

Tiny Update on: Deep Learning Studies with the CALICE AHCAL Technological Prototype

Erik Buhmann¹



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG



¹ eMail: erik.buhmann@desy.de

The last few days

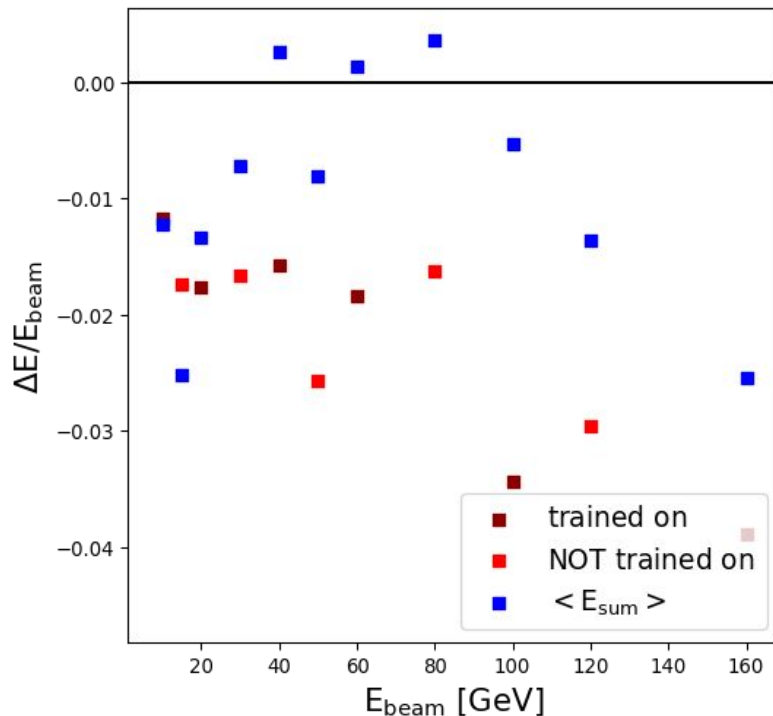
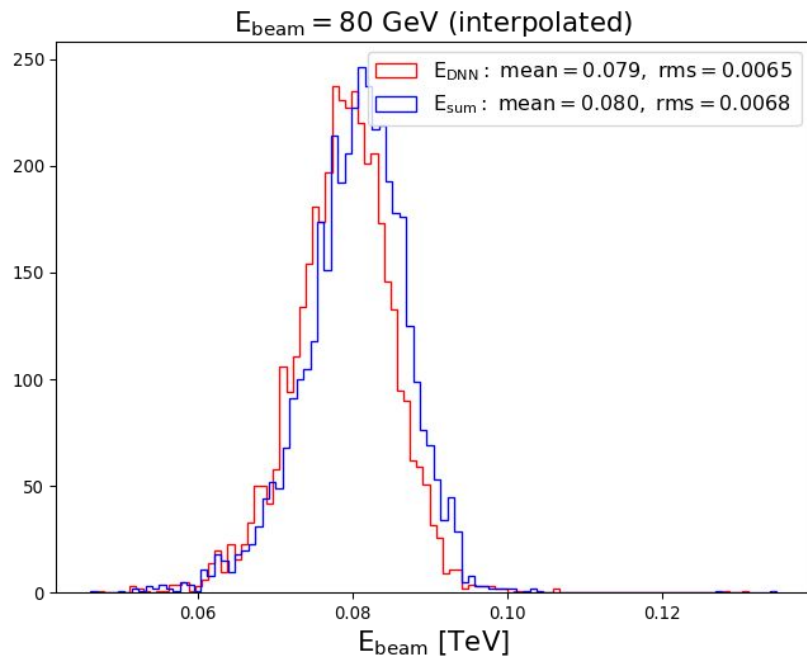


- ✓ Rewrote generator to load images into RAM
 - Increases epoch speed from 60 s to 7 s → whole training from i.e. 30 min to 5 min
- ✓ Implemented generator to flip & rotate event display images
 - Data augmentation 'on-the-fly' on CPU during training
 - Increases image samples by a factor 12
- ✓ Investigated shift of mean in reconstructed energy
 - Changed loss function (metric for performance evaluation of training)
- ✓ Bonus: Hit time in simulation
 - Thanks to Christian for showing me the right steering file entry!

3D Conv. Filter Size = (24,24,38)



- **Shift in mean energy visible**
→ Maybe due to loss function (MSRE)



Example 1) Mean Squared Error (MSE):

$$L(E_{i,true}, E_{i,pred}) = \frac{1}{N} \sum_i (E_{i,pred} - E_{i,true})^2$$

Example 2) Mean Squared Relative Error (MSRE):

$$L(E_{i,true}, E_{i,pred}) = \frac{1}{N} \sum_i \left(\frac{E_{i,pred} - E_{i,true}}{E_{i,true}} \right)^2$$

Used so far

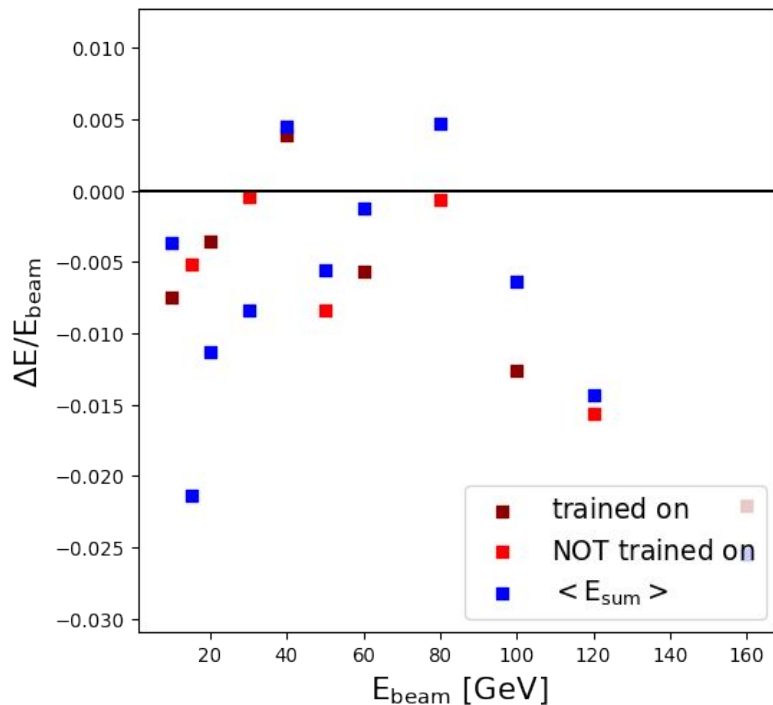
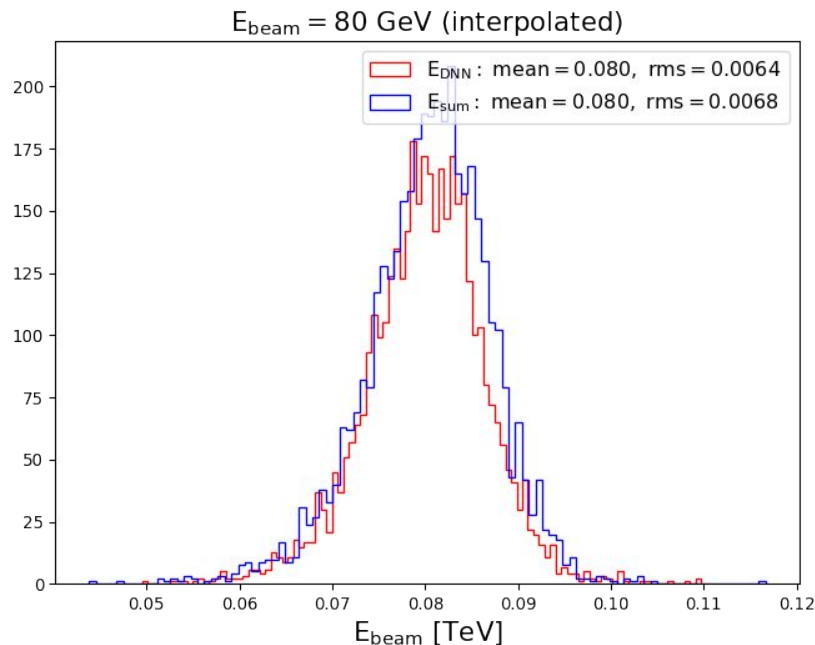
New Example 3) Mean Relative Squared Error (MRSE):

$$L(E_{i,true}, E_{i,pred}) = \frac{1}{N} \sum_i \left(\frac{E_{i,true} - E_{i,pred}}{\sqrt{E_{i,true}}} \right)^2$$

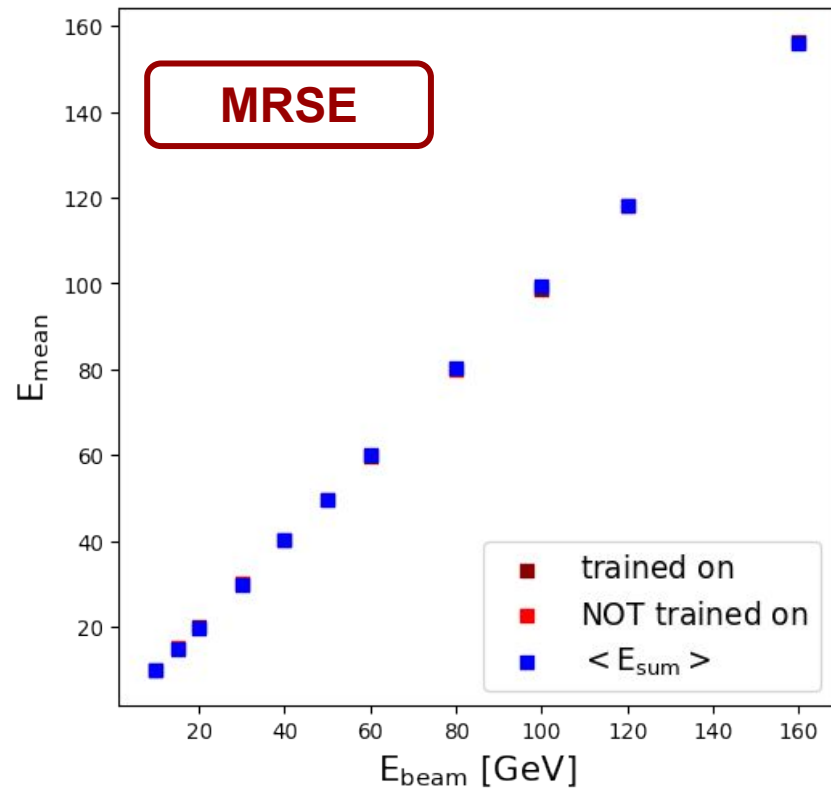
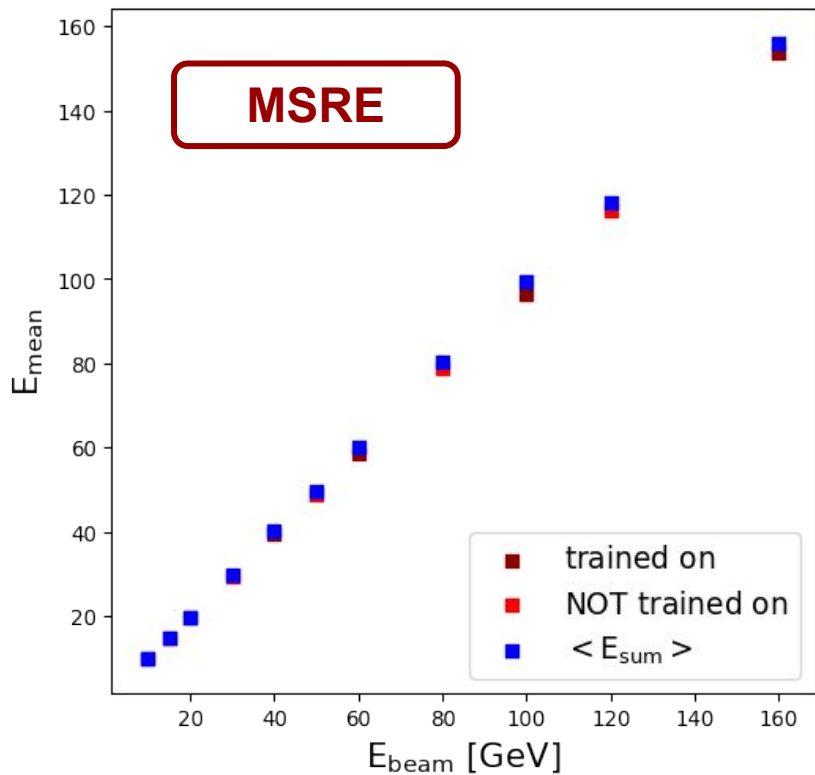
3D Conv. Filter Size = (24,24,38)



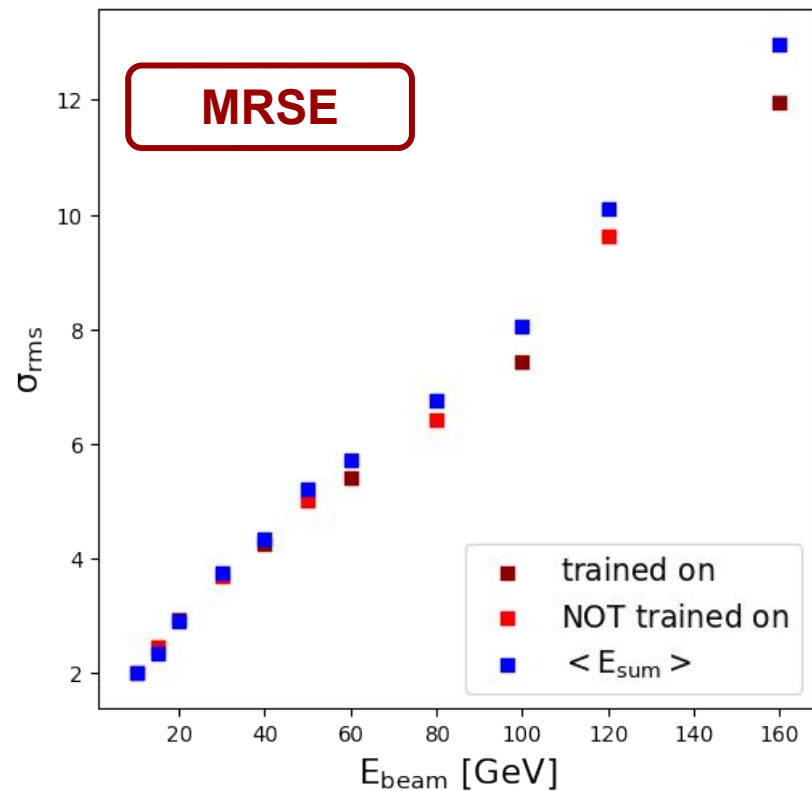
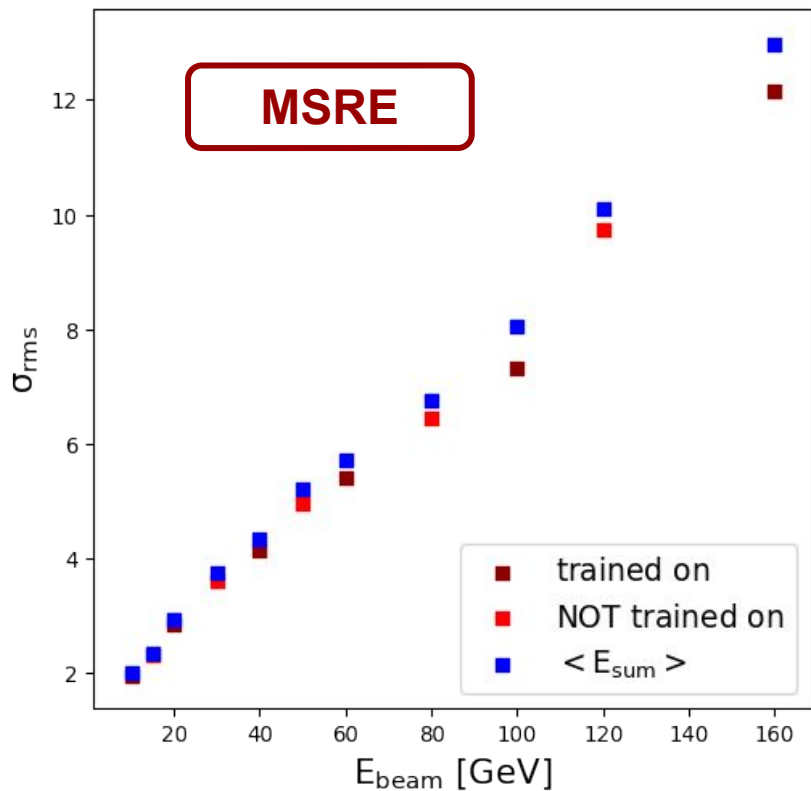
- Shift in mean energy reduced
→ Due to loss function (MRSE)



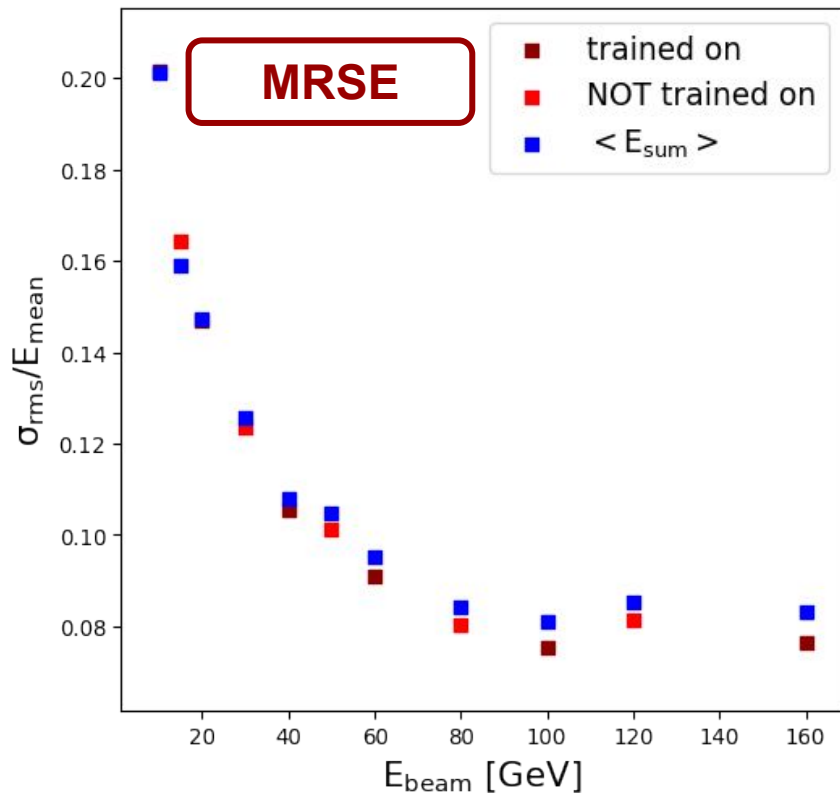
MSRE vs MRSE



MSRE vs MRSE



Summary & Outlook



- Training on simulation and adding time information as new dimension
- Locally Connected Layer for per-channel-calibration
- Multi particle separation, starting with Muons

Bonus Slides

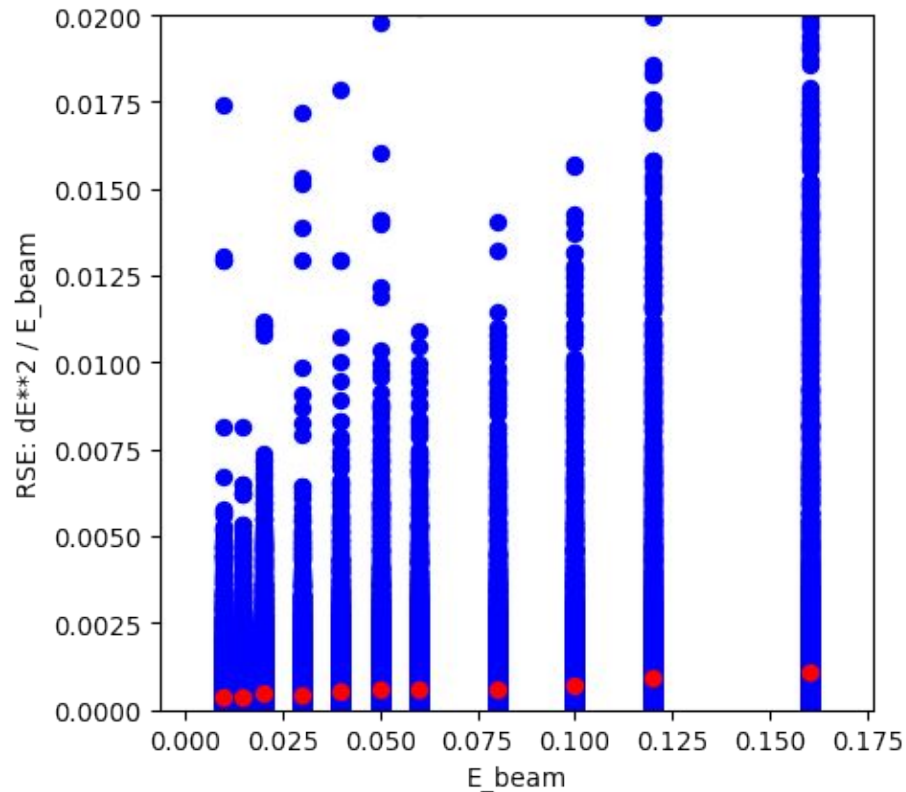
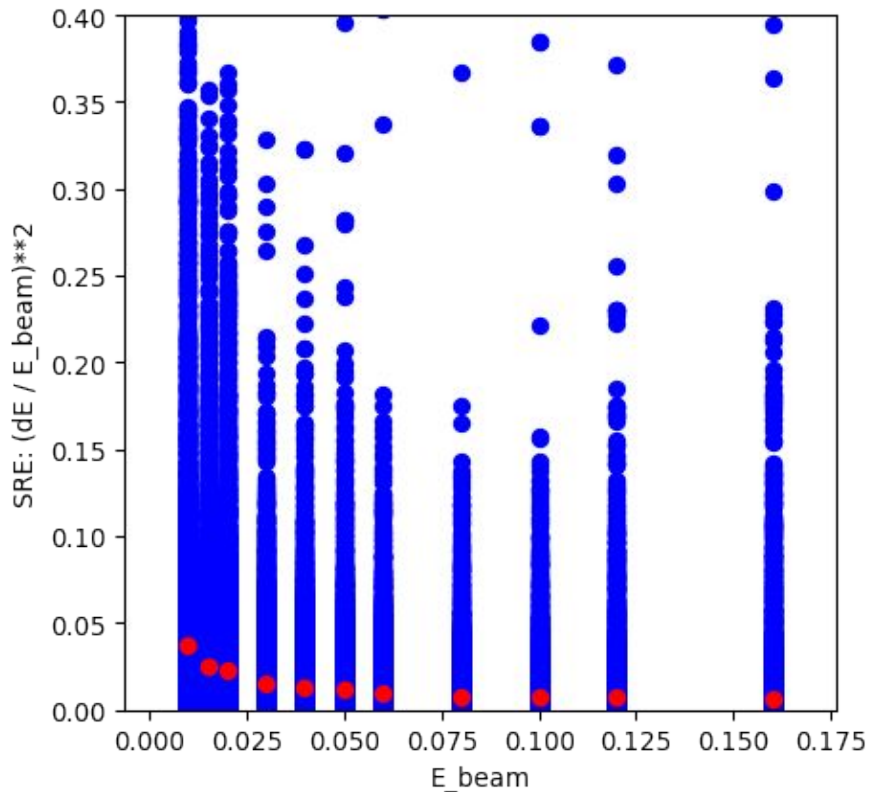


Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG



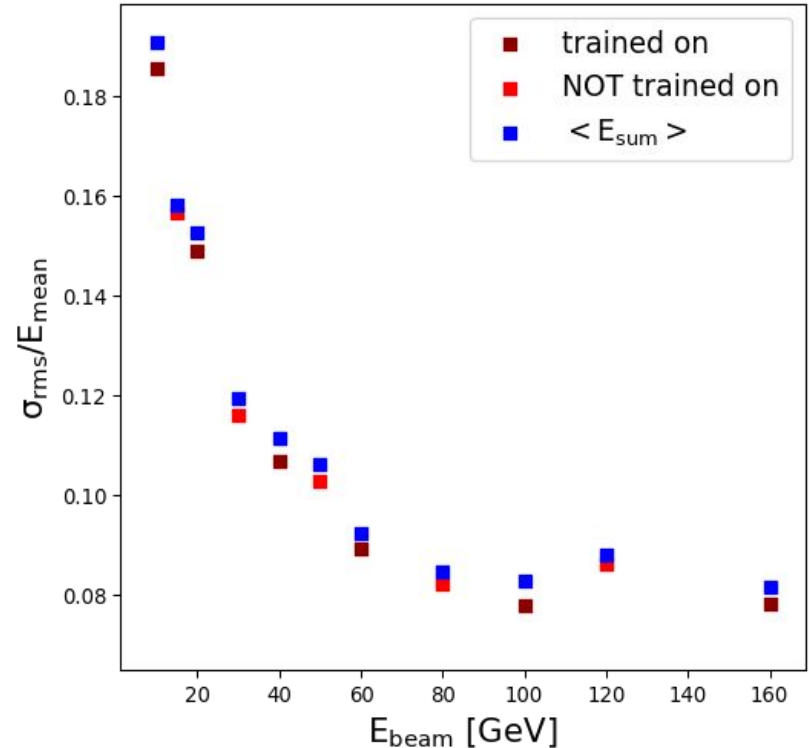
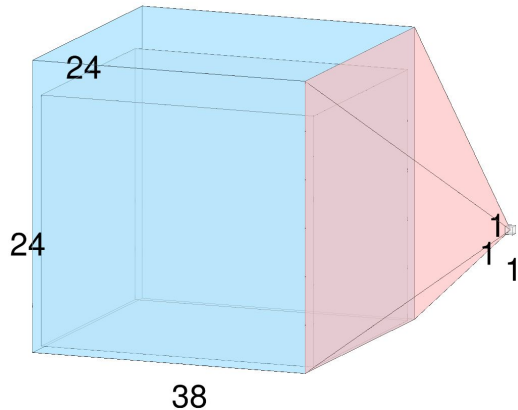
SRE vs RSE of training sample



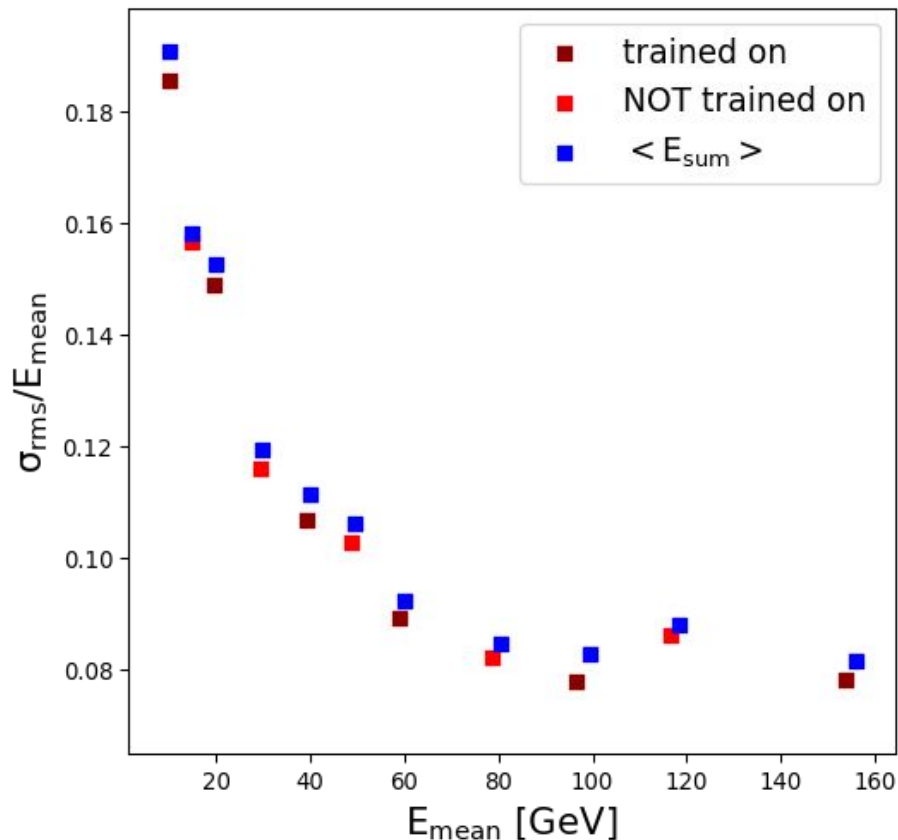
3D Conv. Filter Size = (24,24,38)



- One trainable convolutional layer with size of detector
- Output: Weighted sum of all channels
- Better performance than energy sum
- No systematic difference between “trained on” and “NOT trained on” test samples



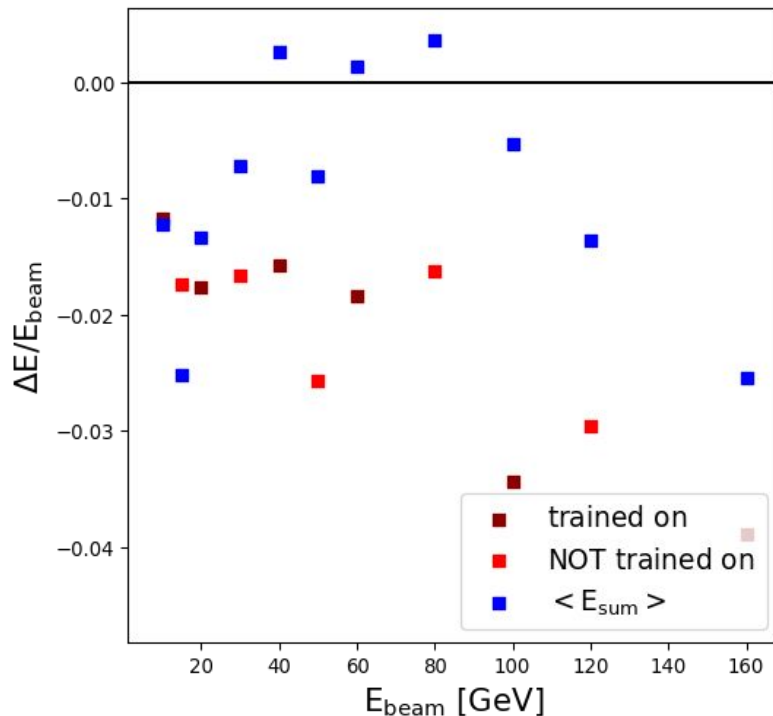
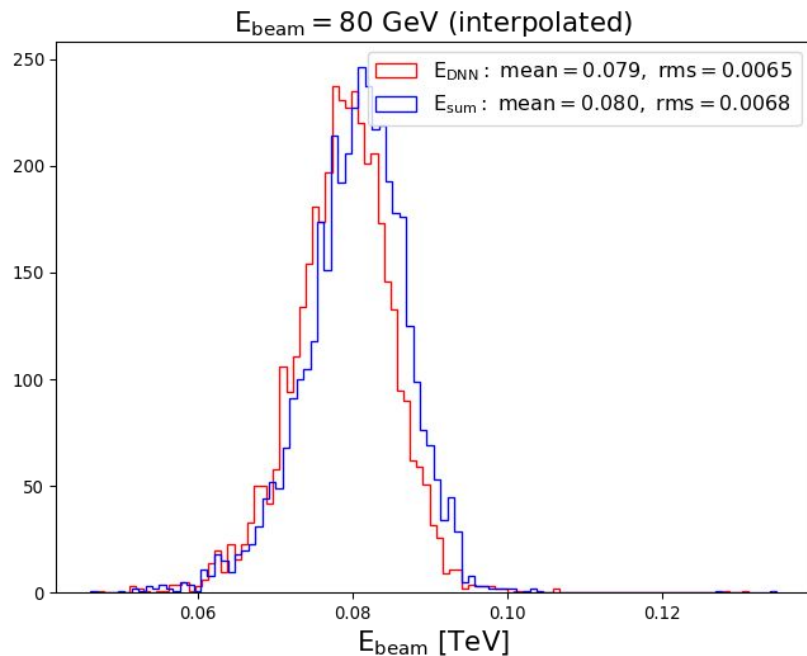
3D Conv. Kernel Size = (24,24,38)



3D Conv. Filter Size = (24,24,38)



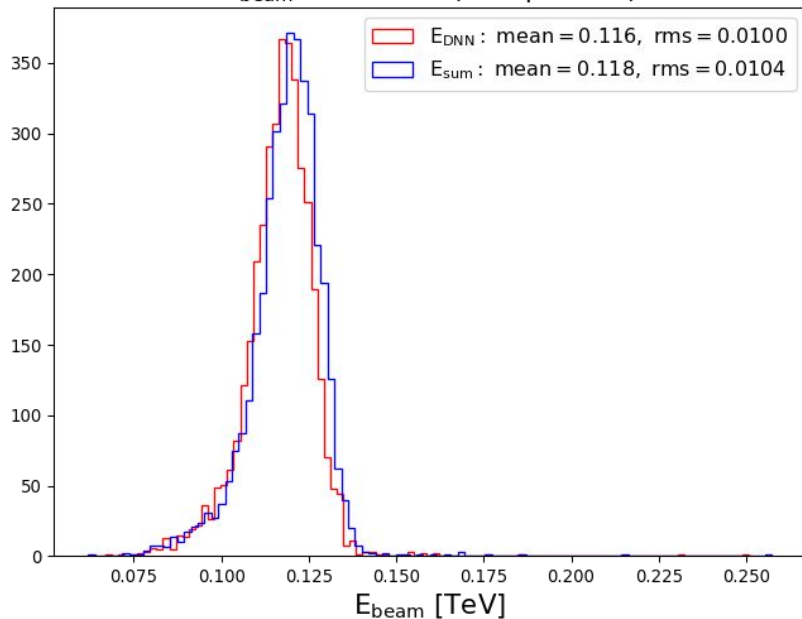
- **Shift in mean energy visible**
→ Maybe due to loss function (MSRE)



3D Conv. Kernel Size = (24,24,38)



$E_{\text{beam}} = 120 \text{ GeV}$ (interpolated)



$E_{\text{beam}} = 160 \text{ GeV}$

