PANE Journal of Physics P. J. Phys. 02 (01), 047 (April 2026)

# EXPLORING PATTERNS IN NUCLEAR BINDING ENERGY: ISOTOPIC, ISOTONIC AND ISOBARIC VARIATIONS

Yubaraj Sharma*[1]

[1]*Department of Physics, Don Bosco College, Tura, Meghalaya, India*

*Corresponding Author: yubaraj.physics@donboscocollege.ac.in

(Received 22 September 2025; revised 12 December 2025; accepted 28 December 2025; published 6 April 2026)

Abstract: We explore patterns in the nuclear binding energy data and seek their explanation in the Semi-Empirical Mass Formula (SEMF). The variations in the binding energy across isotopic, isotonic and isobaric sequences is analysed via standard deviation; the result shows a general decreasing trend, with the steepest decline for isotopic sequence and the gentlest decline for isobaric sequence. We have also analysed the correlation between the number of nuclides and the standard deviation. Transformation of the data into a time-series by arranging it in specific order in terms of the associated proton, neutron and atomic mass numbers reveal interesting patterns. Differencing of the generated time-series is found to have similarities and strong correlation with nucleon separation energies.

Keywords: Binding Energy per Nucleon; Isobars; Isotopes; Isotones; time-series

PACS: 21.10.-k, 21.10.Dr, 29.85.-c

## 1 Introduction

Nature and consequently real-world data are unruly, far from the elegant mathematical portraits drawn on university blackboards. The nuclear binding energy $(\mathrm{BE})$ data may debatably be a good example of this familiar observation, and the simple and elegant formula used to model this data, called the Semi-Empirical Mass Formula (SEMF), although impressive, has limitations. Currently, there exists $\mathrm{BE}$ data of more than $3500$ nuclides [1, 2], although not huge by the "big data" standards of today, still substantial enough to permit data and statistical analysis attempted in the current article.

The chronological journey of the nuclear binding energy concept starts arguably with the laying of the foundations of atomic theory by John Dalton. Observations at the time revealed that atomic masses were very close to whole numbers [1], but small deviations were also noticed, initially attributed to lack of experimental precision [3]. After the discovery of the atomic nucleus by Rutherford, it became abundantly clear that most of the mass of the atom was concentrated in the very small region of the nucleus [4]. With more precise measurements, particularly using high-precision mass-spectrographs, the "mass defect" or the deviation from whole numbers became undeniable [5]. Explanations were provided for this apparent deviation, but were not entirely satisfactory, as the structure of the nucleus was unclear at that time [6, 7], which only settled with the discovery of the neutron [8]. Bethe and Backer [9] clearly indicated the direct correspondence between the mass defect and the binding energy and even calculated the conversion factor from atomic mass units to MeV as $931.05 \pm 0.15\ \mathrm{MeV}$[2]. They redefined packing fraction as the ratio of the mass defect to the mass number[3] and formulated what is now known as the Semi-Empirical Mass Formula (SEMF) based on an earlier work by Weizsäcker[4].

Since the current state of knowledge of the nuclear structure places only protons and neutrons as constituents of the nucleus, it is obvious that the identity of a nuclide and therefore its binding energy is tied to the number of

[1]For example, see Prout's hypothesis or Aston's whole number rule.

[2]The notation used in the original publication was MV.

[3]This is slightly different from the definition of Aston [6], who had an additional factor of 10,000.

[4]Some literature cite this as the Bethe–Weizsäcker formula, crediting Weizsäcker also for the discovery, but the unfamiliarity of the author with the language of the original article [10] has prevented further comments.

these "elementary" constituents of the nucleus. It is, however, customary to set aside the neutron number $(N)$ and use the proton number [5] $(Z)$ and the mass number $(A)$ to represent a nuclide. Consequently, in this representation, the binding energy $\mathrm{BE}(A, Z)$ of a nuclide is defined as a function of the mass number $A$ and proton number $Z$ with the defining equation [4, 11], given by

$$\mathrm{BE}(A, Z) = Zm_p + (A - Z)m_n - M(A, Z) \tag{1}$$

where $m_p$, $m_n$ and $M(A, Z)$ represent the mass of a free proton, a free neutron and the actual mass of the nuclide respectively. All masses are expressed in energy units (usually $\mathrm{MeV}$ or $\mathrm{keV}$), or they can be converted to $\mathrm{MeV}$ units from mass units (amu) using the more recent conversion factor from the CODATA 2022 [12] values:

$$931.49410372 \ \frac{\mathrm{MeV}}{\mathrm{amu}}.$$

The SEMF is given by [11, 13, 14],

$$\mathrm{BE}(A, Z) = a_v A - a_s A^{2/3} - a_c \frac{Z^2}{A^{1/3}} - a_{as} \frac{(A - 2Z)^2}{A} + \delta \tag{2}$$

where $a_v$, $a_s$, $a_c$ and $a_{as}$ are constant coefficients of the volume term, the surface term, the Coulomb term, and the asymmetry term respectively, while the pairing term $\delta$ has two general formulations with $A^{-3/4}$ or $A^{-1/2}$; the sign of the associated constant depends on proton–neutron pairing. The first three terms of SEMF correspond to an aspect of the liquid drop model of the nucleus, while the other two are quantum mechanical corrections; however, all the coefficients are derived empirically from data. These five coefficients are updated regularly as more experimental data becomes available, still, differences [6] persist [13, 15, 16]. These differences between the SEMF predictions and the experimental results, though small, have been the cause of some investigation [17, 18]. The reason for the nomenclature "Semi-Empirical Mass Formula" is historical, as the formula was used to determine the mass of nuclides [7]. However, to facilitate comparison between different nuclides, it is conventional to use the ratio of $\mathrm{BE}$ to the mass number $(A)$, i.e., the binding energyper nucleon; we will hereafter use the shorter notation $\mathrm{BE}/A$ for binding energy per nucleon.

In the present paper, we explore patterns in the $\mathrm{BE}$ data made evident by data analysis techniques and seek explanations from the SEMF. Time-series analysis [19, 20] has a unique set of tools to extract patterns from data. While it is far more prevalent in the fields of finance and meteorology, it has had some applications in physics [21, 22]. Our approach is marginally different from traditional usage — the nuclear binding energy requires two independent variables, usually $A$ and $Z$ (see equations (1) and (2)). We flattened these two dimensions into a single dimension by ordering/sorting one variable followed by the other variable. This allows us to apply traditional methods of time-series analysis to the transformed data.

## 2 Computational aspect

To carry out the calculations and visualizations, two updated data sources were identified: the AME2020 [1, 23] and NuDat3 [2]; since NuDat3 also sources most of its data from AME2020, we opted for the more user-friendly NuDat3 dataset. For computation, we used the open-source Python programming language, with the `Pandas` and `SciPy` libraries/packages for most of the data analysis tasks, and the `matplotlib` package for drawing graphs [24–27].

To calculate [8] the theoretical values from the semi-empirical mass formula, we have used the coefficients and consequently the formula used by Benzaid *et al.* [13]; note that these authors have used $Z^2$ instead of the theoretically more palatable $Z(Z - 1)$ for the Coulomb term. Minor data cleaning was carried out before actual analysis, viz., removal of zero and negative $\mathrm{BE}$ values.

Primarily, we have a dependent variable $(\mathrm{BE}/A)$ and a set of three variables $(Z,\ N,\ A)$, only two of which are independent due to the relation/definition $A = Z + N$. To make it shorter and easier to refer to them, we shall call this collection of three variables the 'nuktriads' [9]. Each set of $(\mathrm{BE}/A)$ values with a common nuktriad value

[5] The number of protons in a nucleus is generally called atomic number, but we use proton number in consonance with the term neutron number

[6] See LDM disparity graph (defined later in the article) in Figs. 4-9 for an indication of the discrepancy

[7] Substituting equation 1 in equation 2 to get the formula for mass

[8] NuDat3 also provides their own SEMF values, however, we choose to calculate values directly

[9] Short for nuclear-triads (collection of three nuclear variables)

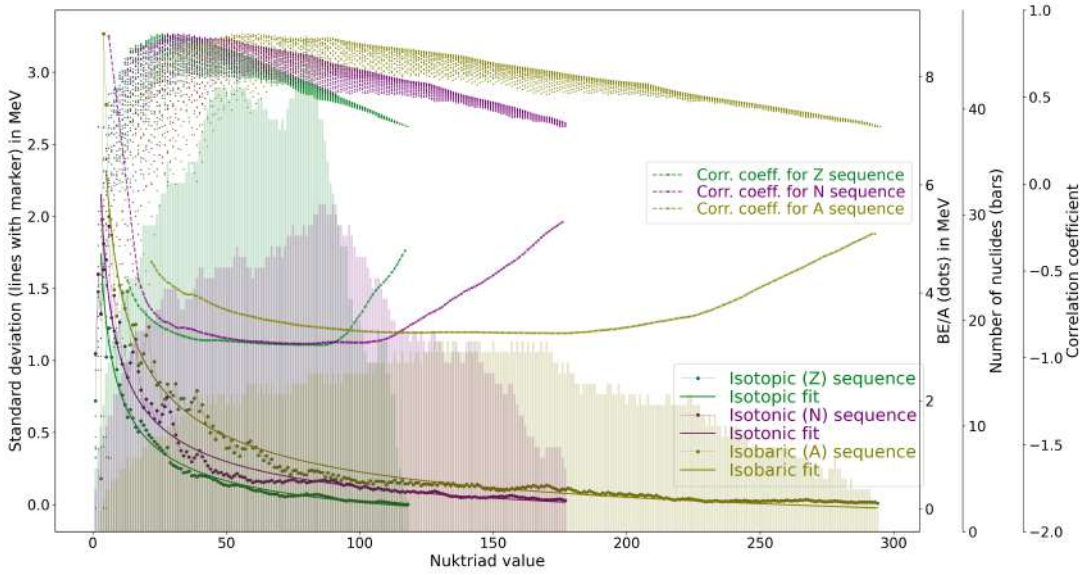



Figure 1: Plot of standard deviation of BE/A variation across isobaric, isotopic and isotonic sequences along with regression fit. Other plots include: Scatter-plot of actual BE/A; bar-graph of number of nuclides for each nuktriad value; correlation between and number of nuclides in the corresponding sequence. All plots have common x-axis.

will be called a sequence. Thus, an isotopic sequence is the set of nuclides with the same value of $Z$, an isotonic sequence is the set of nuclides with the same $N$ value, and an isobaric sequence is the set of nuclides with the same $A$ value.

## 3 Results and Discussion

We first look at the variations in the binding energy data within the isotopic, isotonic and isobaric sequences, quantified using the standard deviation $(\sigma)$ as a measure; these are plotted in Fig. 1. We have used the sample standard deviation, defined as [28]:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{3}$$

where $x_i$ represents the $i^{th}$ value of the variable $x$ having $N$ data points, and $\bar{x}$ is its mean value. This basic definition is specified here for easy reference and reproducibility.

To make optimum use of space and to substantiate the relationships between different groupings, we have included multiple graphs (five sets) in Fig. 1. The common $x$-axis in the figure represents the value of the nuktriad variables — for example, in the figure of $(\mathrm{BE}/A)$ and other associated data plotted against $Z$, the $x$-axis represents the corresponding $Z$ values. Similarly, when the same set of data is plotted against $N$, the $x$-axis represents the corresponding $N$ values and so on. The colours in the figure are matched to corresponding groupings (one specific colour for a specific nuktriad variable). The primary data of our interest are the standard deviations, for which the $y$-scale is given on the left side of the graph. To corroborate the nature of the standard deviation curves obtained, we have included scatter plots of the $(\mathrm{BE}/A)$, arranged as per the three nuktriad sequences ($y$-scale attached to the right side of the graph).

We have also included bar graphs showing the number of nuclides within each nuktriad sequence. We observe that these bar graphs expectedly show broadly a Gaussian/bell-shaped pattern. This is expected because at lower values of the nuktriad variables there are not many stable combinations possible, while nuclides with high nuktriad values have fewer reported data due to their unstable nature and experimental difficulties.

Another aspect that we have packed into the figure is the Pearson correlation coefficient. This statistic $(r)$ gives us a measure of the linear association between two time-series; it is given by [28]:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{4}$$

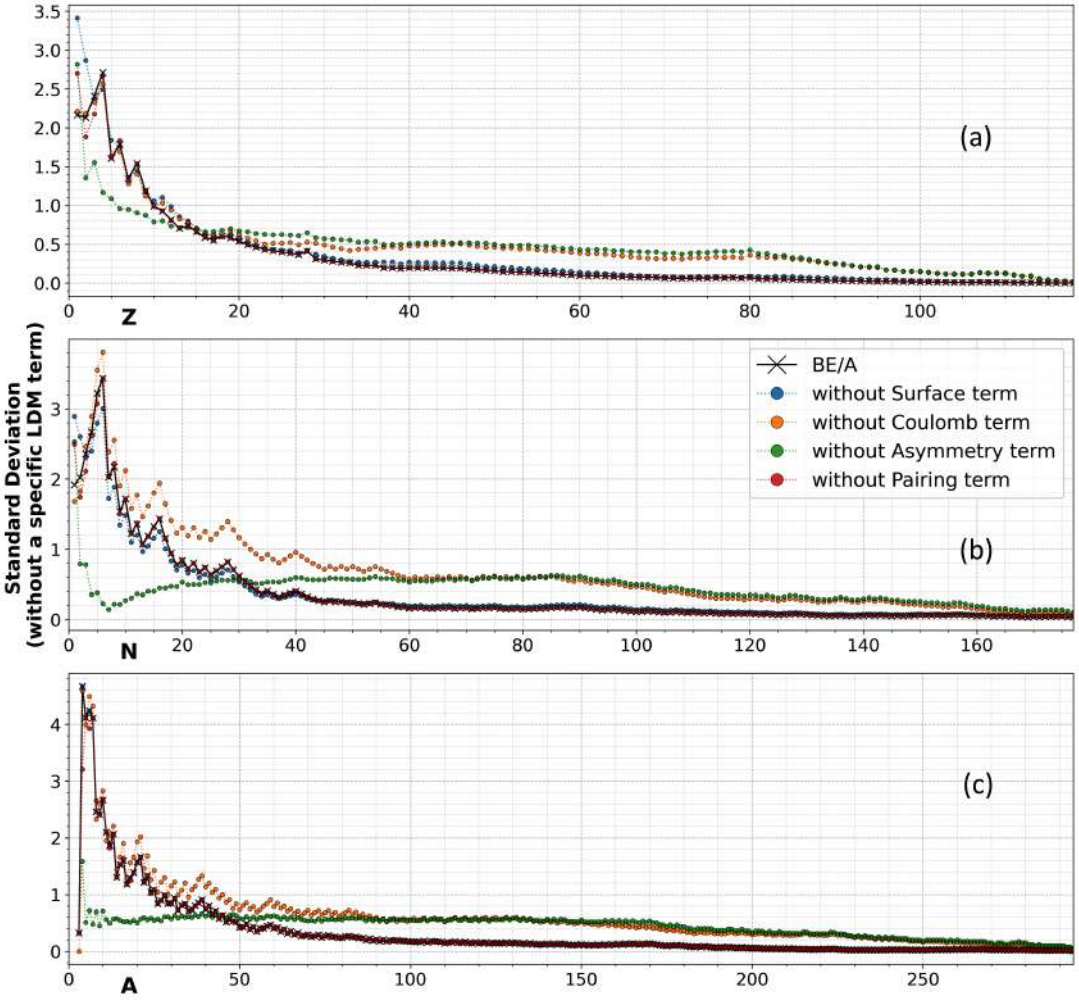



Figure 2: Standard deviation calculated from SEMF with only the nuclides present in the used dataset; the curve with black line contains all terms, other curves have one specific term of the SEMF excluded from the calculation. The x-axis label in three subplots represents the nuktriad sequence associated with the calculated standard deviation.

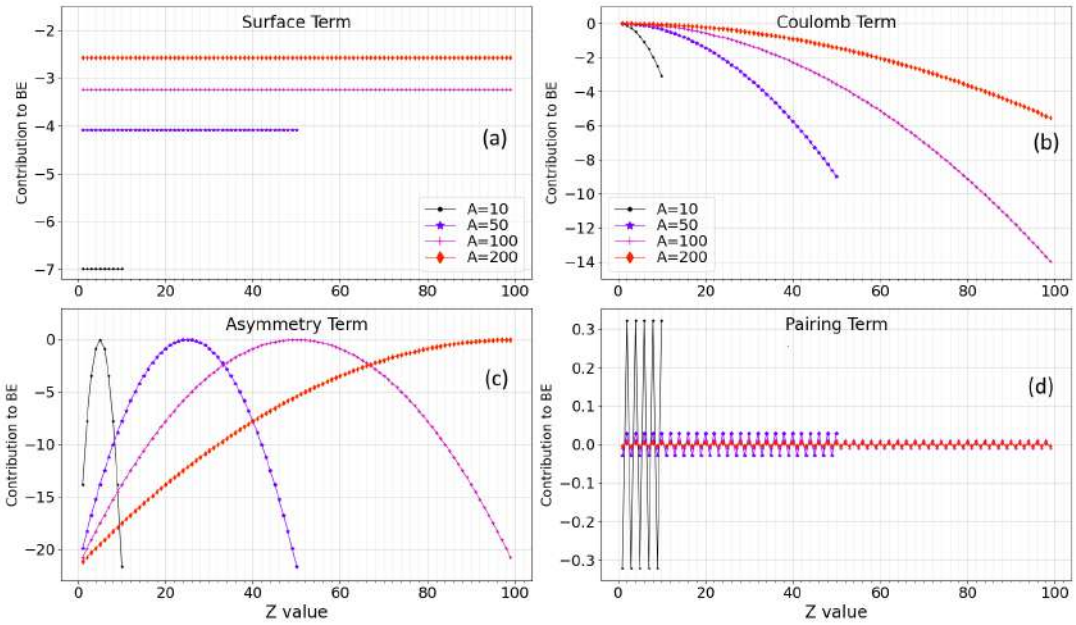



Figure 3: Variation of each term of the semi-empirical mass formula with values (range: 0 to 100) for four specific values of the atomic mass . Volume term is not included as it is a constant.

where $x_i$ and $y_i$ represent the $i^{th}$ value of two correlated variables $x$ and $y$ respectively; $\bar{x}$ and $\bar{y}$ are the mean values of the respective variables.

The correlation plots given in Fig. 1 represent the coefficients of the Pearson correlation between standard deviation values and the number of nuclides; each point (say at $x$-axis or nuktriad value $n$) of the graph represents the correlation between the two variables from the beginning, i.e., nuktriad value 1, of the dataset up to nuktriad value $n$. Only points with statistically significant correlations, with $p$-value $< 0.05$, are plotted. We find that these curves broadly exhibit a U shape. This shape can broadly be explained as follows. The first part of a bell curve, where it is increasing, correlates negatively with the decreasing standard deviation curve; the other part of the bell curve (beyond the peak), where it is also decreasing, gives positive correlation — cumulatively and in conjunction with finer variations, they produce the U curve.

It is not straightforward to find a simple explanation from SEMF for the decreasing trend of the standard deviation ($\sigma$) in all the three nuktriad sequences. We employ the graphical route for explanation. In Fig. 2, we have shown [10] the standard deviation curves of partial SEMF values, i.e., SEMF calculated without a particular term, to observe the impact of the absent term on $\sigma$. The figure clearly shows that the Coulomb and the Asymmetry terms contribute the most to the overall $\sigma$, the latter showing significant drop during the initial part of the curve, and then (about $1/6$th of the way) it rises above the overall $\mathrm{BE}/A$ curve.

The remaining SEMF terms — the Pairing term and to a lesser extent the Surface term — have a barely noticeable individual contribution, as seen in the figure. To explain these observed behaviours of the SEMF terms, we have drawn a representative graph of these terms in Fig. 3: the explanation can be found in the relatively high variation of the Coulomb and the Asymmetry terms with $Z$ (as well as with $A$) compared to the Pairing and Surface terms. The high fluctuations in the Pairing term at low values of $A$ [11] show up in Fig. 2, around the first few points of the $x$-axis, but require close scrutiny.

Another pattern that emerges out of Fig. 1 relates to the grouping: isotopic sequences have generally the lowest $\sigma$ values, followed by isotonic and isobaric sequences. To quantify this behaviour, we decided to perform a statistical fit/regression with some function of known characteristics. The shape of the $\sigma$ curves shows some similarity to a power-law function; finally, we settled on the following function (after some hit-and-trial [12]):

$$f(x) = a\,x^{-1} + b \tag{5}$$

[10] We have not included the Volume term in the plot as it a constant and therefore will not contribute to $\sigma$ (see Eq. (4))

[11] In Fig. 3d, in the subplot for Pairing term, see the curve for A=10

[12] The choice of function was based on regression with several values for the power of x, leading to the subjective conclusion that taking a square root had the best combination of 'r-squared' goodness of fit values, optimum number of parameters and elegance of the fitting function

| Time-series | $\mathbf{a}$ | $\mathbf{b}$ | $r$-squared |
|---|---|---|---|
| Isotopic ($Z$) | 3.588 | -0.342 | 0.956267 |
| Isotonic ($N$) | 4.251 | -0.303 | 0.91717 |
| Isobaric ($A$) | 6.016 | -0.375 | 0.97079 |

Table 1: Coefficients of a regression model (Eq. (5)) of standard deviation in $\mathrm{BE}/A$ data as a function of corresponding nuktriad variables.

where $a$ and $b$ are parameters to be determined from the fit. The values of the fit for different sequences are given in table 1. Now, the parameter $a$ in the fitting function controls the gentleness of bending in the decreasing curve; higher the value of $a$, the gentler (or less steep) is the decrease of $f(x)$, and obviously the parameter $b$ controls the vertical location of the curve. Values obtained for parameter $a$ (table 1) show that the sharpest decrease is for isotopic sequence, while the isobaric sequence has the gentlest decrease. Fit lines are also plotted alongside the $\sigma$ curves.

The various aspects of $\mathrm{BE}/A$ data as well as the SEMF appear prominently when we try to reimagine the data as time-series (instead of a scatter plot, c.f. Fig. 1) by sorting the entire dataset, first by one of the nuktriad variables, followed by another. The independent variable, or the $x$-axis, thus represents simply the order of arrangement of the dataset. Out of the six permutations possible for ordering any two of the nuktriad variables, we have chosen, for economy, three arrangements to represent isotopic, isotonic and isobaric sequences: $(Z, N)$, $(N, Z)$ and $(A, Z)$. The choice stemmed from the assumed order of importance of the nuktriad variables in respect of nuclear binding energy: $Z > N > A$ [13]. For brevity, we shall continue with the same notation but with bold font, i.e., we shall denote as $(\mathbf{Z}, \mathbf{N})$ the sorting of data first along increasing $Z$ values followed by increasing $N$ values. This order of sorting, viz., $(\mathbf{Z}, \mathbf{N})$, leads to $\mathrm{BE}/A$ graph being arranged in segments, where each segment represents an isotopic sequence (a set of constant $Z$ values) within which the neutron number $(N)$ increases monotonically. The other two orders of arrangement, viz., $(\mathbf{N}, \mathbf{Z})$ and $(\mathbf{A}, \mathbf{Z})$, have similar interpretations — isotonic and isobaric sequences respectively — within which $Z$ increases monotonically. Each segment in all of the three $\mathrm{BE}/A$ time-series analysed can be identified visually by their distinct downward-opening parabolic portions.

The time-series of $\mathrm{BE}/A$ data with different orders of sorting are shown in Figs. 4–6. Within the same subplot, we have included predictions from SEMF as well as the difference between the two. For reference, we shall call the plot of the difference between SEMF and data as the LDM disparity curve[14] (represented as $(\mathrm{SEMF} - \mathrm{data})$ in the figure).

We have included a zoomed section of the time-series ($x$-axis value from 2700 to 3000) to show detailed structures present in the plotted data. Corresponding nuktriad variables for the figures are plotted in the lower subplots. We have also included similarly transformed time-series of the various individual terms of the SEMF, so that the nature of their contribution to the total value can be appreciated. We have inset in the SEMF figure itself, a zoomed section of the plot.

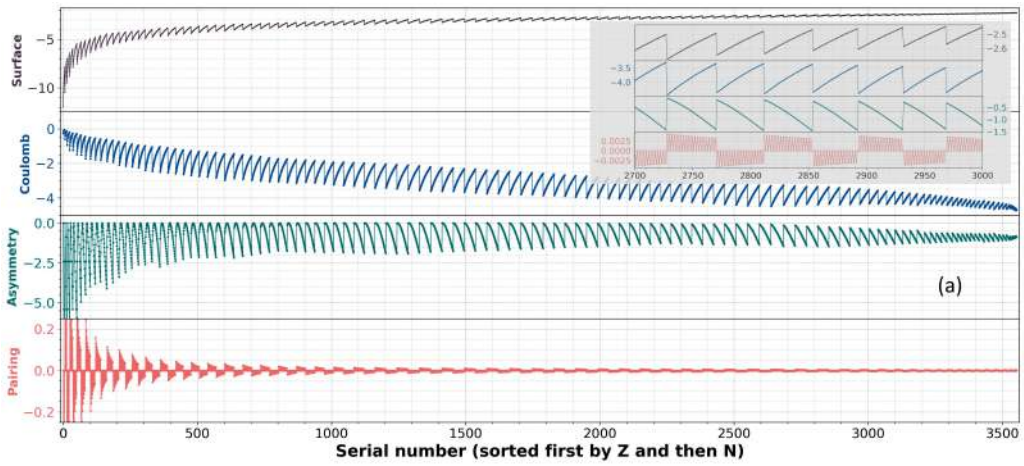
The shapes of the $A$, $Z$ and $N$ scatter plots are trivial to explain, arising out of the time-series ordering and their simple relationship $A = Z + N$. Fig. 4a containing plots of the $(\mathbf{Z}, \mathbf{N})$ sorting clearly shows that it is the interplay between the Coulomb and the Asymmetry terms that gives rise to the parabolic segments in the time-series. We also see that the Surface term (see inset of Fig. 4a) strengthens the Coulomb term due to their co-directional trend. Obviously, they have a direct correspondence with Fig. 3. For $(\mathbf{Z}, \mathbf{N})$ time-series (Fig. 4b), each segment implies constant $Z$ and increasing $N$, which leads to increasing $A$, and in Fig. 3, increasing $A$ for a constant value of $Z$ for Surface and Coulomb terms clearly indicates increasing energy values, which explains the corresponding increasing trend of these terms in Fig. 4a.

The Asymmetry term in Fig. 3c shows parabolas for constant $A$, the vertex of which occurs at $Z = N$. Before the vertex $(Z < N)$, on looking along a constant value of $Z$, the energy associated with the Asymmetry term increases with increasing $A$; and beyond the vertex $(Z > N)$, the energy starts decreasing — this results in a parabola for constant $Z$ and increasing $A$ also. These $A$-parabolas can be observed in the asymmetry term in Fig. 4a roughly before 700 units of the $x$-axis.

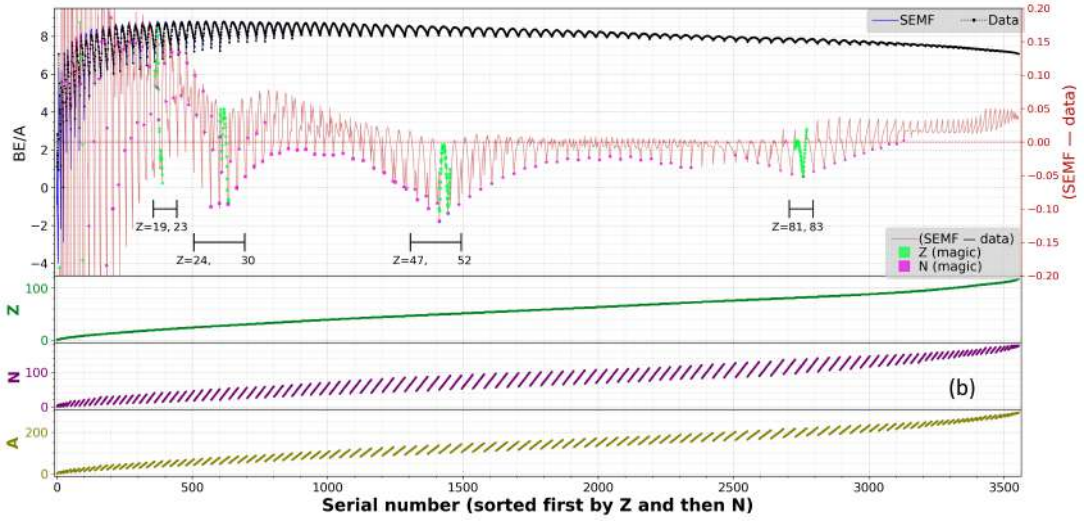
The parabolic segments in $\mathrm{BE}/A$ time-series in $(\mathbf{N}, \mathbf{Z})$ and $(\mathbf{A}, \mathbf{Z})$ sorting are also caused by the interplay between the Coulomb and the Asymmetry terms (Figs. 5a & 6a); the Surface term however changes affiliations here, acting collaboratively with the Asymmetry term and against the Coulomb term for $(\mathbf{N}, \mathbf{Z})$ sorting, while

[13] See table 1 and accompanying discussion; higher variation of $\mathrm{BE}/A$ within a nuktriad is assumed to imply more impact

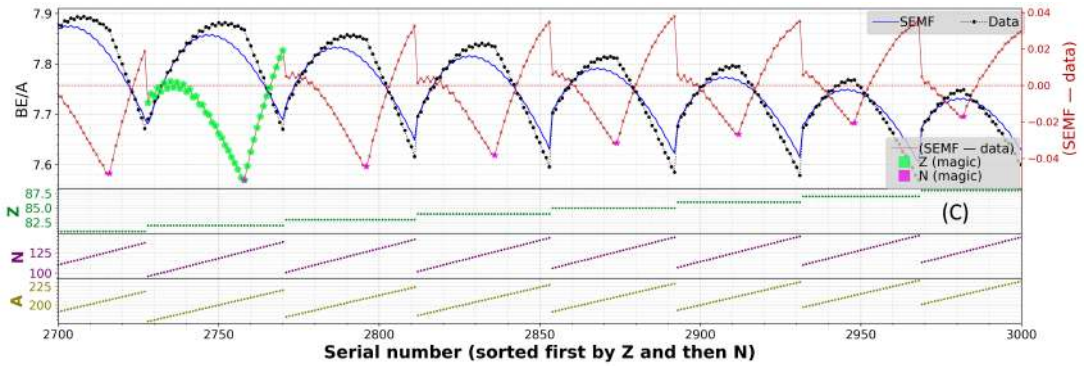
[14] LDM for Liquid Drop Model




(a)

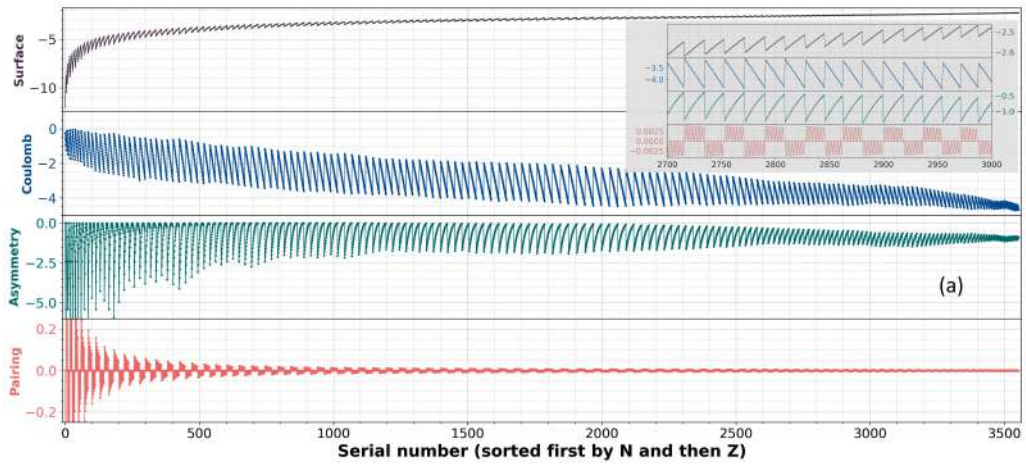



(b)



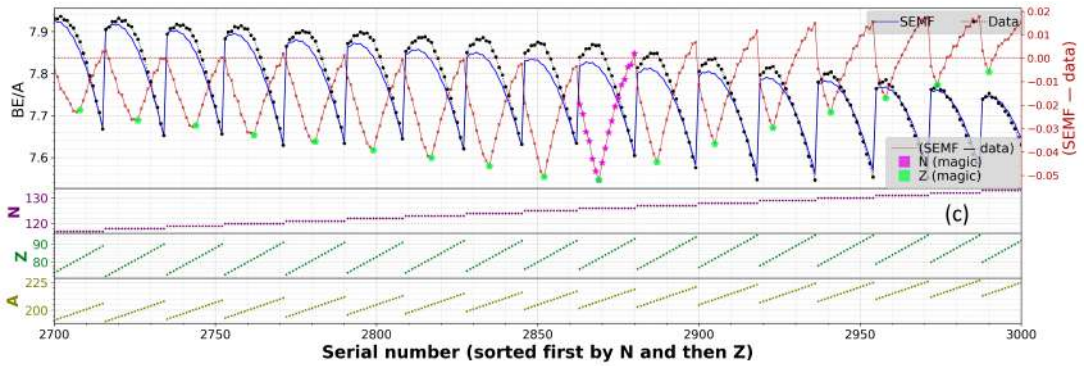

(c)

Figure 4: Time-series plots with $(Z, N)$ sorting. Part (a) shows the various terms of the SEMF. Part (b) shows the $\mathrm{BE}/A$ time-series and associated graphs, and part (c) is a zoomed version of part (b) from 2700 to 3000 on the $x$-axis. The detailed legend for magic numbers is given in Fig. 10.



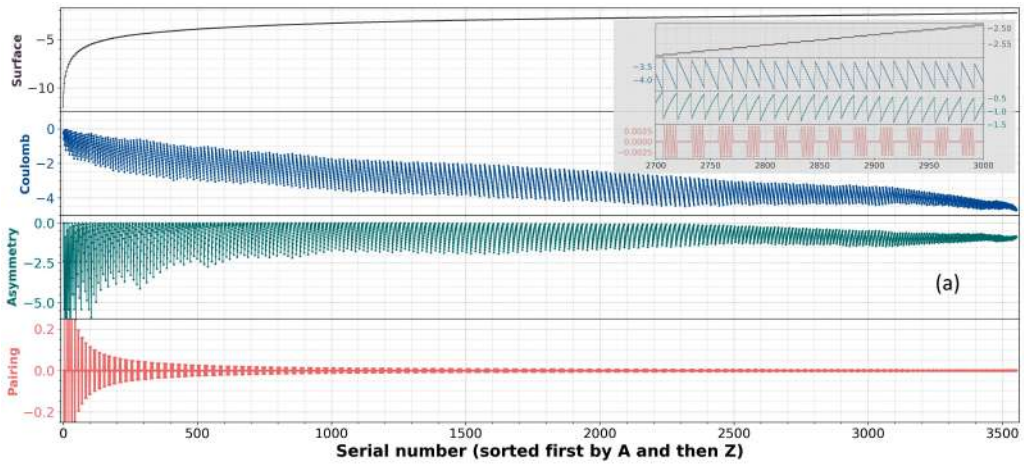

(a)


(b)

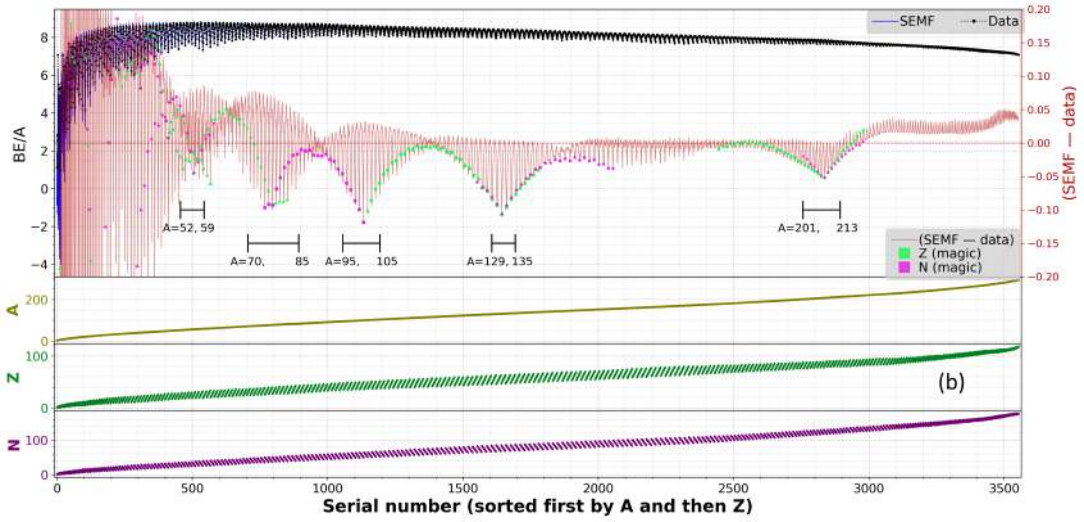



(c)

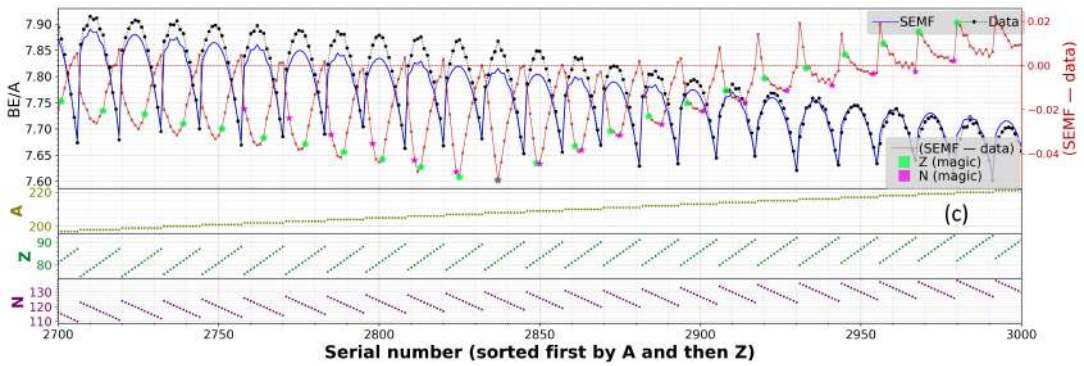
Figure 5: Time-series plots with $(N, Z)$ sorting. Part (a) shows the various terms of the SEMF. Part (b) shows the $\mathrm{BE}/A$ time-series and associated graphs, and part (c) is a zoomed version of part (b) from 2700 to 3000 on the $x$-axis. The detailed legend for magic numbers is given in Fig. 10.




(a)




(b)




(c)

Figure 6: Time-series plots with $(\mathbf{A}, \mathbf{Z})$ sorting. Part (a) shows the various terms of SEMF. Part (b) shows the $\mathrm{BE}/A$ time-series and associated graphs, and part (c) is a zoomed version of part (b), from 2700 to 3000 of the $x$-axis. Detailed legend for magic numbers is given in Fig. 10.

remaining non-aligned/constant for the $(\mathbf{A}, \mathbf{Z})$ sorting. The necessity for shell correction in the SEMF clearly shows up in the LDM disparity curve patterns for all the three time-series (Figs. 4b, 5b & 6b), where we observe significant dips around the nuclear magic numbers $2, 8, 20, 28, 50, 82$, and $126$ [29]. These are much more clearly visible in the zoomed plots (Figs. 4c, 5c & 6c). In the zoomed plot with $(\mathbf{Z}, \mathbf{N})$ sorting as well as the $(\mathbf{N}, \mathbf{Z})$ sorting (Figs. 4c & 5c), we see that the dips in the LDM disparity curve occur at magic numbers, with the deepest dip occurring for a doubly magic nuclide. The parabolic segments of $\mathrm{BE}/A$ plots are clearly not smooth; however, for these two orders of sorting, it appears as if two different curves are joined to form each segment [15], and the magic number appears at the point of joining.

For the $(\mathbf{A}, \mathbf{Z})$ sorting also, the deepest dip is a double magic number and beyond it, the dips in the LDM disparity curve also correspond to magic numbers. But, prior to the double magic number, the dips do not occur at magic numbers. In all three cases, the $\mathrm{BE}/A$ plots and the LDM disparity curve have a mirrored relation with respect to a horizontal line, i.e., the peaks of one occur at the trough of the other (reminiscent of sine and cosine curves). There is a slight systematic shift of this mirrored relationship (the peaks and troughs get misaligned) on one side as we move away from the double magic number.

We now explore one more tool in the time-series arsenal: differencing. We have plotted differencing for the aforementioned three arrangements of $\mathrm{BE}/A$ time-series in Figs. 7, 8 & 9, along with the corresponding proton and neutron separation energies (single and double) as subplots for comparison. The subplot for $\mathrm{BE}/A$ also includes the LDM disparity curve for comparison. A zoomed section of the figure (from serial number $2700$ to $3000$) is also given. Since our focus in this figure is the differencing curve, we have attached the markers for magic numbers to it instead of the LDM disparity curve as was done previously.

Differencing ($\Delta x$) is the simple process of taking the difference between successive terms of a time series, i.e.,

$$\Delta x_i = x_i - x_{i-1}. \tag{6}$$

Differencing with higher lags can also be taken; however, we confine ourselves to the first–order differencing. For the $(Z, N)$ sorting, differencing will give (for each segment)

$$\Delta x_{(Z,N)} = \frac{BE(Z,N)}{A} - \frac{BE(Z,N-1)}{A-1}. \tag{7}$$

At the boundary, $Z$ jumps to the next higher value (increases by 1), while $N$ drops by an uncertain amount, obstructing a precise formula at the boundary. The above equation looks very similar to the formula for neutron separation energy ($S_n$) written in terms of binding energy [30]:

$$S_n(Z,N) = BE(Z,N) - BE(Z,N-1). \tag{8}$$

Similarly, the differencing of $(N, Z)$ time-series resembles proton separation energy ($S_p$) [31]:

$$\Delta x_{(N,Z)} = \frac{BE(N,Z)}{A} - \frac{BE(N,Z-1)}{A-1}, \tag{9}$$

$$S_p(N,Z) = BE(N,Z) - BE(N,Z-1). \tag{10}$$

These similarities can be verified in Figs. 7 and 8.

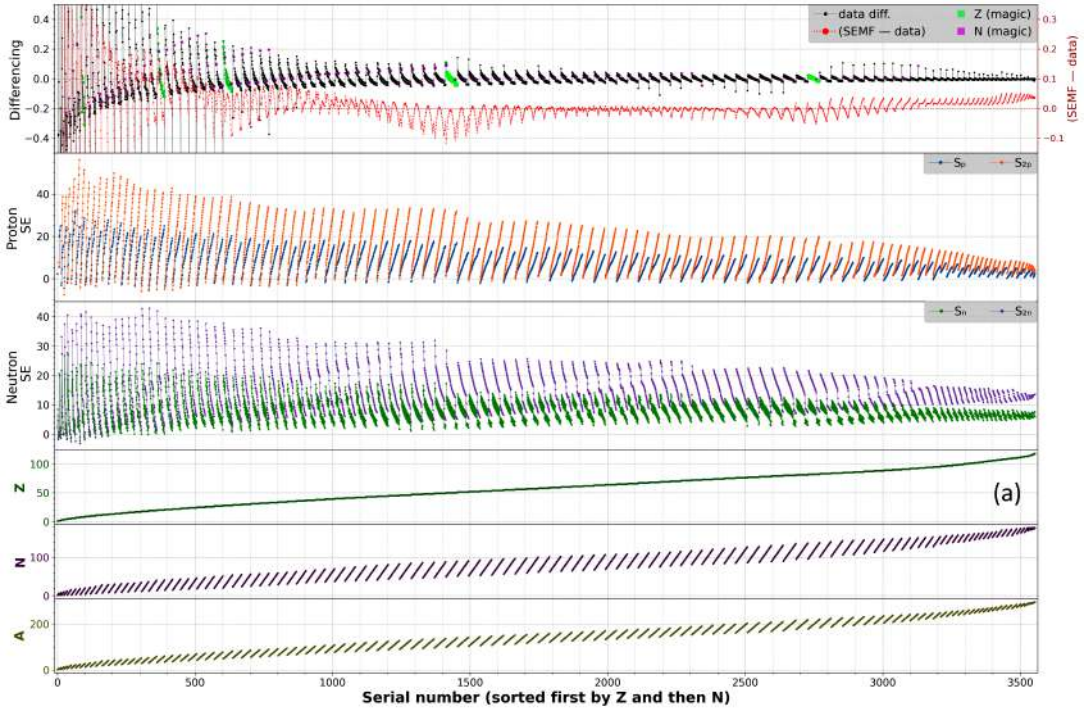
For the differencing of $(A, Z)$ time-series, each segment consists of a constant $A$ and increasing $Z$. Thus, we obtain

$$\Delta x_{(A,Z)} = \frac{1}{A}\left[BE(A,Z) - BE(A,Z-1)\right]. \tag{11}$$

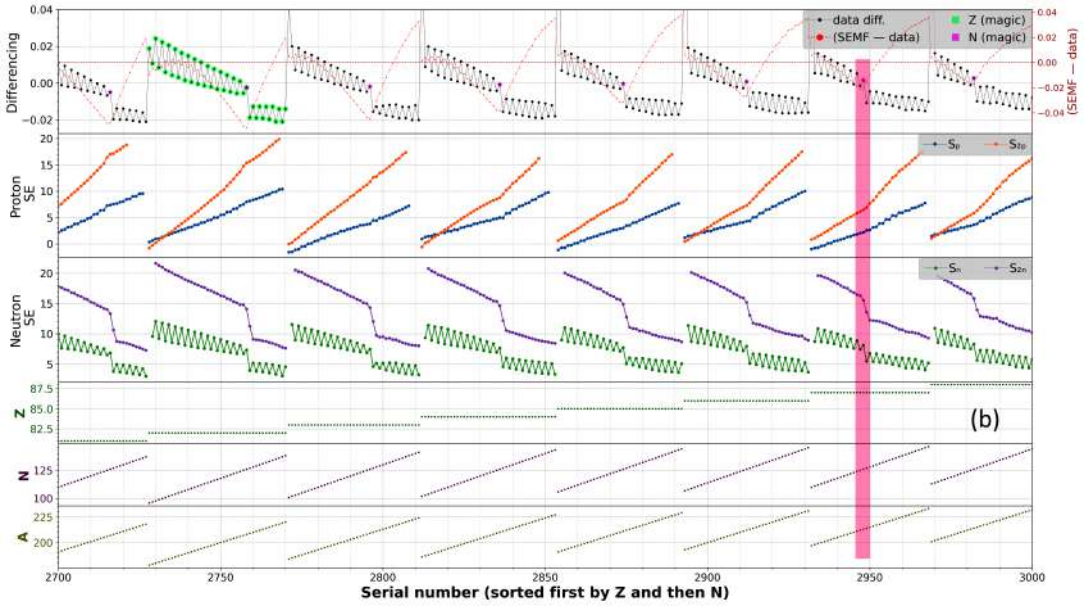
Thus, for a segment with even $A$, both $Z$ and $N$ will be either odd or even, alternating between the two possibilities as they move forward. This causes sharp changes in the binding energy, producing the zigzag behaviour for even $A$ observed in Fig. 9b. For odd $A$, both $Z$ and $N$ will have opposite even–odd character, and therefore a smoother change is observed for the differencing. For the other two time-series, $(\mathbf{Z}, \mathbf{N})$ and $(\mathbf{N}, \mathbf{Z})$, no matter the nature of the first variable (odd or even), the other variable alternates between odd and even when moving forward, thus causing the observed zigzag behaviour for all segments. The $(A, Z)$ differencing has a relation to $\beta$-decay energetics [30]. Clearly, if we had taken only BE instead of $\mathrm{BE}/A$, the connection between differencing and separation energies would have been more direct.

With respect to magic numbers, we find that in $(\mathbf{Z}, \mathbf{N})$ and $(\mathbf{N}, \mathbf{Z})$ time-series (Figs. 7b & 8b), there is a comparatively larger dip in the differencing plot at the position of the magic number that is also mimicked in the

[15] This behaviour is much more prominent for $(\mathbf{Z}, \mathbf{N})$ graphs.

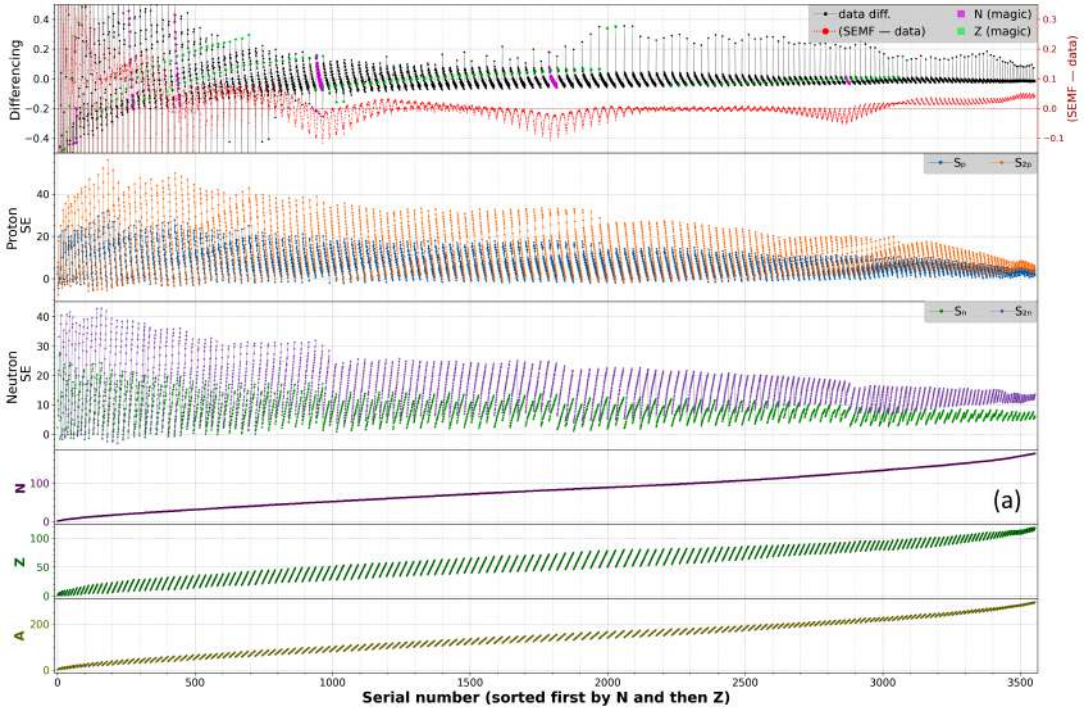



(a)

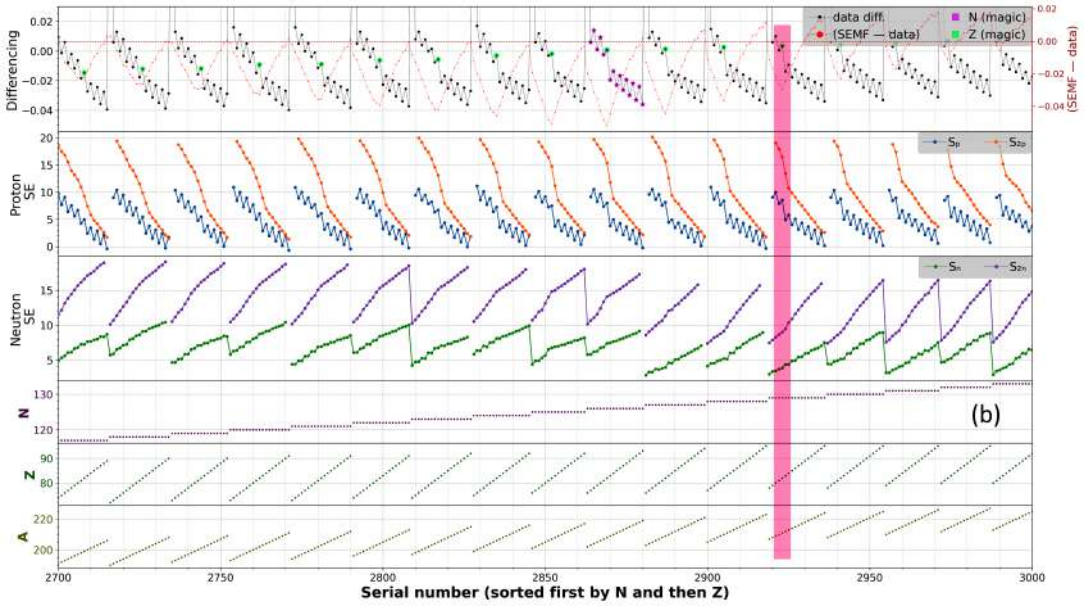



(b)

Figure 7: Time-series plots with (Z, N) sorting. Part (a) shows differencing and LDM disparity curve, along with magic numbers and part (b) is a zoomed version of part (a), from 2700 and 3000 of the x-axis. Detailed legend for magic numbers is given in Fig. 10.

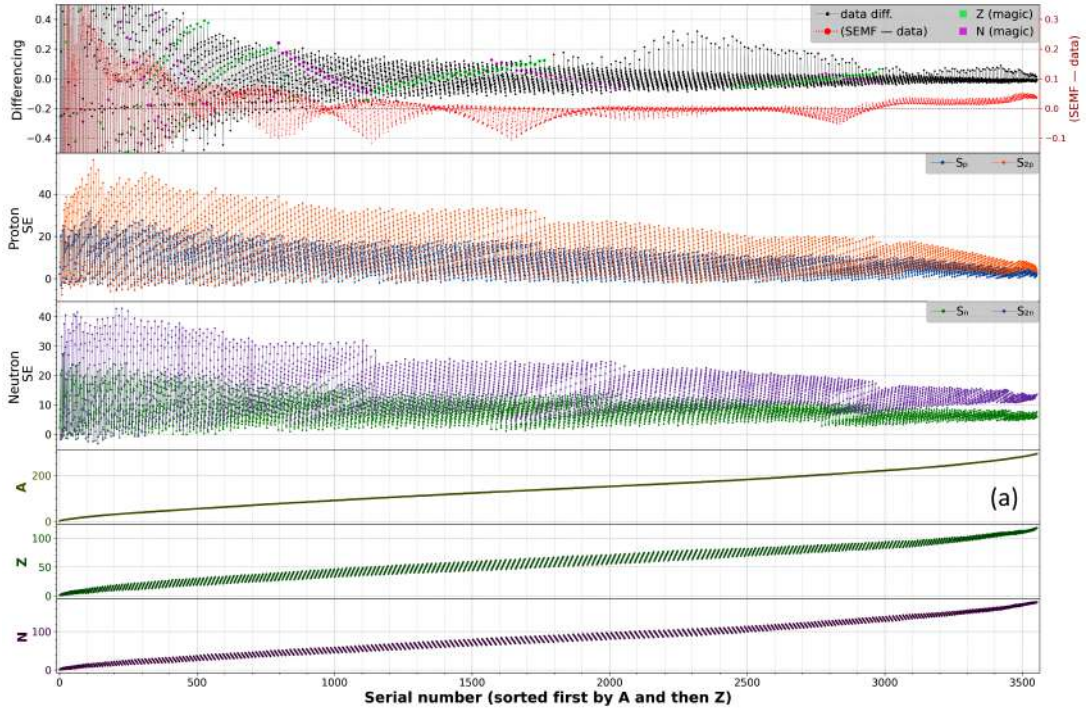



(a)

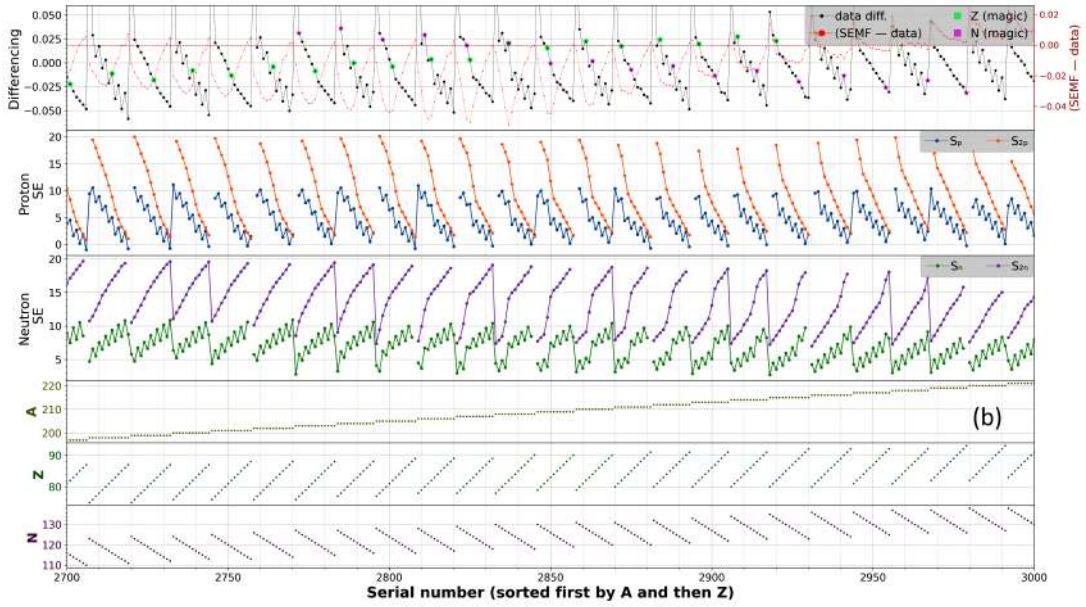



(b)

Figure 8: Time-series plots with (N, Z) sorting. Part (a) shows differencing and LDM disparity curve, along with magic numbers and part (b) is a zoomed version of part (a), from 2700 and 3000 of the x-axis. Detailed legend for magic numbers is given in Fig. 10.




(a)




(b)

Figure 9: Time-series plots with (A, Z) sorting. Part (a) shows differencing and LDM disparity curves, along with magic numbers, and part (b) is a zoomed version of part (a), from 2700 and 3000 on the x-axis. A detailed legend for magic numbers is given in Fig. 10.

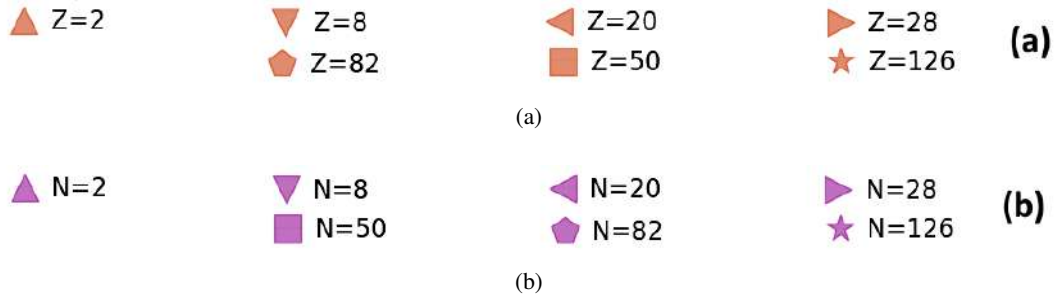



Figure 10: Plot legend for different magic numbers used in Figs. 4-9.

$S_n$ and $S_{2n}$ curves (and the $S_p$ and $S_{2p}$ for $(\mathbf{N}, \mathbf{Z})$); we have highlighted a single example in the corresponding figures for easy identification. Behavior of $(\mathbf{A}, \mathbf{Z})$ in relation to the magic numbers is discernibly without any pattern.

As mentioned previously, there is a correspondence between differencing and nucleon separation energies; we have tried to quantify this using the Pearson correlation. We have also calculated the correlation between differencing and the LDM disparity curve to estimate the linear relationship between them. Although correlation is a cruder tool that compresses the rich interplay in a pair of variables to a single number, it can be useful for summarizing. Table 2 lists the values of the correlation coefficients; all the reported coefficients have extremely low $p$-values indicating the correlations are statistically significant (not surprising given the substantial number of data points).

For each nuktriad variable, we see a pair of positive and a pair of negative correlations with the separation energies. The correlation with the LDM disparity curve (i.e., $(\mathrm{SEMF}-\mathrm{data})$) is found to be positive and relatively high across all three sorting arrangements. This indicates that there exists systematic variation in the data that is not accounted for by the SEMF; otherwise, it would have been nearer zero, implying no correlation. It is tempting to hypothesise that there might be more terms than just the shell corrections that might be contributing to the correlation, but a coherent supporting argument is beyond the expertise of the author; hence, we leave it at that.

| Variable | **Z** | **N** | **A** |
|---|---|---|---|
| $(\mathrm{SEMF}-\mathrm{data})$ | 0.232 | 0.369 | 0.289 |
| $S_p$ | -0.162 | 0.577 | 0.523 |
| $S_{2p}$ | -0.142 | 0.622 | 0.559 |
| $S_n$ | 0.535 | -0.164 | -0.327 |
| $S_{2n}$ | 0.542 | -0.144 | -0.343 |

Table 2: Coefficient values for Pearson Correlation of BE/A time-series differencing with LDM disparity curve and single and double nucleon separation energies.

Besides the features discussed above, these time series have a host of other interesting features, investigation of which may take up more room than traditionally delimited for a journal article; hence, we also leave it for now. To quote an example of the interesting features not enumerated, we observe that for $x$-axis values roughly between 2000 and 2500 (Figs. 4b, 5b & 6b), high-frequency dips in the LDM disparity curve for all three sets of $\mathrm{BE}/A$ time-series do not have any associated magic numbers — this inconsistency may lead to interesting discussions.

The other aspect of the current paper is appreciation of visual approaches to understanding data, which can also be helpful in clarifying the connections to the underlying equations; these kinds of visual and sensory approaches to learning have also been explored by a few authors [32].

# 4 Conclusion

We have explored patterns in $\mathrm{BE}/A$ from the NuDat3/AME2020 dataset. On analysing the standard deviations of isotopic, isotonic and isobaric sequences, we found that variation within isotopic sequence is largest compared to the others, while isobaric sequence has the least variation. Time-series transformation of the $\mathrm{BE}/A$ data revealed interesting pattern and connection with other variables; some of which were identified and discussed in the light

of the semi-empirical mass formula. We found that differencing of the time-series can be related to proton/neutron separation energies; computed correlations indicate paired responses. Besides exploration, these transformation and visualization tools can be useful in pedagogy.

## References


[1] M. Wang, W. J. Huang, F. G. Kondev, G. Audi, and S. Naimi, Chinese Physics C **45**, 030003 (2021).

[2] National Nuclear Data Center (NNDC), Nudat 3.0, `https://www.nndc.bnl.gov/nudat3/`.

[3] T. H. Levere, *Transforming Matter: A History of Chemistry from Alchemy to the Buckyball* (Johns Hopkins University Press, 2003).

[4] I. Kaplan, *Nuclear Physics* (Narosa Publishing House, New Delhi, 2020).

[5] F. W. Aston, Proceedings of the Royal Society of London. Series A **115**, 487 (1927).

[6] F. W. Aston, Nature **120**, 956 (1927).

[7] W. D. Harkins and T. H. Liggett, Journal of Physical Chemistry **28**, 74 (1924).

[8] J. Chadwick, Proceedings of the Royal Society of London. Series A **136**, 692 (1932).

[9] H. A. Bethe and R. F. Bacher, Reviews of Modern Physics **8**, 82 (1936).

[10] C. F. von Weizsäcker, Zeitschrift für Physik **96**, 431 (1935).

[11] B. R. Martin and G. Shaw, *Nuclear and Particle Physics: An Introduction*, 3rd ed. (Wiley, 2019).

[12] P. J. Mohr, D. B. Newell, B. N. Taylor, and E. Tiesinga, Reviews of Modern Physics **97**, 025002 (2025).

[13] D. Benzaid, S. Bentridi, A. Kerraci, and N. Amrani, Nuclear Science and Techniques **31**, 9 (2020).

[14] B. Pandey *et al.*, AIP Advances **14**, 105228 (2024).

[15] D. Cha and B. Kim, Journal of the Korean Physical Society **56**, 1546 (2010).

[16] S. Gora, O. S. K. S. Sastri, and S. K. Soni, European Journal of Physics **43**, 035802 (2022).

[17] A. Saad, Z. I. Elghobary, and M. Taha, Bulletin of the Faculty of Science, Zagazig University , 66 (2025).

[18] S. Lakshminarayana, Preprint (2025).

[19] T. W. Anderson, *The Statistical Analysis of Time Series* (John Wiley & Sons, 2011).

[20] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications* (Springer, 2017).

[21] P. Song and C. T. Russell, Space Science Reviews **87**, 387 (1999).

[22] S. Li and A. Adelmann, Physical Review Accelerators and Beams **26**, 024801 (2023).

[23] W. J. Huang, M. Wang, F. G. Kondev, G. Audi, and S. Naimi, Chinese Physics C **45**, 030002 (2021).

[24] W. McKinney, in *Proceedings of the 9$^{th}$ Python in Science Conference* (2010) pp. 56–61.

[25] The pandas development team, pandas-dev/pandas: Pandas 2.2.3 (2025).

[26] P. Virtanen *et al.*, Nature Methods **17**, 261 (2020).

[27] J. D. Hunter, Computing in Science & Engineering **9**, 90 (2007).

[28] A. Holmes, S. Dean, and A. H. Barbar, *Introductory Business Statistics 2e* (OpenStax, 2023).

[29] D. Steppenbeck *et al.*, Nature **502**, 207 (2013).

[30] J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (Springer, 1979).

[31] S. Athanassopoulos, E. Mavrommatis, K. A. Gernoth, and J. W. Clark, Preprint (2005).

[32] M. Hein, A. Pusch, and S. Heusler, European Journal of Physics **43**, 035801 (2022).