Los Alamos
NATIONAL LABORATORY

Nuclear masses learned from a probabilistic neutral network

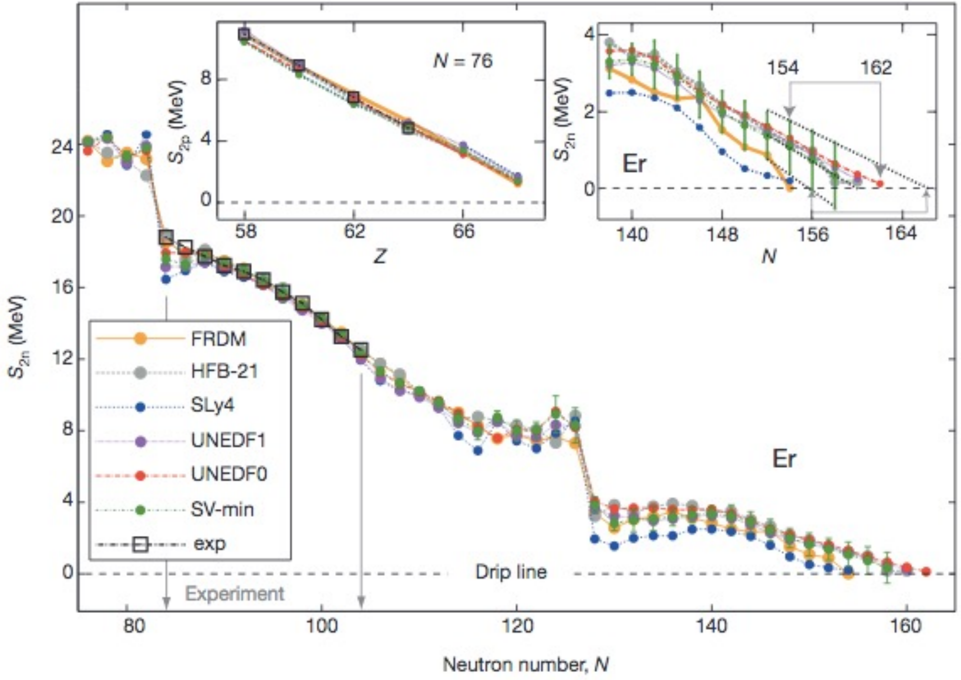
A.E. Lovell, with A.T. Mohan, T.M. Sprouse, and M.R. Mumpower

ISNET-9
May 24, 2023

LA-UR-23-25313

Lovell, *et al.*, *PRC* **106**, 014305 (2022)

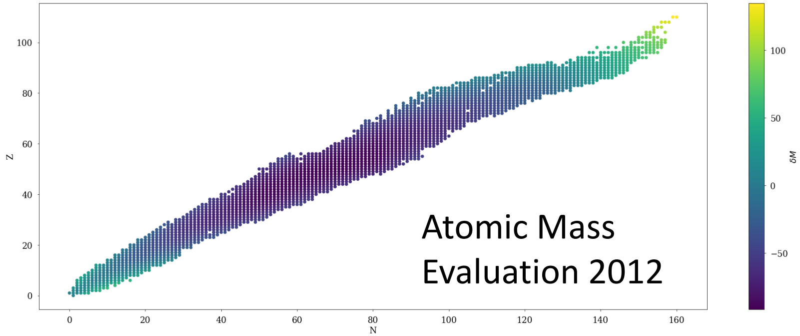
Masses are important input for a variety of nuclear physics calculations, the experimental reach is limited



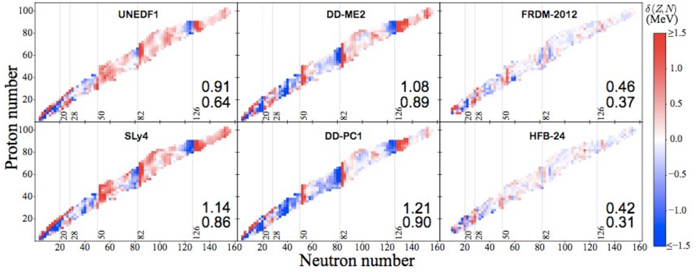
J. Erler, et al., *Nature* **486**, 509 (2012)

$$Q = (m_A + m_B) - (m_c + m_D)$$

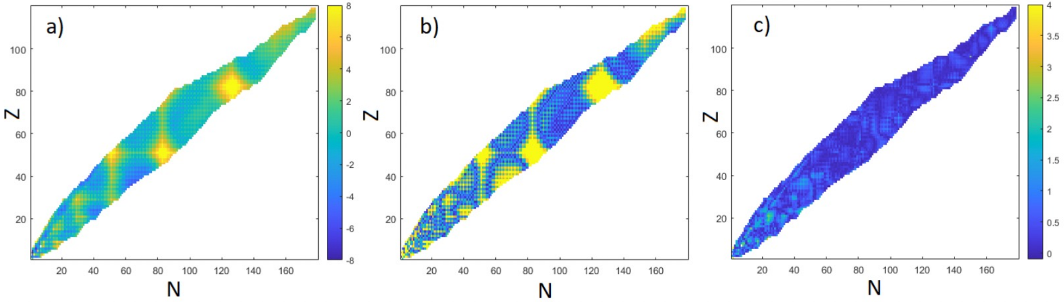
$$BE = Zm_p + Nm_n - \delta M - (Z+N)u$$



Numerous groups have used machine learning to predict nuclear masses

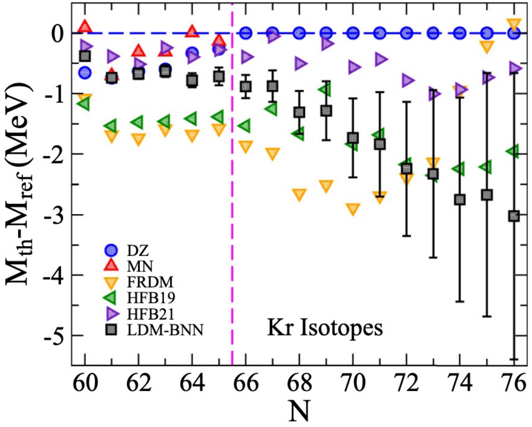


L. Neufcourt*, et al.*, *PRC* **98**, 034318 (2018)

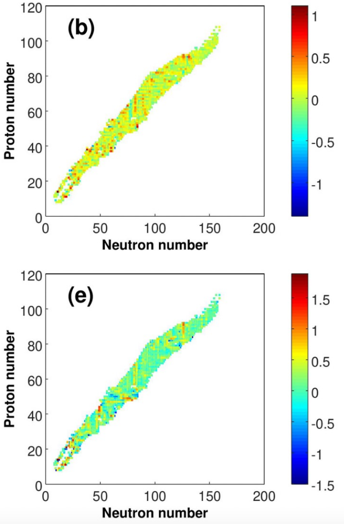


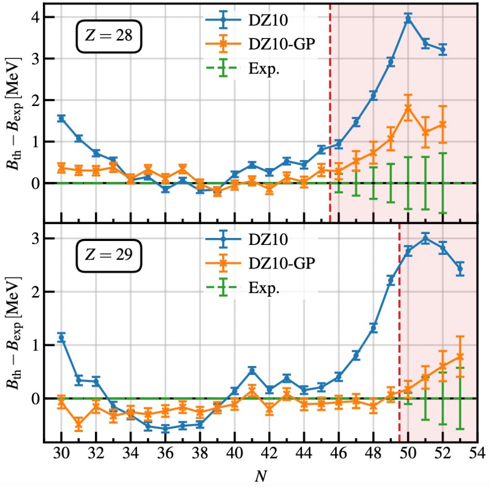
B. Dellen, *et al.*, arXiv:2305.04675

R. Utama, *et al.*, *PRC* **93** 014311 (2016)



A. Sharma, *et al.*, *PRC* **105** L031306 (2022)





M. Shelley and A. Pastore, *Universe* **7**, 131 (2021)

Brief overview of the Mixture Density Network (MDN)

Standard neural network

Input → output

$$y = f(x)$$

$$f(\mathbf{x}) = \alpha_1 \mathcal{N}(\mu_1, \sigma_1) + \alpha_2 \mathcal{N}(\mu_2, \sigma_2) + ... + \alpha_n \mathcal{N}(\mu_n, \sigma_n)$$

Neural network learns the Gaussian variables instead of the mapping between x and y directly

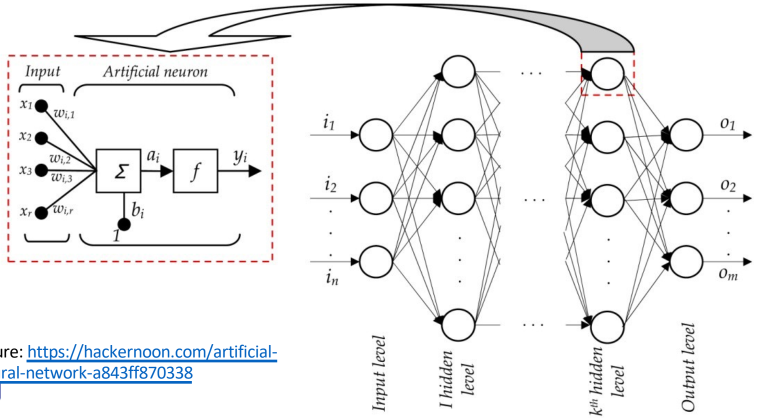
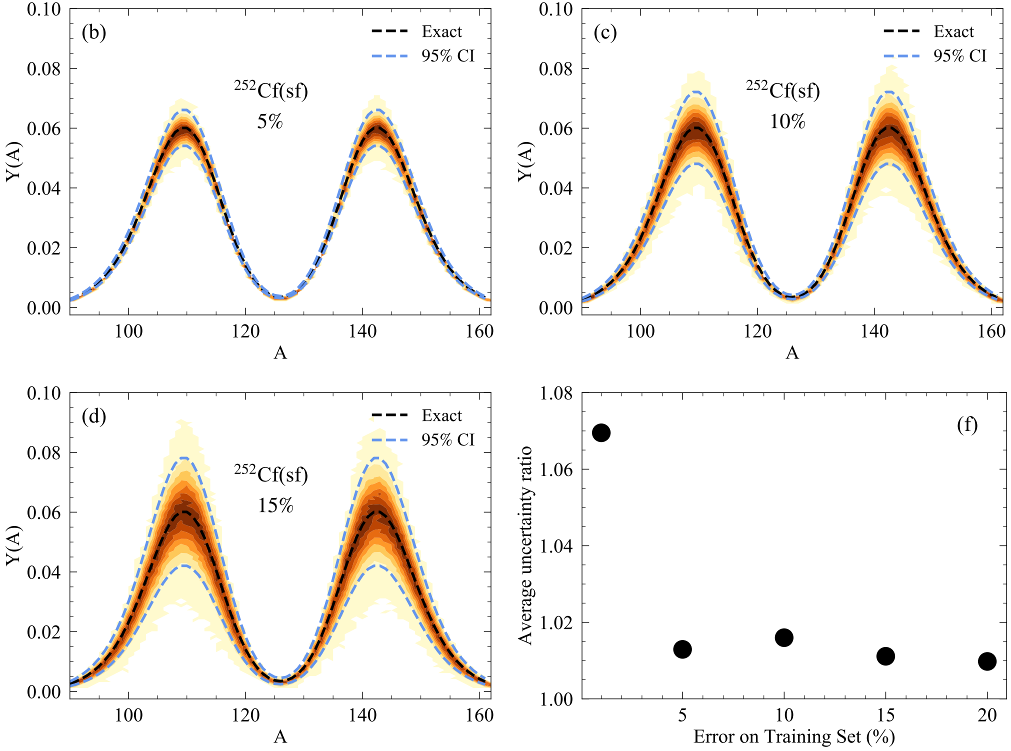


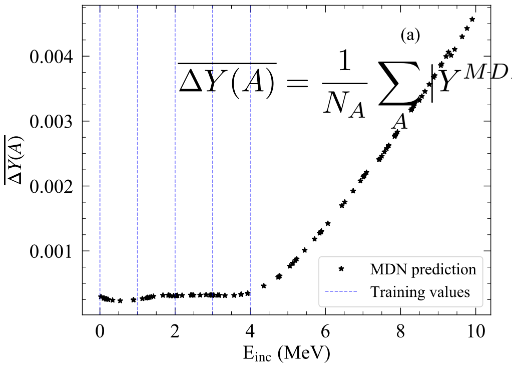
Figure: https://hackernoon.com/artificial-neural-network-a843ff870338

C.M. Bishop, Neural Computing Research Group Report NCRG/94/004 (1994)

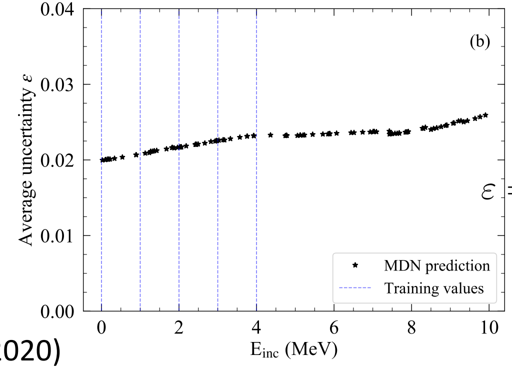
Gaussian mixtures allow uncertainties to be taken into account and included in the predictions. The shape of the posterior is not assumed.

Propagating uncertainties from training to predictions with the MDN, interpolation and extrapolation





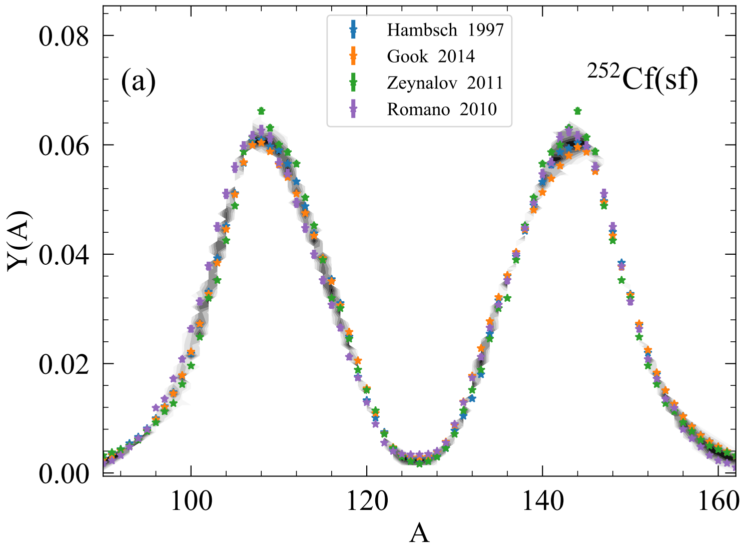
$$\overline{\Delta Y(A)} = \frac{1}{N_A}\sum_A \left|Y^{MDN}(A) - Y^{Exact}(A)\right|$$

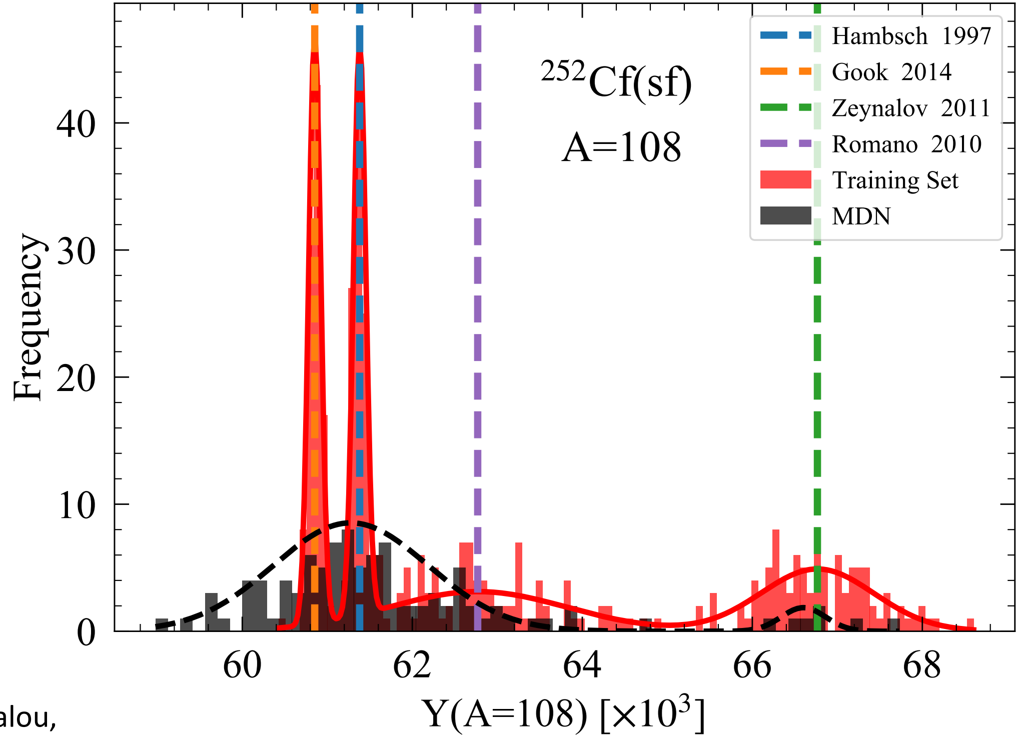


$$\varepsilon = \frac{1}{N}\sum_{i=1}^{N}\frac{\sigma(A_i)}{\overline{Y}(A_i)}$$

A.E. Lovell, A.T. Mohan, and P. Talou, *J. Phys. G* **47**, 114001 (2020)

The MDN provides a potentially complicated predicted posterior distribution





A.E. Lovell, A.T. Mohan, and P. Talou, *J. Phys. G* **47**, 114001 (2020)

Los Alamos NATIONAL LABORATORY

Some numerical details

$$\{\mathbf{x}\} \rightarrow \{\delta M, S_n\}$$

What is **x**?

1 layer with 6 nodes

1 Gaussian mixture per output dimension

25 training points per input vector (takes the uncertainties into account)

Z<20 removed from the training set

$$\{N, Z, A, \qquad \}$$

Some numerical details

$$\{\mathbf{x}\} \to \{\delta M, S_n\}$$

What is **x**?

Take some inspiration from the Liquid Drop model for the binding energies

1 layer with 6 nodes

1 Gaussian mixture per output dimension

25 training points per input vector (takes the uncertainties into account)

Z<20 removed from the training set

$$E_B = a_V A - a_S A^{2/3} - a_C \frac{Z(Z-1)}{A^{1/3}} - a_A \frac{(N-Z)^2}{A} \pm \delta(A,Z)$$

$$\{N, Z, A, \qquad A^{2/3}, Z(Z-1)/A^{1/3}, (N-Z)^2/A, \qquad \}$$

Some numerical details

$$\{\mathbf{x}\} \rightarrow \{\delta M, S_n\}$$

What is **x**?

Also include information about odd-even effects

1 layer with 6 nodes

1 Gaussian mixture per output dimension

25 training points per input vector (takes the uncertainties into account)

Z<20 removed from the training set

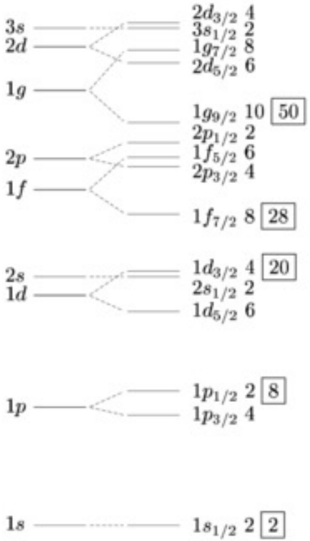
$$\{N, Z, A, \quad A^{2/3}, Z(Z-1)/A^{1/3}, (N-Z)^2/A, \quad N_{EO}, Z_{EO}\}$$

Some numerical details

$$\{\mathbf{x}\} \to \{\delta M, S_n\}$$

What is **x**?

Include information about the magic numbers (shell closures)



1 layer with 6 nodes

1 Gaussian mixture per output dimension

25 training points per input vector (takes the uncertainties into account)

Z<20 removed from the training set

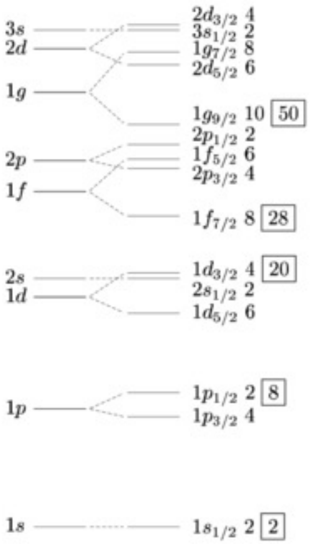
$$\{N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}, (N-Z)^2/A, N_{shell}, Z_{shell}, N_{EO}, Z_{EO}\}$$

Some numerical details

$$\{\mathbf{x}\} \rightarrow \{\delta M, S_n\}$$

What is **x**?

Additional information about the different shells



1 layer with 6 nodes

1 Gaussian mixture per output dimension

25 training points per input vector (takes the uncertainties into account)

Z<20 removed from the training set

$$\{N, Z, A, N_{SM}, Z_{SM}, A^{2/3}, Z(Z-1)/A^{1/3}, (N-Z)^2/A, N_{shell}, Z_{shell}, N_{EO}, Z_{EO}\}$$

We want to assess the impact of including more information into the training set

Model name	Feature space	Output
M2	N, Z	δM
M6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}, (N-Z)^2/A$	δM
M8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	δM
M10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z$	δM
M12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z, N_{\mathrm{shell}}, Z_{\mathrm{shell}}$	δM
MS2	N, Z	$\delta M, S_n$
MS6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}, (N-Z)^2/A$	$\delta M, S_n$
MS8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	$\delta M, S_n$
MS10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z$	$\delta M, S_n$
MS12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z, N_{\mathrm{shell}}, Z_{\mathrm{shell}}$	$\delta M, S_n$

Several training sets are constructed, with increasing physics information in the input

Additionally, we test whether our predictions are better when only mass excess is predicted or both mass excess and the one-neutron separation energy

$$S_n(N,Z) = \Delta_n - \delta M(N,Z) + \delta M(N-1,Z)$$

Model results overview

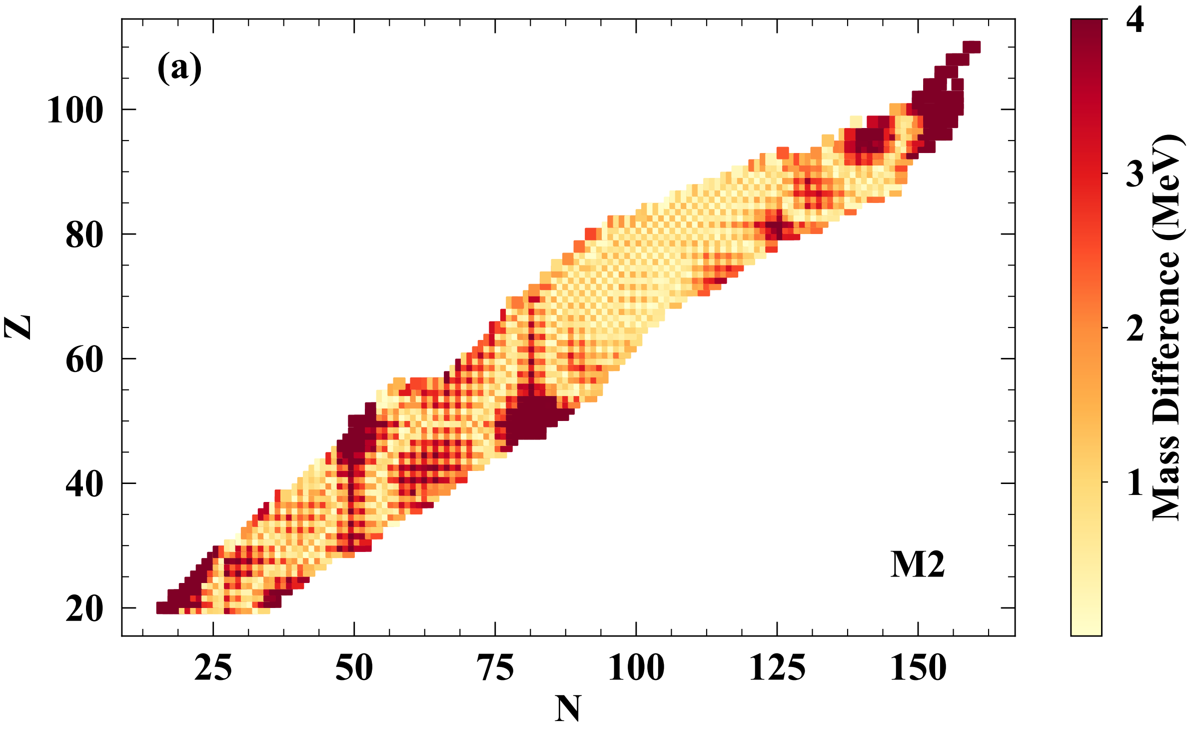
Model	δM σ_{RMS} (MeV)	S_n σ_{RMS} (MeV)
M2	3.90	
MS2	2.43	1.25
M6	1.57	
MS6	2.07	1.21
M8	1.66	
MS8	2.21	0.57
M10	0.58	
MS10	0.76	0.57
M12	0.56	
MS12	0.64	0.47

Generally, the difference between the AME and MDN decreases with the larger input space

$$\sigma_{\mathrm{RMS}} = \sqrt{\frac{\sum_i (\delta M_i^{\mathrm{AME}} - \delta M_i^{\mathrm{MDN}})^2}{K}}$$

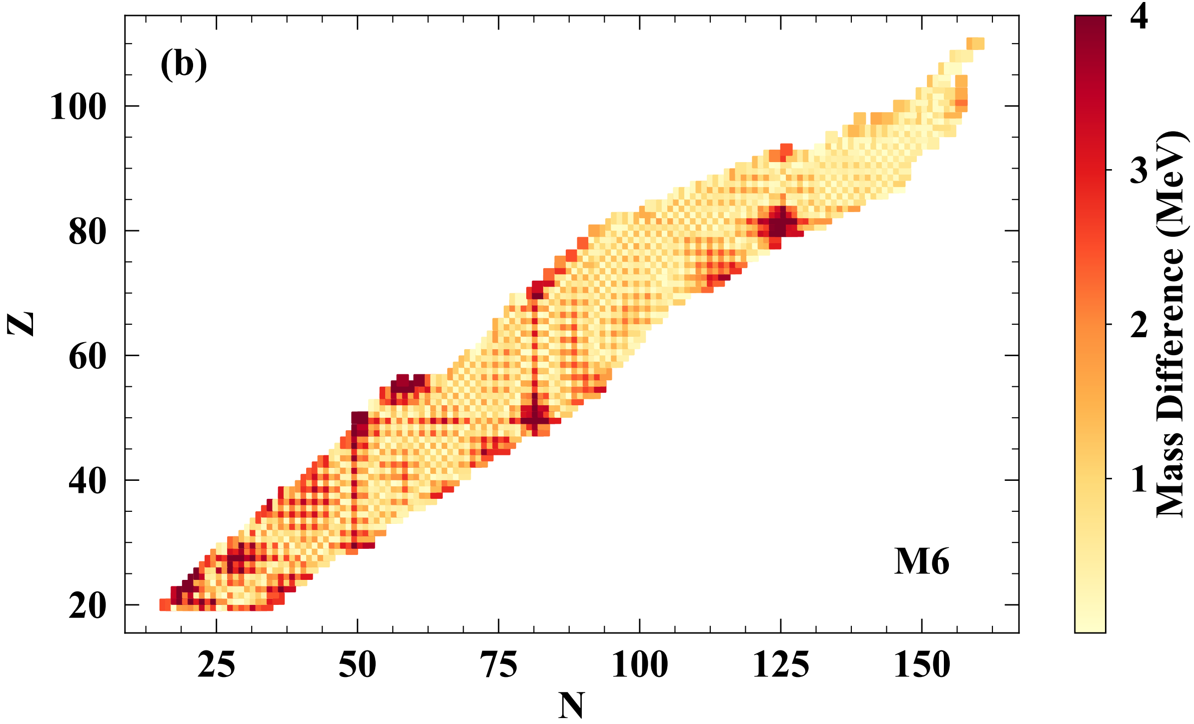
Training on only the mass excess with N and Z as input

Model name	Feature space	Output
M2	N, Z	δM
M6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}$, $(N-Z)^2/A$	δM
M8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	δM
M10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z$	δM
M12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z, N_{\text{shell}}, Z_{\text{shell}}$	δM
MS2	N, Z	$\delta M, S_n$
MS6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}$, $(N-Z)^2/A$	$\delta M, S_n$
MS8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	$\delta M, S_n$
MS10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z$	$\delta M, S_n$
MS12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z, N_{\text{shell}}, Z_{\text{shell}}$	$\delta M, S_n$



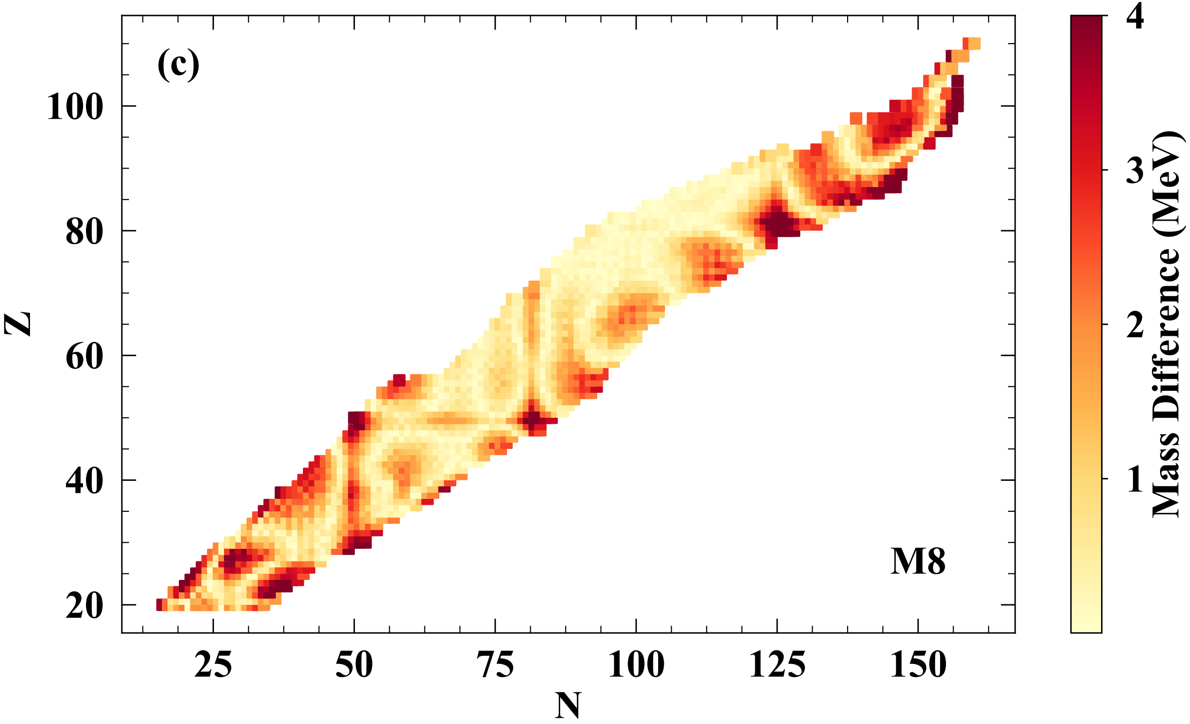
Training on only the mass excess, including macroscopic quantities in the input

Model name	Feature space	Output
M2	N, Z	δM
M6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}, (N-Z)^2/A$	δM
M8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	δM
M10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z$	δM
M12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z, N_{\text{shell}}, Z_{\text{shell}}$	δM
MS2	N, Z	$\delta M, S_n$
MS6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}, (N-Z)^2/A$	$\delta M, S_n$
MS8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	$\delta M, S_n$
MS10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z$	$\delta M, S_n$
MS12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z, N_{\text{shell}}, Z_{\text{shell}}$	$\delta M, S_n$



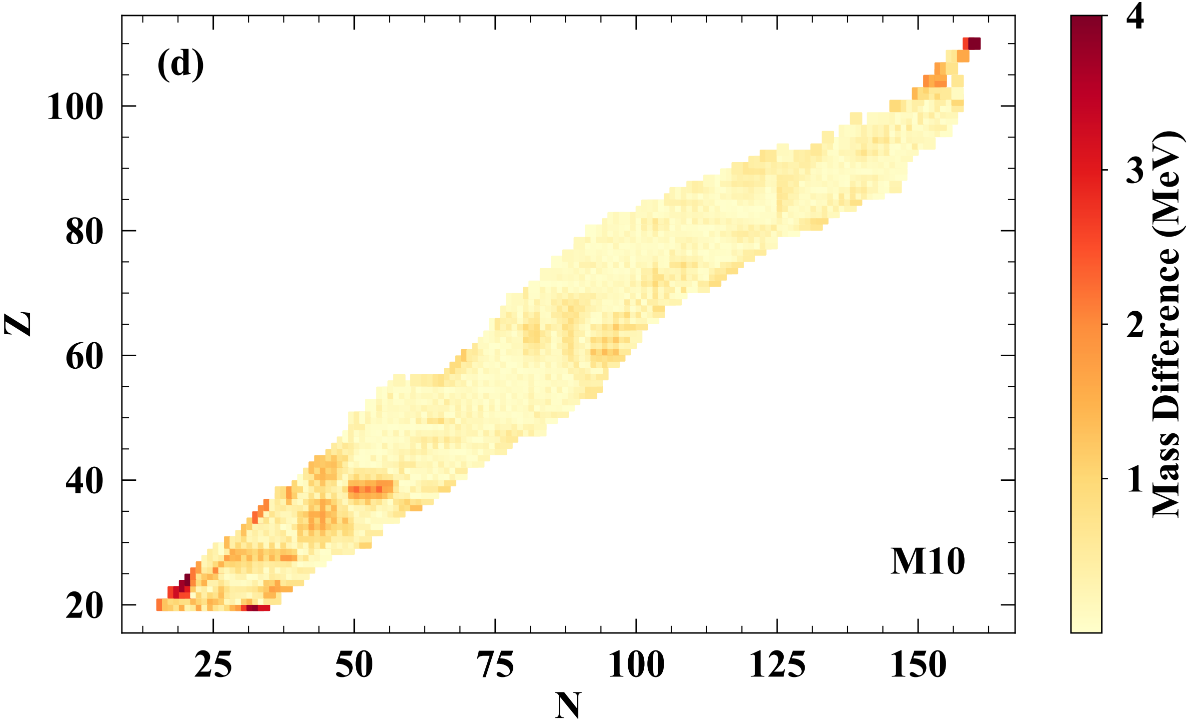
Training on only the mass excess, including information about odd-even staggering

Model name	Feature space	Output
M2	N, Z	δM
M6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}$, $(N-Z)^2/A$	δM
M8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	δM
M10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z$	δM
M12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z, N_{\text{shell}}, Z_{\text{shell}}$	δM
MS2	N, Z	$\delta M, S_n$
MS6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}$, $(N-Z)^2/A$	$\delta M, S_n$
MS8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	$\delta M, S_n$
MS10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z$	$\delta M, S_n$
MS12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z, N_{\text{shell}}, Z_{\text{shell}}$	$\delta M, S_n$



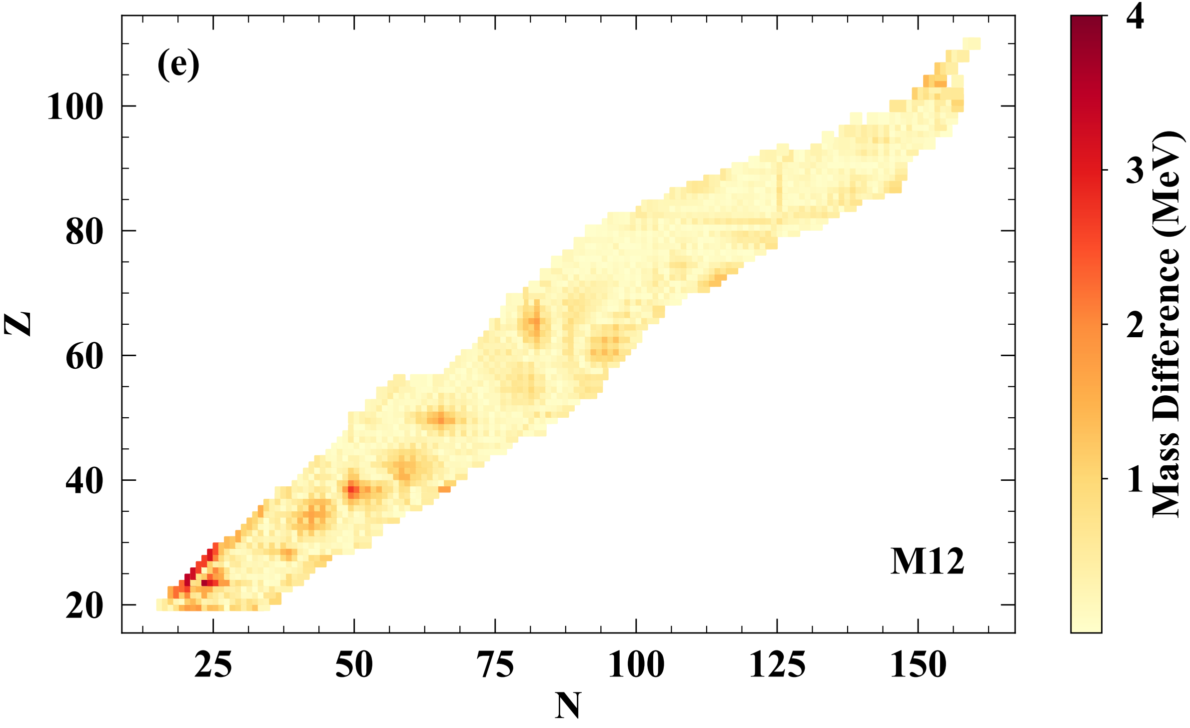
Training on only the mass excess, including information about the distance from shell closures

Model name	Feature space	Output
M2	N, Z	δM
M6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}$, $(N-Z)^2/A$	δM
M8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	δM
M10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z$	δM
M12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z, N_{\rm shell}, Z_{\rm shell}$	δM
MS2	N, Z	$\delta M, S_n$
MS6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}$, $(N-Z)^2/A$	$\delta M, S_n$
MS8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	$\delta M, S_n$
MS10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z$	$\delta M, S_n$
MS12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z, N_{\rm shell}, Z_{\rm shell}$	$\delta M, S_n$



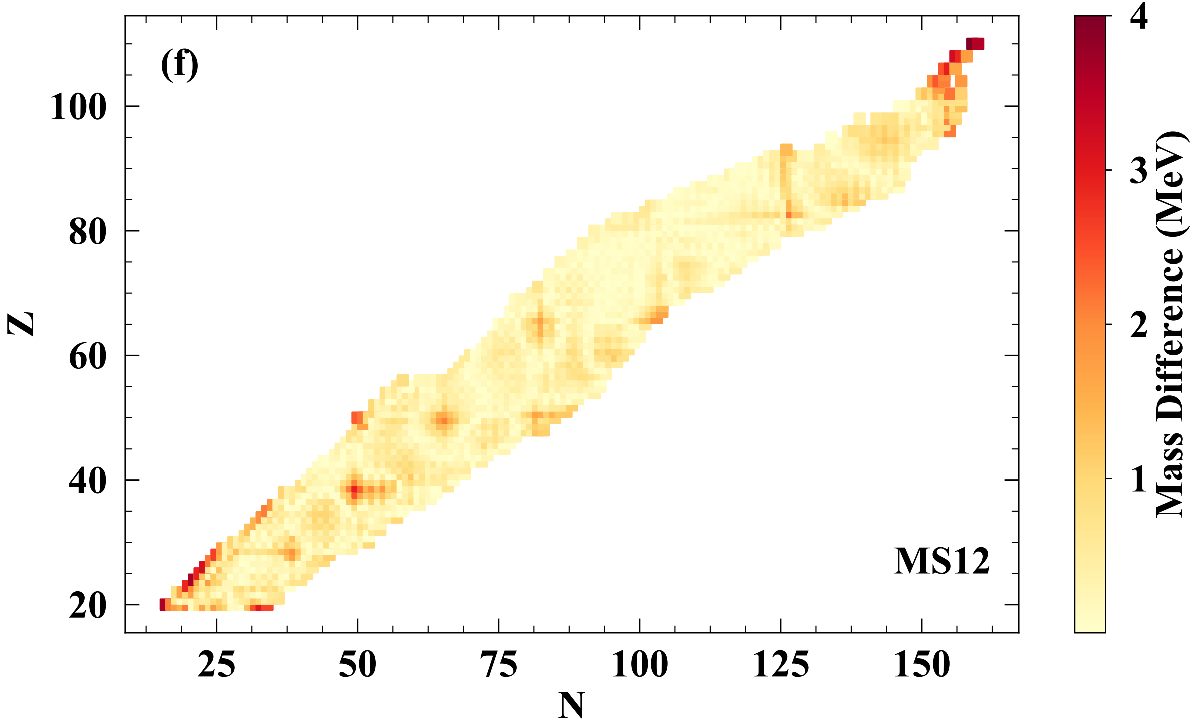
Training on only the mass excess, including information about which shells the nuclei are in

Model name	Feature space	Output
M2	N, Z	δM
M6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}$, $(N-Z)^2/A$	δM
M8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	δM
M10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ ΔN, ΔZ	δM
M12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ ΔN, ΔZ, N_{shell}, Z_{shell}	δM
MS2	N, Z	$\delta M, S_n$
MS6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}$, $(N-Z)^2/A$	$\delta M, S_n$
MS8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	$\delta M, S_n$
MS10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ ΔN, ΔZ	$\delta M, S_n$
MS12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ ΔN, ΔZ, N_{shell}, Z_{shell}	$\delta M, S_n$

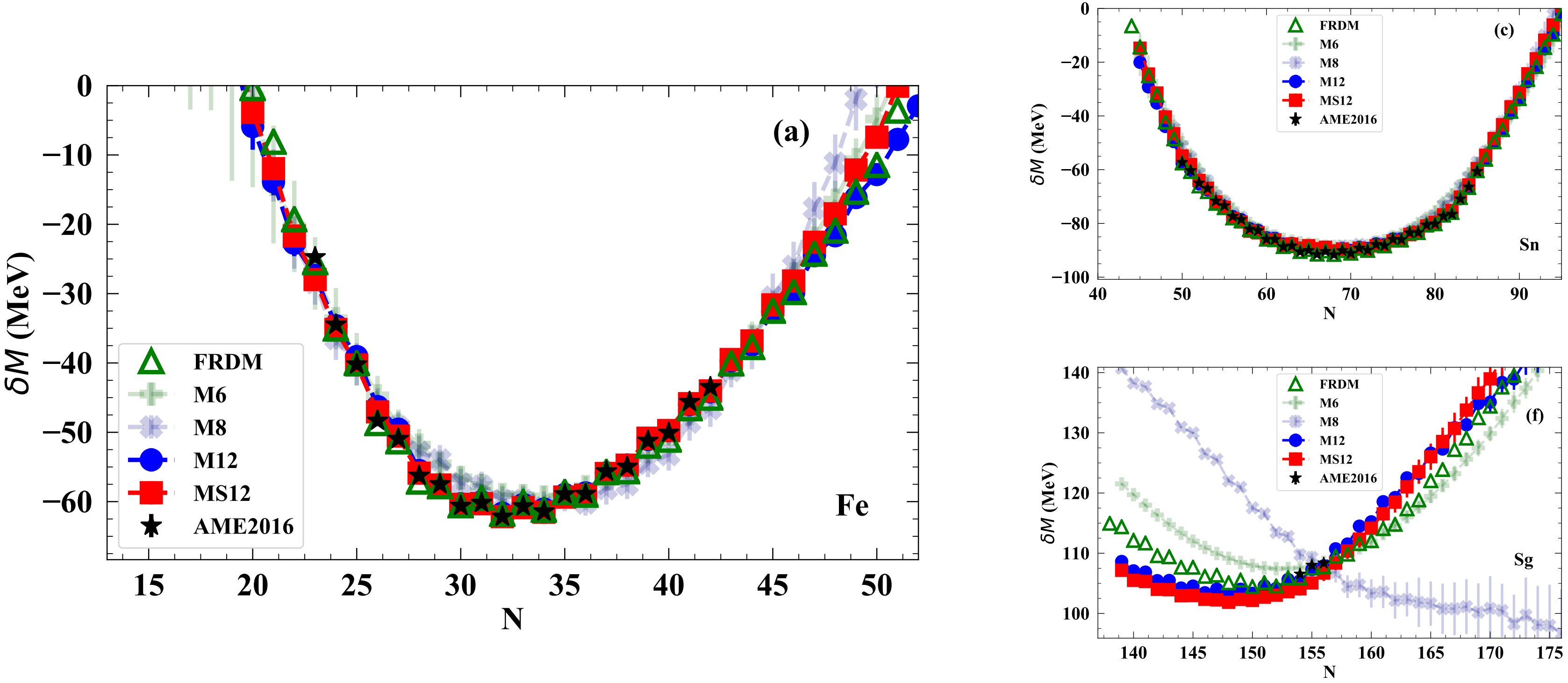


Training on the mass excess and one-neutron separation energy

Model name	Feature space	Output
M2	N, Z	δM
M6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}$, $(N-Z)^2/A$	δM
M8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	δM
M10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z$	δM
M12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z, N_{\text{shell}}, Z_{\text{shell}}$	δM
MS2	N, Z	$\delta M, S_n$
MS6	$N, Z, A, A^{2/3}$, $Z(Z-1)/A^{1/3}$, $(N-Z)^2/A$	$\delta M, S_n$
MS8	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$	$\delta M, S_n$
MS10	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z$	$\delta M, S_n$
MS12	$N, Z, A, A^{2/3}, Z(Z-1)/A^{1/3}$, $(N-Z)^2/A, Z_{EO}, N_{EO}$ $\Delta N, \Delta Z, N_{\text{shell}}, Z_{\text{shell}}$	$\delta M, S_n$

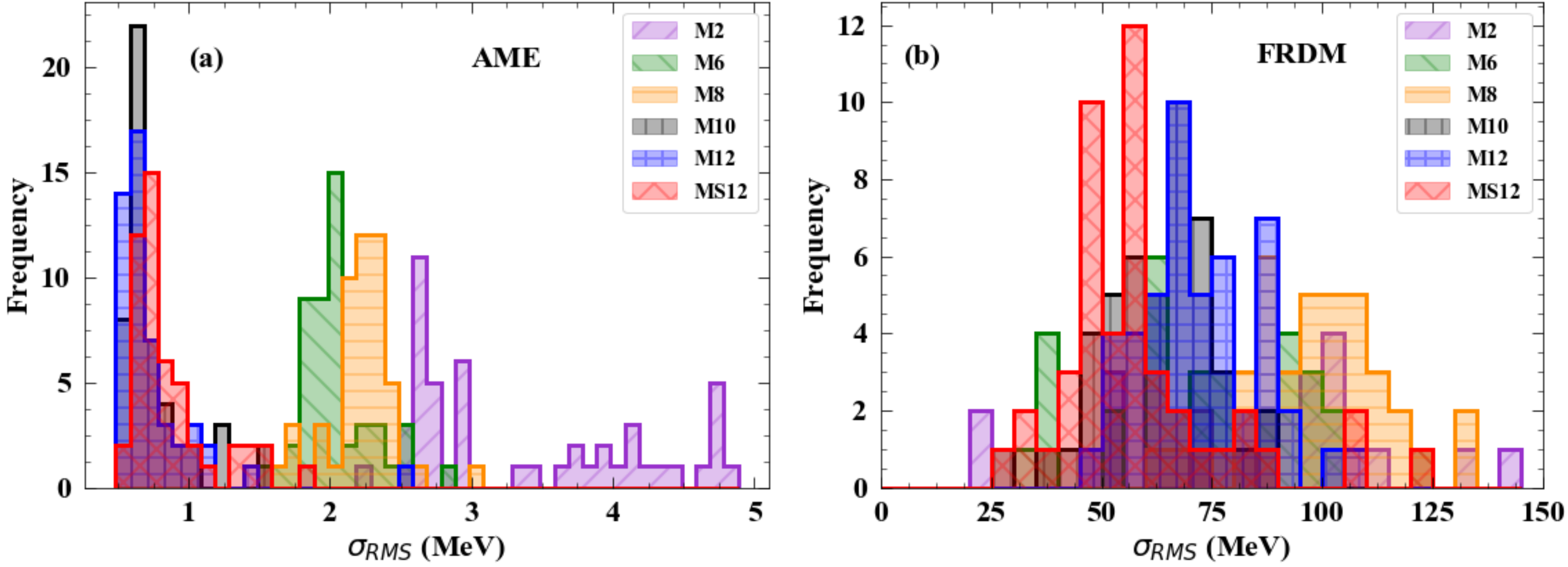


More physics information in the input provides for a more physical extrapolation



Los Alamos
NATIONAL LABORATORY

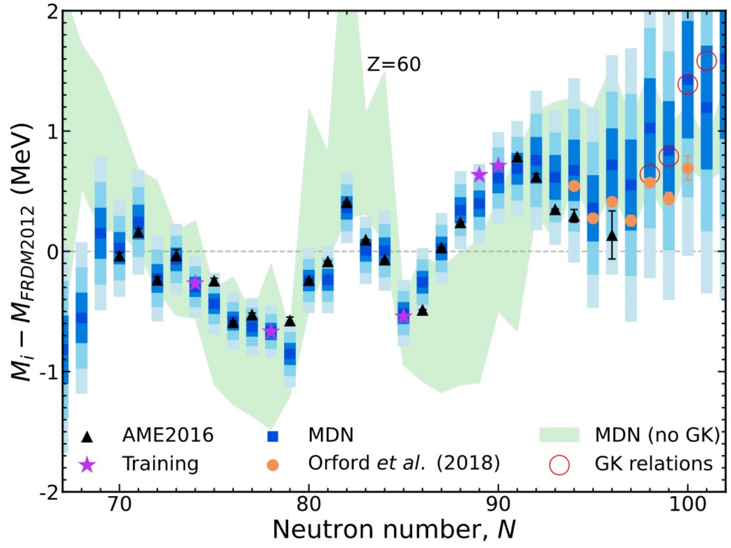
More physics information in the input provides for a more robust extrapolation



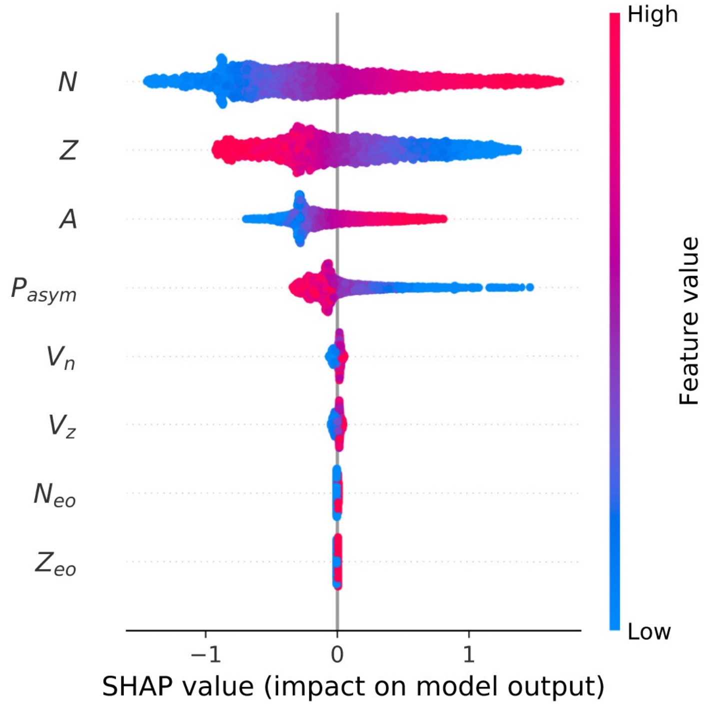
The same network was trained 50 times to understand the spread in the predictions

Next steps

Including more physics constraints and fewer training points



Understanding feature importance (reflects what we know about mass models)



M. R. Mumpower, *et al.*, *PRC* **106**, L021301 (2022)

Summary and conclusions

- Machine learning algorithms that build in well-quantified uncertainties are being used regularly within the nuclear theory community (separate from emulators to quantify e.g. parametric uncertainties)
- We have been investigating a probabilistic algorithm, the Mixture Density Network, to make predictions with uncertainties that reflect the uncertainty on the training data
- Here, we show that including more physics information in the input vector (e.g. nuclear information beyond neutron and proton numbers) can lead to more robust predictions within the training set and extrapolated beyond it
- Similar studies are ongoing to determine what fraction of the training data is needed for robust predictions, to investigate metrics to understand the most important features of the input, and construct covariances between predictions

A brief plug

Applied Nuclear Theory Postdoctoral Researcher

The applied nuclear theory group has seven staff members. We conduct research on a broad set of topics, including developing phenomenological and microscopic models of nuclear reactions, nuclear data evaluations and uncertainty quantification, nuclear many-body theory, and high performance/quantum computing with applications in basic science (nuclear astrophysics and cosmology), nuclear energy (fusion and fission) and nuclear safety and security. The group also actively engages in collaborations with theorists and experimentalists at LANL and worldwide.

Questions? Contact Amy Lovell and Ionel Stetcu at t2-pdsearch-2023@lanl.gov

https://lanl.jobs/search/jobdetails/applied-nuclear-theory-postdoctoral-researcher/42459486-0522-4335-9b37-ad27f570de72 (LANL Jobs, IRC 120807)