

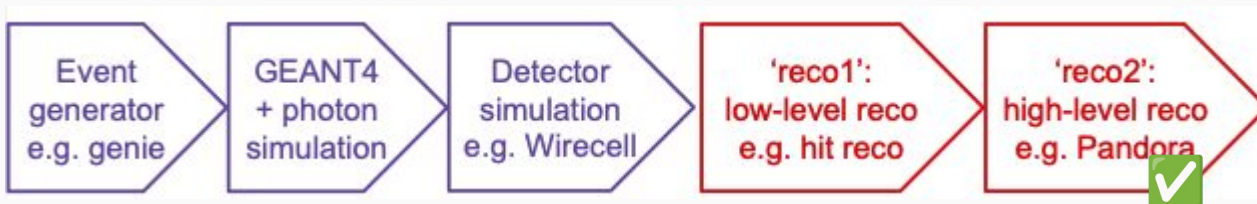
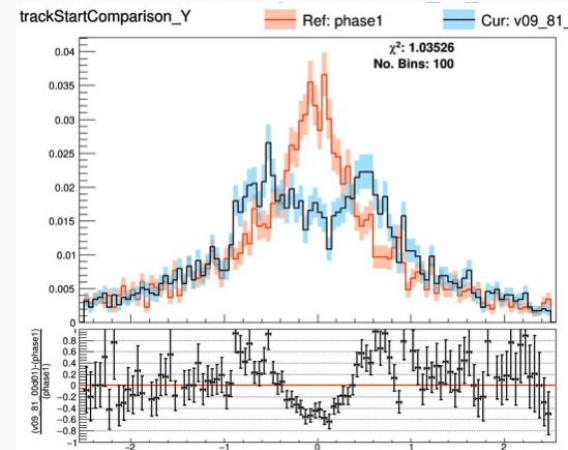
Automated Validation System

Overview and Status



Overview

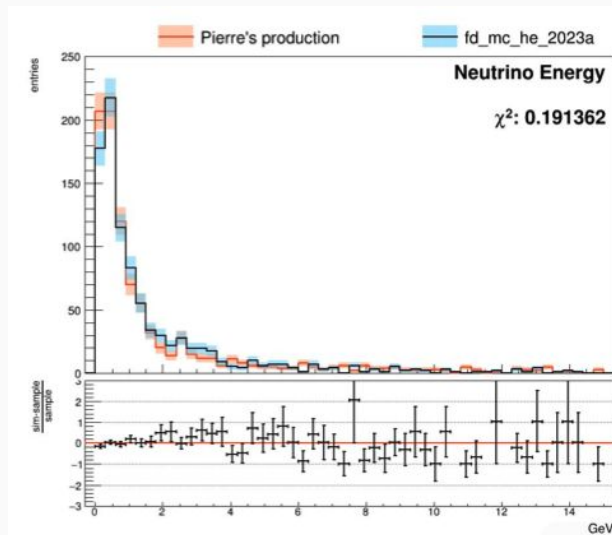
- **Automated validation system:**
 - Running high-stats jobs to compare various distributions of 2 samples: 'reference' vs 'new'
 - Invaluable in spotting bugs during development stages and new production campaigns!
- **Status in January 2024**
 - It was tailored to do reco2/Pandora output comparison plots



- **Since then, more packages have been added, to make:**
 - Gen+Det+Hit plots
 - Plots relevant to the atm nu samples
 - Plots for PDS-related checks

New: Gen+Det+Hits comparisons

- **GenRecoValidator analyzer added. It allows comparisons for:**
 - Generation level variables: true vertex position, nu azimuth angle, nu energy, etc..
 - Detector level variables: detected photons for different flavors, etc..
 - Hit level info: eg. hit charge, hit peak time, hit width, etc...
- **Full list of available comparisons:**
 - [Google sheet](#)
 - No other plots envisioned to be added at this time
 - if other needed, add desired plots to googleSheet, and ping Matteo
- **Status for users:**
 - Added to the CI system, pending a merge



New: atm/beam analyses

	DISTRIBUTIONS					DISTRIBUTIONS					
		Type of plot	Useful for Atm	Useful for LBL			Type of plot	Useful for Atm	Useful for LBL		
46		x	histogram	x	x	61	Flash purity	histogram	x		
47	Reco Nu vertex delta X, Y, Z	y	histogram	x	x	62	Nu energy resolution VS (split per channel and interaction mode)	Etrue	1D	x	x
48		z	histogram	x	x	63		Theta_x true	1D	x	
49		CVN score numu	histogram	x	x	64		Theta_y true	1D	x	
50	CVN score nue	histogram	x	x	65	Theta_z true		1D	x		
51	Reco - True T0 from flash	histogram	x		66	Vertex_x true		1D	x	x	
52	Number of PFP space points	histogram	x	x	67	Vertex_y true		1D	x	x	
53	track dE/dx	histogram	x	x	68	Vertex_z true		1D	x	x	
54	shower dE/dx	histogram	x	x	69	Direction resolution (x, y, z) VS (split per channel and interaction mode)		Etrue	1D	x	
55	Track mom by range (possibly only for true mu)	histogram	x	x	70			Theta_x true	1D	x	
56	Track mom by MCS (possibly only true mu)	histogram	x	x	71			Theta_y true	1D	x	
57	Nu energy reco numu	histogram	x	x	72		Theta_z true	1D	x		
58	Nu energy reco nue	histogram	x	x	73		Vertex_x true	1D	x		
59	Nu energy reco nc	histogram	x	x	74		Vertex_y true	1D	x		
60	Number of PEs per OpFlash	histogram	x	x	75	Vertex_z true	1D	x			
61	Flash purity	histogram	x		76						

- **Full list of available comparisons:**

- [Google sheet](#)
- No other plots envisioned to be added at this time
 - if other needed, add desired plots to googleSheet, and ping Pierre

- **Status for users:**

- Added to the CI system, pending a merge

New: PDS-related plots

Reference workflow adapted from **Low-Energy WG**
[MARLEY events with a flat energy spectrum]

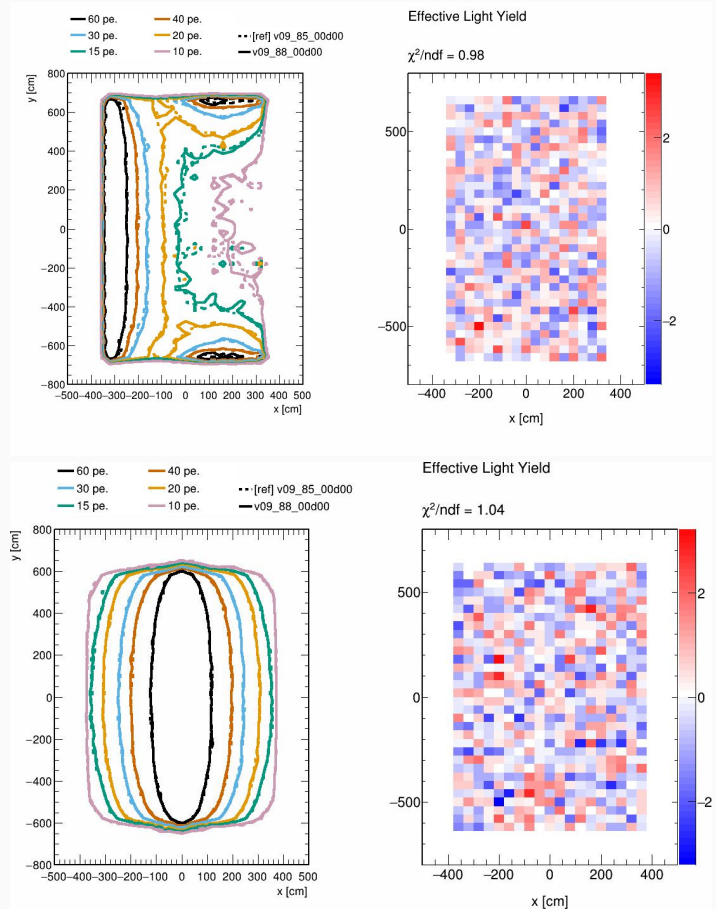
[G4 level] Scintillation time profile, light transport
[OpHit level] OpHit reconstruction, charge measurement
[OpFlash] Number of flashes, purity, charge

- **Full list of available comparisons:**

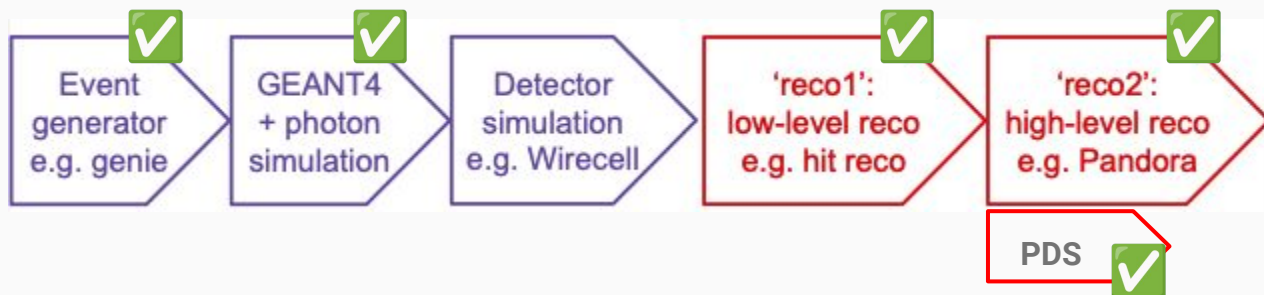
- [Google sheet](#)
- TODO: check backtracker record information
- New requests? Get in touch with Daniele

- **Status for users:**

- Added in the CI system: No.
- Workflow for validation and reference generation in **lar_ci:feature/mib_pds_ci** (both FD-VD and FD-HD)
- Pending pull request for extra analyzer in duneana
- Reference: [slides from May CM](#)

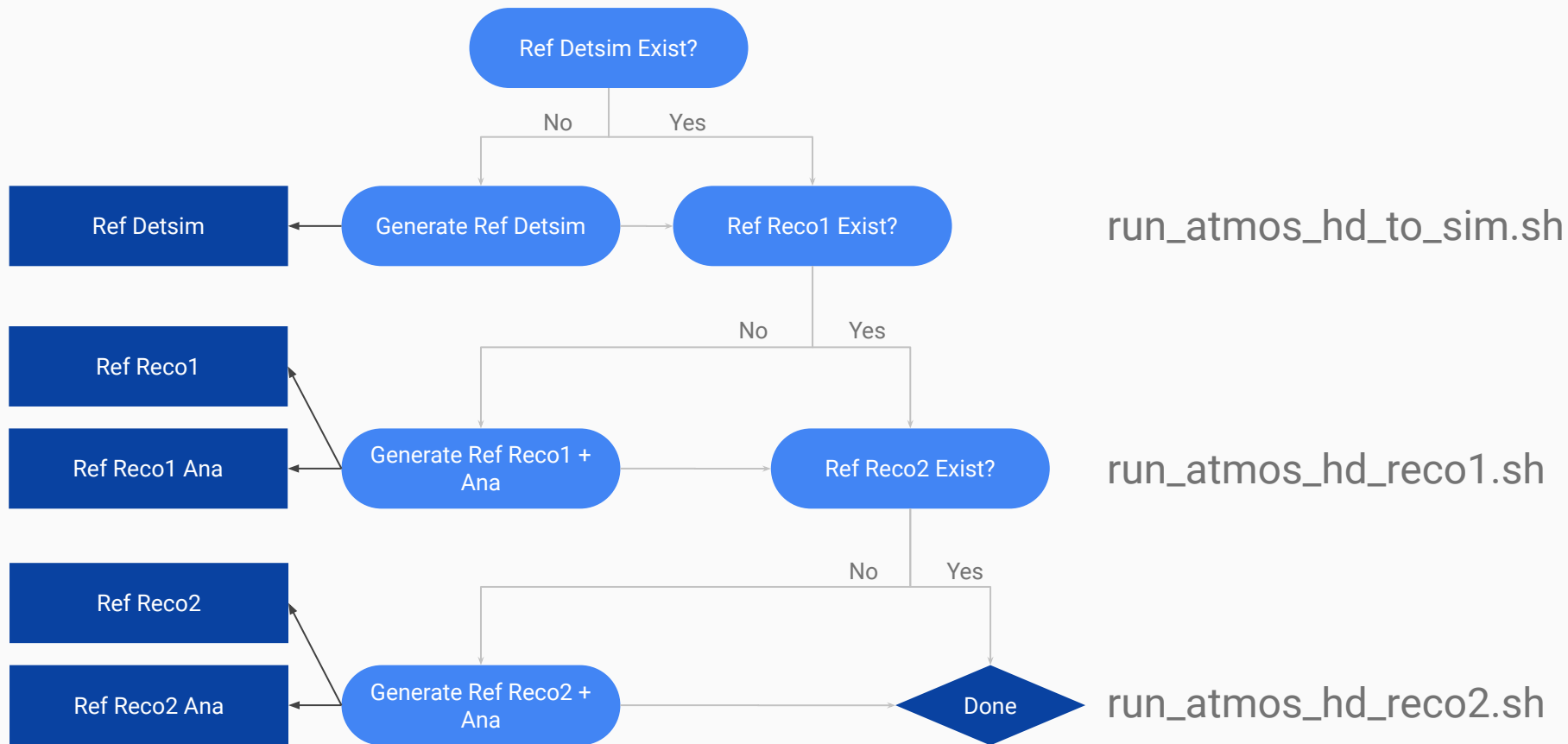


Status



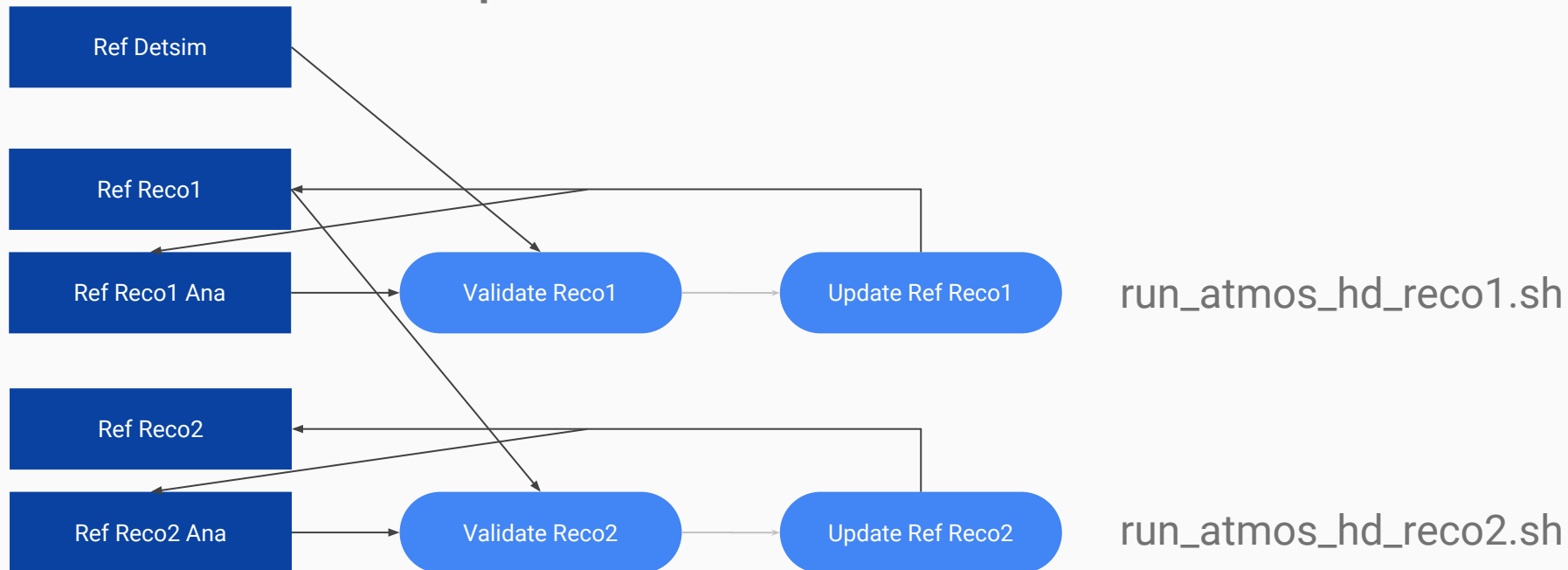
- **Comparisons for all parts of the chain are now possible after the reco1/2 stages**
- **Limitations:**
 - The CI system runs directly from git/redmine, which makes modifying it pretty painful
 - For configuration changes you end up in a commit, push, test cycle, instead of test, commit, push
 - Users should be relatively insulated from this if they are just running the standard chain
 - 2000 event validation (reviewable, but 100K+ runs in the CI system stretch its intended function)
- **Status of integration in CI:**
 - Updated branch with Matteo and Pierre's changes (plus my technical edits) is ready
 - Pending resolution of upstream token issue for final tests and merge (hopefully very soon)
 - PDS updates from Daniele are in a separate branch at the moment, as noted previously
- **A set of instructions for the whole process will be made available (A&E wiki?)**
 - For now, I have a brief guide later in these slides

Workflow overview - reference files



Workflow overview - validation

Update of the reference reco1 and reco2 files is manual



Operation

- **A reference set of run scripts are available in `lar_ci/cfg/dune/atmos`**
 - `run_atmos_hd_to_sim.sh`
 - `run_atmos_hd_reco1.sh`
 - `run_atmos_hd_reco2.sh`
- **But the next few slides will summarise the usage of the above scripts**

Setting up the CI system

- Once you're running in your SL7 container (for now), you'll want to set up a script with the following commands

```
kx509
```

```
source /cvmfs/dune.opensciencegrid.org/products/dune/setup_dune.sh  
source /cvmfs/fermilab.opensciencegrid.org/products/common/etc/setups.sh  
source /exp/dune/app/users/vito/CI/setup_ci.sh  
sh /exp/dune/app/users/vito/CI/store_vtoken.sh # if you haven't run recently
```

- This will enable you to run the CI trigger command
- **Note, this relies on a script historically maintained by Vito Di Benedetto**
 - Vito won't be managing the CI moving forward (though as ever, is being extremely helpful pending a replacement - I don't know what this will look like)
 - Vito is looking to provide a CVMFS-based CI setup to alleviate this issue
 - Until then, you may want to replicate this setup_ci.sh script and the associated infrastructure that it references, but you will need to maintain an up-to-date version of lar_ci

run_atmos_hd_to_sim.sh

```
trigger_token --testmode --ci-tests ci_gen_regression_test_dunefd --token -E dune --build-delay 0 \
--workflow CI_VALIDATION_DUNE_lite --jobname dune_ci_test \
--gridwf-cfg cfg/dune/atmos/grid_workflow_DUNE_atmos_hd_gen_to_sim.cfg \
--quals e26:prof -e DUNEmodules_extra="LArSoft/larsoft DUNE/dunesw DUNE/duneana" \
--revisions "LArSoft/larsoft@v09_91_00 DUNE/dunesw@v09_91_00d00 AndyChappell/duneana>duneana@feature/val_sys" \
--version feature/chappell_vd_ci
```

- **The structure**

- The first three lines above should be very stable and should not change without upstream changes to the CI system itself
 - You may need to update “quals” according to the dunes software version
- The “DUNEmodules_extra” should typically contain the larsoft and dunes references, other repositories will need to be added for any custom branches (in this case duneana, which temporarily uses a custom branch)
- “revisions” indicate the specific versions of the packages specified by DUNEmodules_extra
 - These should reflect either the appropriate tagged release, or a custom feature branch
 - Version consistency is essential
- “version” refers to the lar_ci branch to run, you only need this if you have custom CI code
 - This is temporary pending a merge of this branch into lar_ci master

- **You must copy the detsim files to**

- /pnfs/dune/persistent/users/\${CI_USER}/atmos_validation/ref_atmos/sim/
- CI_USER is your FNAL username

run_atmos_hd_reco1.sh

```
trigger_token --testmode --ci-tests ci_gen_regression_test_dunefd --token -E dune --build-delay 0 \
  --workflow CI_VALIDATION_DUNE_lite --jobname dune_ci_test \
  --gridwf-cfg cfg/dune/atmos/grid_workflow_DUNE_atmos_hd_reco1_${cfg_suffix}.cfg \
  --quals e26:prof -e DUNEmodules_extra="LArSoft/larsoft DUNE/dunesw DUNE/duneana" \
  --revisions "LArSoft/larsoft@v09_91_00 DUNE/dunesw@v09_91_00d00 AndyChappell/duneana>duneana@feature/val_sys" \
  --version feature/chappell_vd_ci
```

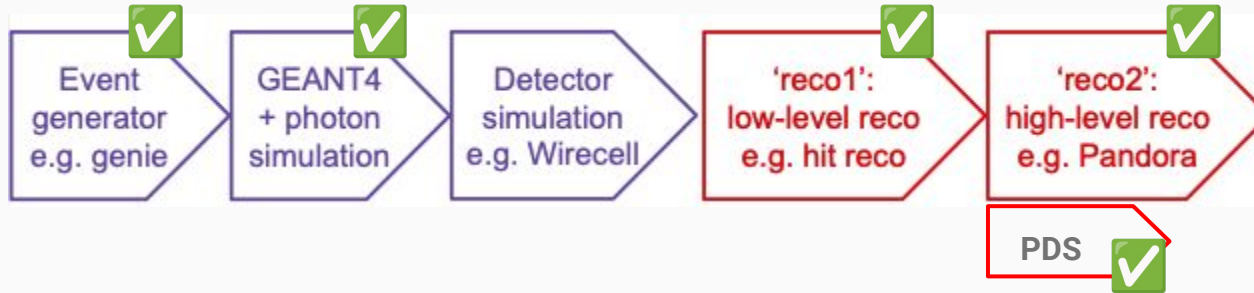
- **This is very similar to the command from the previous slide and all of the comments apply here**
- **The key difference is that this script can run in either reference or validation mode**
 - Providing the argument `ref` or `val` determines which mode runs
 - Mechanics (not shown) in the script set the `${cfg_suffix}` variable to run the right configuration
- **Whether you're running reference or validation step, the output files from this step will be stored in your personal scratch area**
 - Under the relative directory `ci_validation/atmos_hd/${DUNESW_VERSION}`
- **If you are creating reference files, you must copy the reco1 files to**
 - `/pnfs/dune/persistent/users/${CI_USER}/atmos_validation/ref_atmos/reco1/`
- **The mergeana file (pandoraAnalysis_merged.root) should be copied (and renamed) to**
 - `/pnfs/dune/persistent/users/${CI_USER}/atmos_validation/ref_atmos/validation/referencePandoraAnalysis.root`

run_atmos_hd_reco2.sh

```
trigger_token --testmode --ci-tests ci_gen_regression_test_dunefd --token -E dune --build-delay 0 \
--workflow CI_VALIDATION_DUNE_lite --jobname dune_ci_test \
--gridwf-cfg cfg/dune/atmos/grid_workflow_DUNE_atmos_hd_reco2_${cfg_suffix}.cfg \
--quals e26:prof -e DUNEmodules_extra="LArSoft/larsoft DUNE/dunesw DUNE/duneana" \
--revisions "LArSoft/larsoft@v09_91_00 DUNE/dunesw@v09_91_00d00 AndyChappell/duneana>duneana@feature/val_sys" \
--version feature/chappell_vd_ci
```

- **Essentially identical to the previous slide, except for a reco2, rather than reco1 configuration**
- **If you are creating reference files, you may want to copy the reco2 files to**
 - /pnfs/dune/persistent/users/\${CI_USER}/atmos_validation/ref_atmos/reco2/
 - Technically the reco2 files are only needed if you want to rerun the ana validation stage without rerunning reco2
- **The mergeana file (pandoraAnalysis_merged.root) should be copied (and renamed) to**
 - /pnfs/dune/persistent/users/\${CI_USER}/atmos_validation/ref_atmos/validation/referencePandoraAnalysis.root
- **Note, currently this is the same location as the reco1 files due to a technicality of how the final CI processing runs, so it's worth making a suitably named copy of the reco1 mergeana file in persistent first**
 - Addressing this is on my ToDo list

Future developments



- What else might it be nice to have?