Class 1: Introduction to Time Series Models

Andrew Parnell
andrew.parnell@mu.ie

https://andrewcparnell.github.io/TSDA/

PRESS RECORD

Introductions

- Tell us who you are, what you are working on, and what you hope to get out of the week
- Timetable for the week
- A quick note about pre-requisites

How this course works

- This course lives on GitHub, at github.com/andrewcparnell/TSDA which means anyone can see the slides, code, etc, and make comments on it
- The timetable html document provides links to all the pdf slides, handouts, data and practicals
- Let me know if you spot mistakes, as these can be easily updated on the GitHub page
- There is an issues page if you find mistakes, or use Slack to ask questions

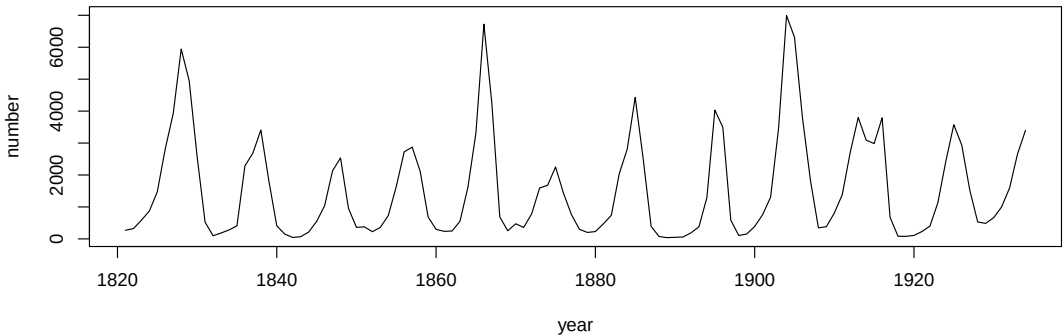
R code, slides, and practicals

- All the slides and practicals are available in pdf format for you to annotate
- In the background, the slides and the practicals are written in `Rmarkdown` format, which means you can load them up in Rstudio and see how everything was created
- When you have spare time, feel free to load up the `.Rmd` files and run the code in the background

R code in slides

- Many of the slides contain R code and output (some of which may be hidden in the `.Rmd` file)
- An example:

```
lynx = read.csv(file = '../../data/lynx.csv')
with(lynx, plot(year, number, type = 'l'))
```



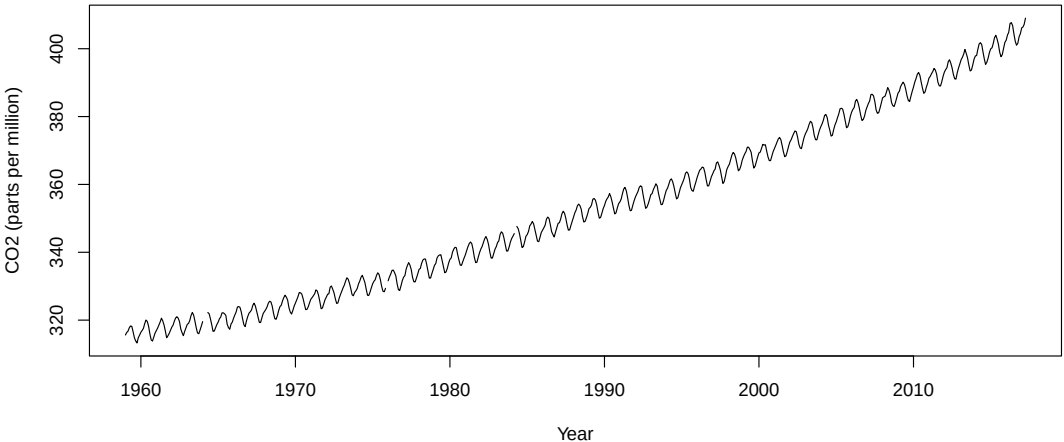
Course format and other details

- Lectures will take place in the morning, practical classes in the afternoon
- Please ask lots of questions
- Some good books:
  - *Forecasting: Principles and Practice* by Hyndman and Athanasopoulos
  - *Hierarchical Modeling and Inference in Ecology* by Royle and Dorazio
  - *Bayesian Methods for Ecology*. by McCarthy
  - *Bayesian Data Analysis* by Gelman et al
- Looking for data? Try the `tsdl` R package
- (see also sources in Practical 3)

What is a time series?

- A time series is any set of data where the response variable is measured over time
- There may be other variables included too (covariates)
- Time may be discrete (1, 2, 3, 4, ...) or continuous (1.7, 2.53, 7.12, ...)
- There may be missing values or outliers
- Occasionally there may be more than one response variable (multivariate time series)

A time series plot: CO2 data



General features of a time series

- Trend: long term behaviour. May be a straight line or something more complicated
- Seasonal: repeated behaviour. May be yearly, monthly, daily, etc. Likely to be dependent on the time resolution
- Error: Leftover uncertainty beyond the trend and seasonal behaviour. May have interesting statistical patterns.

Writing time series mathematically

- If we write y_t as the value of the response variable at time t then the series can be *decomposed* as:

$$y_t = \text{trend}_t + \text{seasonality}_t + \text{error}_t$$

- Most time series models concentrate on the error structure
- Time series analysis is usually harder if you need to identify the seasonality too

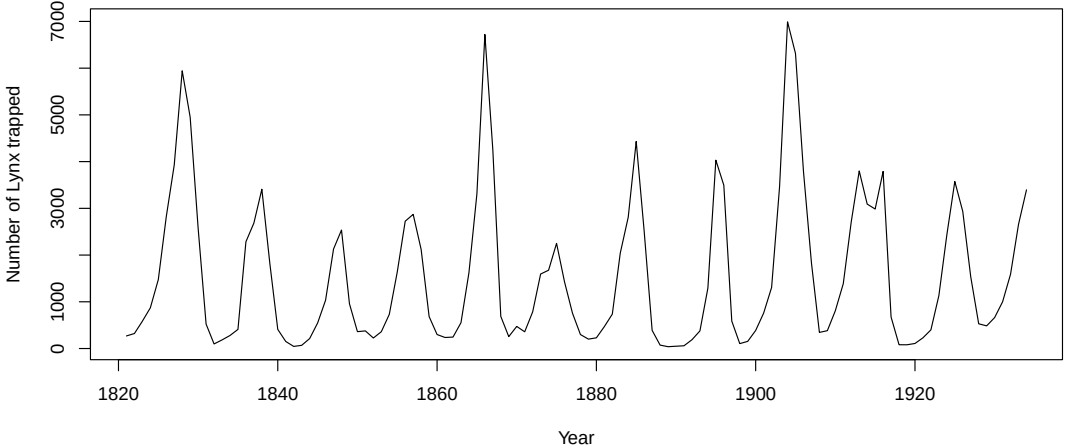
Participation exercise

- I want to you to look at the time series on the following slides and...
- ... identify the trend. Is it linear or non-linear?
- ... identify the seasonality (if any). Can you estimate the frequency?
- ... look at the residual errors after accounting (in your head) for trend and seasonality. Can you spot any patterns or strange observations?

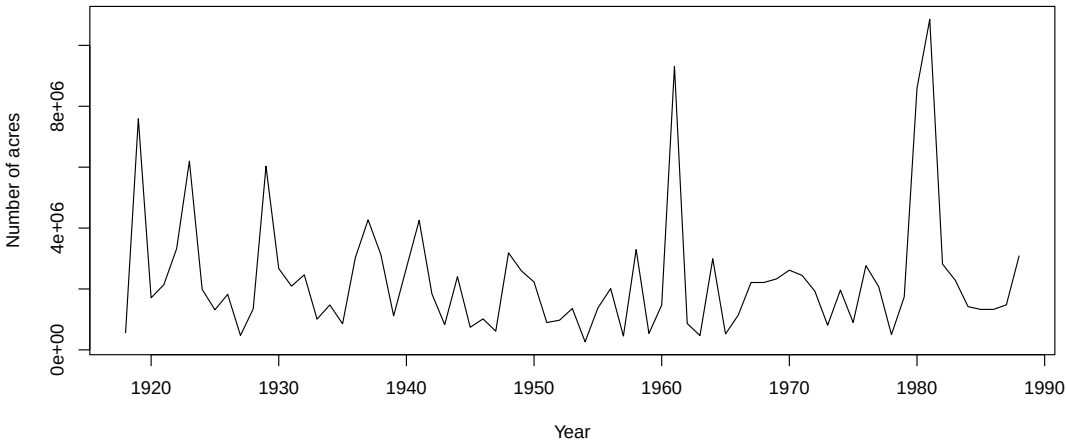
Data set 1: Wheat production in Canada

Data set 2: Sheep numbers in Asia

Data set 3: Lynx trappings in Canada

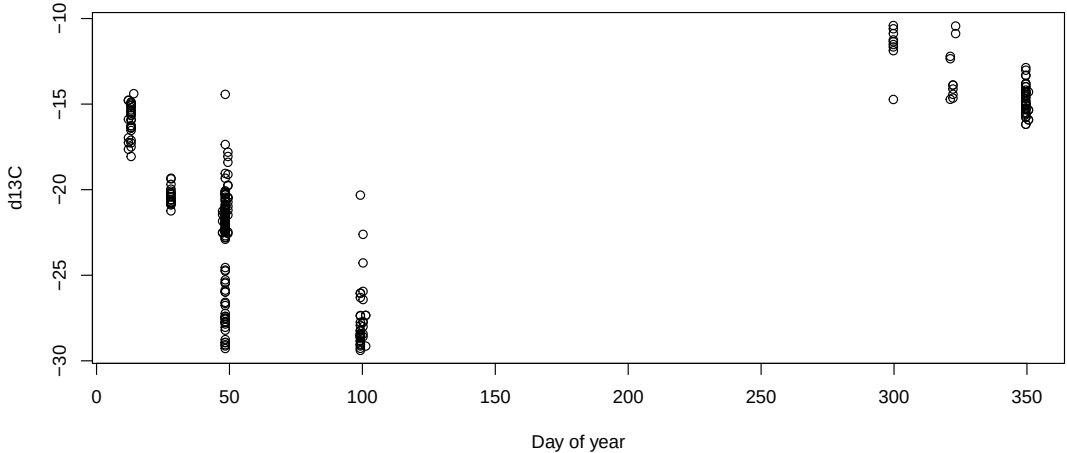


Data set 4: Forest fires in Canada

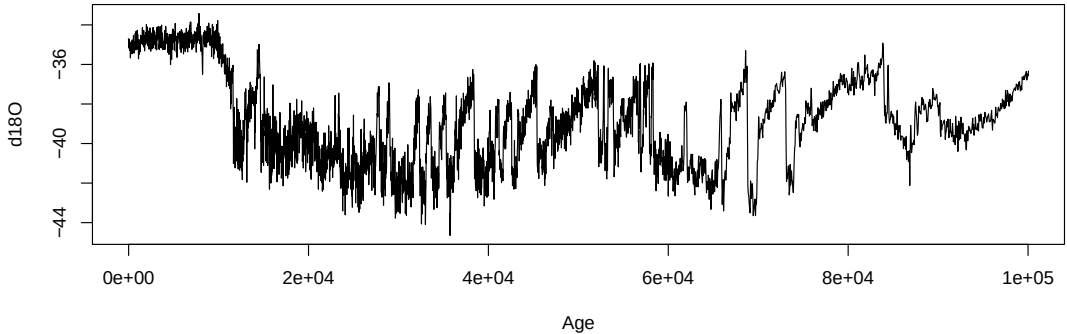


Data set 5: Geese isotopes

Data set 5: Geese isotopes again



Data set 6: Oxygen isotopes in ice




```
table(diff(ice$Age))
```

```
##
##   20   40   50   60  100  120
## 3035    2  779    2    1    1
```

Goals of time series analysis

- Predict future values of the response variable
- Interpolate or smooth the response variable for missing or non-measured times
- Explain which factors are causing the time series to change
- Understand the underlying behaviour of the time series

Summary

This course takes a practical approach, and should help you:

- Understand modern time series modelling techniques
- Get and use tools for thinking about and dealing with uncertainty
- Fit time series models, and make predictions
- Understand your time series data, and the process that generates it