PandAna Validation

Analysis Task Force Meeting

Derek Doyle
September 10, 2021

- PandAna has been used by NOvA for a few years now, mainly for ML applications

- Validation has been done on the fly and secondary to many of these studies

- These slides serve as documented validation of the PandAna event selection framework and guidance for validation through iterative changes

Details

- Working version of PandAna is head of https://github.com/HEPonHPC/pandana

- NOvA-specific Loaders and Spectrum defined in https://github.com/grohmc/NOvAPandAna in addition to 3flavor and NuebarCCInc Cuts and Vars

- Validation scripts committed to NOvAPandAna/validation
  - Definition: pandana_validation_nd_rhc_sumdecaf_20210909
    - "defname: prod_sumdecaf_development_nd_genie_N1810j0211a_nonswap_rhc_nova_v08_full_ndphysics_contain_v1 with limit 2"
  - H5CAFs made using `caf_to_h5` converter
  - CAFAna Cuts: nuebarccinc::decaf::kPreselection
  - PandAna Cuts: NOvAPandAna.Vars.nuebarcc_vars.kDecafPreselection

Event Selection

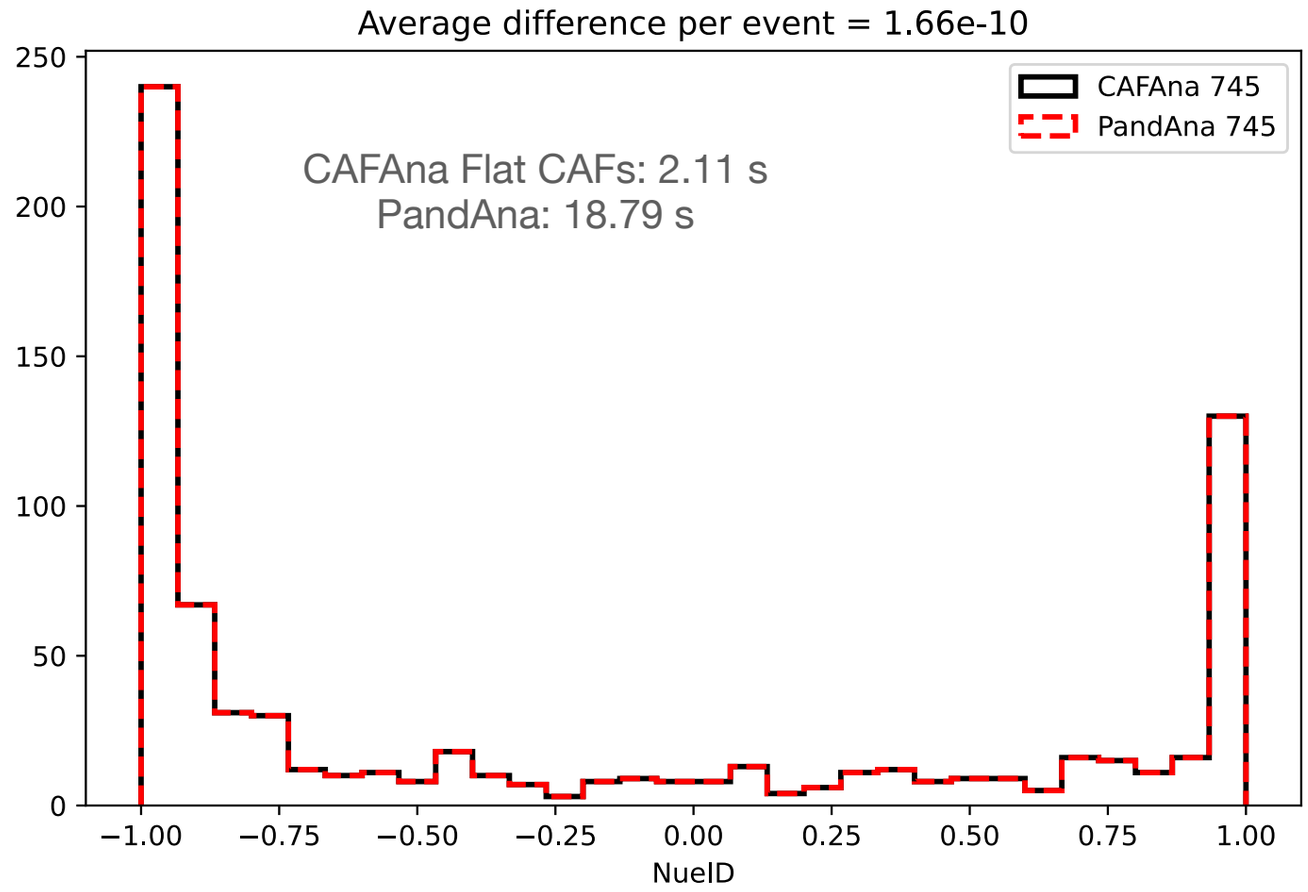
NOvA GPVMs (No MPI)

1. In CAFAna: Make event list file using `ana::MakeEventTTreeFile` including full event index (run/subrun/cycle/batch/event/slice)

2. In PandAna: Convert event tree to a pandas dataframe

3. Fill a NOvASpectrum with same cuts/vars

New to NOvAPandAna — TMVA BDTs!
Small numerical error from TMVA



Event Selection

NOvA GPVMs (With MPI!)

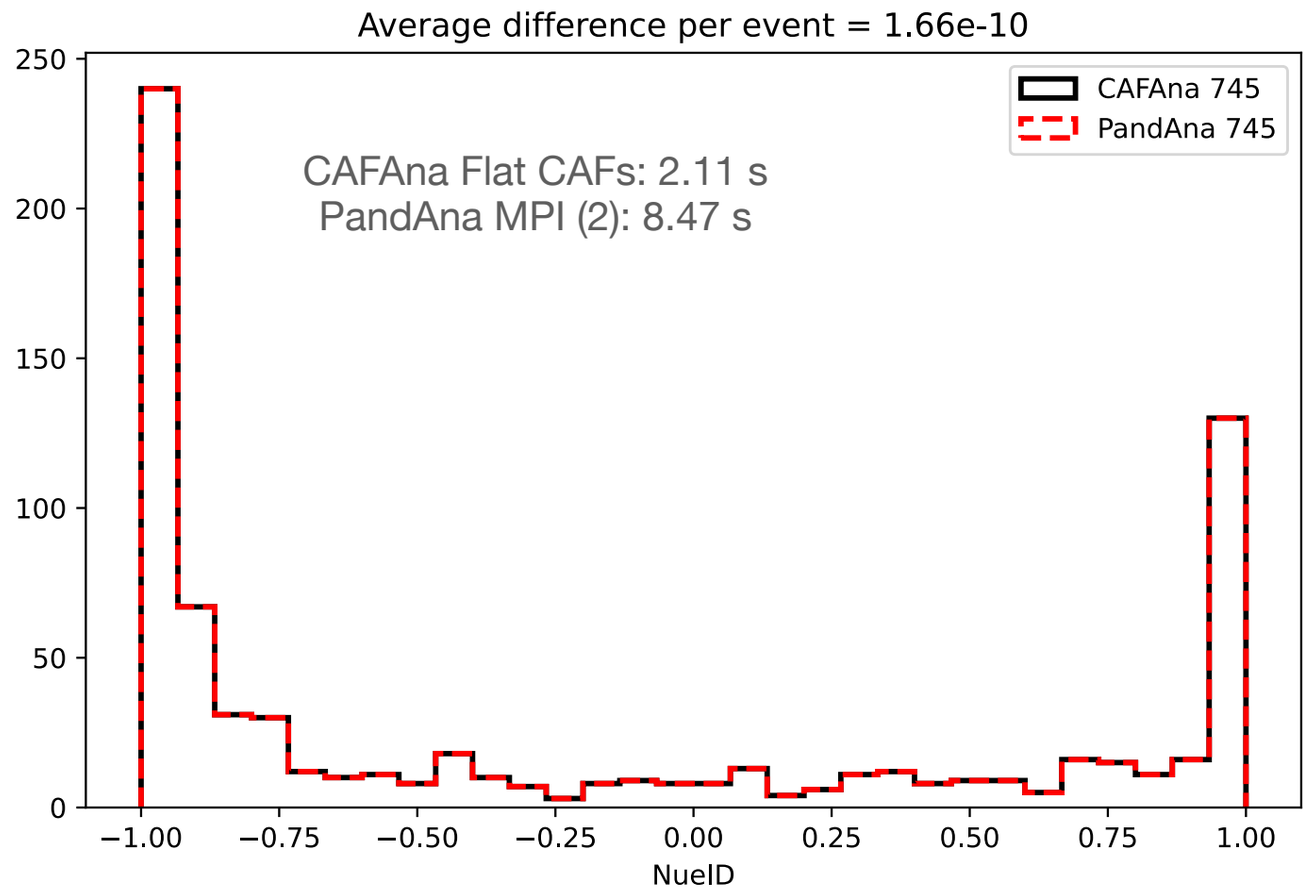
1. In CAFAna: Make event list file using `ana::MakeEventTTreeFile` including full event index (run/subrun/cycle/batch/event/slice)

2. In PandAna: Convert event tree to a pandas dataframe

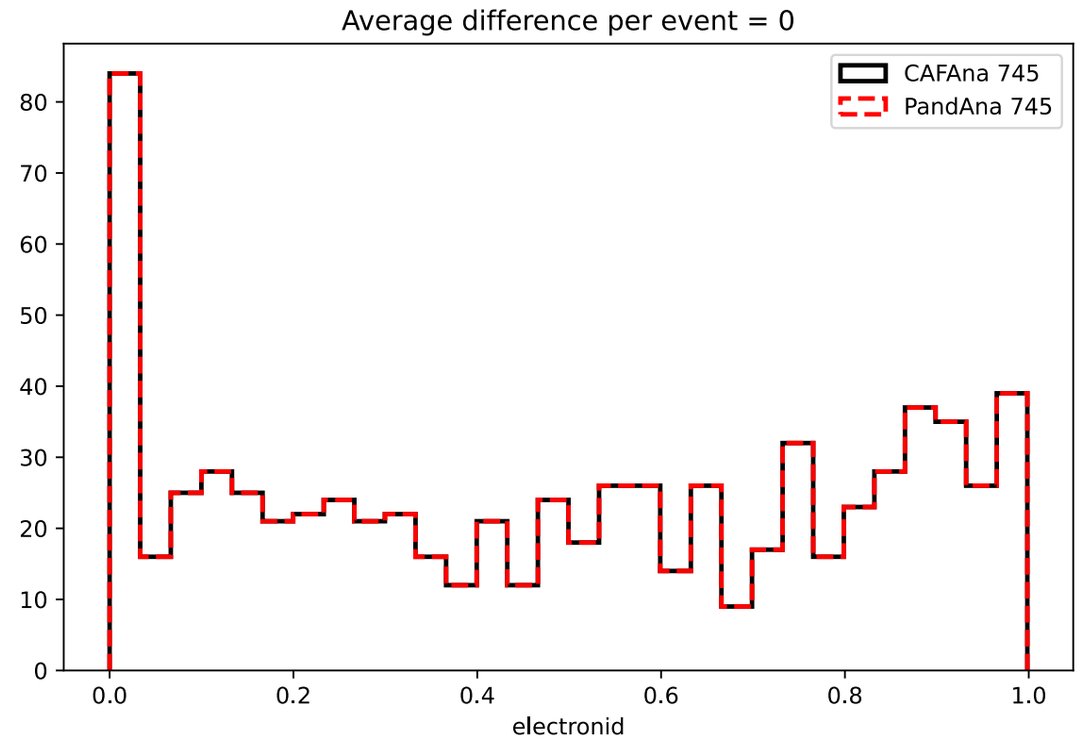
3. Fill a NOvASpectrum with same cuts/vars

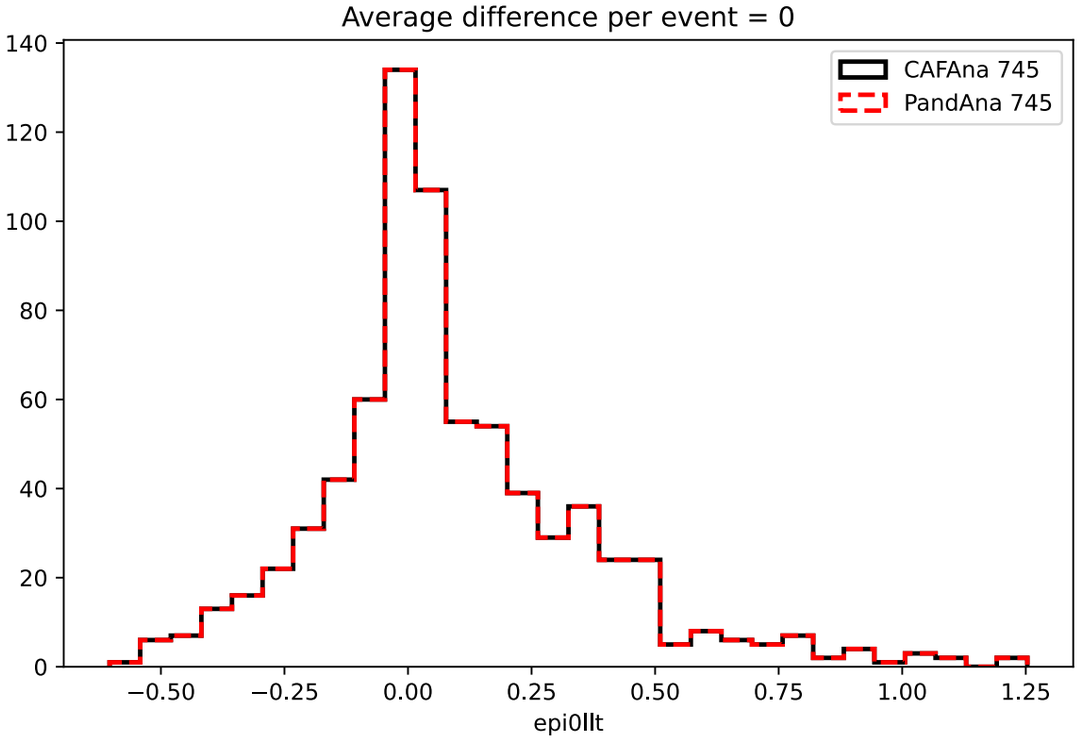
4. MPI reduce

New to NOvAPandAna — TMVA BDTs!
Small numerical error from TMVA

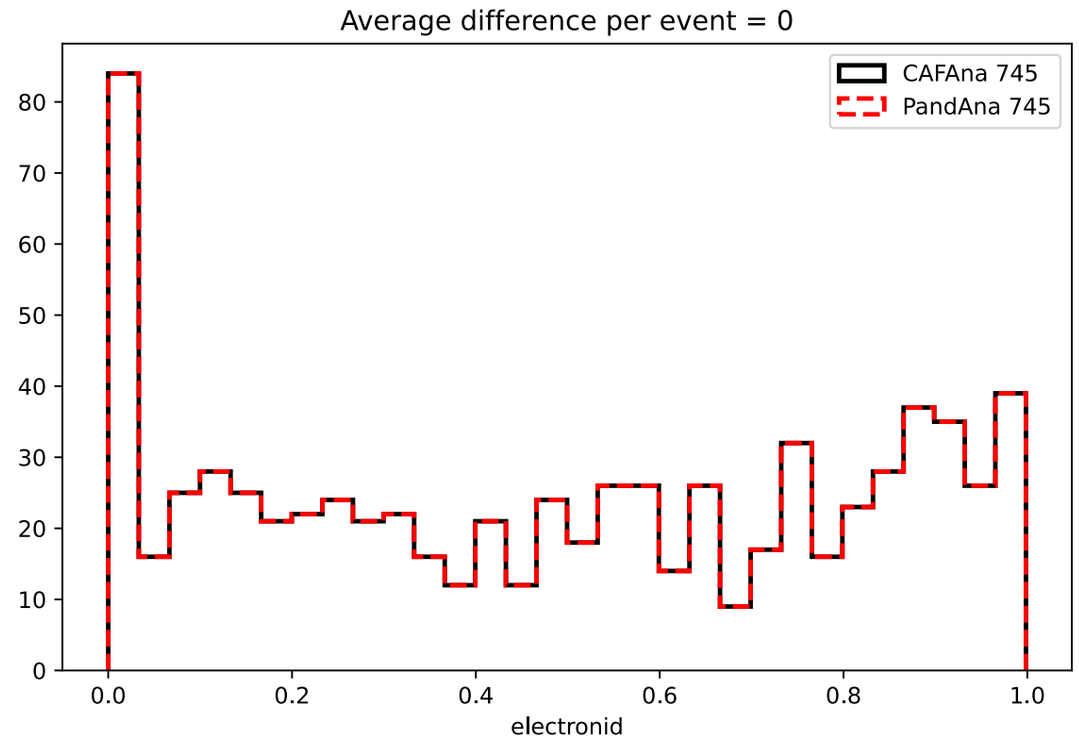


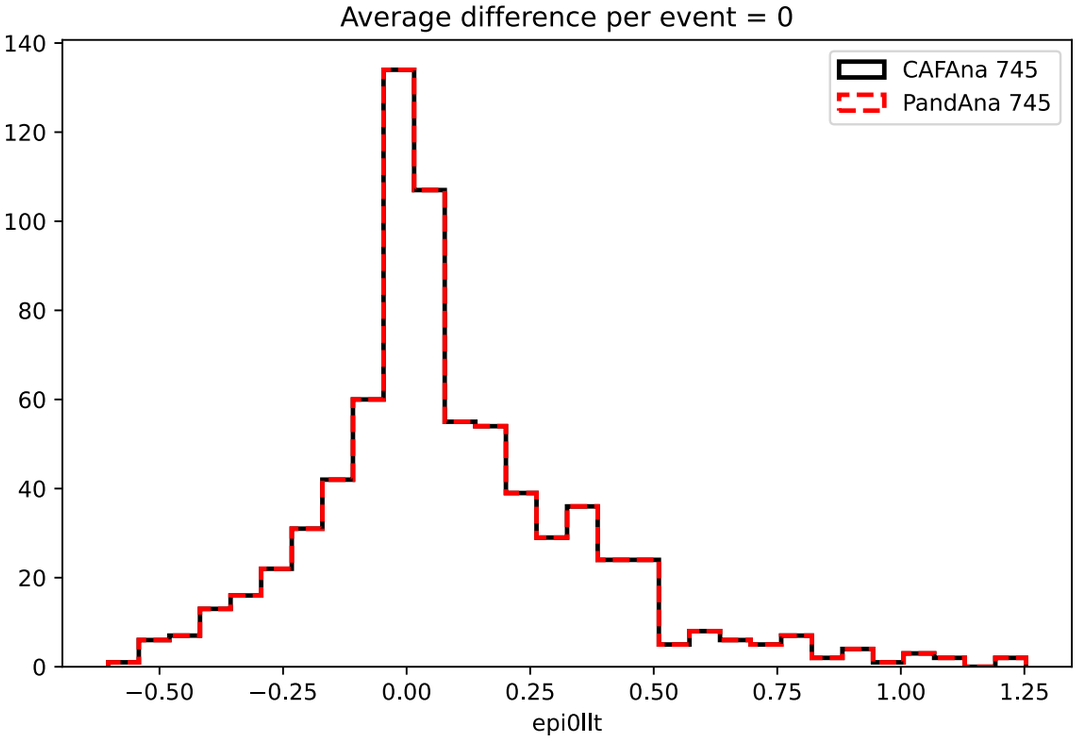
Serial PandAna



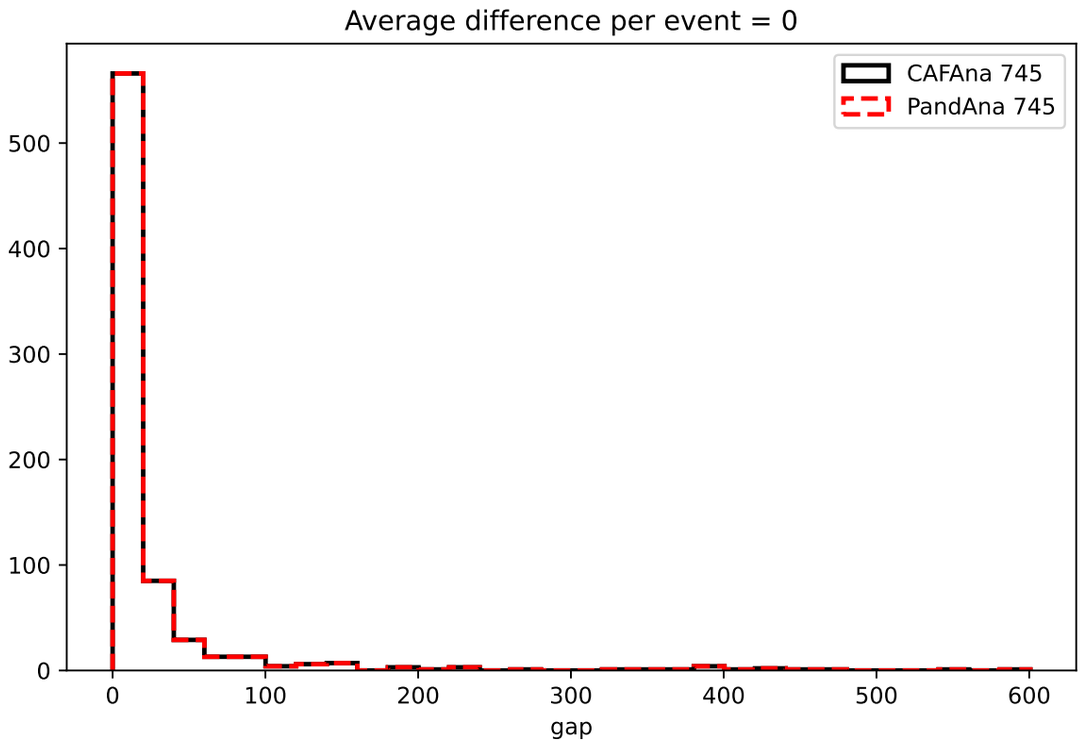


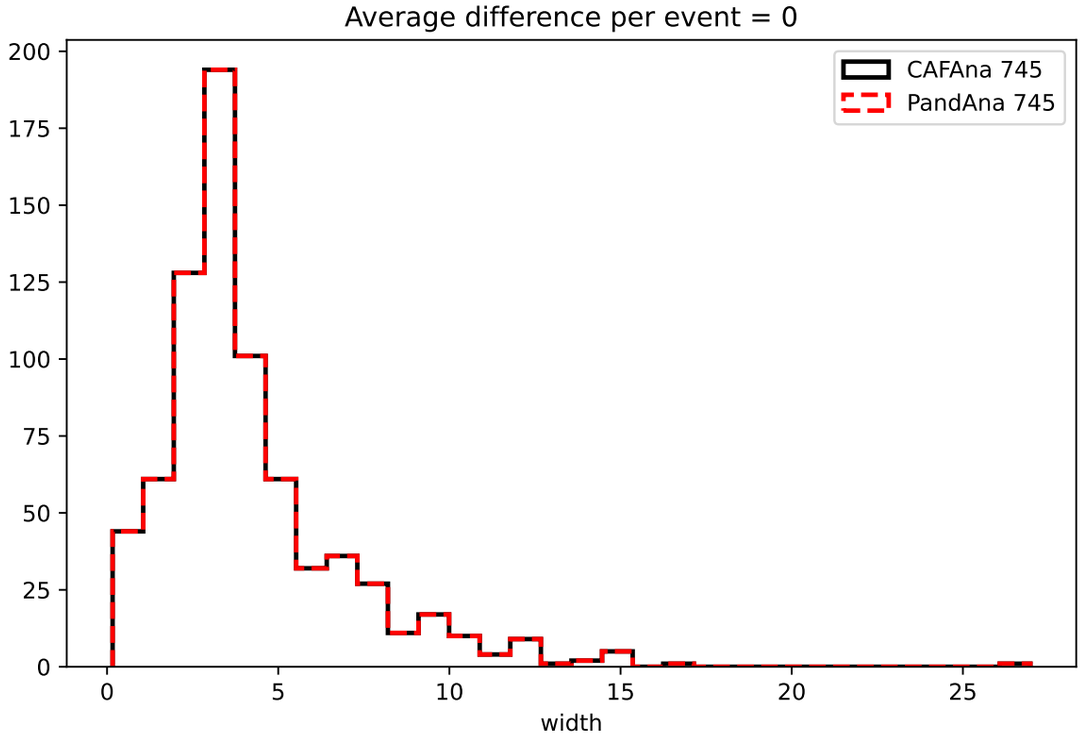
MPI PandAna



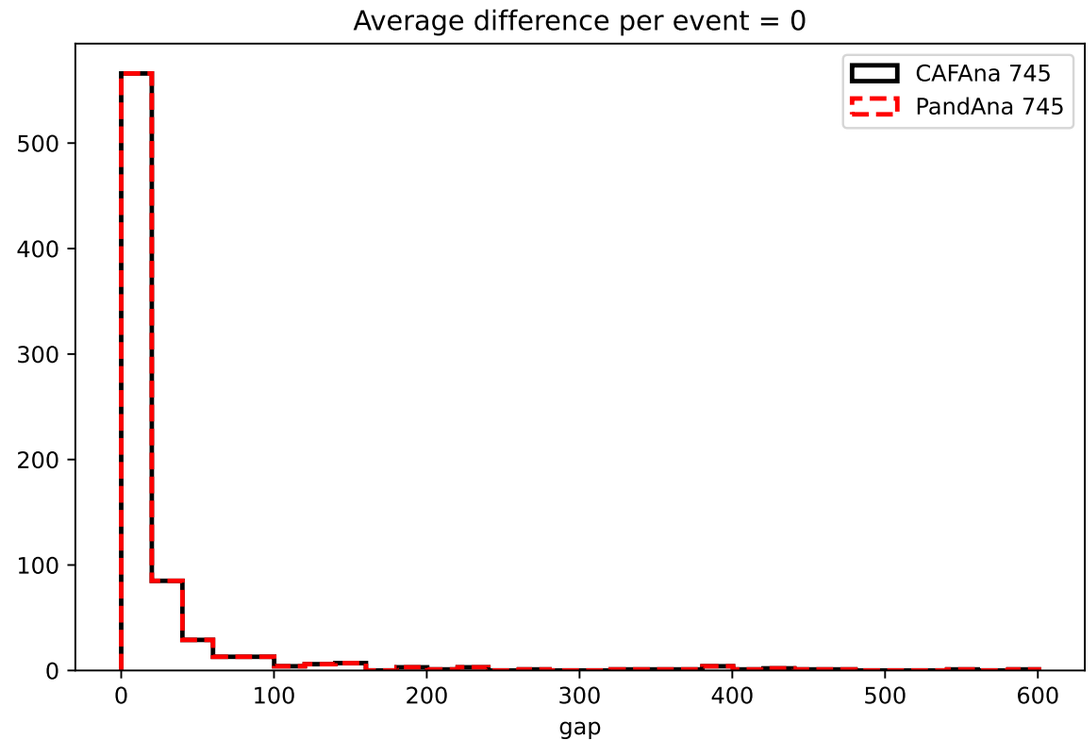


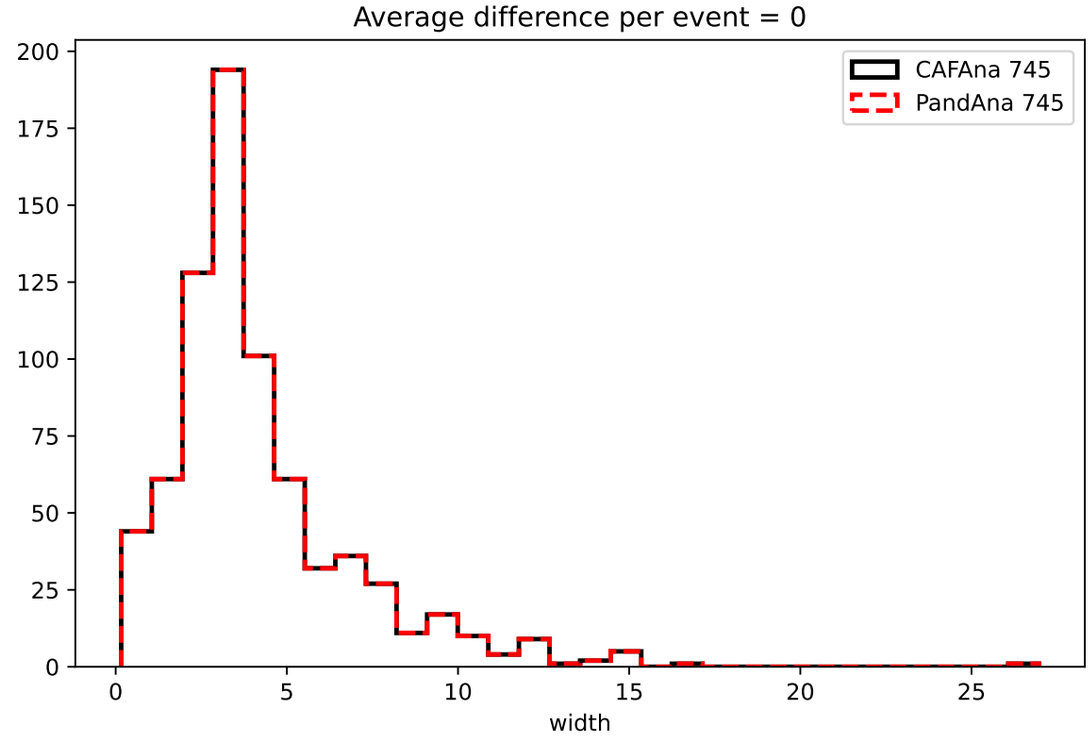
Serial PandAna





MPI PandAna





PandAna3 Validation

- pandana3 is the next iteration of pandana currently in development (M. Paterno, S. Sehrish)
  - Cuts/Vars build an abstract syntax tree for intelligent distribution of work
  - General event selection framework that NOvA can build on (like NOvAPandAna)
- Being developed with a suite of unit tests including selecting rows from a nova-like columnar hdf5 file