Break/Change Point Detection - Automark

Thomas C. M. Lee

Department of Statistics
University of California, Davis

Joint work with Hans M. Gunther, Vinay L. Kashyap, David van Dyk, Raymond K. W. Wong, Cong Xu and Andreas Zezas

August 3, 2022

First paper: original version of Automark

The Annals of Applied Statistics
2016, Vol. 10, No. 2, 1107–1134
DOI: 10.1214/16-AOAS933
© Institute of Mathematical Statistics, 2016

DETECTING ABRUPT CHANGES IN THE SPECTRA OF HIGH-ENERGY ASTROPHYSICAL SOURCES[1]

BY RAYMOND K. W. WONG[*], VINAY L. KASHYAP[†,2], THOMAS C. M. LEE[‡,3] AND DAVID A. VAN DYK[§,4]

Iowa State University[*], Harvard Smithsonian Center for Astrophysics[†], University of California, Davis[‡] and Imperial College London[§]

Variable-intensity astronomical sources are the result of complex and often extreme physical processes. Abrupt changes in source intensity are typically accompanied by equally sudden spectral shifts, that is, sudden changes in the wavelength distribution of the emission. This article develops a method for modeling photon counts collected from observation of such sources. We embed change points into a marked Poisson process, where photon wavelengths are regarded as marks and both the Poisson intensity parameter and the distribution of the marks are allowed to change. To the best of our knowledge, this is the first effort to embed change points into a *marked* Poisson process. Between the change points, the spectrum is modeled nonparamet-

Second paper: 4D version of Automark

The Astronomical Journal, 161:184 (17pp), 2021 April https://doi.org/10.3847/1538-3881/abe0b6
© 2021. The American Astronomical Society. All rights reserved.

Change-point Detection and Image Segmentation for Time Series of Astrophysical Images

Cong Xu[1], Hans Moritz Günther[2], Vinay L. Kashyap[3], Thomas C. M. Lee[1], and Andreas Zezas[4]
[1] Department of Statistics, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA; cngxu@ucdavis.edu
[2] MIT, Kavli Institute for Astrophysics and Space Research, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[3] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[4] Department of Physics, University of Crete, 710 03 Heraklion, Crete, Greece
Received 2020 December 25; revised 2021 January 25; accepted 2021 January 26; published 2021 March 15

Abstract

Many astrophysical phenomena are time-varying, in the sense that their intensity, energy spectrum, and/or the spatial distribution of the emission suddenly change. This paper develops a method for modeling a time series of images. Under the assumption that the arrival times of the photons follow a Poisson process, the data are binned into 4D grids of voxels (time, energy band, and *x*-*y* coordinates), and viewed as a time series of non-homogeneous Poisson images. The method assumes that at each time point, the corresponding multiband image stack is an unknown 3D piecewise constant function including Poisson noise. It also assumes that all image stacks between any two adjacent change points (in time domain) share the same unknown piecewise constant function. The proposed method is designed to estimate the number and the locations of all of the change points (in time domain), as well as all of the unknown piecewise constant functions between any pairs of the change points. The method applies the minimum description length principle to perform this task. A practical algorithm is also developed to solve the corresponding complicated optimization problem. Simulation experiments and applications to real data sets show that the proposed method enjoys very promising empirical properties. Applications to two real data sets, the XMM observation of a flaring star and an emerging solar coronal loop, illustrate the usage of the proposed method and the scientific insight gained from it.

The problem in the first paper

The problem in the first paper

- we deal with high-spectral/high-temporal resolution grating data

The problem in the first paper

- we deal with high-spectral/high-temporal resolution grating data
- these data are obtained as lists of photons

The problem in the first paper

- we deal with high-spectral/high-temporal resolution grating data
- these data are obtained as lists of photons
- for each photon we know

The problem in the first paper

- 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)

The problem in the first paper

- 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 and hence its energy (E)

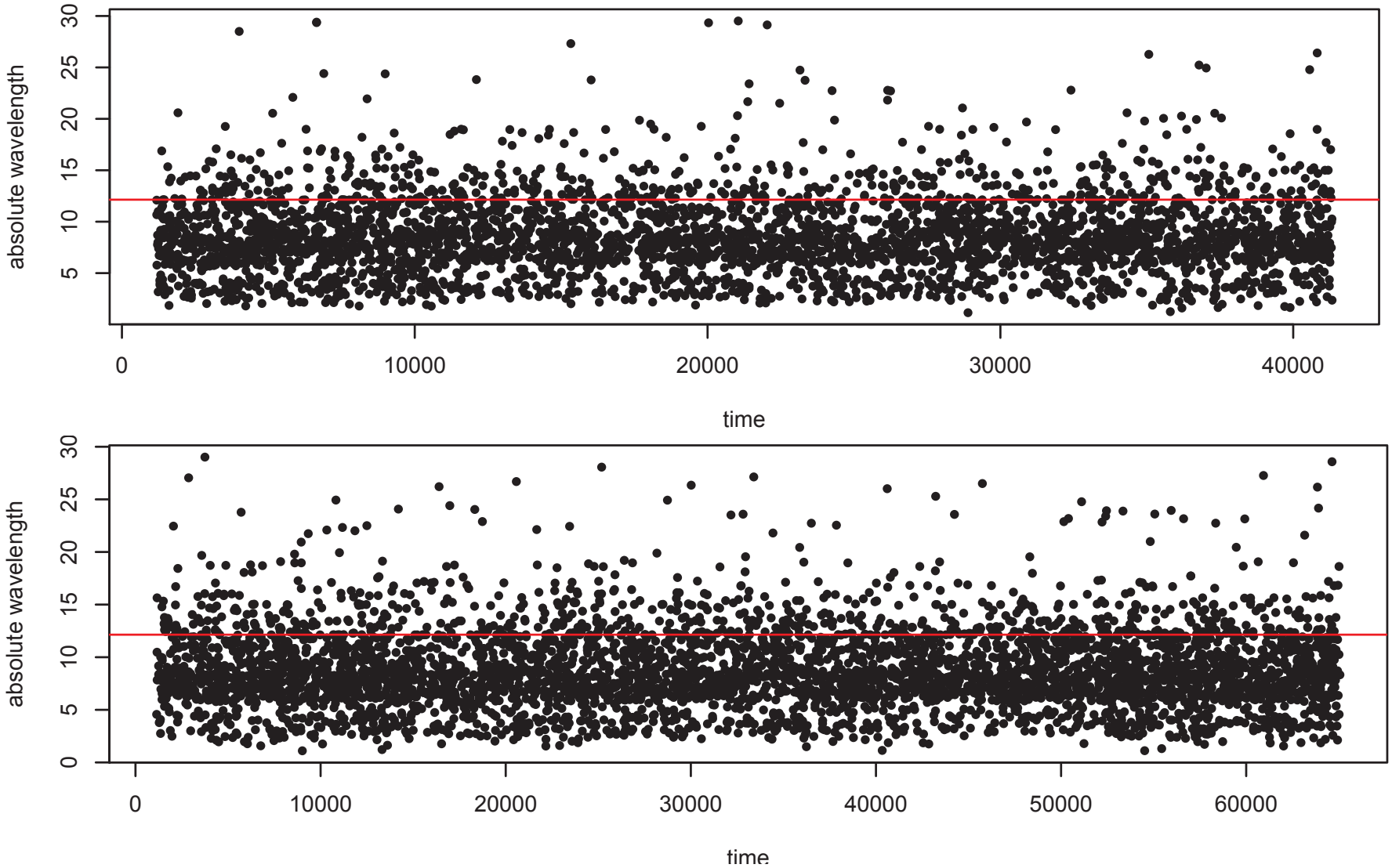
The problem in the first paper

- 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 and hence its energy (E)
  3. its coordinate (x and y, not used in the first paper)

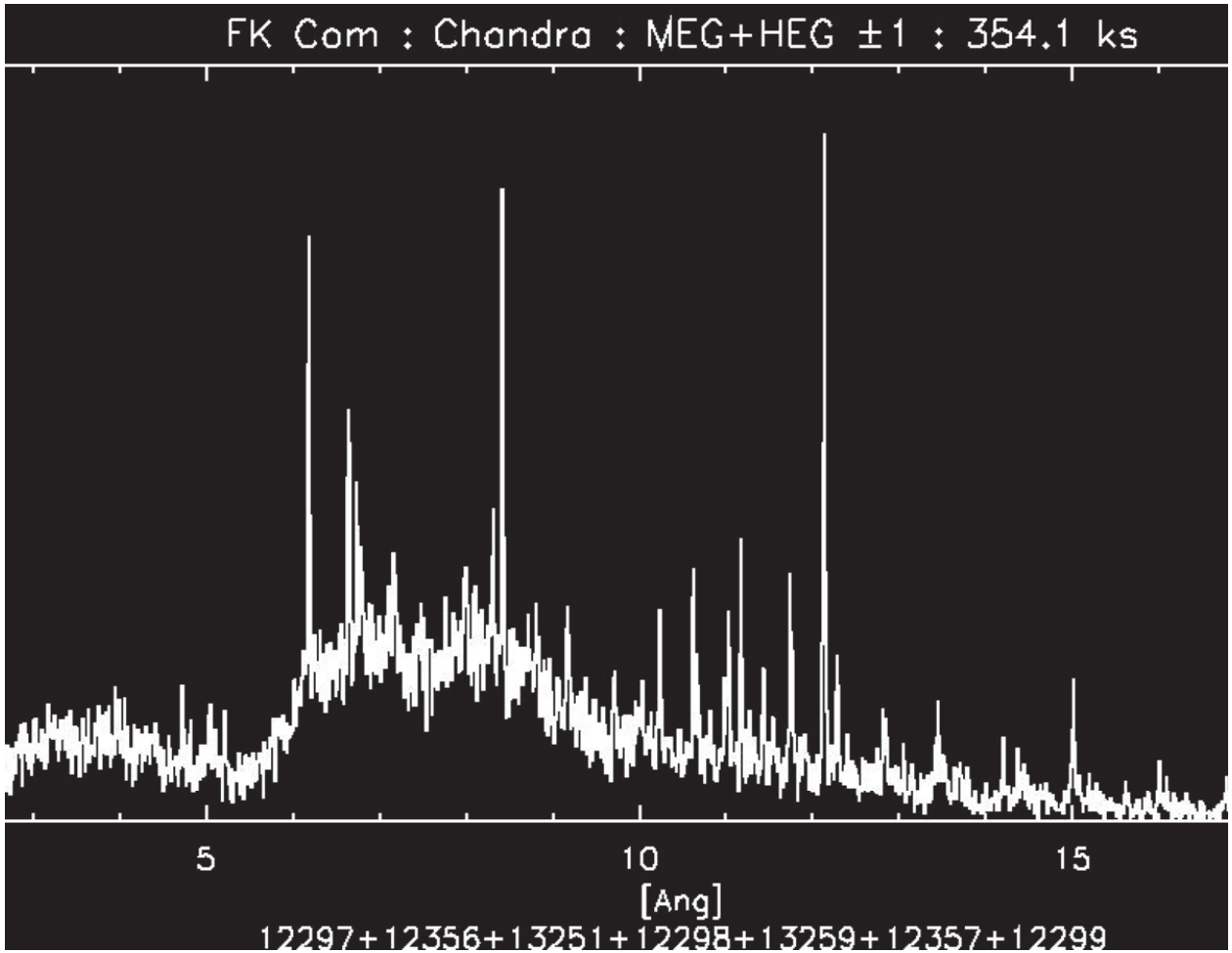
The problem in the first paper

- 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 and hence its energy (E)
  3. its coordinate (x and y, not used in the first paper)
- one interesting question: does the distribution of energy change over time?

Two typical data sets (forget about the red lines)

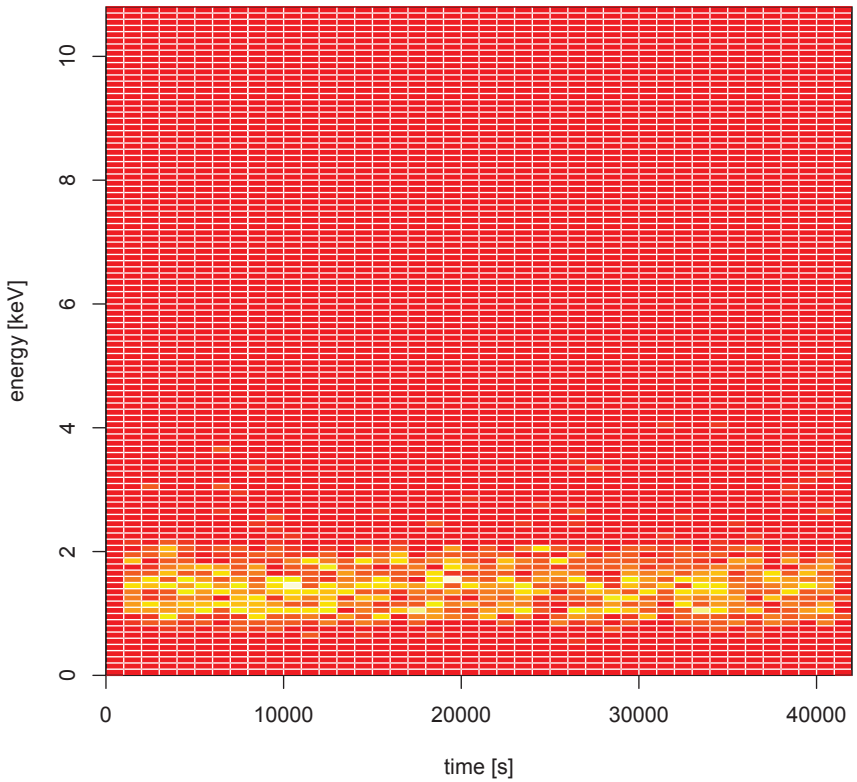


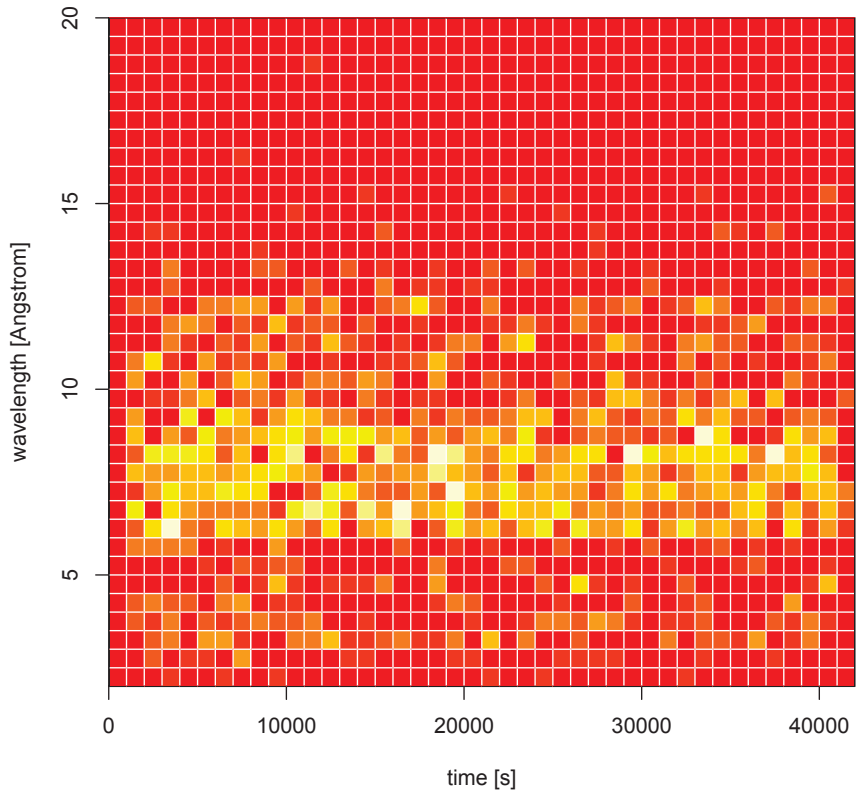
Spectrum at a fixed time t



Note the emission lines

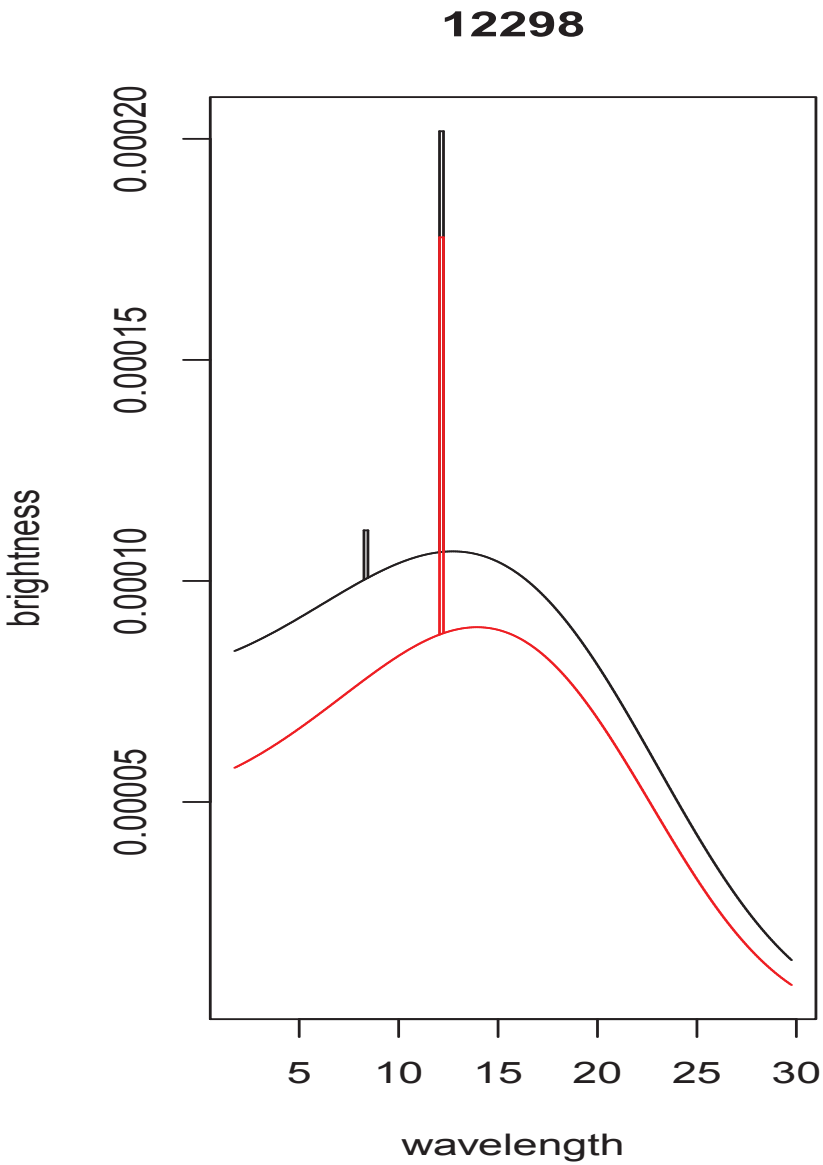
Preprocessing: binning

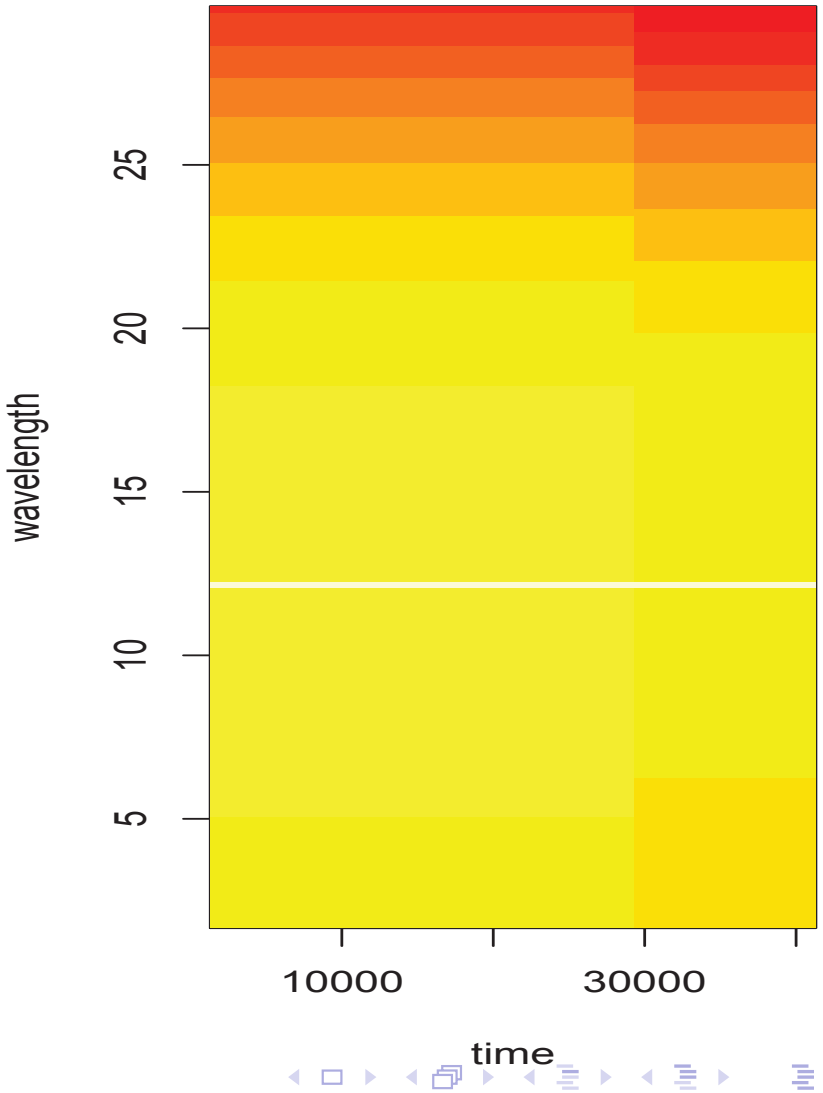




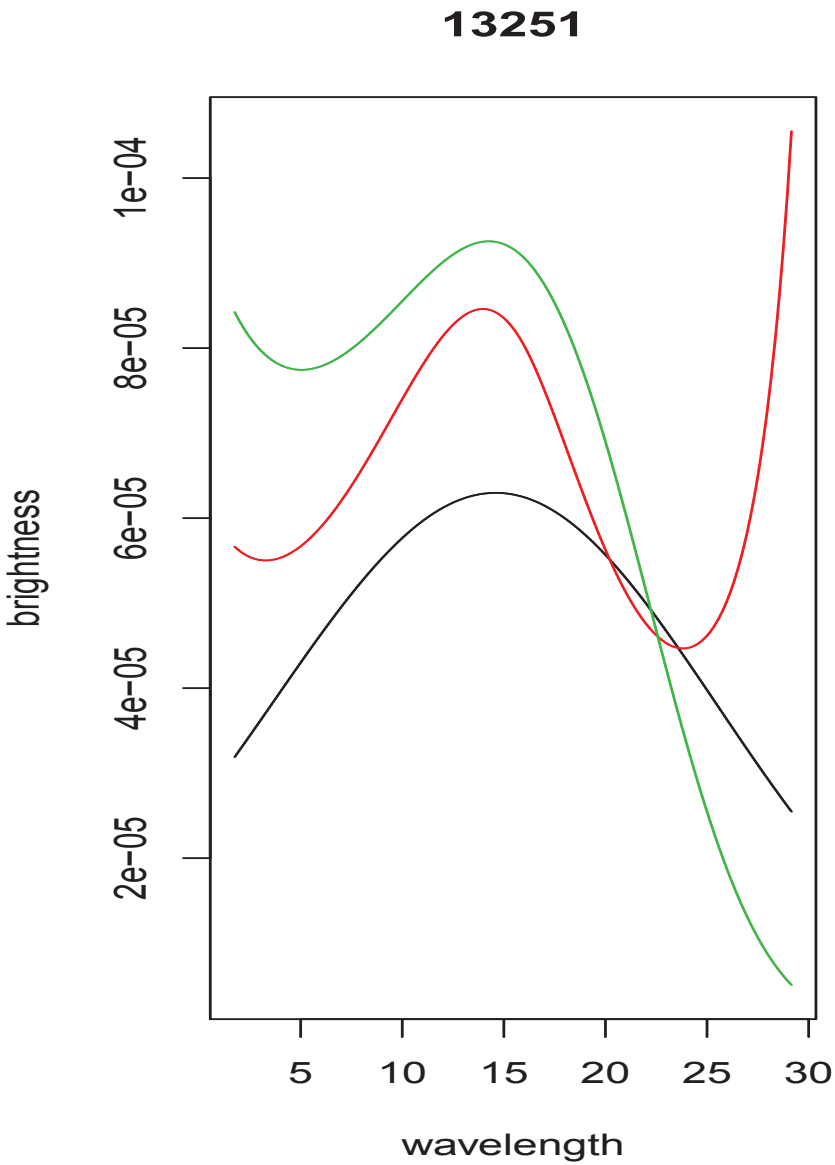
task: estimate/smooth the spectra and detect break points

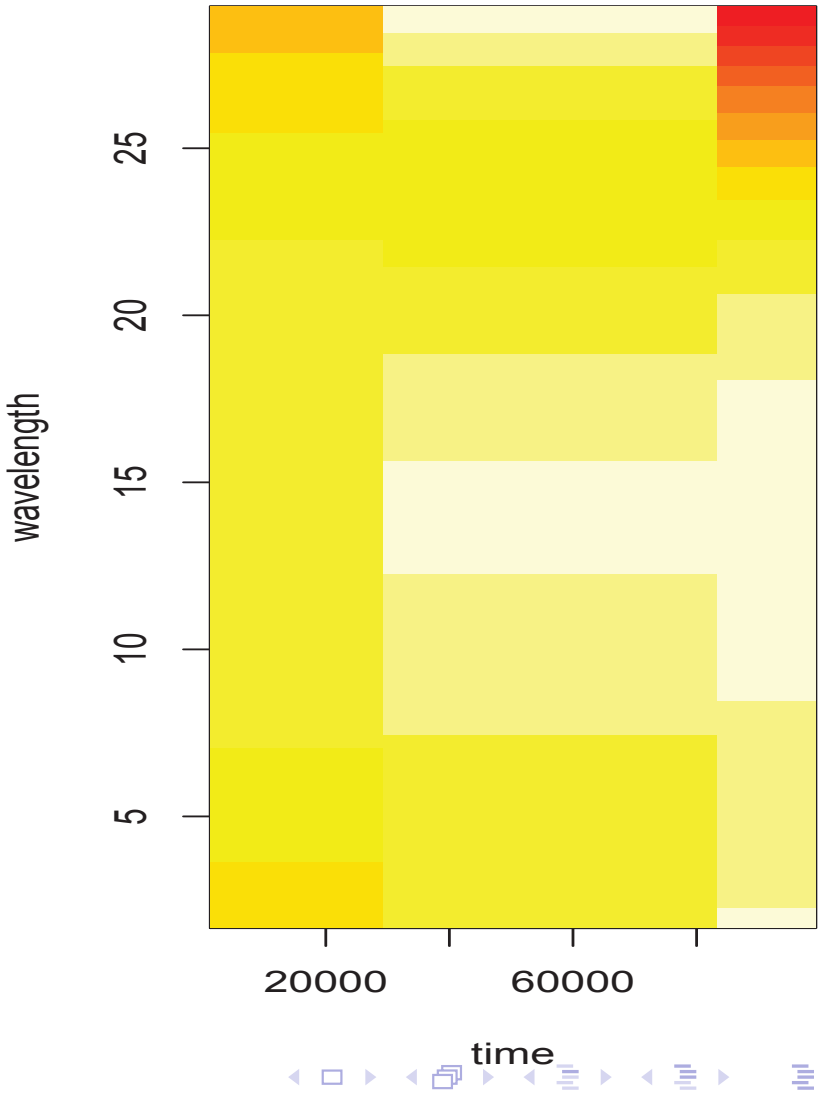
Results: data set 1, two segments (black, red)





Results: data set 2, three segments (black, red, green)





The problem in the second paper

The problem in the second paper

- use all photon attributes: t, E, x and y

The problem in the second paper

- use all photon attributes: t, E, x and y
- do binning spatially (x and y) and temporally (t)

The problem in the second paper

- use all photon attributes: t, E, x and y
- do binning spatially (x and y) and temporally (t)
- a time series of images

The problem in the second paper

- use all photon attributes: t, E, x and y
- do binning spatially (x and y) and temporally (t)
- a time series of images
- (show movie)

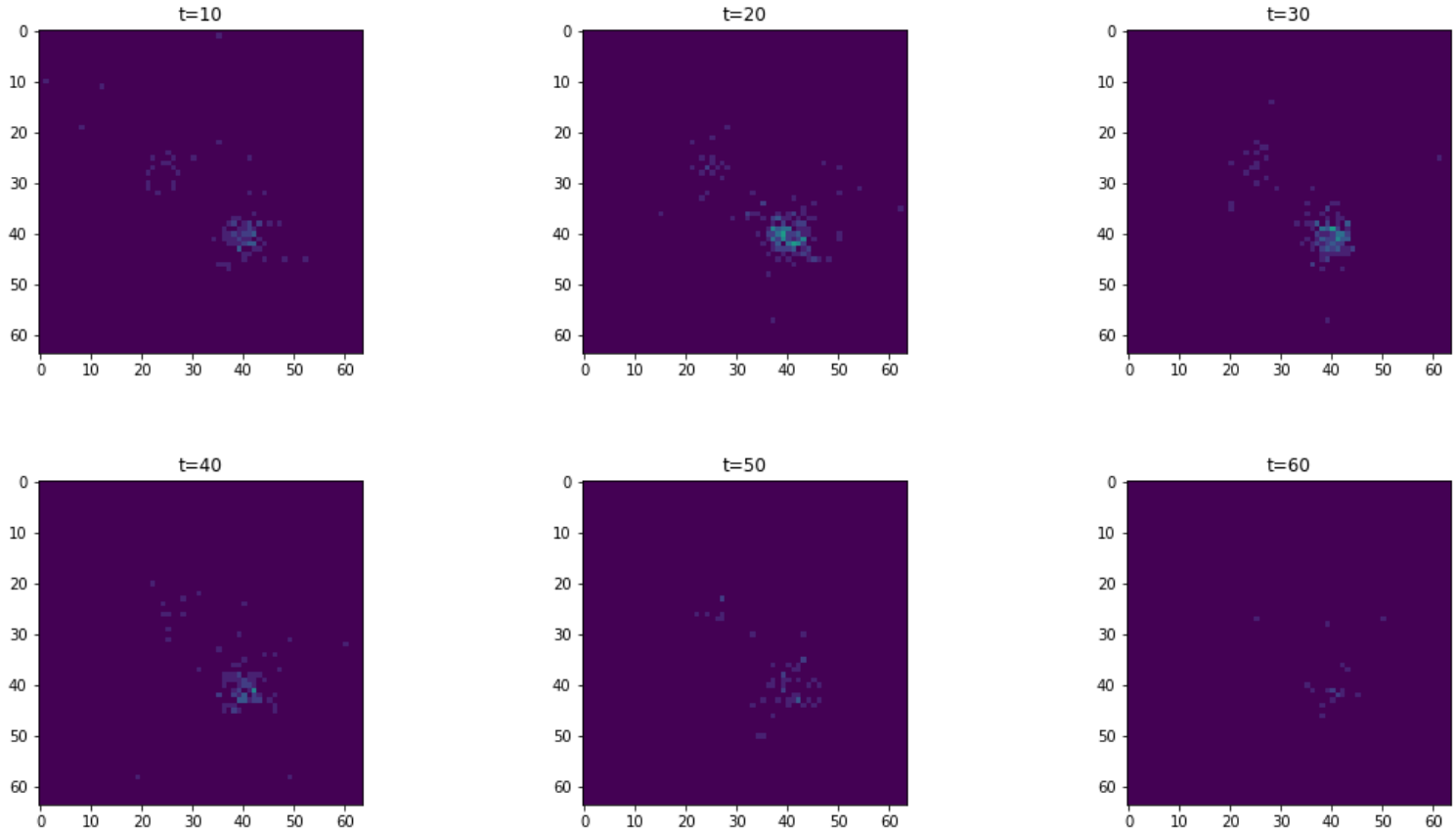
The problem in the second paper

- use all photon attributes: t, E, x and y
- do binning spatially (x and y) and temporally (t)
- a time series of images
- (show movie)
- break/change point over time?

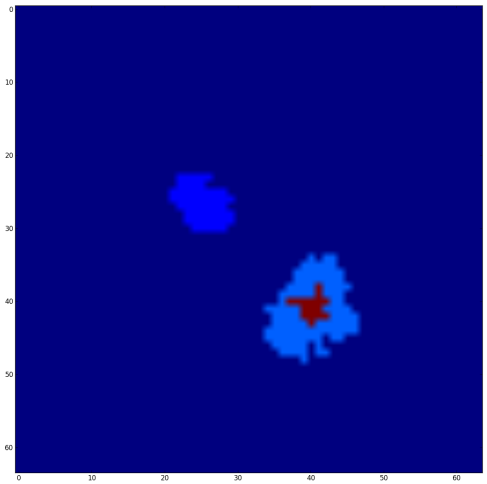
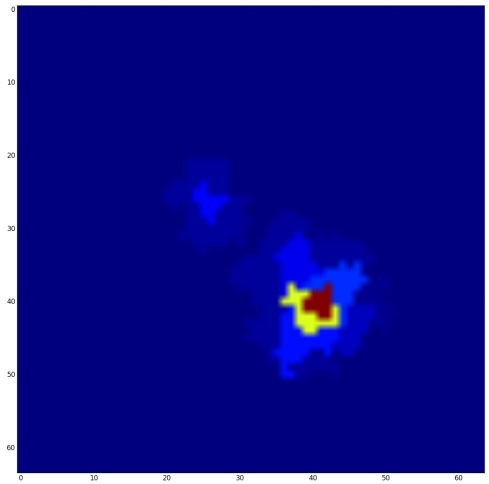
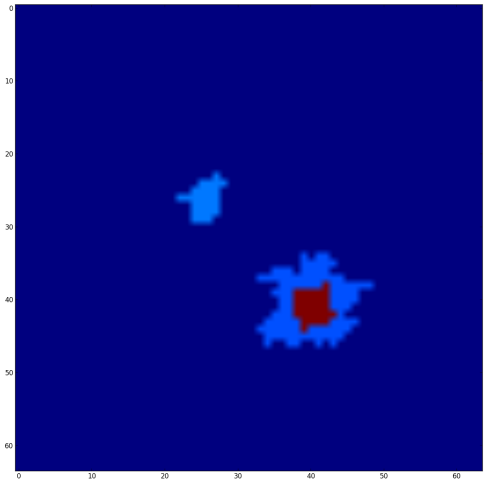
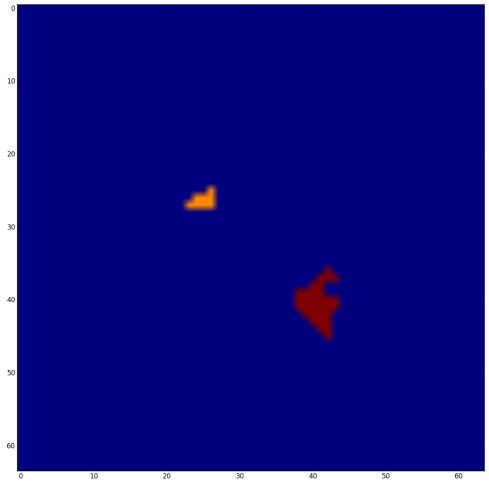
The problem in the second paper

- use all photon attributes: t, E, x and y
- do binning spatially (x and y) and temporally (t)
- a time series of images
- (show movie)
- break/change point over time?
- task: estimate/smooth the images and detect break/change points

Application to real data



Detected breakpoints at $t = 12, 42, 57$

Concluding remarks

- speed is a big issue (Axel and Cong are on it)

Concluding remarks

- speed is a big issue (Axel and Cong are on it)
- without binning?