School of Mathematical Sciences | Lancaster University

Beyond regression: time series and changepoints

Rebecca Killick(r.killick@lancs.ac.uk)
Rewilding, Edinburgh Oct 2025

Process vs Statistical

- Different underlying paradigms
- Practically can look very similar

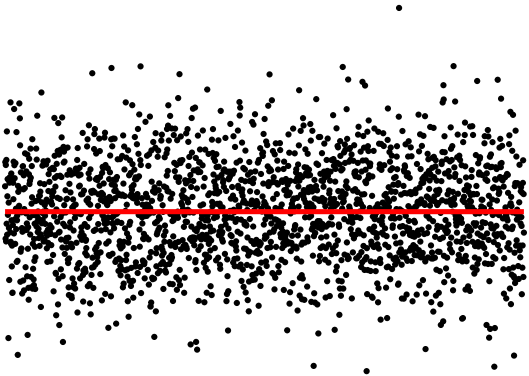
In my brain, a key difference is:

Statistical models require data to develop, process models require data to validate.

Outline

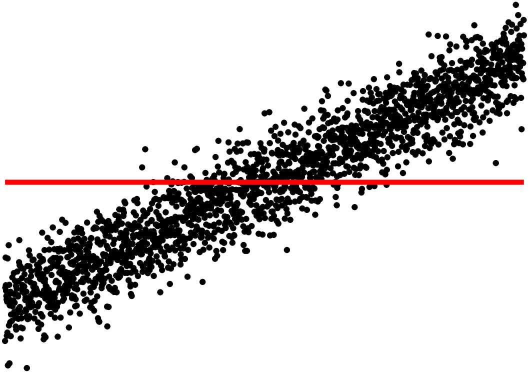
- Recap regression (linear models)
- Time series as regression
- Linear Mixed Models (Tom's talk)
- Changepoint modelling
- Monitoring for change

Model development

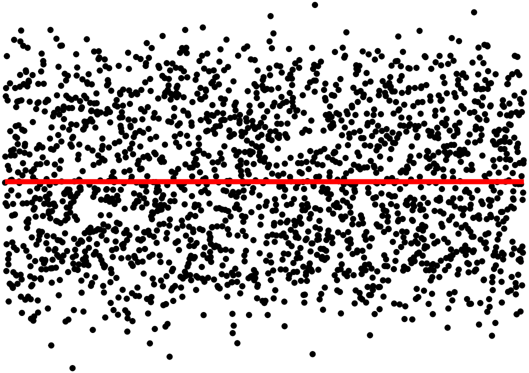


Model development

$y_t = \mathbb{E}(\mathbf{y}) = \mu$

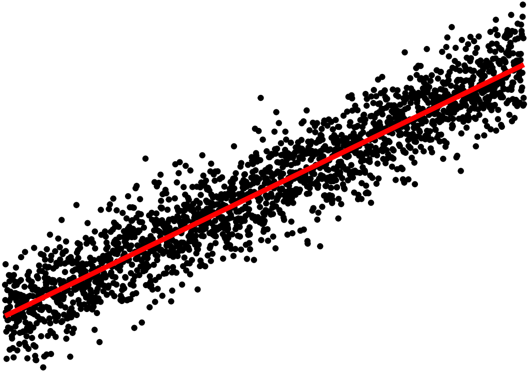


Model development



Linear model

$y_t = a + bt = \beta_1 + \beta_2 t$



Regression

$$\mathbf{y} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ \vdots & \vdots \\ 1 & n-1 \\ 1 & n \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \boldsymbol{\epsilon}$$

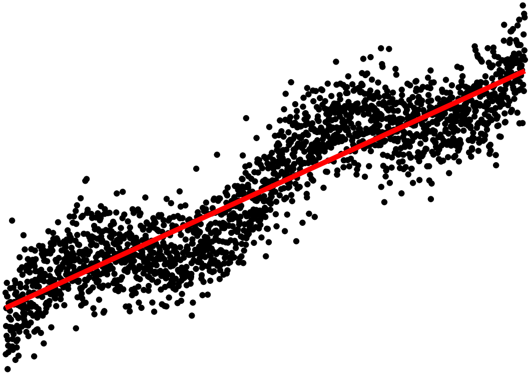
Often $\epsilon \sim \mathrm{N}(0, \sigma^2)$.

More generally

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}.$$

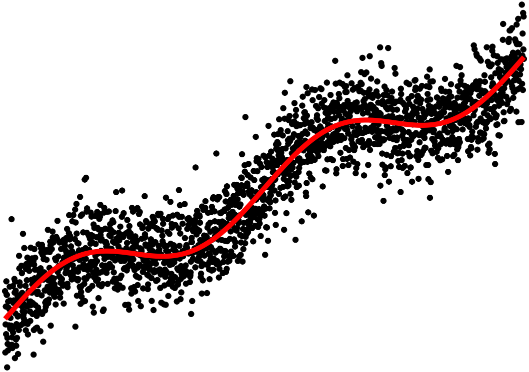
Fit using `lm` in `R` or `OLS` from `statsmodels` in `Python`.

Regression



Regression

$y_t = \beta_1 + \beta_2 t + \beta_3 ?$



Seasonality

$$\mathbf{y} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 1 & 3 & 0 & 1 & 0 \\ \vdots & \vdots & & & \\ 1 & n-1 & 0 & 1 & 0 \\ 1 & n & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{pmatrix} + \epsilon$$

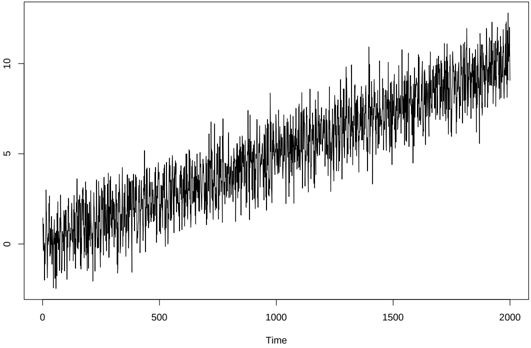
β_3 is the *change* from quarter 1 to quarter 2.

β_4 is the *change* from quarter 1 to quarter 3.

β_5 is the *change* from quarter 1 to quarter 4.

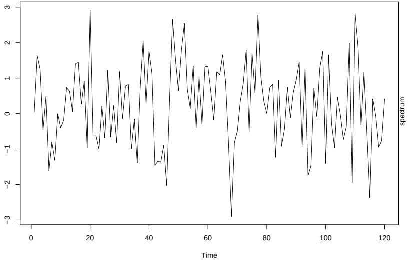
Seasonality

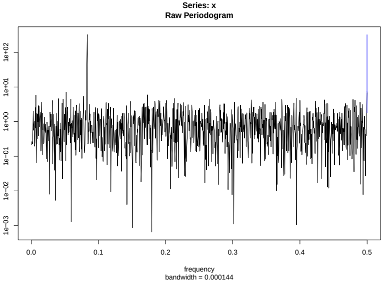
$$y_t = \beta_1 + \beta_2 t + \beta_3 \sin(\omega t) + \epsilon_t$$



Seasonality

$$y_t = \beta_1 + \beta_2 t + \beta_3 \sin(\omega t) + \epsilon_t$$





Mean Modelling

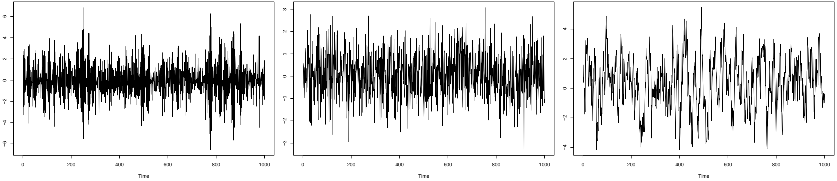
Many different things you can do here, depending on how you model.

- Traditional regression
- Time dependence through trend and seasonality
- Non-parametric modelling
- Spatial modelling

We have been thinking of this for continuous measurements but other types of measurements can use this model through GLM (Generalized Linear Model).

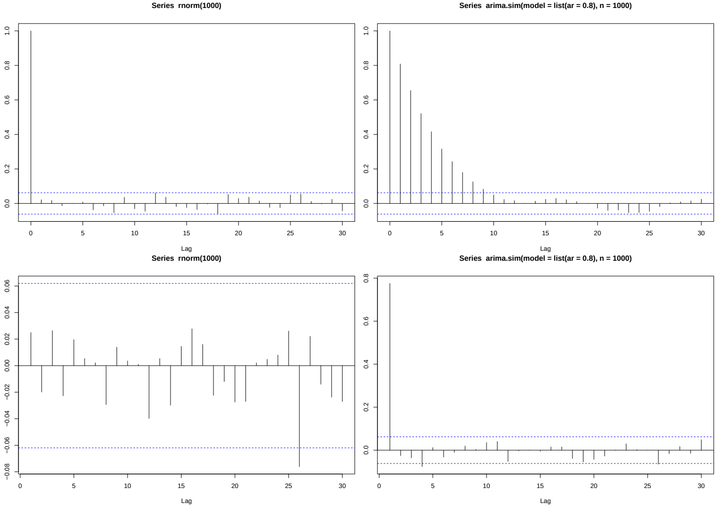
Second orders

$$\Gamma(t,s) = \text{Cov}(Y_t, Y_s) = \mathbb{E}[(Y_t - \mathbb{E}(Y_t))(Y_s - \mathbb{E}(Y_s))]$$



Check Second Order

Weighted Portmanteau Test (`WeightedPortTest` in R)



GLMER

Small amount of time points? Use a Generalized Linear Mixed Effects Regression (GLMER) model (`lme4` in `R`).

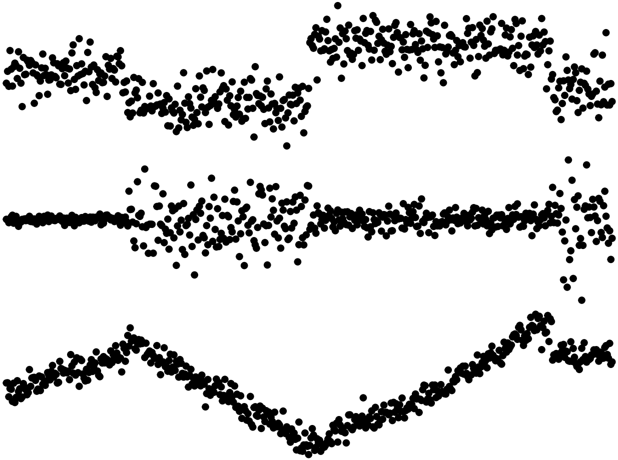
$$y_i|b \sim \mathrm{Distr}\left(\mu_i, \sigma^2\right) \tag{1}$$

$$g(\mu) = X\beta + Zb, \tag{2}$$

- X - external regressors with coefficients β
- Distr - conditional distribution
- Z - random effects with coefficients b

Non-stationarity

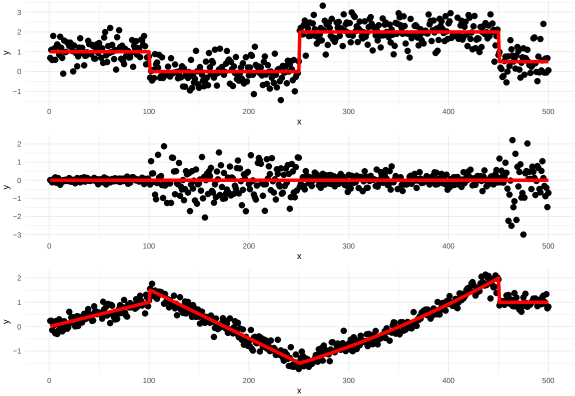
Model development



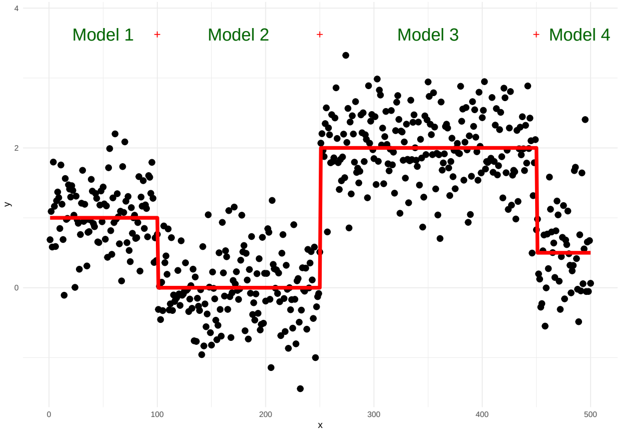
What are changepoints?

For data $y_1, \ldots, y_n$, if a changepoint exists at τ, then $y_1, \ldots, y_\tau$ differ from $y_{\tau+1}, \ldots, y_n$ in some way.

There are many different types of change.

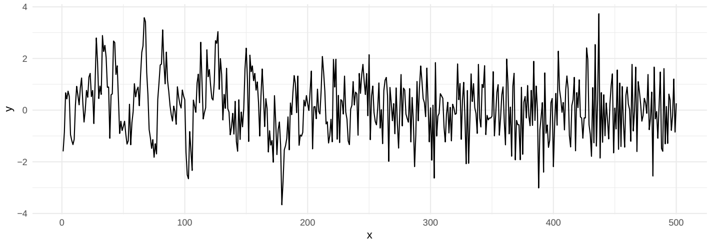


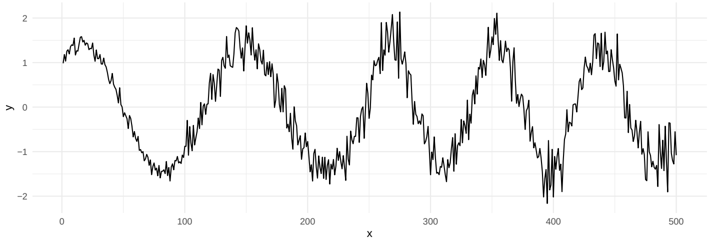
Changepoint Problem



- How many changes? (`changepoint` package in `R`)
- Where are the changes? 2^{n-1} possible solutions!

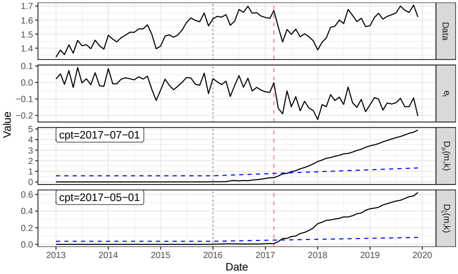
More complicated change

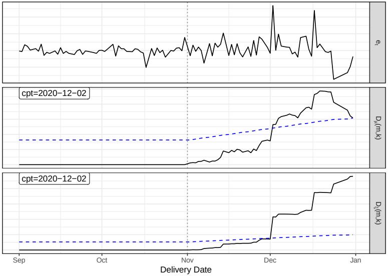




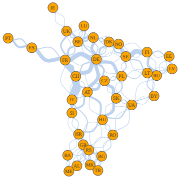
Concept drift

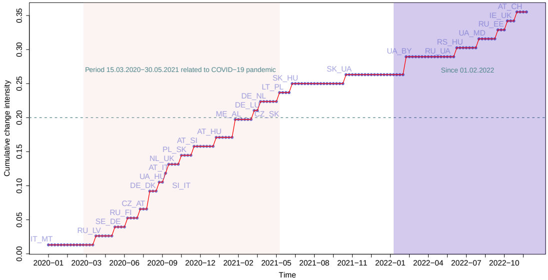
Monitor models for concept drift - works with any type of model online fit/forecast





Concept drift in time series on networks





Summary

- Took a modelling approach to introducing time series
- We learnt about mixed effects models, traditional ARMA models, and changepoint models
- Model diagnostics and visually (even parts of) data are important.
- Many more models out there; spatiotemporal, threshold, functional, locally stationary, . . .
- I enjoy working with messy real life data . . .
- . . . as often it sparks my next research challenge.

Code

R packages: lme4, forecast, changepoint, WeightedPortTest

Links or stars from: https://www.github.com/rkillick

Online monitoring: https://github.com/grundy95/changepoint.forecast

Multivariate:

Covariance: https://github.com/s-ryan1/Covariance_RMT_simulations

Mean/Var: https://github.com/grundy95/changepoint.geo