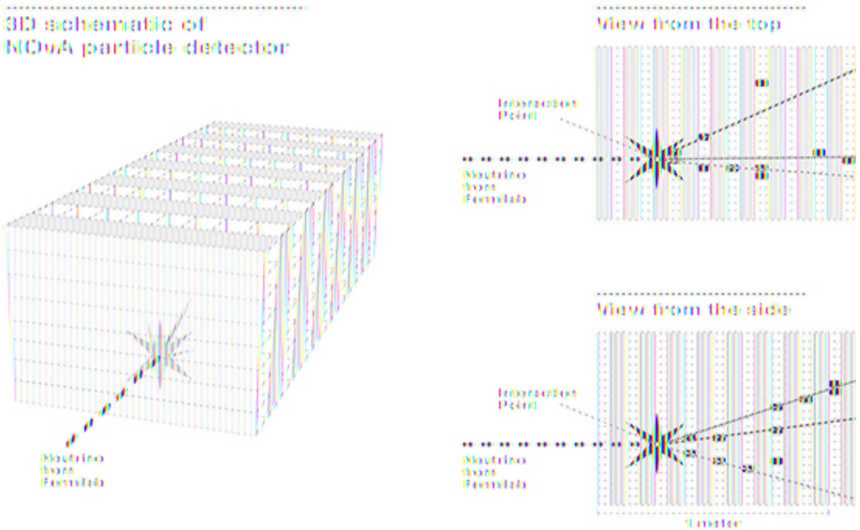
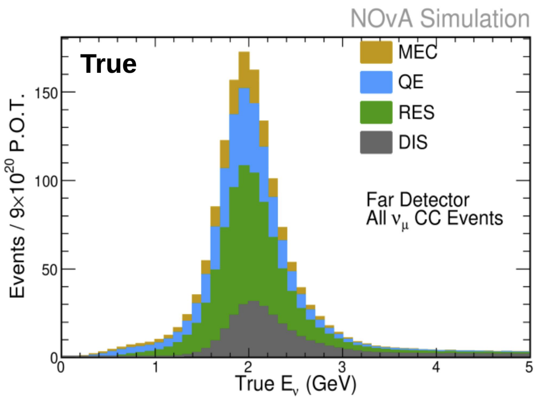
Regression CNNs for Energy Reconstruction in the NOvA Experiment

Ben Jargowsky
University of California, Irvine
For the NOvA Collaboration

NOvA



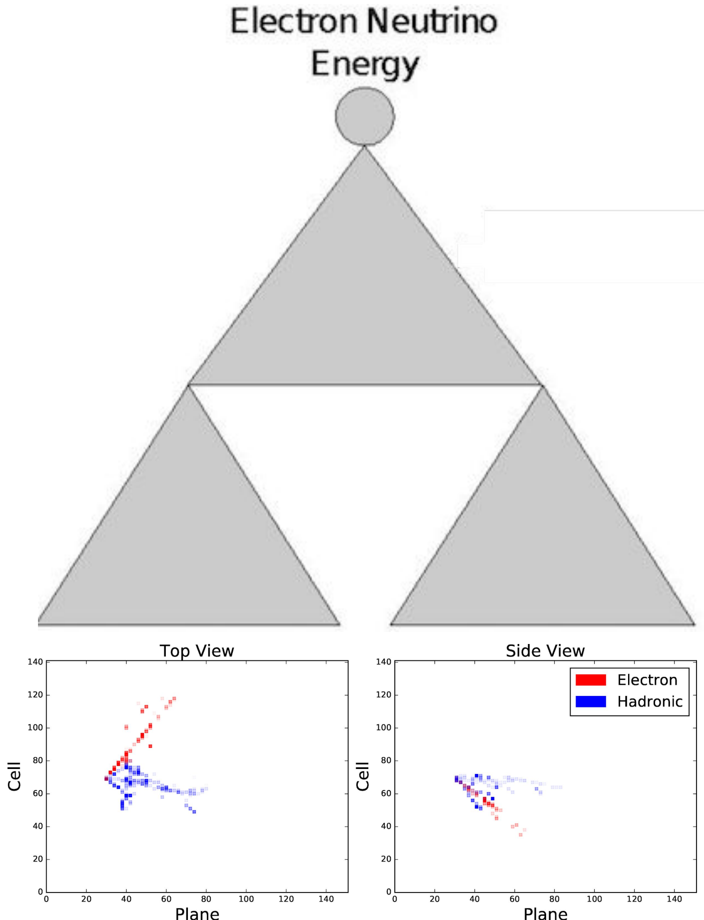
- NOvA is a neutrino oscillation experiment with an 810 km baseline
- NOvA's detector uses stacked PVC cells filled with liquid scintillator
- NOvA studies ν_μ oscillations to ν_e (and $\overline{\nu}_\mu$ oscillations to $\overline{\nu}_e$)
- Precise energy resolution is needed to make measurements of oscillation parameters
- With many interaction types to consider, we must have energy estimation for events with complicated hadronic energy components



Introduction to CNNs

- CNNs are convenient for taking images as inputs, as they use translationally invariant filters to look for features
- CNNs can be utilized in 2 ways:
    - Classification: Particle ID, Event ID
        - Here the output layer will have a sigmoid activation function, which outputs a number between 0 and 1, or softmax, it's generalization to an arbitrary number of classes
    - Regression: Particle Momentum/Energy, Event Energy, Vertex
        - Here the output layer has a linear activation function
- Here I will focus on regression CNNs for NOvA, specifically for particle and event energy reconstruction

Regression CNNs for Energy Estimation



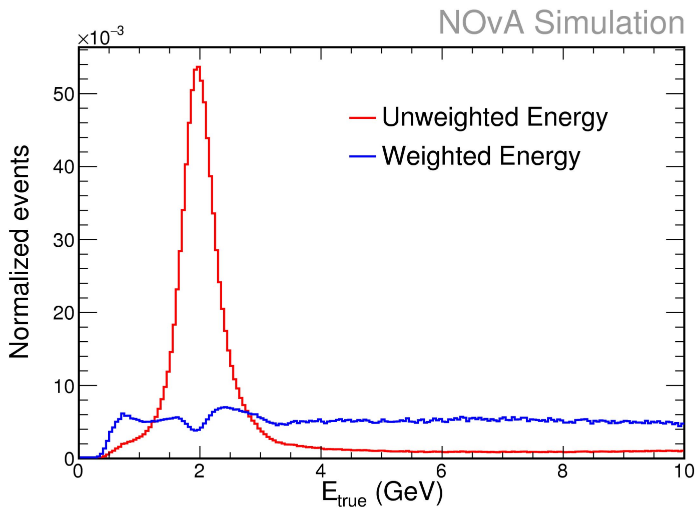
- We feed the 2 images of the event, one for each view, to the 2 inputs of the CNN which estimate event energy
- We use a “Siamese” architecture with two inputs
- Each view is a cropped 80x100 image (80 cells by 100 planes)
- The CNN architecture used is an adapted ResNet

Loss Function

- We use a weighting scheme so the loss function sees a flat distribution

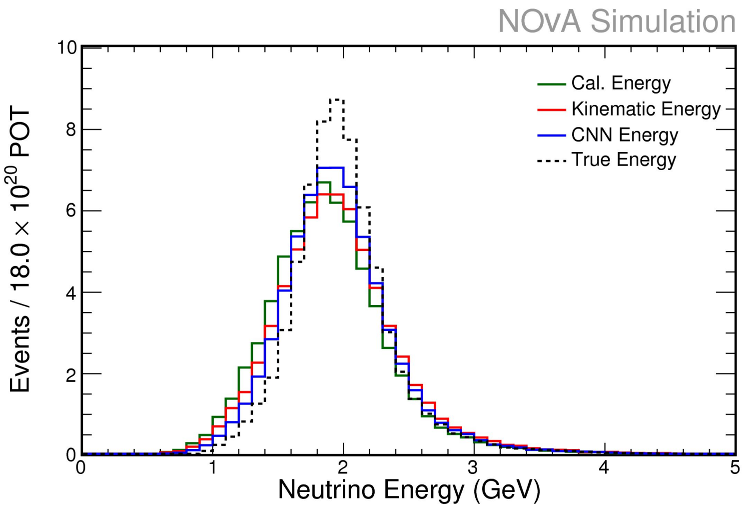
$$L(\mathbf{W}, \{\mathbf{x}_i, y_i\}_{i=1}^{n}) = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{f_{\mathbf{W}}(\mathbf{x}_i) - y_i}{y_i} \right|$$

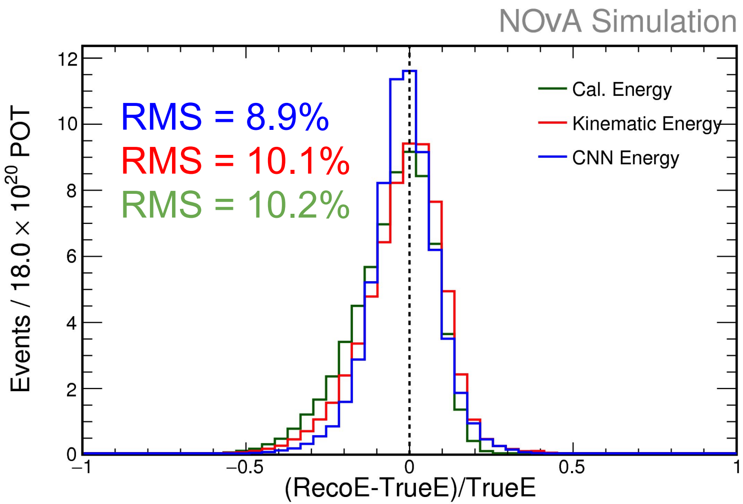
- We use mean absolute percentage error instead of square of errors to decrease the effects of outliers



ν_e Energy Performance

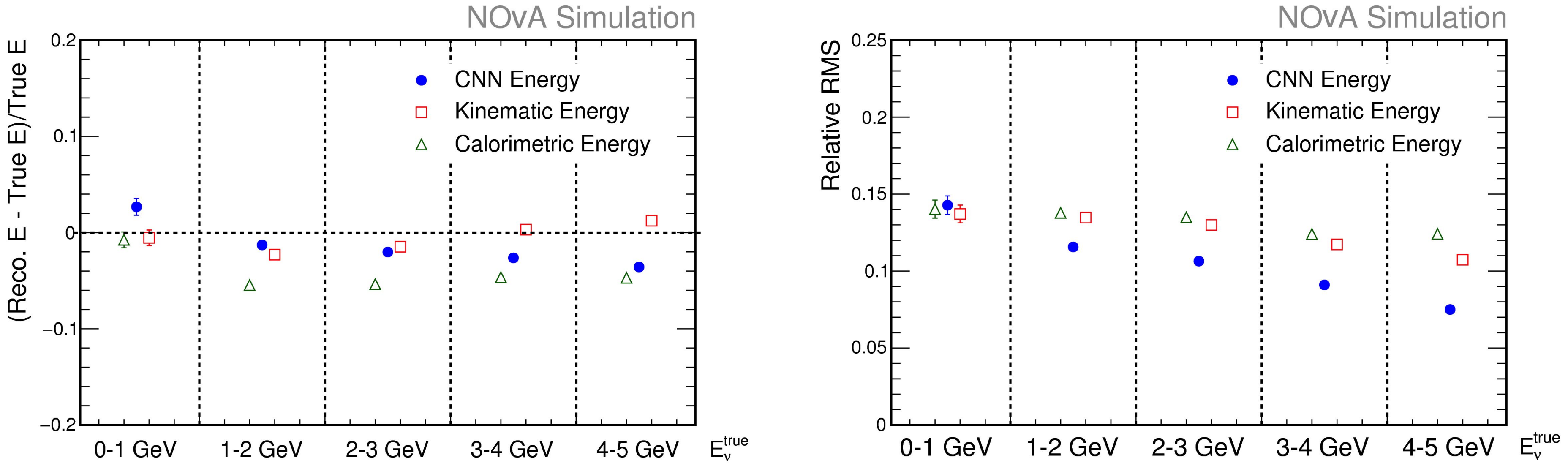
- Energy regression results for ν_e CC results are shown here
- Calorimetric method is based on adding up calorimetric energy per cell and applying an overall factor
- The kinematic method is based on a fit using electromagnetic and hadronic clustering:
- $E(\nu_e) = A\,E_{EM} + B\,E_{hadron} + C\,E_{EM}^2 + D\,E_{hadron}^2$





ν_e Energy Performance

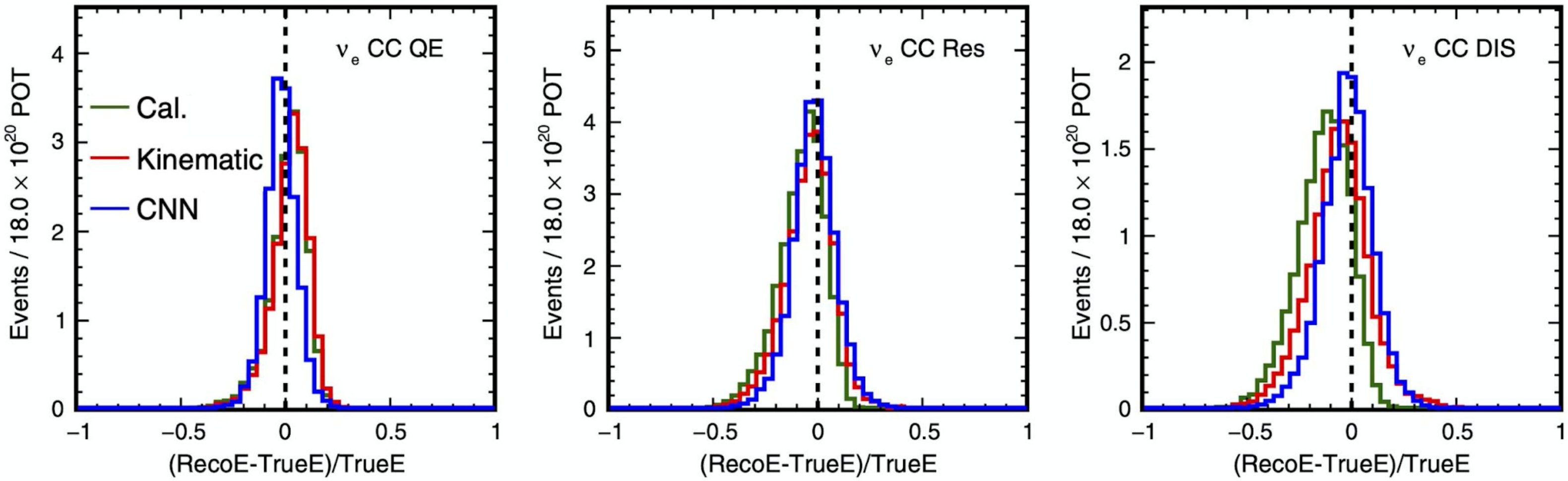
- The CNN resolution keeps the edge over calorimetric and kinematic based energy estimators from 1 GeV and up



ν_e Energy Performance

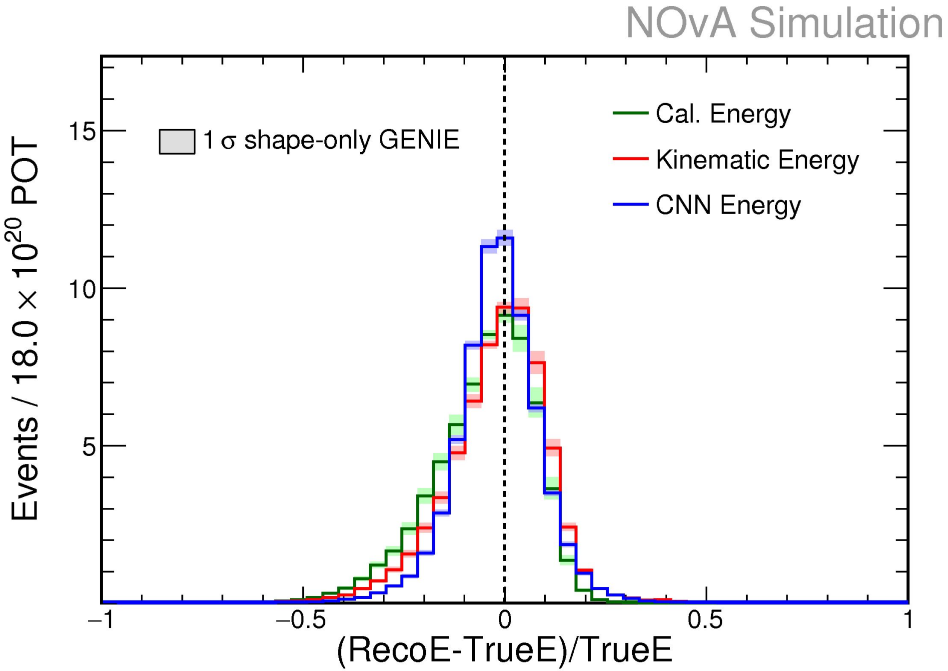
- CNN based method shows good stability over different interaction types
- This shows robustness a CNN model can provide with a large degrees of freedom

NOvA Simulation

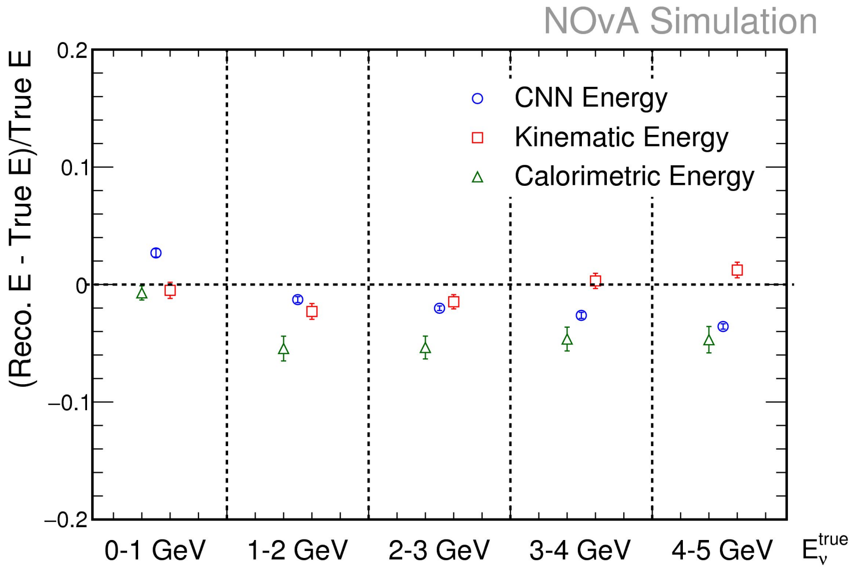


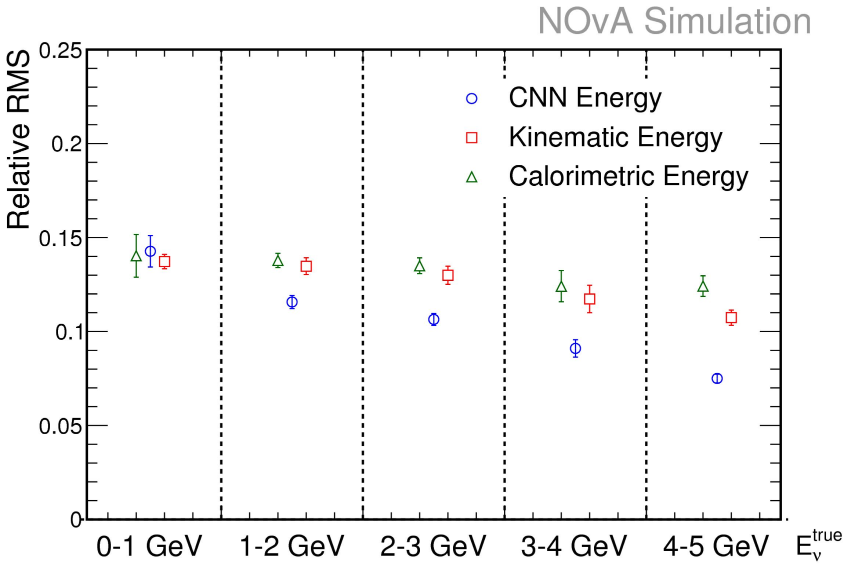
ν_e Systematics

- To quantify systematic uncertainties arising from simulation, we vary cross section uncertainties
- CNN event energy estimator is also very robust to cross section systematics, as well as the better performance



ν_e Systematics



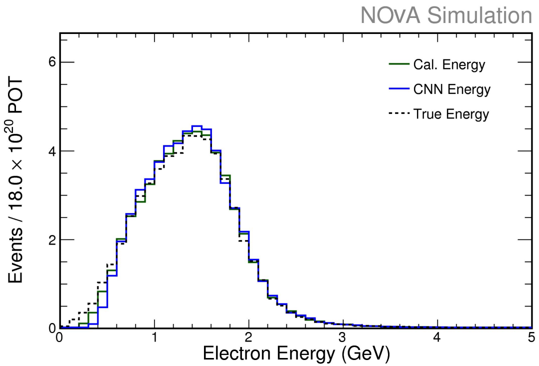


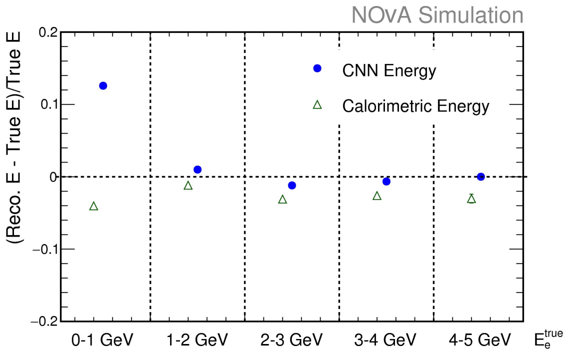
- Energy bars here are systematic uncertainties
- CNN is more robust to cross section systematics from 1 GeV onward

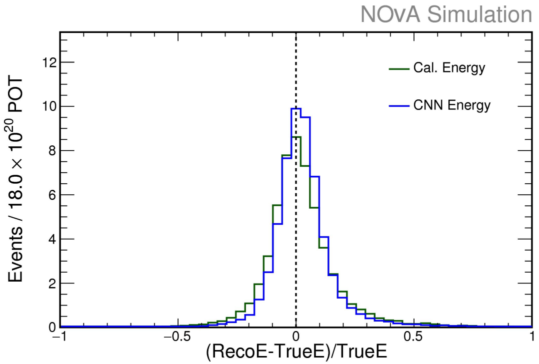
Leptonic or Hadronic Energy

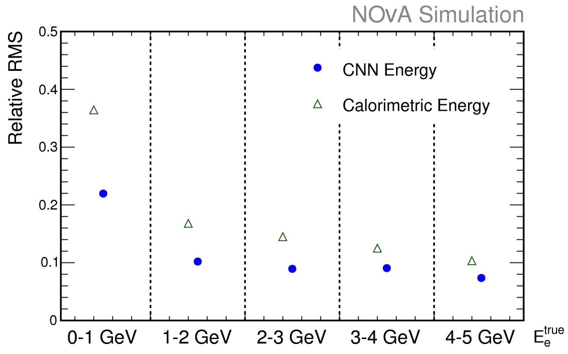
- Besides event energy, we can also use CNNs to predict energy of the leptonic or hadronic part (in ν_e CC or ν_μ CC events)
- $E(\nu) = E_{leptonic} + E_{hadronic}$
- This is useful to access kinematic quantities of the event, such as momentum exchanged in an event
- We can also use this to develop event energies (as will be shown for ν_μ CC events)

ν_e CC Electron Energy





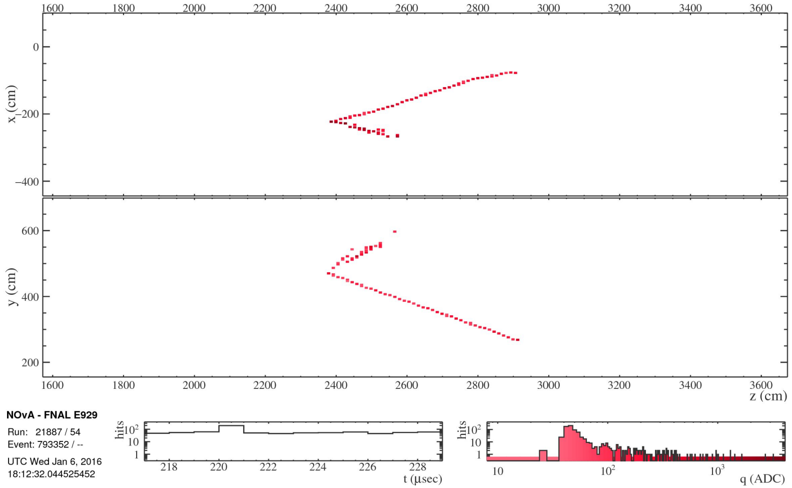




- An electron energy estimator was also developed
- The input images for this estimator are only hits from the electron part of ν_e CC events
- The CNN based estimator sees an improvement over the traditional method based on summing up calibrated hit energy

ν_μ CC Events

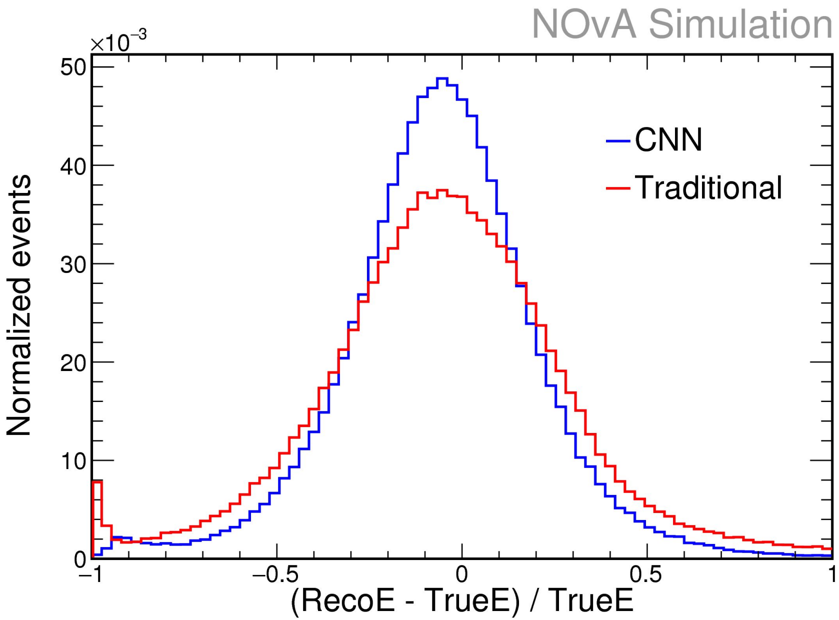
- We can also apply regression CNNs to ν_μ CC events
- Muon tracks are long, so the ν_e pixelmap size would not contain the full event
- We can specialize to estimating the hadronic part of energy, then use:
- $E(\nu_\mu) = E_\mu + E_{hadronic}$
- Then we combine our CNN hadronic energy estimator with the traditional, length based muon energy estimator to estimate ν_μ event energy



Here is a simulated ν_μ CC event, with a long muon track and hadronic shower

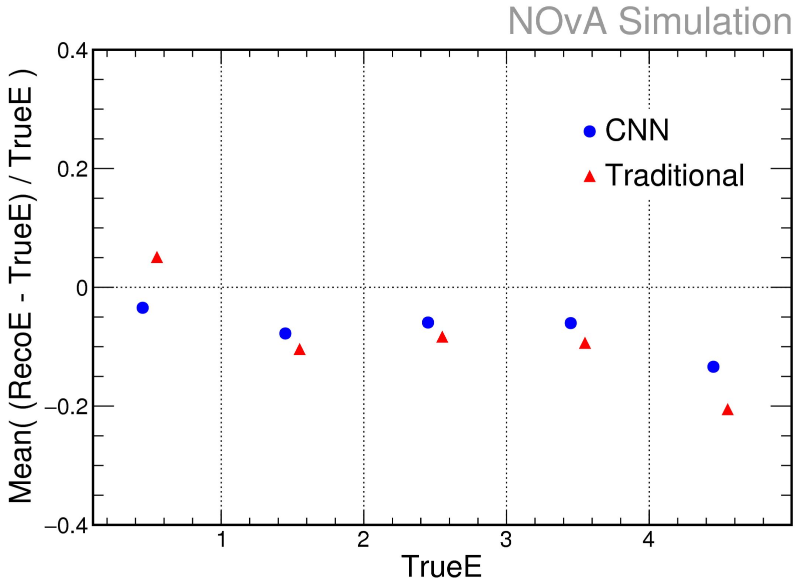
Hadronic Energy

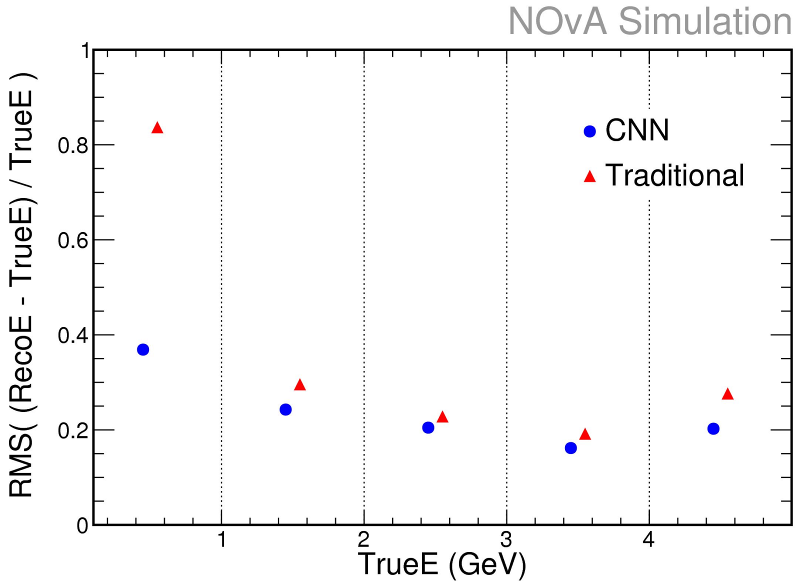
- Here we show the resolution of the hadronic energy
- We apply a weighting to the hadronic energy spectrum
- The traditional method is found by adding up visible calorimetric energy from the hadron, then applying a calibration factor



Hadronic Energy

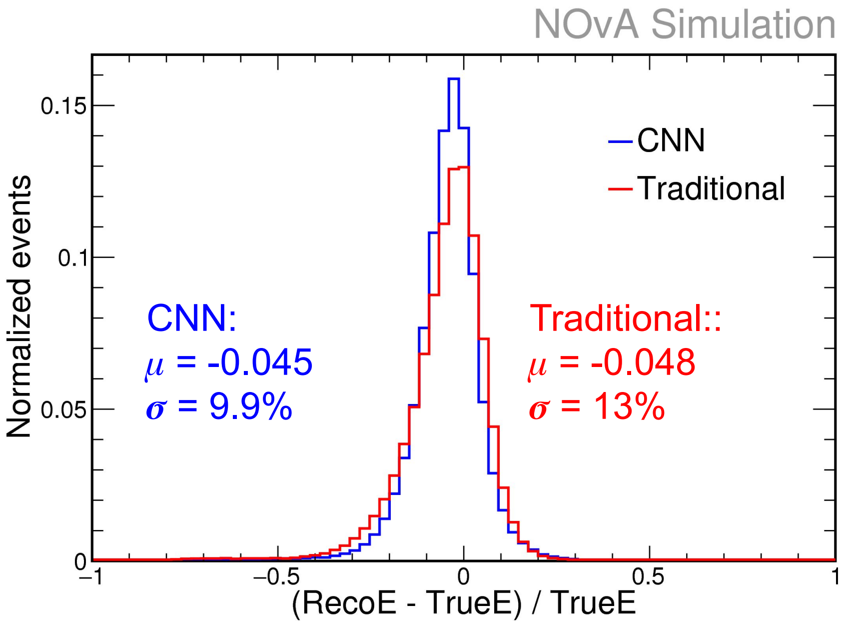
- Here we show the mean and RMS of the hadronic energy resolution sees an improvement over all regions of true hadronic energy



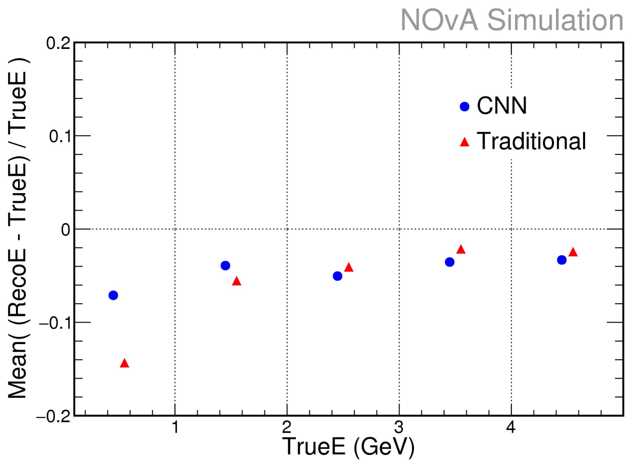


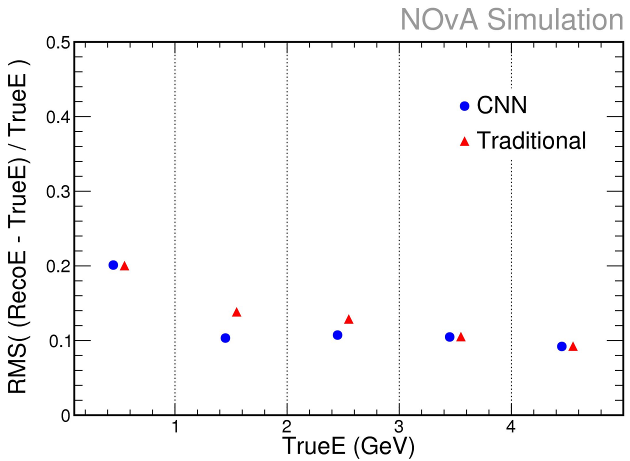
ν_μ Energy

- We estimate the ν_μ event energy with:
- $E(\nu_\mu) = E_\mu + E_{hadronic}$
- Traditional method of muon energy is based on length of muon prong
- We do this using both the traditional and CNN reconstructed hadron energy



ν_μ Energy





- Here we look at Mean and RMS of ν_μ energy resolution over ranges of true ν_μ energy
- The CNN has a smaller bias at lower energies
- The CNN's RMS does about as well or better everywhere

Conclusions

- NOvA has used regression CNNs to develop energy estimators that take raw cell outputs as an input image of the event
- The ν_e energy estimator has improved bias and resolution over tradition ν_e estimators, as well as more stable systematics
- Regression CNNs are also effective for leptonic and hadronic energy
- NOvA is currently working towards using CNNs to develop a better hadronic energy estimator for application to ν_μ event energy
- Better energy regression performance will give us better measurements of oscillation parameters