



# Taming nuclear mass models with Gaussian processes

Wen-Jia Huang<sup>1,2</sup> · Keisuke Fujii<sup>3</sup> · Hao-Zhao Liang<sup>4,5,6</sup>

Received: 19 May 2025 / Revised: 9 September 2025 / Accepted: 12 September 2025 / Published online: 22 May 2026  
© UT-Battelle, LLC and Wen-Jia Huang, Hao-Zhao Liang 2026

## Abstract

We propose a new set of nuclear mass predictions based on multiple theoretical mass models. By employing Gaussian process regression with the Matérn kernel, we achieved root-mean-square (rms) deviations below 100 keV for the training dataset. The best-performing mass models achieved rms deviations below 150 keV for the new precise mass data from AME2020, whereas the ensemble average showed robust performance across the nuclear chart. Our approach uniquely combines: (1) systematic refinement of eight mass models through their residuals, (2) physics-informed features, including magic numbers, nucleon parity numbers, neutron excess, and nuclear collectivity, and (3) theory-to-theory validation demonstrating robust extrapolation capability. We find that the Matérn kernel provides superior uncertainty quantification compared to the RBF kernel, with a length-scale analysis revealing enhanced inter-nuclei correlations. We provide complete mass predictions for all unknown nuclides in AME2020, offering valuable constraints for nuclear structure studies and astrophysical modeling when used with proper uncertainty propagation.

**Keywords** Gaussian processes · Mass models · Predictive power ·  $r$ -process

## 1 Introduction

The search for the origin of heavy elements in the Universe remains one of the most formidable challenges in scientific inquiry [1, 2]. While it is widely accepted that approximately half of the elements heavier than iron are produced through the rapid neutron-capture process ( $r$ -process), the realistic astrophysical sites responsible for  $r$ -process are still under debate [3–5]. Calculations of  $r$ -process nucleosynthesis involve the properties of thousands of nuclei, spanning

from the line of  $\beta$  stability to the neutron dripline. Among these properties, the nuclear mass is arguably the most crucial. On the one hand, they affect the energy generation and the reaction path; on the other hand, they influence various properties of nuclei that enter the astrophysical calculations, such as  $\beta$ -decay rates, neutron-capture rates, and  $\beta$ -delayed neutron emission probabilities [6]. Sensitivity study underscores that mass uncertainties of 0.5 MeV significantly impact  $r$ -process abundance predictions [7]. A recent study explored how statistical uncertainties in mass models propagate to simulations of  $r$ -process nucleosynthesis [8].

Despite considerable efforts to measure the masses of nuclei near the neutron dripline [9, 10], many of which along the  $r$ -process path remain inaccessible even with the

This work was supported by the Key R&D Program of China (No. 2021YFA1601500), the Talent Support Project of Guangdong Program (No. 2023TQ07A872), and the U.S. D.O.E contract DE-AC05-00OR22725.

✉ Wen-Jia Huang  
huangwj@impcas.ac.cn

✉ Keisuke Fujii  
fujii@ornl.gov

✉ Hao-Zhao Liang  
haozhao.liang@phys.s.u-tokyo.ac.jp

<sup>1</sup> Advanced Energy Science and Technology Guangdong Laboratory, Huizhou 516007, China

<sup>2</sup> Key Laboratory of High Precision Nuclear Spectroscopy and Center for Nuclear Matter Science, Institute of Modern

Physics, Chinese Academy of Sciences, Lanzhou 730000, China

<sup>3</sup> Oak Ridge National Laboratory, Oak Ridge, TN 37831-6169, United States of America

<sup>4</sup> Department of Physics, Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan

<sup>5</sup> Quark Nuclear Science Institute, The University of Tokyo, Tokyo 113-0033, Japan

<sup>6</sup> Interdisciplinary Theoretical and Mathematical Sciences Program, RIKEN, Wako 351-0198, Japan

most advanced radioactive ion beam facilities [11]. Theoretical mass models, although originating from different approaches, are thus the last resort for obtaining mass information on unknown nuclei [12–14], forming the basis for nucleosynthesis calculations. Although most mass models fit well with experimental mass data, they diverge dramatically when extrapolating toward unknown regions [12]. As pointed out in Ref. [15], the accuracy and predictive power of mass models do not always correlate well. Furthermore, the uncertainty in mass models is difficult to quantify, hindering a comprehensive evaluation of their performance in extrapolation. Quantifying uncertainties to assess the predictive capability of mass models is highly demanding [16, 17]. In the 1990 s, it was demonstrated that neural networks could be employed to predict the masses and other properties of nuclei [18, 19]. Along this line, neural networks have started to play an important role in predicting nuclear masses [20–25]. Currently, different machine learning approaches equipped with Bayesian methods have gained much attention and are widely used in the nuclear physics community (see Refs. [26–28] and references therein). Other machine learning algorithms, such as kernel ridge regression, have also begun to play an important role in the description of nuclear masses and separation energies [29, 30].

Bayesian-based methods can largely avoid overfitting by incorporating priors over parameters [31]. Bayesian analysis presents its results as a posterior probability distribution over parameters and variables, allowing the mean and variance to be calculated using statistical approaches. Bayesian Neural Networks (BNNs) have been successfully applied to refine model predictions of nuclear masses [32–35],  $\alpha$ -decay energies [36], and  $\beta$ -decay half-lives [37]. However, BNNs are often not easy to use in practice because they require many decisions regarding activation functions and learning rates. Moreover, they involve hundreds of parameters that are difficult to interpret. An alternative is to use Gaussian Processes (GPs) [38], which have only a few parameters. GPs have been used to improve model predictions of separation energies [39], nuclear masses [40], and to estimate the limits of the nuclear landscape [41]. Beyond mass models, Bayesian Monte Carlo frameworks have been employed for nuclear data evaluation [42].

Recently, GPs have been extensively employed to refine nuclear mass predictions [43–45]. While these studies demonstrated the power of GPs for mass modeling, our work extends beyond single-model refinement by systematically improving eight commonly used theoretical mass models through their residuals. This ensemble approach combines the strengths of each model while compensating for their individual limitations, which is a crucial advancement for astrophysical applications where extrapolation reliability is essential. To rigorously test the predictive power of our GP,

we introduce a novel theory-to-theory validation framework using FRDM12 as “pseudo-experimental” data. This allows us to evaluate the extrapolation performance far beyond the known regions. Furthermore, we provide complete mass predictions for all unmeasured nuclei listed in AME2020 for individual GP-refined mass models, which can serve as a comprehensive source for both astrophysical modeling and future experimental research.

This paper is organized as follows: In Sect. 2, we discuss the current mass models that are widely used in nuclear physics and their comparisons with experimental data. In Sect. 3, we outline the concept of GP regression and its implementation in mass models. In Sect. 4, we illustrate the performance of the GP models. In Sect. 5, we fine-tune our GP model and test its predictive capability. Finally, we conclude our work in Sect. 6 and offer future perspectives.

## 2 Nuclear mass model

Current global mass models can be categorized into three classes [12]: macroscopic–microscopic, microscopic, and phenomenological. The first nuclear mass model, the Bethe–Weizsäcker formula, treats the nucleus as a charged liquid drop [46, 47]. While it can predict the binding energies for most nuclei, it fails to describe those near the magic numbers. Based on the images of liquid drops, macroscopic–microscopic mass models incorporate shell corrections and pairing effects. Two widely used models in this category are the Finite-Range Droplet Model (FRDM) [48] and the Weizsäcker–Skyrme (WS) mass model [49]. The microscopic mass model is primarily based on the mean-field theory using energy density functionals. For example, self-consistent Hartree–Fock–Bogoliubov (HFB) calculations with Skyrme and Gogny interactions [50] provide the most accurate results, achieving a root-mean-square (rms) deviation of 0.5 MeV [51]. A more complex microscopic model, which further improves the description of fission barriers and large deformations, provides global descriptions of both ground-state and excited-state properties of nuclei, albeit with a modest accuracy of 1–4 MeV [52]. Phenomenological mass models, on the other hand, use a set of parameters fitted to experimental mass data, achieving an rms deviation of approximately 0.4 MeV [53]. Improving the accuracy of these mass models is crucial for meeting the requirements of astrophysical calculations.

Figure 1 displays the residuals between different mass models and the experimental data in the Atomic Mass Evaluation AME2016 [54]: the macroscopic–microscopic mass models FRDM95 [55], FRDM12 [48], and WS3 [56], microscopic mass models HFB27 [51], UNEDF1 [52], and RMF [57], and the phenomenological models DZ28 [53] and KTUY05 [58]. Most models generally described the

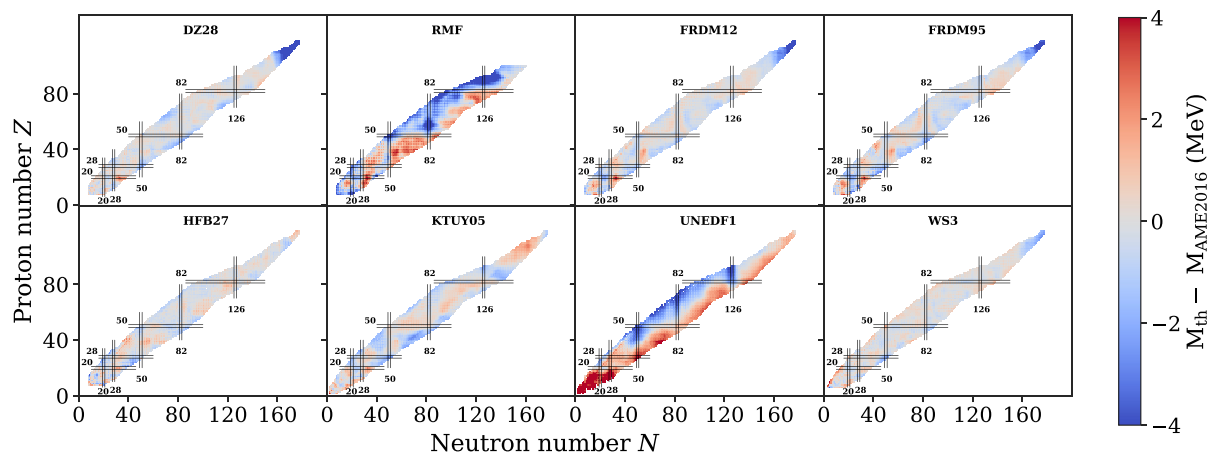


**Fig. 1** (Color online) Residuals (in MeV) between theoretical mass models and experimental data from AME2016

mass data well, particularly DZ28, WS3, and HFB27. However, the microscopic mass models RMF and UNEDF1 exhibit prominent systematic patterns in the nuclear chart. For UNEDF1, there appears to be a separation where the mass-excess values of proton-rich nuclei are underestimated, whereas those of neutron-rich nuclei are overestimated. This suggests that some important physics may be missing or not considered in their effective interactions. The models DZ28, FRDM95, FRDM12, and WS3 appear to underpredict the mass values for nuclei with  $Z > 100$ . The predictive power of all mass models diminishes as they approach the driplines.

The residuals between the mass models and experimental data not only quantify the model accuracy but also provide valuable insights into the model deficiencies. A detailed study of these residuals can help improve the mass models. GPs are a powerful tool for capturing systematic patterns and underlying structures in the data. By carefully designing the framework for GPs, the model accuracy can be significantly improved.

### 3 Gaussian-process regression

The Gaussian-Process regression is a non-parametric Bayesian technique that provides exceptional flexibility in modeling complex functions of arbitrary forms [38]. By employing a kernel function, it can effectively capture the underlying trends and quantify uncertainties, establishing correlations between any two points in the input space. In this work, we model the residuals  $\delta = M_{\text{th}} - M_{\text{exp}}$  between theoretical mass predictions ( $M_{\text{th}}$ ) and experimental data ( $M_{\text{exp}}$ ) using GP regression. The GP-refined predictions of mass models can be easily recovered by using  $M_{\text{th}}^{\text{GP}} = M_{\text{th}} - \delta^{\text{GP}}$ . The mathematical derivations of Gaussian-Process regression are detailed in the Supplementary

Material (SM) [59]. Our GP model employs two kernel functions to describe spatial dependencies between nuclei: the Matérn (MAT32) and the RBF kernel (see Eqs. 4–5 in SM for definitions [59]). The input space includes proton and neutron numbers ( $Z, N$ ), the parity number of proton  $\delta_Z = (-1)^Z$  and neutron  $\delta_N = (-1)^N$ , the distance to the nearest proton magic number  $v_Z = Z - Z_0$  and neutron magic number  $v_N = N - N_0$ , where  $Z_0 = 2, 8, 20, 28, 50, 82$  and  $N_0 = 2, 8, 20, 28, 50, 82, 126$ , the neutron-excess number  $NE = N - Z$ , and the nuclear collectivity indicator  $C = \frac{v_Z v_N}{v_Z + v_N}$  [60]. The GP model with the normal two-dimensional input space ( $Z, N$ ) serves as a baseline in our work and is denoted as GP-2D, and the model with the full input space ( $Z, N, \delta_Z, \delta_N, v_Z, v_N, NE, C$ ) is denoted as GP-8D.

The training dataset consists of 2272 nuclei from AME2016 with uncertainties  $\sigma^{\text{exp}} < 100$  keV and  $Z, N \geq 8$ . To validate the extrapolation of our GP model, we tested the model on new masses from AME2020 [61] that were not included in AME2016. This non-random partitioning ensured a rigorous evaluation of the predictive capability. Our framework was implemented using the *GPy* software developed by the Sheffield Machine Learning Group [62] and applied to refine the residuals of the mass models.

### 4 Results and validation

In this study, the mean and variance of our estimator  $\delta^{\text{GP}}$  are calculated analytically from the posterior distribution under the assumption of a Gaussian-distributed likelihood function. The approach yields results that are approximately equal to those obtained via MCMC sampling but with significantly reduced computation costs [39].

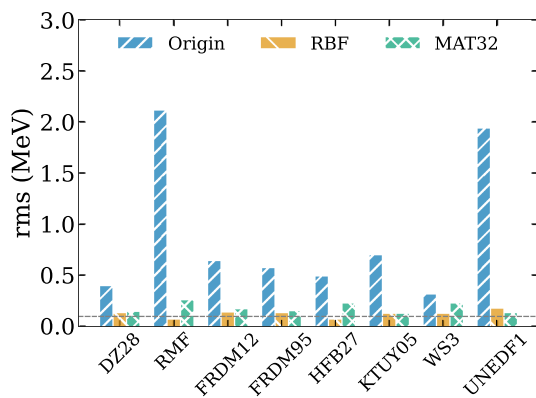


**Fig. 2** (Color online) Root-mean-square (rms) deviations (in MeV) of mass models before and after Gaussian-Process refinement using  $(Z, N)$  input (GP-2D). The results for both the RBF and MAT32 kernels are shown, with a gray dashed line indicating 100 keV accuracy

Figure 2 displays the rms deviations of the mass models for the training dataset. The GP-2D regression with only the two-dimensional input space  $(Z, N)$  markedly enhanced the accuracy of all mass models.  $\sigma^{\text{GP}}$  is consistently lower for the RBF kernel than that for the MAT32 kernel in most cases, indicating that it might produce more accurate mass predictions. The greatest enhancements were observed in the two microscopic mass models, RMF and UNEDF1, where the rms deviations dropped from the original values of approximately 2 MeV to approximately 0.1 MeV. For the other models, the accuracy improved by factors of 2–6 compared to their original values.

In Bayesian analysis, the credibility interval (CI) is the interval within which an estimate is expected to fall with a particular probability. As pointed out in Ref. [39], the rms deviation is not the sole quantity used to address the performance of a mass model. Therefore, we evaluate two empirical coverage probabilities for CIs for our estimator  $\delta^{\text{GP}}$ , which correspond to 68% (1CI) and 95% (2CI) posterior probabilities of lying for the MAT32 and RBF kernels used in this study. The results are summarized in Table 1.

The credibility intervals for both the MAT32 and RBF kernels fell below their nominal values, suggesting a potential underestimation of the uncertainties in these GP regressions. Ideally, we expect a model whose credibility interval matches the expected values, although this is not always the case in practice. The 1CI and 2CI empirical coverage probabilities show that the MAT32 kernel often performs closer to the expected 68% and 95% ranges, which may indicate better uncertainty quantification in some instances. Furthermore, the RBF kernel might oversmooth the mass surface, while the Matérn kernel can capture sharp jumps in the residuals and avoid the “concentration of measure” problem in high-dimensional spaces [63]. Regarding error estimation, the GP with the MAT32 kernel performed better than the RBF

**Table 1** Root-mean-square (rms) deviations (in MeV) between eight mass models and the experimental masses in AME2016 before ( $\sigma^{\text{org}}$ ) and after GP refinement ( $\sigma^{\text{GP-2D}}$ ) in  $(Z, N)$  input space. The last two columns list the empirical coverage probabilities for 68% (1CI) and 95% (2CI) credibility intervals. The numbers before the parentheses are the results from the MAT32 kernel, and the numbers inside the parentheses are from the RBF kernel

| Model  | $\sigma^{\text{org}}$ | $\sigma^{\text{GP-2D}}$ | 1CI (%)    | 2CI (%)    |
|--------|-----------------------|-------------------------|------------|------------|
| DZ28   | 0.400                 | 0.150(0.135)            | 50.0(48.8) | 81.9(79.6) |
| RMF    | 2.112                 | 0.261(0.074)            | 60.3(91.2) | 85.6(98.5) |
| FRDM95 | 0.645                 | 0.177(0.147)            | 59.3(58.1) | 87.0(85.9) |
| FRDM12 | 0.576                 | 0.159(0.139)            | 59.1(55.2) | 85.4(85.0) |
| HFB27  | 0.494                 | 0.231(0.077)            | 50.2(90.9) | 80.7(99.0) |
| KTUY05 | 0.700                 | 0.133(0.132)            | 52.4(44.2) | 84.4(77.4) |
| WS3    | 0.319                 | 0.229(0.133)            | 48.6(47.1) | 76.1(79.2) |
| UNEDF1 | 1.940                 | 0.136(0.182)            | 52.6(40.7) | 84.2(67.1) |

kernel. Given its more conservative uncertainty estimates and better empirical coverage, we selected the MAT32 kernel for subsequent GP model refinement.

Figure 3 illustrates the performance of our MAT32 kernel regression for the Sn isotopic chain. After GP-2D refinement, all mass models, regardless of their theoretical foundations, showed excellent agreement with the experimental data. For example, the two bulges from the HFB27 model are flattened, and the downsloping trends of the RMF and UNEDF1 predictions toward the neutron-deficient side of Sn are significantly corrected. This demonstrates that the GP can capture the systematic patterns inherent in each model’s residuals, ultimately improving the model’s accuracy (see Fig. 1 in SM for a full residual map [59]). More importantly, the inclusion of statistical uncertainties with each prediction provides valuable insight into the model’s reliability, which is particularly crucial for estimating the masses of exotic nuclei far from stability, which are key to astrophysical studies.

## 5 Fine-tuning and predictive power

While the GP-2D model has already demonstrated significant improvements, further enhancement of its predictive capability requires the incorporation of additional nuclear structure features. Thus, we conducted GP regression by integrating the expanded set of physics-informed inputs, as mentioned above. As shown in Table 2, the full 8D model, GP-8D, outperforms the baseline model GP-2D. The rms deviations in GP-8D for RMF, FRDM95, HFB27, and WS3 decreased by 72.8%, 50.3%, 64.1%, and 63.3%, respectively. These enhancements demonstrate that our GP framework further captures subtle nuclear interactions that are often missed by conventional mass models. The

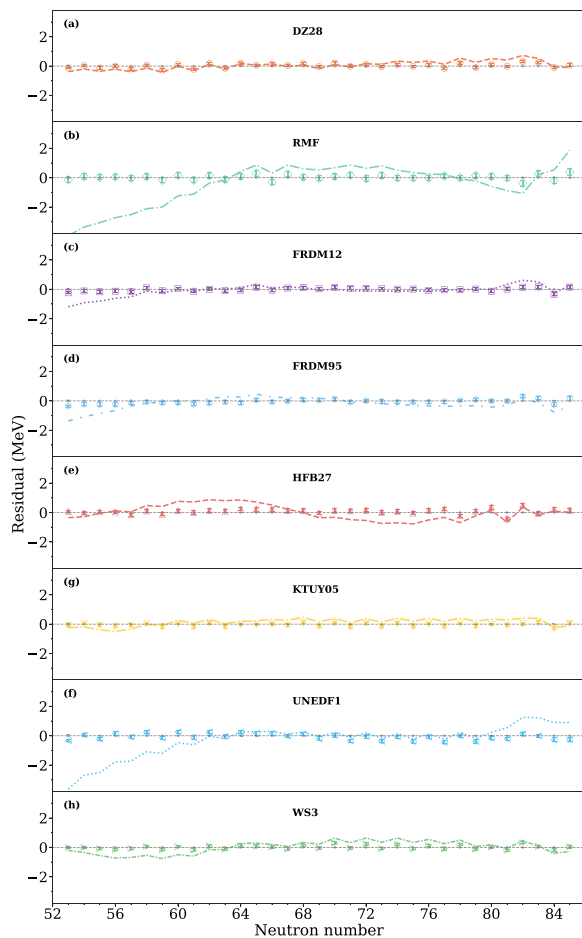


**Fig. 3** (Color online) Residuals for the Sn isotopic chain before (lines) and after (symbols with error bars) the application of the MAT32 kernel in the  $(Z, N)$  input space (GP-2D) in the training dataset

accuracy achieved in this study is comparable to the mass predictions using the 12 Garvey-Kelson relations [64] and the mass relations of mirror nuclei [65]. One of us (H.Z.L.) used BNN to learn the mass surface of even-even nuclei

**Table 2** Comparisons of GP-2D and GP-8D models.  $\sigma$  denotes the rms deviations (in MeV) and  $\Delta$  indicates the rms deviation reductions from  $\sigma^{\text{GP-2D}}$  to  $\sigma^{\text{GP-8D}}$  in percentage.  $l_p$  and  $l_n$  represent the length scales for proton and neutron numbers, respectively

| Model  | $\sigma^{\text{GP-2D}}$ | $\sigma^{\text{GP-8D}}$ | $\Delta(\%)$ | $(l_p, l_n)^{\text{GP-2D}}$ | $(l_p, l_n)^{\text{GP-8D}}$ |
|--------|-------------------------|-------------------------|--------------|-----------------------------|-----------------------------|
| DZ28   | 0.150                   | 0.082                   | 45.3         | (4.2, 5.8)                  | (25.9, 28.3)                |
| RMF    | 0.261                   | 0.074                   | 72.8         | (3.1, 9.0)                  | (48, 89)                    |
| FRDM95 | 0.177                   | 0.088                   | 50.3         | (4.1, 5.2)                  | (21.1, 14.8)                |
| FRDM12 | 0.159                   | 0.087                   | 45.2         | (3.9, 5.3)                  | (28.8, 24.1)                |
| HFB27  | 0.231                   | 0.083                   | 64.1         | (2.5, 3.9)                  | (10.9, 9.2)                 |
| KTUY05 | 0.133                   | 0.081                   | 39.1         | (4.0, 7.7)                  | (20.0, 25.5)                |
| WS3    | 0.229                   | 0.084                   | 63.3         | (13.8, 13.5)                | (136.8, 123.0)              |
| UNEDF1 | 0.136                   | 0.073                   | 46.2         | (1.8, 3.9)                  | (10.0, 13.1)                |

and the correlation energies, reaching a similar accuracy of 84 keV [35]. The predictions for the masses of the Sn isotopic chain in the full input space are displayed in Fig. 2 in SM [59].

The exceptional accuracy of our refined GP models highlights the importance of examining the length-scale parameters in the GP framework (the definition of length scales is seen in Eqs. 4–5 in SM [59]). These parameters govern the decay of correlations between nuclei across the  $(Z, N)$  plane, directly affecting the model's ability to extrapolate to regions far from the experimental data. As shown in Table 2 (last two columns), the length scales for  $Z$  and  $N$  exhibited significant changes when transitioning from GP-2D to GP-8D. For models that utilize only  $(Z, N)$ , these parameters typically fall within the range of 2 to 14. However, upon the implementation of more physical inputs, these length-scale values increase significantly. For example, the length scales for  $Z$  and  $N$  for FRDM12 increase from 4 and 5 to 29 and 24, respectively. In GP regression, larger length scales indicate a slower decay in the correlations between data points as their spatial separation increases. This behavior aligns with the nature of the GP model, reflecting stronger interrelations among the expanded set of input features. We favor larger length-scale values in the GP model because of their robustness against outliers. A shorter length scale could overly rely on neighboring nuclei for predictions, potentially compromising accuracy if the last few nuclei are measured inaccurately. Although it is not the scope of this study to discuss the impact of the length scale in detail, its critical role in model performance merits further investigation.

To evaluate the predictive power of our refined GP models, we tested them against the 74 new masses adopted in the AME2020 (see Fig. 4a). The residuals of each mass model before and after the GP-8D refinements are compared in Fig. 4(c–j). The original models exhibited large deviations, whereas the GP-8D refinement substantially reduced the residuals across all cases. Specifically, the refined models demonstrated a closer alignment with the experimental data, as evidenced by the narrowing of the residual distributions and reduction in the rms deviations. If we constrain the test set for nuclei with uncertainties smaller than 100 keV, the rms deviations are further reduced. For example, the rms deviations for DZ28 and WS3 were reduced to 149 keV, crossing the 150 keV level. It should be noted that the obtained 150 keV rms deviation represents only short-range (local) extrapolation capability, which cannot be extrapolated to the  $r$ -process region without additional validation. We quantify the ensemble performance of our GP models by calculating the mean and standard deviation across all GP-refined

models, which are defined as  $\overline{M}^{\text{GP}} = \sum_{i=1}^m M_{\text{th},i}^{\text{GP}}/m$  and

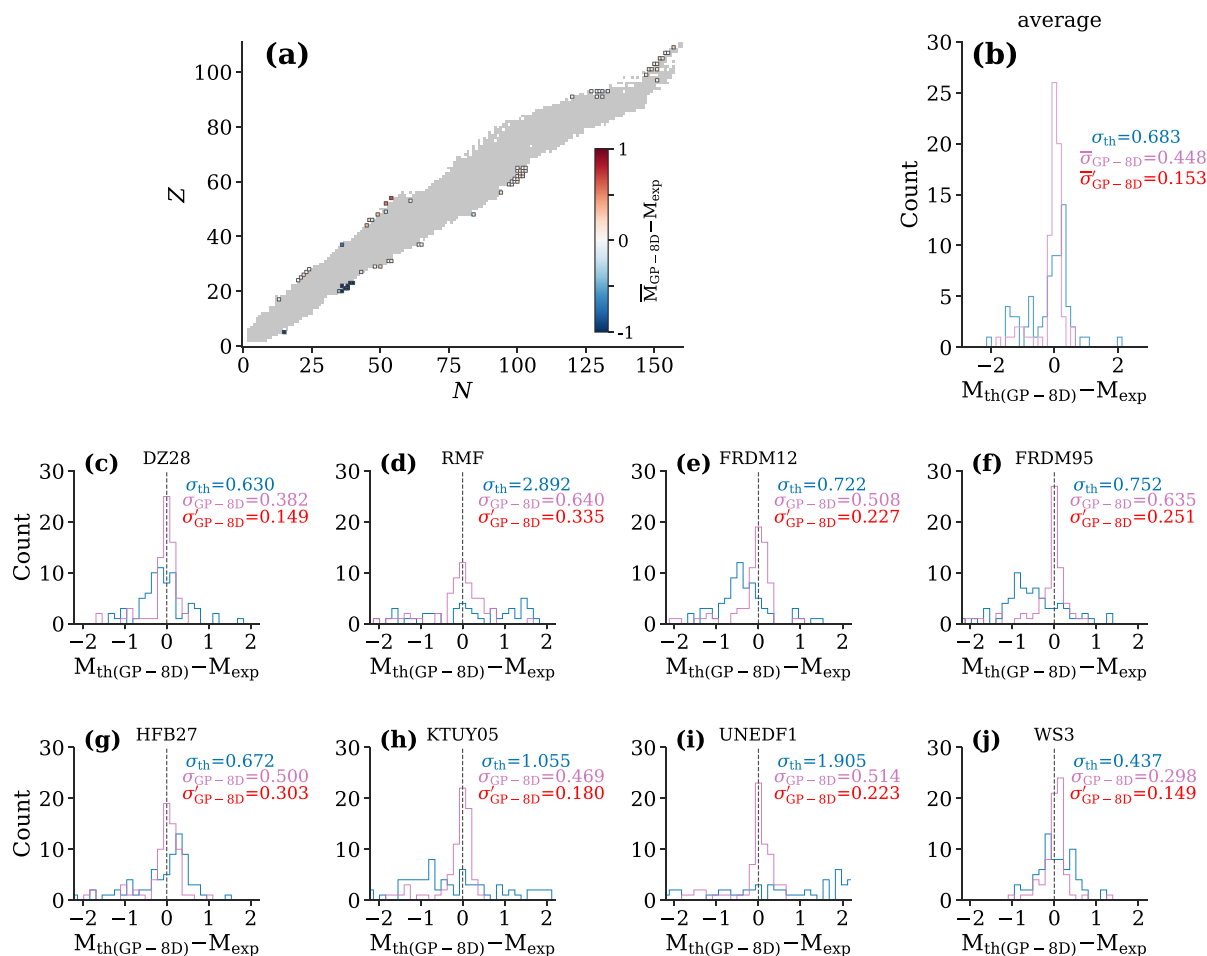


**Fig. 4** (Color online) (a) Nuclear chart showing the new masses from AME2020 (colored squares) and AME2016 training data (gray). The color shows the residuals between the simple average of the GP-8D estimates and the new mass values (see text for details). (b) Histogram of residuals between the model average and new masses in

AME2020. (c–j) Histograms of the residuals for each mass model. The numbers in the histograms show the calculations of the standard deviations of the original mass models ( $\sigma_{th}$ ), GP-8D refined models ( $\sigma_{GP}$ ), and GP-8D refined models compared to masses with uncertainties smaller than 0.1 MeV ( $\sigma'_{GP-8D}$ )

$\sigma'_{GP} = \sqrt{\frac{\sum_{j=1}^n (M_j - \bar{M}_{GP})^2}{n-1}}$  ( $m = 8$  and  $n = 74$ ), respectively. As shown in Fig. 4b, the mean predictions match the accuracy of the best-performing models (DZ28 and WS3).

To rigorously assess the extrapolation capability of our GP ensemble approach, we performed a theory-to-theory validation using FRDM12 as “pseudo-experimental” data (see SM for details [59]). This validation revealed three characteristic regimes in the trends of the rms deviation as a function of the extrapolation distance  $r$  (cf. Figure 3 in SM [59]). In the short-range regime ( $r < 5$ ), the rms deviations of the averaged GP models increase with  $r$ . This behavior reflects the expected uncertainty growth in the GP predictions. In the intermediate regime ( $5 \leq r \leq 13$ ), the rms deviation first decreases and then reaches a plateau at approximately 200 keV. This behavior indicates that the GP

can better capture systematic trends even far from stability. For long-range extrapolation ( $r > 13$ ), the rms deviation starts to increase again and appears to reach a plateau at approximately 300 keV. The theory-to-theory validation provides quantitative confidence intervals for GP predictions in the unknown regions of the nuclear chart.

Figure 5 compares the original and average GP-refined mass models for the Sc and Pr isotopic chains. The new mass values for  $^{54-60}\text{Sc}$  were taken from Refs. [66–69] and for Pr taken from Ref. [70]. Our GP estimates show good agreement with experimental data and the new mass data (within  $2\sigma$ ). The emergence of systematic deviations for neutron-rich  $^{56-60}\text{Sc}$  should be interpreted with caution. These masses come from  $B\rho$ -ToF measurements [68, 69] and have larger uncertainties ( $>100$  keV). In AME2020, five of the 15 results in Ref. [69] were excluded. Specifically, AME2020 suggests  $^{60}\text{Sc}$  should

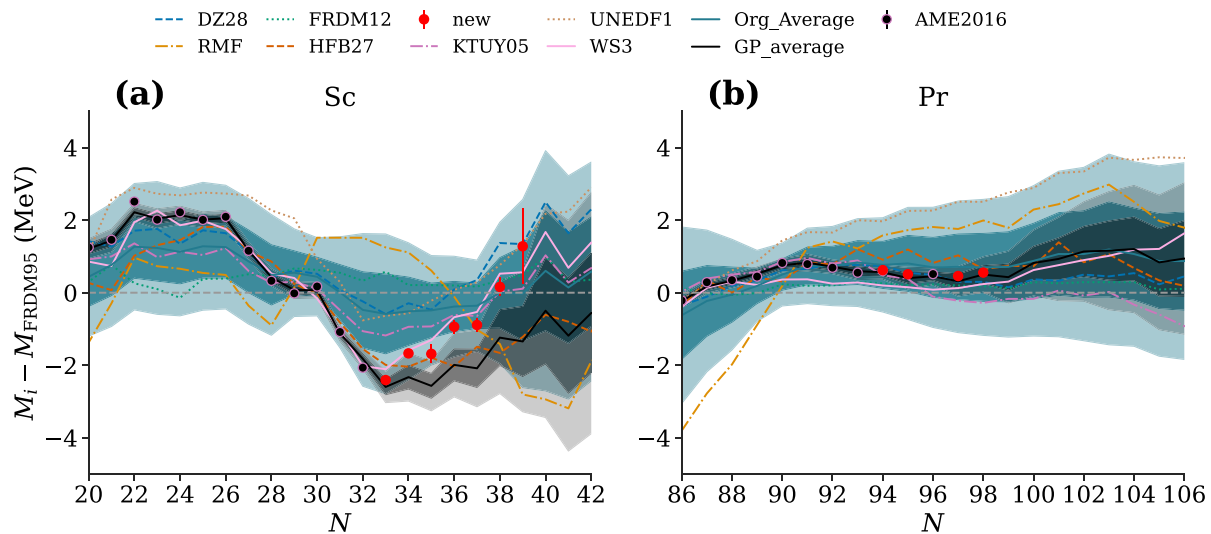


**Fig. 5** (Color online) Mass differences between different mass models relative to FRDM95 for Sc **(a)** and Pr **(b)** isotopic chains. The black solid points correspond to the experimental data in AME2016, while the red ones are new masses taken from [66–70]. The dark gray

(blue) and light gray (blue) bands show the  $1\sigma$  and  $2\sigma$  standard deviations of the simple average of GP-8D (original mass models), respectively

be 1 MeV more bound. For  $^{54,55}\text{Sc}$ , they were measured using high-precision MR-ToF [66, 67] and our GP predictions agree with their results. The deviations for neutron-rich Sc isotopes may reflect the limitations of the  $B\rho$ -ToF technique rather than the GP's extrapolation deficiency. High-precision mass measurements of Sc isotopes are required for more critical validation.

We also note in Fig. 5 that the averaged uncertainty bands for Pr and Sc isotopes do not grow at the same pace—the uncertainties of Pr grow more slowly than those of Sc isotopes. This is mainly due to the fact that the mass data around neutron-rich Pr isotopes are more constrained (with more higher precision experimental data) than that of neutron-rich Sc isotopes. The Pr isotopes are located in the heavier region, where more nuclides are available than those of Sc. Because the GP is completely driven by the data, we obtain a more reliable extrapolation if the surrounding training dataset is more accurate. Another possibility is the way in which each model describes the nuclear structure effects in its respective region. While the uncertainty of the averaged GP-refined models increases steadily as they move away from the known region, this represents a limitation of the current experimental constraints rather than inherent model deficiencies.

## 6 Conclusion

The present study demonstrates that Gaussian Processes significantly improve the accuracy of nuclear mass models. By capturing the inherent and systematic biases rooted in each model, GPs enhance the reliability of mass predictions, particularly for nuclei far from stability, which are critical for astrophysical modeling. Through the incorporation of physics-informed features, our refined models achieved a sub-100 keV accuracy for nuclei with  $Z, N \geq 8$ , representing a substantial improvement over the traditional mass models. By comparing with high-precision ( $\sigma^{\text{exp}} < 100$  keV) mass data from AME2020, we confirmed that our GP-refined mass models achieved an rms deviation below 150 keV for short-range extrapolation, with the Matérn kernel offering particularly robust uncertainty quantification compared to the RBF kernel. While the GP-refined models show robust performance near known regions, their reliability for  $r$ -process regions requires further validation. Through theory-to-theory validation, we noticed that the rms deviations first increased as one moved away from the known region. However, the flat area in the intermediate region ( $5 \leq r \leq 13$ ) indicates that our GP model can better capture systematic

patterns even at larger extrapolation distances. A comprehensive set of mass predictions (Table 1 in SM [59]) for unknown masses in AME2020 serves as a valuable benchmark for future mass measurements and astrophysical calculations. Given the inherent capability of GP for interpolation, we can also use our models to identify outliers in the existing mass data. Future studies should explore the impact of length-scale parameters and validate models using high-precision data.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s41365-026-01939-w>.

**Author contributions** All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Wen-Jia Huang, Keisuke Fujii, and Hao-Zhao Liang. The first draft of the manuscript was written by Wen-Jia Huang and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Funding** Open Access funding provided by The University of Tokyo.

**Data availability** The data that support the findings of this study are openly available in Science Data Bank at 31253.11.sciencedb.j00186.01007 and 10.57760/sciencedb.j00186.01007.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. G. Mathews, J. Cowan, New insights into the astrophysical r-process. *Nature* **345**, 491–494 (1990). <https://doi.org/10.1038/345491a0>
2. F.-K. Thielemann, A. Arcones, R. Käppeli et al., What are the astrophysical sites for the r-process and the production of heavy elements? *Prog. Part. Nucl. Phys.* **66**, 346–353 (2011). <https://doi.org/10.1016/j.pnpnp.2011.01.032>
3. M. Arnould, S. Goriely, K. Takahashi, The r-process of stellar nucleosynthesis: Astrophysics and nuclear physics achievements and mysteries. *Phys. Rep.* **450**, 97–213 (2007). <https://doi.org/10.1016/j.physrep.2007.06.002>
4. T. Kajino, W. Aoki, A. Balantekin et al., Current status of r-process nucleosynthesis. *Prog. Part. Nucl. Phys.* **107**, 109–166 (2019). <https://doi.org/10.1016/j.pnpnp.2019.02.008>
5. D.M. Siegel, r-Process nucleosynthesis in gravitational-wave and other explosive astrophysical events. *Nature Rev. Phys.* **4**, 306–318 (2022). <https://doi.org/10.1038/s42254-022-00439-1>
6. M. Mumpower, R. Surman, G. McLaughlin et al., The impact of individual nuclear properties on r-process nucleosynthesis. *Prog. Part. Nucl. Phys.* **86**, 86–126 (2016). <https://doi.org/10.1016/j.pnpnp.2015.09.001>
7. M.R. Mumpower, R. Surman, D.-L. Fang et al., Impact of individual nuclear masses on r-process abundances. *Phys. Rev. C* **92**, 035807 (2015). <https://doi.org/10.1103/PhysRevC.92.035807>
8. T.M. Sprouse, R. Navarro Perez, R. Surman et al., Propagation of statistical uncertainties of Skyrme mass models to simulations of r-process nucleosynthesis. *Phys. Rev. C* **101**, 055803 (2020). <https://doi.org/10.1103/PhysRevC.101.055803>
9. D. Atanasov, P. Ascher, K. Blaum et al., Precision mass measurements of  $^{129-131}\text{Cd}$  and their impact on stellar nucleosynthesis via the rapid neutron capture process. *Phys. Rev. Lett.* **115**, 232501 (2015). <https://doi.org/10.1103/PhysRevLett.115.232501>
10. R. Orford, N. Vassh, J.A. Clark et al., Precision mass measurements of neutron-rich neodymium and samarium isotopes and their role in understanding rare-earth peak formation. *Phys. Rev. Lett.* **120**, 262702 (2018). <https://doi.org/10.1103/PhysRevLett.120.262702>
11. T. Yamaguchi, H. Koura, Y. Litvinov et al., Masses of exotic nuclei. *Prog. Part. Nucl. Phys.* **120**, 103882 (2021). <https://doi.org/10.1016/j.pnpnp.2021.103882>
12. D. Lunney, J.M. Pearson, C. Thibault, Recent trends in the determination of nuclear masses. *Rev. Mod. Phys.* **75**, 1021–1082 (2003). <https://doi.org/10.1103/RevModPhys.75.1021>
13. X.Y. Xu, L. Deng, A.X. Chen et al., Improved nuclear mass formula with an additional term from the Fermi gas model. *Nucl. Sci. Tech.* **35**, 91 (2024). <https://doi.org/10.1007/s41365-024-01450-0>
14. M.-T. Wan, O. Li, A. Liu et al., Properties of the drip-line nucleus and mass relation of mirror nuclei. *Nucl. Sci. Tech.* **36**, 26 (2025). <https://doi.org/10.1007/s41365-024-01633-9>
15. A. Sobczewski, Y.A. Litvinov, Predictive power of nuclear-mass models. *Phys. Rev. C* **90**, 017302 (2014). <https://doi.org/10.1103/PhysRevC.90.017302>
16. J.D. McDonnell, N. Schunck, D. Higdon et al., Uncertainty quantification for nuclear density functional theory and information content of new measurements. *Phys. Rev. Lett.* **114**, 122501 (2015). <https://doi.org/10.1103/PhysRevLett.114.122501>
17. A. Idini, Statistical learnability of nuclear masses. *Phys. Rev. Res.* **2**, 043363 (2020). <https://doi.org/10.1103/PhysRevResearch.2.043363>
18. S. Gazula, J. Clark, H. Bohr, Learning and prediction of nuclear stability by neural networks. *Nucl. Phys. A* **540**, 1–26 (1992). [https://doi.org/10.1016/0375-9474\(92\)90191-L](https://doi.org/10.1016/0375-9474(92)90191-L)
19. K. Gernoth, J. Clark, J. Prater et al., Neural network models of nuclear systematics. *Phys. Lett. B* **300**, 1–7 (1993). [https://doi.org/10.1016/0370-2693\(93\)90738-4](https://doi.org/10.1016/0370-2693(93)90738-4)
20. S. Athanassopoulos, E. Mavrommatis, K. Gernoth et al., Nuclear mass systematics using neural networks. *Nucl. Phys. A* **743**, 222–235 (2004). <https://doi.org/10.1016/j.nuclphysa.2004.08.006>
21. T. Bayram, S. Akkoyun, S.O. Kara, A study on ground-state energies of nuclei by using neural networks. *Ann. Nucl. Energ.* **63**, 172–175 (2014). <https://doi.org/10.1016/j.anucene.2013.07.039>
22. H.F. Zhang, L.H. Wang, J.P. Yin et al., Performance of the levenberg-marquardt neural network approach in nuclear mass prediction. *J. Phys. G Nucl. Part. Phys.* **44**, 045110 (2017). <https://doi.org/10.1088/1361-6471/aa5d78>
23. R.-D. Lasserri, D. Regnier, J.-P. Ebran et al., Taming nuclear complexity with a committee of multilayer neural networks. *Phys. Rev. Lett.* **124**, 162502 (2020). <https://doi.org/10.1103/PhysRevLett.124.162502>

24. A. Sharma, A. Gandhi, A. Kumar, Learning correlations in nuclear masses using neural networks. *Phys. Rev. C* **105**, L031306 (2022). <https://doi.org/10.1103/PhysRevC.105.L031306>
25. X.-C. Ming, H.-F. Zhang, R.-R. Xu et al., Nuclear mass based on the multi-task learning neural network method. *Nucl. Sci. Tech.* **33**, 48 (2022). <https://doi.org/10.1007/s41365-022-01031-z>
26. A. Boehnlein, M. Diefenthaler, N. Sato et al., Colloquium: Machine learning in nuclear physics. *Rev. Mod. Phys.* **94**, 031003 (2022). <https://doi.org/10.1103/RevModPhys.94.031003>
27. Z.-P. Gao, Q.-F. Li, Studies on several problems in nuclear physics by using machine learning. *Nuclear Techniques (in Chinese)* **46**, 080009 (2023). <https://doi.org/10.11889/j.0253-3219.2023.hjs.46.080009>
28. K.Z. Tan, W.Q. Wan, J. Liu, Continuous Bayesian probability estimator in predictions of nuclear masses. *Nuclear Techniques (in Chinese)* **48**, 250104 (2025). <https://doi.org/10.11889/j.0253-3219.2025.hjs.48.250104>
29. X.H. Wu, P.W. Zhao, Predicting nuclear masses with the kernel ridge regression. *Phys. Rev. C* **101**, 051301(R) (2020). <https://doi.org/10.1103/PhysRevC.101.051301>
30. Y.-Y. Guo, X.-Q. Tang, H.X. Liu et al., Nucleon separation energy predictions with the Kernel ridge regression. *Nuclear Techniques (in Chinese)* **48**, 250096 (2025). <https://doi.org/10.11889/j.0253-3219.2025.hjs.48.250096>
31. R.M. Neal, *Bayesian learning of neural networks*, Springer. N. Y. (1996). <https://doi.org/10.1007/978-1-4612-0745-0>
32. R. Utama, J. Piekarewicz, H.B. Prosper, Nuclear mass predictions for the crustal composition of neutron stars: a Bayesian neural network approach. *Phys. Rev. C* **93**, 014311 (2016). <https://doi.org/10.1103/PhysRevC.93.014311>
33. R. Utama, J. Piekarewicz, Refining mass formulas for astrophysical applications: A Bayesian neural network approach. *Phys. Rev. C* **96**, 044308 (2017). <https://doi.org/10.1103/PhysRevC.96.044308>
34. Z. Niu, H. Liang, Nuclear mass predictions based on Bayesian neural network approach with pairing and shell effects. *Phys. Lett. B* **778**, 48–53 (2018). <https://doi.org/10.1016/j.physletb.2018.01.002>
35. Z.M. Niu, H.Z. Liang, Nuclear mass predictions with machine learning reaching the accuracy required by  $r$ -process studies. *Phys. Rev. C* **106**, L021303 (2022). <https://doi.org/10.1103/PhysRevC.106.L021303>
36. U.B. Rodr  guez, C.Z. Vargas, M. Gon  salves et al., Bayesian neural network improvements to nuclear mass formulae and predictions in the superheavy elements region. *Euro. Phys. Lett.* **127**, 42001 (2019). <https://doi.org/10.1209/0295-5075/127/42001>
37. Z.M. Niu, H.Z. Liang, B.H. Sun et al., Predictions of nuclear  $\beta$ -decay half-lives with machine learning and their impact on  $r$ -process nucleosynthesis. *Phys. Rev. C* **99**, 064307 (2019). <https://doi.org/10.1103/PhysRevC.99.064307>
38. C.E. Rasmussen, C.K. Williams, *Gaussian processes for machine learning*. MIT Press (2006)
39. L. Neufcourt, Y. Cao, W. Nazarewicz et al., Bayesian approach to model-based extrapolation of nuclear observables. *Phys. Rev. C* **98**, 034318 (2018). <https://doi.org/10.1103/PhysRevC.98.034318>
40. M. Shelley, A. Pastore, A new mass model for nuclear astrophysics: Crossing 200 keV accuracy. *Universe* **7**, 131 (2021). <https://doi.org/10.3390/universe7050131>
41. L. Neufcourt, Y. Cao, S.A. Giuliani et al., Quantified limits of the nuclear landscape. *Phys. Rev. C* **101**, 044307 (2020). <https://doi.org/10.1103/PhysRevC.101.044307>
42. E. Alhassan, D. Rochman, A. Vasiliev et al., Iterative Bayesian Monte Carlo for nuclear data evaluation. *Nucl. Sci. Tech.* **33**, 50 (2022). <https://doi.org/10.1007/s41365-022-01034-w>
43. E. Y  ksel, D. Soydaner, H. Bahtiyar, Nuclear mass predictions using machine learning models. *Phys. Rev. C* **109**, 064322 (2024). <https://doi.org/10.1103/PhysRevC.109.064322>
44. I. Bentley, J. Tedder, M. Gebran et al., High precision binding energies from physics-informed machine learning. *Phys. Rev. C* **111**, 034305 (2025). <https://doi.org/10.1103/PhysRevC.111.034305>
45. Z.-Y. Yuan, D. Bai, Z. Wang et al., Reliable calculations of nuclear binding energies by the gaussian process of machine learning. *Nucl. Sci. Tech.* **35**, 105 (2024). <https://doi.org/10.1007/s41365-024-01463-9>
46. C.F. v. Weizs  cker, Zur theorie der kernmassen. *Z. Physik* **96**, 431–458 (1935). <https://doi.org/10.1007/BF01337700>
47. H.A. Bethe, R.F. Bacher, Nuclear physics A. stationary states of nuclei. *Rev. Mod. Phys.* **8**, 82–229 (1936). <https://doi.org/10.1103/RevModPhys.8.82>
48. P. M  ller, A.J. Sierk, T. Ichikawa et al., Nuclear ground-state masses and deformations: Frdm (2012). *Atomic Data and Nuclear Data Tables* 109–110, 1–204 (2016). <https://doi.org/10.1016/j.adt.2015.10.002>
49. N. Wang, M. Liu, X. Wu, Modification of nuclear mass formula by considering isospin effects. *Phys. Rev. C* **81**, 044322 (2010). <https://doi.org/10.1103/PhysRevC.81.044322>
50. S. Goriely, S. Hilaire, M. Girod et al., The Gogny-Hartree-Fock-Bogoliubov nuclear-mass model. *Eur. Phys. J. A* **52**, 202 (2016). <https://doi.org/10.1140/epja/i2016-16202-3>
51. S. Goriely, N. Chamel, J.M. Pearson, Hartree-Fock-Bogoliubov nuclear mass model with 0.50 MeV accuracy based on standard forms of skyrme and pairing functionals. *Phys. Rev. C* **88**, 061302 (2013). <https://doi.org/10.1103/PhysRevC.88.061302>
52. M. Kortelainen, J. McDonnell, W. Nazarewicz et al., Nuclear energy density optimization: Large deformations. *Phys. Rev. C* **85**, 024304 (2012). <https://doi.org/10.1103/PhysRevC.85.024304>
53. J. Duflo, A. Zuker, Microscopic mass formulas. *Phys. Rev. C* **52**, R23–R27 (1995). <https://doi.org/10.1103/PhysRevC.52.R23>
54. M. Wang, G. Audi, F. Kondev et al., The AME2016 atomic mass evaluation. *Chin. Phys. C* **41**, 030003 (2017). <https://doi.org/10.1088/1674-1137/41/3/030003>
55. P. M  ller, J. Nix, W. Myers et al., Nuclear ground-state masses and deformations. *At. Data Nucl. Data Tables* **59**, 185–381 (1995). <https://doi.org/10.1006/adnd.1995.1002>
56. M. Liu, N. Wang, Y. Deng et al., Further improvements on a global nuclear mass model. *Phys. Rev. C* **84**, 014333 (2011). <https://doi.org/10.1103/PhysRevC.84.014333>
57. L. Geng, H. Toki, J. Meng, Masses, deformations and charge radii-nuclear ground-state properties in the relativistic mean field model. *Prog. Theoret. Phys.* **113**, 785–800 (2005). <https://doi.org/10.1143/PTP.113.785>
58. H. Koura, T. Tachibana, M. Uno et al., Nuclidic mass formula on a spherical basis with an improved even-odd term. *Prog. Theoret. Phys.* **113**, 305–325 (2005). <https://doi.org/10.1143/PTP.113.305>
59. See Supplementary Material at <https://doi.org/10.1007/s41365-026-01939-w>
60. R.F. Casten, D.S. Brenner, P.E. Haustein, Valence p-n interactions and the development of collectivity in heavy nuclei. *Phys. Rev. Lett.* **58**, 658–661 (1987). <https://doi.org/10.1103/PhysRevLett.58.658>
61. M. Wang, W. Huang, F. G. Kondev et al., The AME 2020 atomic mass evaluation (ii). tables, graphs and references. *Chinese Phys. C* **45**(3), 030003 (2021). <https://doi.org/10.1088/1674-1137/abddaf>
62. The GPy authors, GPy: A gaussian process framework in python. <http://github.com/SheffieldML/GPy> (2012–2015)

63. D. P. Dubhashi, A. Panconesi, Concentration of measure for the analysis of randomized algorithms. Cambridge University Press (2009)
64. J. Barea, A. Frank, J.G. Hirsch et al., Garvey-kelson relations and the new nuclear mass tables. *Phys. Rev. C* **77**, 041304 (2008). <https://doi.org/10.1103/PhysRevC.77.041304>
65. C. Xu, M. Bao, Improved mass relations of mirror nuclei. *Nucl. Sci. Tech.* **35**, 157 (2024). <https://doi.org/10.1007/s41365-024-01501-6>
66. E. Leistenschneider, E. Dunling, G. Bollen et al., Precision mass measurements of neutron-rich scandium isotopes refine the evolution of  $n = 32$  and  $n = 34$  shell closures. *Phys. Rev. Lett.* **126**, 042501 (2021). <https://doi.org/10.1103/PhysRevLett.126.042501>
67. S. Iimura, M. Rosenbusch, A. Takamine et al., Study of the  $n = 32$  and  $n = 34$  shell gap for Ti and V by the first high-precision multireflection time-of-flight mass measurements at bigrips-slowri. *Phys. Rev. Lett.* **130**, 012501 (2023). <https://doi.org/10.1103/PhysRevLett.130.012501>
68. Z. Meisel, S. George, S. Ahn et al., Nuclear mass measurements map the structure of atomic nuclei and accreting neutron stars. *Phys. Rev. C* **101**, 052801 (2020). <https://doi.org/10.1103/PhysRevC.101.052801>
69. S. Michimasa, M. Kobayashi, Y. Kiyokawa et al., Mapping of a new deformation region around  $^{62}\text{Ti}$ . *Phys. Rev. Lett.* **125**, 122501 (2020). <https://doi.org/10.1103/PhysRevLett.125.122501>
70. R. Orford, N. Vassh, J.A. Clark et al., Searching for the origin of the rare-earth peak with precision mass measurements across Ce-Eu isotopic chains. *Phys. Rev. C* **105**, L052802 (2022). <https://doi.org/10.1103/PhysRevC.105.L052802>