An estimation of the muon signal in the SD using Deep Neural Networks

A. Guillén [1]   A. Bueno [2]   **J.M. Carceller*** [2]   J.C. Martínez-Velázquez [2]
G. Rubio [2]   C.J.T. Peixoto [3]   P. Sánchez-Lucas [2]

[1]Dpto. de Arquitectura y Tecnología de Computadores, Universidad de Granada, Granada.

[2]Dpto. de Física Teórica y del Cosmos & C.A.F.P.E., Universidad de Granada, Granada.

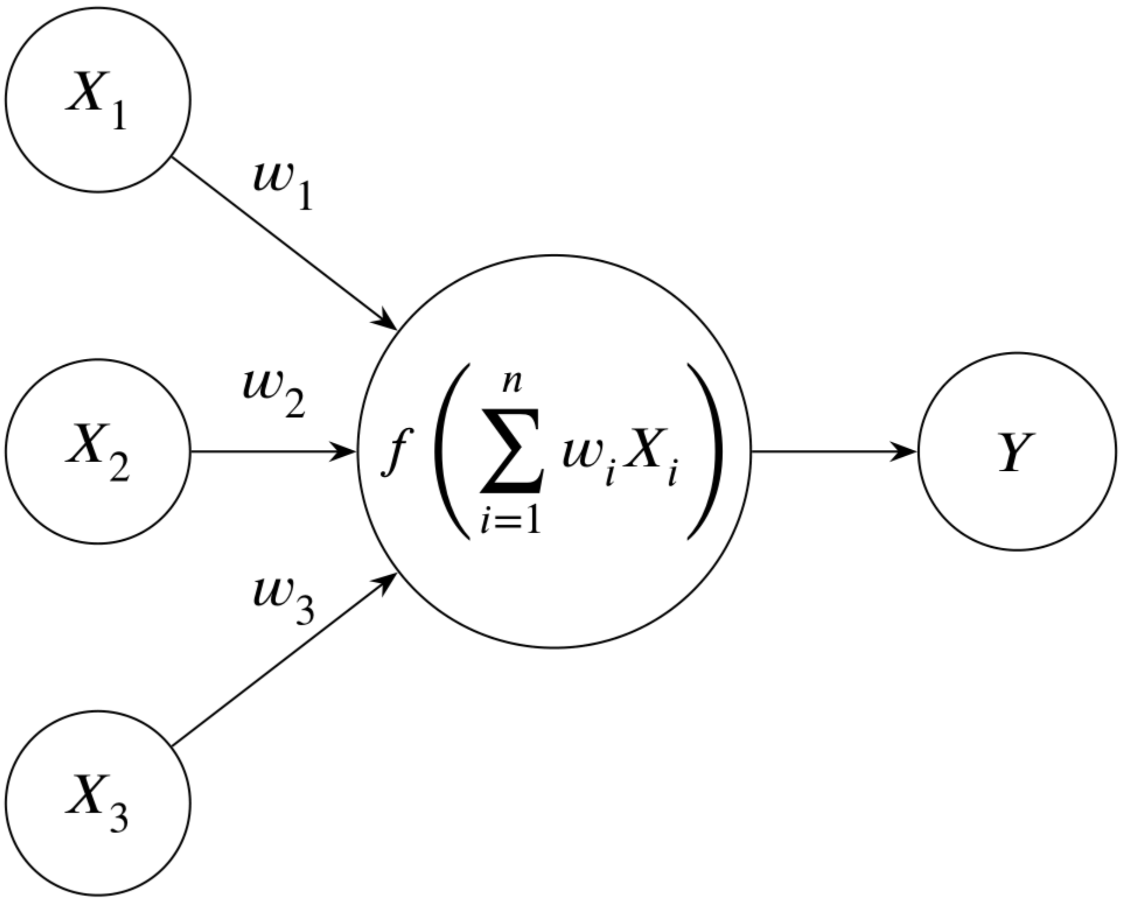
[3]Depto. de Ciências Básicas e Ambientais, Escola de Engenharia de Lorena, U. of São Paulo.

June 8, 2018

*Speaker

Artificial Neural Networks (NNs)

- Artificial Neural networks are loosely based on human neural networks.
- At each neuron, the value of the activation function f is computed based on the values of some inputs and the weights (free parameters) w that connect one neuron to another.

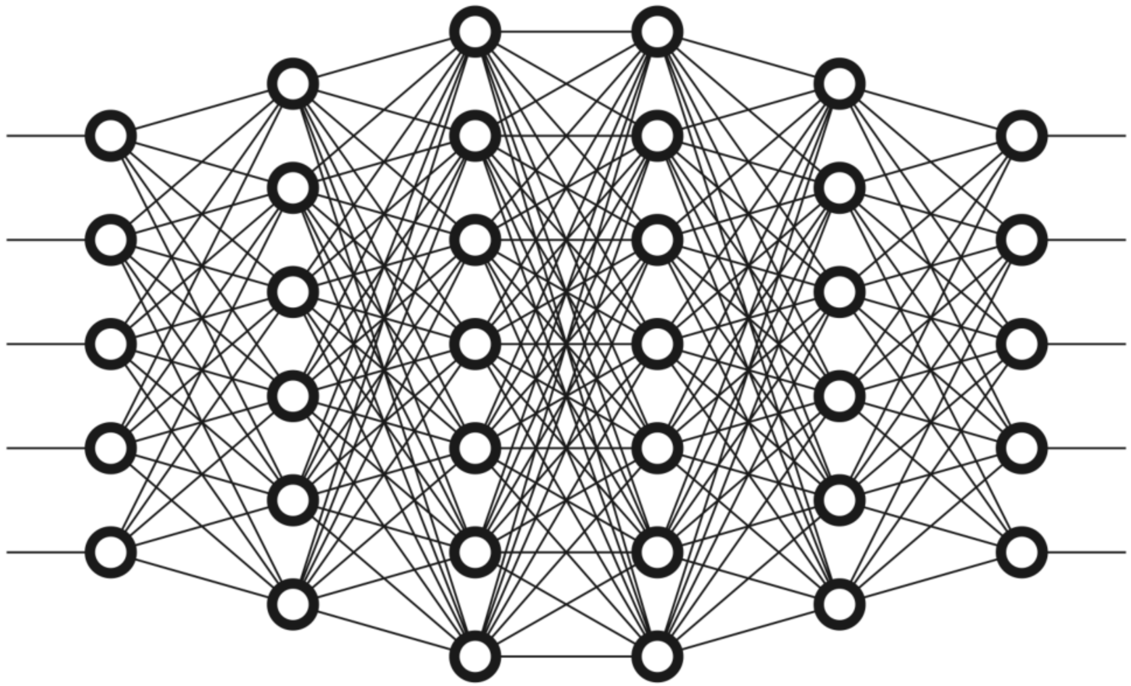


Example: Single neuron with three inputs X_1, X_2, X_3 and one output Y

f is a function of the linear combinations of the weights w and the inputs X

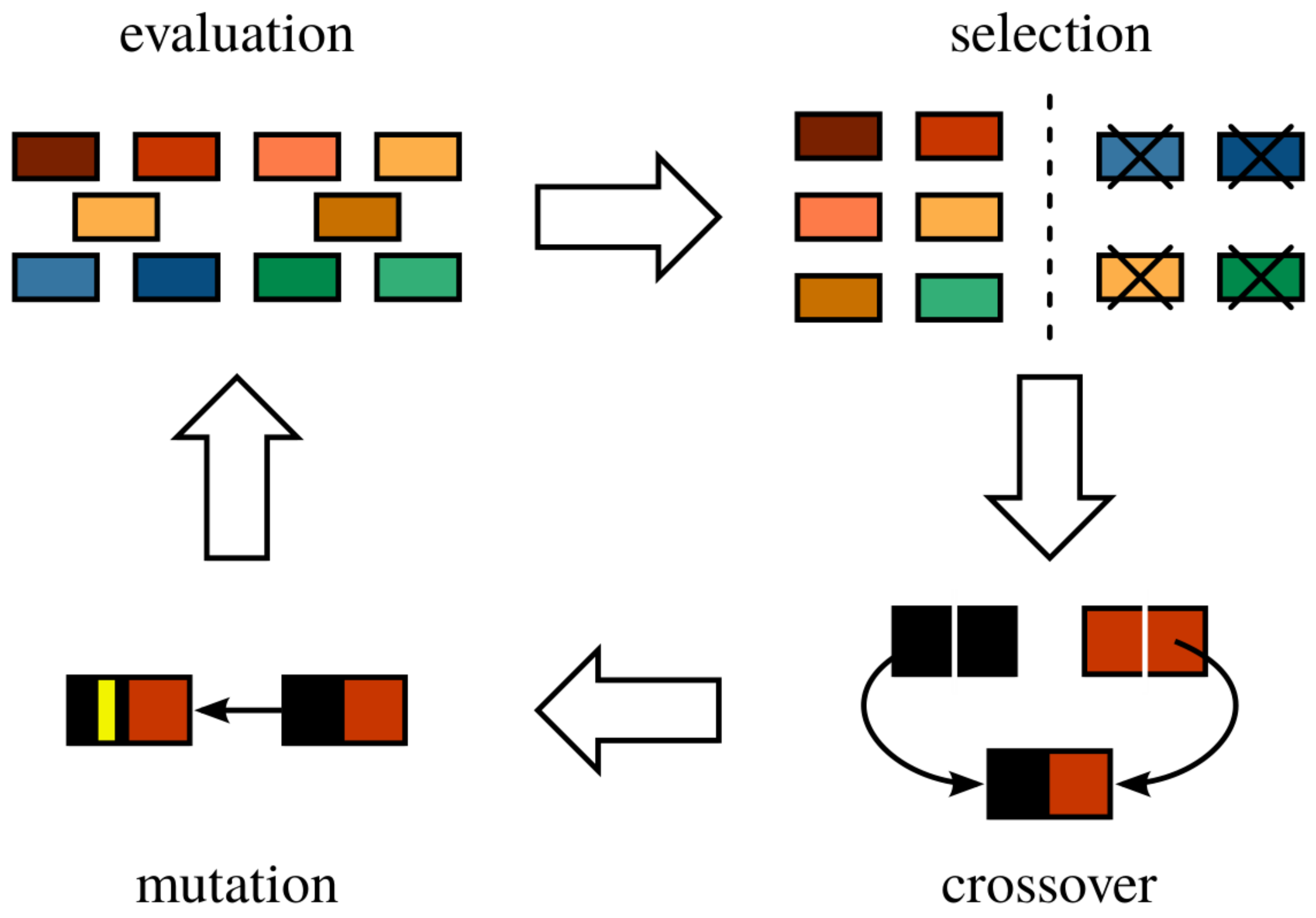
How does a neural network learn?

- The neural networks performs a <u>fit on the training data</u>, minimizing the average mean squared error between the prediction and the true values.
- During the training, the weights w of the neural network are modified so that the prediction minimizes the mean squared error.



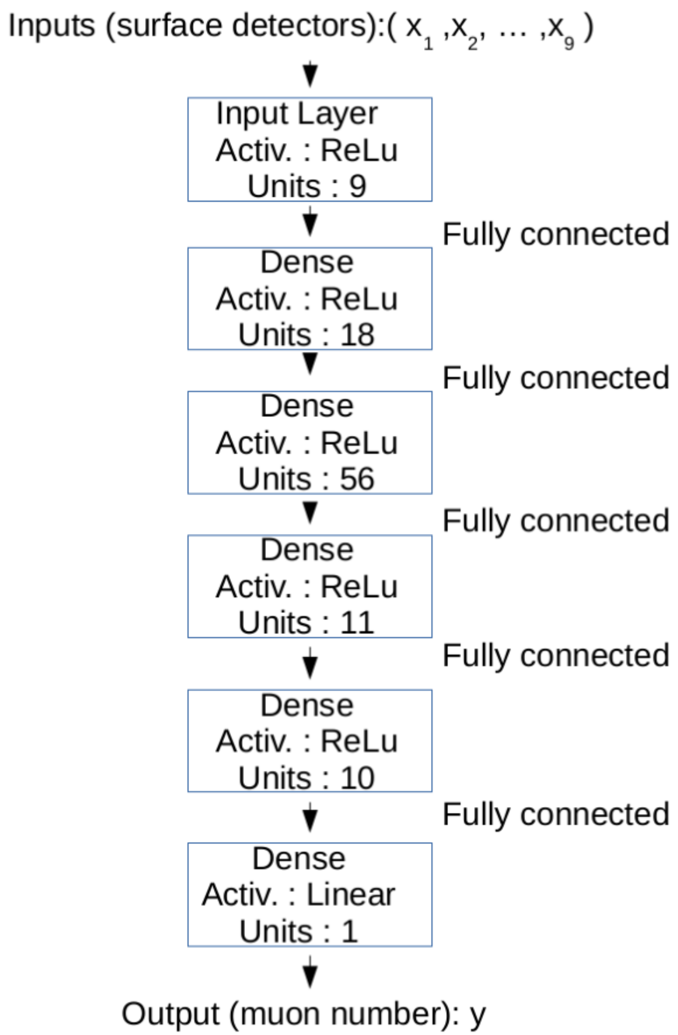
Choosing a Neural Network

- There are several free parameters that can be fined tuned to achieve better performance. For example, the number of layers of the NN or the number of neurons in each layer.
- We have used genetic algorithms to find a good NN

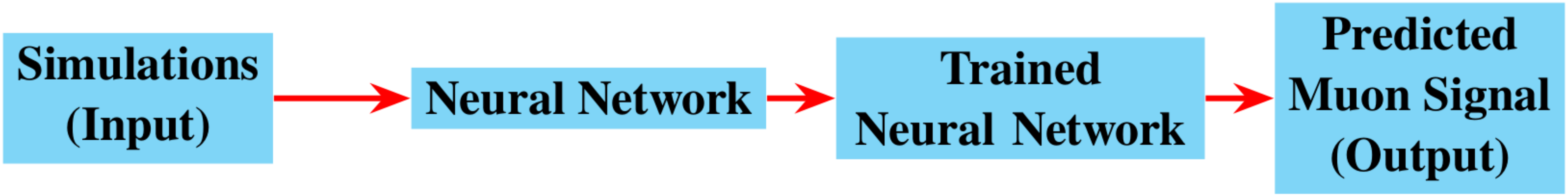


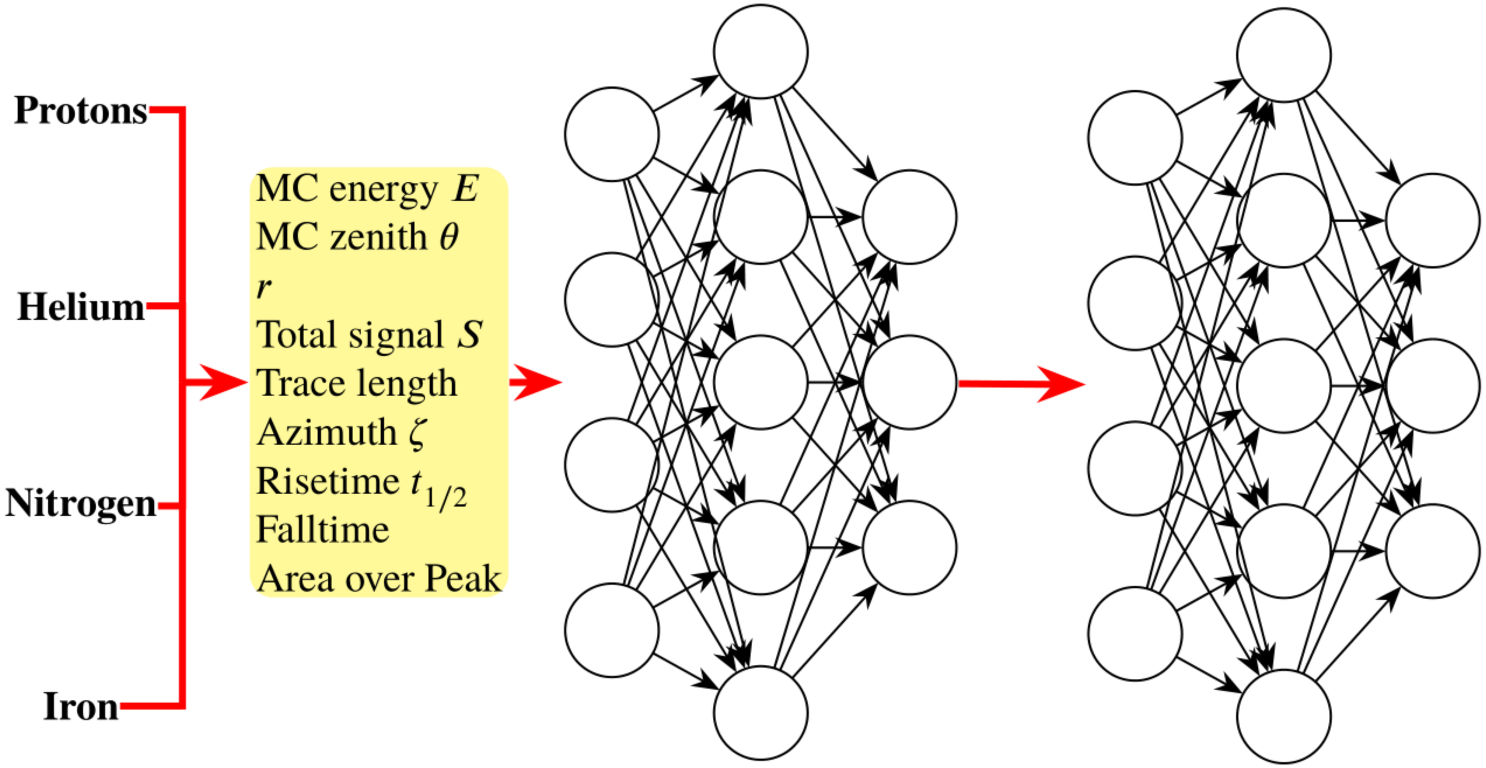
Our Neural Network

- The neural network used for this work has 6 layers, 2 of them being the input layer and the output layer.
- The number of neurons in each layer is 9, 18, 56, 11, 10, 1



Neural Networks: From input to output





- We use more simulations (that the net has not seen) to test the prediction.

- **One prediction for each station: S_{μ}^{true}**

How did we train the NN?

Training
25% Proton, Helium, Nitrogen and Iron (QGSJet)
Roughly 20000 stations of each primary.

Variables used in training

1. MonteCarlo energy E	In $\log_{10}$
2. MonteCarlo zenith θ	In $\sec\theta$
3. Distance to the core r	In m
4. Total signal S	In VEM
5. Trace length	
6. Azimuth ζ	In radians
7. Risetime $t_{1/2}$	In ns
8. Falltime	In ns
9. Area over Peak	

Remark: S_{μ}^{true} is not strictly given as input to the net, it is used to guide the training process.

⚠ Before training, the input is normalized to avoid operating with numbers that are very different in size.

Number of simulations

- Summary of the simulated events and detectors used in this work. Notice that the batches of stations used for training, validation and test correspond to distinct sets of detectors.

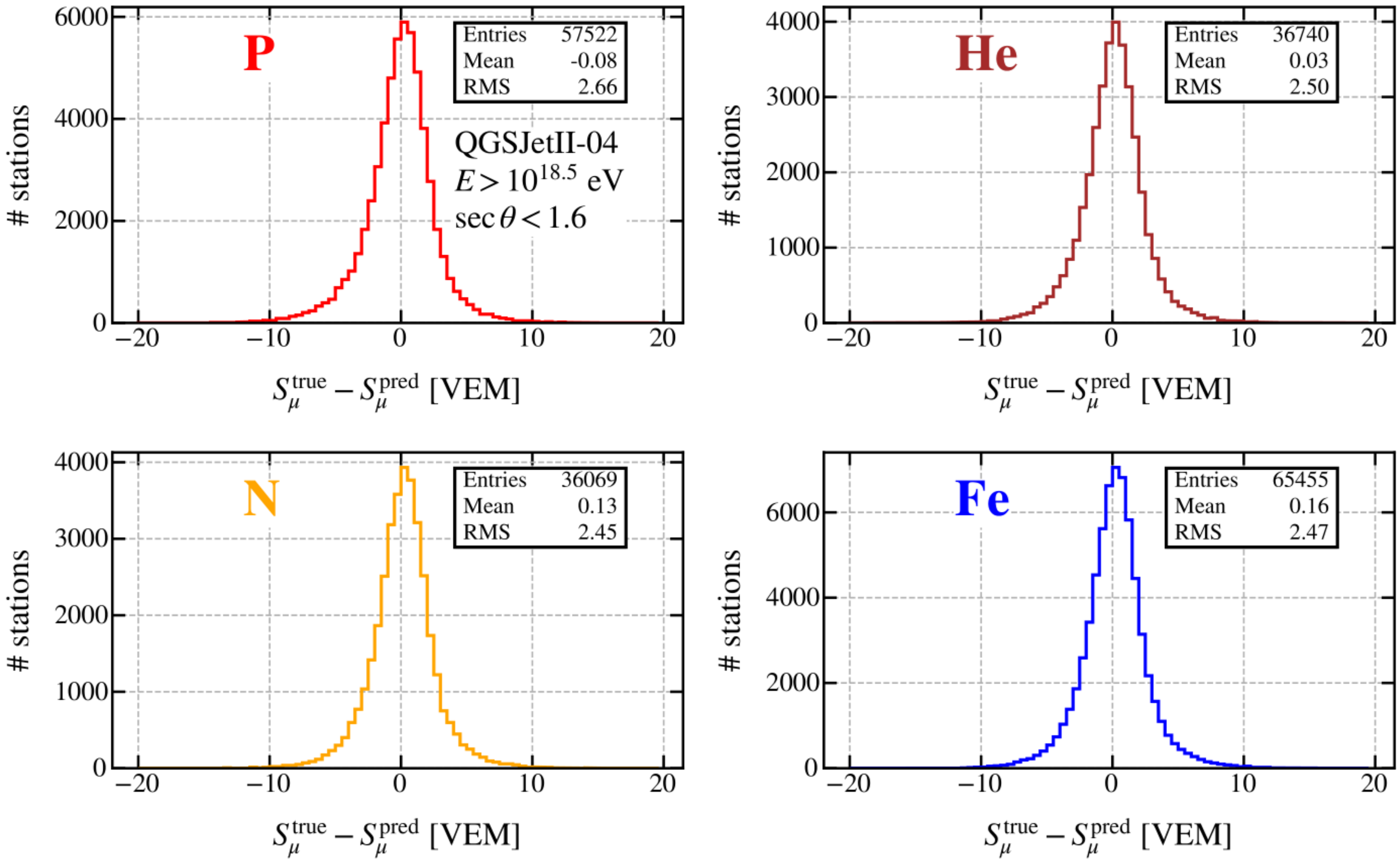
QGSJETII-04				
		Training	Validation	Test
Primary	# of events	# of detectors		
Proton	19362	16088	4022	57522
Helium	12341	15960	3989	36740
Nitrogen	12201	16071	4017	36069
Iron	19478	16076	4018	65455
EPOS-LHC				
		Training	Validation	Test
Primary	# of events	# of detectors		
Proton	18456	–	–	78063
Iron	18779	–	–	86862

Before starting

- The trace used is the average of the traces recorded by each of the active PMTs of the station.
- We do not use stations that have saturation in **any** of the two channels.
- We only use stations with total signal S_{total} above 10 VEM.
- We restrict to energies between $10^{18.5}$ eV and 10^{20} eV and zenith angles between 0 and 45 degrees.

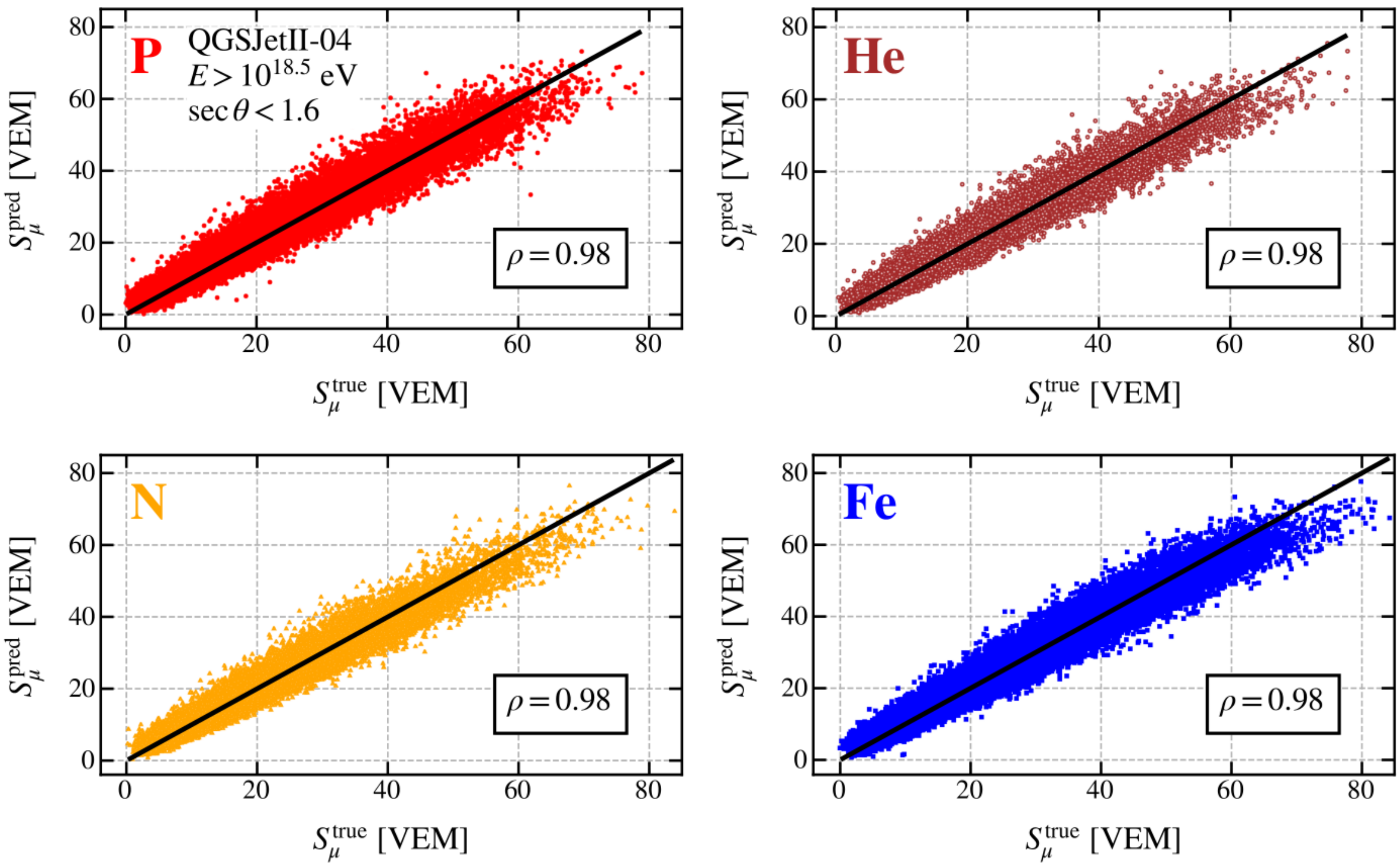
S_μ^{true}-S_μ^{pred} (station level)

- The difference between the real S_μ^{true} and predicted S_μ^{pred} signals is a few VEMs.
- The mean of the distribution is very close to zero, changing slightly as the mass of the primary increases.
- The RMS of the distribution is ~ 2.5 VEM.



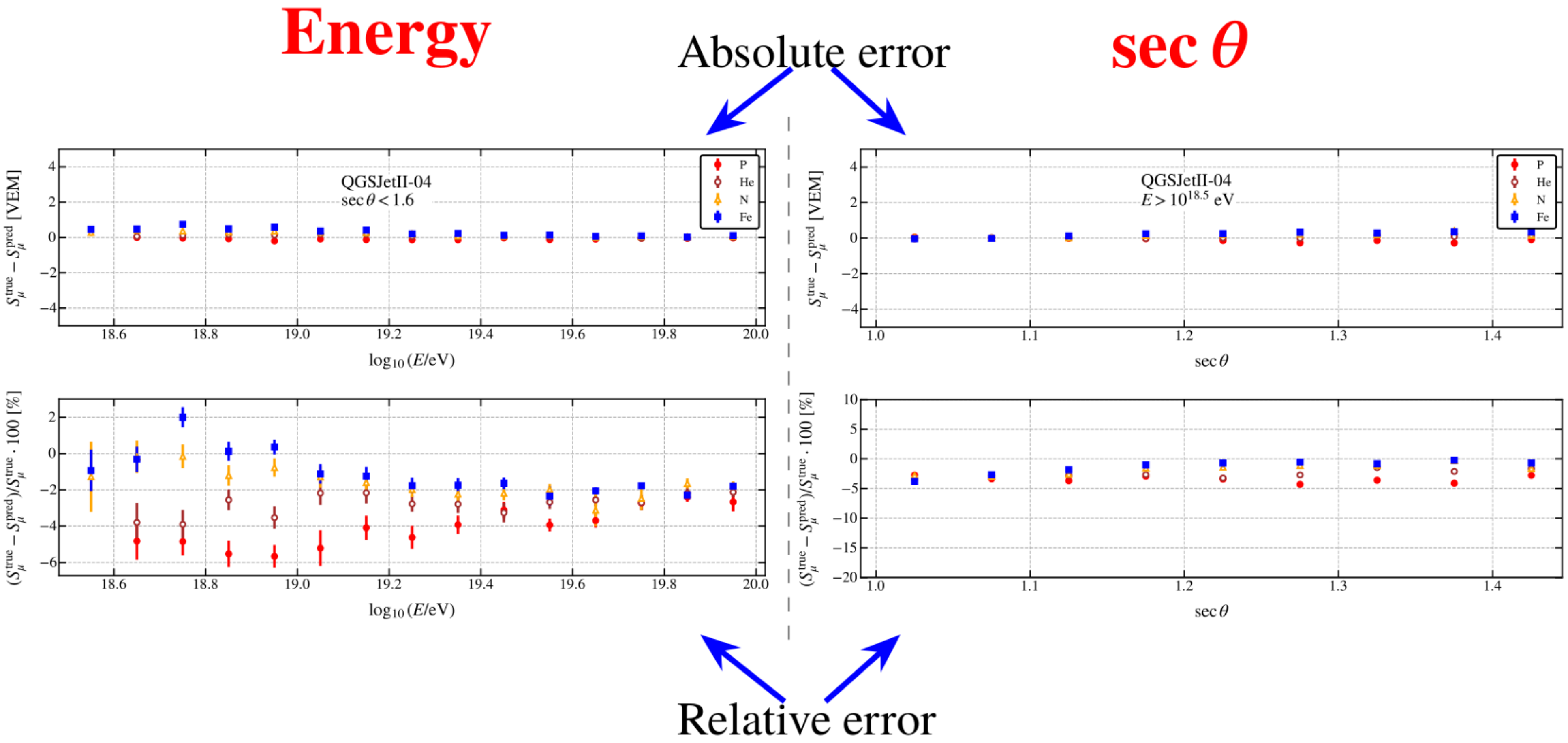
S_μ^{true}-S_μ^{pred} (station level)

- The correlation between S_μ^{true} and S_μ^{pred} is very high, with a Pearson correlation coefficient $\rho = 0.98$.



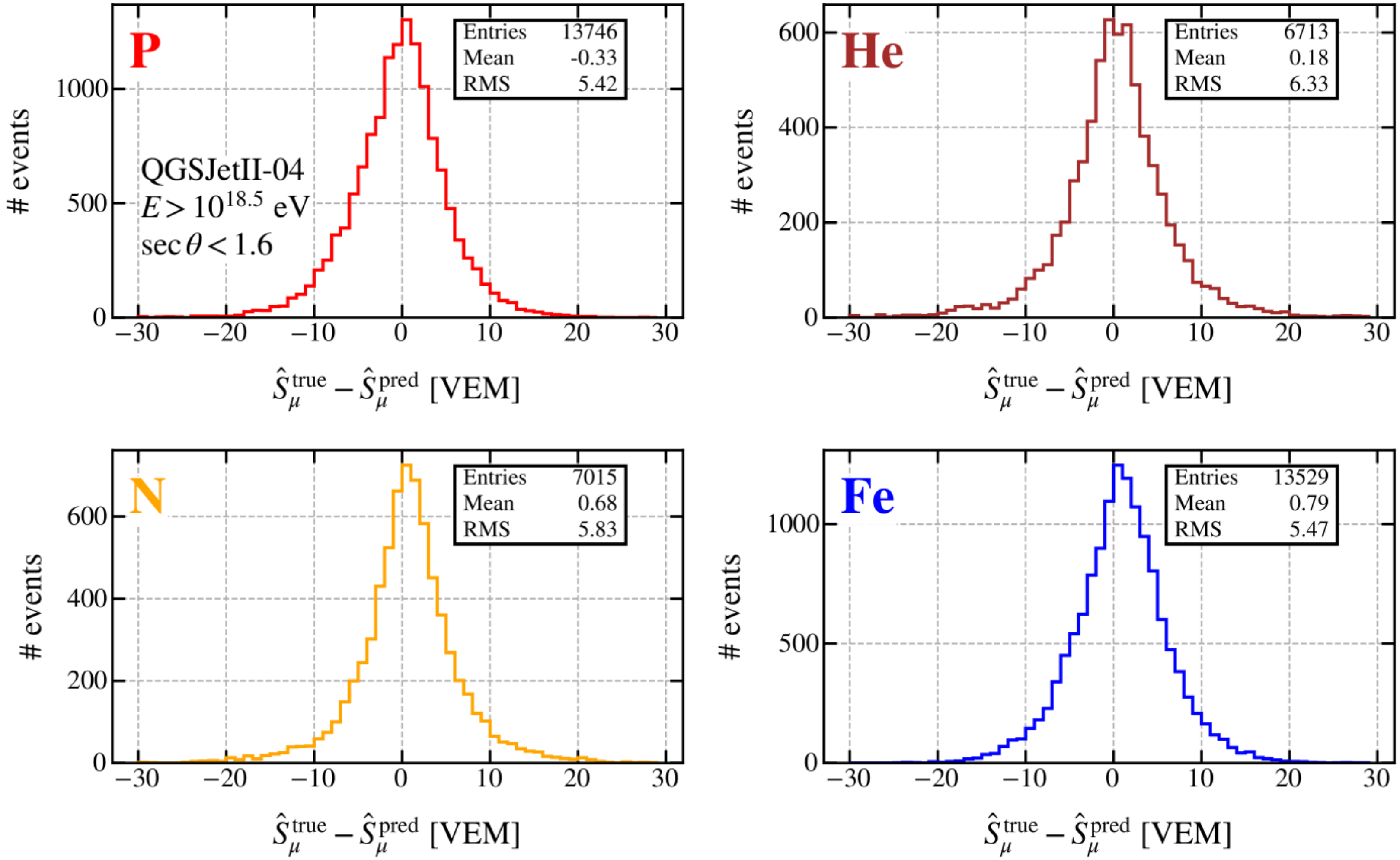
$S_\mu^{\text{true}} - S_\mu^{\text{pred}}$ as a function of the energy E and $\sec\theta$

- The bias with E and $\sec\theta$ is very small, being, at most, 1 VEM for the absolute bias and below 10% for the relative bias.



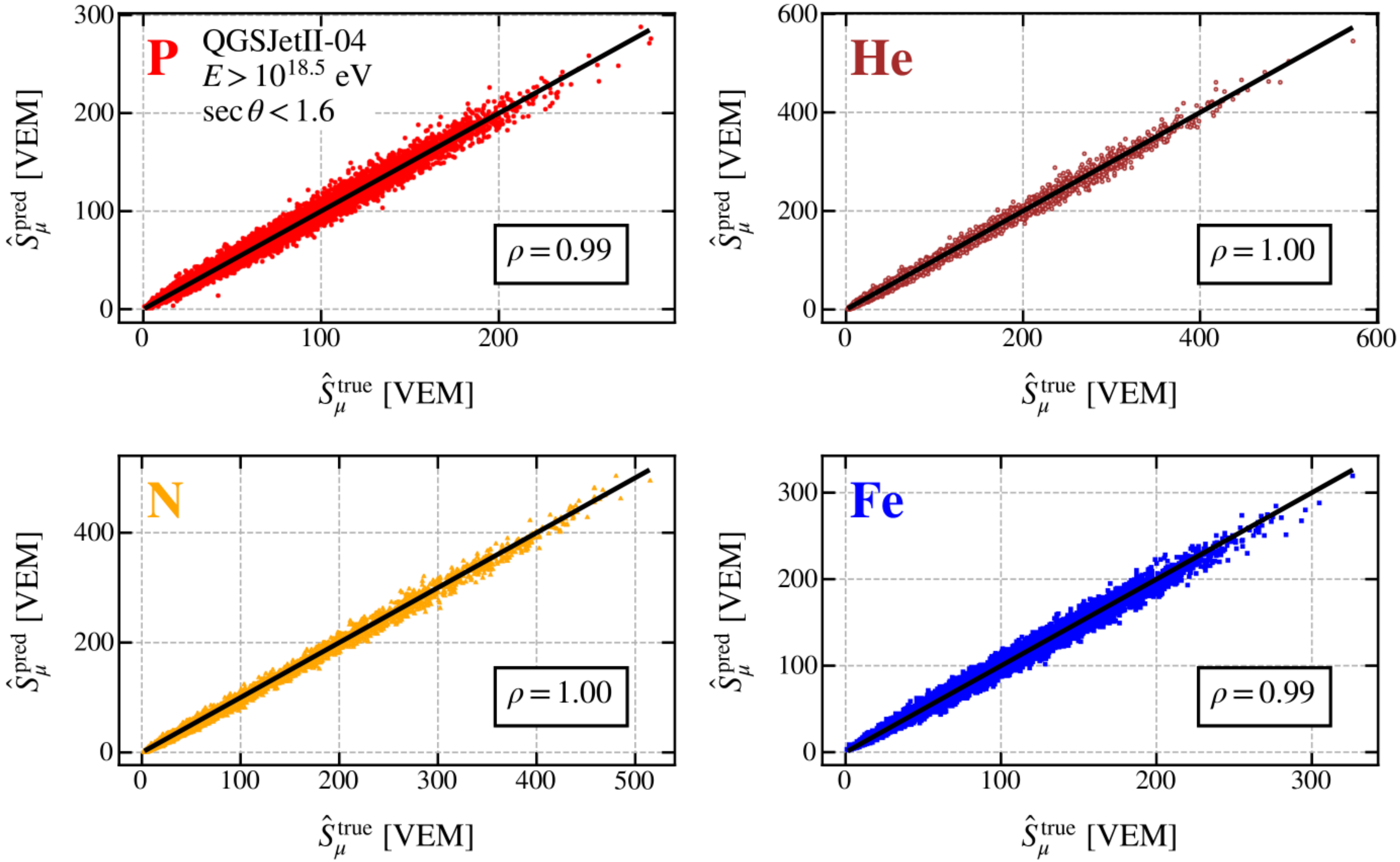
$\hat{S}_\mu^{\text{true}}$-$\hat{S}_\mu^{\text{pred}}$ (event level)

- **Important**: The variables at the event level will be represented with a hat (ˆ) on top of them. $\hat{S}_\mu^{\text{true}}$ is the sum of S_μ^{true} for each station in the same event.
- The mean is very close to zero and the RMS ∼ 6 VEM.



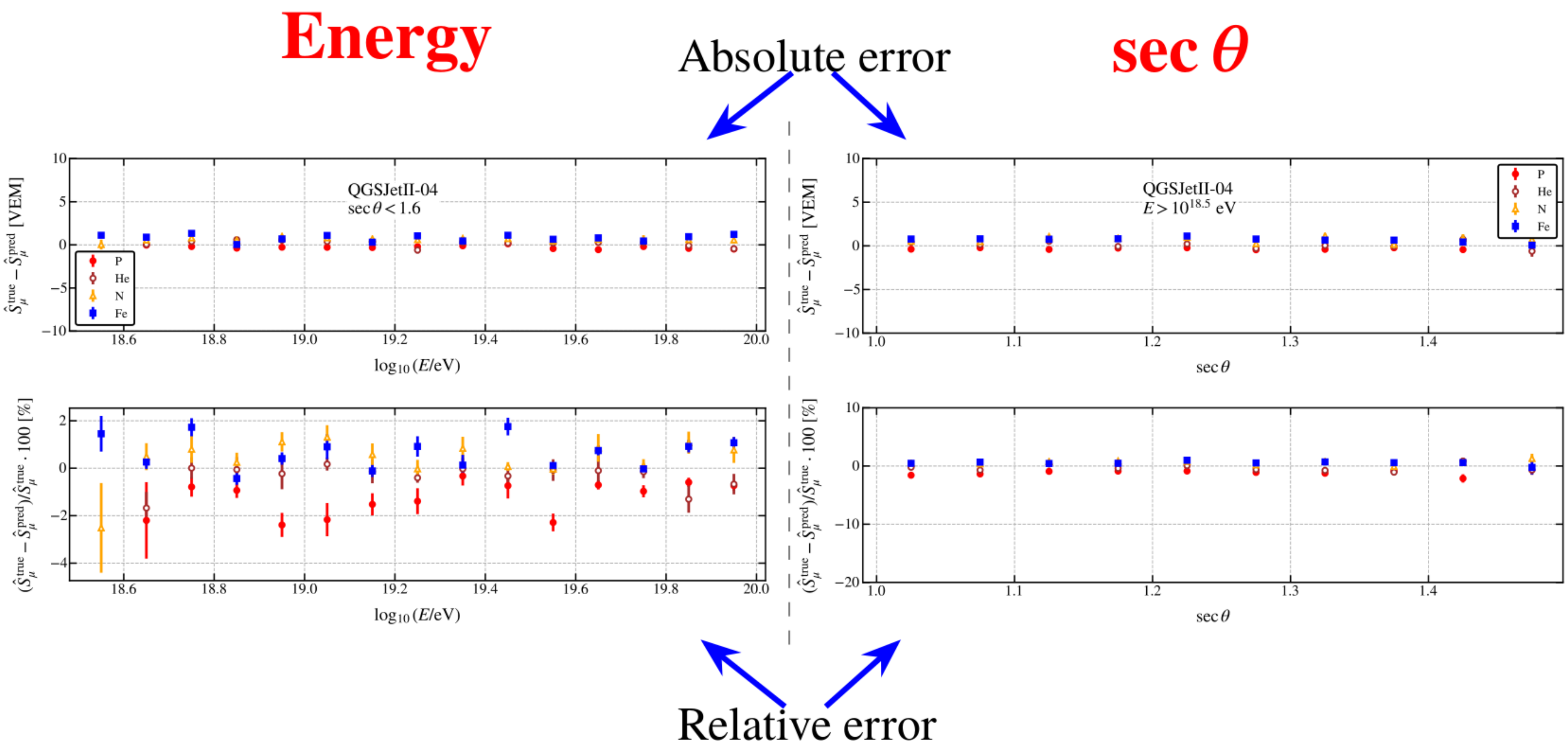
$\hat{S}_\mu^{\text{true}}$-$\hat{S}_\mu^{\text{pred}}$ (event level)

- The correlation between S_μ^{true} and S_μ^{pred} is very high, with a Pearson correlation coefficient $\rho = 0.99$.



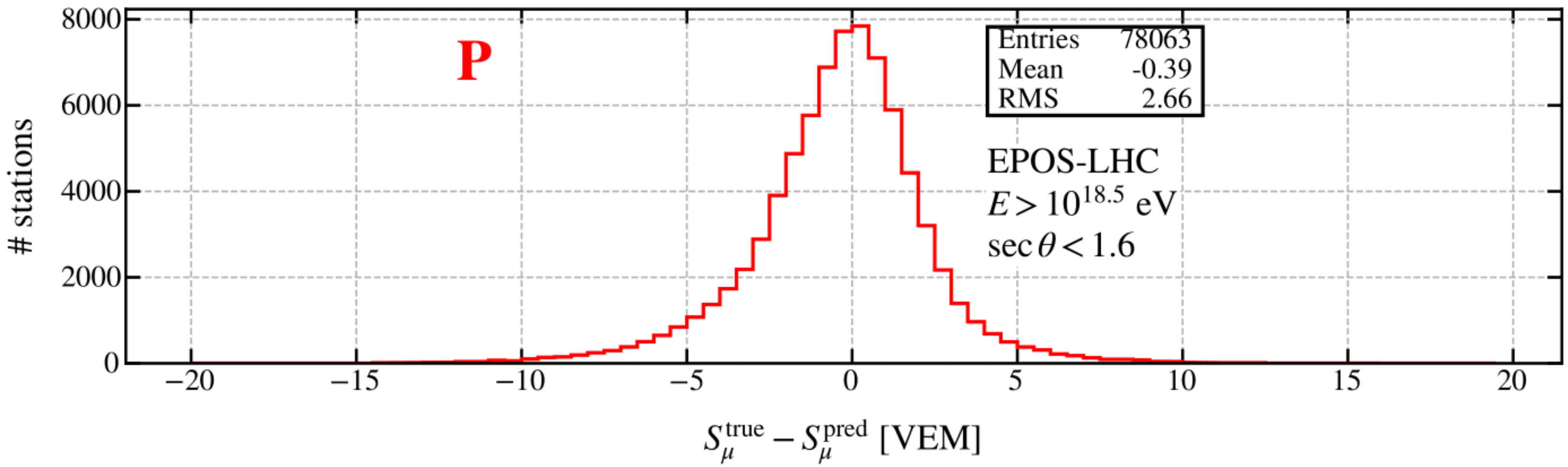
$\hat{S}_\mu^{\textbf{true}}$-$\hat{S}_\mu^{\textbf{pred}}$ vs E and $\sec\theta$ (event level)

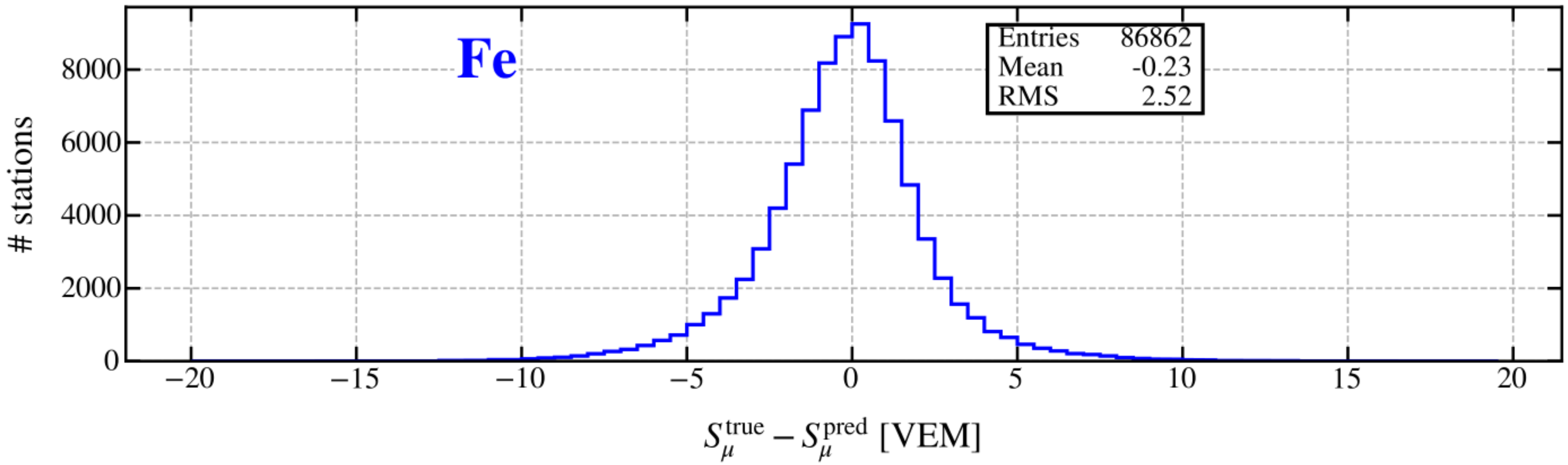
- The bias with E and $\sec\theta$ at most, 2 VEM for the absolute bias and below 6% for the relative bias.



S_μ^{true}-S_μ^{pred} for EPOS (station level)

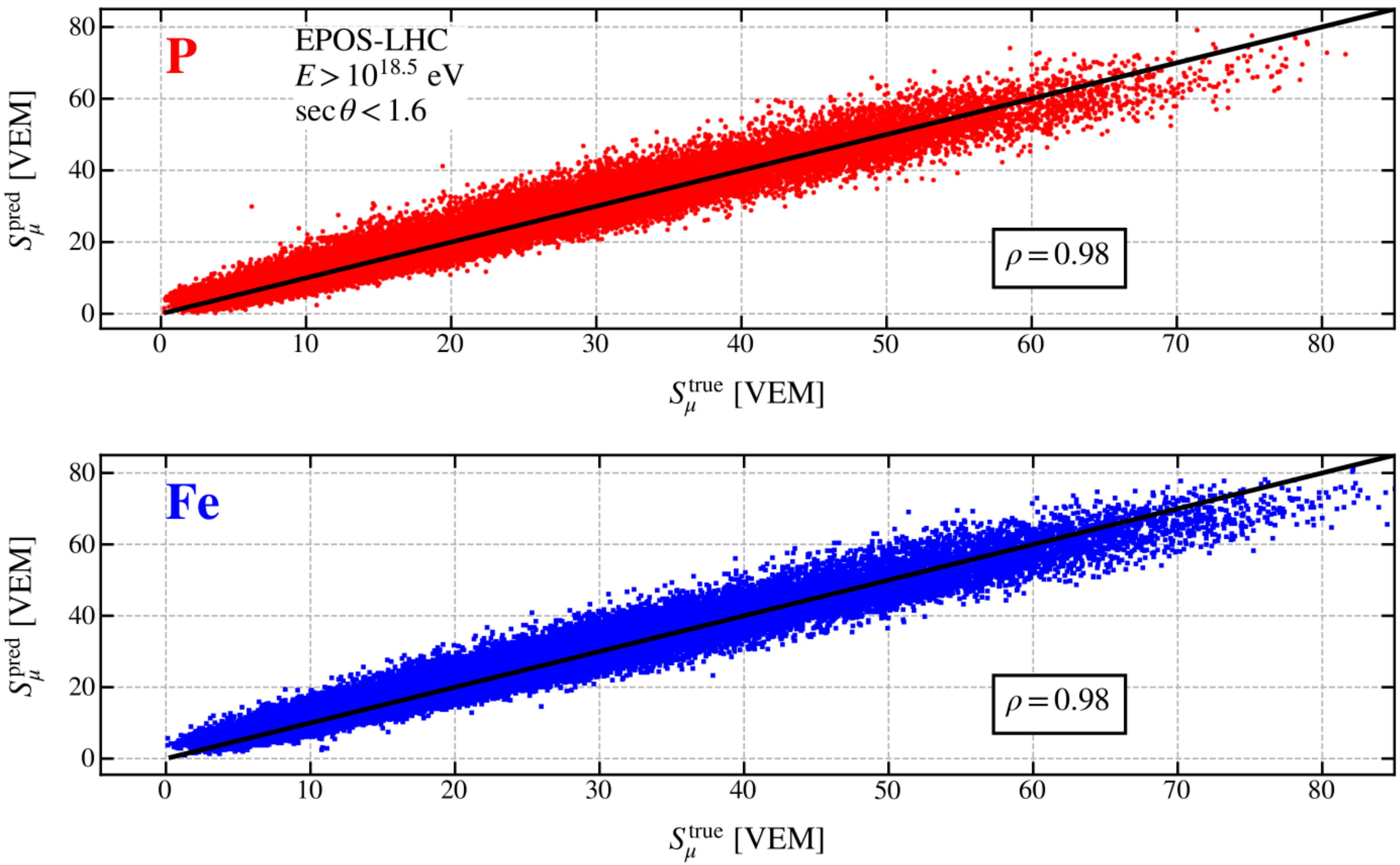
- The estimation performs very similarly to the one done for QGSJet.
- The mean of the distribution is slightly biased towards less muons while the RMS of the distribution is still ~ 2.5 VEM.





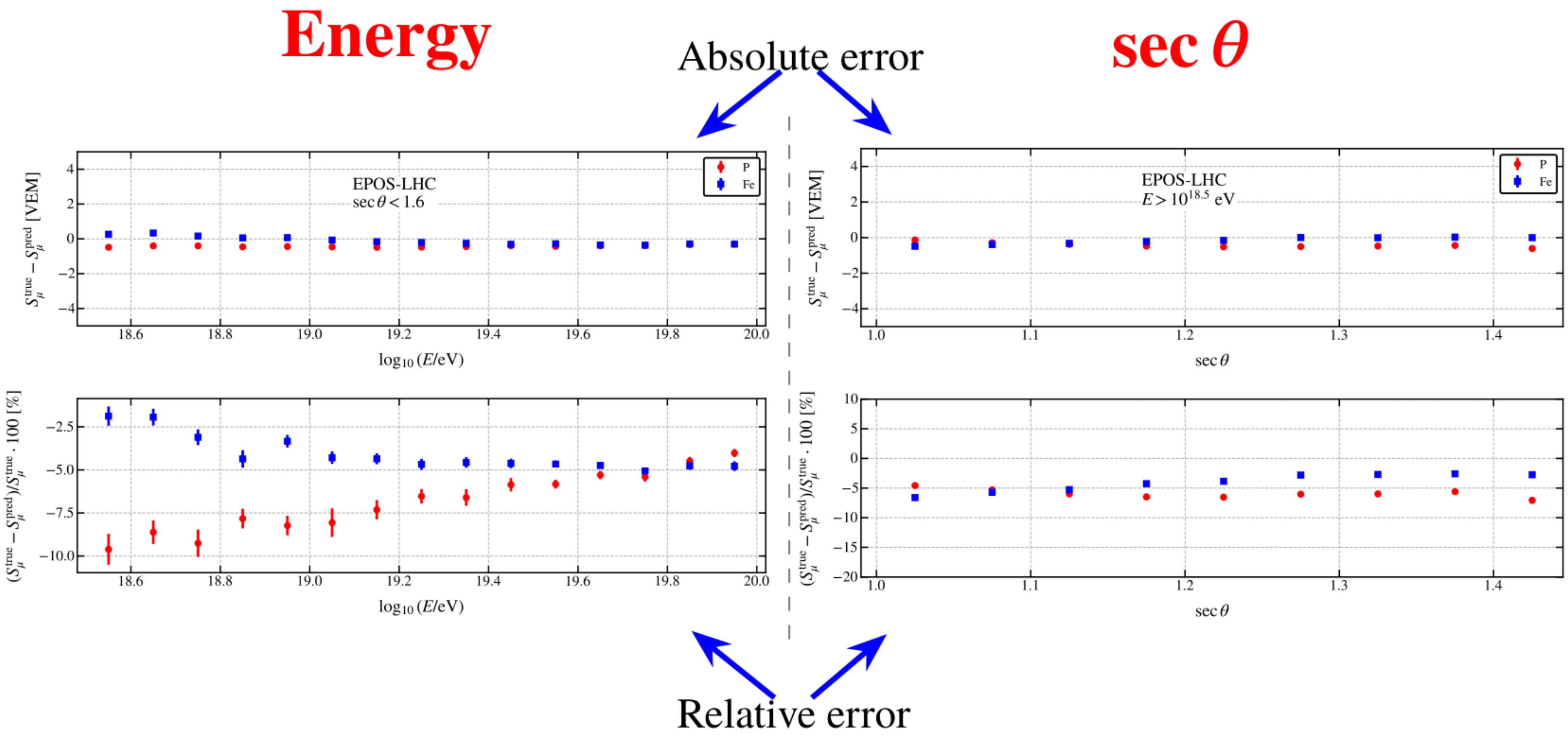
S_μ^{true}-S_μ^{pred} for EPOS (station level)

- The correlation between S_μ^{true} and S_μ^{pred} is very high, with a Pearson correlation coefficient $\rho = 0.98$.



S_μ^{true}-S_μ^{pred} for EPOS vs E and $\sec\theta$ (station level)

- The biases with E and $\sec\theta$ are below 1 VEM and a $\sim 10\%$ at most for the relative difference.



Summary of the first part

- We have proved that using Neural Networks, one can predict the muon signal with very low uncertainty.
- The muon signal can be consistently predicted across a large range of energies and zenith angle, within 2 VEMs of the true value in each station.
- The estimation shows a very small bias with distance to the core r, energy E, zenith angle $\sec\theta$ and **hadronic model** used.

GAP2018_14