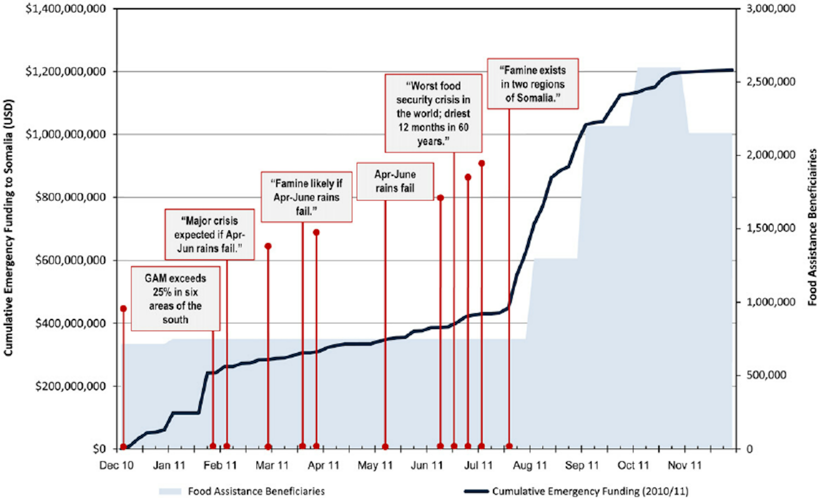
Pre-emptive disaster action using Bayesian online change point detection

Dan, Gianluca and Paolo

University of Bath ITT9

1st February 2019

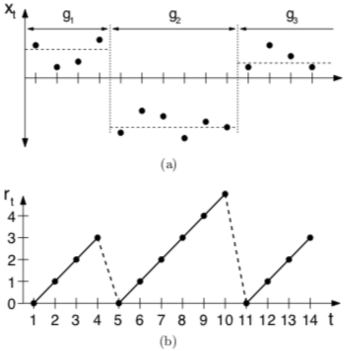
Why identifying a change point is relevant



Source : Hillbruner and Moloney [1]

Method assumptions and notation

- $r_t =$ "run starting at t" $=$ number of steps at time t since previous change point.



Source : Adams and MacKay [2]

- $x_t \sim p(x_t \mid \boldsymbol{\eta}_\rho)$, $\boldsymbol{\eta}_\rho$ i.i.d. for $\rho = 1, 2, \ldots$. Data points conditionally independent in different partitions.
- 'Change point detection is the identification of abrupt changes in the generative parameters of sequential data.' (Adams and MacKay [2])

Detecting a change point : marginal predictive distribution

- Time t is a change point $\Rightarrow$ significant change in series behaviour $\Rightarrow$ x_t is extremely unlikely under the (now outdated) model.
- Credibility of unseen x_{t+1} can be assessed by how well it fits with previous data $x_{1:t}$.
- Marginal predictive distribution :

$$p(x_{t+1} \mid \boldsymbol{x}_{1:t}) = \sum_{r_t=0}^{t} \underbrace{p(x_{t+1} \mid r_t, \boldsymbol{x}_t^{(r)})}_{\textit{Predictive}} \underbrace{p(r_t \mid \boldsymbol{x}_{1:t})}_{\textit{Posterior}}, \quad (1)$$

where $\boldsymbol{x}_t^{(r)}$ is data on run r_t.

Detecting a change point : predictive distribution

Assume

$$x_{t+1} = \mathcal{F}_{\boldsymbol{x}^{(r)}}(\boldsymbol{\eta}) + \xi \qquad \xi \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}).$$

for a chosen $\mathcal{F}$, e.g. neural network.

Then

$$p(x_{t+1} \mid r_t, \boldsymbol{x}_t^{(r)}) = \int p(x_{t+1} \mid r_t, \boldsymbol{x}_t^{(r)}, \boldsymbol{\eta}) \underbrace{p(\boldsymbol{\eta} \mid r_t, \boldsymbol{x}_t^{(r)})}_{SMC, MCMC} d\boldsymbol{\eta}.$$

Can also use conjugate prior.

Detecting a change point : posterior distribution

For the posterior, construct the iterative formula

$$p(r_t, \boldsymbol{x}_{1:t}) = \sum_{r_{t-1}=0}^{t-1} p(r_t \mid r_{t-1}) p(x_t \mid r_{t-1}, \boldsymbol{x}_t^{(r)}) p(r_{t-1}, \boldsymbol{x}_{1:t-1})$$

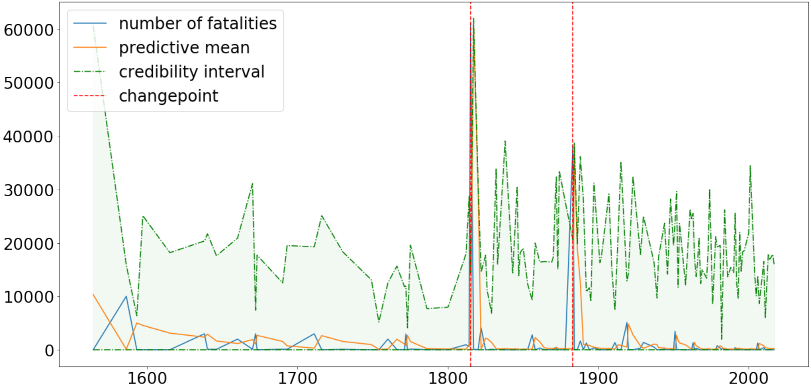
where

$$p(r_t \mid r_{t-1}) = \begin{cases} H & \text{if } r_t = 0 \\ 1 - H & \text{if } r_t = r_{t-1} + 1 \\ 0 & \text{otherwise} \end{cases}$$

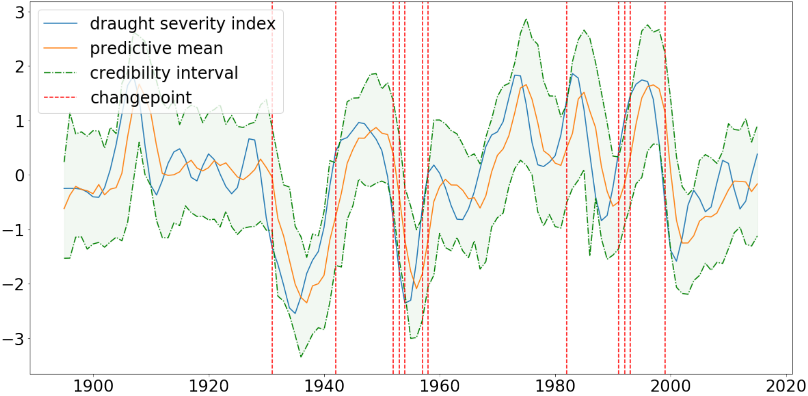
for known H. Then

$$p(r_t \mid \boldsymbol{x}_{1:t}) = \frac{p(r_t, \boldsymbol{x}_{1:t})}{p(\boldsymbol{x}_{1:t})} = \frac{p(r_t, \boldsymbol{x}_{1:t})}{\sum_{r_t=0}^{t} p(r_t, \boldsymbol{x}_{1:t})}.$$

Results : 1815 eruption of Mount Tambora



Results : Palmer Drought Severity Index USA



Source : NOAA (National Oceanic and Atmospheric Administration)

References

[1] Hillbruner, Chris and Moloney, Grainne.
When early warning is not enough—Lessons learned from the 2011 Somalia Famine.
2012

[2] Adams, Ryan Prescott and MacKay, David JC.
Bayesian online changepoint detection.
2007