year.

The corresponding plots for men are depicted in Figs. 9.4 and 9.5. While male mortality is higher than female mortality at any age—at least in highly developed countries, the seasonal characteristics are rather similar between the two sexes: The proportion of excess deaths during winter varies between 5% at age 50 and 15% at age 90 with no apparent period effect. Also the part of the year when deaths peak among men occurs at the end of January. Those *seasonal* mortality similarities between women and men are not only present for all-cause mortality but also for most causes of death. That is why we restricted ourselves to show only the results for women but they apply equally to men. We show the results for men only in the case of motor vehicle accident because much less women die of that cause.

The largest subcategory analyzed by us in this chapter is death from heart diseases (see Figs. 9.6 and 9.7, pp. 107–108). Up to 1,300 deaths were recorded at a single age during a single month of a given year. As we can infer from the
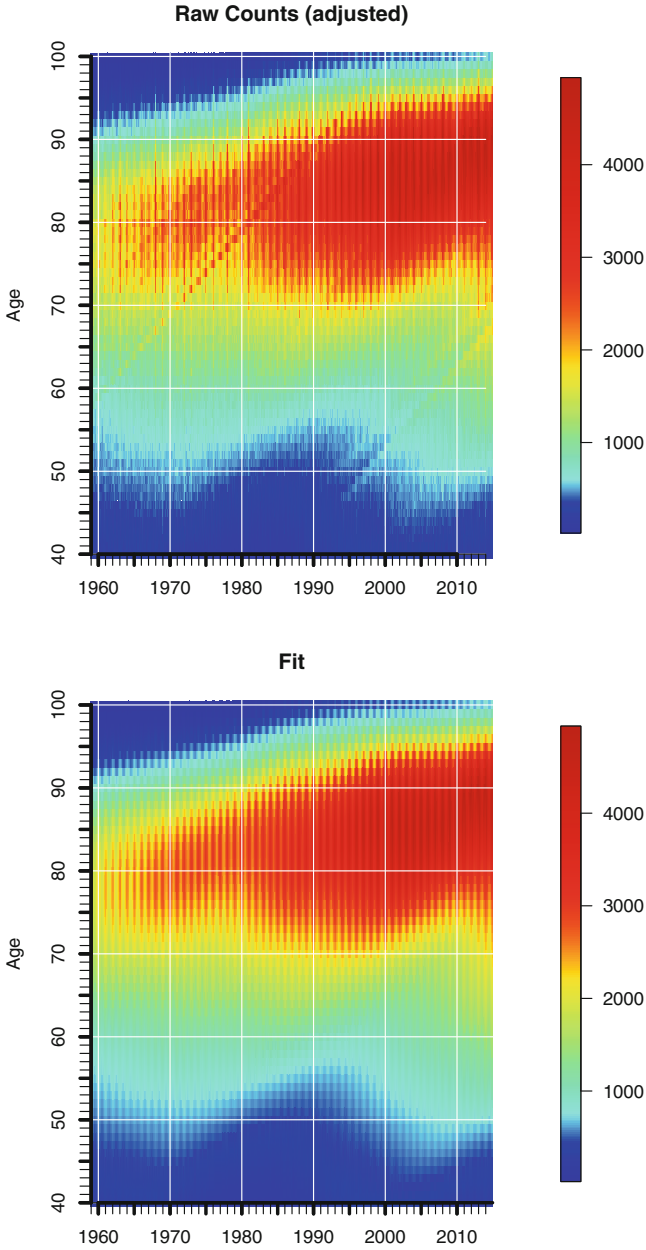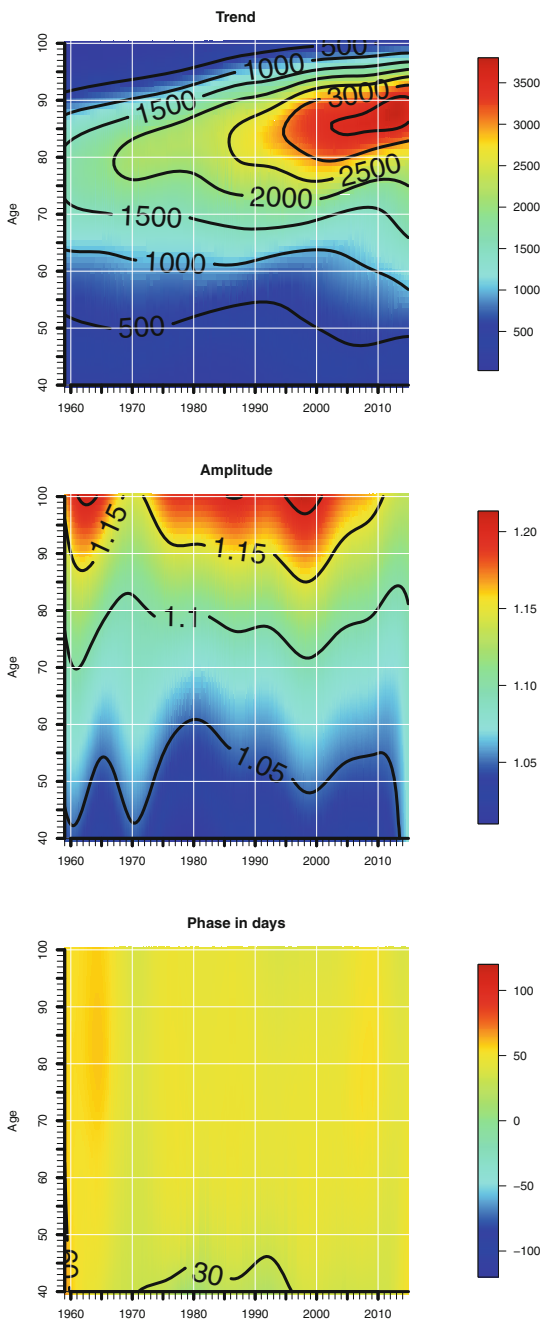
Fig. 9.2 Seasonality of mortality from all causes in the United States, 1959–2014, women, raw counts (adjusted for length of month) and fitted model (Data source: Human Mortality Database)

Fig. 9.3 Seasonality of
mortality from all causes in
the United States, 1959–2014,
women, estimated trend
surface (*top panel*), amplitude
(*middle panel*), and phase
(*bottom panel*) (Data source:
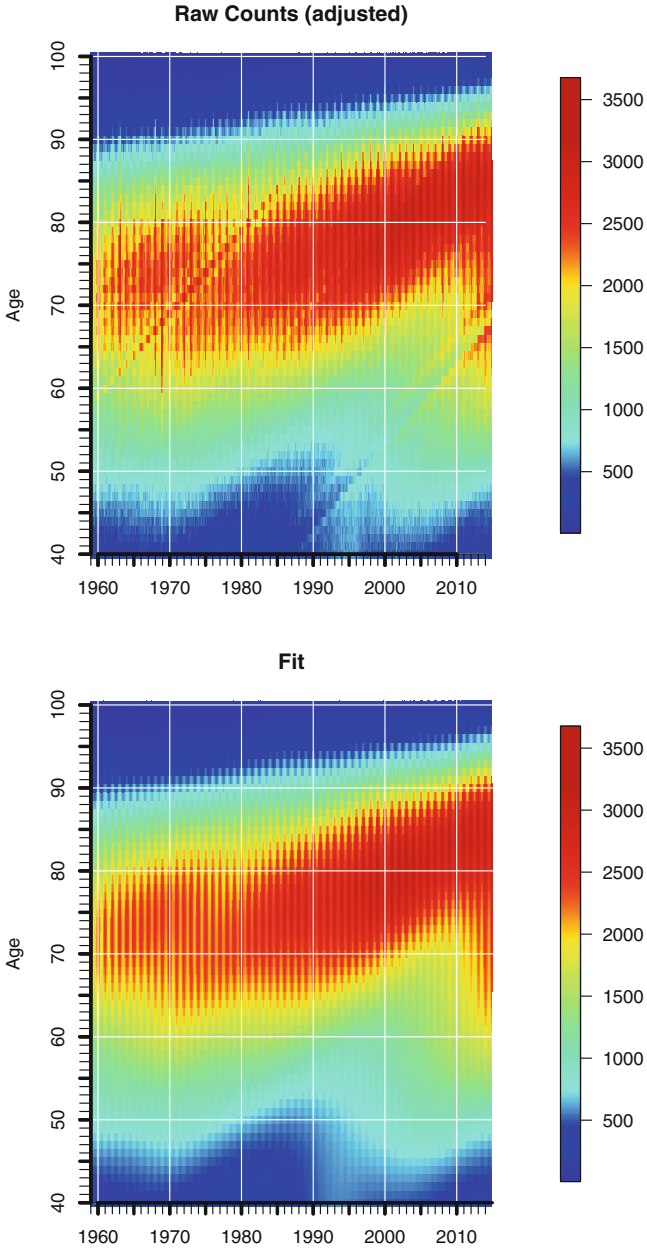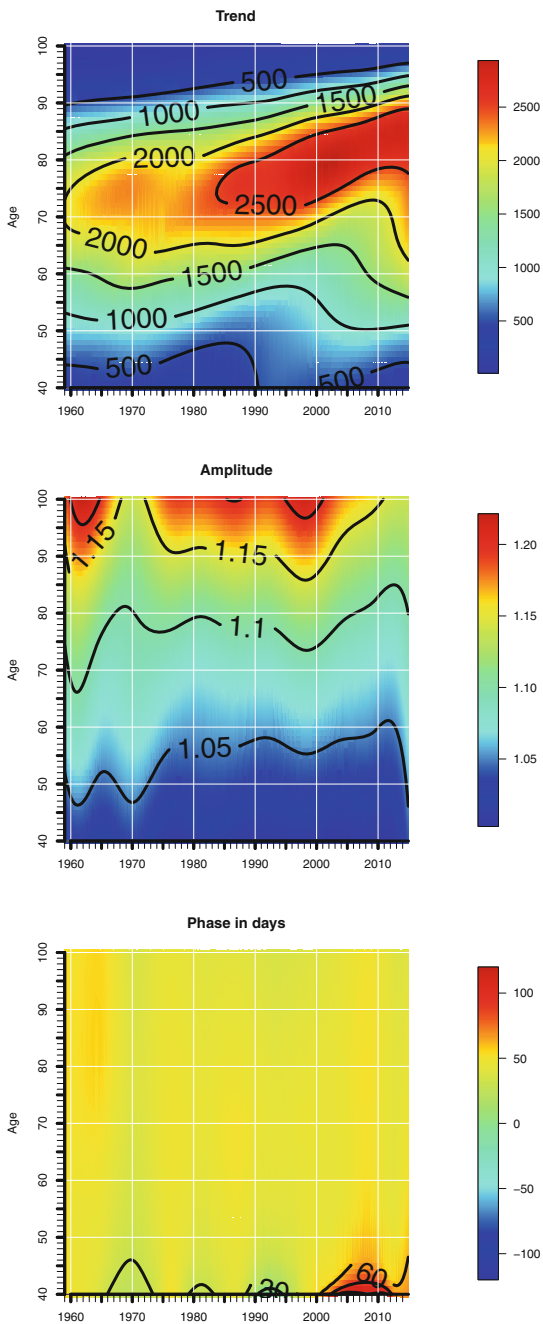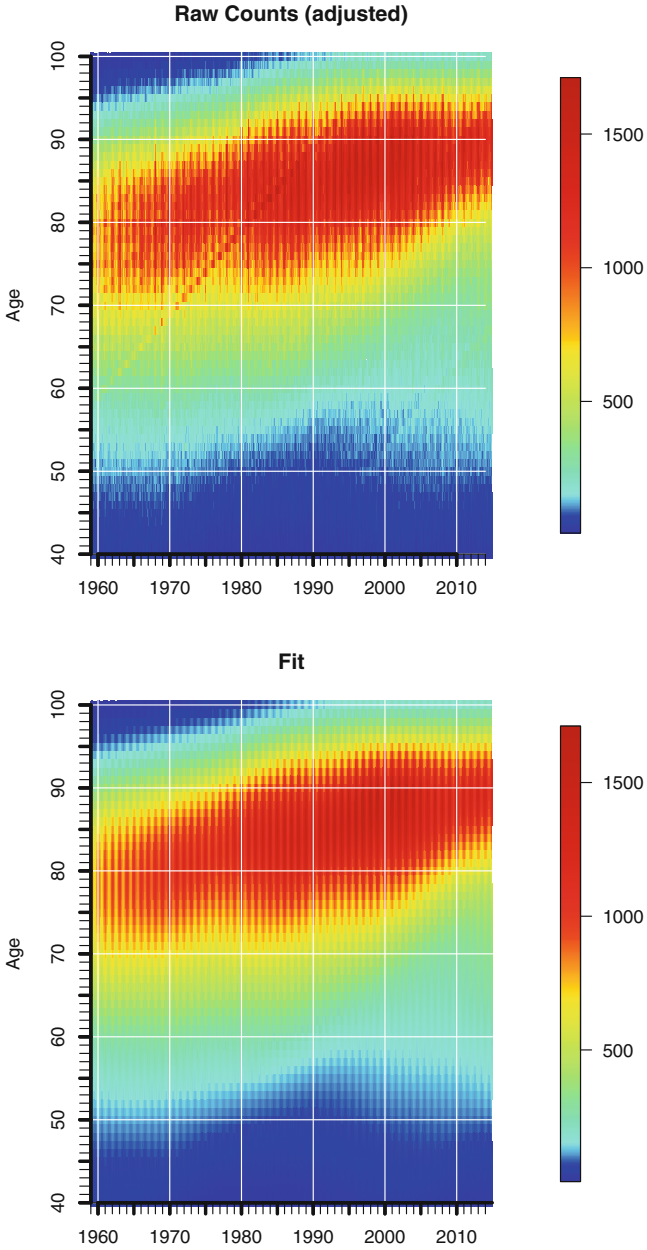Human Mortality Database)

Fig. 9.4 Seasonality of mortality from all causes in the United States, 1959–2014, men, raw counts (adjusted for length of month) and fitted model (Data source: Human Mortality Database)

Fig. 9.5 Seasonality of
mortality from all causes in
the United States, 1959–2014,
men, estimated trend surface
(*top panel*), amplitude
(*middle panel*), and phase
(*bottom panel*) (Data source:
Human Mortality Database)

Fig. 9.6 Seasonality of mortality from heart diseases in the United States, 1959–2014, women, raw counts (adjusted for length of month) and fitted model (Data source: Human Mortality Database)

Fig. 9.7 Seasonality of mortality from heart diseases in the united states, 1959–2014, women, estimated trend surface (*top panel*), amplitude (*middle panel*), and phase (*bottom panel*) (Data source: Human Mortality Database)
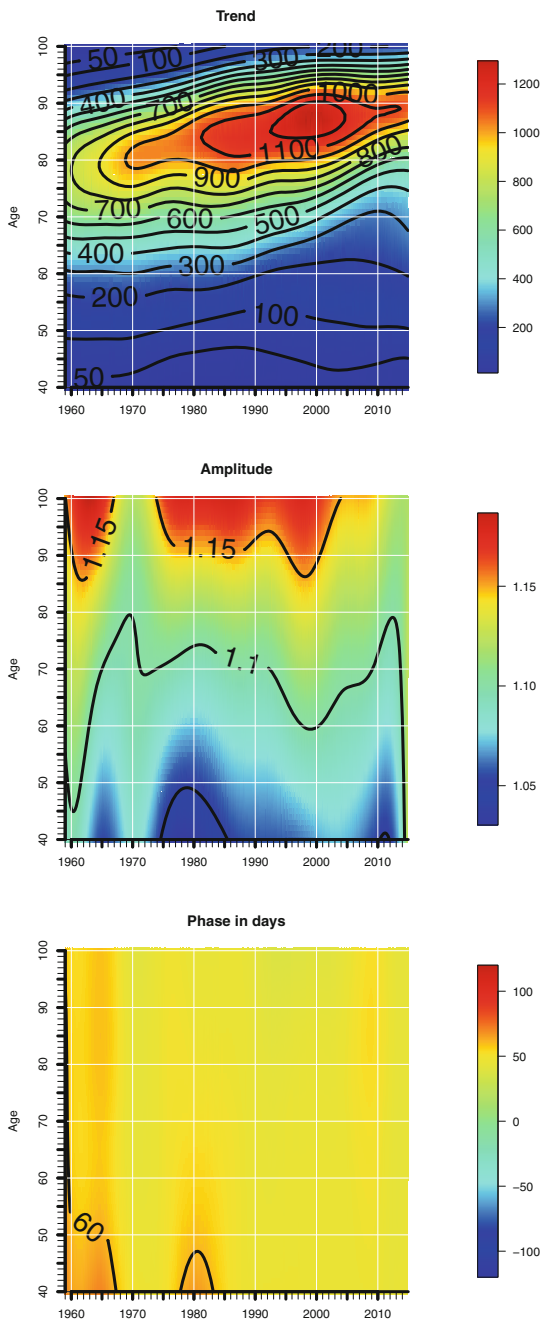
seasonal decomposition, this is the outcome of about 10 to 15% of excess deaths during the peak season. Also here we can not detect any period effects. In contrast to all-cause mortality with its peak at the end of January, deaths from heart diseases are highest at the end of February since the colors indicate a value of slightly below 60.

Most deaths from circulatory diseases can be attributed either to heart diseases or to cerebrovascular diseases. We analyzed the seasonal pattern of the latter category for women in Figs. 9.8 and 9.9 for men on pages 112–113. Comparable to heart diseases, the corridor with the largest number of deaths is moving to higher ages; the actual numbers are much smaller than for the other category, though. The extent of the seasonal pattern is remarkably similar to heart diseases. The amplitude is elevated again by about 10% around age 70 with larger fluctuations at higher ages and smaller fluctuations at younger ages. A clear trend over time is again not visible. Cerebrovascular diseases peak a bit earlier than heart diseases as suggested by the lower panels of Figs. 9.9 and 9.11. The highest number of deaths can be typically observed before the 30th day of the year, i.e., sometime between the middle and the end of January.

The Eurowinter group investigated the impact of cold temperatures on mortality about 20 years ago (e.g., Eurowinter Group 1997). They looked at ischaemic heart disease, cerebrovascular diseases, and respiratory diseases. As those three categories are mainly responsible for the seasonal pattern, we also analyzed the pattern for respiratory diseases, please see Figs. 9.12 and 9.13 on pages 114 & 115. The observed number of deaths is a bit higher than for cerebrovascular diseases. The seasonal decomposition on the second page shows that this is primarily the outcome of large seasonal fluctuations. Even the highest values in the trend surface on top are smaller than the corresponding values for cerebrovascular diseases. Excess deaths are, however, not only 10 to 15% higher in winter than throughout the year in general. The middle panel clearly illustrates that deaths from diseases such as pneumonia, influenza, COPD, etc. are at least 30% higher during peak season, which occurs at the end of February as the plot for the phase at the bottom illustrates. In contrast to the previously discussed two groups of circulatory diseases, the darker shades of blue during more recent years in the plot of the amplitude for respiratory diseases suggest that seasonal fluctuations became smaller over time.

Although motor vehicle accidents are by no means a major cause of death category, we decided nevertheless to include it. In the worst case 200 people of a given age died during a single month. The raw counts and fitted counts in Fig. 9.14 and the trend surface in Fig. 9.15 demonstrate that the period with the highest numbers of deaths is (thankfully) over. It occurred during the 1970s and 1980s to men aged around 20 years. The same plots show also that those men, born between 1950 and about 1965 suffer from a higher number of deaths also at higher ages. Since we are not looking at mortality per se but at death counts, this cohort effect is not necessarily the outcome of higher mortality; it could also be caused by the high number of births during those years ("baby boomers"). It is interesting to note, however, that we can also here detect a pattern on the 45° line for the seasonal amplitude and for the phase, which should be unaffected by
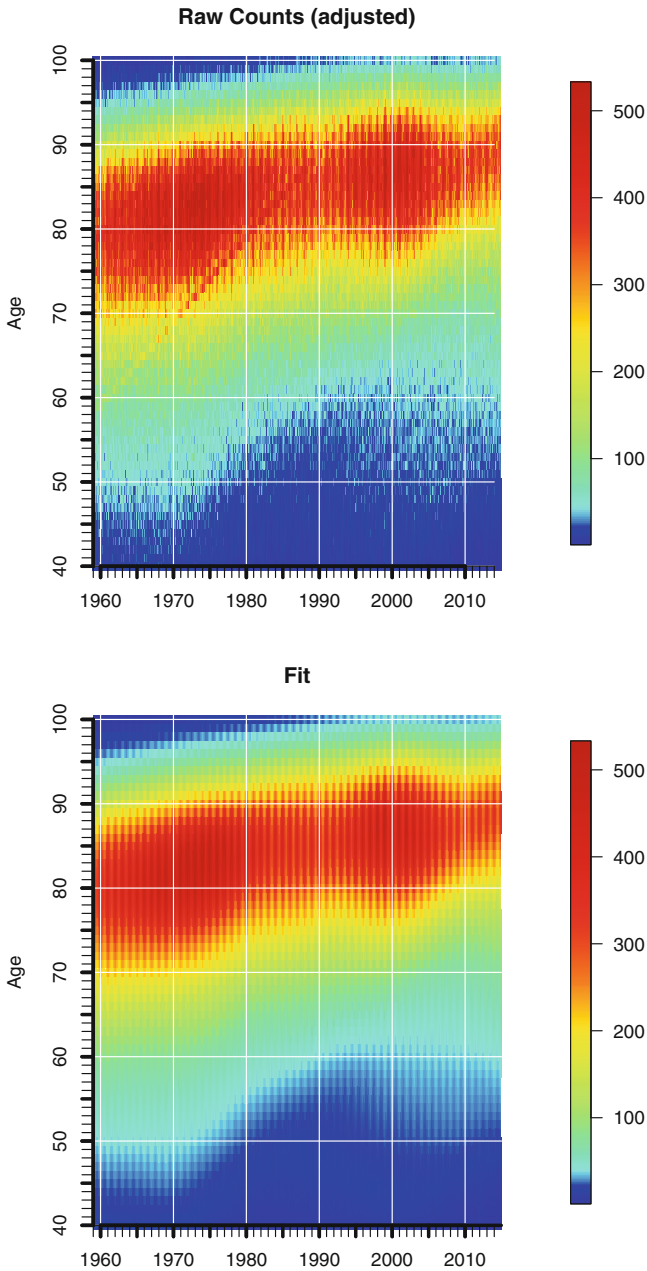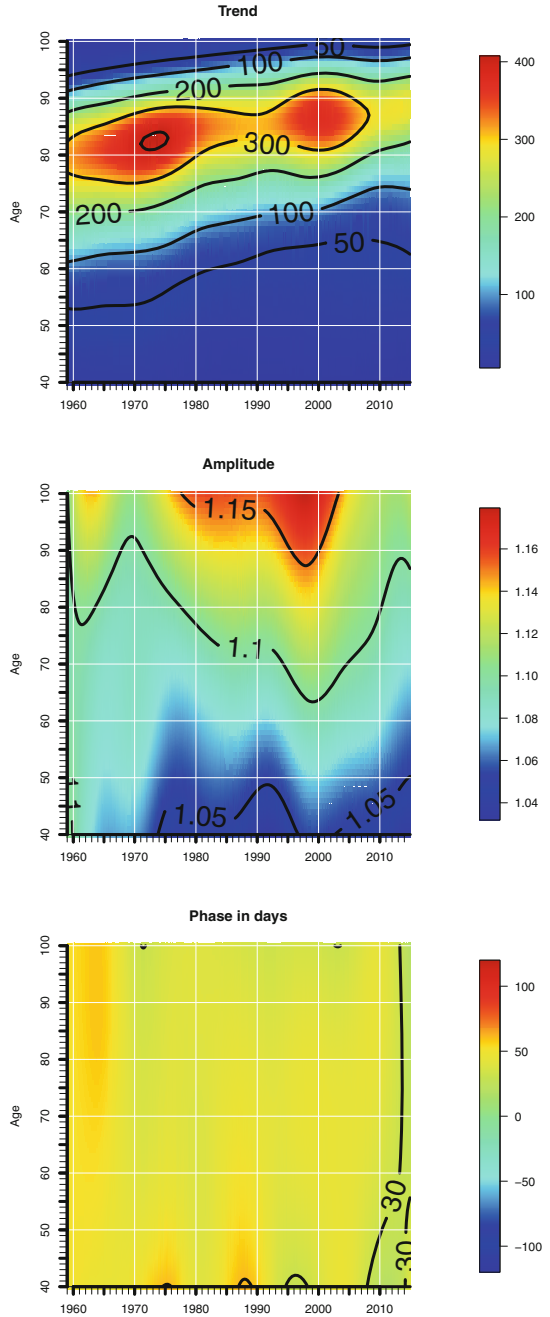
Fig. 9.8 Seasonality of mortality from cerebrovascular diseases in the United States, 1959–2014, women, raw counts (adjusted for length of month) and fitted model (Data source: Human Mortality Database)

Fig. 9.9 Seasonality of
mortality from
cerebrovascular diseases in
the United States, 1959–2014,
women, estimated trend
surface (*top panel*), amplitude
(*middle panel*), and phase
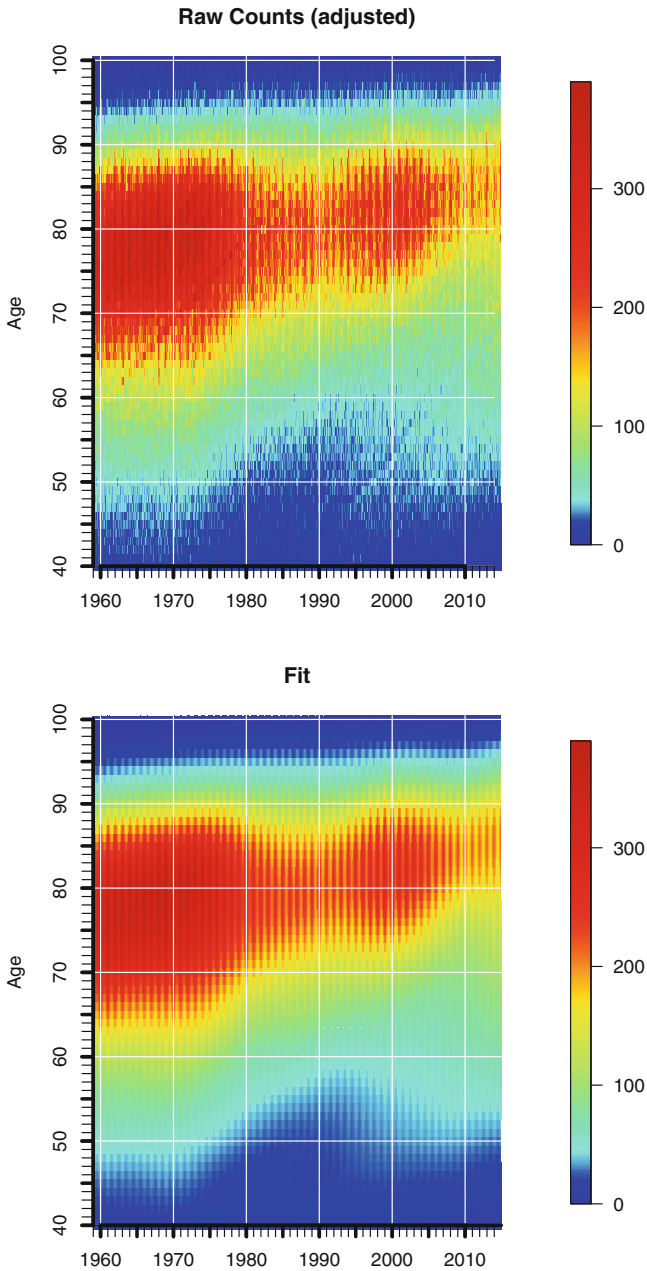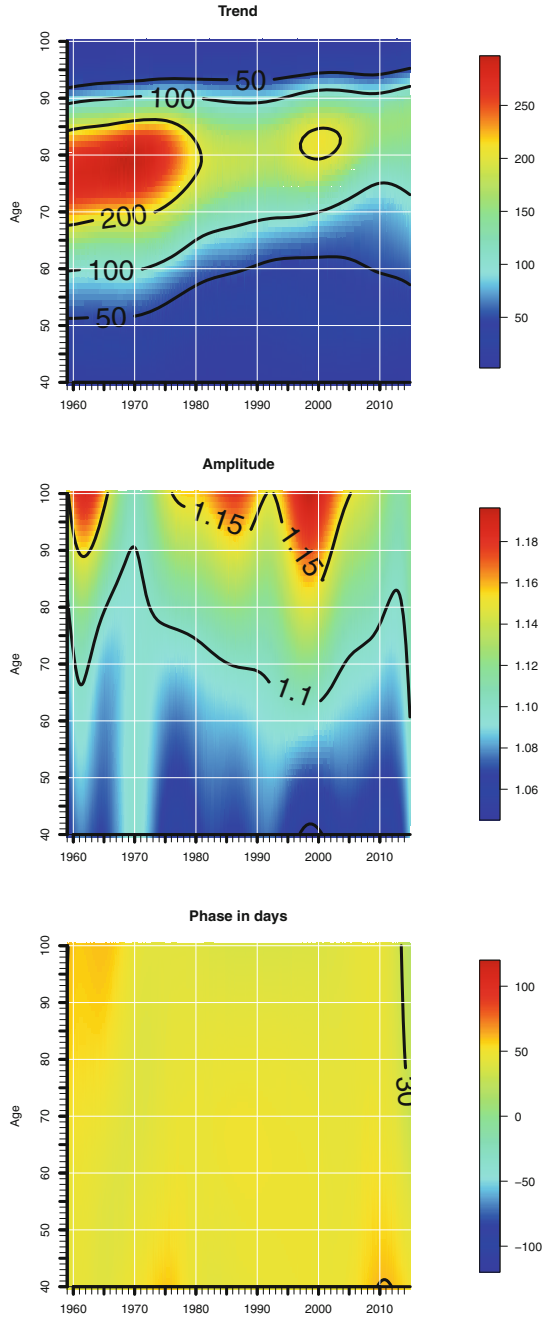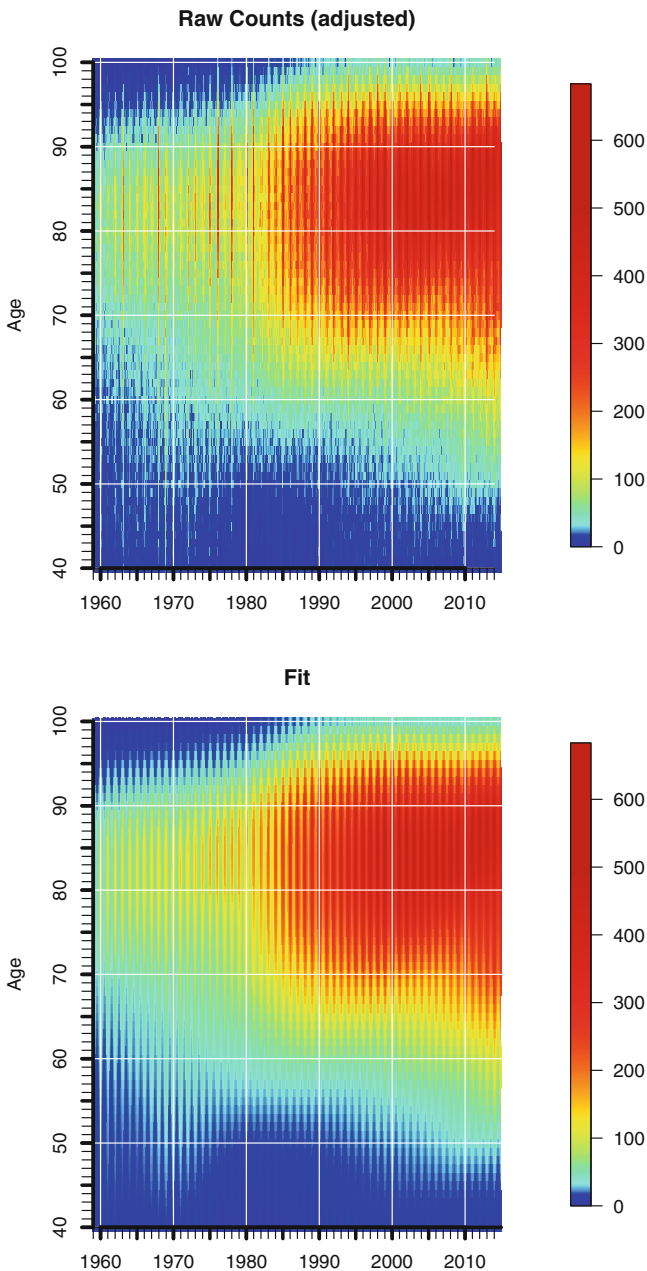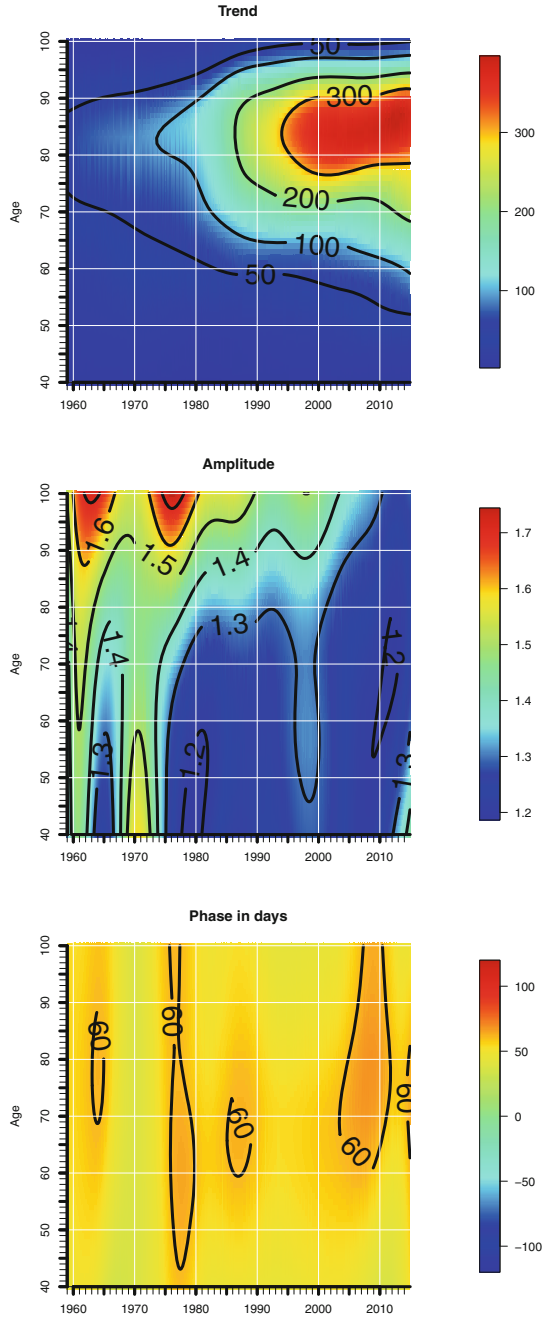(*bottom panel*) (Data source:
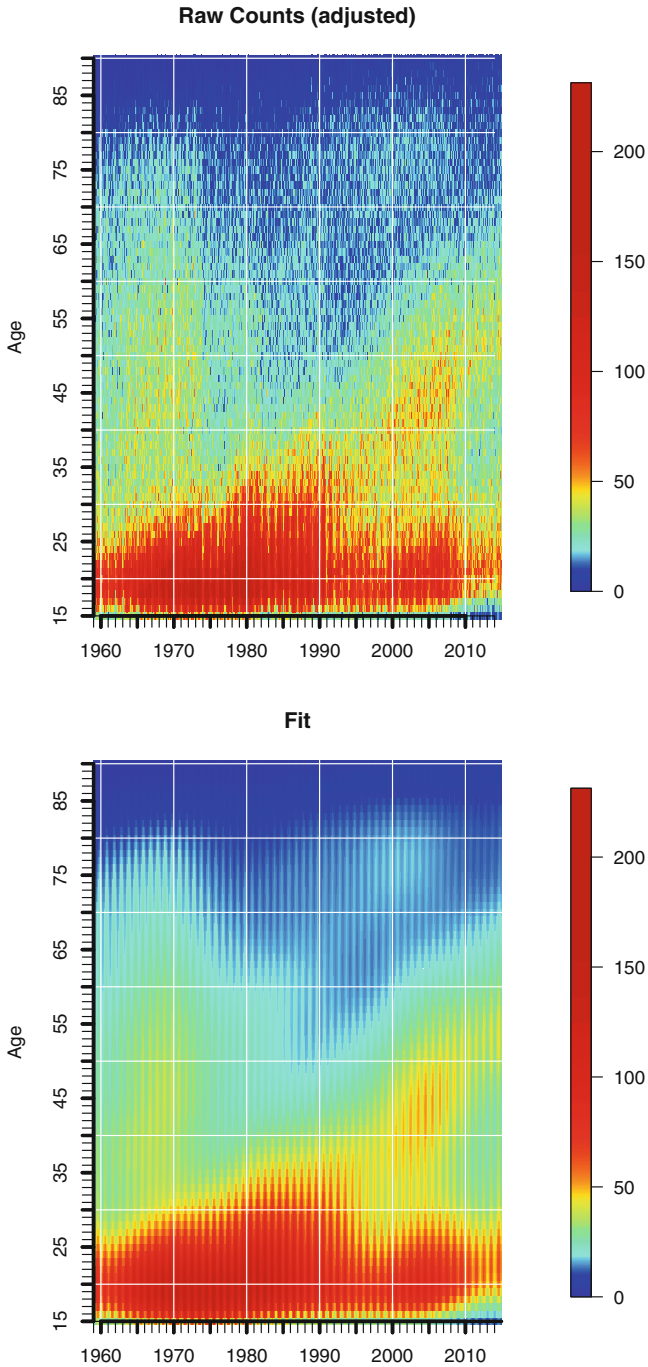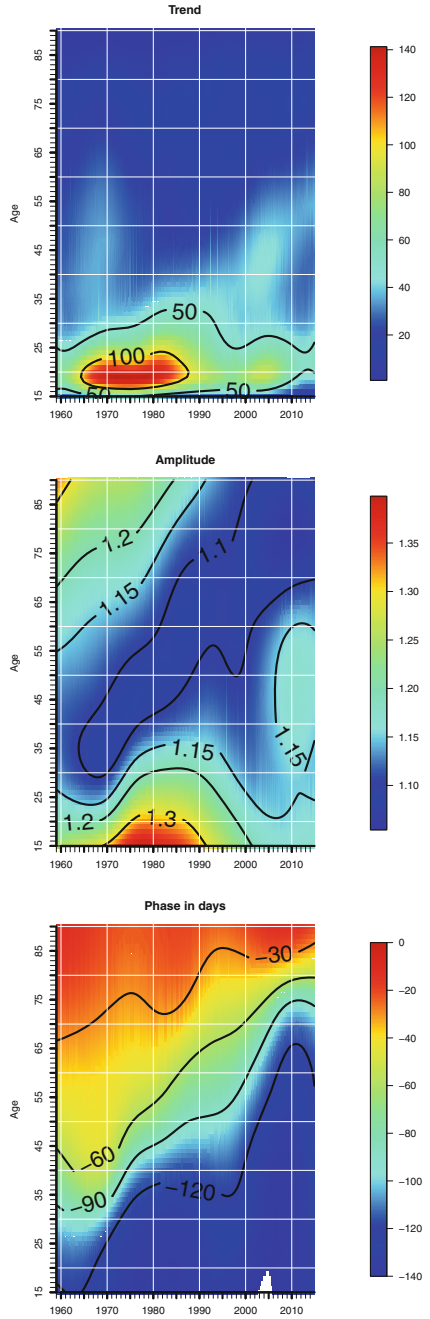Human Mortality Database)

Fig. 9.10 Seasonality of mortality from cerebrovascular diseases in the United States, 1959–2014, men, raw counts (adjusted for length of month) and fitted model (Data source: Human Mortality Database)

Fig. 9.11 Seasonality of
mortality from
cerebrovascular diseases in
the United States, 1959–2014,
men, estimated trend surface
(*top panel*), amplitude
(*middle panel*), and phase
(*bottom panel*) (Data source:
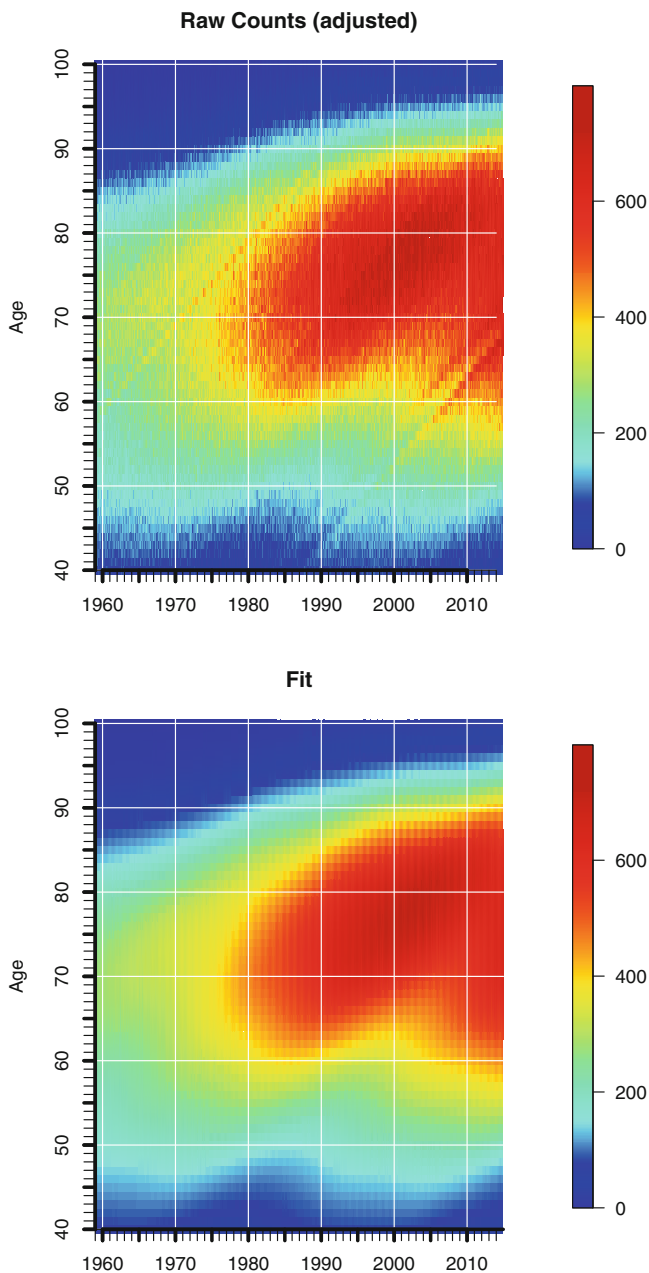Human Mortality Database)

Fig. 9.12 Seasonality of mortality from respiratory diseases in the United States, 1959–2014, women, raw counts (adjusted for length of month) and fitted model (Data source: Human Mortality Database)

Fig. 9.13 Seasonality of mortality from respiratory diseases in the United States, 1959–2014, women, estimated trend surface (*top panel*), amplitude (*middle panel*), and phase (*bottom panel*) (Data source: Human Mortality Database)
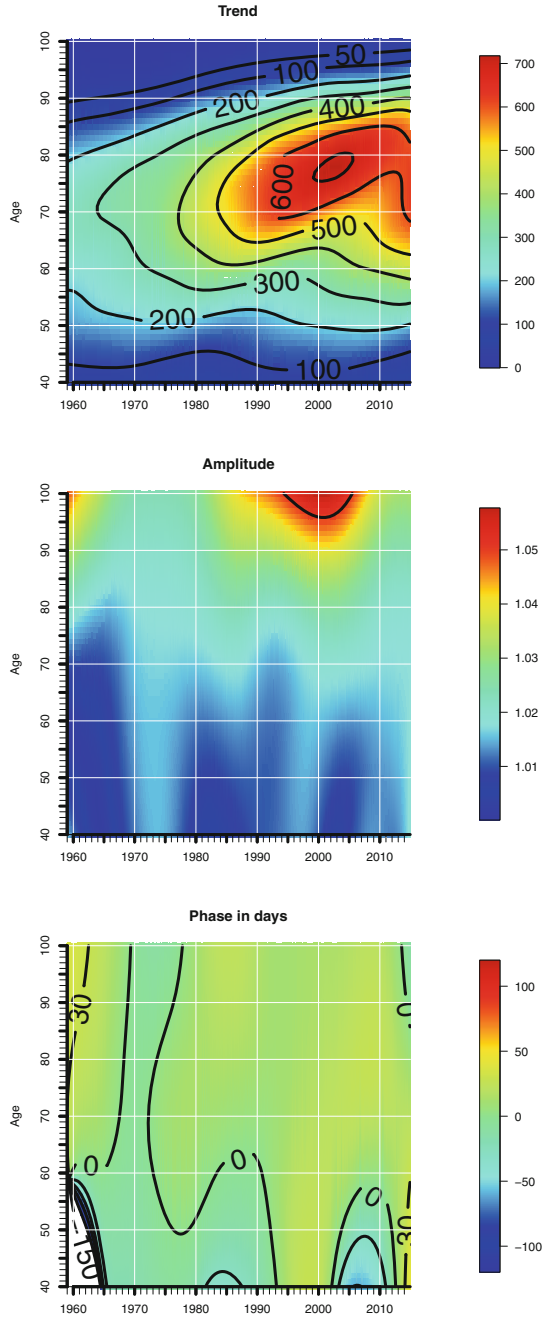
Raw Counts (adjusted)

Fit

Fig. 9.14 Seasonality of motor vehicle accidents in the United States, 1959–2014, men, raw counts (adjusted for length of month) and fitted model (Data source: Human Mortality Database)

Fig. 9.15 Seasonality of
motor vehicle accidents in the
United States, 1959–2014,
men, estimated trend surface
(*top panel*), amplitude
(*middle panel*), and phase
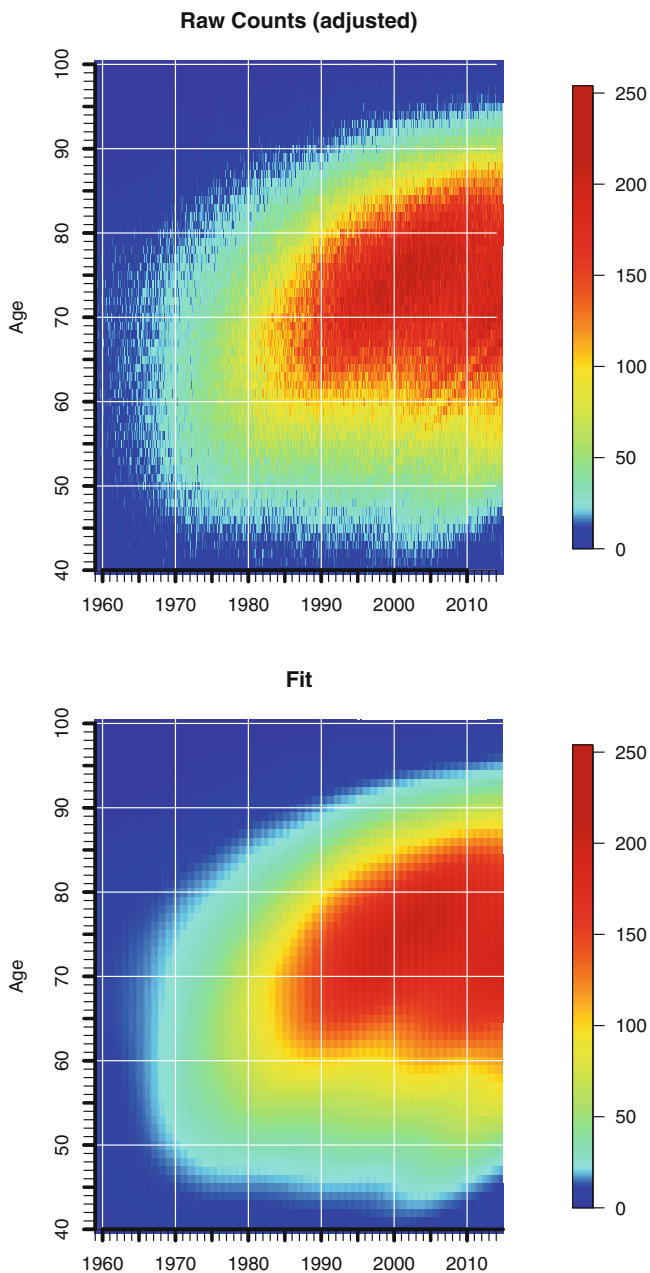(*bottom panel*) (Data source:
Human Mortality Database)

Fig. 9.16 Seasonality of mortality from all cancers in the United States, 1959–2014, women, raw counts (adjusted for length of month) and fitted model (Data source: Human Mortality Database)

Fig. 9.17 Seasonality of
mortality from all cancers in
the United States, 1959–2014,
women, estimated trend
surface (*top panel*), amplitude
(*middle panel*), and phase
(*bottom panel*) (Data source:
Human Mortality Database)

Fig. 9.18 Seasonality of mortality from lung cancer in the United States, 1959–2014, women, raw counts (adjusted for length of month) and fitted model (Data source: Human Mortality Database)

Fig. 9.19 Seasonality of
mortality from lung cancer in
the United States, 1959–2014,
women, estimated trend
surface (*top panel*), amplitude
(*middle panel*), and phase
(*bottom panel*) (Data source:
Human Mortality Database)
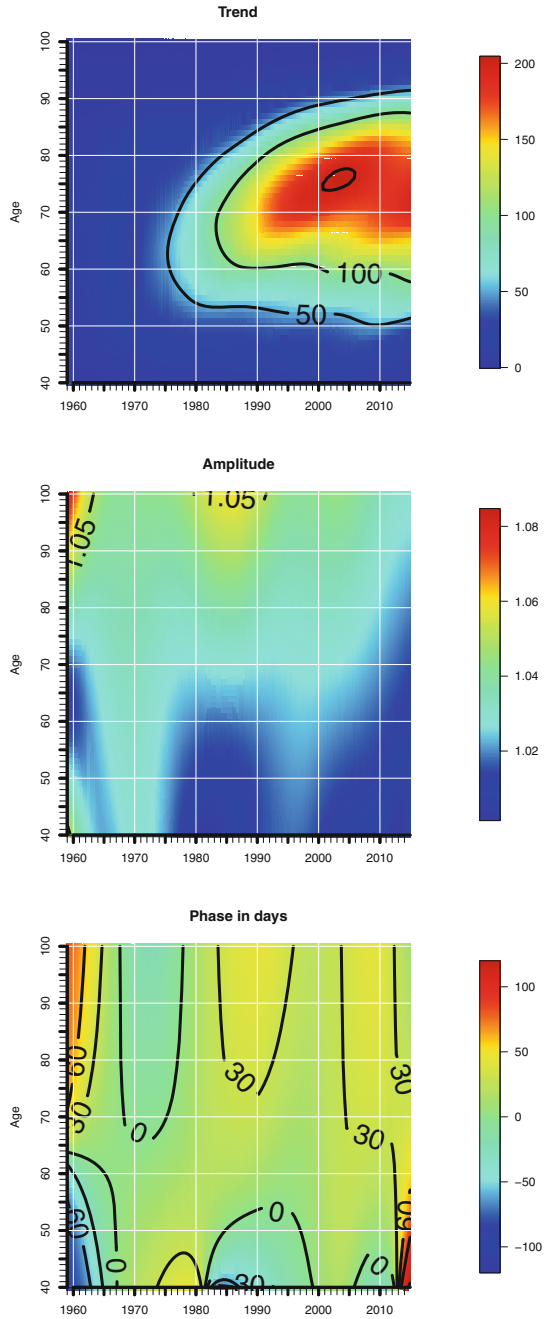
a larger population at risk since the trend surface accounts for it. The panel in the middle of Fig. 9.15 shows lowest seasonality for the birth cohort born before the baby boomers mentioned above. Also the change of the period when most deaths from motor vehicle accidents occur throughout a year features a cohort pattern. Whereas deaths from car accidents and similar causes peaked late in fall for older cohorts, the highest number of deaths for baby boomers and later generations are recorded at least 120 days before the 1st of January, which corresponds to August of a year.

We want to conclude this chapter by showing that cancers in general (see Figs. 9.16 and 9.17 on pages 118–119) and lung cancer (see Figs. 9.18 and 9.19 on pages 120–121) are examples of non-seasonal diseases. Clearly the fluctuations throughout a year are barely noticable as the middle panels of Figs. 9.17 and 9.19 illustrate.