A biometric analysis of infant mortality and temperature, northern Sweden 1895-1950

Göran Broström and Tommy Bengtsson

2022-02-10 19:04:47

Abstract

The effect of extreme temperatures on infant mortality in the Umeå and Skellefteå regions 1895-1950 is studied in a biometric analysis setting. More precisely, the effect of climate and weather, measured by temperature, average and extremes, on infant mortality is investigated. It turns out that climate (average) is more important than weather (extremes), low average temperatures are more important than temporary dips in temperature, but effects are different in neonatal and postneonatal settings.

Contents

1 Introduction

The impact of ambient temperature variations on infant mortality is studied for two northern Sweden areas, the Umeå coastal region and Skellefteå coastal and inland regions, during the first half of the twentieth century. Two recent papers (Junkka et al., 2021; Karlsson et al., 2021) studied neonatal mortality and temperature variations in this geographical area during the years 1880–1950. Climate and mortality in general is a research area that has generated great interest over the last years, see Bengtsson and Broström (2010) for an example of our own efforts.

The effect of seasonal variation and the occurrence of extreme monthly temperatures is studied and interacted with social class and time period. Studies are performed separately for neonatal and postneonatal mortality, and for winter and summer seasons.

Figure 1 shows the study area within Sweden, with the weather stations marked. The map is taken (with permission) from the paper by Junkka et al. (2021).

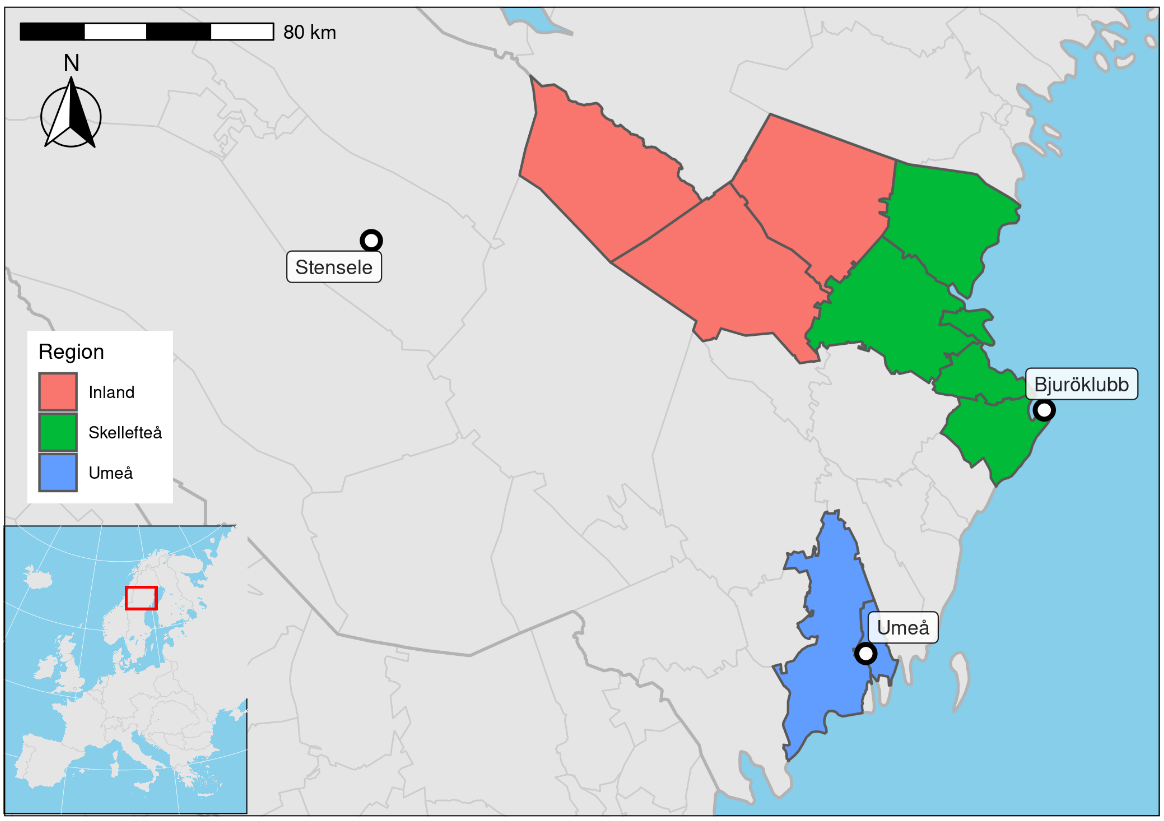

Figure 1: Umeå, Skellefteå (orange) and its inland (green).

2 Data

We have two sources of data which we combine into one data set suitable for our purpose. The first is demographic data obtained from the *Centre for Demographic and Ageing Research* (CEDAR, https://cedar.umu.se), the second is daily temperature measurements obtained from the *Swedish Meteorological and Hydrological Institute* (SMHI, https://www.smhi.se).

2.1 Infant mortality

Individual data with all births between 1 January 1895 and 31 December 1950 in two coastal and one inland areas of north Sweden, Skellefteå (51560 births) and Umeå (31213 births). They were followed until death or age one year, whichever came first. The following *static* characteristics were observed on each child:

birthdate Date of birth.

sex Girl or boy.

exit Number of days under observation.

event Logical, *TRUE* if a death is observed.

socBranch Working branch of father (if any).

illeg Mother unmarried?

parity Order among siblings.

Some crude statistics about infant, neonatal, and postneonatal mortality are shown in Figures.

Figure 2 shows the average weekly crude infant mortality, and a clear seasonal pattern is visible.

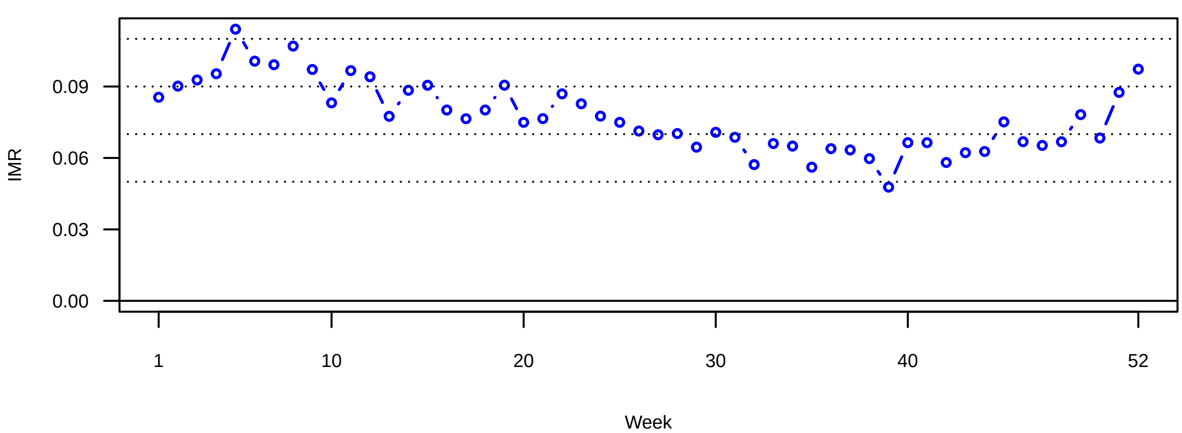

Figure 2: Crude infant mortality by week of year, Umeå/Skellefteå 1895–1950.

The average weekly neonatal mortality is shown in Figure 3.

The seasonal pattern is similar to the one we found above for infant mortality.

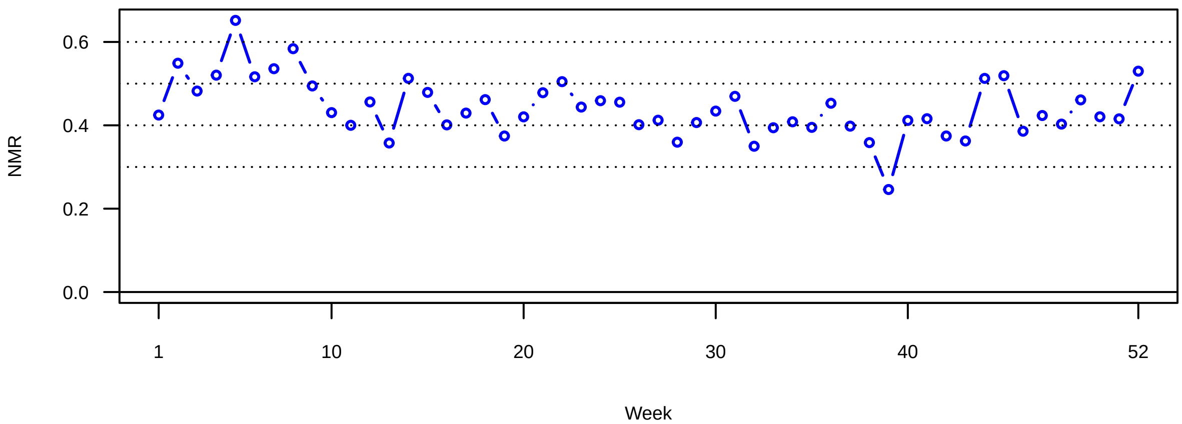

Figure 3: Crude neonatal mortality by week of year, Umeå/Skellefteå 1895–1950.

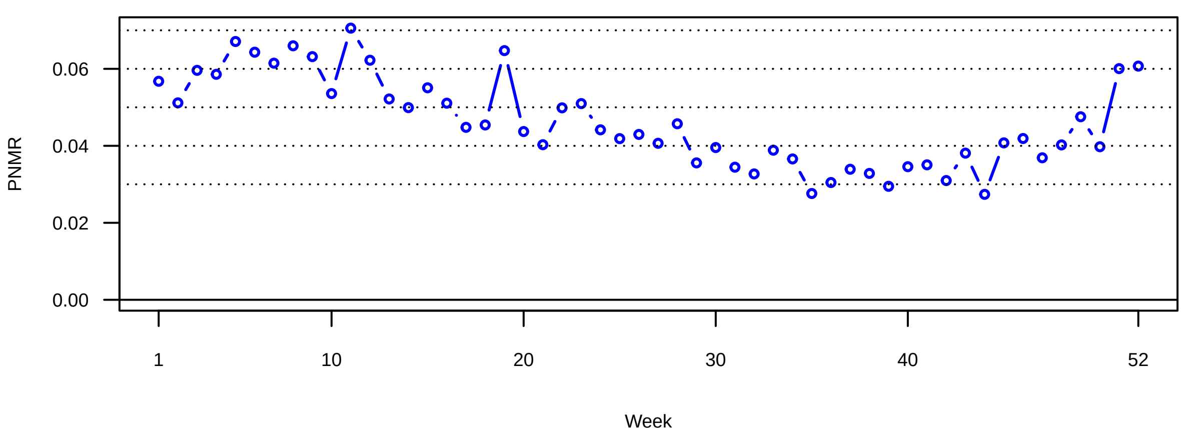

Figure 4: Crude postneonatal mortality by week of year, Umeå/Skellefteå 1895–1950.

The average weekly postneonatal mortality is shown in Figure 4.

The seasonal pattern is once again similar to the one we found for infant mortality. Next, the decline over the years in Figures 5 and 6.

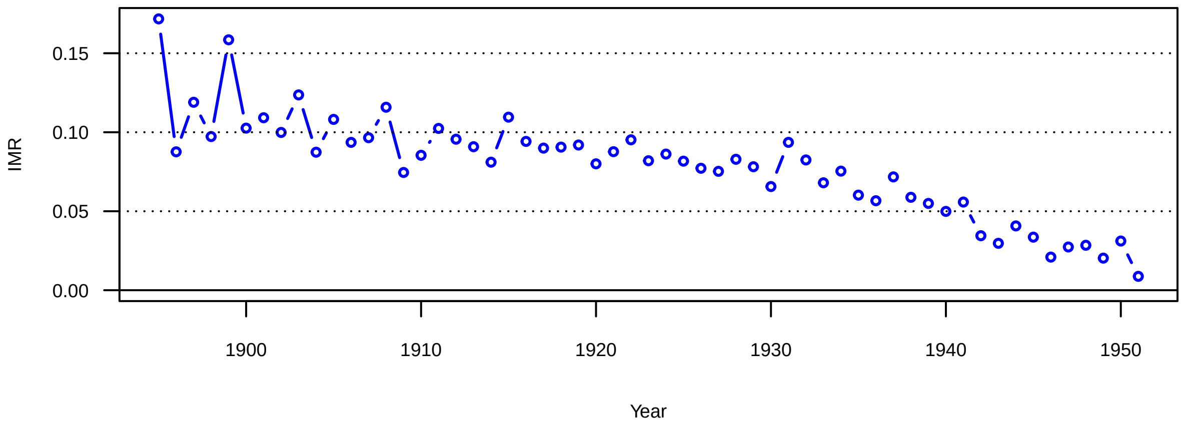

Figure 5: Crude IMR by year, Umeå-Skellefteå 1895–1950.

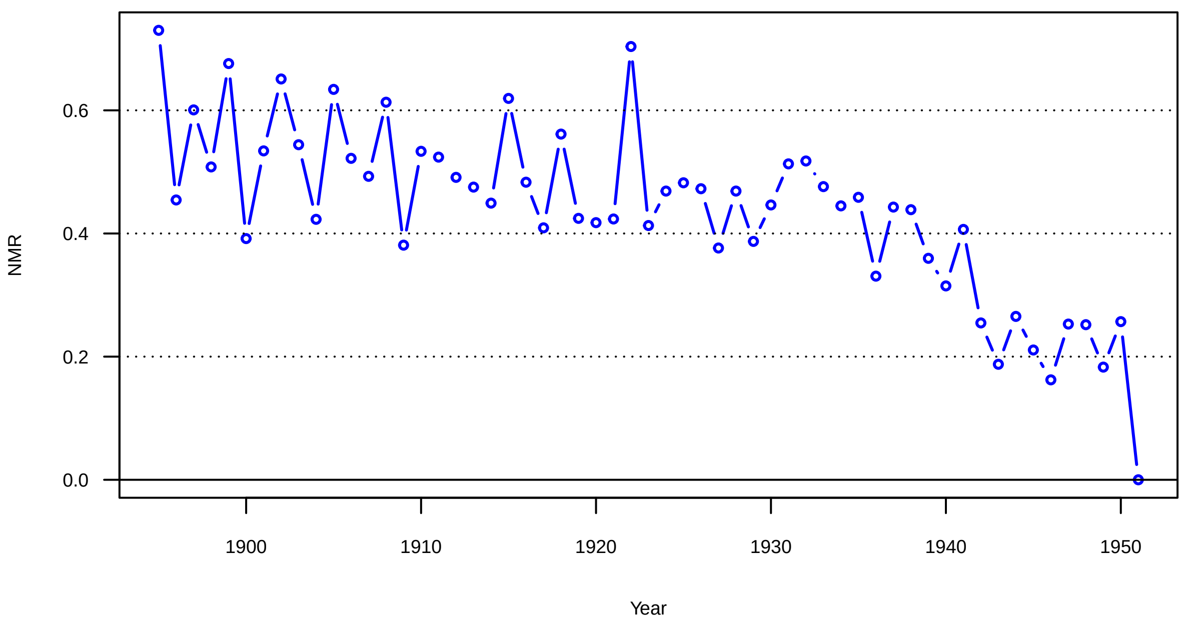

Figure 6: Crude NMR by year, Umeå-Skellefteå 1895–1950.

Note that the estimates of neonatal and postneonatal mortality are non-standard here: They are calculated as *occurrence-exposure rates*, that is, the number of deaths divided by *exposure time*. They are thus crude *hazard rates*, not probabilities, which is more common in definitions of mortality.

Table 1: Raw temperature data from first week of 1923, Umeå.

Date	**Time**	**Temperature**	**Quality**
1923-01-01	07:00:00	0.4	G
1923-01-01	13:00:00	0.6	G
1923-01-01	20:00:00	0.0	G
1923-01-02	07:00:00	-1.4	G
1923-01-02	13:00:00	-1.4	G
1923-01-02	20:00:00	-1.2	G
1923-01-03	07:00:00	0.4	G
1923-01-03	13:00:00	0.8	G
1923-01-03	20:00:00	1.2	G
1923-01-04	07:00:00	1.4	G
1923-01-04	13:00:00	1.2	G
1923-01-04	20:00:00	1.0	G
1923-01-05	07:00:00	-1.4	G
1923-01-05	13:00:00	-3.2	G
1923-01-05	20:00:00	-3.4	G
1923-01-06	07:00:00	1.0	G
1923-01-06	13:00:00	0.4	G
1923-01-06	20:00:00	0.4	G
1923-01-07	07:00:00	0.6	G
1923-01-07	13:00:00	0.4	G
1923-01-07	20:00:00	0.4	G

Table 2: Weekly summarized temperature data: Umeå 1923, first week.

week	**year**	**mintemp**	**maxtemp**	**meantemp**	**emintemp**	**emaxtemp**	**emeantemp**
1	1923	-3.4	1.4	-0.1	-17.73	-0.36	-7.54

2.2 Temperature

Temperature data are collected from three weather stations, *Umeå*, *Bjuröklubb* (used with population data from Skellefteå coastal area), and *Stensele* (Inland). All stations deliver daily temperature data covering our time period, usually three measures per day, morning, noon, and evening. In Table 1, the Umeå data from the week 1–7 January, 1923 is shown.

There are three measurements per day, or 21 per week. In the forthcoming analyses, the weekly data are summarized in a few measurements, see Table 2. Our rule for week numbering is that week No. 1 always start on January 1. Week No. 52 always ends at December 31, and so will be eight days long, except for leap years, when it will be nine days long.

Weekly averages (`mintemp`, `maxtemp`, `meantemp`) are calculated by week and year, and deviations from the averages (`emintemp`, `emaxtemp`, `emeantemp`) of the weekly averages are used as time-varying *communal covariates*. As an example, see Figure 7, where the variation around the average minimum temperature (`emintemp`) week 1 is shown.

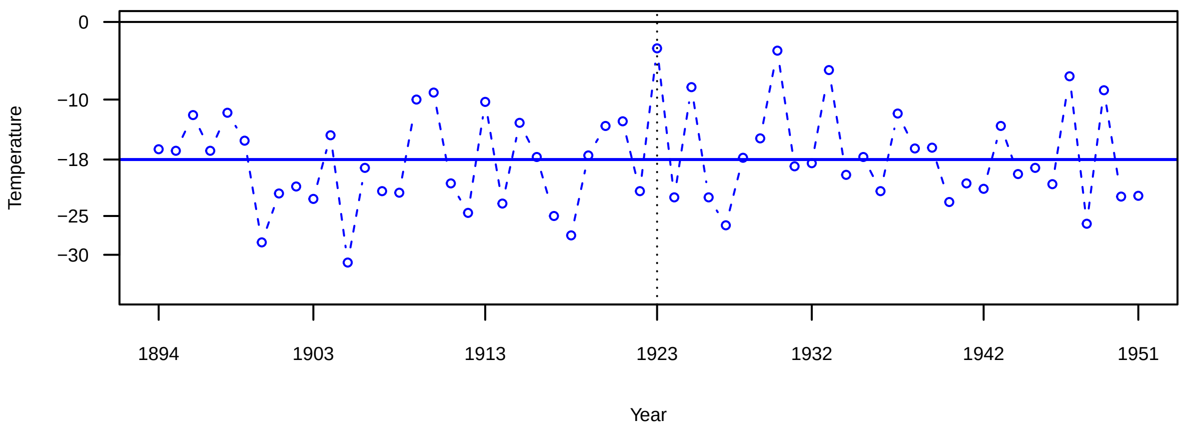

Figure 7: Minimum temperature the first week of each year.

Curiously, our randomly selected year 1923 turns out contain the warmest first week of all years, see Figure 7.

Figure 8 shows the average monthly distribution over all years. The subregional patterns and levels are very similar.

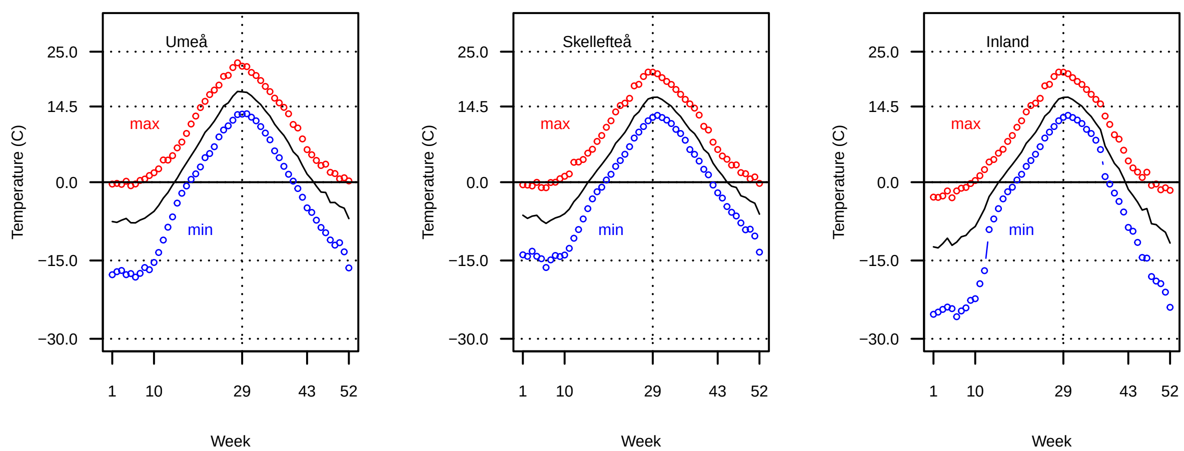

Figure 8: Weekly max, mean, and min temperature averages, 1895–1950.

Time trends of yearly average temperatures, see Figure 9.

2.3 Temperature as communal covariates

The two data sets, mortality and weather, are combined into one by treating temperature data as a communal covariate and incorporate it as such in the mortality data set. The function *make.communal* in the **R** (R Core Team, 2021) package *eha* (Broström, 2021a,b) is used for that purpose. Resulting data drame is partly shown in Table 3.

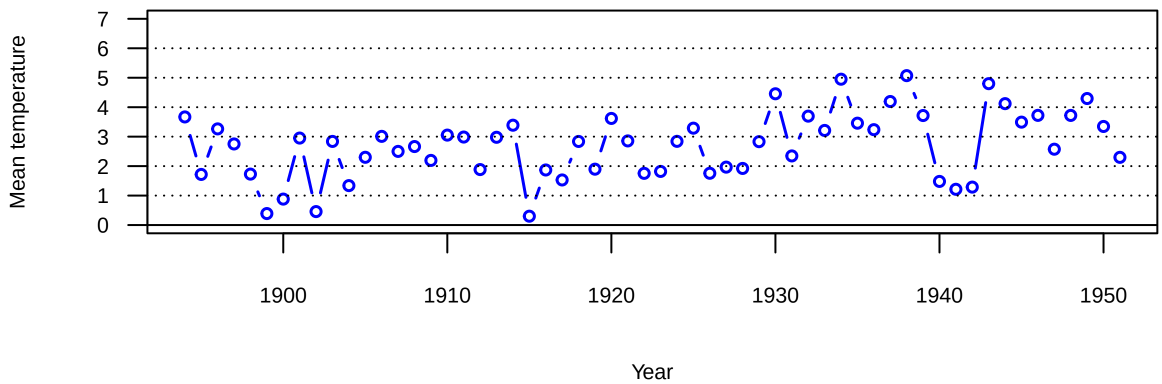

Figure 9: Yearly average temperatures, Umeå and Skellefteå.

Table 3: Data with communal covariates.

enter	exit	event	lowTemp	highTemp	extemp	emeantemp	week	year
0.0000000	0.0180327	0	FALSE	FALSE	0	12.584483	35	1900
0.0180327	0.0372634	0	FALSE	FALSE	-4	11.286207	36	1900
0.0372634	0.0564942	0	FALSE	FALSE	-1	10.086207	37	1900
0.0564942	0.0757250	0	FALSE	FALSE	1	6.910345	38	1900
0.0757250	0.0949557	0	FALSE	FALSE	-2	5.315517	39	1900
0.0949557	0.1141865	0	FALSE	FALSE	0	3.610345	40	1900

3 Statistical modelling

It turns out that extremely low temperature (`lowTemp`) is bad during all seasons except summer, and extremely high temperature (`highTemp`) is bad during summer, but good otherwise. So we group season into two categories, *summer* and *not summer*. In each case separate analyses for neonatal and postneonatal mortality are performed.

The *summer* half-year is the weeks 14–39, about 1 April to 30 September, and the *winter* half-year is the rest, weeks 1-13 and 40–52, 1 January to 31 March and 1 October to 31 December. This is the division made in Karlsson et al. (2021), and we keep it for comparability reasons.

4 Results

The results regarding neonatal mortality is much in accordance with the results found by Junkka et al. (2021). However, they used temperature in a "hockey-stick" regression with a breakpoint at 14.5 degrees Celsius and a negative slope (decreasing risk) to the left and a positive slope (increasing risk) to the right. Instead, we are using the average weekly temperature for the 52 weeks of a year, for each week averaging over all the years in the study, as our "reference points" ("climate"), adding deviances up and down ("weather") as "short-term temperature stress". This is similar to the way prices and mortality were related in for instance Bengtsson and Broström (2011), that is, a time series split into long time trend and short term variation.

Comments on other candidate covariates:

Birth month is left out in the analyses despite that fact that it is an important factor in neonatal mortality. However, we include *time of year* in terms of *winter* and *summer* as a time-varying covariate, and in the neonatal case it will coincide to a great extent with birth period. For the postneonal case the situation is different, but it turns out that for those infants who have survived the first month of life, birth month does not matter much. We separate the investigation into two parts, *neonatal* and *postneonatal* mortality. But first, a joint analysis.

Socioeconomic status is divided into two factor covariates: *socBranch* and *socStatus*. The latter should be seen relative to actual *socBranch*. We have information on whether the infant was *illegitimate* (mother unmarried), but we incorporate those cases in the category *none* of *socBranch* and *unknown* as category of *socStatus*.

Covariates may affect neonatal and postneonatal mortality differently, and one way to investigate that is to stratify the infant mortality data into two age intervals, one from birth to one month of age and the other from one month to one year of age. Then the interaction between the stratum variable and other variables of interest is investigated, with this result:

```
Single term deletions

Model:
oe(event, exposure) ~ strata(ageIvl) * (extemp + extemp.1 + emeantemp +
    socBranch + period)
                        Df   AIC    LRT  Pr(>Chi)
<none>                     41722
strata(ageIvl):extemp     1 41720  0.024  0.877627
strata(ageIvl):extemp.1   1 41722  2.244  0.134141
strata(ageIvl):emeantemp  1 41739 18.969 1.329e-05 ***
strata(ageIvl):socBranch  3 41730 13.818  0.003163 **
strata(ageIvl):period     2 41795 77.203 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We can see that the variables *emeantemp*, *socStatus* and *period* all have different effects on mortality in the neonatal case compared to the postneonatal case. In terms of *infant mortality*, we can say that these caovariates have non-proportional effects.

4.1 Neonatal mortality

The analyses are split into two parts by season, *winter* is one, and *summer* the other. Summer covers the months April to September, and winter the rest of the months.

4.1.1 Winter

This period refers to the months *October to March*. A Cox regression involves as interesting variables *highTemp*, an indicator of temperature at least four degrees above the expected for at least two weeks in a row, *emeantemp* the expected temperature the actual weeek, and *extemp* the *excess temperature* the actual week.

Table 4: Neonatal mortality, October to March.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
extemp	−0.024	−0.005	0.995	0.006	0.444
extemp.1	−0.026	−0.012	0.988	0.006	0.052
emeantemp	−4.140	−0.020	0.980	0.006	0.001
socBranch					0.446
office	0.105	0	1	(reference)	
farming	0.471	0.067	1.070	0.091	
none	0.024	0.111	1.118	0.168	
worker	0.400	0.127	1.135	0.090	
period					0.000
(1895,1914]	0.354	0	1	(reference)	
(1914,1935]	0.372	−0.088	0.915	0.052	
(1935,1951]	0.275	−0.620	0.538	0.070	
subreg					0.289
ume	0.296	0	1	(reference)	
ske	0.487	0.057	1.059	0.056	
inland	0.217	−0.038	0.963	0.072	
Events	1806	TTR	3899		
Max. logLik.	−2171				

4.1.2 Summer

This period refers to the months *April to September*. A Cox regression involves as interesting variables *highTemp*, an indicator of temperature at least four degrees above the expected for at least two weeks in a row, *emeantemp* the expected temperature the actual weeek, and *extemp* the *excess temperature* the actual week.

4.2 Postneonatal mortality

4.2.1 Winter

Interaction tests with temperature and period, postneonatal mortality:

```
Single term deletions

Model:
oe(event, exposure) ~ period * (extemp + extemp.1 + emeantemp)
                 Df   AIC    LRT Pr(>Chi)
<none>              17721
period:extemp     2 17722 4.2212   0.1212
period:extemp.1   2 17719 1.4788   0.4774
period:emeantemp  2 17720 2.3455   0.3095
```

Table 5: Neonatal mortality, April to September.

Covariate	Mean	Coef	H.R.	S.E.	L-R p
extemp	0.022	0.001	1.001	0.012	0.911
extemp.1	0.019	0.007	1.007	0.012	0.546
emeantemp	9.817	−0.006	0.994	0.005	0.163
socBranch					0.016
office	0.107	0	1	(reference)	
farming	0.460	0.229	1.258	0.096	
none	0.026	0.452	1.572	0.157	
worker	0.407	0.160	1.173	0.095	
period					0.000
(1895,1914]	0.333	0	1	(reference)	
(1914,1935]	0.370	−0.137	0.872	0.055	
(1935,1951]	0.297	−0.552	0.576	0.071	
subreg					0.371
ume	0.295	0	1	(reference)	
ske	0.485	0.081	1.084	0.059	
inland	0.220	0.038	1.039	0.071	
Events	1689	TTR	4026		
Max. logLik.	−2027				

This shows that the temperature effects do not vary much with time period, but we continue with separate analyses for the time periods (1895,1914], (1914,1935], (1935,1951]. Just to make sure, will be joined when appropriate.

Table 6: Postneonatal mortality, October to March 1895-1914. Adjusted for sex and parity.

Covariate	Mean	Coef	H.R.	S.E.	L-R p
extemp	−0.273	−0.018	0.982	0.007	0.017
extemp.1	−0.237	−0.016	0.984	0.008	0.036
emeantemp	−4.154	−0.050	0.951	0.008	0.000
socBranch					0.002
office	0.063	0	1	(reference)	
farming	0.557	0.264	1.303	0.146	
none	0.022	0.366	1.442	0.248	
worker	0.358	0.443	1.558	0.147	
subreg					0.000
ume	0.305	0	1	(reference)	
ske	0.476	−0.382	0.683	0.068	
inland	0.218	−0.516	0.597	0.089	
Events	1111	TTR	15642		
Max. logLik.	−3884				

4.2.1.1 The time period 1895–1913 Table 6 shows that the effects of extreme temperatures have the expected directions, extra high temperatures is *positive* (remember that this concerns winter conditions), but not statistically significant. *Climate* (emeantemp), on the other hand shows to be very important, where increasing temperature by one degree lower mortality by 0.7 per cent. The normal short term fluctuations (extemp) are not so important, but in the expected direction.

The importance of social class (socBranch) in explaining the effects of climate and weather is negligible, but there is variation between the regions regarding *climate*:

Table 7: Postneonatal mortality, Umeå, October to March 1895-1913.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
extemp	−0.223	−0.014	0.986	0.013	0.264
extemp.1	−0.183	−0.016	0.984	0.013	0.209
emeantemp	−3.621	−0.013	0.987	0.012	0.291
socBranch					0.001
office	0.099	0	1	(reference)	
farming	0.425	0.597	1.817	0.227	
none	0.024	0.827	2.286	0.360	
worker	0.452	0.775	2.171	0.224	
Events	436	TTR	4777		
Max. logLik.	−1429				

Table 8: Postneonatal mortality, Skellefteå with inland, October to March 1895-1913.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
extemp	−0.294	−0.019	0.981	0.009	0.036
extemp.1	−0.260	−0.016	0.985	0.009	0.098
emeantemp	−4.388	−0.063	0.939	0.009	0.000
socBranch					0.149
office	0.046	0	1	(reference)	
farming	0.616	−0.066	0.936	0.190	
none	0.021	−0.075	0.927	0.344	
worker	0.317	0.125	1.133	0.194	
Events	675	TTR	10865		
Max. logLik.	−2444				

We can conclude from Tables 7 and 8 that the main difference between Umeå and Skellefteå is that *climate* has more severe effect in Skellefteå (with inland), while *weather* is less important.

4.2.1.2 The time period 1914–1934 Table 9 shows that the weather effect is lagged one week (*extemp.1*), and that climate continues to be very important. Let's look at regional differences.

4.2.1.3 The time period 1935–1950 Table 12 shows that

4.2.2 Summer

The result in Table 15 shows that *climate* (`emeantemp`) is more important than *weather* (`exTemp`). Moreover, no signs of interaction between weather or climate and the rest of covariates (not shown).

Table 9: Postneonatal mortality, October to March 1914-1934. Adjusted for sex, parity, and subregion.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
extemp	−0.020	0.004	1.004	0.008	0.587
extemp.1	−0.044	−0.025	0.975	0.008	0.001
emeantemp	−4.170	−0.056	0.946	0.009	0.000
socBranch					0.012
office	0.098	0	1	(reference)	
farming	0.497	0.165	1.179	0.125	
none	0.026	0.259	1.296	0.234	
worker	0.380	0.341	1.407	0.124	
subreg					0.000
ume	0.265	0	1	(reference)	
ske	0.508	−0.292	0.747	0.076	
inland	0.227	−0.382	0.683	0.096	
Events	967	TTR	17087		
Max. logLik.	−3594				

Table 10: Postneonatal mortality, Umeå, October to March 1914-1934.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
extemp	−0.118	0.018	1.019	0.015	0.221
extemp.1	−0.138	−0.022	0.978	0.015	0.128
emeantemp	−3.619	−0.061	0.941	0.016	0.000
socBranch					0.005
office	0.185	0	1	(reference)	
farming	0.388	0.552	1.738	0.197	
none	0.041	0.358	1.430	0.363	
worker	0.385	0.644	1.903	0.195	
Events	303	TTR	4522		
Max. logLik.	−1069				

5 Conclusion

Remains to be written, especially the discussion about temperature and mortality.

Used processing time (seconds):

```
    user   system  elapsed 
4048.438 3237.694 1067.373 
```

Done: 2022-02-10 19:22:35.

References

Bengtsson, T. and Broström, G. (2010). Mortality crisis in rural southern Sweden 1766–1860. In Kurosu, T., Bengtsson, T., and Campbell, C., editors, *Demographic Response to*

Table 11: Postneonatal mortality, Skellefteå with inland, October to March 1914-1934.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
extemp	0.015	−0.001	0.999	0.010	0.908
extemp.1	−0.011	−0.026	0.974	0.009	0.005
emeantemp	−4.369	−0.050	0.951	0.009	0.000
socBranch					0.059
office	0.066	0	1	(reference)	
farming	0.536	−0.167	0.846	0.158	
none	0.020	0.128	1.137	0.307	
worker	0.378	0.052	1.053	0.160	
Events	664	TTR	12566		
Max. logLik.	−2513				

Table 12: Postneonatal mortality, October to March 1935-1950. Adjusted for sex, parity, and subregion.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
extemp.1	0.394	−0.007	0.993	0.013	0.595
emeantemp	−4.129	−0.057	0.944	0.016	0.000
socBranch					0.027
office	0.172	0	1	(reference)	
farming	0.313	0.612	1.845	0.228	
none	0.028	0.757	2.132	0.388	
worker	0.487	0.532	1.702	0.217	
subreg					0.886
ume	0.318	0	1	(reference)	
ske	0.473	0.017	1.017	0.151	
inland	0.210	0.085	1.089	0.181	
sex					0.137
boy	0.516	0	1	(reference)	
girl	0.484	−0.182	0.833	0.123	
Events	268	TTR	13484		
Max. logLik.	−1270				

Economic and Environmental Crisis, pages 1–16. Reitaku University Press, Kashiwa.

Bengtsson, T. and Broström, G. (2011). Famines and mortality crises in 18th to 19th century southern Sweden. *Genus*, 67:119–139.

Broström, G. (2021a). *eha: Event History Analysis*. R package version 2.9.0. https://CRAN.R-project.org/package=eha.

Broström, G. (2021b). *Event History Analysis with R, Second Edition*. Chapman & Hall/CRC, Boca Raton.

Junkka, J., Karlsson, L., Lundevaller, E., and Schumann, B. (2021). Climate vulnerability of Swedish newborns: Gender differences and time trends of temperature-related neonatal mortality, 1880–1950. *Environmental Research*, 192. article id 110400.

Table 13: Postneonatal mortality, Umeå, October to March 1935-1950.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
extemp	0.376	0.014	1.014	0.031	0.658
extemp.1	0.364	0.002	1.002	0.031	0.943
emeantemp	−3.615	−0.036	0.964	0.031	0.229
socBranch					0.302
office	0.278	0	1	(reference)	
farming	0.226	0.579	1.785	0.347	
none	0.047	−0.146	0.864	0.757	
worker	0.450	0.419	1.521	0.317	
Events	73	TTR	4281		
Max. logLik.	−358				

Table 14: Postneonatal mortality, Skellefteå with inland, October to March 1895-1913.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
extemp	0.407	−0.002	0.998	0.017	0.886
extemp.1	0.408	−0.011	0.989	0.017	0.520
emeantemp	−4.368	−0.067	0.936	0.017	0.000
socBranch					0.030
office	0.122	0	1	(reference)	
farming	0.354	0.704	2.022	0.312	
none	0.019	1.312	3.713	0.479	
worker	0.504	0.636	1.889	0.307	
Events	195	TTR	9203		
Max. logLik.	−901				

Karlsson, L., Junkka, J., Schumann, B., and Häggström Lundevaller, E. (2021). Socioeconomic disparities in climate vulnerability: neonatal mortality in northern Sweden, 1880–1950. *Population and Environment*, 10.1007.

R Core Team (2021). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.

Table 15: Postneonatal mortality, April to September. Adjusted for sex, parity, time period, and subregion.

Covariate	**Mean**	**Coef**	**H.R.**	**S.E.**	**L-R p**
extemp	0.006	0.007	1.007	0.011	0.520
extemp.1	0.003	−0.012	0.988	0.011	0.274
emeantemp	9.963	−0.020	0.980	0.004	0.000
socBranch					0.000
office	0.107	0	1	(reference)	
farming	0.464	0.365	1.441	0.099	
none	0.025	0.449	1.566	0.171	
worker	0.403	0.490	1.632	0.098	
period					0.000
(1895,1914]	0.340	0	1	(reference)	
(1914,1935]	0.370	−0.225	0.798	0.050	
(1935,1951]	0.291	−1.054	0.349	0.076	
Events	1914	TTR	45958		
Max. logLik.	−7656				