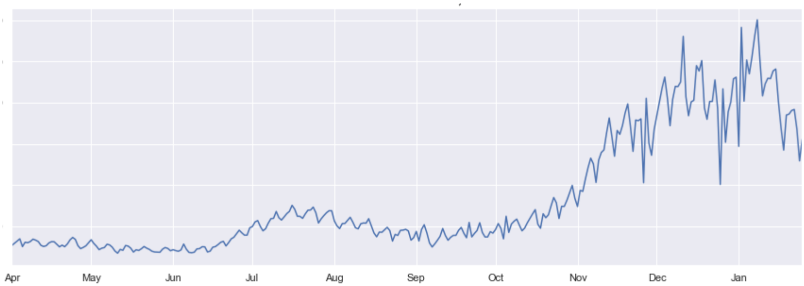
Identifying sequential changes in mean and variance within more complex model structures

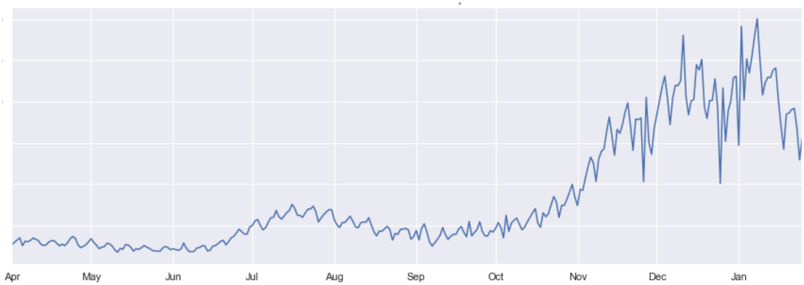
Rebecca Killick (r.killick@lancs.ac.uk)
Joint work with: Thomas Grundy, Ivan Svetunkov (Lancaster University)
and Zhao Liu, Shan Islam and Michael Taylor (Royal Mail)
CMAF Oct 2021

Motivation



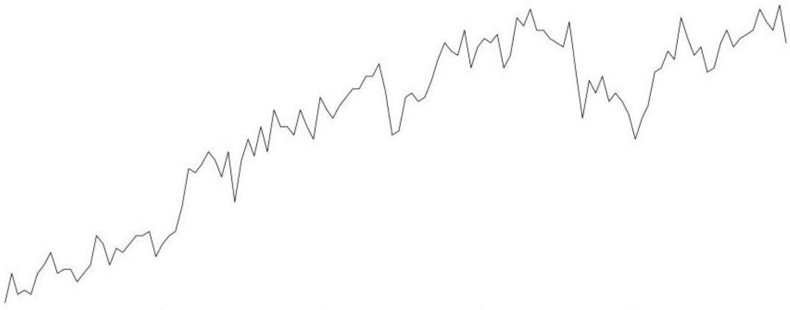
Motivation

Data Science | Lancaster University



- Multiple seasonal frequencies
- Regressors for holiday periods
- Explanatory series
- Lagged dependencies

Motivation



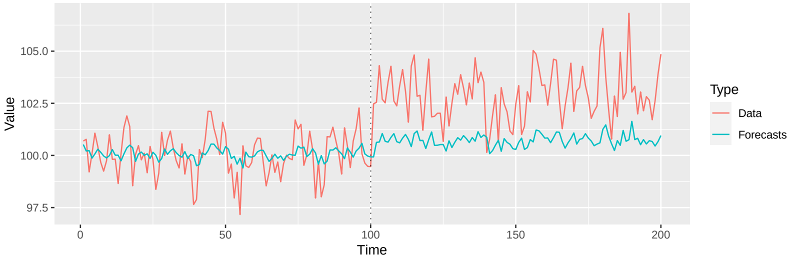
Problems

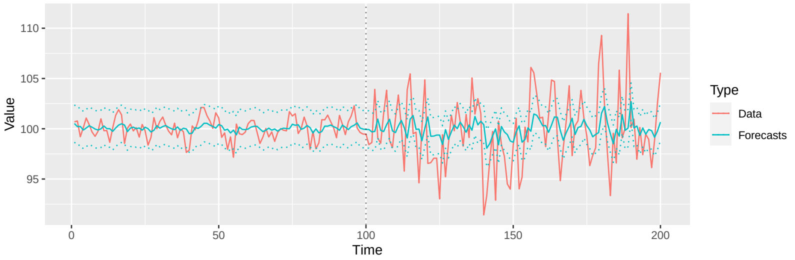
Automatically identifying changes in data requiring complex modelling is problematic:

- Most methods need to be able to fit the model after the change ...
- ... leading to delays in online detection
- ... poor model estimation in short segments
- ... identifiability issues with changepoints close together
- ... inability to fit long seasonal frequencies (e.g. yearly)

- Those that don't are often restricted to mean/linear trend series only.

Allied problem





Causes of Poor Forecasts

A change in the underlying process (data) could cause our forecasting model to behave poorly!

Intuitively:

- Mean change in raw data $\rightarrow$ forecasts become biased.
- Variance change in raw data $\rightarrow$ prediction intervals become poor.
- Other changes in raw data $\rightarrow$ potential combination of biased and inaccurate prediction intervals.

Our Aim

Aim: Create a framework for detecting changes in complex models which doesn't require fitting the model to the post-change data.

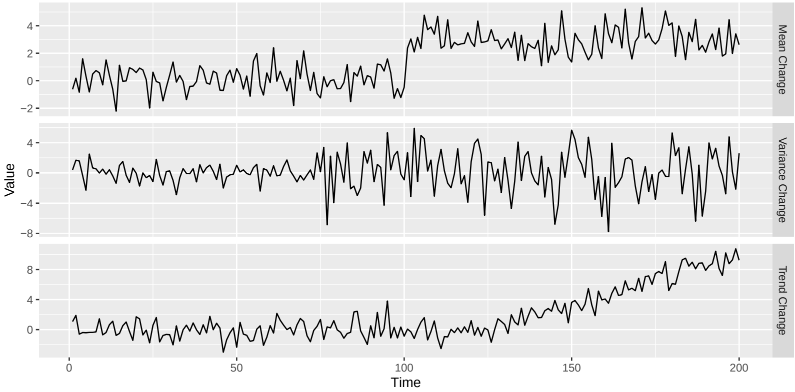
Intuitively, changes in the raw data = forecasts become poor.

Solution: Use sequential forecast errors.

Bonus: This also provides a framework for identifying changes in forecast performance.

Changepoints

A changepoint is an abrupt structural change in the data, e.g.



Online Changepoints

Aim to detect change in ordered data as new data becomes available.

- Balance two performance measures:
  1. Detect changes as quickly as possible.
  2. Minimise false alarms.
- Algorithms are split into 2 phases:
  1. Training period - data is assumed stable/stationary and we estimate model parameters (e.g. the mean).
  2. Monitoring period - new data is arrives and we compare these to the training data to see if statistically different (e.g. a different mean).

Detecting Changes

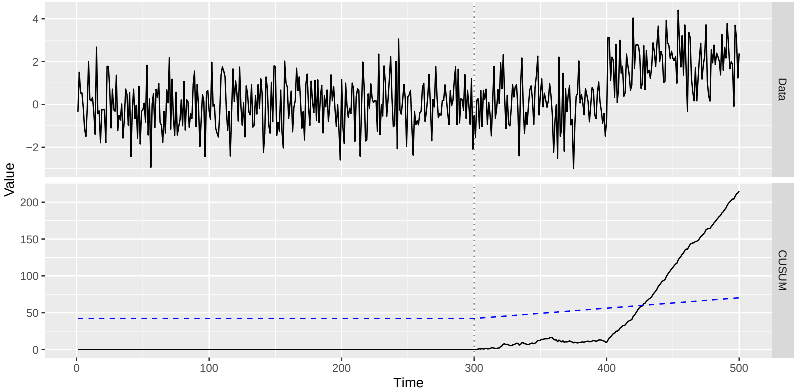
Detector:

- Measures the difference between mean in training period and mean in monitoring period.
- CUSUM Detector: $Q(m,k) = \sum_{t=m+1}^{k} y_t - \frac{k}{m}\sum_{t=1}^{m} y_t$.
- Page's CUSUM: $D(m,k) = \max_{0\le t\le k} |Q(m,k) - Q(m,t)|$

Stopping Rule:

- Defines a rule for when to flag a change.
- Typically when the detector exceeds some threshold.
- Threshold controls false positive rate.
- $\tau = \min\{k \ge 1 : D(m,k) \ge c\hat{\sigma}_m g(m,k)\}$.

Example



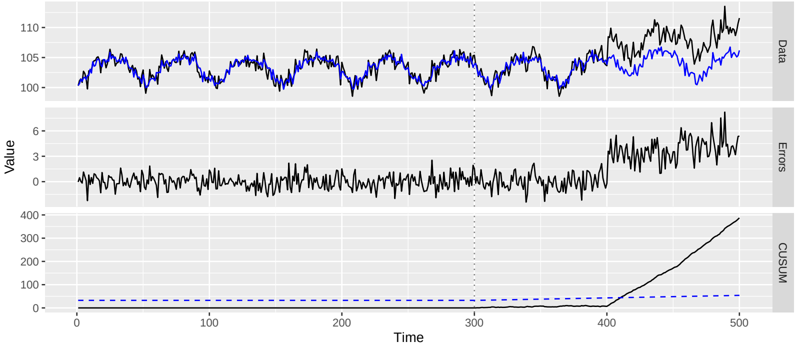
Value
Data
CUSUM
Time
4
2
0
−2
200
150
100
50
0
0
100
200
300
400
500

Solution

What about the forecast errors?

$$e_t = Y_t - \hat{y}_t(1) \;, \qquad Q(m,k) = \sum_{t=m+1}^{k} e_t - \frac{k}{m}\sum_{t=1}^{m} e_t$$

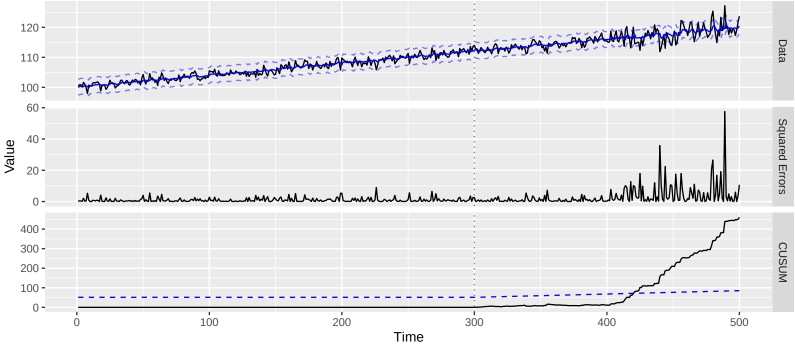
- Forecasting model accounts for data complexities.



Variance Changes

Variance change in errors = Mean change in squared errors.

$$Q(m,k) = \sum_{t=m+1}^{k} (e_t - \mu_m)^2 - \frac{k}{m}\sum_{t=1}^{m}(e_t - \mu_m)^2 \ .$$



Theory & Simulations

Theory: Under certain assumptions,

- Mean change in raw data $\rightarrow$ mean change in forecast errors.
- Mean and/or variance change in raw data $\rightarrow$ mean change in squared forecast errors.

Simulations:

- Using forecast errors performs better than raw CUSUM.
- Can also be used to detect changes in,
  - Trend.
  - Dependence structure.
  - Error structure.

Which detector is best?

CUSUM of Raw Forecast Errors:

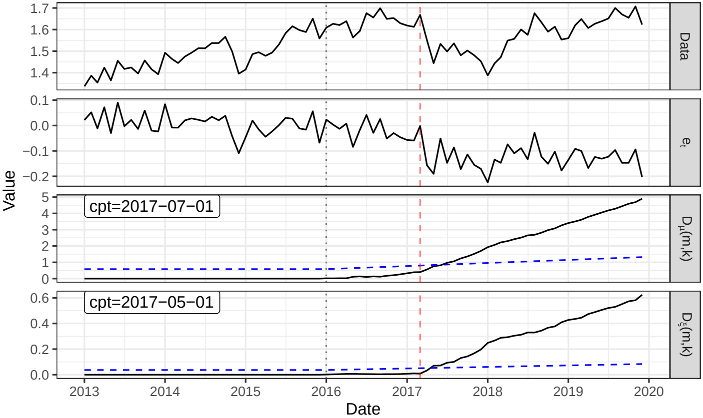
- First-order changes **ONLY** e.g.,
  - Mean changes.
  - Trend changes

CUSUM of Squared Forecast Errors:

- Potential second-order changes,
  - Variance changes.
  - Mean and Variance changes.
  - Dependence structure changes.

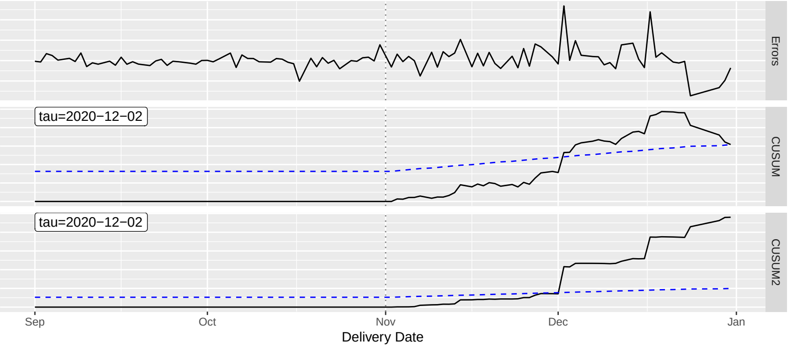
A&E Admissions

- A&E Admissions with Gallstone related pathology
- Predicting admissions helps to manage workload
- Elective surgeries can be scheduled for quieter periods

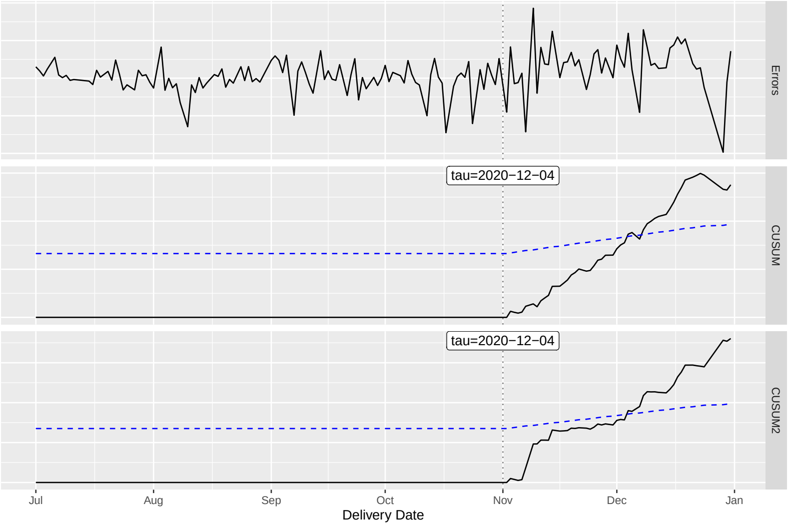


TTP Project

- Forecast the no. of parcels being delivered from each delivery office.
- In each year and delivery office the Xmas rush starts at different times.
- We can use this framework to detect when the Xmas rush begins.



TTP Project



Errors
CUSUM
CUSUM2
tau=2020−12−04
tau=2020−12−04
Jul
Aug
Sep
Oct
Nov
Dec
Jan
Delivery Date

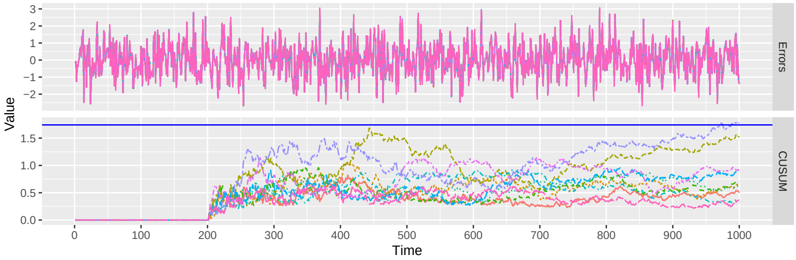
Threshold Choice

Data Science | Lancaster University

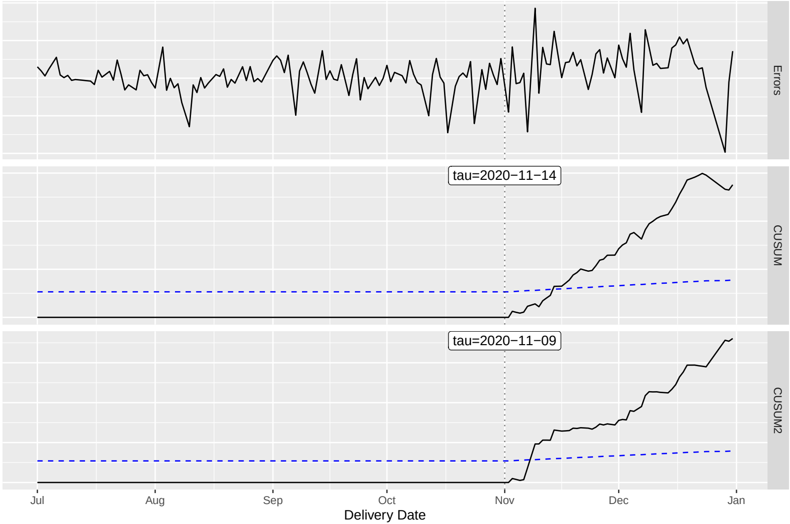
In many scenarios, the theoretical threshold is too conservative.

Potential Solution:

1. Simulate from known forecasting model.
2. Run algorithm on simulated null data.
3. Take the 95% quantile of the max CUSUM statistics.



TTP Project



Errors
CUSUM
CUSUM2
tau=2020−11−14
tau=2020−11−09
Jul
Aug
Sep
Oct
Nov
Dec
Jan
Delivery Date

Conclusion

Data Science | Lancaster University

Summary:

- Created a framework for detecting when/if forecasts become poor.
- Shown using forecast errors is better than the raw data for detecting changes.

Future work:

- Extension to multivariate forecasts.
- Consider alternative online changepoint methods.