Cluster-izing in AFNI

The Times They Are a-Changin'

Bob (not Dylan) Cox

In Response to

Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. by Eklund, Nichols, & Knutsson

http://www.pnas.org/content/113/28/7900.full

The line it is drawn
The curse it is cast
The slow one now
Will later be fast
As the present now
Will later be past
The order is rapidly fadin'
And the first one now will later be last
For the times they are a-changin'

RWC: Feb 2017

NIMH National Institute of Mental Health

NIH

The "Classic" Method (AFNI)

- Estimate spatial smoothness of noise
  - Assume spatial auto-correlation function (ACF) has a Gaussian shape (1 parameter = FWHM)
  - Estimate FWHM from correlation of 1st spatial differences in noise (Forman *et al* 1995)
- Simulate noise-only 3D data with that smoothness, determine cluster-size threshold that goes with a given voxel-wise threshold, to give global False Positive Rate (FPR)

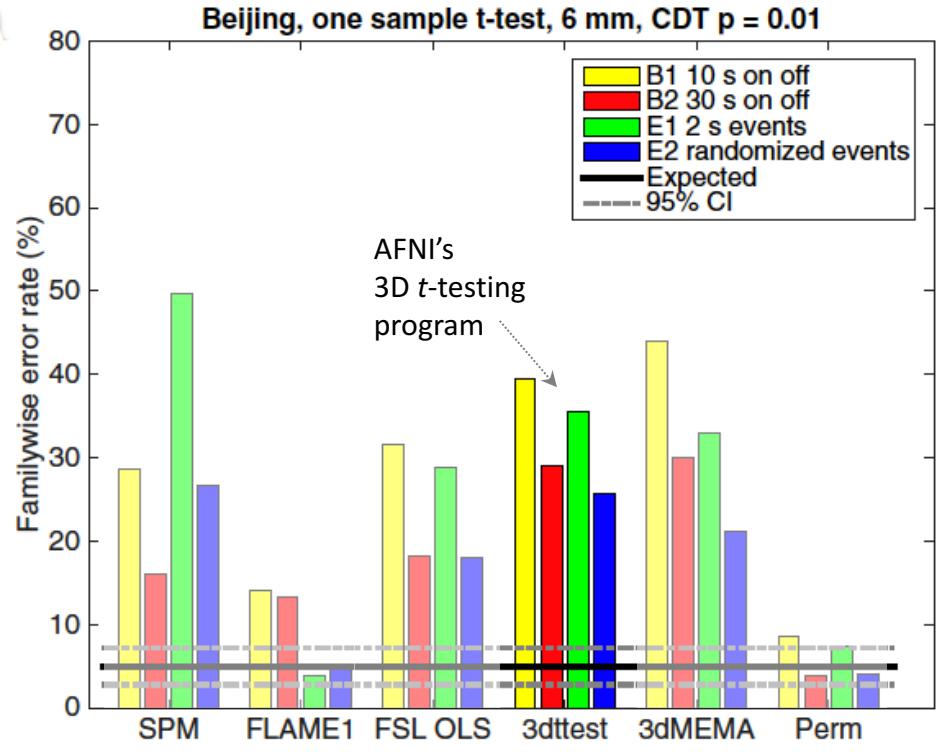
RWC: Feb 2017

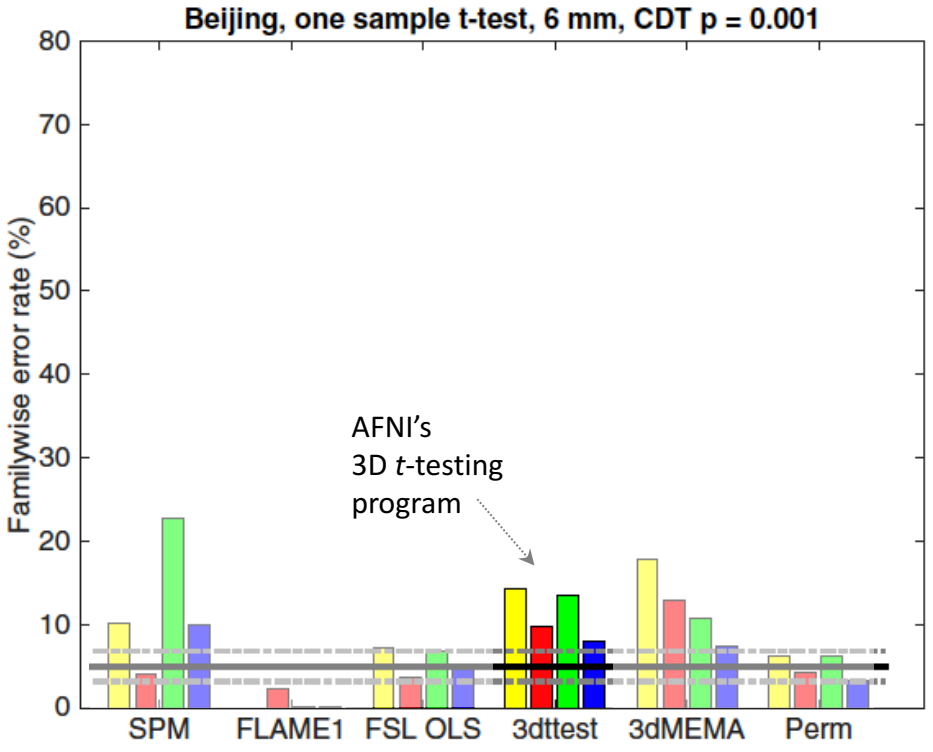
Testing *Some* Method

- Eklund *et al*: use rs-FMRI (FCON-1000) as null data
  - Analyze each of 198 x 2 subject collections (Beijing and Cambridge) with fake task timings
    - 2 x Block design, 2 x Event-related design
    - 4 x spatial blur levels (4, 6, 8, 10 mm)

    } 16 basic cases
  - Carry out 1- and 2-sample *t*-tests between subsets of these collections – 1000 random subsets (per case, per collection, per diverse variations)
  - Count clusters surviving, get FPR estimate
- Their scripts and results are available on GitHub

Some Results from Eklund *et al*

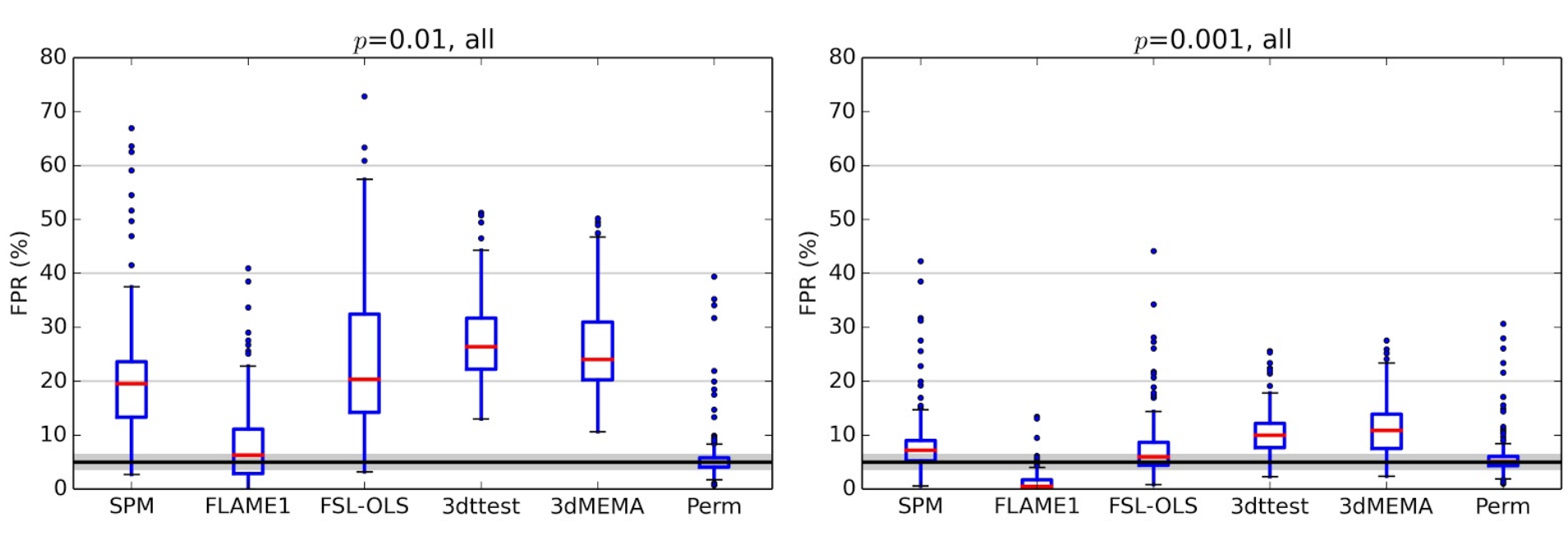




- *SPM and FSL and 3dMEMA results not further discussed herein*
- Each bar is the FPR estimated from 1000 1-sample *t*-test runs
- Nominal 5% band (3.6% to 6.4% = 95% interval for 1000 samples)
- Per voxel p=0.010 on left, p=0.001 on right

RWC: Feb 2017

All Their Results Summarized



- Box plots across all cases: 1- & 2-sample, various sample sizes, various "stimuli", various data sources
- "Up to 70%" FPR (triply-used quote from Eklund *et al*) is not a decent summary of the situation.

RWC: Feb 2017

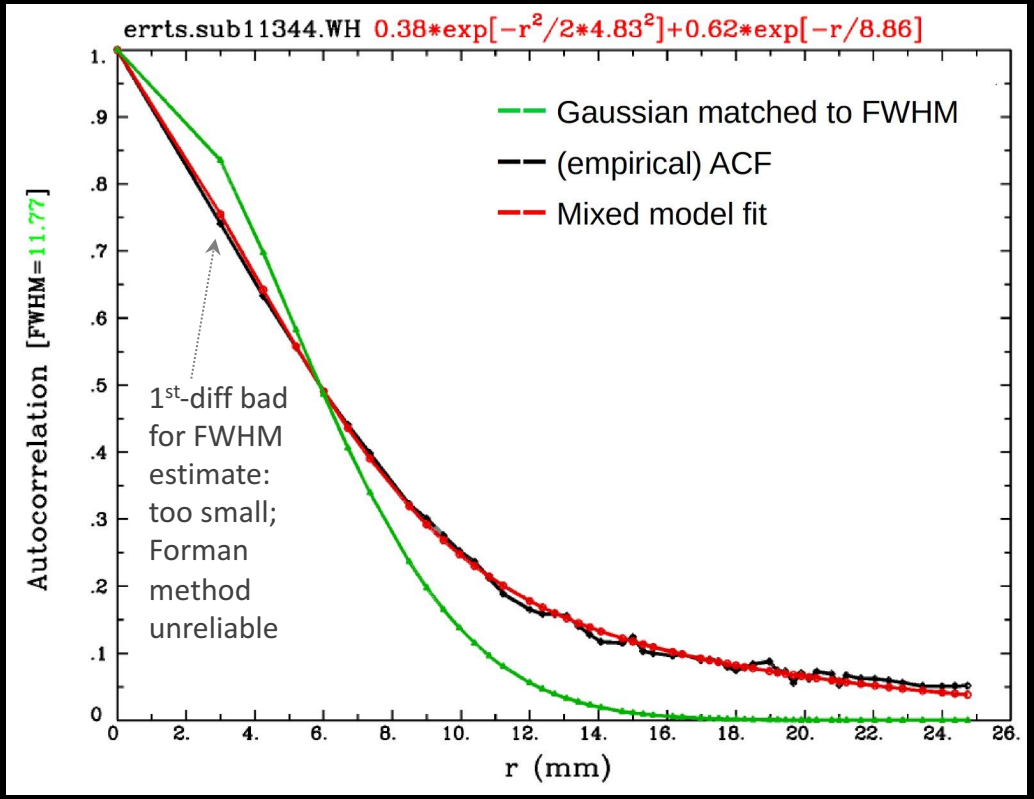
Bugs and Flaws

- AFNI's cluster-size threshold calculating program (3dClustSim) had a bug
  - Is a big deal in the PNAS paper (and popular press)
  - Was not actually that important (*cf* 3 slides ahead)
  - Using the Forman method was another flaw
- However, there was/is a bigger flaw
  - Shared with FSL and SPM
  - Assumption of Gaussian shape for spatial autocorrelation function (ACF) of the noise
    - Describes how noise in one voxel is correlated with noise in another voxel (distance *r* away)

RWC: Feb 2017

NonGaussianity in ACF

- ACF from single subject datasets has long tails – nonGaussian shape + 1st difference fail

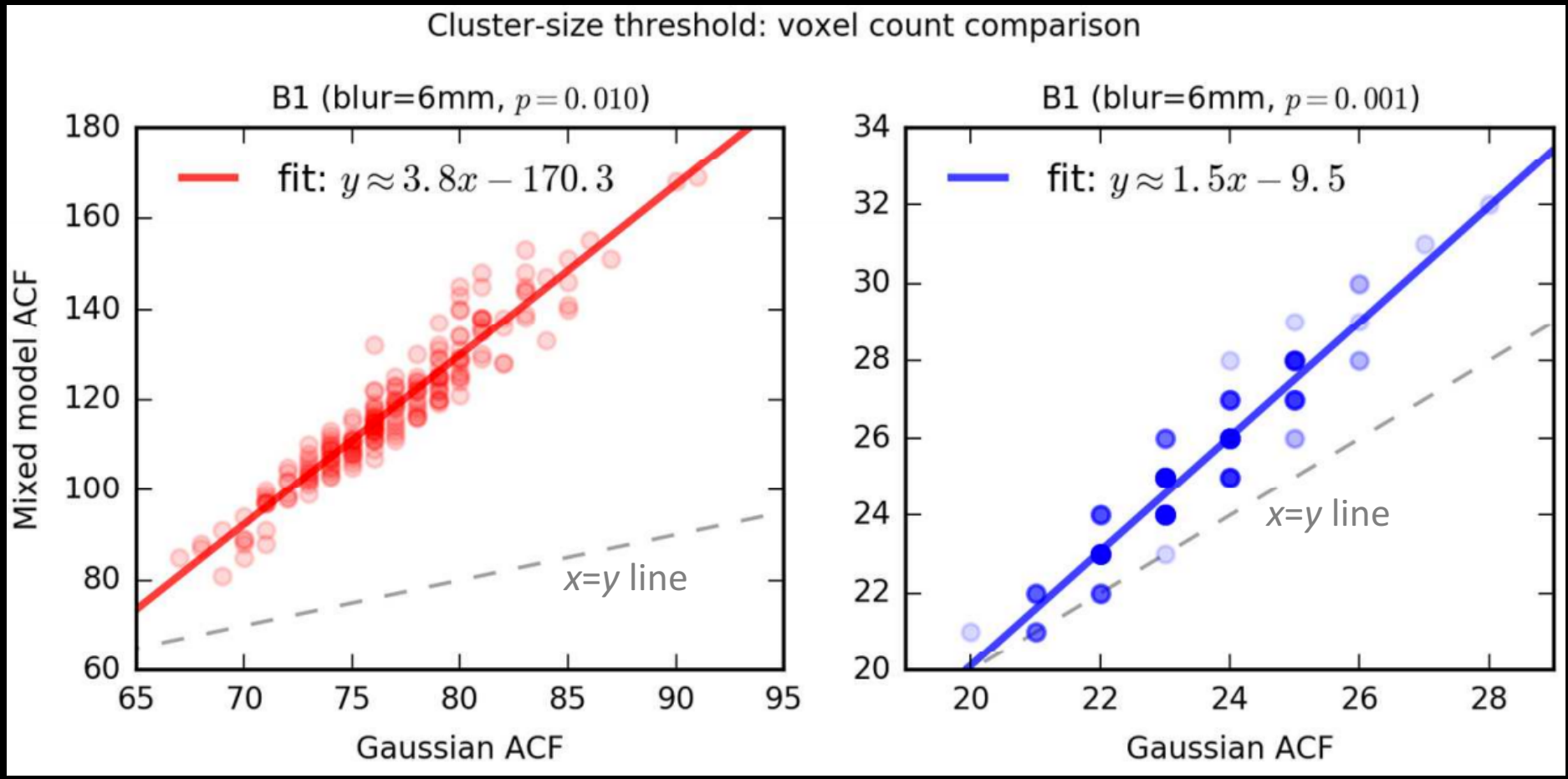


Modify 3dClustSim to use mixed ACF model (Gaussian plus mono-exponential) with 3 parameters (*a*,*b*,*c*) instead of 1 (FWHM)

$$a\ \exp[-r^2/(2b^2)]+(1-a)\exp[-r/c]$$

RWC: Feb 2017

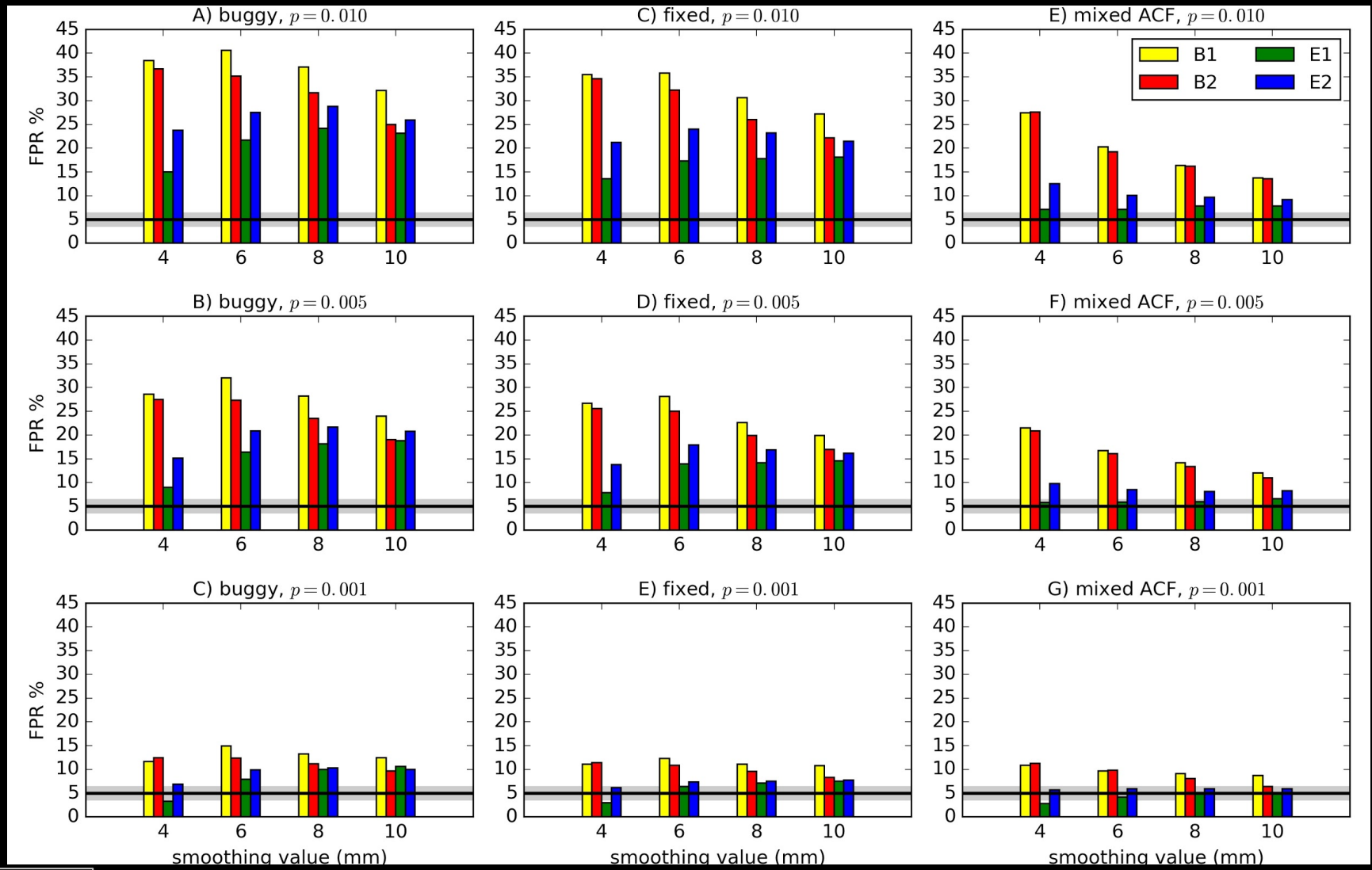
Do Long Tails Matter? *Yes*



- Compare cluster-size thresholds for 198 subjects
- Computed via 3dClustSim using 2 different ACF models
- In words: don't use Gaussian ACF for FMRI (as usually done)
  - *NB*: Gaussian FWHM taken from mixed model ACF (*cf* previous slide)

RWC: Feb 2017

AFNI Results Redux



RWC: Feb 2017

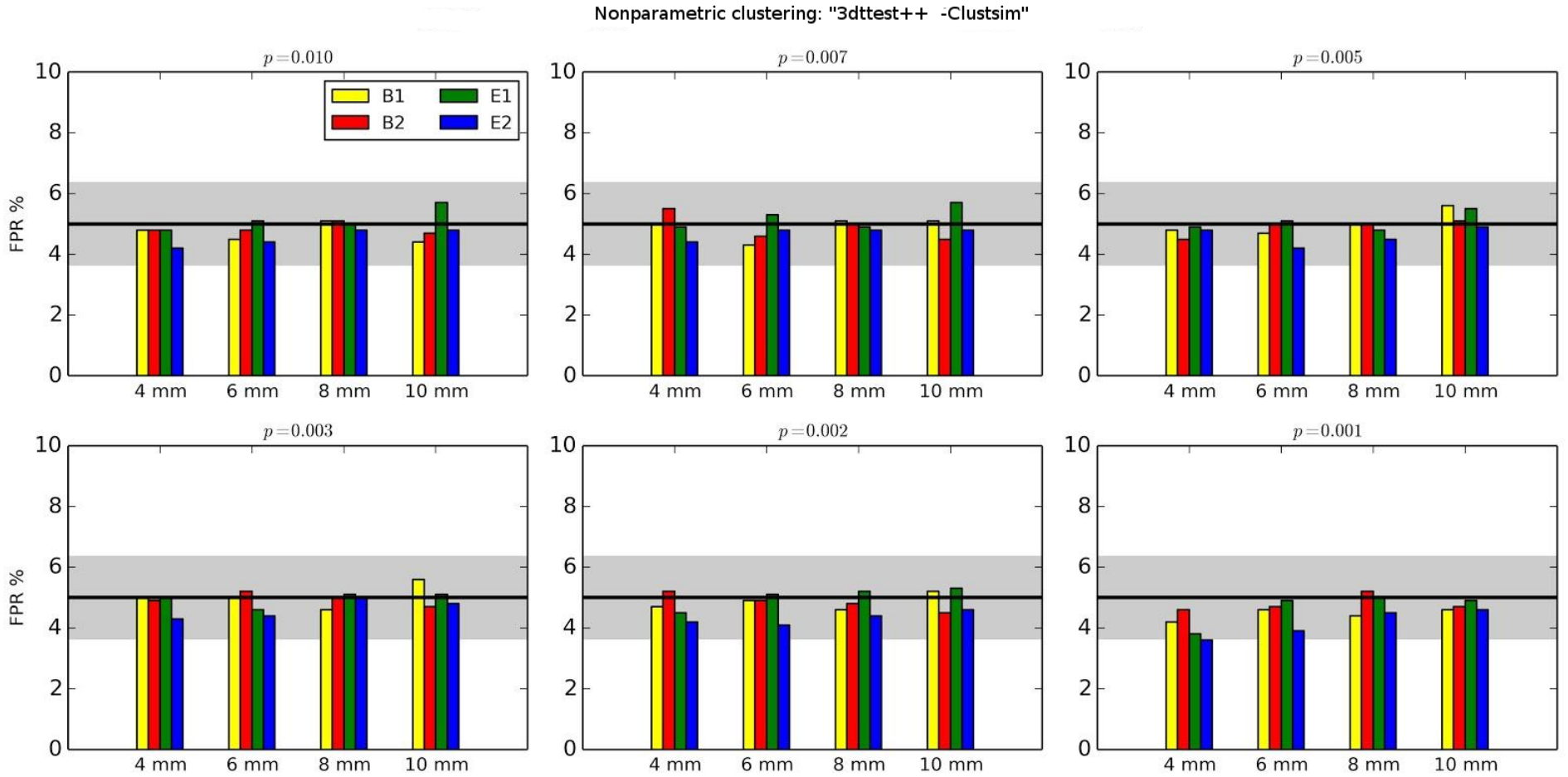
How to: ACF method

- Run 3dFWHMx with the '-acf' option to get the ACF parameters for each subject, from the residuals dataset errts*+tlrc.HEAD
  - This is done now in afni_proc.py
- Average each of the 3 ACF parameters across subjects (not automatic)
- Use 3dClustSim with the '-acf' option (giving the 3 averaged parameters) to get the cluster size threshold tables for the group analysis

Why Is Model-Based FPR Still High?

- Using ACF estimates improved results
  - So the wider ACF and longer tails are a part of the original problem – *but not all of it*
- Too short tails in the group *t*-statistics, caused by outlier subjects in the data
  - Also explained a part of it – *but not enough*
- My current thinking
  - Spatial ACF is *not* stationary (same everywhere)
    - Over-wide in some places

Another Solution: Nonparametric Clustering in AFNI



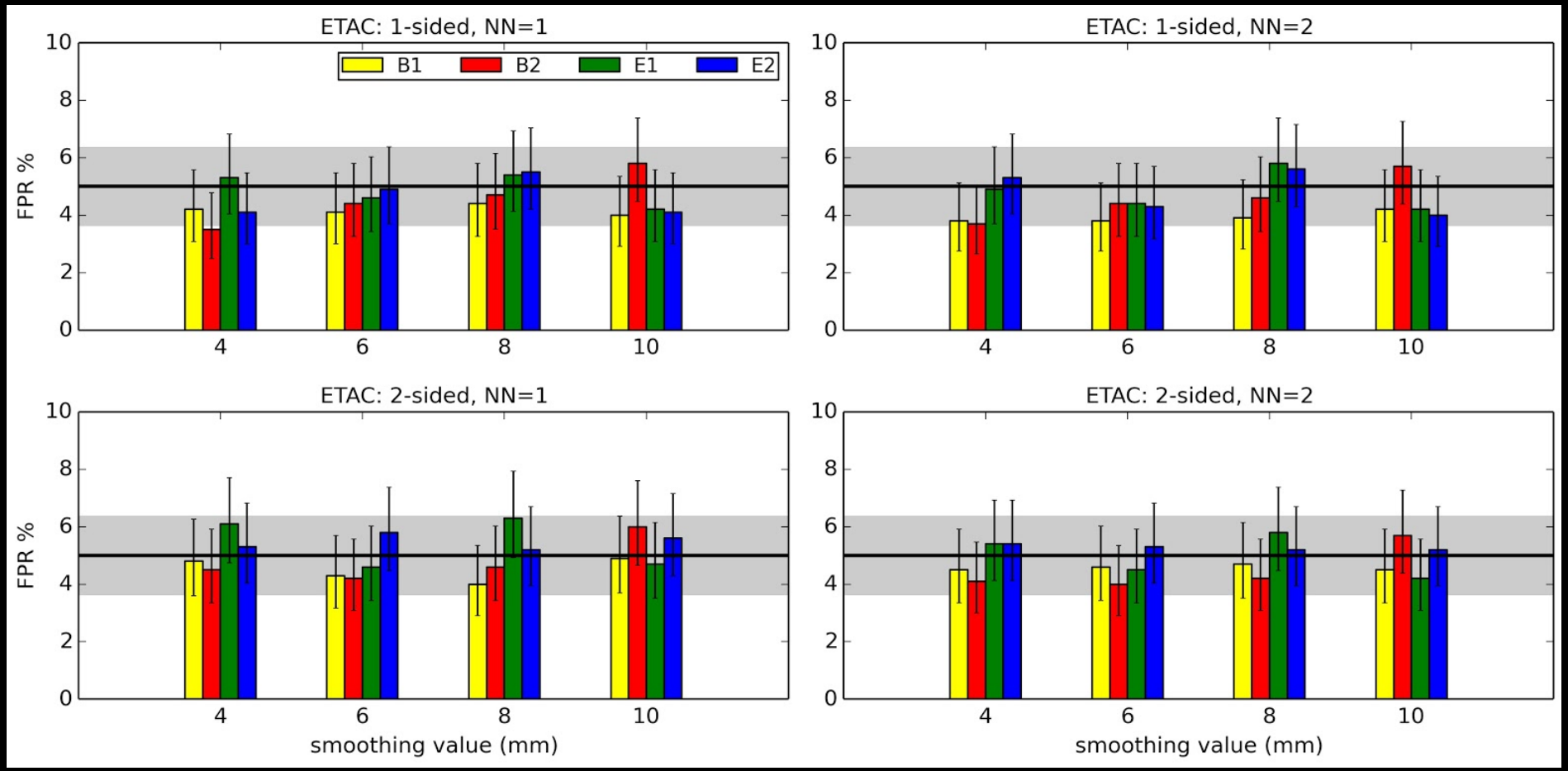
- *t*-test residuals are permuted/randomized (10000 times)
- 10000 re-*t*-tests computed from residuals fed into 3dClustSim

RWC: Feb 2017

How to: Nonparametric Clustering

- Only for *t*-tests at this time
- 3dttest++ with the –Clustsim option
- Gives excellent FPR control ☺
- Has stringently large cluster-size thresholds ☹
  - Seems to be needed to deal with the extra-wide spatial ACF in some regions
  - Cluster-size threshold is very nonlinear in smoothness
- Leads to the idea of making the cluster-size threshold depend on spatial location ……

In progress: **ETAC** (the future*?*)



- **E**quitable **T**hresholding **A**nd **C**lustering [Not the most beautiful name]
- Spatially variable cluster-size threshold
- Also: Uses multiple *p*-value thresholds simultaneously

RWC: Feb 2017

Finally

- At this time (with up-to-date AFNI):
  - 3dClustSim -acf ($p \leq 0.002$) *OR* 3dttest++ -Clustsim
- In the future
  - ETAC (needs more testing, more generalization)
  - A parametric method to allow for spatially variable ACF [*ie*, (a,b,c) as functions of (x,y,z)]*???*
    - Would be useful for group stats too complex to permute (*eg*, 3dLME, 3dMVM)
- Paper accepted in *Brain Connectivity*
- Letter accepted in *PNAS*
  - We'll put them on aRxiv.org *real soon now*

RWC: Feb 2017