MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **95** (2026) 621–639
ISSN: 0340–6253
doi: 10.46793/match.95-3.11325

Predictive Numerical Modeling of Zinc Sulfide Nanocluster Properties

Iman Khanipour, Mohammad B. Ahmadi*

Mathematics Department, Faculty of Science, Shiraz University

ac.ikhanipour@gmail.com, mbahmadi@shirazu.ac.ir

(Received May 14, 2025)

Abstract

We present numerical models for predicting five key electronic properties of zinc sulfide (ZnS) nanoclusters: binding energy per ZnS unit, static dipole polarizability, HOMO-LUMO gap, vertical ionization potential, and chemical hardness. These models employ four graph-based topological indices (Wiener, hyper-Wiener, second-order connectivity, and Szeged) alongside cluster size n. The models offer substantially reduced estimation errors and strong predictive performance. The results demonstrate that numerical models provide efficient routes for estimating ZnS nanocluster properties, facilitating rapid screening and guiding experimental efforts.

1 Introduction

Atomic clusters (which are finite assemblies of atoms) exhibit structural and electronic properties that occupy a unique regime between discrete molecules and extended solids. They often adopt configurations that have no direct analogue in bulk materials; for instance, covalent clusters can form hollow cage geometries such as the C_{60} fullerene, and ionic clusters may only partially mimic the bulk crystal lattice depending on their size

*Corresponding author.

and stoichiometry. These size-dependent structural motifs give rise to novel physical and chemical behaviors, positioning clusters as both subjects of fundamental study and candidates for technological applications such as catalysis, sensing, and nanoelectronics [26].

A key hurdle in cluster science is elucidating how their physicochemical properties change as a function of size. As Jena and Castleman point out, clusters pose fundamental questions such as "at what size does a metal exhibit metallic behavior?"–or more broadly, "when do clusters begin to reproduce bulk characteristics?" [26]. Empirically, properties like stability, electronic gap, and reactivity do not scale monotonically with size or composition but can display non-intuitive trends. Although cluster attributes must eventually approach their bulk limits, the transition is complex: finite clusters often exhibit properties that, while related to bulk values, remain distinctly different, and there is no straightforward theoretical framework to predict bulk parameters directly from small-cluster data [1]. Consequently, constructing reliable predictive models that capture these size-dependent behaviors is essential.

Zinc sulfide (ZnS) nanoclusters vividly illustrate the challenges of size-dependent property prediction. As a wide-bandgap $II-VI$ semiconductor (bulk band gap $\approx 3.6 - 3.7\,eV$), ZnS underpins numerous optoelectronic technologies, including UV-emitting diodes and photonic devices. At the nanoscale, ZnS clusters have found applications in electroluminescent displays, solar energy harvesting, and photocatalysis, while their high surface area and tunable band structure also make them attractive for chemical and biosensing platforms. The ability to tailor ZnS's electronic and optical responses through cluster size (a result of quantum confinement and surface effects) lies at the heart of these applications and compels the development of theoretical frameworks capable of accurately predicting cluster properties across varied sizes [27].

Graph theory provides a rigorous framework for abstracting molecular and cluster architectures into mathematical objects. In this approach, each atom corresponds to a vertex and each chemical bond to an edge, yielding a connected graph that captures the essence of the structure without requiring precise spatial coordinates. Originating with Euler's work on

polyhedra, this representation has become indispensable in chemistry and materials science for its ability to translate complex bonding patterns into computable quantities. Among these are topological indices (graph invariants such as the Wiener and hyper-Wiener indices) that distill the entire network geometry into single numerical values. These indices have long been employed in QSPR and QSAR studies to link molecular topology with physicochemical behaviors, enabling efficient prediction and ranking of properties ranging from reactivity to biological activity [27, 28].

Topological indices are particularly attractive because they condense complex structural information into single numerical values that often correlate strongly with macroscopic properties. For instance, the Wiener index and its distance-based generalizations have historically been applied to predict boiling points, thermodynamic stabilities, and other bulk characteristics of organic compounds. In the nanocluster domain, these indices naturally encode both size and connectivity, making them promising descriptors for cluster energetics and electronic behavior [27]. Building on this concept, in [1] it was demonstrated that four distance-based indices (specifically, the Wiener, hyper-Wiener, second-order connectivity, and Szeged measures) could reliably model a variety of electronic properties of ZnS clusters, thereby validating the extension of graph-theoretic tools to nanoscale systems.

In this work, we propose more accurate models for predicting ZnS cluster properties based on a graph-theoretic approach and, alternatively, on cluster size n alone. The paper is organized as follows. In Section 2, we introduce the four selected topological indices (the Wiener index (W), hyper-Wiener index (WW), second-order connectivity index (^2X), and Szeged index (Sz)), together with the five target properties (binding energy per ZnS unit, mean static dipole polarizability, HOMO-LUMO gap, vertical ionization potential, and chemical hardness). Section 3 develops and evaluates several numerical modeling approaches that use these indices and the cluster size n to predict ZnS nanocluster properties. We compare each model's performance against the original power-law fits of Mohajeri et al. [1] to identify the most accurate and robust predictors.

2 Topological descriptors and cluster properties

Graph-based topological indices have long been employed as molecular descriptors that link structural features to physicochemical and biological properties. To date, over a thousand distinct topological indices have been introduced in the literature [3–9]. In this study, however, we restrict our focus to four key indices (Wiener, hyper-Wiener, second-order connectivity, and Szeged) for the purpose of property estimation. Throughout this section, G denotes a simple, undirected molecular graph.

The Wiener index (W) is a foundational topological descriptor introduced by Harry Wiener in 1947. It quantifies molecular structure by summing the shortest-path distances between all pairs of vertices in a molecular graph. Formally, it is expressed as

$$W(G) = \sum_{\{u,v\} \subseteq V(G)} d(u, v),$$

where $V(G)$ represents the set of its vertices, and $d(u, v)$ is the shortest-path distance between vertices u and v. A higher Wiener index typically indicates a molecule with a more extended or branched structure [10, 11].

Milan Randić extended the concept of the Wiener index by introducing the hyper-Wiener index (WW), which incorporates not only the shortest-path distances but also their squares, thereby emphasizing the influence of long-range vertex separations within a graph [12, 13]. This enhanced sensitivity to structural variation is captured mathematically as

$$\mathrm{WW}(G) = \frac{1}{2} \sum_{\{u,v\} \subseteq V(G)} (d(u, v) + (d(u, v))^2).$$

The second-order connectivity index, denoted as ${}^2X$, quantifies the branching of G by summing over all induced P_3 subgraphs (also called 2-paths). An induced P_3 subgraph on three distinct vertices $\{u, v, w\} \subseteq$

$V(G)$ is one in which uv and vw are edges in G, and $u \neq w$. Formally,

$$^2X(G) = \sum_{\substack{\{u,v,w\}\subseteq V(G)\\ uv,\, vw\in E(G)}} \frac{1}{\sqrt{d(u)d(v)d(w)}},$$

where $d(u)$ represents the degree of vertex u in G and $E(G)$ is the edge set of the graph. This topological descriptor is a specific instance of the general m-connectivity index [14–16].

The Szeged index (Sz) extends the concept of the Wiener index by accounting for how vertices are distributed with respect to each edge in the graph. For each edge $e = uv$, the values $n_u(e)$ and $n_v(e)$ represent the number of vertices that are closer to vertex u than to v, and vice versa [11, 17, 18]. The index is given by

$$\mathrm{Sz}(G) = \sum_{e=uv\in E(G)} n_u(e) \cdot n_v(e).$$

In [1], these four topological indices were calculated for each $(ZnS)_n$ nanocluster graph. As structure-based descriptors, they encapsulate essential information about molecular geometry and connectivity while bypassing the need for explicit atomic-level details. This abstraction makes them particularly effective for predicting various physicochemical properties of nanoclusters.

The average binding energy per ZnS unit, $BE(n)$, provides a measure of cluster stability and is defined as

$$BE(n) = \frac{nE(ZnS) - E((ZnS)_n)}{n},$$

where $E((ZnS)_n)$ is the total electronic energy of the n-unit cluster and $E(ZnS)$ is the energy of a single ZnS monomer. In practice, both quantities are obtained from electronic-structure calculations, such as density functional theory. Physically, a larger $BE(n)$ indicates that each ZnS pair is more strongly bound within the cluster. Chemically, $BE(n)$ is analogous to the cohesive energy per formula unit, representing the energy required to dissociate the cluster into isolated monomers. As such, $BE(n)$ serves

as a direct indicator of the cluster's resistance to fragmentation [1, 25].

The static dipole polarizability $< \alpha >$ quantifies the extent to which a cluster's electron cloud deforms in response to a uniform external electric field. It is given by the average of the diagonal elements of the polarizability tensor:

$$< \alpha >= \frac{1}{3}(\alpha_{xx} + \alpha_{yy} + \alpha_{zz}),$$

where α_{ii} are the tensor's principal components. In practice, $< \alpha >$ is computed using finite-field or linear-response techniques within electronic-structure methods. Physically, a larger value of $< \alpha >$ signifies a more easily deformable electron density, which has implications for intermolecular interactions, optical properties, and chemical reactivity [19].

The HOMO-LUMO gap (Δ) is defined as the energy difference between the highest occupied and lowest unoccupied molecular orbitals of the neutral cluster, evaluated at its ground-state geometry. In Kohn-Sham density functional theory, Δ is most often approximated by the difference in Kohn-Sham orbital energies $\varepsilon_{LUMO} - \varepsilon_{HOMO}$. Alternatively, it can be obtained from total-energy differences via $\Delta = E(N+1) + E(N-1) - 2E(N)$, although this is less common in practice. The magnitude of Δ serves as a primary indicator of electronic behavior: a large gap is characteristic of insulating, chemically inert systems, whereas a small gap signals metallic or highly reactive character. Physically, Δ corresponds to the onset of optical absorption and thus influences the color and photophysical properties of the cluster, and it also enters directly into the definition of chemical hardness [20, 21, 23, 24].

The vertical ionization potential (VIP or simply IP) is the energy required to remove an electron from the neutral cluster without allowing its geometry to relax. It is defined as

$$VIP = E^{+} - E^{0},$$

where E^{+} and E^{0} are the total energies of the cationic and neutral species, respectively, both evaluated at the geometry of the neutral cluster. VIP quantifies how strongly electrons are bound within the cluster–a larger VIP implies more tightly held electrons. It can also be interpreted as

a measure of the cluster's tendency to lose a valence electron, and thus reflects its electronic stability [1, 21].

In density-functional theory, the chemical hardness η is rigorously defined as half the second derivative of the ground-state energy E with respect to electron number N at constant geometry. In practical applications, η is approximated by the finite-difference expression

$$\eta \approx \frac{1}{2}(VIP - EA),$$

where VIP and EA denote the vertical ionization potential and electron affinity, respectively. Chemical hardness measures the resistance of a system's electron density to deformation: a **hard** species (large η) is reluctant to undergo charge transfer and tends to be chemically inert, whereas a **soft** species (small η) is more polarizable and reactive. This concept forms the basis of the hard-soft acid-base (HSAB) principle [20–22].

3 Predictive numerical models and interpretation of the results

In [1], the authors calculated four distance-based topological indices for $(ZnS)_n$ clusters ($n = 3, \ldots, 10$) and computed five key electronic properties using DFT at the B3LYP/6-31+G(d) level. To extend their structural analysis to larger clusters ($n = 11, \ldots, 15$), they invoked Euler's theorem (linking the number of faces, edges, and vertices in any simple closed polyhedron) to construct plausible cage graphs and calculate the corresponding indices.

The full dataset of topological descriptors and electronic properties appears in Table 1. For vertical ionization potentials and HOMO-LUMO gaps covering $n = 3, \ldots, 15$, we adopt the results of Burnin et al. [2], summarized in Table 1c.

Mohajeri et al. [1] established power-law regressions of the form

$$y = a\,TI^b, \tag{1}$$

Table 1. Topological indices and electronic properties for $(ZnS)_n$

(a) Topological indices for $(ZnS)_n$ clusters with $n = 3, \ldots, 15$

n	W	WW	Sz	^2X
3	27	42	30	2.121
4	64	120	144	2.828
5	125	275	320	3.535
6	144	158	800	6.928
7	215	380	1384	8.083
8	304	608	1872	9.238
9	411	858	3318	10.392
10	536	1151	4758	11.547
11	685	1529	5814	12.702
12	864	2526	8712	13.856
13	1050	2033	10594	15.011
14	1268	3132	13486	16.166
15	1515	3879	16328	17.321

(b) Electronic properties for $(ZnS)_n$ clusters with $n = 3, \ldots, 10$

n	$<\alpha>$ (Å^3)	η(eV)	$BE(n)$(eV)
3	19.733	3.750	2.860
4	26.797	3.691	3.056
5	34.441	3.651	3.084
6	36.563	3.337	3.187
7	43.258	3.281	3.191
8	49.061	3.315	3.340
9	54.636	3.255	3.403
10	61.285	3.175	3.410

(c) Electronic properties for $(ZnS)_n$ clusters with $n = 3, \ldots, 15$

n	$VIP\,(eV)$	$\Delta\,(eV)$
3	8.86	4.19
4	8.53	4.49
5	8.57	4.59
6	8.45	3.88
7	8.46	3.82
8	8.40	4.06
9	8.38	4.16
10	8.26	4.04
11	8.29	4.17
12	8.47	4.52
13	8.25	4.19
14	8.27	4.23
15	8.29	4.41

Table 2. Coefficients a and exponents b for each property-index pair with $n = 3, \ldots, 10$

TI	W		WW		Sz		^2X	
	a	b	a	b	a	b	a	b
$<\alpha>^{1/3}$	1.7711	0.127	1.7893	0.112	2.4260	0.187	2.0947	0.073
$BE(n)$	2.3554	0.060	2.3732	0.052	2.7254	0.089	2.5448	0.034
η/n	6.3397	-0.476	2.4566	-0.173	2.0386	-0.728	3.4559	-0.278
VIP^{-1}/n	0.1442	-0.394	0.0554	-0.170	0.0557	-0.597	0.0866	-0.229
Δ^{-1}/n	0.2829	-0.385	0.2786	-0.342	0.1071	-0.560	0.1686	-0.221

linking various cluster properties (specifically, the cube root of the mean polarizability ($<\alpha>^{1/3}$), the binding energy per ZnS unit ($BE(n)$), the chemical hardness per unit (η/n), the inverse vertical ionization potential per unit (VIP^{-1}/n), and the inverse HOMO-LUMO gap per unit (Δ^{-1}/n)) to each topological index (TI). The coefficients a and exponents b for each property-index pair are listed in Table 2.

By applying a log-space transformation, they demonstrated strong correlations between the topological indices and cluster properties, and their power-law regressions provided good fits for $(ZnS)_n$ clusters. However, we can introduce alternative modeling approaches that achieve higher accuracy and reliably predict trends as cluster size increases.

In this section, we assess several alternative models (developed through numerical methods) to identify functional relationships that more reliably predict cluster properties from both topological indices and cluster size n.

Our analysis is based on the data compiled in Table 1, represented as the set $\{(TI_i, y_i)\}_{i \in I}$. Here, TI_i denotes the i^{th} topological index (selected from the Wiener, hyper-Wiener, second-order connectivity, or Szeged descriptors) or, in some cases, the cluster size n. The corresponding y_i represents the i^{th} property, chosen from $< \alpha >^{1/3}$, $BE(n)$, η/n, VIP^{-1}/n, or Δ^{-1}/n.

We construct models using the dataset and evaluate their performance based on Mean Absolute Error (MAE), in order to determine which model most accurately estimates and predicts the properties of $(ZnS)_n$ nanoclusters from their topological indices or from the cluster size n.

3.1 Prediction by extrapolation

In this framework, we begin by fitting extrapolation-based models to the subset indexed by I_{extra}, in order to derive candidate predictive functions. Let I_{extra} denote the index set of cluster sizes used for extrapolation, defined separately for each group of target properties as follows:

1. For $< \alpha >^{1/3}$, η/n, and $BE(n)$, we set

$$I_{extra} = \{3, \ldots, 9\}.$$

2. For VIP^{-1}/n and Δ^{-1}/n, we set

$$I_{extra} = \{3, \ldots, 10\}.$$

Each model is then evaluated on the full dataset using Mean Absolute Error (MAE) to assess its accuracy relative to the power-law regressions proposed by [1]. Only the top-performing extrapolation methods (those yielding the lowest error values) are presented in detail. Additionally, we model each property directly as a function of n, enabling trend estimation without the added complexity of computing topological indices. We also describe the mathematical foundations of these extrapolation techniques and provide illustrative implementations using Python's SciPy library.

3.1.1 Piecewise linear and piecewise exponential

A piecewise linear model approximates the target function by connecting each adjacent pair of points with a straight segment and extending the end segments beyond the data range. Given ordered topological indices $\{(TI_i, y_i)\}_{i=1}^{N}$, where $TI_1 < \cdots < TI_N$.

For $i = 1, \ldots, N-1$, we define

$$f(TI) = \begin{cases} y_1 + \dfrac{y_2 - y_1}{TI_2 - TI_1}(TI - TI_1), & \text{if } TI \le TI_1, \\ y_i + \dfrac{y_{i+1} - y_i}{TI_{i+1} - TI_i}(TI - TI_i), & \text{if } TI \in [TI_i, TI_{i+1}], \\ y_N + \dfrac{y_N - y_{N-1}}{TI_N - TI_{N-1}}(TI - TI_N), & \text{if } TI \ge TI_N. \end{cases}$$

A piecewise exponential model fits an exponential curve to each adjacent pair of data points and extends those curves beyond the endpoints for extrapolation. Specifically, for each interval $[TI_i, TI_{i+1}]$ with corresponding values (TI_i, y_i) and (TI_{i+1}, y_{i+1}), we define

$$f(TI) = y_i \exp(\lambda_i (TI - TI_i)), \quad TI \in [TI_i, TI_{i+1}],$$

and enforce the boundary conditions

$$f(TI_i) = y_i, \quad f(TI_{i+1}) = y_{i+1}.$$

Solving these yields the local rate parameter

$$\lambda_i = \frac{1}{TI_{i+1} - TI_i} \ln\left(\frac{y_{i+1}}{y_i}\right),$$

which requires $y_i > 0$ and $y_{i+1} > 0$. For extrapolation, the same formula is applied to the first and last segments, giving the full definition:

$$f(TI) = \begin{cases} y_1 \exp(\lambda_1 (TI - TI_1)), & TI \le TI_1, \\ y_i \exp(\lambda_i (TI - TI_i)), & TI \in [TI_i, TI_{i+1}], \\ y_N \exp(\lambda_{N-1} (TI - TI_N)), & TI \ge TI_N, \end{cases}$$

for $i = 1, \ldots, N-1$. Here, each λ_i governs the exponential growth or

decay rate on its interval: $\lambda_i > 0$ produces an increasing curve, while $\lambda_i < 0$ yields a decreasing curve, ensuring both continuity and a locally exponential behavior.

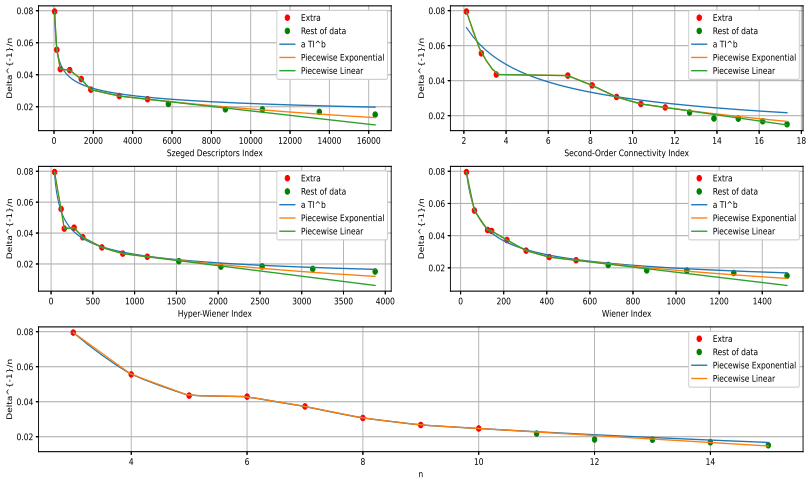

Figure 1. Comparison of piecewise linear and piecewise exponential extrapolations against power-law model for Δ^{-1}/n. The five panels display results for the four topological indices and the direct n-based trend.

We apply both piecewise linear and piecewise exponential extrapolations to the data specified by I_{extra}, using interp1d from SciPy's interpolate module. The resulting MAE values, together with those from the model proposed by [1] (with parameters a and b listed in Table 2), are presented in Table 3, which also includes the MAE for the direct size-based model. These results demonstrate that the piecewise approaches offer strong alternatives to the power-law regressions and are capable of estimating properties using only the cluster size n. As an example, Figure 1 illustrates the extrapolated trends in comparison with the regression predictions for Δ^{-1}/n.

3.1.2 Radial basis function

Radial Basis Function (RBF) extrapolation is a powerful numerical method used to approximate multivariate functions based on scattered data. Given

Table 3. Computed MAE for the piecewise linear model (E_{PL}), piecewise exponential model (E_{PE}), and power-law model (E_M)

TI	W			WW			Sz			^2X			n	
	E_M	E_{PL}	E_{PE}	E_M	E_{PL}	E_{PE}	E_M	E_{PL}	E_{PE}	E_M	E_{PL}	E_{PE}	E_{PL}	E_{PE}
VIP^{-1}/n	0.0005	0.0008	0.0003	0.0047	0.0011	0.0005	0.0009	0.0008	0.0004	0.0015	0.0002	0.0001	0.0002	0.0001
Δ^{-1}/n	0.0011	0.0010	0.0004	0.0016	0.0014	0.0006	0.0027	0.0011	0.0005	0.0051	0.0003	0.0006	0.0003	0.0006
$<\alpha>^{1/3}$	0.0092	0.0010	0.0018	0.0512	0.0011	0.0019	0.0409	0.0019	0.0013	0.0899	0.0018	0.0012	0.0018	0.0012
$BE(n)$	0.0310	0.0083	0.0085	0.0427	0.0084	0.0085	0.0309	0.0070	0.0071	0.0479	0.0070	0.0072	0.0070	0.0071
η/n	0.0347	0.0020	0.0010	0.3088	0.0021	0.0010	0.0286	0.0009	0.0001	0.0392	0.0009	0.0001	0.0009	0.0001

data points $\{(TI_i, y_i)\}_{i=1}^N$, the RBF model takes the form

$$f(TI) = \sum_{i=1}^{N} w_i \Phi(r), \quad r = \mid TI - TI_i \mid,$$

where $w_i \in \mathbb{R}$ are weights to be determined, $\Phi(r)$ is a radial basis function depending only on the distance r. In this study, we focus on two common kernels:

1. Multiquadric
$$\Phi(r) = \sqrt{1 + (\varepsilon r)^2},$$

2. Inverse Multiquadric
$$\Phi(r) = \frac{1}{\sqrt{1 + (\varepsilon r)^2}},$$

where $\varepsilon > 0$ is a shape parameter that can be tuned for optimal predictive performance.

We require that

$$f(TI_i) = y_i, \quad i = 1, \ldots, N,$$

which yields the linear system

$$\sum_{i=1}^{N} w_i \Phi(\mid TI_j - TI_i \mid) = y_j, \quad j = 1, \ldots, N.$$

In matrix form, this becomes $Qw = y$, where $Q_{ji} = \Phi(\mid TI_j - TI_i \mid)$, $w = (w_1, \ldots, w_N)^T$, and $y = (y_1, \ldots, y_N)^T$. Solving the linear system $Qw = y$ yields the weights w. Once the weights w_i are computed, the

function $f(TI)$ can be evaluated for any TI (inside or outside the range), thereby enabling extrapolation beyond the original data domain.

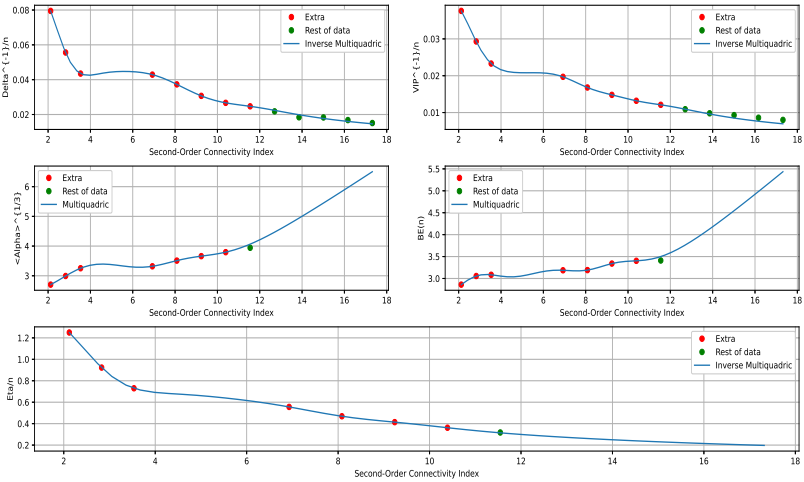

Figure 2. RBF models (multiquadric or inverse multiquadric) for predicting properties using the second-order connectivity index

We implement RBF extrapolation on the data specified by I_{extra} using SciPy's RBF class. Selecting an appropriate shape parameter ε is crucial for performance. For instance, setting $\varepsilon = 2.3$ when extrapolating properties based on the second-order connectivity index yields an excellent fit, as illustrated in Figure 2.

Moreover, our results show that RBF extrapolation with the same $\varepsilon = 2.3$, applied directly to cluster size n, also performs effectively (Figure 3). The corresponding MAE values for the n-based RBF model are provided in Table 4.

3.2 Prediction by curve fitting

The power-law model proposed by [1] generally performs good; however, in some cases, the provided coefficient a and exponent b lead to increased error and reduced accuracy. To address this issue, we re-estimated the parameters a and b using a curve fitting approach.

They used data for $n = 3, \ldots, 10$ in their model. We applied the curve

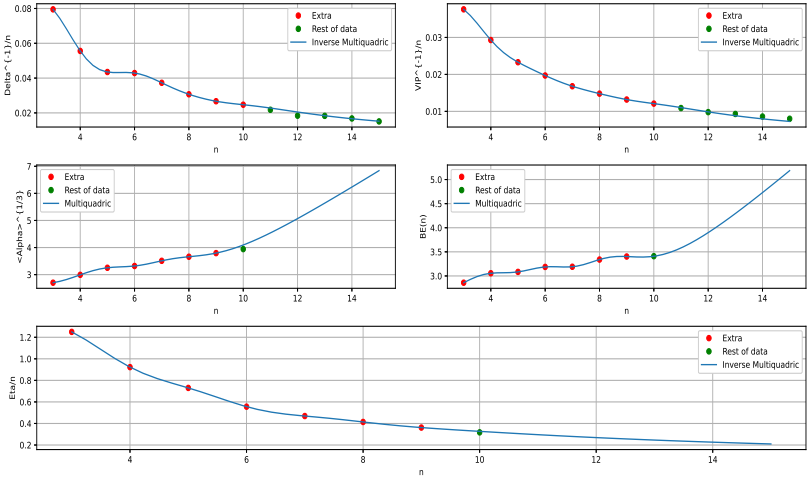

Figure 3. RBF models (multiquadric or inverse multiquadric) for predicting properties based on cluster size n

Table 4. Computed MAE for the RBF multiquadratic model (E_{RBFM}) and the RBF inverse multiquadratic model (E_{RBFIM})

TI	n	
	E_{RBFM}	E_{RBFIM}
$<\alpha>^{1/3}$	0.0188	–
$BE(n)$	0.0002	–
η/n	–	0.0012
VIP^{-1}/n	–	0.0002
Δ^{-1}/n	–	0.0003

fitting approach over the same range and observed improved accuracy. Motivated by this, we aimed to develop a more general predictive model and therefore performed curve fitting on the entire dataset, since the original model shows its performance in prediction. We also found that a power-law relationship holds for cluster size n, so we estimated the corresponding parameters a and b.

Given data $\{(TI_i, y_i)\}_{i=1}^{N}$, the power-law model can be linearized by

taking natural logarithms:

$$\ln(y_i) = \ln(a) + b\ln(TI_i) + \varepsilon_i,$$

where ε_i denotes the residual error. Let

$$u_i = \ln(TI_i), \quad v_i = \ln(y_i), \quad \alpha = \ln(a).$$

We then solve the least squares problem

$$\min_{\alpha,b} \sum_{i=1}^{N} (v_i - (\alpha + bu_i))^2.$$

The optimal parameters have closed-form solutions:

$$b = \frac{\sum_{i=1}^{N}(u_i - \bar{u})(v_i - \bar{v})}{\sum_{i=1}^{N}(u_i - \bar{u})^2}, \quad \alpha = \bar{v} - b\bar{u},$$

where $\bar{u}$ and $\bar{v}$ are the sample means of $\{u_i\}$ and $\{v_i\}$, respectively, and $a = \exp(\alpha)$.

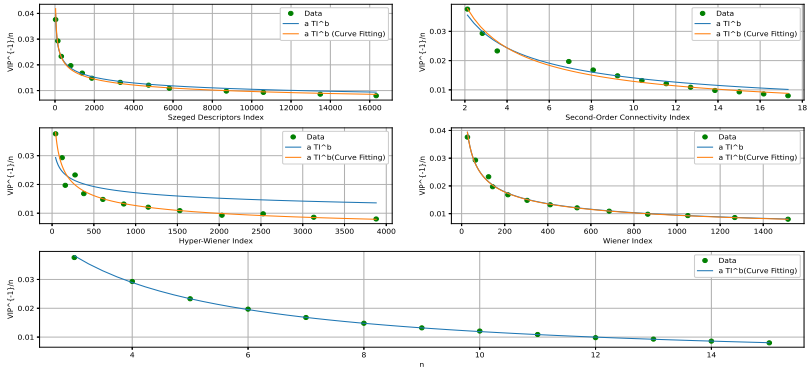

Figure 4. Comparison of curve fitting model against power-law model for VIP^{-1}/n. The five panels display results for the four topological indices and the direct n-based trend.

The estimated parameters for both the topological indices and the cluster size n are provided in Table 5, and their corresponding MAE values, compared to the original model, are listed in Table 6. As an illustrative example, Figure 4 compares the two power-law models for VIP^{-1}/n, highlighting the improvement achieved by the curve-fitting approach.

Table 5. Coefficients a and exponents b of curve-fitted power-law model for $(ZnS)_n$ clusters with $n = 3, \ldots, 15$

TI	W		WW		Sz		^2X		n	
	a	b	a	b	a	b	a	b	a	b
$<\alpha>^{1/3}$	1.7711	0.1269	1.7763	0.1131	2.0946	0.0731	2.4262	0.1872	1.9560	0.3026
$BE(n)$	2.3532	0.0597	2.3426	0.0543	2.5428	0.0346	2.7237	0.0891	2.4628	0.1430
η/n	6.3374	-0.4757	6.4315	-0.4288	3.4548	-0.2776	2.0379	-0.7279	4.4794	-1.1489
VIP^{-1}/n	0.1455	-0.3965	0.1402	-0.3482	0.0989	-0.2527	0.0636	-0.6928	0.1106	-0.9677
Δ^{-1}/n	0.3203	-0.4127	0.3071	-0.3618	0.2105	-0.2607	0.1306	-0.7043	0.2380	-1.0012

Table 6. Computed MAE for power-law model (E_M) and the curve-fitted power-law model (E_{CF})

TI	W		WW		Sz		^2X		n
	E_M	E_{CF}	E_M	E_{CF}	E_M	E_{CF}	E_M	E_{CF}	E_{CF}
VIP^{-1}/n	0.0005	0.0005	0.0047	0.0005	0.0009	0.0008	0.0015	0.0011	0.0001
Δ^{-1}/n	0.0011	0.0009	0.0016	0.0011	0.0027	0.0025	0.0051	0.0039	0.0015
$<\alpha>^{1/3}$	0.0092	0.0093	0.0512	0.0336	0.0409	0.0402	0.0899	0.0899	0.0257
$BE(n)$	0.0310	0.0311	0.0427	0.0189	0.0309	0.0285	0.0479	0.0483	0.0283
η/n	0.0347	0.0349	0.3088	0.0271	0.0286	0.0285	0.0392	0.0392	0.0109

4 Conclusions

We examined four graph-based topological indices and five key electronic properties of ZnS nanoclusters. Subsequently, we proposed and evaluated several numerical models, including re-parameterized power-law regressions, piecewise and radial basis function extrapolations. These models outperformed the original power-law model in predicting and estimating cluster properties. To reduce computational cost, we also developed models based solely on cluster size n, eliminating the need for topological index calculations. For instance, we demonstrated that a simple power-law model based on n reliably reproduces target properties with minimal computational overhead. Overall, our results indicate that these numerical

approaches offer efficient and accurate alternatives for estimating and predicting ZnS cluster properties.

References

[1] A. Mohajeri, M. Alipour, M. B. Ahmadi, A graph theory study on $(\mathrm{ZnS})_n$ ($n = 3$–10) nanoclusters, *Chem. Phys. Lett.* **503** (2011) 162–166.

[2] A. Burnin, E. Sanville, J. J. BelBruno, Experimental and computational study of the $\mathrm{Zn}_n\mathrm{S}_n$ and $\mathrm{Zn}_n\mathrm{S}_n^+$ clusters, *J. Phys. Chem. A* **109** (2005) 5026–5034.

[3] S. C. Basak, D. K. Harriss, V. R. Magnuson, Comparative study of lipophilicity versus topological molecular descriptors in biological correlations, *J. Pharm. Sci.* **73** (1984) 429–437.

[4] H. Hosoya, Topological index. A newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons, *Bull. Chem. Soc. Jpn.* **44** (1971) 2332–2339.

[5] D. Bonchev, N. Trinajstić, Information theory, distance matrix, and molecular branching, *J. Chem. Phys.* **67** (1977) 4517–4533.

[6] B. Ren, Y. Xu, G. Chen, A novel topological index for QSPR/QSAR study of organic compounds, *Acta Chim. Sin.* **57** (1999) 563–571.

[7] F. Yang, X. Yan, L. Ouyang, The improved valence connectivity topological index and its application δ comparison with δ_v in QSAR/QSPR of three systems, *J. Chem. Phys.* **11** (1998) #221.

[8] I. Gutman, W. Linert, I. Lukovits, Ž. Tomović, The multiplicative version of the Wiener index, *J. Chem. Inf. Comput. Sci.* **40** (2000) 113–116.

[9] R. Todeschini, V. Consonni, *Handbook of Molecular Descriptors*, Wiley-VCH, Weinheim, 2008.

[10] H. Wiener, Structural determination of paraffin boiling points, *J. Am. Chem. Soc.* **69** (1947) 17–20.

[11] B. Furtula, Trinajstić index, *Discr. Math. Lett.* **9** (2022) 100–106.

[12] M. Randić, X. Guo, T. Oxley, H. Krishnapriyan, Wiener matrix: Source of novel graph invariants, *J. Chem. Inf. Comput. Sci.* **33** (1993) 709–716.

[13] D. J. Klein, I. Lukovits, I. Gutman, On the definition of the hyper-Wiener index for cycle-containing structures, *J. Chem. Inf. Comput. Sci.* **35** (1995) 50–52.

[14] A. Madanshekaf, M. Ghaneei, The second-order connectivity index of dendrimer nanostars, *Dig. J. Nanomat. Biostruct.* **6** (2011) 433–437.

[15] M. Randić, Characterization of molecular branching, *J. Am. Chem. Soc.* **97** (1975) 6609–6615.

[16] L. Kier, *Molecular Connectivity in Chemistry and Drug Research*, Elsevier, 2012.

[17] M. H. Khalifeh, H. Yousefi-Azari, A. R. Ashrafi, I. Gutman, The edge Szeged index of product graphs, *Croat. Chem. Acta* **81** (2008) 277–281.

[18] I. Gutman, A formula for the Wiener number of trees and its extension to graphs containing cycles, *Graph Theory Notes N. Y.* **27** (1994) 9–15.

[19] A. D. Buckingham, Permanent and induced molecular moments and long-range intermolecular forces, *Adv. Chem. Phys.* **12** (1967) 107–142.

[20] R. G. Parr, Density functional theory of atoms and molecules, in: *Horizons of Quantum Chemistry*, Springer, 1989, pp. 5–15.

[21] R. G. Parr, R. G. Pearson, Absolute hardness: Companion parameter to absolute electronegativity, *J. Am. Chem. Soc.* **105** (1983) 7512–7516.

[22] R. G. Pearson, Absolute electronegativity and absolute hardness of Lewis acids and bases, *J. Am. Chem. Soc.* **107** (1985) 6801–6806.

[23] W. Kohn, L. J. Sham, Self-consistent equations including exchange and correlation effects, *Phys. Rev.* **140** (1965) A1133–A1138.

[24] D. A. McQuarrie, J. D. Simon, *Physical Chemistry: A Molecular Approach*, Univ. Science Books, Sausalito, 1997.

[25] W. A. Harrison, *Electronic Structure and the Properties of Solids*, Freeman, San Francisco, 1980.

[26] P. Jena, A. W. Castleman, Clusters: A bridge across the disciplines of physics and chemistry, *Proc. Natl. Acad. Sci. USA* **103** (2006) 10560–10569.

[27] B. Chunsong, A. Naeem, S. Yousaf, A. Aslam, F. Tchier, A. Issa, Exploring expected values of topological indices of random cyclodecane chains for chemical insights, *Sci. Rep.* **14** (2024) #10065.

[28] N. Idrees, E. Noor, S. Rashid, F. T. Agama, Role of topological indices in predictive modeling and ranking of drugs treating eye disorders, *Sci. Rep.* **15** (2025) #1271.