Joint Spectral-Temporal Analysis of High-Energy Astronomical Sources

Raymond K. W. Wong

Department of Statistics
University of California at Davis

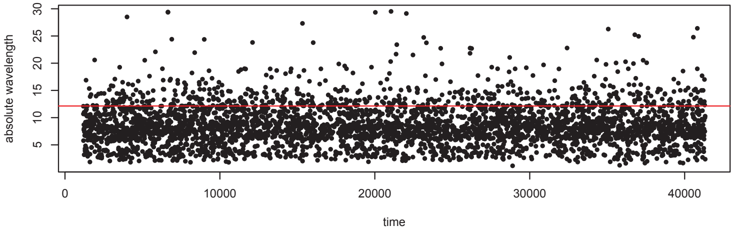
Vinay Kashyap, David A. van Dyk and Thomas C. M. Lee

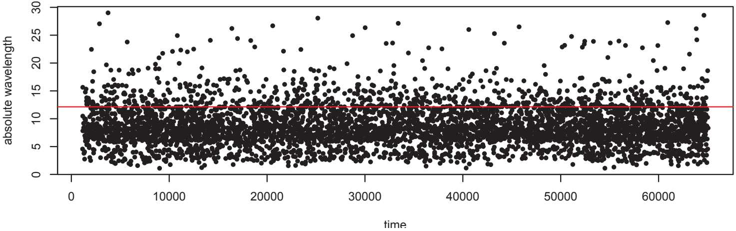
November 26, 2013

the problem

- we deal with high-spectral/high-temporal resolution grating data
- these data are obtained as lists of photons
- for each photon we know
  1. the time at which it was recorded (t)
  2. its wavelength (w) and hence its energy (E)
- one interesting question: does the distribution of energy change over time?

two typical data sets

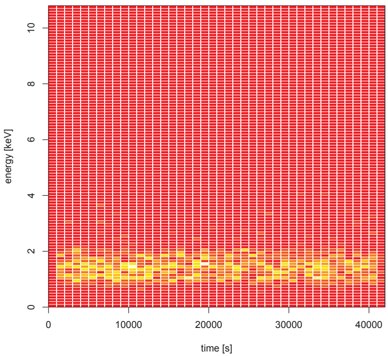


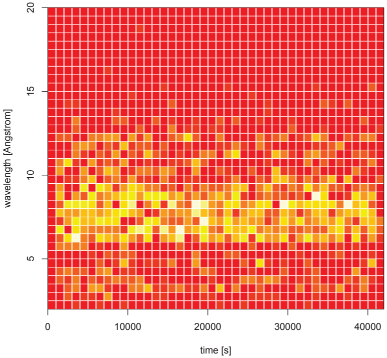


preprocessing: binning the data

- as a first step, we "bin" the data
- i.e., lay a grid over the data and count how many points in each grid box
- Poisson counts in each grid box (or bin)
- size of grid/bin: needs to be carefully chosen

binned data sets





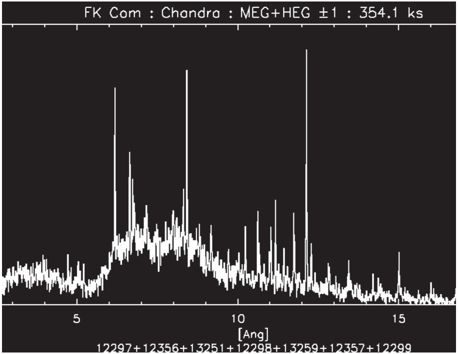
poisson modeling

- each bin is indexed by two quantities:
  1. t: time
  2. w: wavelength
- denote the observed counts as $C(t, w)$
- denote the brightness of a source as $\mu(t, w)$ (expected counts per unit area)
- $$C(t, w) \sim \text{Poisson}\left\{\delta t \times \delta w \times \mu(t, w) \sum_{k=1}^{K} A_k(w)\right\}$$

  K: number of detectors; $A_k(w)$: effective area for the kth detector (all known)

about $\mu(t, w)$

- no simple parametric models, so do nonparametric
- which typically requires smoothness assumption
- with *emission lines*, $\mu(t, w)$ is not completely smooth



modeling of $\mu(t, w)$

- for now assume $\mu(t, w)$ is the same for all t
- i.e., homogeneous across time
- and model the energy spectrum $\mu(w)$
- split $\mu(w)$ into two parts: smooth part + emission lines
    1. smooth part: radial basis expansion
       (use polynomial of power 3: $1, x, x^2, x^3, |x - \text{"knots"}|^3$)
    2. emission lines: delta functions

model for $\mu(w)$

- $$g(\mu(w)) = \sum_{j=1}^{P} \beta_j b_j(w) + \sum_{i=1}^{n} \eta_i I_i(w)$$
- g: link function as in GLM/GAM, for Poisson data
- P: number of basis functions, pre-specified
- b_j: the jth basis (radial basis)
- n: number of bins in the w-direction
- I_i: delta function
- β_j's and η_i's: parameters to be estimated
- note: number of parameters > number of observations

parameter estimation

- $$g(\mu(w)) = \sum_{j=1}^{P} \beta_j b_j(w) + \sum_{i=1}^{n} \eta_i I_i(w)$$
- need to set some β_j's and η_i's to zero
- do L_1 penalty (lasso)
- given tuning parameters γ and ρ, estimate $\boldsymbol{\beta}$ and $\boldsymbol{\eta}$ by minimizing

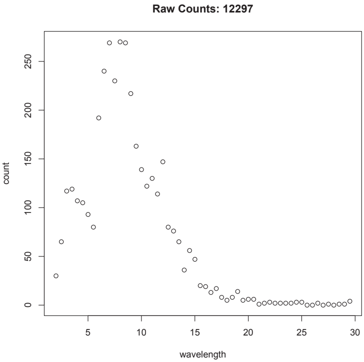
$$-\text{log likelihood} + \gamma\left\{\rho|\boldsymbol{\beta}|_1 + (1-\rho)|\boldsymbol{\eta}|_1\right\}$$

- fast algorithms exist

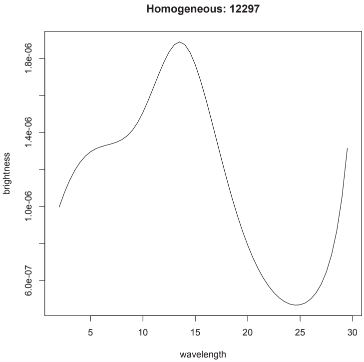
selecting the tuning parameters

- need to choose γ and ρ
- in classical lasso, they can be chosen say by cross-validation, AIC or BIC
- cross-validation: too slow
- AIC/BIC: cannot be blindly used here, as "$p > n$"
- see Chen and Chen (2008, Biometrika), where an Extended BIC criterion is proposed to handle the "$p > n$" issue
- we follow the idea and developed an Extended MDL criterion (Minimum Description Length)

an example (without emission lines)



raw counts



fitted

incorporating time t in modeling

- the energy spectrum typically changes over time
- as a first step, we do piecewise modeling of $\mu(t, w)$
- i.e., $\mu(t, w)$ is the same between any two breakpoints:

$$\begin{aligned}\mu(t,w) = \quad & \mu_1(w) I_{\{t_0 \le t < t_1\}} \\ + \quad & \mu_2(w) I_{\{t_1 \le t < t_2\}} \\ + \quad & \ldots \\ + \quad & \mu_B(w) I_{\{t_{B-1} \le t < t_B\}}\end{aligned}$$

- the number of breakpoints B, and the locations of the breakpoints t_j's, are unknown

selecting B and t_j's

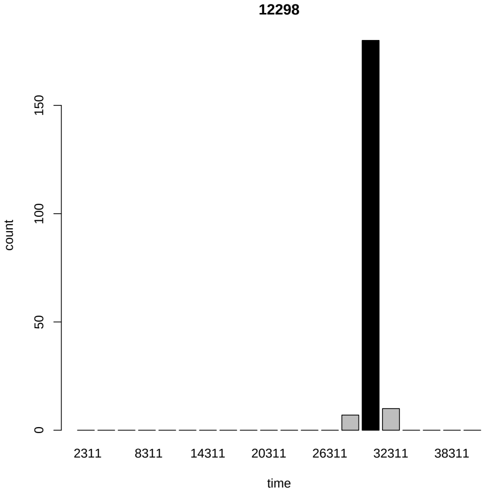
- it is a model selection problem
- MDL has been proven to be very successful in various structural break detection problems
- again, we are in the "$p > n$" scenario
- so direct application of classical MDL won't work here
- as before, we developed an Extended MDL criterion for choosing the final model
- (essentially a penalized likelihood, with 4 penalty terms)

practical fitting

- involves a non-trivial minimization problem
- a possibility is genetic algorithms
- but slow
- we use a “tree growing” strategy
- i.e., at each time step, choose the best location for adding one breakpoint, repeat until a local minimum is found
- (we could certainly do “tree pruning”, and more)

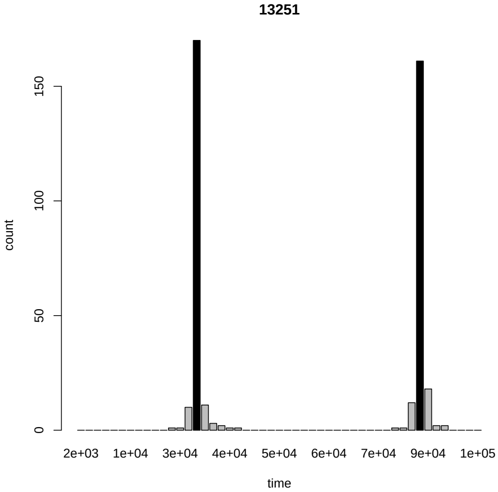
break detection simulation 1

1 true break, with 200 repetitions

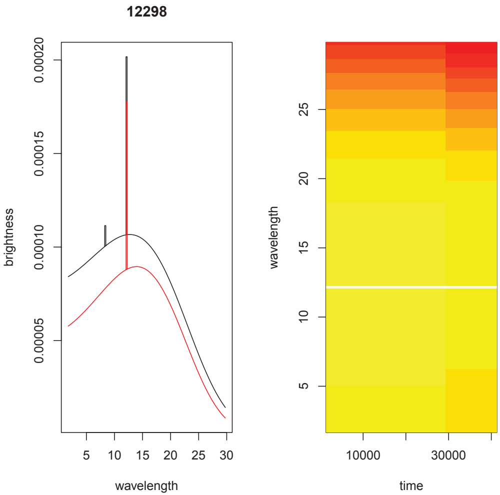


break detection simulation 2

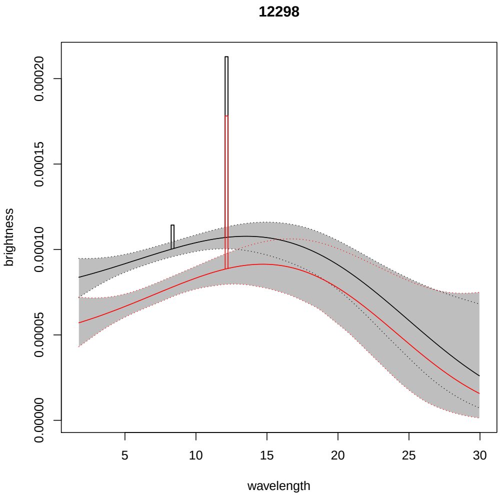
2 true breaks, with 200 repetitions



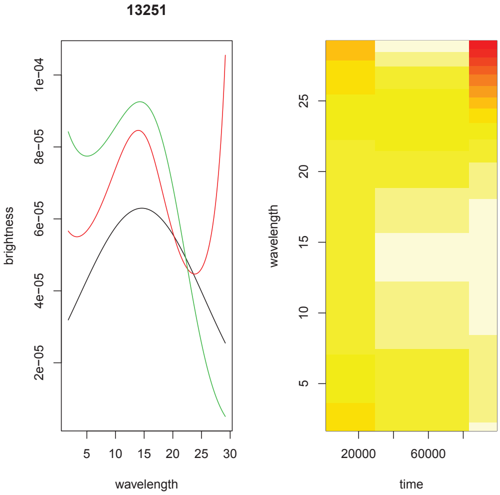
real data set 1



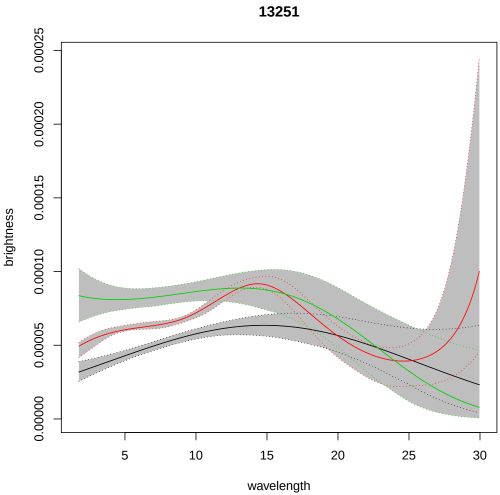
real data set 1



real data set 2



real data set 2



concluding remarks

- presented a method for detecting changes of energy spectrum over time
- modern regression techniques and new model selection criteria are used
- future work:
  1. better modeling in t
  2. theoretical properties of Extended MDL

The end.

Thank you.