

Offline/MC status and issues

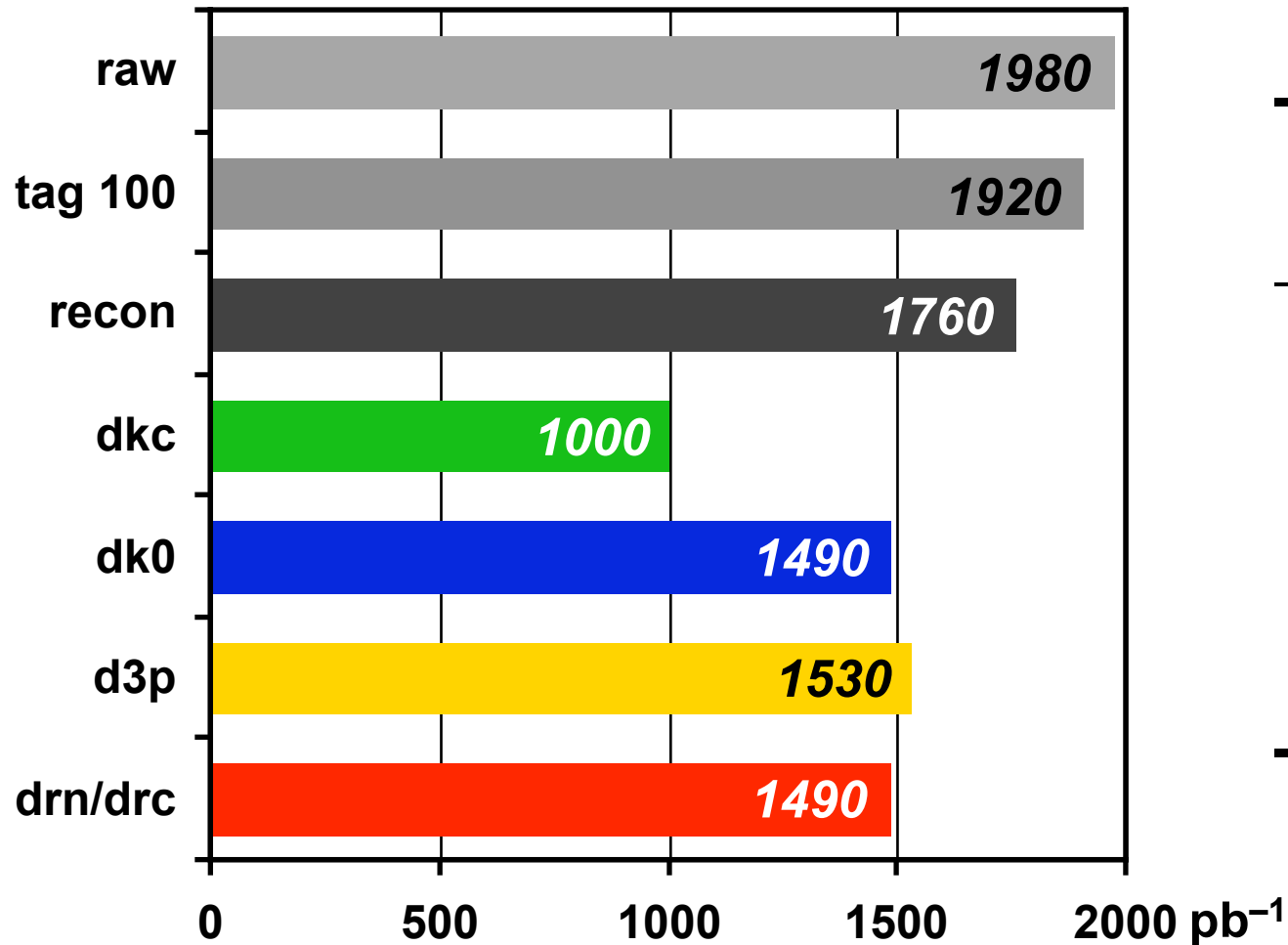
Matthew Moulson & Graziano Venanzoni

KLOE General Meeting

Frascati, 13 December 2005

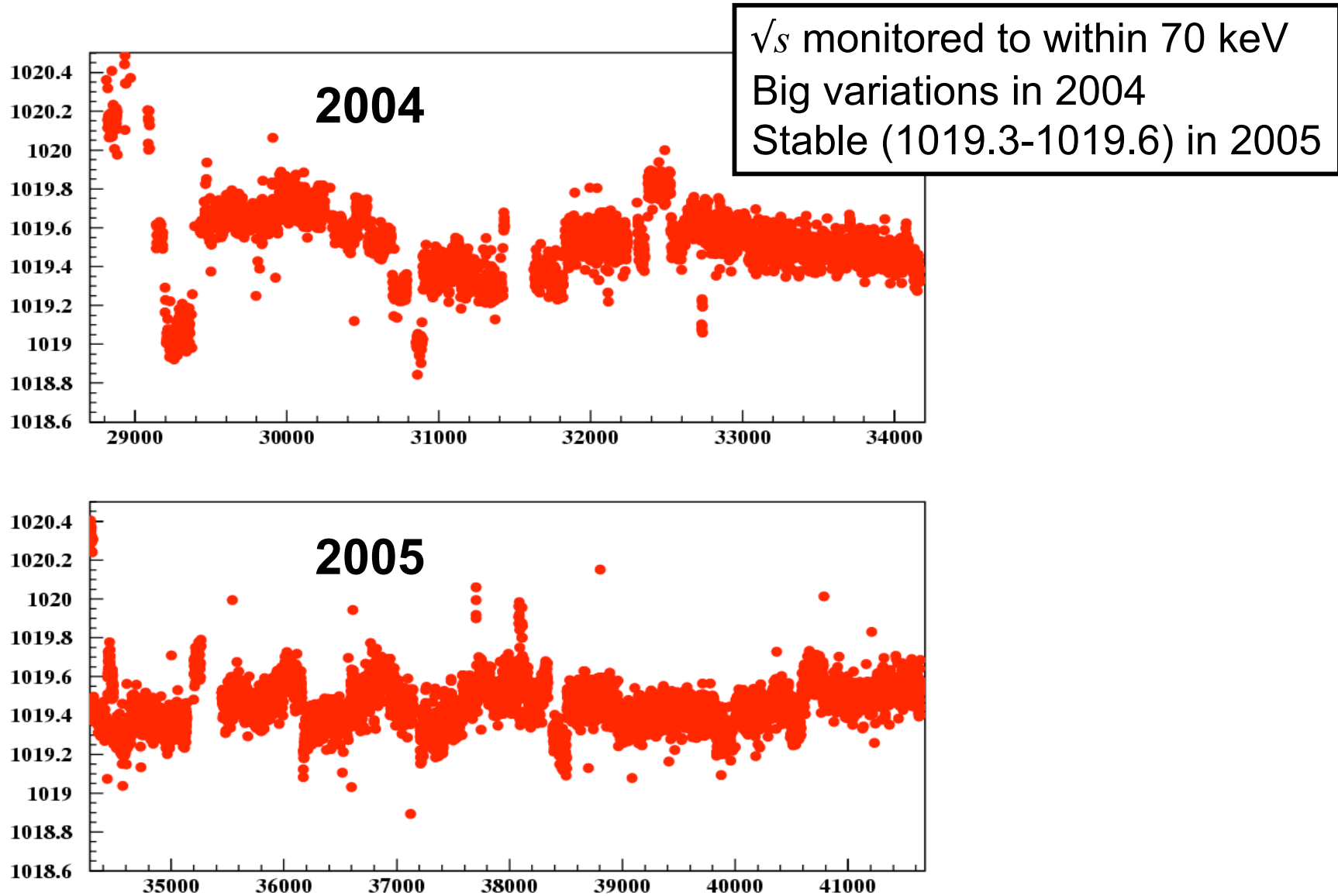
Reconstruction of 2004-2005 data

Runs 28700 (9 May 04) to 41902 (5 Dec 05)

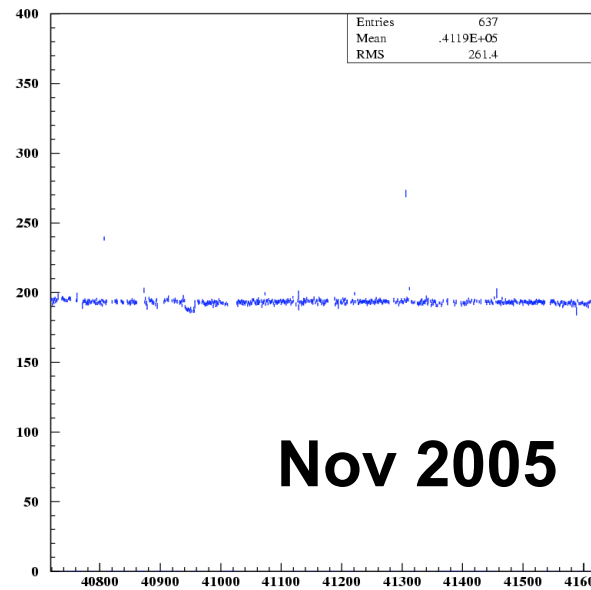
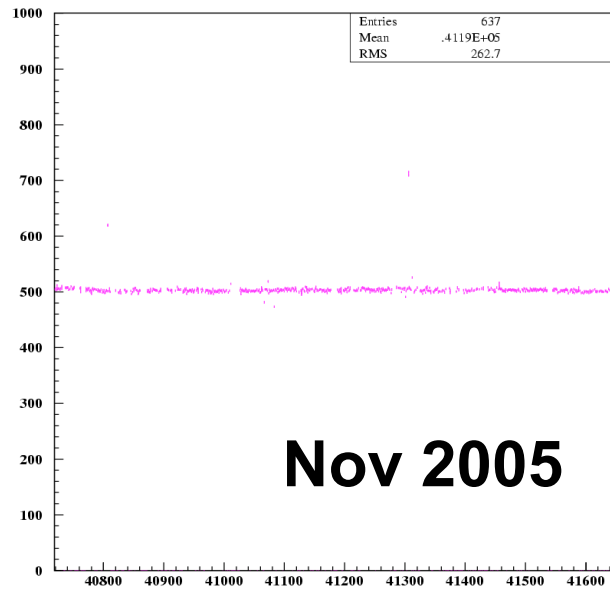
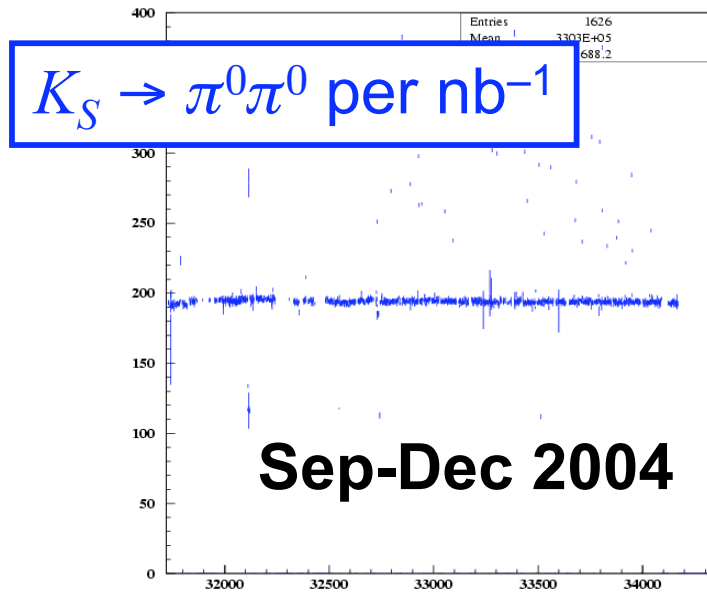
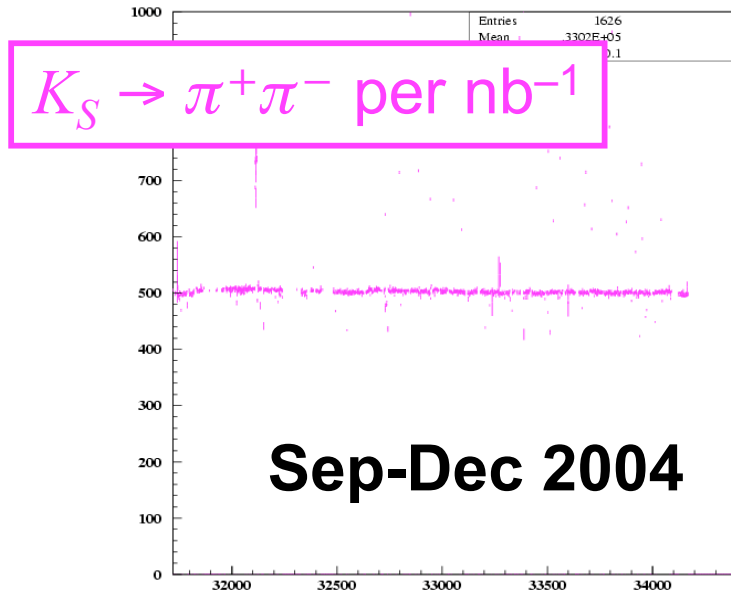


$\int L dt$ (pb ⁻¹)	
tag 100	
2001	168
2002	289
2004	687
2005	1237
total	2381

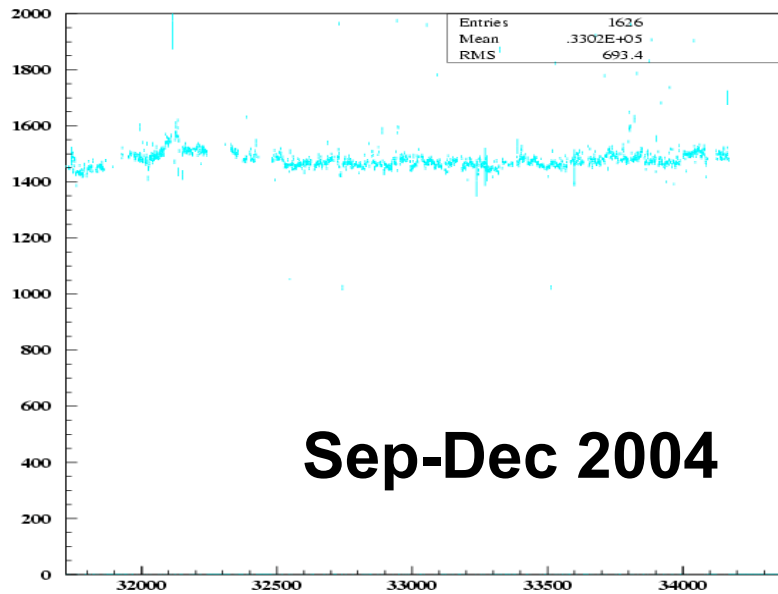
Machine energy stability



Reconstructed event rate stability

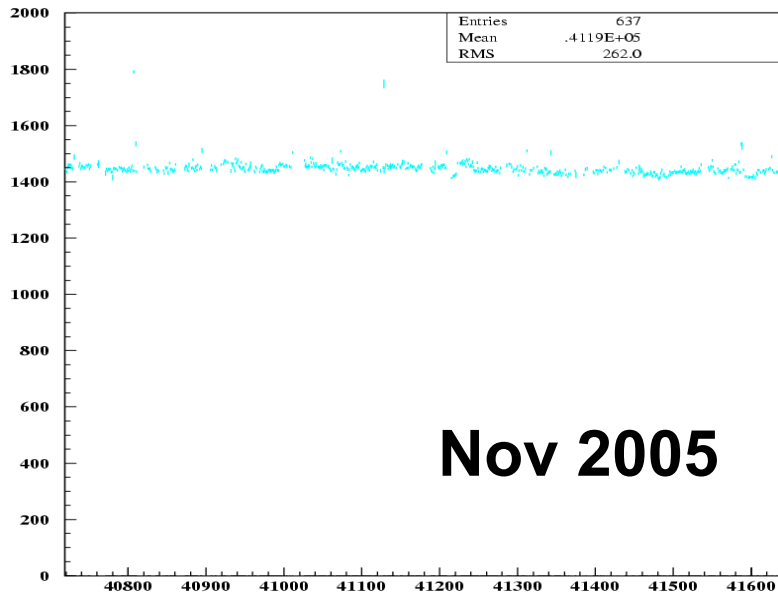


Reconstructed event rate stability

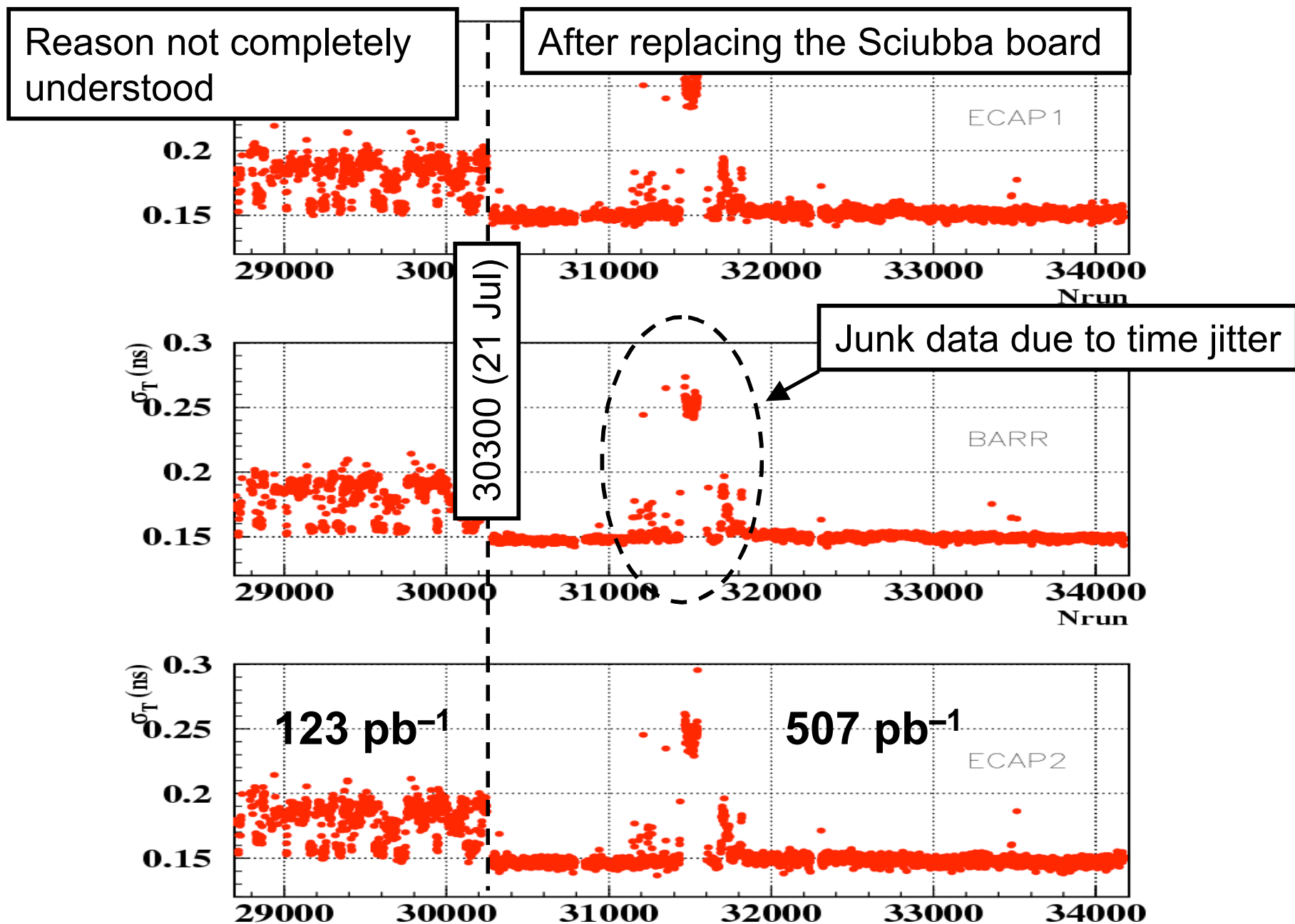


$$\phi \rightarrow K^+ K^- \text{ per nb}^{-1}$$

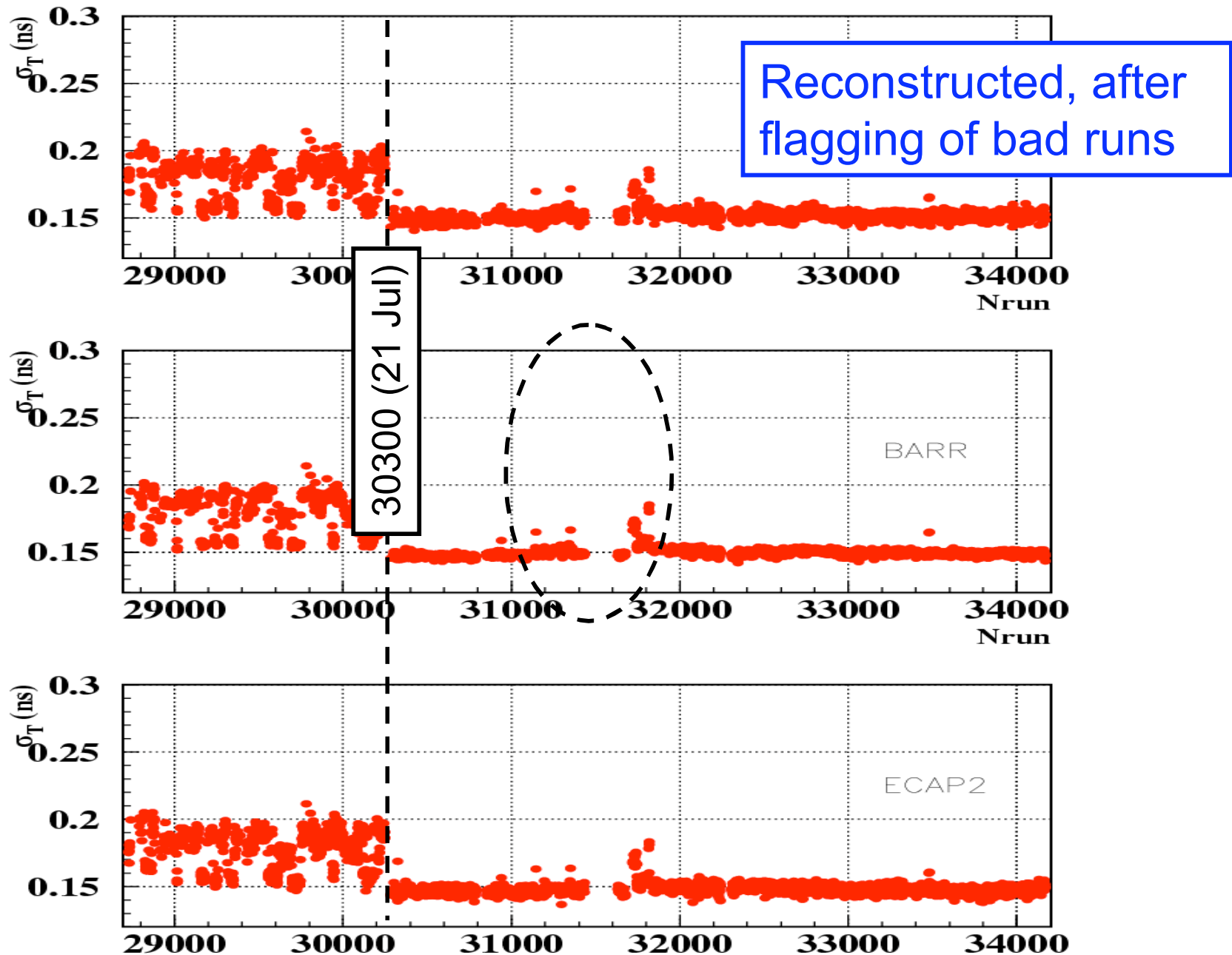
K^\pm counters are more sensitive to machine background



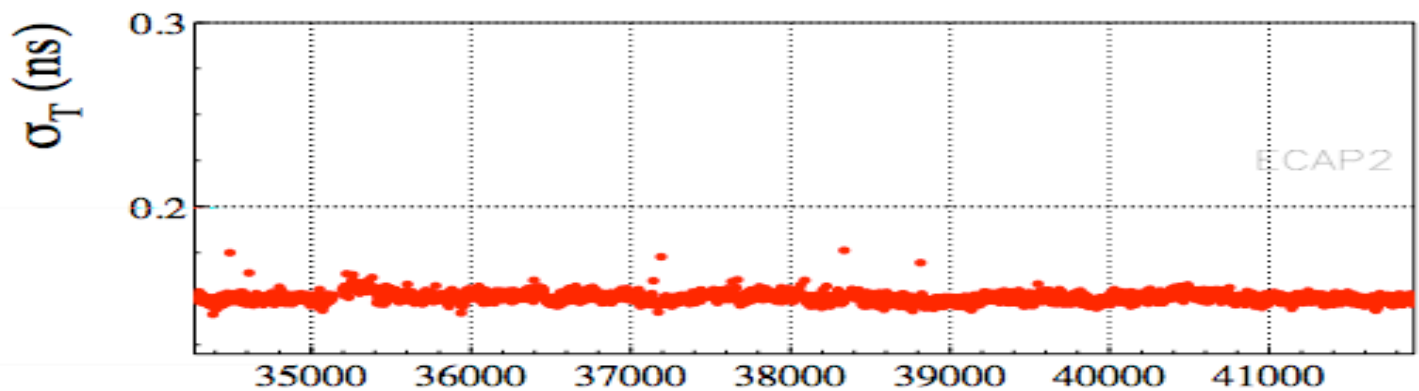
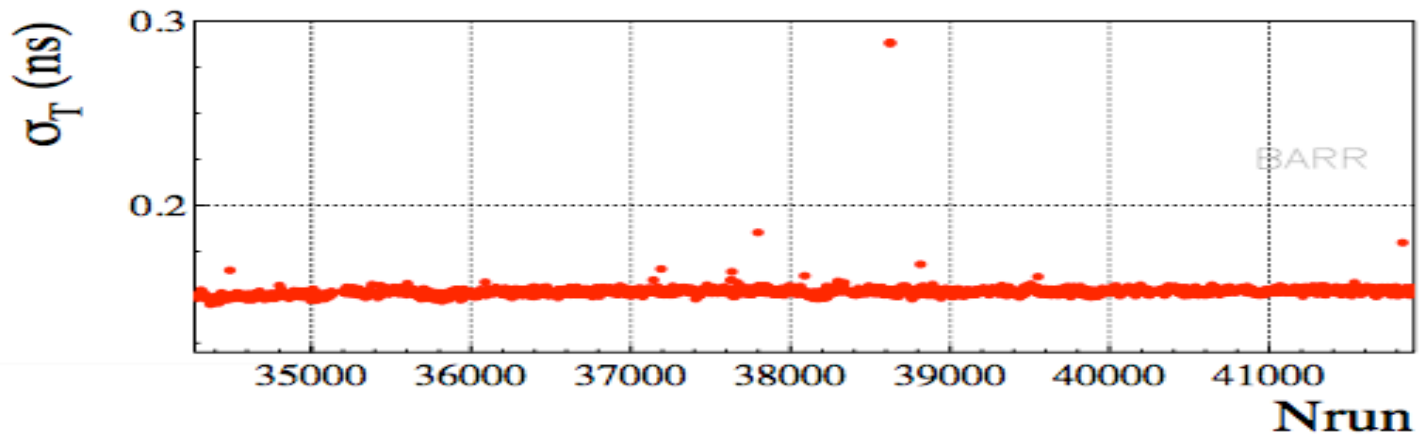
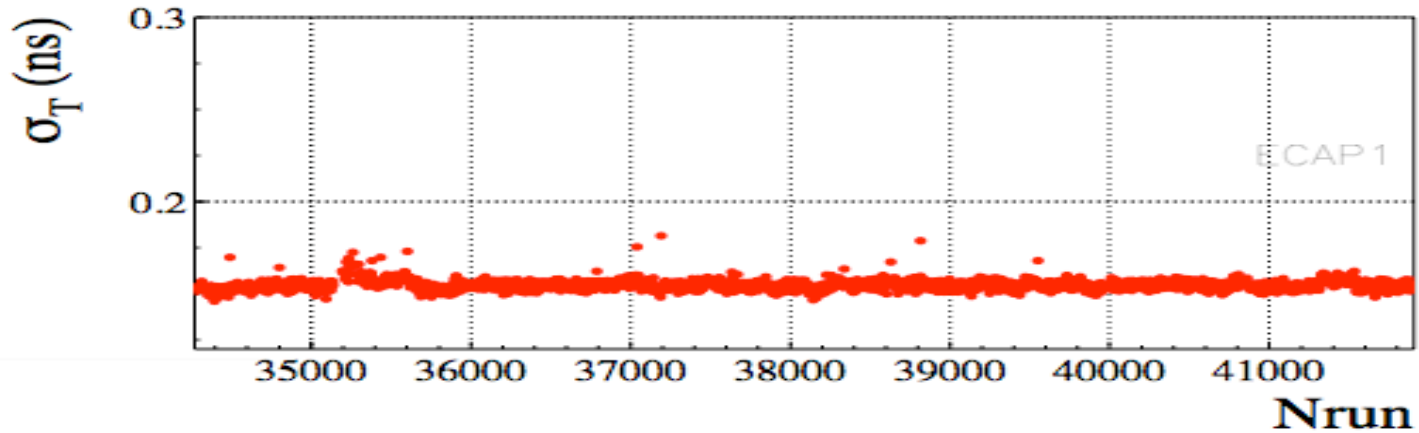
2004 EmC time resolution vs run



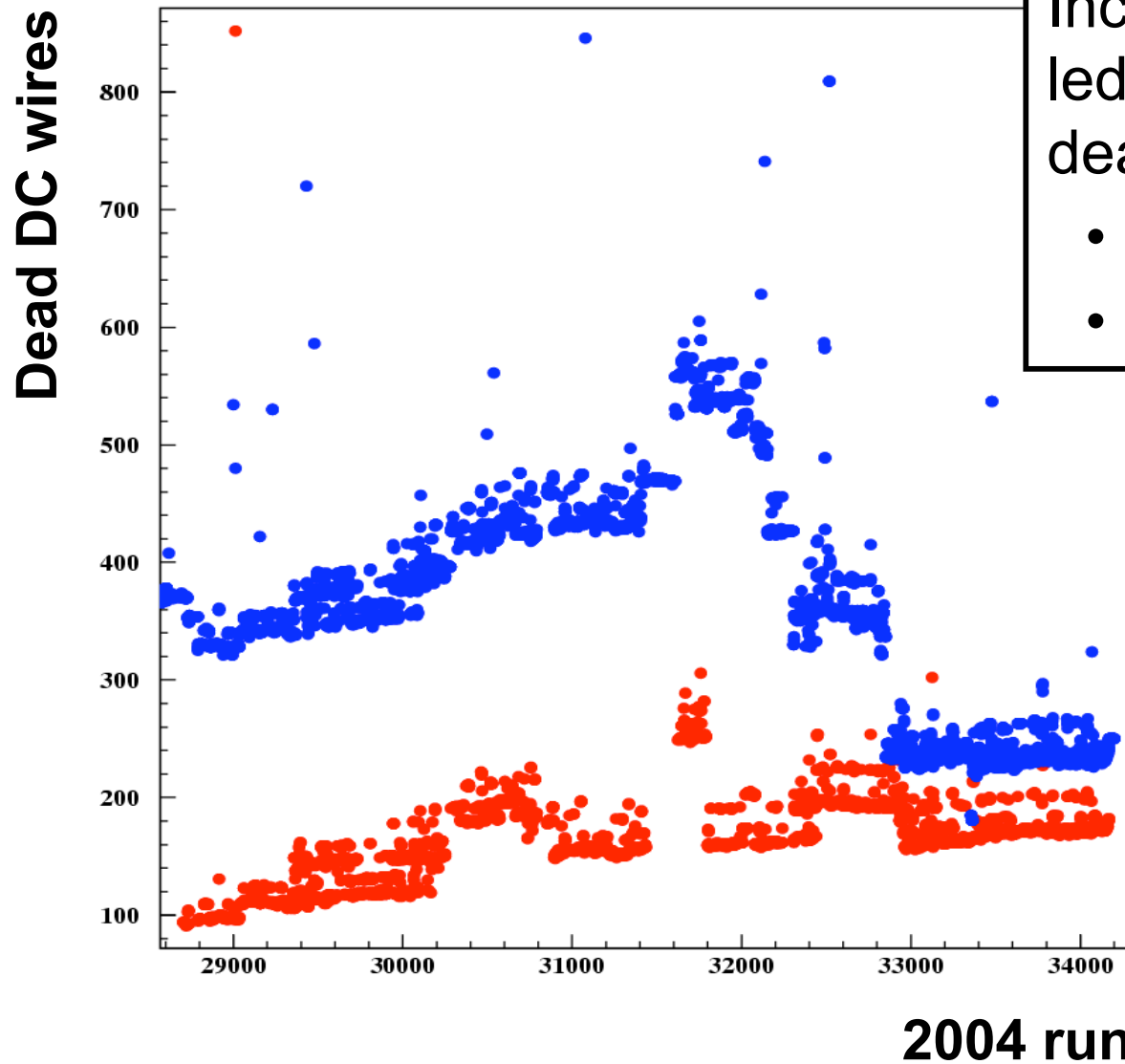
2004 EmC time resolution vs run



2005 EmC time resolution vs run



DC wire maps and efficiencies in DB2



Incorrect DB loading led to falsely high dead-wire counts for:

- 2001-02 MC recon
- 2004 data recon

- Before correction
- After correction

A photograph of a sunset over a body of water. The sun is low on the horizon, creating a bright orange and yellow glow that reflects on the water's surface. The sky transitions from a deep orange near the horizon to a darker, reddish-purple at the top. The water in the foreground is dark with small ripples.

The beginning of the end?

**Need to start thinking about closing out
offline activities and leaving a legacy data set**

Offline/MC tasks for 2006

Reconstruction of 2006 data ($\sqrt{s} = 1000$ MeV)

Data quality and database maintenance

Make sure all reconstructed runs are complete

Make sure bad runs flagged and reconstructed files deleted

Close holes in data reconstruction and DST coverage

DST file size problem = 1 week of scripting work

MC production for 2004-2005

Reprocessing of 2004 data

Is there any other data that still needs reprocessing?

Re-reconstruction of 2001/2002 MC sample

Still interested? 30 pb⁻¹? More?

Improvements to KLOE Monte Carlo

Main improvements for simulation of 2004-2005 data

- Map of 2004 machine and trigger conditions
2005 in progress
- New IR geometry in simulation
- Better parameterizations of EmC response
Time, energy resolution; cluster efficiencies
- Improved simulation of nuclear interactions/regeneration
in DC wall and beam pipe
- New secondary decay generators
 $K_S \rightarrow \pi^+\pi^-e^+e^-$; $\pi \rightarrow e\nu$, $\pi \rightarrow e^+e^-$
- Inserted background from randomly triggered events
- Simulation of dE/dx measurement in DC

Items for production test: Summary

Item	Develop	Implement
New IR geometry	CB	CB
Attenuation lengths in EmC endcaps	PG	CB MM
2002 EmC time resolution adjustments	CG	MM
2002 cluster efficiency parameters	MP TS	MM
Nuclear interactions/regeneration	CB	CB
New secondary decay generators	Various	CB
LSB (background) insertion	MM	MM
SQZ (compression) fix		MM
Cell/cluster index fix		MM
dE/dx simulation	VP	MM

Monte Carlo status

First production test 2 Nov:

25 pb⁻¹ all_phys, LSF = 0.2

Runs 37000-37200 (2005 data)

geanfi 1.7.4, datarec DBV-24

VERT banks corrupted, 1/4 endcaps missing!

Second production test 25 Nov, same sample:

Preliminary results look good, still waiting for feedback

Next step: Preproduction

100 pb⁻¹ all_phys, LSF = 0.2

- Test new cluster efficiency parameters
- Test new EmC time resolution parameters
- Additional EmC studies on energy scale calibration

Additional MC production

2002 MC produced for $\pi\pi\gamma$ group

geanfi 1.7.4, datarec DBV-24: *Caveat emptor!*

pho5mmg:

- $e^+e^- \rightarrow \mu^+\mu^-\gamma$ events with PHOKHARA 5
- 248.5 pb⁻¹ (2002 data)
- LSF = 6, 42M events

filt_3p:

- $\phi \rightarrow \pi^+\pi^-\pi^0$ events with generator-level filter to ensure taggable photon from π^0 at large angle
- LSF = 1 with respect to $\sigma(\phi \rightarrow \pi^+\pi^-\pi^0)$, but filter gives 7× increase in effective statistics
- Not yet started

Items for development: Summary

Item	Develop	Implement
2004 data quality scan (\sqrt{s} , etc)	GV	
2004 trigger: quality, DC thresh, DISH maps	MP BS	
2005 data quality scan (\sqrt{s} , etc)	GV	
2005 trigger: quality, DC thresh, DISH maps	MP BS	
Check EmC energy scale calibration	PG	
Check EmC time resolution parameters	CG	
Check cluster efficiency parameters	TS	
DC tracking resolution: s - t relations	AA PDS MM TS	
DC tracking resolution: MS from wires	SK	
Correlated noise for charged kaons	EDL PDS	
$\phi \rightarrow \rho\pi$ generator	FN	CB

Monte Carlo production plans

Averaged over entire MC sample: 0.21M evts/B80 day = 2.4 Hz
0.41 s/evt (simulation + reconstruction + DST)

2001-2002 MC production

$\phi \rightarrow$ all, scale = 0.2

$K_S K_L$, scale = 1

$K^+ K^-$, scale = 1

ϕ radiative, scale = 5

Other (1M evts/pb⁻¹)

Total: 3.1M evts/pb⁻¹

(about same as number of
 ϕ decays in data)

Estimated time for 2004-2005 MC

2001-2002 450 pb ⁻¹	2004-2005 2000 pb⁻¹
1.85G evts 8800 B80 days	8.25G evts 39000 B80 days

KLOE reconstructed data

DBV	2001	2002	2004	2005	Comment
12-15	161	34			Bad cut in FILFO No bias sample for FILFO No bias sample for rad
19-21			634 bad wires!	225	Various significant ECL mods Minor changes to recon No bias sample for rad
22		236		394	Stable ECL Old bias sample for rad
23+		13		499	Stable ECL ufo as bias sample
analyzed	161	283	634	1118	
total	168	289	687	1237	

Offline resources: CPU

2006 offline projects	B80 days
Online reconstruction of 2006 data	3000
Reprocessing of 2004 data	16000
MC production for 2004-2005 data	39000
Reprocessing 01-02 data	5000
Reprocessing 02-04-05 DBV<23	22000
Minimum total	58000
Maximum total	85000

Current offline farm:

200 B80 CPUs (60 typically reserved for analysis)

Offline work completed in 14-20 months on 140 B80s

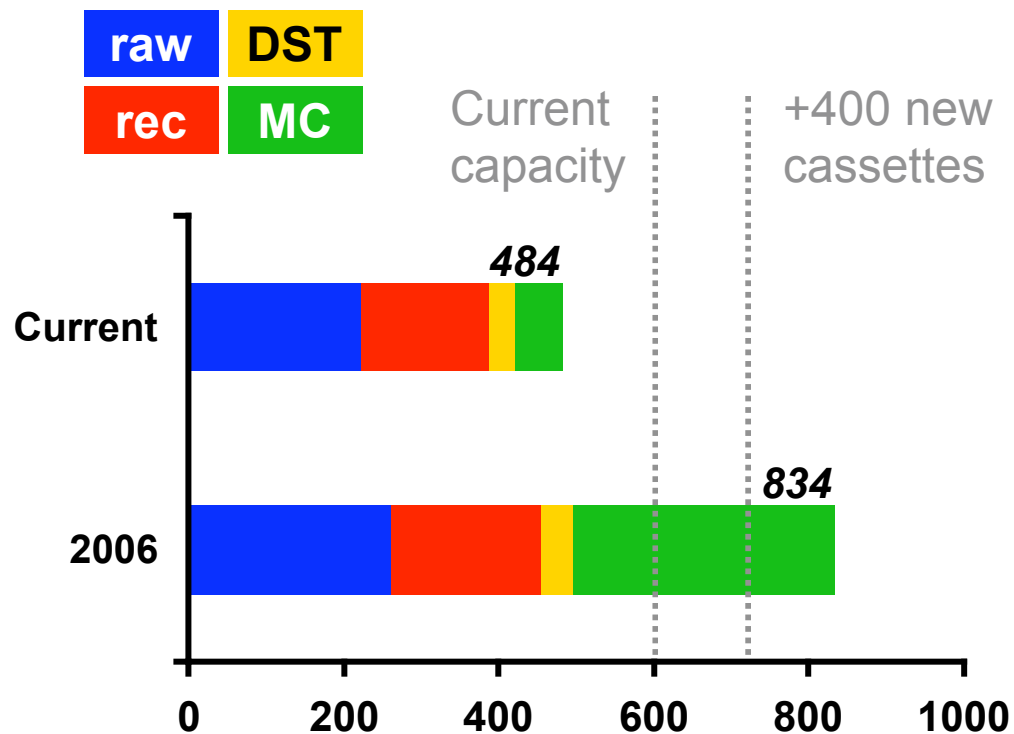
Expansion plans:

Ordering 3 16-way 1.5 GHz Power5 servers

Will add 200 B80 to offline farm → complete work in 7-10 months

Offline resources: Tape library

Tape library usage (TB)



2006 estimate includes:

- Closing of holes in datarec/DST coverage
- Off-peak running in 2006
- MC for '04-'05 data

Library capacity:

- Currently **600 TB**
- Expandable to **1400 TB**
(1000/3600 slots in use in new library)

Order ready for 400 new cassettes (120 TB)

Offline resources: Disk space

DSTs cached on nfs-mounted disks for fast analysis access

DST volume

Current	2006
33 TB data 10 TB MC 43 TB total	40 TB data 41 TB MC 81 TB total

Current DST cache capacity: **13 TB**

Purchase of 21 TB + new controller approved

Will request additional disk space (~30 TB) next year

Outlook and summary

- Need to discuss priorities, especially concerning reprocessing for final data set
- When new CPU arrives, we will have the offline capacity to generate a definitive MC sample and reprocess as necessary on a time scale compatible with end 2006
- Seems to me that MC has highest priority
- Should technically be able to begin cranking out preliminary results before complete conclusion of offline activity