

Table of Contents

Explanation for HTMLPages:	1
Comments to different action items on https://twiki.ppe.gla.ac.uk/bin/view/Main/DatasetTroubleShooting	2
Duplicate file checks.....	2
2. Unique files with wrong num_MV1_70.....	2
3. (num_MV1_70) tests.....	2
4. Numbers of events in triplets of files.....	2
5. Comparison of systematic impact in INT note.....	2
6. Yields that repeat in the dataset.....	2

Explanation for HTMLPages:

- red: uncertainty $> 50\%$
- blue: both up/down variation are $>$ or < 0
-

Comments to different action items on <https://twiki.ppe.gla.ac.uk/bin/view/Main/DatasetTrouble>

Duplicate file checks

- first check: look at sum of all weights for L3 files, $\text{weight} = \text{FinalWeight} * \text{TRFweight}$. There are two ways that two files can have exactly the same yield:
 - ◆ a) they are identical because a problem might have occurred when copying back the file to its correct location (in case two jobs were run at the same time in the same folder). This has been prevented in the last 2 months by fixing the batch-submission script so that each job has its own folder.
 - ◆ b) the impact of the systematic uncertainty is tiny. In the list of files provided by Donny, the values are rounded to 6 digits. I repeated those checks looking at 12 digits. The list of files still having the same yield can be found here:
`/afs/cern.ch/work/d/dquilty/GlaBookkeepingGUI/UsefulLittleScripts/LogDoubleEvents_afterFixLepton`
 - * findings: we see identical yields for two sort of systematics:
 - ◆ `JesEffectiveMix1`, `eer`, `ees`, `musc`, `mums` ---> these systematics have a small impact in general, so it would be just by chance that the systematic does not change the event yield
 - ◆ 5 files for `btag/ctag` syst: checking them now, is it ok if I rerun those?
- the 97 files identified by Donny have been rerun

2. Unique files with wrong `num_MV1_70`

- 29 files have been remade

3. (`num_MV1_70`) tests

- did not find the cause for events with wrong number of btags in AF2 files, however since its only 1 event out of 5M, I am not sure its worth investigating. Assume that it could happen due to numerical rounding, but no proof found.

4. Numbers of events in triplets of files

- 8 affected files were remade.

5. Comparison of systematic impact in INT note

- Here the comparison is not straight forward. What is shown in the INT note are normalisation systematics after smoothing and pruning and hence they cannot be directly compared. What are your plans for smoothing and pruning? At this stage I would suggest to ignore their numbers there since we do not know the raw numbers, and just check for us if the size of the systematic makes sense. If you want to see our pruning script, I copied it here: `/afs/cern.ch/work/a/aknue/public/ForDonny/Pruning`

6. Yields that repeat in the dataset

- see explanation at 1, most of those cases are due to rounding (systematics have sub-percent/permille effects), the only remaining samples with the same yield are listed in the txt file mentioned in point 1.

-- AndreaKnue - 2015-03-27

Comments to different action items on <https://twiki.ppe.gla.ac.uk/bin/view/Main/DatasetTroubleShooting>

This topic: Main > L3ForDonny

Topic revision: r1 - 2015-03-27 - AndreaKnue



Copyright &© 2008-2020 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback