Stat 221, Time Series Analysis. Day 1.

<u>Outline for the day:</u>

1. Syllabus, etc.

2. Examples of time series.

Note that the CCLE or Canvas website for this course is not maintained.

The course website is http://www.stat.ucla.edu/~frederic/221/F22 .

I do not give hw hints in office hours. Conceptual questions only.

Only one question is off limits, and it is "What did we do in class?"

Time series are numerical observations recorded sequentially in time. Usually the times are regularly spaced, though not always. And usually, the observations are *autocorrelated*, meaning two observations close together in time have values that are correlated. Often this correlation is high when the values are very close together in time, and then the correlation gradually wears off as the time between the observations gets bigger.

1. Examples of time series.

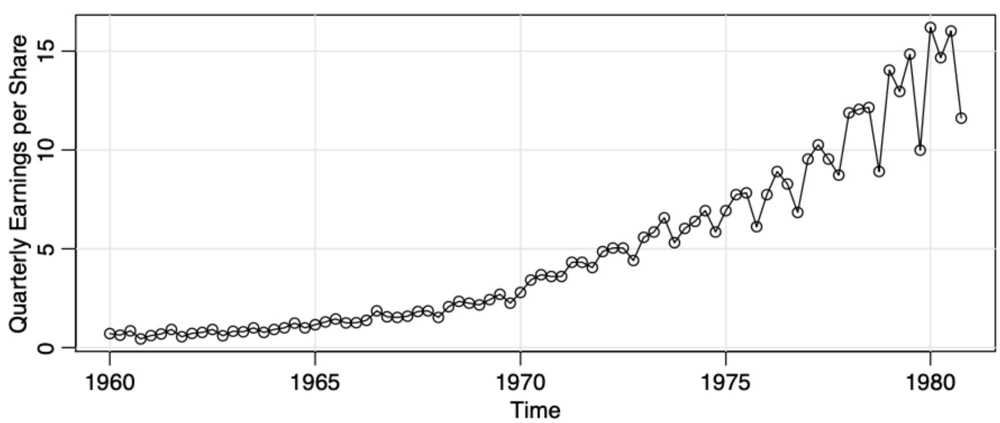


Fig. 1.1. *Johnson & Johnson quarterly earnings per share, 84 quarters, 1960-I to 1980-IV.*

Johnson & Johnson Quarterly Earnings

One great thing about the course textbook is they include the code and data used to make all their plots. But first you have to download and install the astsa library. In *R*, do

install.packages("astsa") ## to download the package

library(astsa) ## to load it into your current *R* session.

plot(jj, type="o", ylab="Quarterly Earnings per Share")

Global Warming

The data are the global mean land–ocean temperature index from 1880 to 2015, with the base period 1951-1980.

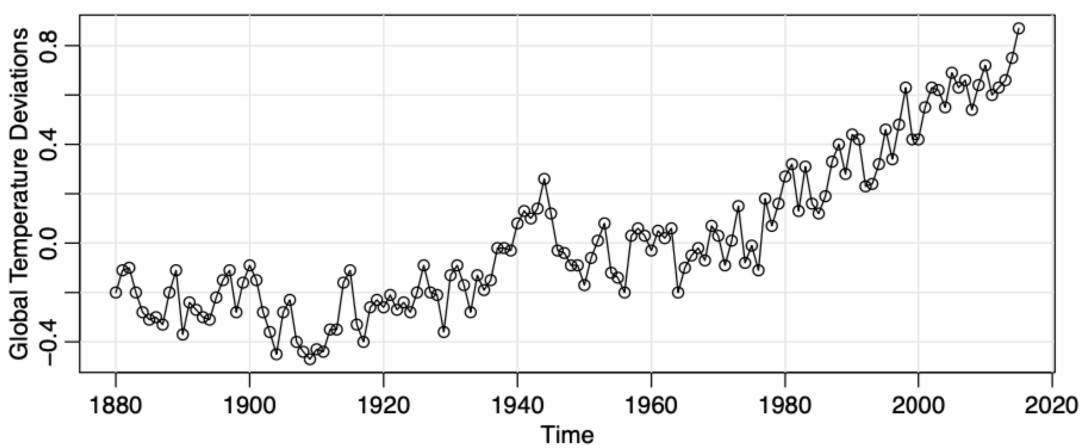


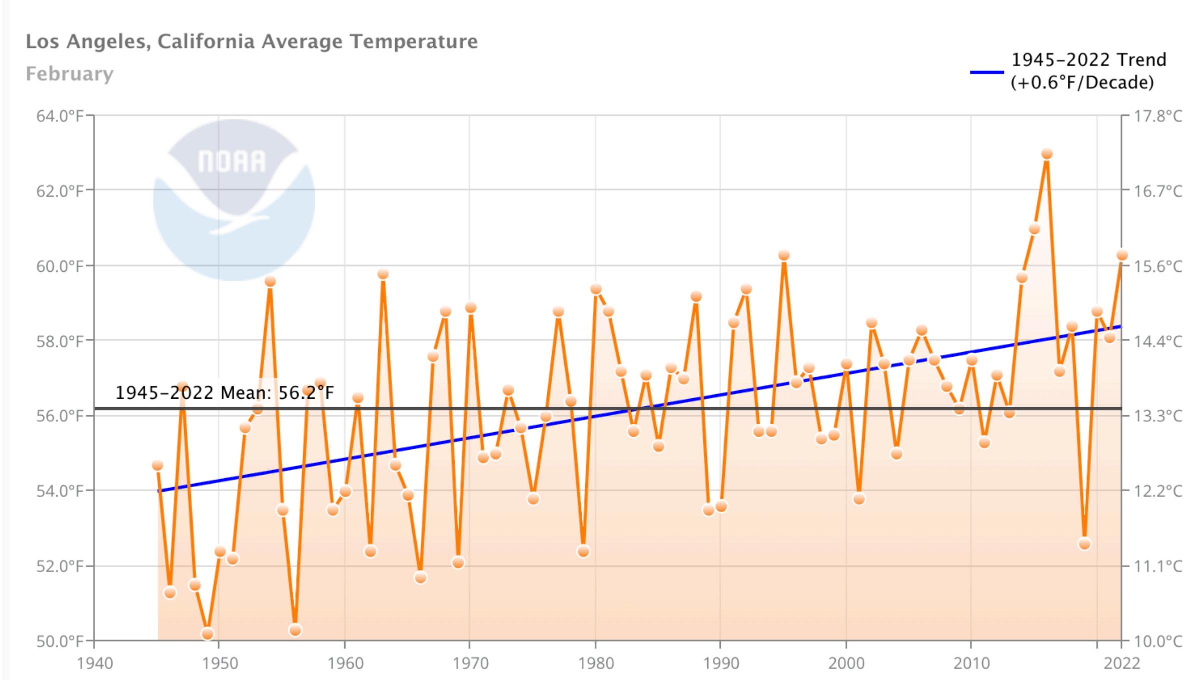
Fig. 1.2. *Yearly average global temperature deviations (1880–2015) in degrees centigrade.*

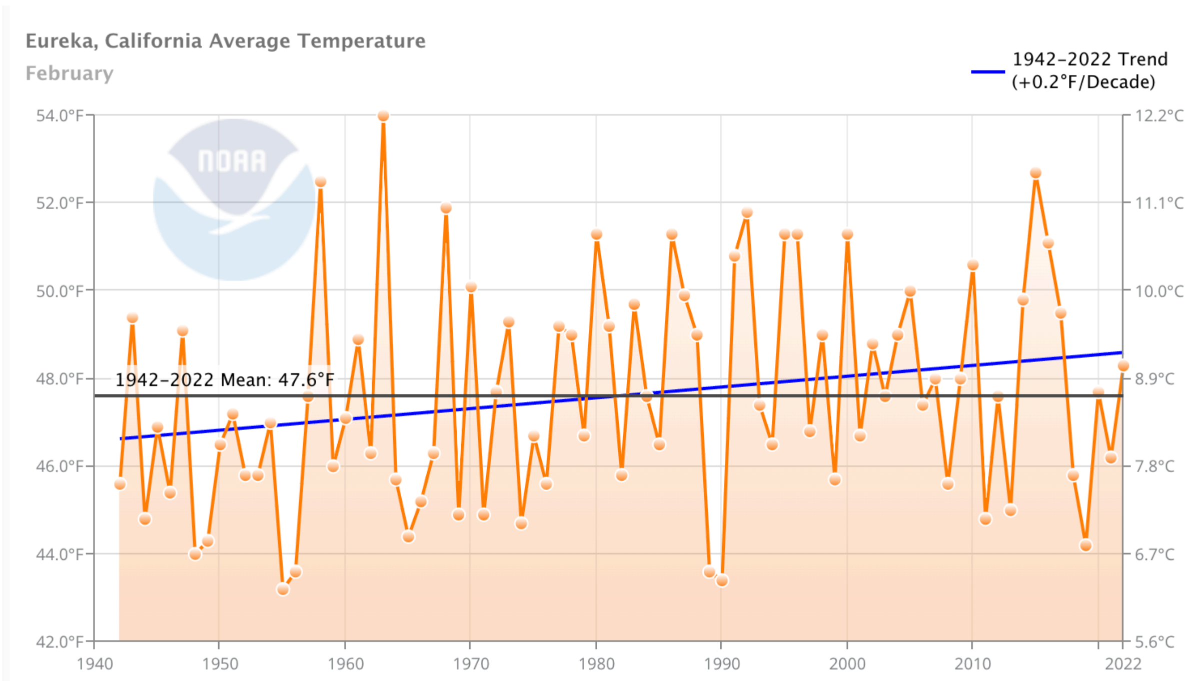
There is an obvious increasing trend. It is unclear if it is linear, or perhaps flat until 1940 or 1950 and then increasing thereafter. There is substantial variation around the trend, but the variation seems stable. In other words, it does not seem to be increasing or decreasing very much over time.

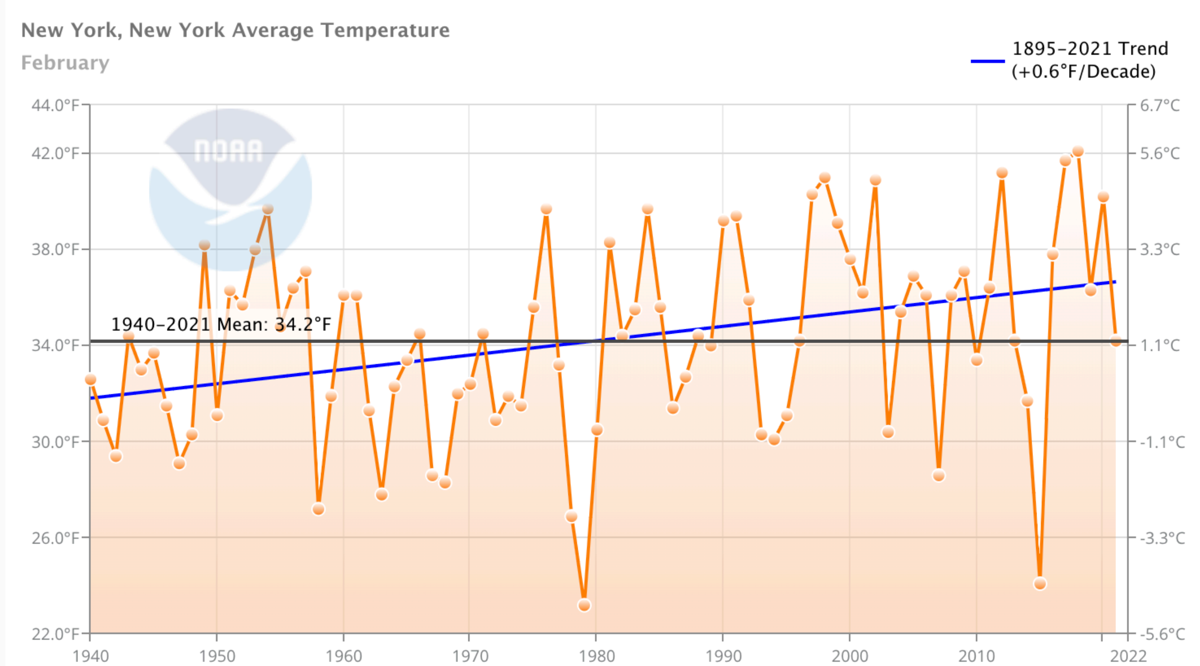
The book says the increasing trend here "has been used as an argument for the global warming hypothesis". One might argue global warming is more than a hypothesis at this point. The book also says "The question of interest for global warming proponents and opponents is whether the overall trend is natural or whether it is caused by some human-induced interface." Again, one might argue this is not really an open question anymore since there seems to be overwhelming scientific evidence at this point. The book goes on to say "Problem 2.8 examines 634 years of glacial sediment data that might be taken as a long-term temperature proxy. Such percentage changes in temperature do not seem to be unusual over a time period of 100 years." This is an often-used argument by those seeking to contend that human-induced global warming is not occurring. The argument is that, because temperatures seem to have changed a lot in the distant past, the current changes may be natural. As a result, they might not be human-induced, and there is also an implicit suggestion that they might not be such a big deal.

This argument can be rebutted, however, by a few different observations. First, we have an overwhelming amount of evidence about the precise mechanism by which humans are currently contributing to global warming, by the combustion of fossil fuels and other greenhouse gases. Second, this mechanism is not merely an explanation made after the fact, in which case it could be argued to be spurious, as such after-the-fact explanations sometimes are. In this case, scientists anticipated fifty years ago, and consistently over the past fifty years, that this temperature increase would occur, and in addition forecast quite accurately the nature of this increase. Third, there is considerably more evidence for human-induced climate change over the last 50 years than merely the global average temperatures. On a local level, the greenhouse effect caused by our use of fossil fuels has led to local temperature changes that are very neatly aligned with the global warming "hypothesis". For example, in urban areas where human populations are more dense and where the combustion of fossil fuels has been higher, local temperature increases over the past 50 years have correspondingly been higher. For example, according to an article in Scientific American, https://www.scientificamerican.com/article/heat-islands-cook-u-s-cities-faster-than-ever , "[i]n two thirds of the cities analyzed (41 of 60), urbanization and climate change appear to be combining to increase summer heat faster than climate change alone is raising regional temperatures. In three quarters (45 of 60) of cities examined, urbanized areas are warming faster than adjacent rural locations." See for example the plots below,

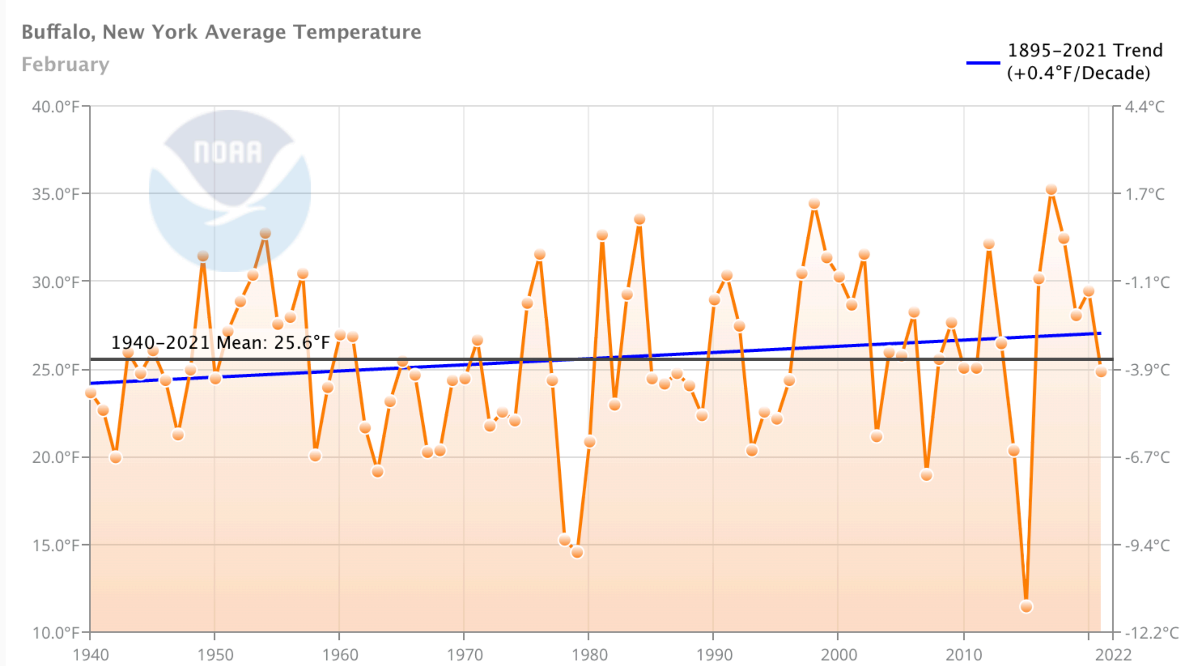
from NOAA, showing average temperatures in Los Angeles, CA, compared to Eureka, CA, or New York, NY, compared to Buffalo, NY. There is a nice tool at https://www.ncdc.noaa.gov/cag/city/time-series where you can input any city you choose and make the corresponding plot. Such local changes were also forecast accurately by scientists, and are nearly impossible to explain by random chance and are not at all in line with the pre-historic global climate changes referred to by the book's authors here. Fourth, there is no sign of comparable pre-historical time when levels of CO2 and other greenhouse gases were increasing at current levels and achieving current levels.


Los Angeles, California Average Temperature
February
1945-2022 Trend (+0.6°F/Decade)
1945-2022 Mean: 56.2°F
64.0°F
62.0°F
60.0°F
58.0°F
56.0°F
54.0°F
52.0°F
50.0°F
17.8°C
16.7°C
15.6°C
14.4°C
13.3°C
12.2°C
11.1°C
10.0°C
1940
1950
1960
1970
1980
1990
2000
2010
2022


Eureka, California Average Temperature
February
1942-2022 Trend (+0.2°F/Decade)
1942-2022 Mean: 47.6°F
54.0°F
52.0°F
50.0°F
48.0°F
46.0°F
44.0°F
42.0°F
12.2°C
11.1°C
10.0°C
8.9°C
7.8°C
6.7°C
5.6°C
1940
1950
1960
1970
1980
1990
2000
2010
2022



New York, New York Average Temperature
February
1895-2021 Trend
(+0.6°F/Decade)
1940-2021 Mean: 34.2°F
44.0°F
42.0°F
38.0°F
34.0°F
30.0°F
26.0°F
22.0°F
6.7°C
5.6°C
3.3°C
1.1°C
-1.1°C
-3.3°C
-5.6°C
1940
1950
1960
1970
1980
1990
2000
2010
2022



Buffalo, New York Average Temperature
February
1895-2021 Trend
(+0.4°F/Decade)
1940-2021 Mean: 25.6°F
40.0°F
35.0°F
30.0°F
25.0°F
20.0°F
15.0°F
10.0°F
4.4°C
1.7°C
-1.1°C
-3.9°C
-6.7°C
-9.4°C
-12.2°C
1940
1950
1960
1970
1980
1990
2000
2010
2022

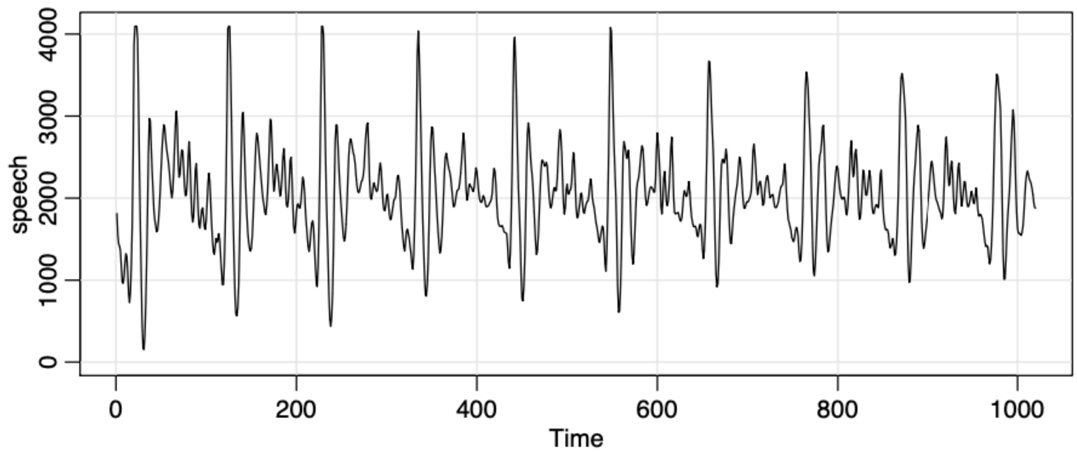


Fig. 1.3. *Speech recording of the syllable aaa ··· hhh sampled at 10,000 points per second with* $n = 1020$ *points.*

Speech.

Time series do not have to span huge lengths of time. An example spanning a very short time is the sample of 0.1 seconds of someone saying "aaahhh" in Figure 1.3. The regular, repeating cycle is readily apparent. We can use spectral analysis on this data to model this cyclic behavior, and remove it to see what is left, which might be indicative of the person's vocal patterns, and also potentially modify the cyclic part to change pitch or volume, or to analyze the cyclic part to have a computer potentially identify what is being said.

plot(speech) will reproduce the figure.

Finance.

An example of daily data is Figure 1.4 which shows the daily percentage changes from the previous day's value, of the Dow Jones Industrial Average (DJIA) from 2006 to 2016. The daily percent changes are called *returns*. The big spikes are from the large financial crisis of 2008. Although the returns seem to have mean zero, it also seems like there are certain time periods when the volatility is very high and other time periods where the volatility is very low. *ARCH* models, *GARCH* models, and *stochastic volatility* models can be used with data like this to try to forecast the volatility in the future.

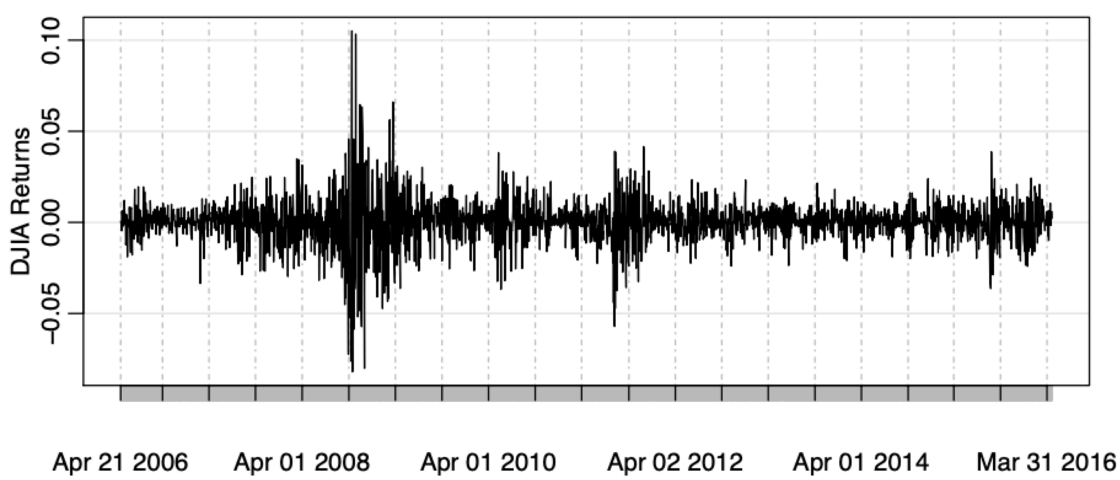


Fig. 1.4. *The daily returns of the Dow Jones Industrial Average (DJIA) from April 20, 2006 to April 20, 2016.*

Bivariate Weather and Fish data.

Sometimes, for each time, you observe two different variables that might be related, and this is called a bivariate time series. An example is the *Southern Oscillation Index* (SOI) and fish data shown in Figure 1.5. For each month out of the 453 months from 1950 to 1987, we see in the top panel the SOI, which measures changes in air pressure related to sea surface temperatures in the central Pacific Ocean. In the bottom panel labelled *recruitment*, for these same months, we see the number of new fish detected, according to a member of the Pacific Environmental Fisheries Group. We know that every three to seven years, the central Pacific gets warmer as a result of the El Niño effect. Both the SOI and the fish data seem to have repeating cycles. In the SOI data, there seem to be two obvious cycles, one seasonal 12-month cycle, and maybe also a 4-year or so cycle. Specral analysis can be useful here to try to pinpoint these cycles and see if they are statistically significant, and to see if the cycles in the two datasets are closely related.

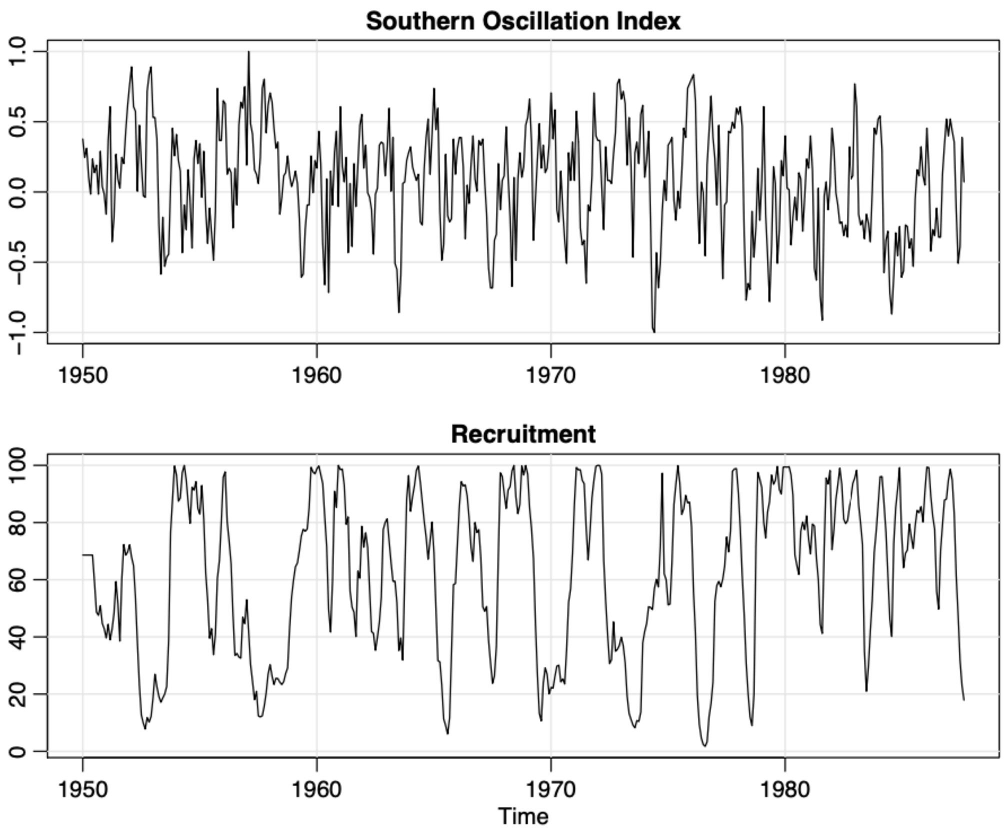


Fig. 1.5. *Monthly SOI and Recruitment (estimated new fish), 1950-1987.*