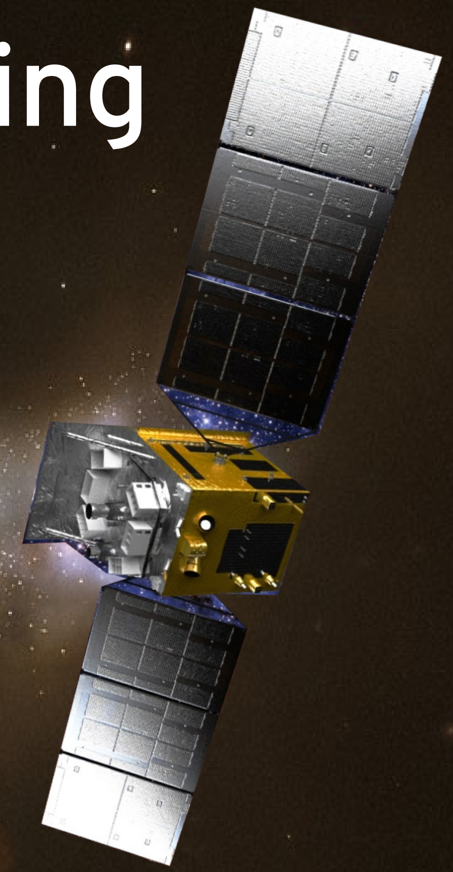
BeppoSAX-GRBM background modelling with machine learning techniques: prospects for HXMT



Cristiano Guidorzi

on behalf of the Uni Ferrara + INAF Bologna GRB team

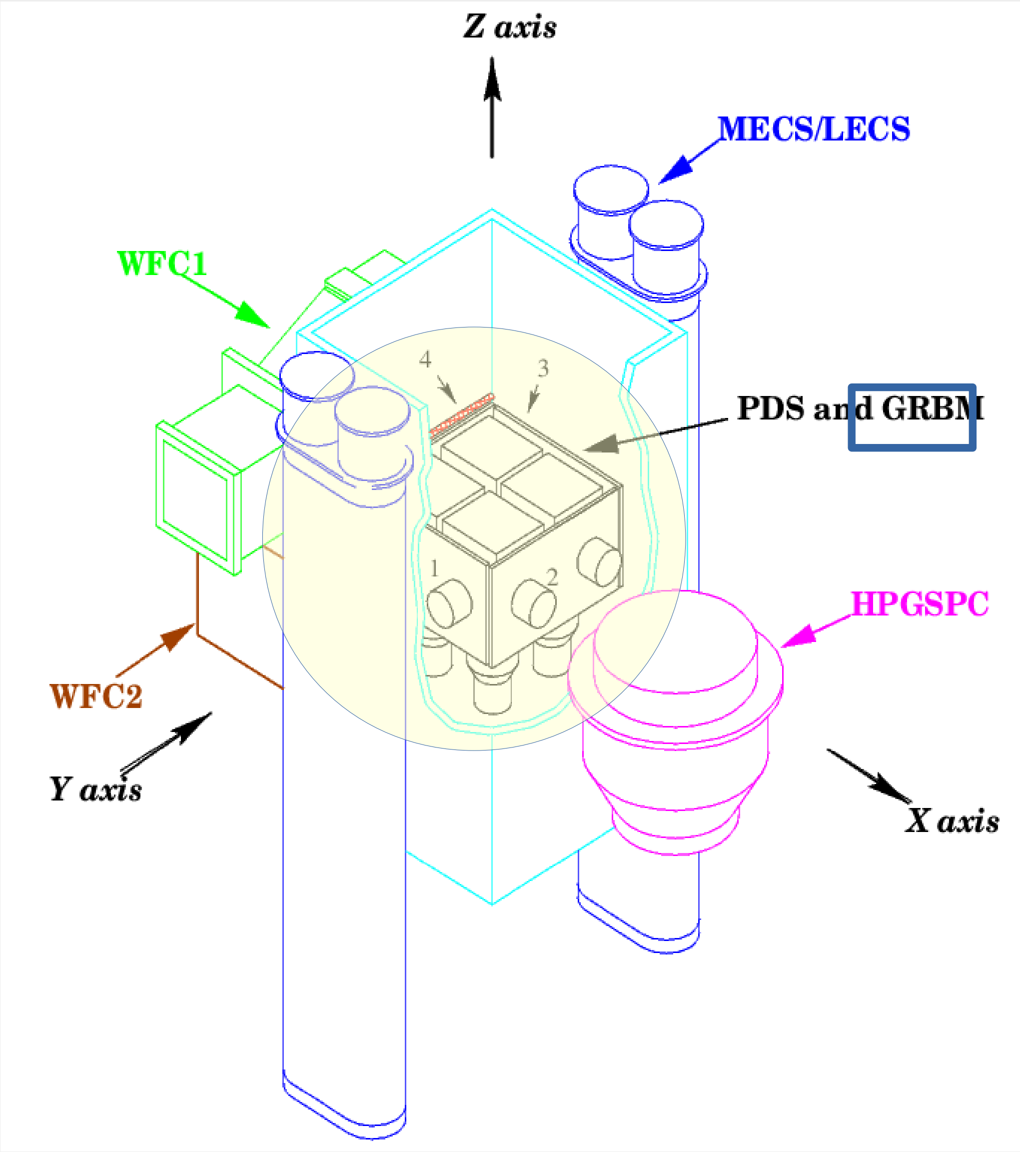
University of Ferrara

agenzia spaziale italiana

INAF - IASF BOLOGNA

ISTITUTO DI ASTROFISICA SPAZIALE
E FISICA COSMICA - BOLOGNA

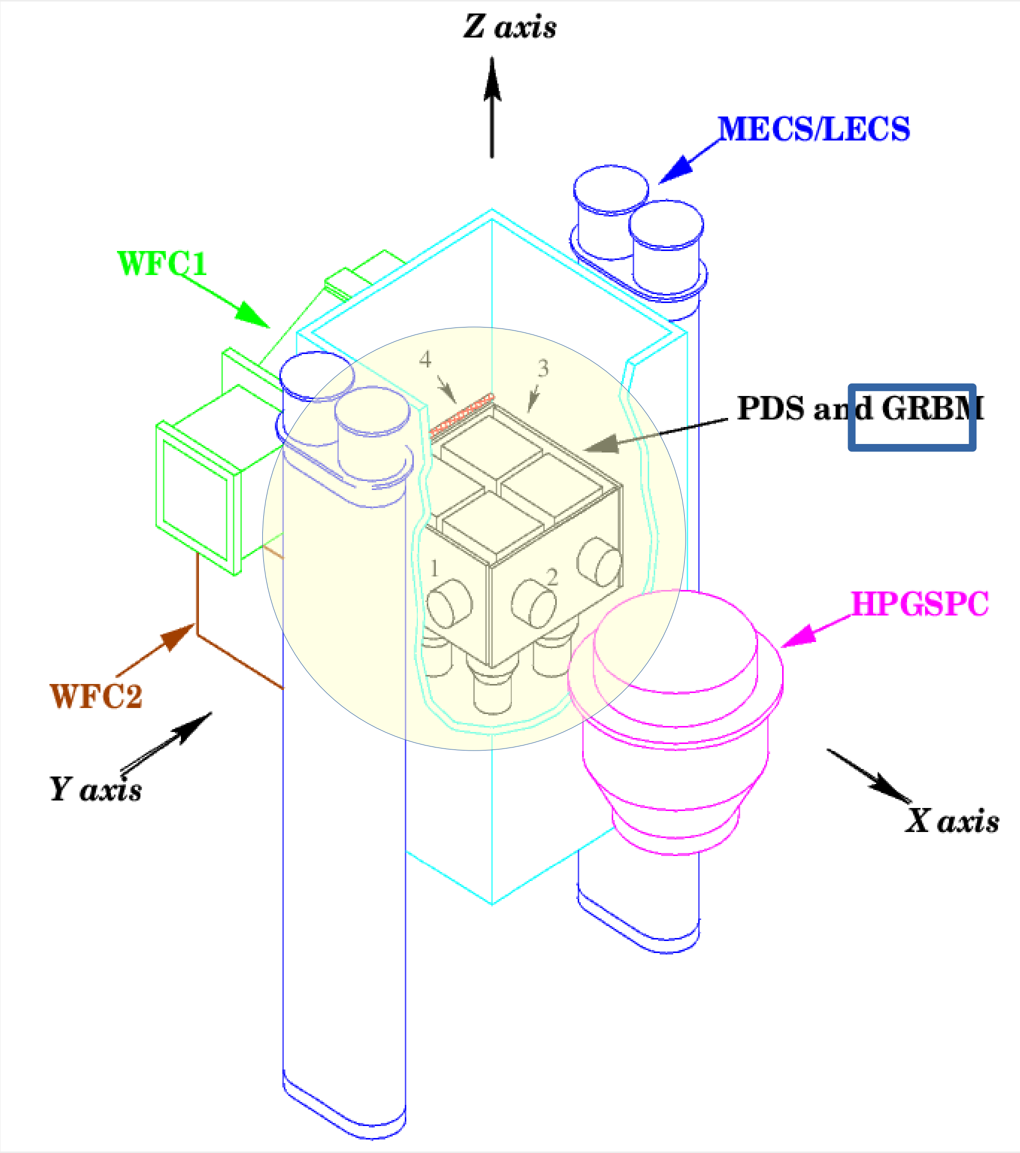
BeppoSAX/GRBM



- 4 independent CsI(Na) slabs (1 x 27.5 x 41.3 cm^3)
- Geom area: 1136 cm^2 each
- 2 energy bands:
  - 40-700 keV
  - > 100 keV

  C/R time resolution:
  - 1 s (continuous)
  - 7.8ms/0.5ms (GRB trigger)
- Low-inclination (~3°)orbit

BeppoSAX/GRBM



- 4 independent CsI(Na) slabs (1 x 27.5 x 41.3 cm^3)
- Geom area: 1136 cm^2 each
- 2 energy bands:
  - 40-700 keV
  - > 100 keV

  C/R time resolution:
  - 1 s (continuous)
  - 7.8ms/0.5ms (GRB trigger)
- Low-inclination (~3°)orbit

BeppoSAX/GRBM background modelling: what for?

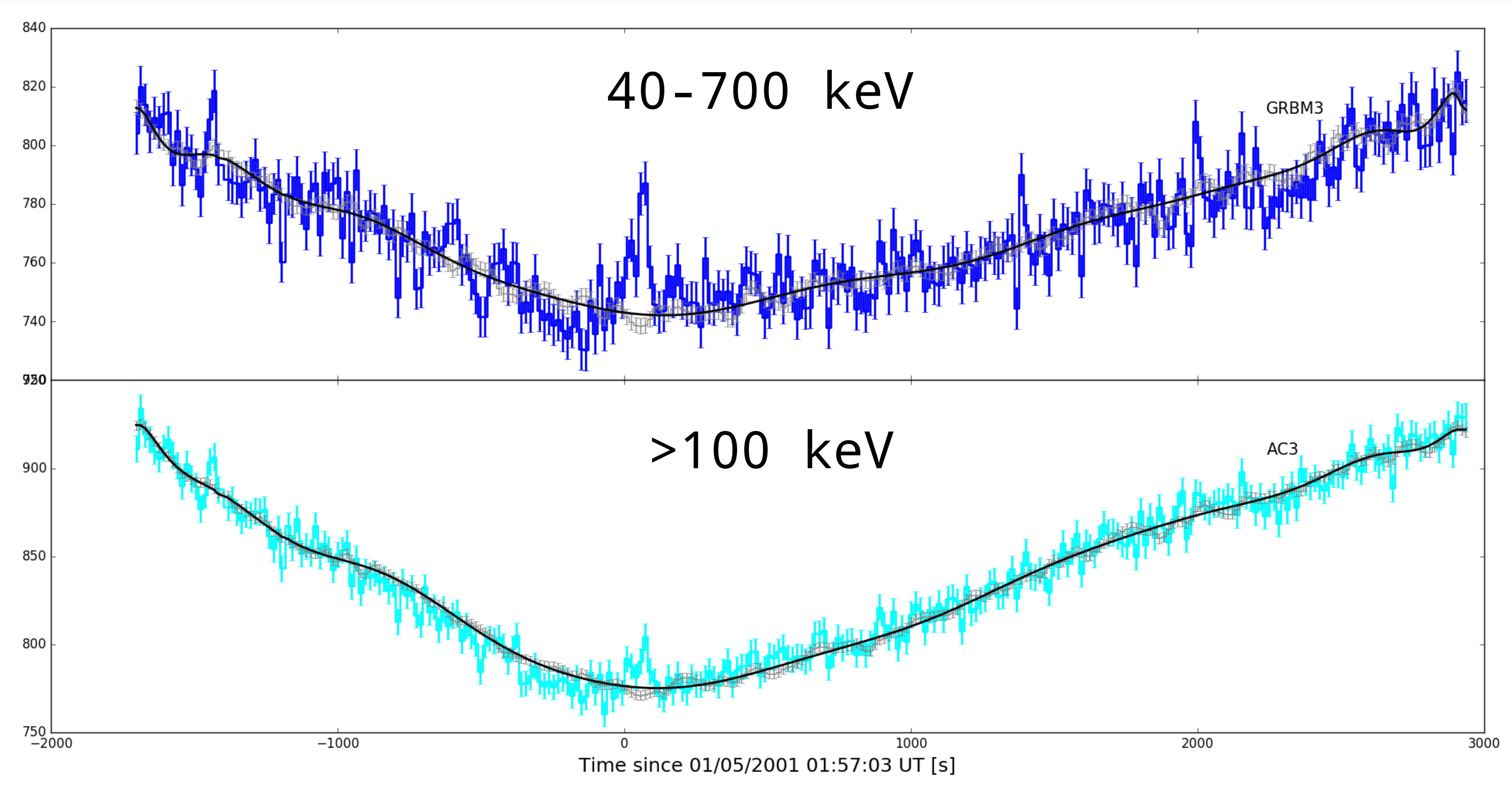
Science Motivations

- Sensitive search for dim, long-lasting (~10^3 s) events so far undetected
  - Cross-check of fast X-ray transients (FXTs) detected in the BeppoSAX Wide Field Cameras
  - Low-luminosity GRBs (possibly different class; shock breakout emission)
  - Possible counterparts to Fast Radio Bursts (DeLaunay+16)
  - Dim, long-lived tails following major events
- Monitoring of bright, hard sources such as CygX-1 in the 40-700 keV

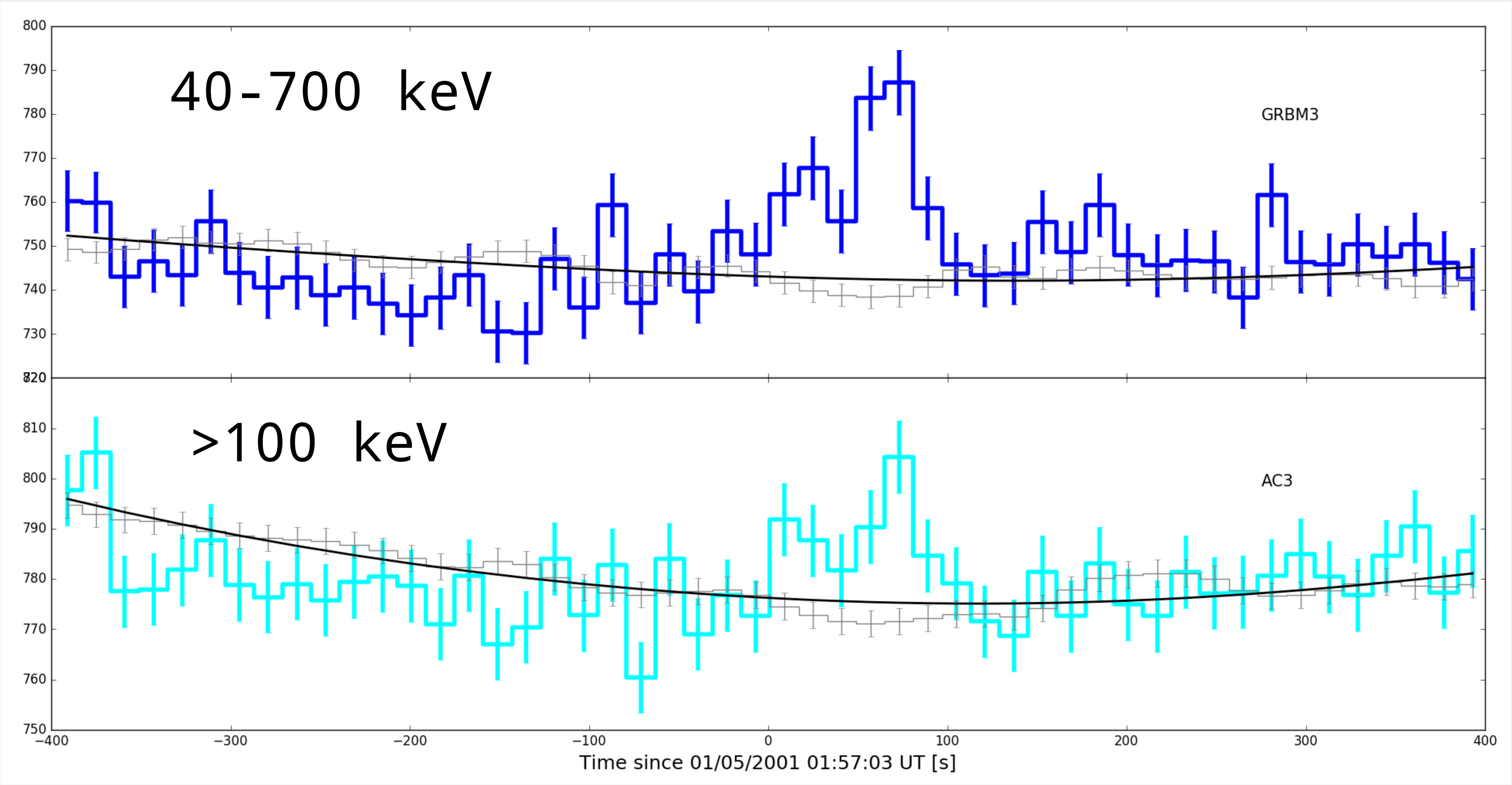
First preliminary attempt prior to adopting machine learning

Example: FXT (GRB?) 010501B

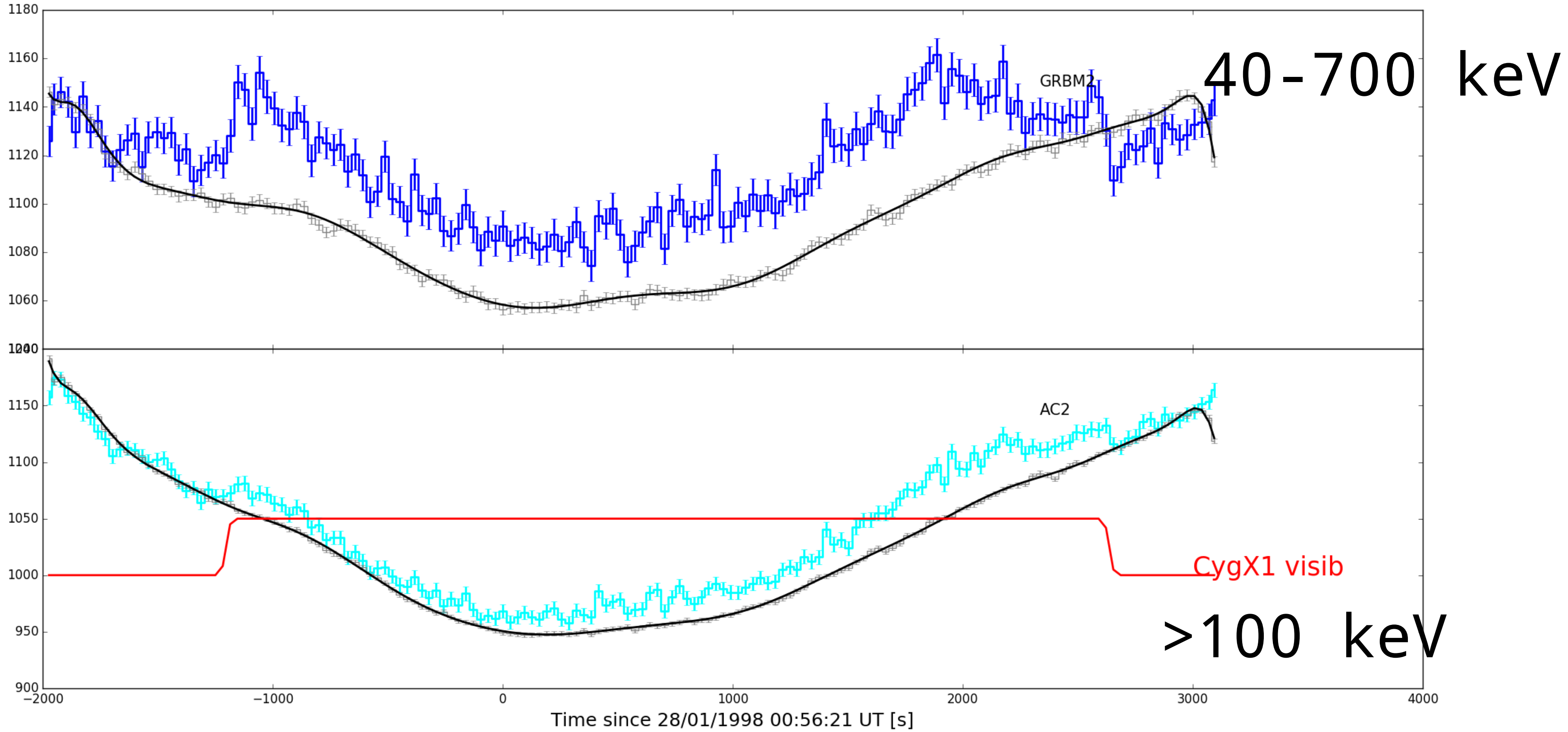
(bkg modelled by averaging out similar orbits)



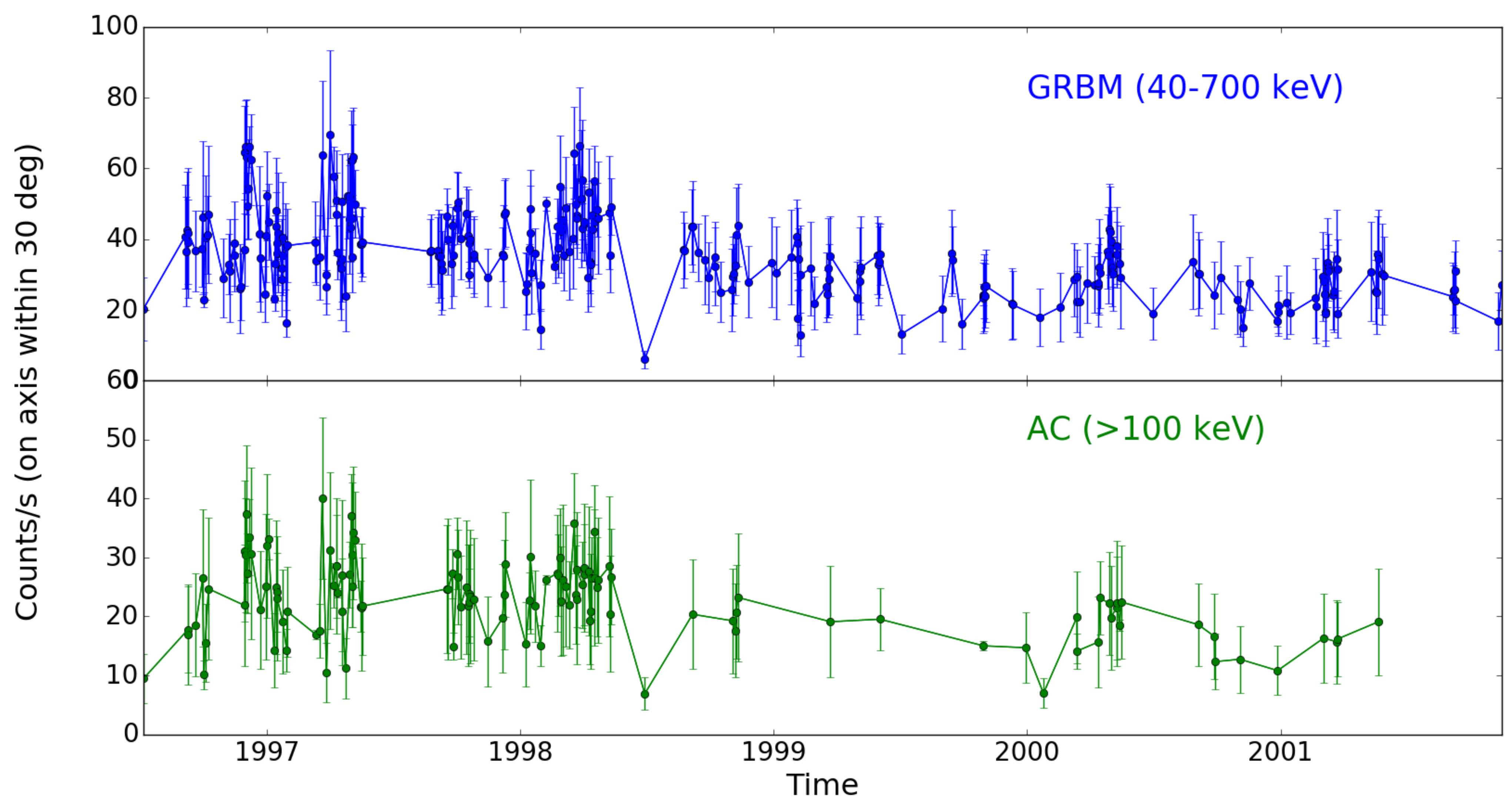
FXT 010501B: close in



Potential for CygX-1



CygX-1: rate history (prelim)



Then we moved on towards a machine learning approach

Core strategy

- Identify a few characteristics (=**features/attributes** in the ML parlance) that are thought/known to mostly contribute to determine the background level and behaviour

Core strategy

- Identify a few characteristics (=**features/attributes** in the ML parlance) that are thought/known to mostly contribute to determine the background level and behaviour
- Partition the entire data set so as to group observations with similar background (based on the attributes that have previously been identified)

Core strategy

- Identify a few characteristics (=**features/attributes** in the ML parlance) that are thought/known to mostly contribute to determine the background level and behaviour
- Partition the entire data set so as to group observations with similar background (based on the attributes that have previously been identified)
- Address a **regression problem** adopting a **supervised learning**:

Core strategy

- Identify a few characteristics (=**features/attributes** in the ML parlance) that are thought/known to mostly contribute to determine the background level and behaviour
- Partition the entire data set so as to group observations with similar background (based on the attributes that have previously been identified)
- Address a **regression problem** adopting a **supervised learning**:
  - Split the data set into a training set and a validation set

Core strategy

- Identify a few characteristics (=**features/attributes** in the ML parlance) that are thought/known to mostly contribute to determine the background level and behaviour
- Partition the entire data set so as to group observations with similar background (based on the attributes that have previously been identified)
- Address a **regression problem** adopting a **supervised learning**:
  - Split the data set into a training set and a validation set
  - Explore the dependence of background on attributes using the training set

Core strategy

- Identify a few characteristics (=**features/attributes** in the ML parlance) that are thought/known to mostly contribute to determine the background level and behaviour
- Partition the entire data set so as to group observations with similar background (based on the attributes that have previously been identified)
- Address a **regression problem** adopting a **supervised learning**:
  - Split the data set into a training set and a validation set
  - Explore the dependence of background on attributes using the training set
    - We currently use a **linear regression** (maximum likelihood estimation)

Core strategy

- Identify a few characteristics (=**features/attributes** in the ML parlance) that are thought/known to mostly contribute to determine the background level and behaviour
- Partition the entire data set so as to group observations with similar background (based on the attributes that have previously been identified)
- Address a **regression problem** adopting a **supervised learning**:
  - Split the data set into a training set and a validation set
  - Explore the dependence of background on attributes using the training set
    - We currently use a **linear regression** (maximum likelihood estimation)
    - Next we plan on using generalized linear models (GLMs)

Core strategy

- Identify a few characteristics (=**features/attributes** in the ML parlance) that are thought/known to mostly contribute to determine the background level and behaviour
- Partition the entire data set so as to group observations with similar background (based on the attributes that have previously been identified)
- Address a **regression problem** adopting a **supervised learning**:
  - Split the data set into a training set and a validation set
  - Explore the dependence of background on attributes using the training set
    - We currently use a **linear regression** (maximum likelihood estimation)
    - Next we plan on using generalized linear models (GLMs)
  - Estimate mean, median, and overall absolute error distributions on validation set

More specifically:

- **Basic data unit definition**:
  - **orbit** (counts within a ~5-ks time window)

More specifically:

- **Basic data unit definition**:
  - **orbit** (counts within a ~5-ks time window)
- **Data selection**:
  - remove orbits affected by known GRBs, solar X-ray flares, data gaps

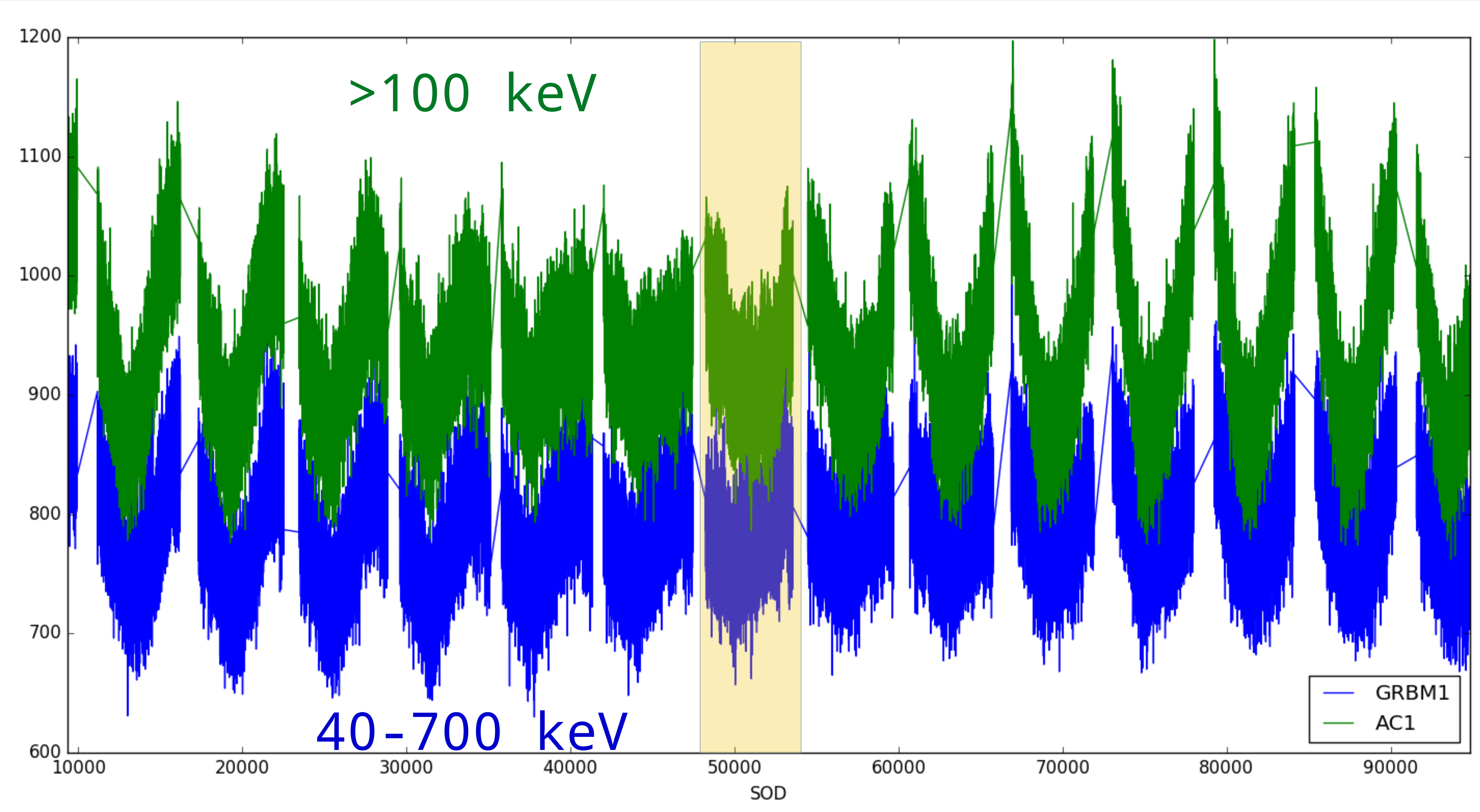
More specifically:

- **Basic data unit definition**:
  - **orbit** (counts within a ~5-ks time window)
- **Data selection**:
  - remove orbits affected by known GRBs, solar X-ray flares, data gaps
- **Starting set of orbit attributes**:
  - Spacecraft (s/c) attitude
  - Portion of visible (=not Earth-blocked) sky
  - Positions (wrt s/c frame) of Crab, CygX-1, Sun (hard X-ray bright sources)
  - Vertical geoagnetic cutoff rigidity and geomagnetic coordinates of s/c
  - Epoch
  - Particle Monitor data

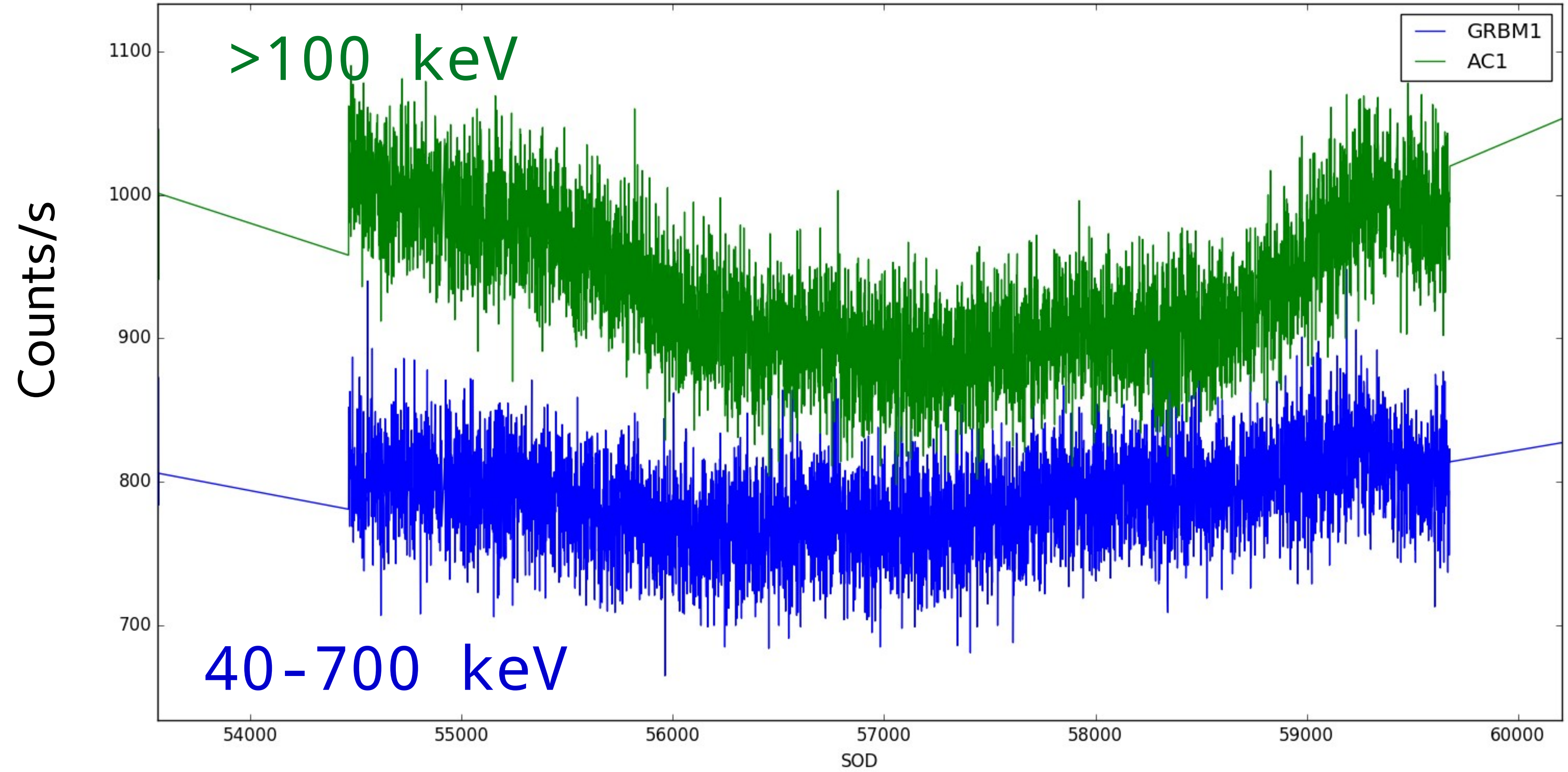
More specifically:

- **Basic data unit definition**:
  - **orbit** (counts within a ~5-ks time window)
- **Data selection**:
  - remove orbits affected by known GRBs, solar X-ray flares, data gaps
- **Starting set of orbit attributes**:
  - Spacecraft (s/c) attitude
  - Portion of visible (=not Earth-blocked) sky
  - Positions (wrt s/c frame) of Crab, CygX-1, Sun (hard X-ray bright sources)
  - Vertical geoagnetic cutoff rigidity and geomagnetic coordinates of s/c
  - Epoch
  - Particle Monitor data
- **Orbit grouping based on s/c attitude**:
  - Composition of a number of separate "**orbit families**"
  - Apply supervised ML regression within each family indepdendently of one another

Several consecutive orbits



Single orbit



Creation of Families

- Each family has a father OP

Creation of Families

- Each family has a father OP
- All members have pointing differing by < 30 degrees from the father (axis by axis)

Creation of Families

- Each family has a father OP
- All members have pointing differing by < 30 degrees from the father (axis by axis)
- Fathers chosen so as to maximise mutual distances (least anisotropic grouping = optimal sky coverage)

Creation of Families

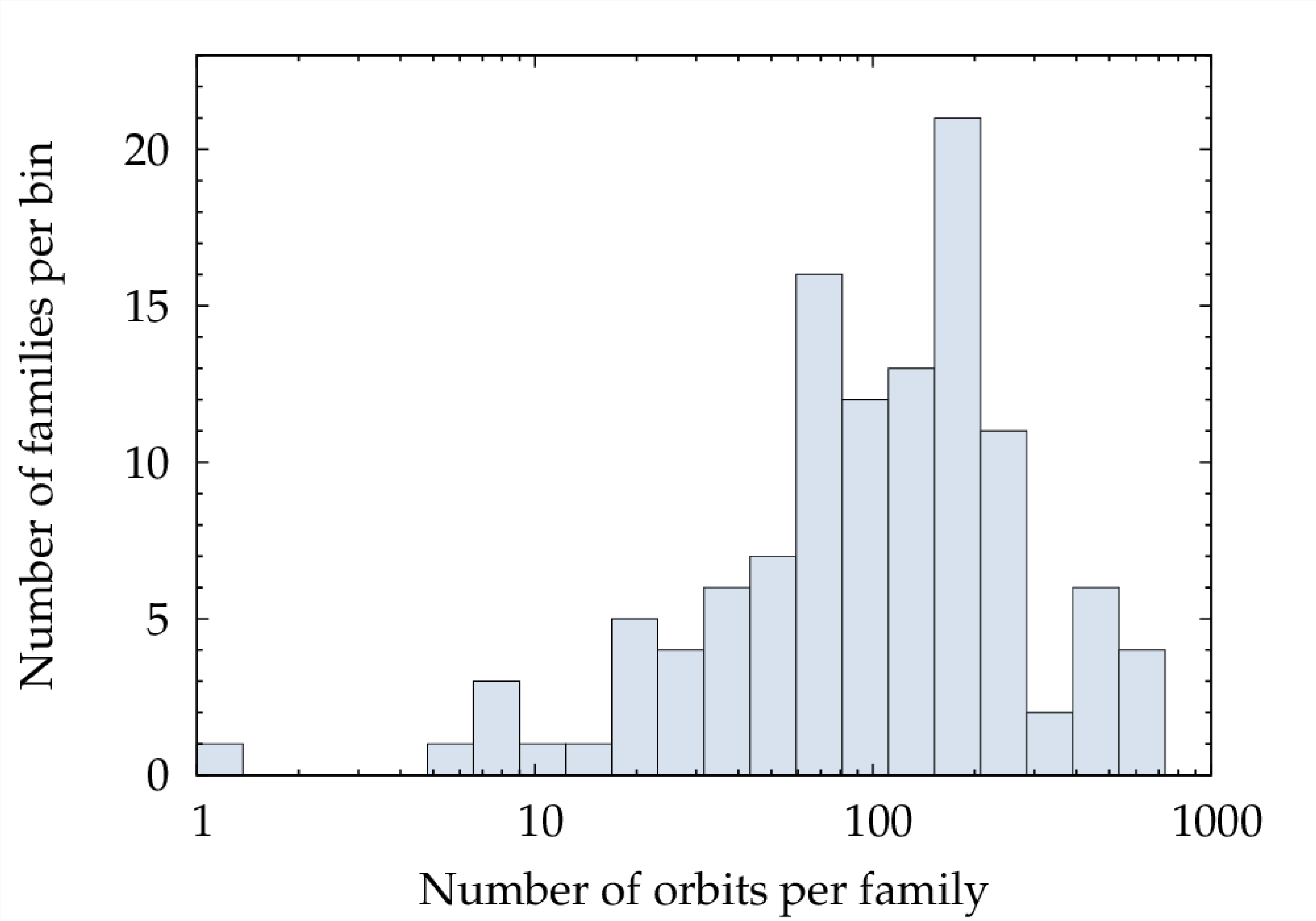
- Each family has a father OP
- All members have pointing differing by < 30 degrees from the father (axis by axis)
- Fathers chosen so as to maximise mutual distances (least anisotropic grouping = optimal sky coverage)
- Results:
    - 115 families
    - 17072 good orbits (after selection)

Creation of Families

- Each family has a father OP
- All members have pointing differing by < 30 degrees from the father (axis by axis)
- Fathers chosen so as to maximise mutual distances (least anisotropic grouping = optimal sky coverage)
- Results:
  - 115 families
  - 17072 good orbits (after selection)

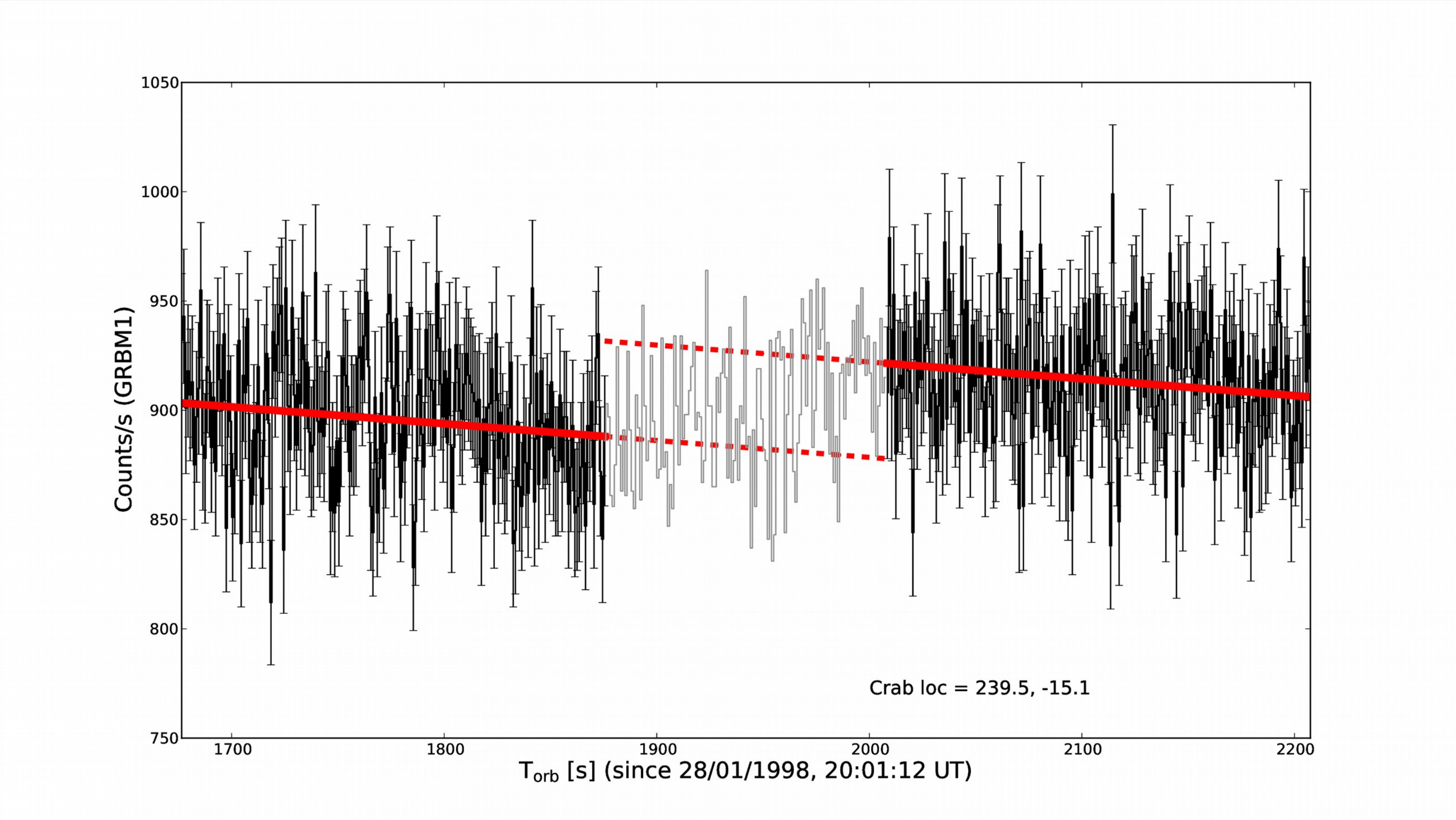
average = 148 orbits/family

Distribution of the number of orbits per family

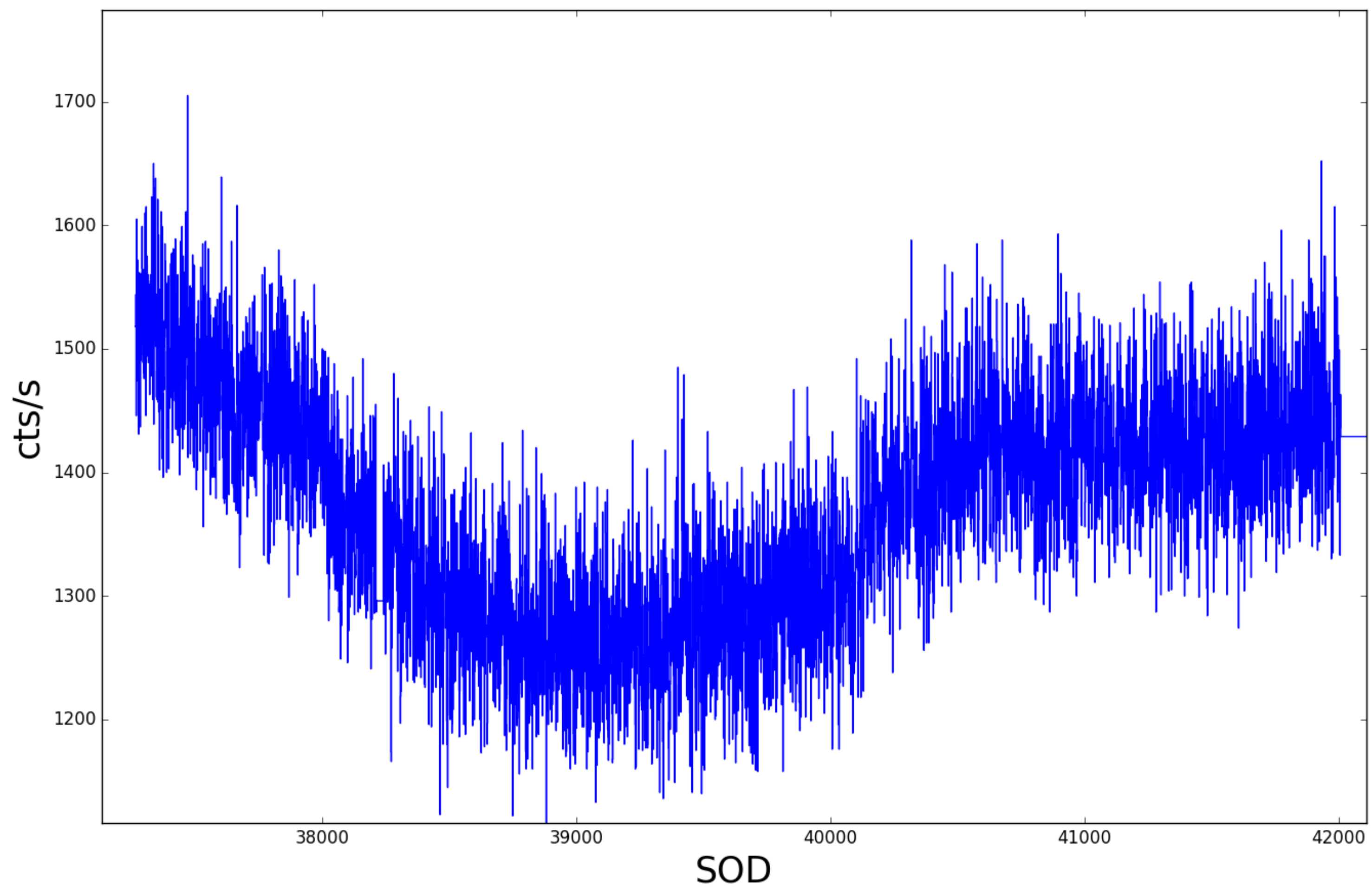


Bright sources removal: occultation technique

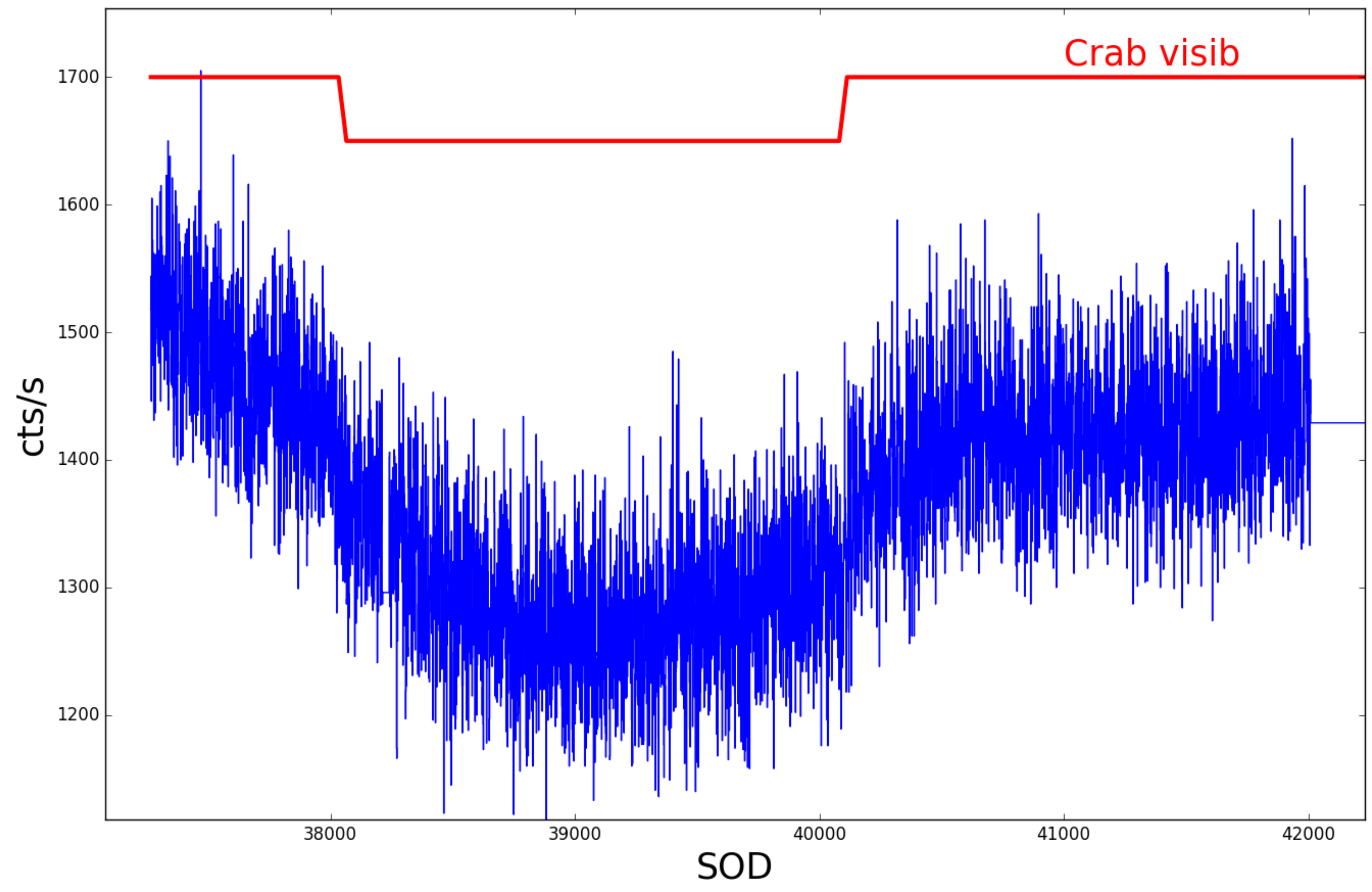
To ease comparison between different orbits, we preliminarily removed the contribution of Crab and of CygX-1 after estimating their rate with the occultation technique



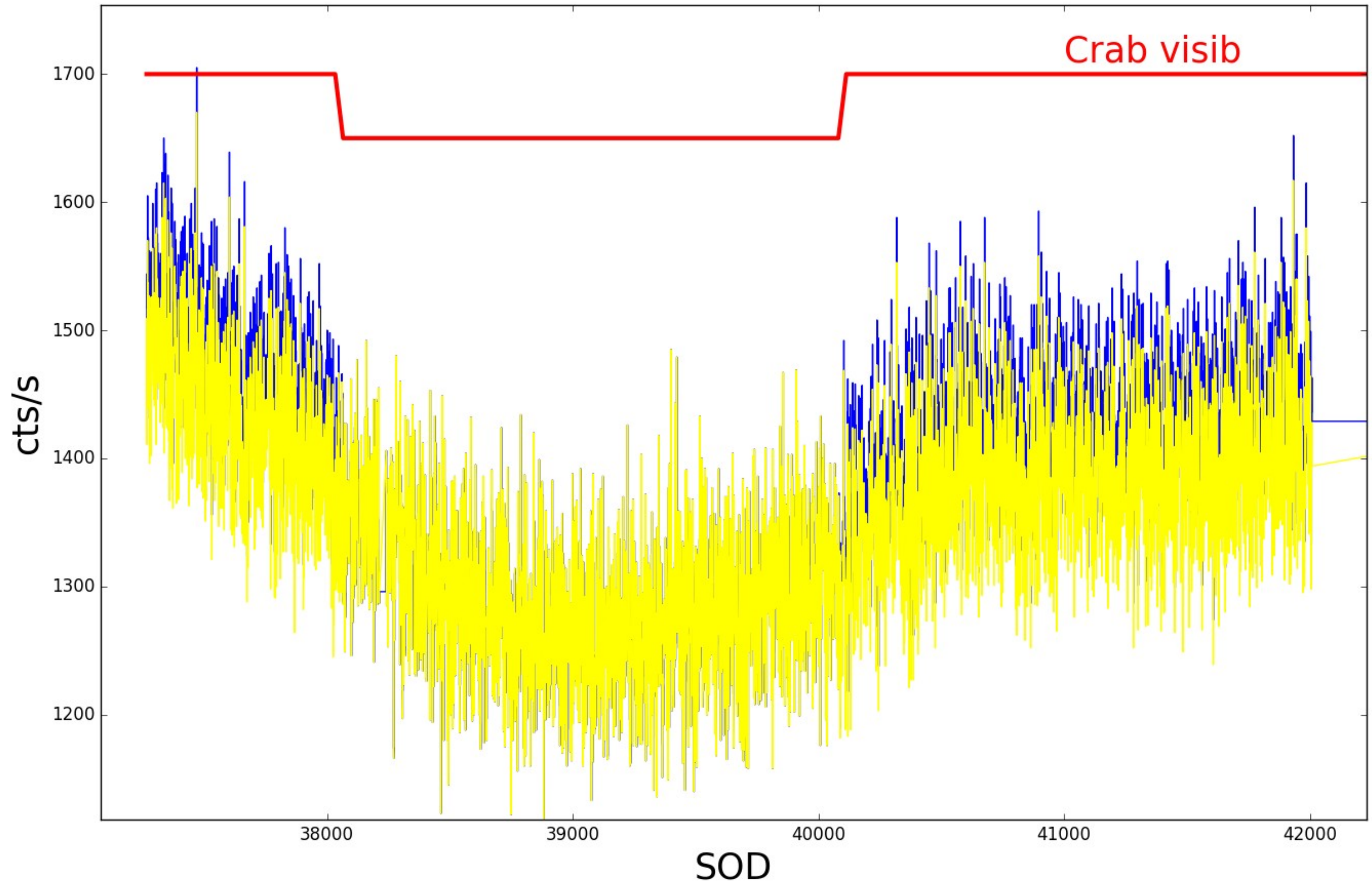
Occultation step: e.g. Crab



Occultation step: e.g. Crab



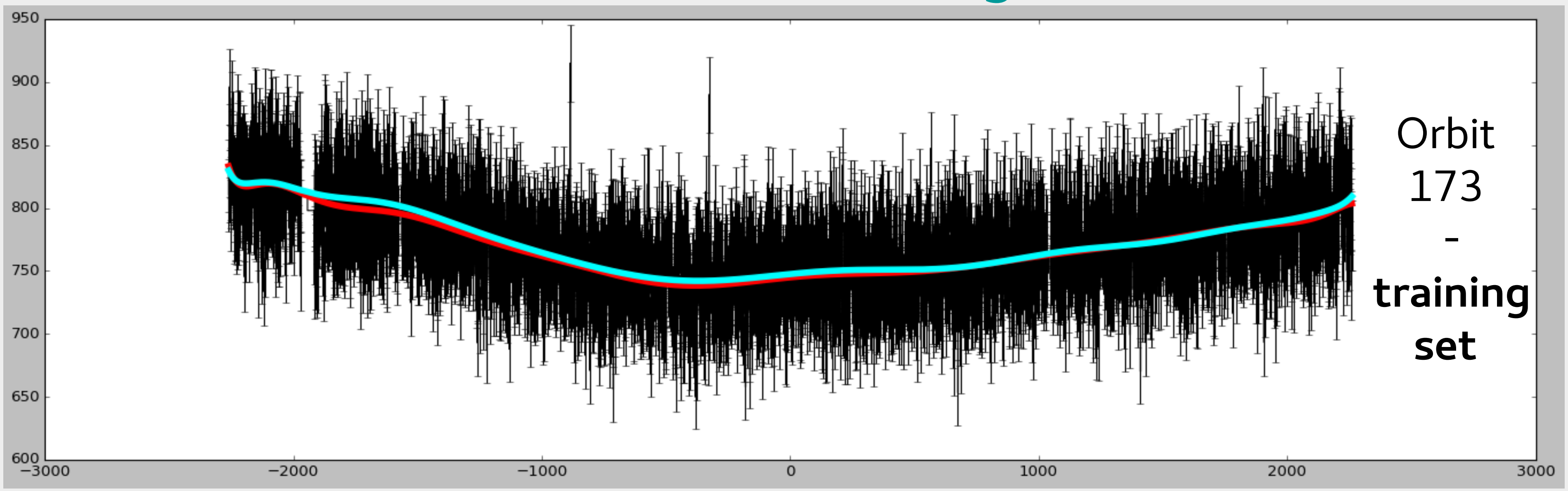
Occultation step: e.g. Crab

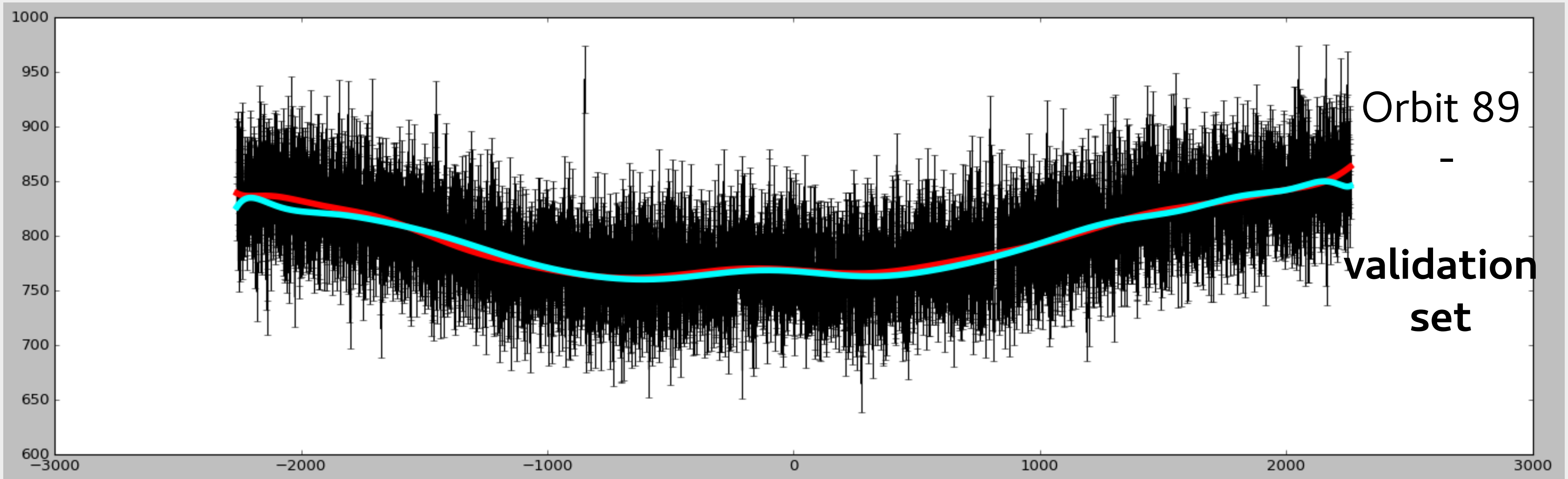


Preliminary results

Example of bkg modelling for generic orbits (randomly selected from Fam 88)

Smoothed LC ML-modelled bkg

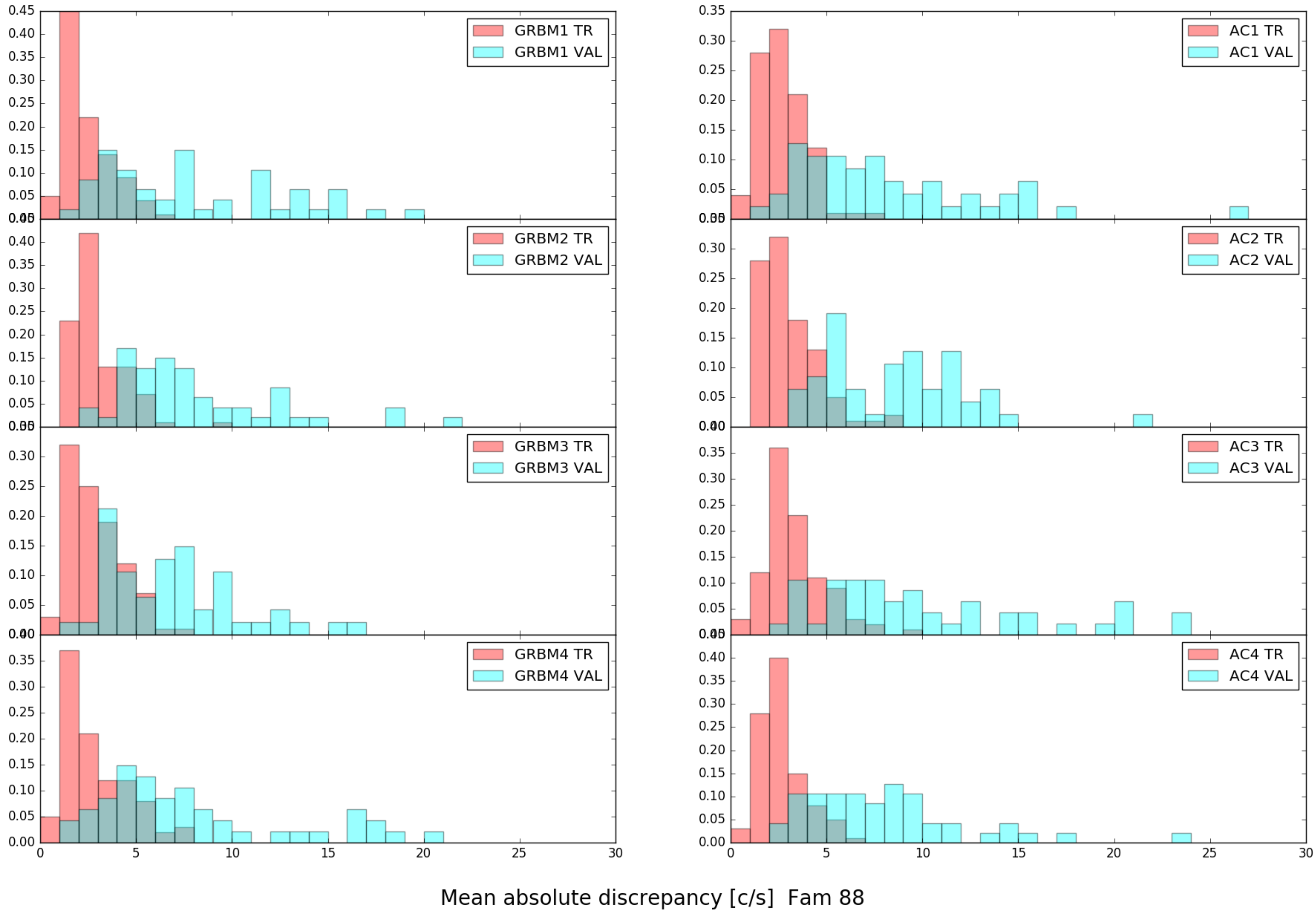




Example1: Training & Validation (orb Fam 88)

(N=207 orbits)

mean abs error distribution



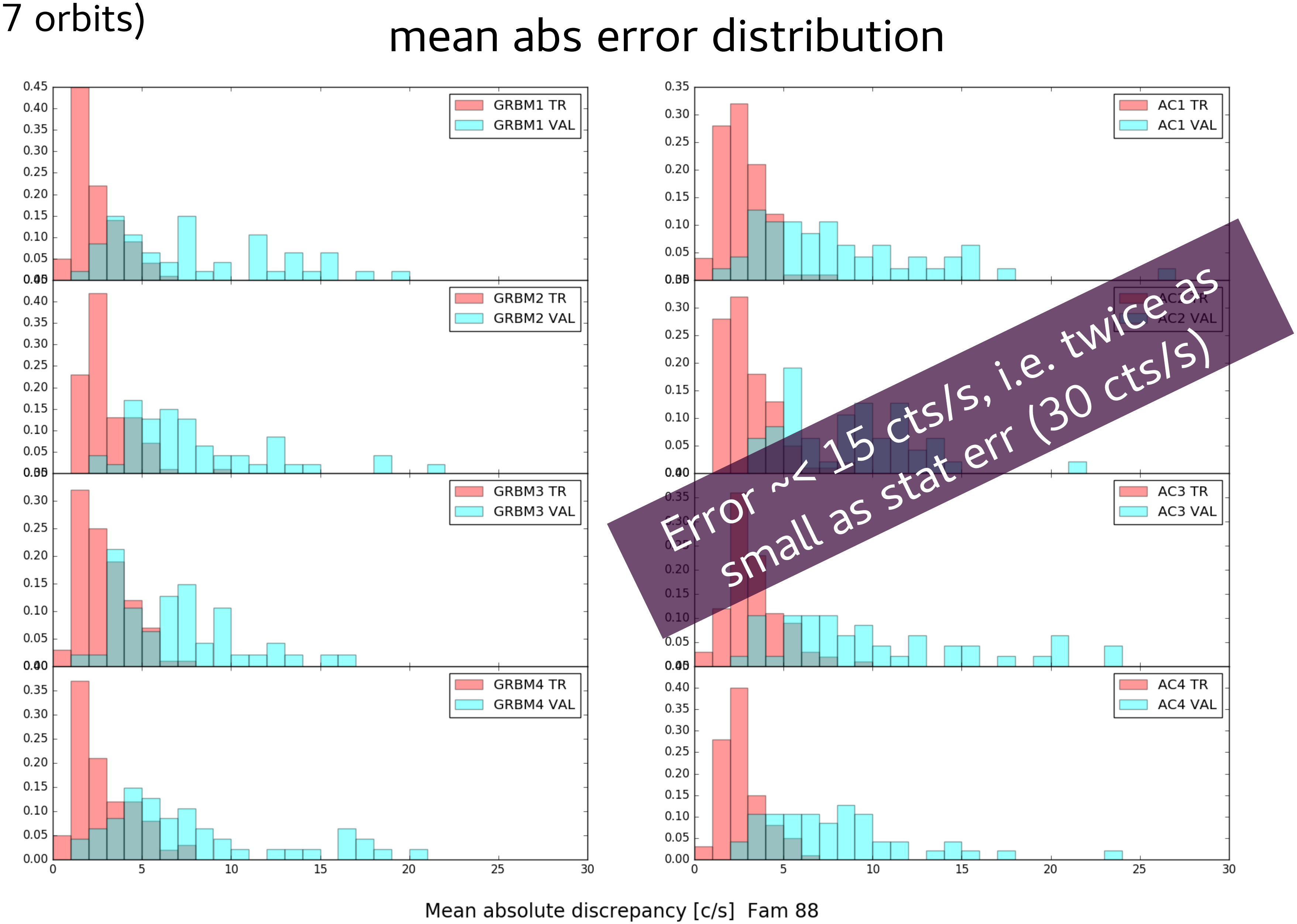
Example1: Training & Validation (orb Fam 88)

(N=207 orbits)

Training errors are < Val errors → slight overfitting

Example1: Training & Validation (orb Fam 88)

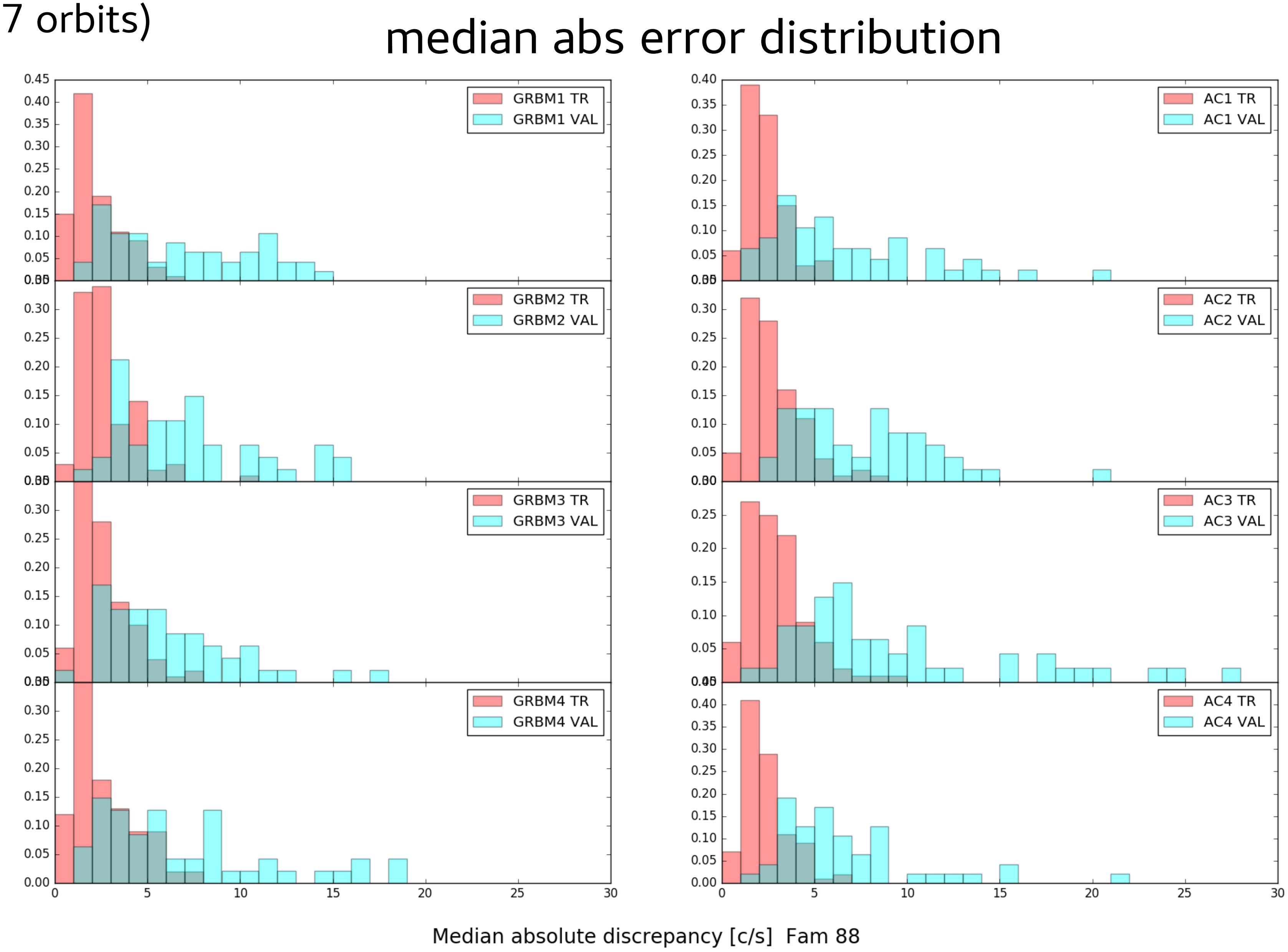
(N=207 orbits)



Training errors are < Val errors → slight overfitting

Example1: Training & Validation (orb Fam 88)

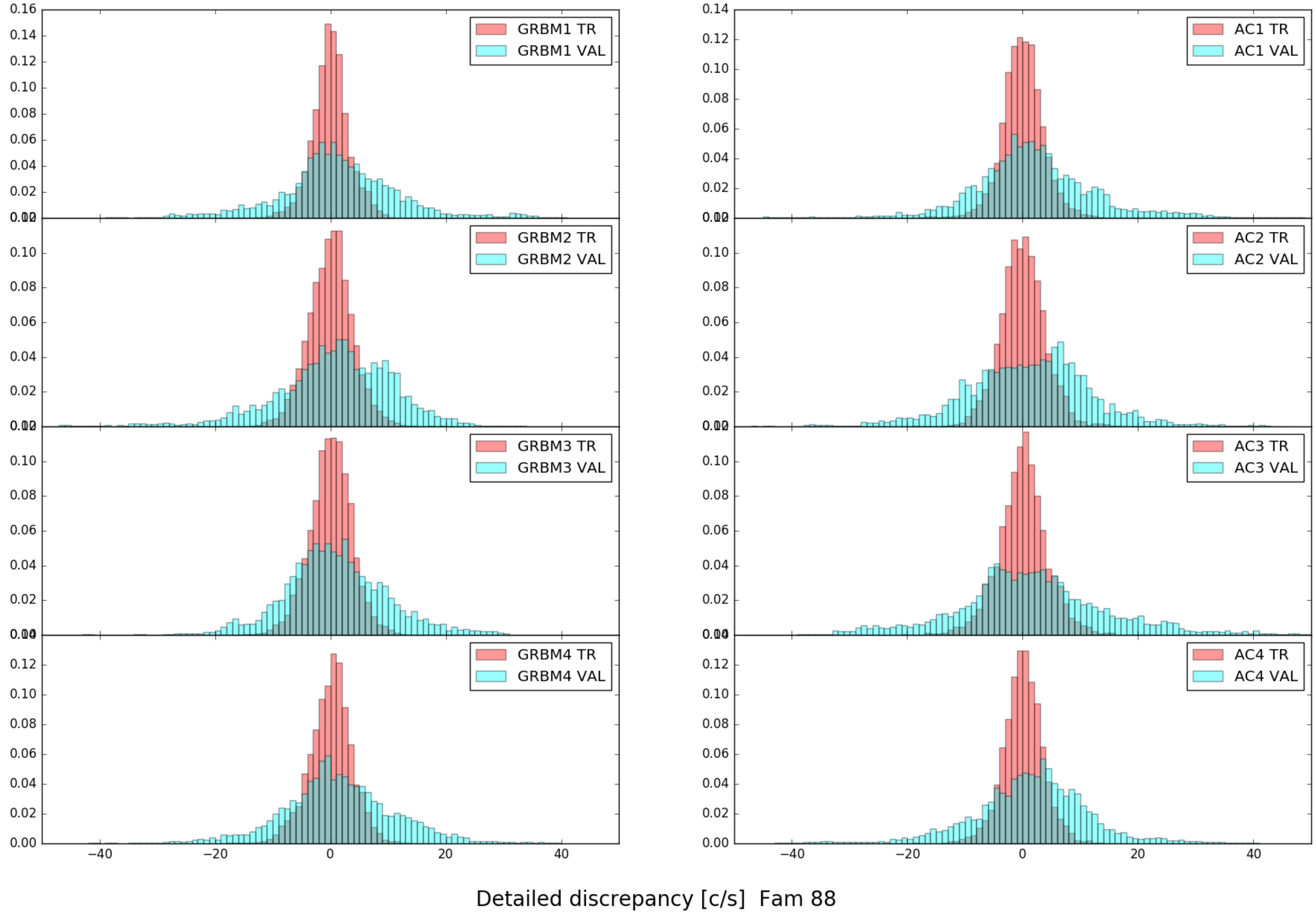
(N=207 orbits)



Example1: Training & Validation (orb Fam 88)

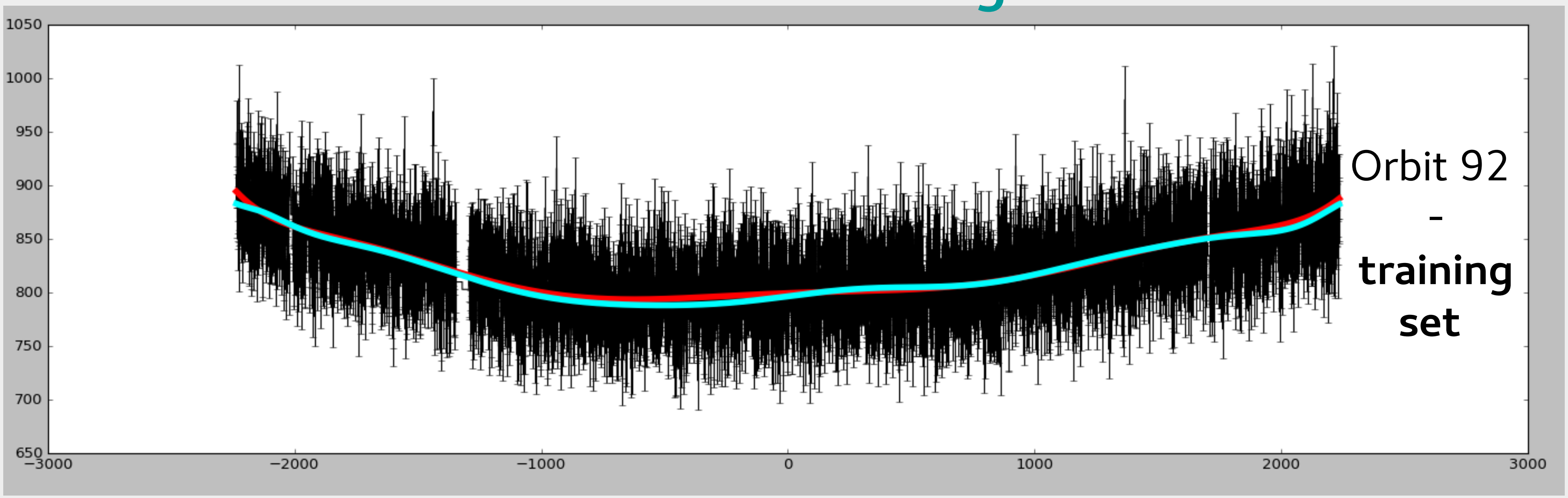
(N=207 orbits)

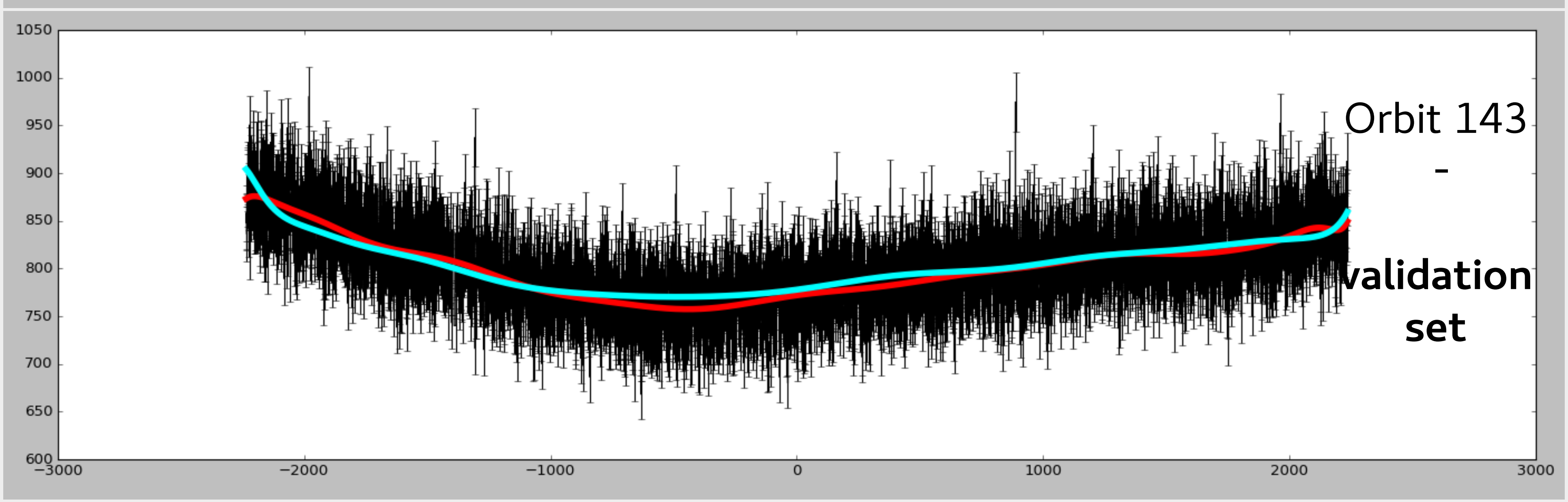
overall abs error distribution



Example of bkg modelling for generic orbits (randomly selected from Fam 78)

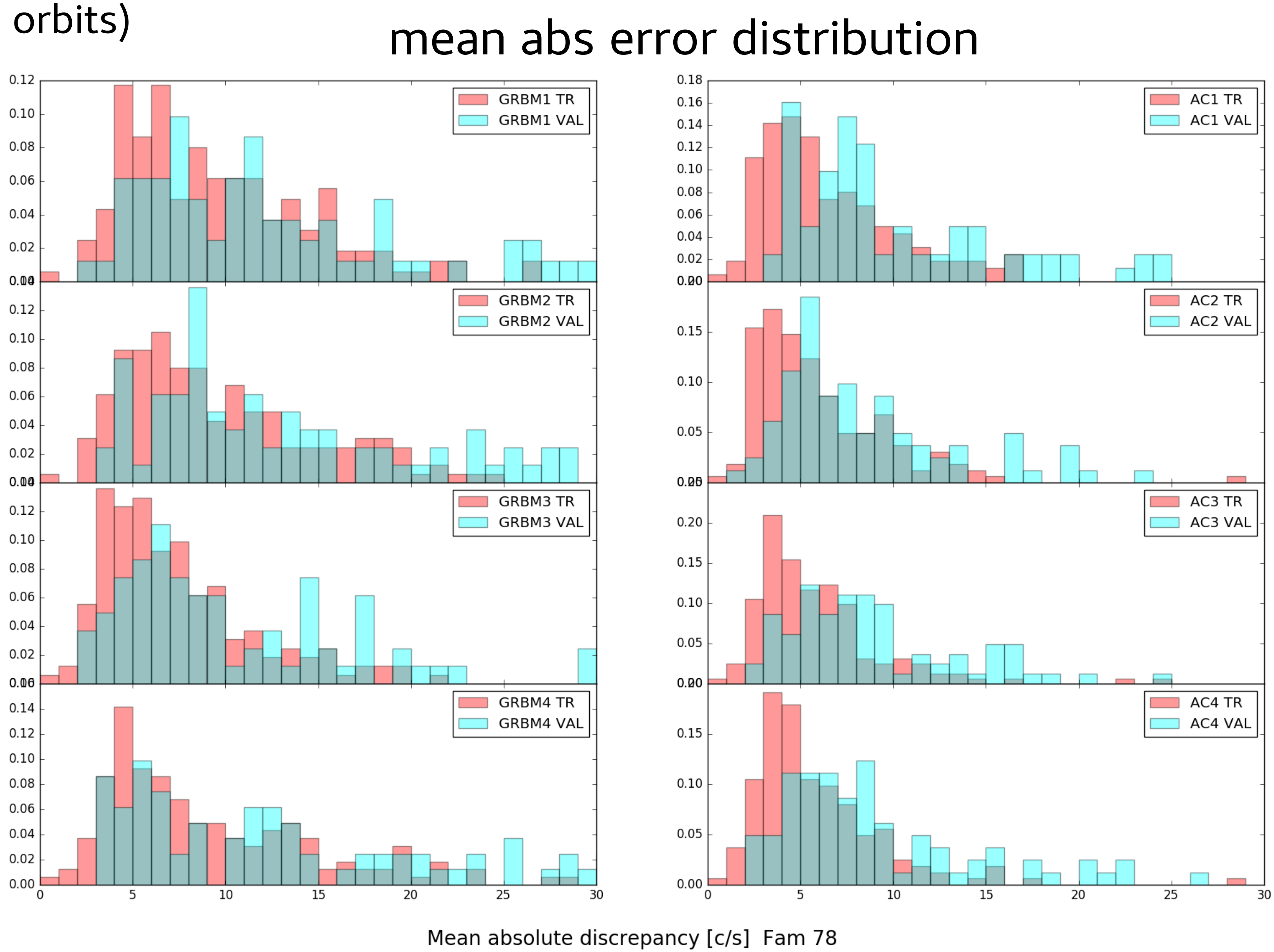
Smoothed LC ML-modelled bkg





Example2: Training & Validation (orb Fam 78)

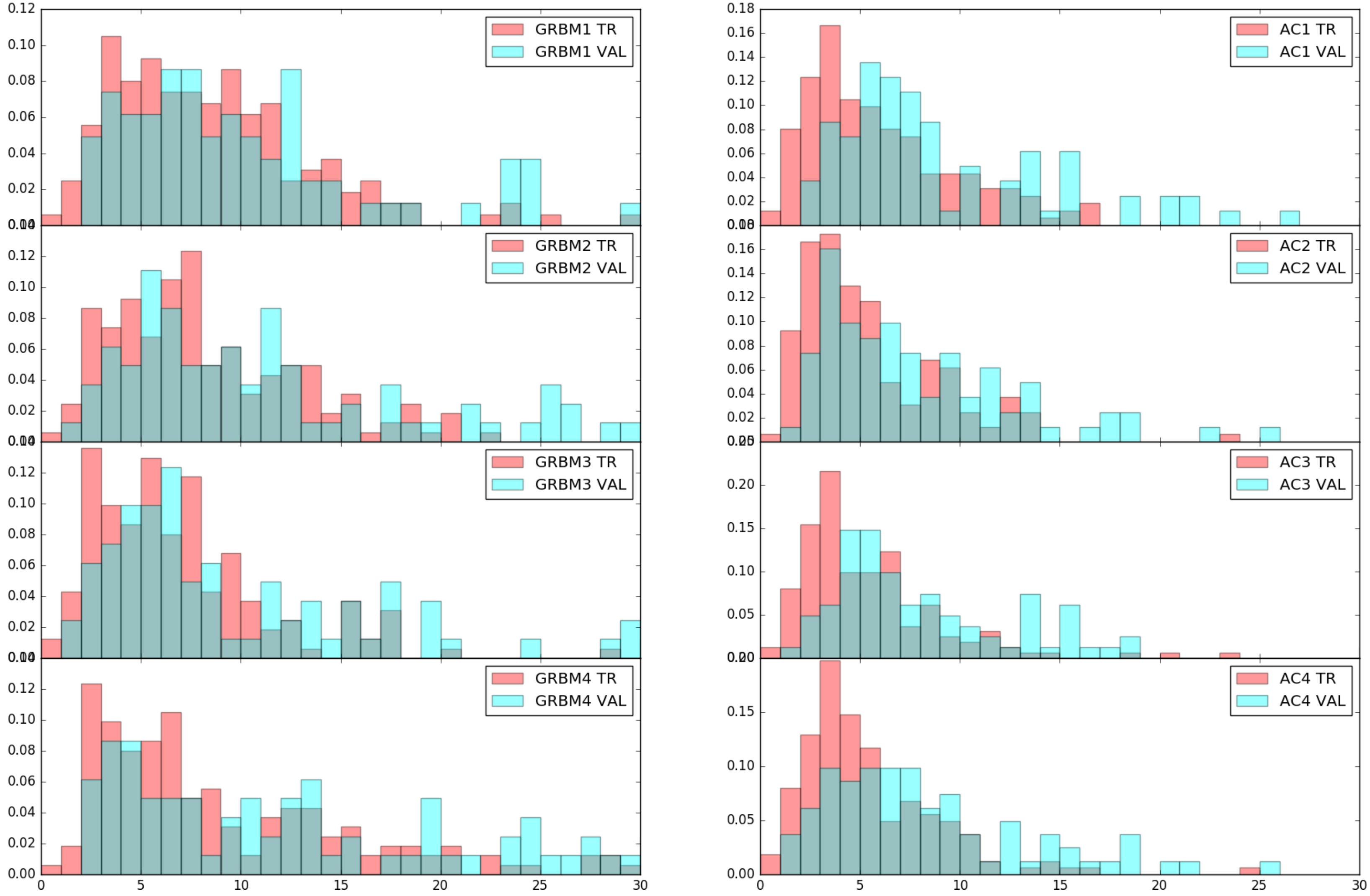
(N=332 orbits)



Example2: Training & Validation (orb Fam 78)

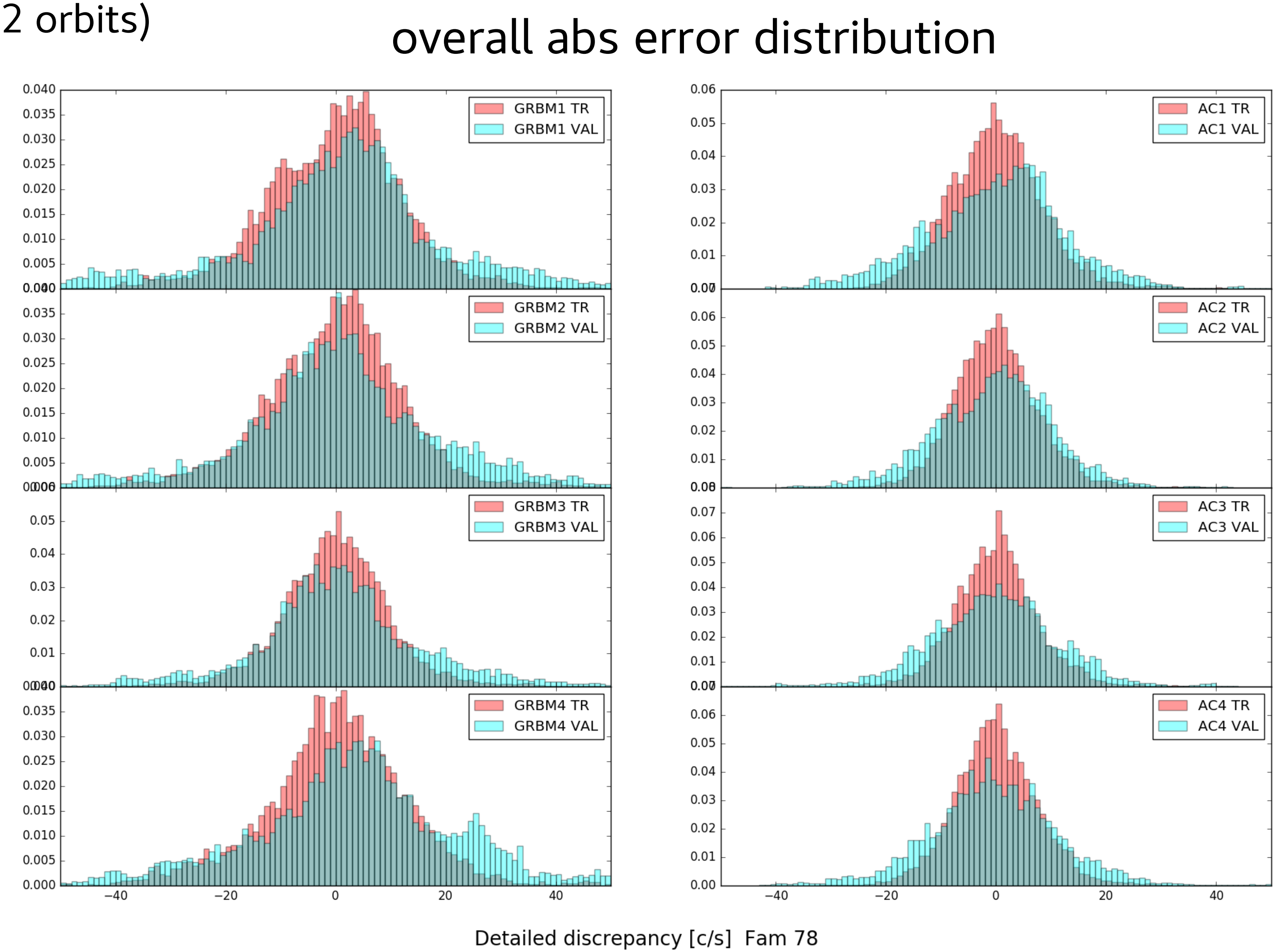
(N=332 orbits)

median abs error distribution



Example2: Training & Validation (orb Fam 78)

(N=332 orbits)



Summing up

- A ML linear regression approach looks promising. We're still in the process of refining the supervised learning tuning the set of attributes that play a role.

Summing up

- A ML linear regression approach looks promising. We're still in the process of refining the supervised learning tuning the set of attributes that play a role.
- Ultimate goal: systematic search for weak, slow-varying events to constrain low-lum GRB rate (e.g. 060218, 100316D, 171205A; paper in prep)

Summing up

- A ML linear regression approach looks promising. We're still in the process of refining the supervised learning tuning the set of attributes that play a role.
- Ultimate goal: systematic search for weak, slow-varying events to constrain low-lum GRB rate (e.g. 060218, 100316D, 171205A; paper in prep)
- Search for Fast Radio Burst putative counterparts

Summing up

- A ML linear regression approach looks promising. We're still in the process of refining the supervised learning tuning the set of attributes that play a role.
- Ultimate goal: systematic search for weak, slow-varying events to constrain low-lum GRB rate (e.g. 060218, 100316D, 171205A; paper in prep)
- Search for Fast Radio Burst putative counterparts
- Extract complete LC (40-700 and >100 keV) of CygX1 and study variability

Summing up

- A ML linear regression approach looks promising. We're still in the process of refining the supervised learning tuning the set of attributes that play a role.
- Ultimate goal: systematic search for weak, slow-varying events to constrain low-lum GRB rate (e.g. 060218, 100316D, 171205A; paper in prep)
- Search for Fast Radio Burst putative counterparts
- Extract complete LC (40-700 and >100 keV) of CygX1 and study variability
- Once the procedure will be optimised on BSAX data, a similar approach can be applied to HXMT data, building upon the knowledge of the role(s) played by the different attributes.

Thank you 谢谢

References

- Guidorzi et al., in prep
- Frontera et al. 1997, A&A Suppl. S., 122, 357
- Bishop, “Pattern Recognition and Machine Learning”, Springer
- Murphy, “Machine Learning: A Probabilistic Perspective”, MIT Press

Back-up Slides