UNIVERSITY OF LEEDS   University of Sheffield   UNIVERSITY of York

Check for updates

Automated bond valence sum analysis for crystallographic database quality assessment: a systematic study of rock salt oxides, halides, and chalcogenides

Nik Reeves-McLaren [ID]

Crystallographic databases are vital for research and increasingly serve as training data for machine learning in materials discovery, yet systematic quality assessment at database scale remains absent. Bond valence sum (BVS) analysis – testing whether calculated bond valences match expected oxidation states – provides one automated quality metric, but has not been systematically applied to quantify validation rates, characterise failure modes, or identify parameter inadequacies. Eir is an automated Python tool for high-throughput BVS analysis, applied here to 840 rock salt structures from the Inorganic Crystal Structure Database. Of 613 assessable structures, 50.8% validated under ambient data collection conditions, with systematic examination leaving no unexplained outliers. 'Failed' structures stratify into six categories (Types 1–6): Type 1, systematic parameter inadequacy (107 structures, 21.3%); Type 2, missing BVS parameters (154 structures, excluded from validation); Types 3–5, methodological limitations from diffraction-averaged geometries (65 structures, 12.9%); and Type 6, database quality issues (12 structures, 2.4%). Three main findings emerged: (1) alkaline earth oxides exhibit systematic parameter inadequacy (CaO, SrO, BaO: 100% failure, <1.5% variance); (2) oxide–chalcogenide validation inversion demonstrates anion-specific parameter quality; (3) multi-phase refinement contamination shows structures passing peer review may prove inappropriate as computational references. The open-source tool provides validated infrastructure for maintaining database integrity as crystallographic repositories serve AI-driven materials discovery at unprecedented scale.

1 Introduction

The rapid expansion of crystallographic databases has highlighted a need for automated quality assessment tools, particularly as these repositories increasingly serve as training data for machine learning models in materials discovery.[1,2] The Inorganic Crystal Structure Database (ICSD) contains over 325 000 entries with thousands of new structures added annually – a scale vastly exceeding the capacity of traditional peer review methods.[3] Manual examination of individual structures through visual inspection of refinement statistics cannot detect systematic errors that manifest across multiple related structures.

In diverse fields, numerous studies have demonstrated that published experimental data contain more errors than commonly assumed: X-ray photoelectron spectroscopy (XPS) peak fitting exhibits approximately 40% error rates, accurate mass measurements in organic synthesis show only 40% full compliance with reporting guidelines, and over 800 fabricated metal–organic framework (MOF) structures from a systematic research papermill passed peer review despite being entirely fictitious.[4–9] The integration of artificial intelligence into materials discovery workflows also makes database integrity assessment critically urgent.[10] Systematic errors in training data propagate through machine learning models and amplify into predictions at scale; database quality assessment tools that can identify and stratify structural errors are therefore essential not merely for archival integrity, but for ensuring the reliability of future AI-driven materials discovery and development.[2]

Bond valence sum (BVS) analysis provides one possible metric for automated quality assessment of inorganic crystal structures. The method, based on the empirical observation that the sum of bond valences around a cation should equal its formal oxidation state, offers a consistency check analogous to the formula-mass matching employed in Christmann's analysis of high-resolution mass spectrometry.[11–13] Unlike

School of Chemical, Materials and Biological Engineering, University of Sheffield, Sir Robert Hadfield Building, Mappin Street, Sheffield, S1 3JD, UK.
E-mail: nik.reeves-mclaren@sheffield.ac.uk

refinement statistics, which assess model-data agreement, BVS analysis tests the physical plausibility of atomic arrangements: structures exhibiting large systematic deviations between calculated bond valence sums (S_i) and expected formal valences (V_i) indicate either incorrect oxidation state assignments, inadequate BVS parameters, or genuine structural anomalies requiring investigation. The method has found widespread application in individual structure validation, identification of mixed-valence compounds, and detection of coordination geometry errors.[14,15] However, BVS analysis has not been systematically applied at database scale to assess the prevalence of structural quality issues or to identify systematic inadequacies in the BVS parameter compilations themselves.

Previous efforts to automate crystallographic data validation have focused primarily on refinement quality metrics and format compliance. Goodman *et al.* (2004) developed an experimental data checker for organic chemistry that demonstrated high accuracy in published experimental data, with errors still occasionally perpetuated.[16] Similar tools exist for protein crystallography (MolProbity, WHAT_CHECK) and small-molecule structures (CheckCIF), but these primarily flag outliers in geometry, displacement parameters, and refinement statistics rather than fundamental physical inconsistency.[17–19] The Materials Project and other high-throughput computational initiatives employ BVS analysis as a quality filter during structure curation, but systematic studies quantifying validation success rates, characterising failure modes, and identifying parameter inadequacies across large structural databases remain absent from the literature.[20]

To address this gap, I developed Eir, an automated Python tool for high-throughput bond valence sum analysis of crystallographic databases. The software performs systematic quality assessment by extracting structural data from Crystallographic Information Files (CIF), calculating coordination environments through symmetry operations and periodic boundary handling, computing bond valence sums using the Gagné & Hawthorne (2015) and Brown (2020) parameter compilations (1385 cation–anion pairs), and classifying structures according to validation reliability.[21–23] The hierarchical confidence framework distinguishes well-validated structures suitable for automated acceptance from those requiring manual review due to borderline deviations, structures exhibiting heterovalent disorder where spectroscopic validation is essential regardless of deviation magnitude, and structures with systematic failures indicating parameter inadequacy or database quality issues. Diagnostic reports identify specific error types when validation fails.

I applied this tool to 841 rock salt (space group $Fm\bar{3}m$) structures retrieved from the ICSD. The dataset includes binary oxides, halides, chalcogenides, and mixed-occupancy solid solutions. Rock salt was selected as a test case due to its simple cubic symmetry, well-defined octahedral coordination, and prevalence across diverse chemical systems, permitting assessment of BVS parameter quality for multiple anion types ($X = O^{2-}$, F^-, Cl^-, Br^-, S^{2-}, Se^{2-}, Te^{2-}) within a controlled structural framework with clear valence assignments. Rock salts

also exhibit sufficient complexity – through solid solution formation, vacancy disorder, and mixed-valence phenomena – to test the robustness of automated classification algorithms.

This work addresses three fundamental questions. First, can systematic deviations distinguish genuine parameter inadequacies from experimental artefacts? Structures exhibiting consistent failures across multiple laboratories and decades potentially indicate where BVS parameters require revision. Second, do validation patterns depend on measurement technology evolution? Temporal analysis of compounds measured across multiple decades permits evaluation of whether apparent validation in older structures reflects genuine parameter adequacy or simply insufficient measurement precision to detect systematic errors. Third, what methodological limitations emerge at database scale that individual structure validation may not reveal? The analysis of diverse chemical systems spanning multiple anion types, mixed-occupancy solid solutions, and structures with varying disorder characteristics enables identification of systematic failure modes, each requiring distinct approaches: parameter revision for reproducible inadequacies, spectroscopic validation for heterovalent disorder systems, and metadata enhancement for database contamination.

The open-source release of Eir provides the crystallographic community with means for systematic quality assessment at database scale, enabling automated workflows where well-validated structures proceed without intervention whilst structures exhibiting significant deviations receive appropriate scrutiny.

2 Methods

2.1 Dataset selection and preparation

Rock salt (space group $Fm\bar{3}m$, NaCl-type) structures were retrieved from the Inorganic Crystal Structure Database (ICSD), with datasets spanning 1916–2025. This structure type was selected due to its simple cubic symmetry, well-characterised coordination geometry (octahedral, coordination number 6), and prevalence across diverse chemical systems, permitting assessment of BVS parameter quality for multiple anion types within a controlled structural framework, and with clear cationic and anionic oxidation state assignments.

The full dataset comprised 767 rock salt structures (unique ICSD collection codes), yielding 841 total measurements when including structures measured under multiple pressure–temperature conditions. Of these, 336 (43.9%) are oxides, 197 (25.7%) are halides, and 233 (30.4%) are chalcogenides. Duplicates at different conditions were most common for oxides (55) and halides (19); chalcogenides had none. Eighty structures (10.4%) exhibit site disorder: 9 with mixed anion sites, 60 with homovalent mixed cation sites, and 11 with heterovalent mixed cation sites.

Mixed-occupancy systems included both homovalent disorder ($Ni_{1-x}Zn_xO$, $Fe_{1-x}Mn_xS$, $Ca_{1-x}Cd_xO$ solid solutions where both cations exhibit identical oxidation states) and heterovalent disorder systems ($Li_xCo_{1-x}O_2$, $Li_xFe_{1-x}O_2$, Li_3TaO_4,

Li_3NbO_4 where cations exhibit valence differences $\Delta V = 2$ or $\Delta V = 4$). These systems tested automated classification of Bosi-type artefacts arising from diffraction-averaged bond lengths in mixed-valence environments.[24] Substitutional solid solutions with mixed anion sites ($K(Br_xCl_{1-x})$, $Na_{1-x}K_x$ chlorides) and vacancy-disordered nonstoichiometric compounds were included to assess methodological limitations.

Of the full dataset, 154 structures (18.3% of the complete dataset) lacked BVS parameters in the Gagné & Hawthorne and Brown compilations and were therefore excluded from further analyses. The largest parameter gap affected lanthanide sulphides (100/135 structures, 74%), followed by early transition metal oxides (Ti^{2+}–O^{2-}, Nb^{2+}–O^{2-}, Sc^{2+}–O^{2-}), silver halides (Ag^+–Br^-, Ag^+–I^-), and 4d metal sulphides (Zr^{4+}–S^{2-}). The end result was a dataset of 613 assessable structures for systematic analysis.

2.2 Bond valence sum theory: physical basis

Bond valence sum (BVS) analysis uses an empirical relationship between interatomic bond length (R) and bond strength (s), quantified through the Brown–Altermatt exponential relationship:

$$s_{ij} = \exp[(R_0 - R_{ij})/B] \qquad (1)$$

where s_{ij} represents the bond valence contribution from bond j to cation i, R_{ij} is the observed bond length, R_0 is an empirically fitted bond valence parameter characteristic of the cation–anion pair, and B is a universal softness parameter (typically 0.37 Å). The valence sum rule states that the sum of bond valences around any atom should equal its formal oxidation state:

$$S_i = \sum s_{ij} \approx V_i \qquad (2)$$

where V_i is the formal valence of cation i. The discrepancy factor $d_1 = (S_i - V_i)$ quantifies deviations: $d_1 > +0.05$ v.u. indicates overbonding (compressed coordination), $d_1 < -0.05$ v.u. indicates underbonding (expanded coordination), and $|d_1| < 0.05$ v.u. Suggests balanced bonding. The Global Instability Index $\left(\text{GII} = \sum_i |S_i - V_i| / \sum_i |V_i| \right)$ provides a structure-wide metric, with GII < 0.1 v.u. typically indicating well-refined structures and GII > 0.2 v.u. suggesting structural instability or parameter inadequacy.[25]

The physical interpretation follows Pauling's second rule: cations with $S_i < V_i$ occupy coordination polyhedra with bonds longer than optimal, whilst those with $S_i > V_i$ experience compressed coordination.[26] Systematic deviations across multiple independent measurements may suggest parameter inadequacy, with isolated outliers pointing to structural defects or poor refinement quality.

2.3 Eir: automated BVS analysis tool

Eir is an open-source Python application developed specifically for high-throughput bond valence sum analysis of crystallographic data. The software addresses the computational chal-

lenges inherent in database-scale validation: robust parsing of diverse Crystallographic Information File (CIF) formats, accurate coordination environment determination across periodic boundaries, efficient BVS calculation using comprehensive parameter compilations, and intelligent error classification to distinguish correctable parameter errors from methodological limitations. The software was developed using a large language model (LLM, Claude 3.5 Sonnet, Anthropic) assisted workflow, with systematic validation steps to ensure code integrity.

2.3.1 Batch processing and data handling. The software analyses entire directories of CIF files, generating both per-structure diagnostic reports and cross-dataset summary statistics. CIF parsing utilises PyCifRW (4.4.1) and gemmi (0.6.5) libraries to extract atomic coordinates, site occupancies, space group symmetry, and lattice parameters from CIF files.[27,28] The parser handles common format variations and provides informative error messages when encountering malformed structural data.

2.3.2 Coordination environment detection. Eir implements symmetry-aware bond finding *via* generation of $3 \times 3 \times 3$ supercells with full application of space group operations, followed by distance-based primary coordination sphere identification with gap analysis requiring minimum 0.05 Å separation to avoid numerical precision artefacts. This approach correctly handles periodic boundary effects and symmetry-equivalent positions in high-symmetry structures, which simple distance cutoffs fail to address. Bond multiplicities are determined by clustering distances within 0.05 Å tolerance after rounding to five decimal places, allowing for symmetry-equivalent positions.

2.3.3 Parameter management and mixed occupancy handling. The tool integrates the Gagné & Hawthorne (2015)[21] compilation of bond valence parameters, supplemented by Brown (2020)[23] updates, providing coverage for 1385 cation–anion pairs across oxides, halides, and chalcogenides. Missing parameter pairs are flagged as non-assessable rather than triggering calculation errors, preventing misclassification of structures. For sites exhibiting fractional occupancy, the tool automatically classifies disorder type (pure sites, homovalent disorder, heterovalent disorder) and calculates valence differences (ΔV) to identify Bosi-type artefacts.[24] For heterovalent sites, the tool computes bond valence sums for each constituent element independently and evaluates whether deviations follow predicted patterns for averaging artefacts, enabling distinction between mathematical artefacts and genuine structural strain.

2.3.4 Confidence classification framework. A hierarchical decision framework stratifies structures according to validation reliability and methodological appropriateness, enabling integration into database curation workflows (detailed decision logic in Section 2.5). Well-validated structures with minimal deviations in ordered systems are accepted automatically, whilst structures exhibiting moderate deviations from homovalent size mismatch effects are flagged for review to confirm the deviation source is geometric rather than indicative of parameter inadequacy. Structures with large systematic deviations

or substantial heterovalent disorder where diffraction-averaged bond lengths cannot distinguish mathematical artefacts from genuine strain are routed for spectroscopic validation, whilst isolated anomalies receive case-by-case manual review. This stratification recognises that different failure modes demand different remediation strategies: parameter revision for reproducible inadequacies across multiple independent structures, complementary experimental techniques for methodologically ambiguous cases, and detailed investigation for outliers potentially reflecting synthesis defects or poor refinement quality.

2.3.5 Diagnostic output and visualisation. For each structure, Eir generates human-readable reports detailing BVS deviations with statistical summaries (mean, standard deviation, range), bonding state classifications (balanced/underbonded/overbonded for each cation site), confidence assignments with justifications referencing specific decision criteria, and actionable recommendations when validation fails. Integration with py3Dmol (2.0.4) provides basic interactive 3D structure viewing when deviations exceed 0.30 v.u., facilitating rapid visual inspection of anomalous coordination environments to identify potential causes.[29]

2.3.6 Development and validation. The tool was developed over approximately 40 iteration cycles addressing crystallographic edge cases: handling of partial occupancies in solid solutions, detection of coordination gaps in low-symmetry distorted structures, management of isotope specifications in CIF files, and robust error handling for malformed structural data. Each iteration was validated against manually calculated reference structures to ensure algorithmic correctness, with particular attention to structures exhibiting known systematic errors from literature sources. Performance testing on a relatively standard gaming laptop (Intel i7-11800H @ 2.30 GHz, 32 GB RAM) demonstrated processing speeds of approximately 1–2 structures per second for rock salt geometries, scaling to larger unit cells and lower symmetries with proportional computational cost.

Eir 1.2 is released under an open-source licence and available at [https://github.com/SuperBladesman/Eir]. The software requires Python 3.9+ with dependencies managed through standard package managers (pip, conda). Comprehensive documentation includes installation instructions and thorough documentation. Eir 1.2 is designed to be run within a Jupyter Notebook environment. The tool is named after Eir, the Norse goddess of healing in Germanic mythology, reflecting its diagnostic role in identifying and characterising structural validation issues in crystallographic data.

2.3.7 Classification framework. Validation employs two criteria to identify structures where BVS analysis functions reliably: percentage deviations below 10% ensure consistent assessment across different oxidation states (Li^+ vs. Ta^{5+}), whilst absolute discrepancy $|d_1| < 0.10$ v.u. ensures deviations remain below thresholds where bonding strain becomes physically significant. This combined approach prevents both false positives (small absolute errors appearing disproportionately large for monovalent cations) and false negatives (large absolute errors masked by high formal valence in multivalent

systems). Using either criterion alone proves insufficient: percentage-only thresholds miss systematic relative errors scaling with ionic potential, whilst absolute-only cutoffs flag acceptable size-mismatch effects in solid solutions as failures.

Structures falling outside these thresholds require interpretation. Intermediate deviations (10–15%, $|d_1| < 0.15$ v. u.) indicate borderline cases where either modest parameter inadequacy or expected geometric effects (homovalent solid solutions, high-pressure compression) could explain observed behaviour – manual review distinguishes these scenarios. Larger deviations typically signal significant issues. For database-scale automated validation, combining absolute ($|d_1|$) and relative (%) thresholds ensures consistent quality standards across different oxidation states, as shown by valence-dependent trends in validation metrics, Fig. S1.

The choice of site-specific validation criteria ($|d_1|$ and % deviation) rather than structure-averaged metrics like GII merits explanation. GII provides a useful structure-wide measure of overall bonding strain but cannot identify which cation sites are problematic or why validation fails. For example, structures with heterovalent disorder where all sites deviate modestly (e.g., 5% across all sites) produce high GII but may represent genuine mixed-valence states requiring spectroscopic validation rather than automatic rejection. Conversely, structures where a single site exhibits severe deviation (e.g., one site at 15%, others at 2%) produce low GII but indicate clear parameter inadequacy for specific cation–anion pairs. The site-specific approach enables identification of which cation–anion combinations require parameter revision and distinction between systematic parameter failure and measurement quality issues. GII is reported for all structures (Data Files S1 and S3–S6) and provides useful context, but the primary validation workflow uses site-specific metrics to enable systematic parameter improvement.

The classification framework recognises that different disorder types demand different validation approaches. Compositionally ordered structures with small deviations validate automatically. Homovalent disorder systems (identical oxidation states, modest size differences) receive relaxed thresholds (up to 15%) because diffraction-averaged lattice parameters inherently produce geometric mismatch even when parameters are adequate. Heterovalent disorder systems ($\Delta V \geq 3$) were not assessed for BVS validation entirely and were flagged as required spectroscopic confirmation beyond the scope of this work, as diffraction-averaged bond lengths struggle to distinguish mathematical averaging artefacts from genuine structural strain when substantially different valences occupy equivalent sites.

2.3.8 Distinguishing parameter inadequacy from experimental scatter. Systematic parameter failures exhibit two diagnostic signatures: reproducible deviations across multiple independent measurements (different laboratories, synthesis methods) combined with exceptionally low inter-structure variance (<5% of mean deviation magnitude). This pattern likely indicates genuine parameter inadequacy rather than experimental error, synthesis defects, or coordination environment

determination failures, all of which would produce substantially larger variance. Isolated outliers from otherwise well-validated compound classes instead were manually reviewed for case-specific assessment.

3 Results

3.1 Overall dataset statistics and validation patterns

For the 613 assessable structures, systematic analysis stratified by measurement conditions revealed condition-dependent validation behaviour. Of 502 structures measured under ambient conditions ($T < 350$ K, $P < 0.5$ GPa), 255 structures (50.8%) exhibited bond valence sum deviations below 10% with discrepancy factors $|d_1| < 0.10$ valence units (v.u.). These empirical thresholds correspond to well-refined structures where calculated bond valences match expected formal valences within normal experimental precision, whilst the discrepancy factor $d_1 = S_i - V_i$ provides a site-specific measure of bonding strain that enables distinction between balanced coordination environments ($|d_1| < 0.05$ v.u.) and structures requiring investigation for parameter inadequacy, coordination environment errors, or genuine structural anomalies ($|d_1| > 0.10$ v.u.).

The automated stratification system partitioned the 502 ambient structures into hierarchical categories requiring different validation approaches. Of these, 255 structures (50.8%) exhibited deviations permitting automated acceptance, 68 structures (13.5%) fell into the borderline category requiring manual review to distinguish size-mismatch effects from parameter inadequacy, and 179 structures (35.7%) exceeded validation thresholds.

Systematic examination of the 179 failed structures revealed complete traceability: 175 structures (34.9%) partition into six identified failure modes (Types 1–6), whilst 4 structures (0.8%) lacked sufficient independent measurements ($n < 3$) or exhibited quality indicators preventing reliable classification:

• Type 1: systematic parameter inadequacy (107 structures, 21.3%).

• Type 2: missing BVS parameters (154 structures, excluded from validation; discussed in Section 2.1).

• Types 3–5: methodological limitations where diffraction-averaged structures prove inappropriate for BVS analysis (65 structures, 12.9%).

• Type 6: database quality issues arising from measurement contexts or synthesis conditions (12 structures, 2.4%).

This stratification places BVS analysis as a diagnostic tool that identifies what requires attention and why, Table 1.

Validation rates also varied with measurement conditions. High-pressure/temperature structures showed substantially reduced validation (19.4%, 19/98 structures) relative to ambient conditions (49.5%, 283/572 structures), consistent with pressure-induced bond compression producing overbonding that exceeds validation thresholds. Elevated-temperature structures at ambient pressure exhibited unexpectedly high validation (81.2%, 13/16 structures), though the small sample size prevents definitive interpretation. The systematic pressure-dependence demonstrates that current parameter compilations implicitly assume ambient geometry.

The systematic variation in validation rates by anion type may provide insight into differential BVS parameter adequacy across compound classes. Fluorides (92.0%) and bromides (88.0%) validate at rates indicating high parameter quality, likely reflecting adequate representation in the reference structure sets employed during parameter fitting by Gagné & Hawthorne and earlier compilations.

Table 1 Bond valence sum validation statistics for 613 rock salt structures organised by (a) measurement condition (ambient, or high pressure and/ or temperature), and (b) compound class. Validation requires percentage deviation <10% and $|d_1| < 0.10$ v.u. Ambient conditions: $T < 350$ K, $P < 0.5$ GPa

(a) Validation by measurement condition (all assessable structures)

Condition	Structures	Validated, n (%)	Borderline, n (%)	Failed,[a] n (%)
Ambient	502	255 (50.8%)	68 (13.5%)	179 (35.7%)
High T	16	13 (81.2%)	3 (18.8%)	0 (0.0%)
High T/P	95	18 (18.9%)	22 (23.2%)	55 (57.9%)
TOTAL	**613**	**286 (46.7%)**	**93 (15.2%)**	**234 (38.2%)**

(b) Validation by compound class (ambient conditions only)

Class	Structures	Validated, n (%)	Borderline, n (%)	Failed,[a] n (%)
Fluoride	24	23 (95.8%)	1 (4.2%)	0 (0.0%)
Bromide	25	22 (88.0%)	0 (0.0%)	3 (12.0%)
Sulphide	109	82 (75.2%)	3 (2.8%)	24 (22.0%)
Iodide	9	5 (55.6%)	0 (0.0%)	4 (44.4%)
Chloride	78	35 (44.9%)	15 (19.2%)	28 (35.9%)
Oxide	255	87 (34.1%)	49 (19.2%)	119 (46.7%)
Other	2	1 (50.0%)	0 (0.0%)	1 (50.0%)
TOTAL	**502**	**255 (50.8%)**	**68 (13.5%)**	**179 (35.7%)**

[a] Failed structures include categorised failures (Types 1–6, 28.5%) and structures with insufficient data (9.6%).

The high fluoride validation rate (95.8%, 23/24 structures) is particularly significant given fluorine's strong electronegativity and the potential for ionic radius variations with coordination number, suggesting that available F^- parameters capture these effects adequately for rock salt coordination environments across the complete alkali metal series from Li^+ ($r = 0.76$ Å, CN = 6) to Cs^+ ($r = 1.67$ Å, CN = 6). Parameter quality remains excellent even for the largest cation radius combinations. It should be noted that the fluoride dataset comprises only those alkali metal cations that adopt the rock salt structure (Li^+, Na^+, K^+, Cs^+); RbF crystallises in the CsCl structure type (CN = 8) and is therefore excluded from this analysis. In contrast, alkaline earth oxides adopt rock salt geometry across the complete series (MgO to BaO), providing a more complete test of parameter adequacy across a wider cation size range.

Sulphides occupy an intermediate position (75.2%), with failures concentrated in early lanthanides (LaS, NdS, ErS) exhibiting unusual Ln^{2+} oxidation states under extreme reducing conditions; alkaline earth and late 3d sulphides validated reliably. This pattern may reflect either limited high-quality reference structures for S^{2-} parameter fitting or genuine chemical complexity in sulphide bonding that challenges rigid-ion assumptions. Iodides show further reduced validation (55.6%), with all three LiI structures failing. This may reflect both parameter inadequacy for Li^+–I^- combinations and limited reference data for the largest, most polarisable halide.

The most significant finding concerns the poor validation rates observed for chlorides (39.6%) and oxides (36.7%). For oxides, this represents a fundamental challenge to the conventional assumption that O^{2-} parameters achieve high quality given oxygen's prevalence in inorganic crystal chemistry and the extensive literature on oxide structures. The observation that chlorides validate as poorly as oxides is equally unexpected, given that alkali chlorides such as NaCl function as textbook examples in crystallographic education and might therefore be assumed to validate reliably. The convergence of oxide and chloride validation rates suggests common limitations potentially related to reference structure quality, systematic errors in historical diffraction measurements, or inadequacy of the functional form employed for these particular anion types.

The subsequent sections analyse the failure categories to distinguish genuine parameter inadequacies (requiring future systematic revision of published parameters) from methodological limitations (where diffraction-averaged structures prove inappropriate for BVS analysis regardless of parameter quality) and database quality issues (where structures lacking proper refinement or deposited from inappropriate synthesis contexts produce spurious deviations).

3.2. Systematic deviations and parameter inadequacy

3.2.1 Alkaline earth and related divalent oxides.
Systematic parameter inadequacy, defined as reproducible failure across multiple independent structures with exceptionally low inter-

Table 2 Systematic validation failures in divalent oxides. Mean BVS deviation and standard deviation calculated for all ambient measurements for each oxide. MgO failure rate (26.4%) includes 19 high-pressure/temperature measurements. Ambient-only validation: 97.1% (33/34 structures); high P/T: 31.6% (6/19 structures)

Oxide	Structures, n	Failure rate	Mean BVS deviation (Std Dev)/%	Mean cation radius/Å
MgO	53	26.4%	−2.6 (±3.2)	0.72
ZnO	4	100%	−7.9 (±0.6)	0.74
CdO	17	100%	−10.9 (±1.3)	0.95
CaO	15	100%	−11.0 (±1.0)	1.00
EuO	12	100%	−16.8 (±0.3)	1.17
SrO	11	100%	−17.5 (±1.2)	1.18
BaO	5	100%	−21.4 (±0.9)	1.35

structure variance, emerged most clearly for alkaline earth oxides and related divalent systems. Table 2 presents validation statistics and mean deviations for binary oxides spanning the series MgO → CaO → SrO → BaO alongside the lanthanide divalent oxide EuO and transition metal oxides CdO and ZnO.

Three features establish systematic parameter inadequacy, Fig. 1. First, all alkaline earth oxides with cations larger than Mg^{2+} (Ca, Sr, Ba) across multiple synthesis methods and measurement decades. Second, mean deviations increase sys-
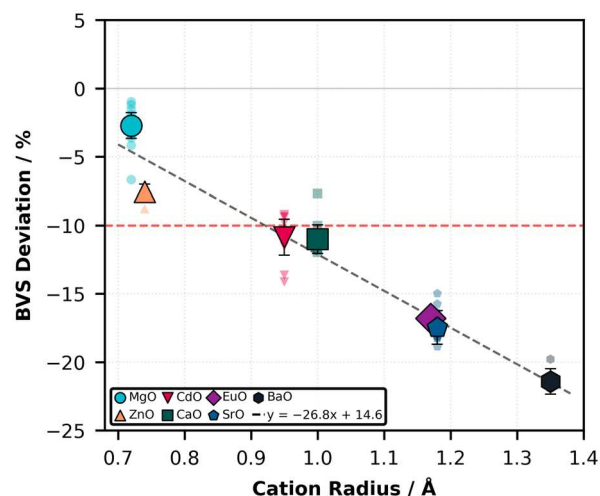

Fig. 1 Bond valence sum deviation *versus* cation radius for divalent oxides under ambient conditions ($T < 350$ K, $P < 0.5$ GPa). Small symbols represent individual measurements; large symbols show mean values with error bars indicating standard deviation. Circles: alkaline earth oxides (MgO, CaO, SrO, BaO); triangles: late transition metal oxides (ZnO, CdO); square: lanthanide oxide (EuO). Red dashed lines indicate the ±10% validation threshold. The systematic increase in negative deviation with increasing cation radius ($r = -0.526$), coupled with exceptionally low inter-structure variance (<1.5% for Ca, Sr, Ba, Eu), demonstrates genuine parameter inadequacy rather than experimental scatter. MgO exhibits transitional behaviour, validating under ambient conditions but failing systematically under pressure. All oxides with $r > 0.95$ Å show 100% validation failure regardless of measurement decade (1922–2023) or technique, indicating that parameter revision rather than improved measurements is required.

tematically with cation radius from −11.0% for CaO to −21.4% for BaO. Finally, low inter-structure variance (<1.5% standard deviation excluding MgO) distinguishes genuine parameter inadequacy from experimental scatter or coordination environment errors that would cause larger variance.

The systematic under-bonding could reflect either parameter inadequacy or genuinely unfavorable coordination, as alkaline earth oxides larger than MgO thermodynamically prefer higher coordination numbers (CN 8–12) over rock salt octahedral geometry. However, the oxide-sulphide validation inversion (Section 3.2.2) provides definitive evidence: BaS and CaS validate whilst BaO and CaO fail systematically despite identical cation sizes and coordination numbers, demonstrating anion-specific parameter inadequacy rather than cation-specific structural instability.

MgO occupies a transitional position, validating in 97.1% of ambient structures (−2.6 ± 3.2%) but failing systematically under high pressure (P > 10 GPa: 100% failure, deviations exceeding −15%). This condition-dependent behaviour indicates that Mg^{2+}–O^{2-} parameters achieve adequate accuracy within their ambient training range but extrapolate poorly to compressed geometries, providing additional evidence that negative deviations in larger alkaline earth oxides reflect parameter inadequacy rather than experimental artefacts.

CdO and ZnO extend the systematic inadequacy to late transition metals. CdO (100% failure, −10.9 ± 1.3%) exhibits deviation magnitude comparable to CaO despite having different electronic configuration ($4d^{10}$ *vs.* noble gas), whilst ZnO (100% failure, −7.9 ± 0.6%) falls between MgO and CdO in both radius and deviation magnitude. These oxides were not flagged in preliminary analyses with smaller datasets; comprehensive database-scale analysis permitted recognition of reproducible patterns in systems initially represented by insufficient structures (n < 3) to assess systematicity.

Temporal spans in these data range from 68 years (BaO) to 107 years (MgO), moving from film-based powder diffraction (1920s) through scintillation counter diffractometry (1960s–1980s) to modern CCD and synchrotron techniques (2000s–2020s). CaO provides a particularly instructive case: 15 structures spanning 1922–2023 yield −11.0 ± 1.0%, with structures measured in 2023 not significantly improved from those measured in 1922, indicating that parameter revision rather than improved measurements is required.

3.2.2 Oxide-chalcogenide inversion: evidence for anion-specific parameter inadequacy. Systematic comparison of validation behaviour across anion types revealed that parameter adequacy depends on both anion identity and cation size, Fig. 2.

Oxide validation exhibits negative correlation with cation radius (r = −0.509, p = 0.13, n = 10), though not reaching statistical significance. MgO (0.72 Å) validates reliably (97% validation across 34 structures), whilst validation degrades through MnO (0.83 Å, 22% validation across 36 structures) to complete systematic failure for r ≥ 0.95 Å.

Fluorides validate uniformly across the complete cation radius range examined (0.76–1.52 Å, mean validation 92.0%,
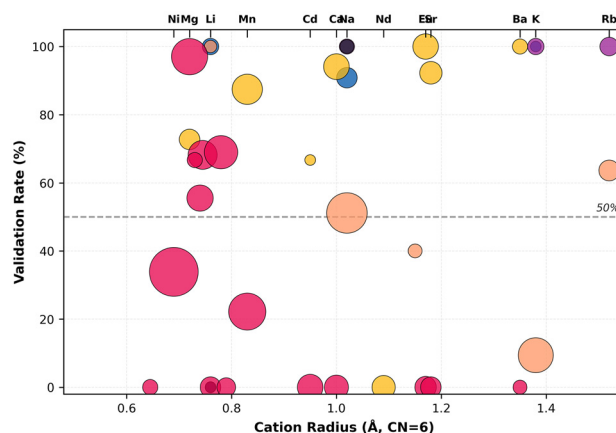

Fig. 2 Validation rate *versus* cation radius stratified by anion type for cation–anion combinations with n ≥ 3 structures. Bubble size represents number of structures; bubble colour indicates anion type: O^{2-} (red), F^- (blue), Cl^- (orange), S^{2-} (yellow), Br^- (purple), I^- (black). Top axis shows representative cation positions for reference.

r = −0.406, p = 0.59, n = 4), demonstrating size-independent parameter adequacy. Chlorides exhibit behaviour intermediate between oxides and fluorides, showing modest positive correlation (r = +0.785, p = 0.22, n = 4) with increased validation for large cations. Bromides display similar positive behaviour (r = +0.834, p = 0.17, n = 4), validating reliably for large cations whilst showing reduced validation for small cations.

Direct oxide-sulphide comparison for cations represented in both compound classes (Ca, Ba, Eu, Mg, Cd, Mn) provides the strongest evidence for anion-specific parameter inadequacy. Calcium sulphide validates at 94.1% (16 of 17 structures) whilst calcium oxide fails completely (0%, 15 structures). Barium and europium exhibit identical inversions: BaS and EuS validate whilst BaO and EuO fail systematically. Magnesium exhibits no inversion (73.6% oxide, 72.7% sulphide), consistent with transitional adequacy at the small-radius limit. If systematic underbonding in alkaline earth oxides reflected errors in cation radius values, coordination environment determination failures, or inappropriate treatment of cation polarisability, these failures should appear in sulphides. The observation that Ba^{2+}–S^{2-} and Ca^{2+}–S^{2-} parameters validate reliably whilst Ba^{2+}–O^{2-} and Ca^{2+}–O^{2-} parameters fail systematically indicates anion-specific inadequacy rather than cation-specific issues.

The alkaline earth sulphide measurements span 1923–2006, with no structures measured post-2006 representing an 18–44 years gap during which diffractometer technology advanced substantially. The MgS case demonstrates that early measurements can mask inadequacies: 1990s structures validated whilst 2023 structures failed systematically. Modern high-precision measurements of CaS, SrS, and BaS using synchrotron diffraction with refinement quality comparable to recent MgS determinations represent the critical experiment distinguishing measurement precision evolution from genuine anion-specific parameter adequacy.
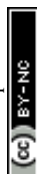
3.3 Mixed-occupancy systems

3.3.1 Homovalent solid solutions: size-mismatch effects. Three $Ni_{1-x}Zn_xO$ solid solutions (x = 0.2, 0.3, 0.4, measured 1954) exhibited deviations of −8.4%, −10.1%, and −10.7%, straddling the validation threshold. Diffraction reports averaged metal–oxygen distances reflecting weighted means of Ni–O and Zn–O bonds. BVS analysis applied to this averaged geometry produces modest deviations (8–11%) likely reflecting inadequacy of linear parameter averaging: local Ni–O (Ni^{2+}, 0.69 Å) and Zn–O (Zn^{2+}, 0.74 Å) distances differ from the crystallographic average of the solid solution composition.

The borderline validation status suggests that homovalent disorder with modest size differences ($\Delta R < 0.1$ Å) produces manageable deviations potentially attributable to size-mismatch effects rather than fundamental parameter inadequacy. The absence of large-magnitude failures (>20%) or systematic directional bias across the compositional series indicates that BVS analysis functions semi-quantitatively for homovalent solid solutions, though precision degrades relative to pure end-member compounds. This contrasts sharply with heterovalent disorder cases where fundamentally different oxidation states produce qualitatively different failure patterns.

3.3.2 Heterovalent disorder: Bosi-type artefacts. Structures containing cations of substantially different oxidation state on crystallographically equivalent sites (heterovalent disorder) exhibited systematic validation failure characterised by bidirectional deviations: lower-valence cations consistently overbonded whilst higher-valence cations are underbonded. This pattern, termed a Bosi-type artefact following Bosi (2014),[24] arises from fundamental incompatibility between diffraction-averaged structures and site-specific bond valence sum analysis when valence differences $\Delta V \geq 2$ exist.

Table 3 presents validation statistics for heterovalent systems spanning $\Delta V = 2$ (Li^+/M^{3+} in $LiMO_2$) and $\Delta V = 4$ (Li^+/M^{5+} in Li_3MO_4). All 22 structures exhibited 100% failure. When Li^+ (0.76 Å) and Co^{3+} (0.55 Å low-spin) occupy the same crystallographic site in $LiCoO_2$, X-ray diffraction reports a weighted average metal–oxygen distance (~2.05 Å) reflecting the 50 : 50 site occupancy. This averaged distance is too short for Li^+ (producing overbonding, $S_i > V_i$) and too long for Co^{3+}

(underbonding, $S_i < V_i$). The diffraction-averaged geometry thus satisfies neither cation's bonding requirements individually despite the actual local structures being chemically reasonable. Deviation magnitude scales with valence difference: $\Delta V = 2$ systems exhibit deviations of 25–32%, whilst $\Delta V = 4$ systems show 38–60%, reflecting greater disparity between required bond lengths when larger valence differences are forced into a single averaged distance.

The high standard deviations in Table 3 (±20 to ±36%) contrast with the exceptionally low variances observed for systematic parameter inadequacy cases (Table 2). These large variances reflect genuine structural and compositional variation in a small dataset rather than simple experimental scatter.

The $LiFeO_2$ dataset (12 structures, 1964–2022) spans multiple synthesis routes (solid state, hydrothermal, ion exchange), variable stoichiometry (some structures report non-unity Li : Fe ratios), and potential oxygen nonstoichiometry ($LiFeO_{2+\delta}$). The $LiCoO_2$ dataset includes structures measured at different temperatures (293 K, 373 K) and pressures (ambient, high-pressure), conditions known to affect Li/Co site occupancy in layered oxides. The variance thus encodes real chemical and structural diversity rather than measurement imprecision.

Methodological implications. The 100% failure rate for heterovalent systems with $\Delta V \geq 2$ indicates that these rock salt structures should not be subjected to conventional BVS validation criteria developed for ordered single-cation systems. The failures do not reflect inadequate parameters – the Gagné & Hawthorne compilation includes well-validated $Li^+–O^{2-}$, $Co^{3+}–O^{2-}$, $Fe^{3+}–O^{2-}$, $Ta^{5+}–O^{2-}$, and $Nb^{5+}–O^{2-}$ parameters that function reliably in ordered structures. Rather, the failures reflect fundamental inappropriateness of applying site-specific BVS analysis to diffraction-averaged bond lengths when substantial valence heterogeneity exists.

Structures exhibiting substantial heterovalent disorder ($\Delta V \geq 2$) should be flagged for spectroscopic validation regardless of BVS deviation magnitude, as diffraction alone cannot confirm whether reported averaged geometries reflect genuine structure or represent refinement artefacts. Techniques providing site-specific oxidation state information, such as X-ray absorption near-edge spectroscopy or Mössbauer spectroscopy for appropriate elements, are required to validate that expected valence states exist locally even when diffraction-averaged bond lengths appear chemically unreasonable. Users requiring local coordination environment information should seek techniques providing site-specific probes (pair distribution function analysis, spectroscopy, neutron diffraction with high Q-resolution) rather than accepting diffraction-averaged structures uncritically. The $LiCoO_2$ and $LiFeO_2$ structures analysed here have been cited extensively in battery materials literature and employed as input for density functional theory calculations, yet their reported M–O distances represent averages over Li^+ and M^{3+} sites that cannot simultaneously satisfy both cations' bonding requirements.

3.3.3 Compositional disorder with large size mismatch. Beyond heterovalent disorder (Section 3.3.2), systematic fail-

Table 3 Bond valence sum validation statistics for heterovalent disorder systems exhibiting Bosi-type artefacts. ΔV = valence difference between mixed cations. All structures showed 100% validation failure. Mean deviations represent averaged percentage deviations across all measurements for each compound. The bidirectional deviation pattern (lower-valence cations overbonded, higher-valence cations underbonded) demonstrates, despite the small sample size, systematic mathematical artefacts from diffraction-averaged bond lengths

Oxide	Structures, n	ΔV	Mean BVS deviation (Std Dev)/%	Failure rate
$LiFeO_2$	12	2	+3.7 (±20.5)	100%
$LiCoO_2$	4	2	+5.6 (±35.6)	100%
Li_3TaO_4	4	4	−6.5 (±28.3)	100%
Li_3NbO_4	2	4	−6.0 (±33.7)	100%
TOTAL	22	—	(±20–36%)	100%

ures emerged when atoms of identical formal charge but different size occupy crystallographically equivalent positions. These compositional disorder cases (Type 5) represent methodological limitations distinct from parameter inadequacy.

Mixed anion solid solutions (Type 5A) exhibited deviations of +15% to +52% when two halide ions occupied the same site. Seven $Na(Br_xCl_{1-x})$ structures showed systematic compositional dependence: deviation decreased from +51.9% at $x = 0.10$ (near-NaCl) to +15.2% at $x = 0.70$ (near-NaBr). The mechanism mirrors heterovalent disorder but with size differences (Cl^- 1.81 Å $vs.$ Br^- 1.96 Å) rather than valence driving the incompatibility.

Mixed cation solid solutions (Type 5B) represented the largest failure category amongst compositionally disordered systems (60 structures, 10.5% of ambient dataset, 67% failure rate). The thirteen $Na_{1-x}K_xCl$ structures exhibited the most systematic behaviour: deviations ranged from −19% to +64% and correlated strongly with composition ($r = -0.96$, $p < 0.001$), reflecting substantial size mismatch (Na^+ 1.02 Å, K^+ 1.38 Å). Mixed 3d transition metal oxides ($Ni_{1-x}Zn_xO$, $Fe_{1-x}Mn_xO$, 41 structures) showed smaller deviations (−1% to +19%) due to more modest cation size differences (Ni^{2+}–Zn^{2+} $\Delta R = 0.05$ Å).

General principle. Deviation magnitude scales with size mismatch. Small differences ($\Delta R < 0.10$ Å) produce borderline deviations; large differences ($\Delta R > 0.30$ Å) cause systematic failures exceeding 50%, demonstrating that BVS analysis is inappropriate for structures with substantial compositional disorder. These structures are legitimate for phase identification but unsuitable for applications requiring local bonding information (BVS validation, DFT calculations, ML training). Databases could assist users by flagging disorder type so that structures can be selected as appropriate for their purpose.

3.3.4 Vacancy disorder: nonstoichiometry and oxygen deficiency. Two structures exhibited deviations attributable to vacancy disorder: $Nd_{1.00}S_{0.89}$ (+49.2%) and $V_{1.00}O_{0.90}$ (+48.6%). Both represent nonstoichiometric phases with anion deficiency, refining to compositions departing from ideal MX stoichiometry. The positive deviations (overbonding) reflect how bond valence sum analysis interprets vacancy-containing structures: observed cation–anion distances remain approximately constant whilst anion site occupancy decreases, producing fewer bonds per cation site. BVS analysis sums over existing bonds without accounting for vacant sites, thereby calculating valence sums exceeding formal valence.

These failures represent a fundamental limitation of conventional BVS implementation rather than parameter inadequacy. The method assumes full site occupancy and sums valence contributions from all nearest-neighbour bonds identified through geometric criteria. Vacancy disorder violates this assumption: cations possess fewer oxygen neighbours than coordination number alone would suggest, but standard BVS calculation identifies all nearest-neighbour positions (both occupied and vacant) and sums their contributions.

Correct treatment would require modification of summation procedures to weight contributions by anion site occupancy factors, essentially computing bond valence sums over averaged coordination environments accounting for statistical vacancy distribution. The small number of vacancy-disordered structures (2 of 572 ambient structures, 0.3%) indicates this represents a minor contribution to overall failure statistics, but the qualitatively distinct failure mechanism warrants separate classification. Nonstoichiometric phases are important in functional materials applications (oxygen storage materials, ionic conductors, battery cathodes) and users retrieving such structures from databases should recognise that conventional BVS validation will systematically fail for structures with substantial vacancy disorder regardless of parameter quality or measurement precision. Future development of Eir will look to implement this.

3.4 Database quality assessment: structure appropriateness for computational applications

Systematic examination of failed structures revealed multiple cases where validation failures arose not from parameter inadequacy or methodological limitations but from database entries representing structures of questionable appropriateness for reference applications. Fourteen structures (2.4% of ambient dataset) exhibited deviations traceable to measurement techniques, sample contexts, or synthesis conditions rendering them unsuitable as computational reference structures despite having passed peer review and database curation. These cases are partitioned into four categories documenting distinct database quality issues that systematic BVS analysis helps to successfully identify.

Type 6A: Pre-modern crystallographic techniques (8 structures, 1.4% of dataset). Structures measured 1920–1970 using film-based powder diffraction, reporting lattice parameters to 1–4 decimal places without R-factors, thermal displacement parameters, or estimated standard deviations, represent the historical foundation of crystallographic databases but lack quality indicators required for computational applications. Examples include:

• **NiO (CollCode 53930, 1920):** Lattice parameter to one decimal place, deviation −8.2%.[30]

• **MgS (CollCode 53939, 1923):** Film technique, no diffraction patterns shown, deviation +12.2%.[31]

• **SmO (CollCode 77682, 1956):** Synthesis study in JACS reporting lattice parameters only from material sublimed at 1125–1300 °C under argon, no R-value or complete thermal parameters, deviation +8.0%.[32]

• **Ag_xNa_yCl ($x = 0.1, 0.2$; $y = 0.9, 0.8$; CollCodes 60281–60282, 1970):** film data, lattice parameters only, deviations −16%.[33]

A $K_{0.3}Rb_{0.7}Cl$ structure (CollCode 22171, 1969) merits particular emphasis: ICSD metadata explicitly states "composition has largest deviation from Vegard's Law", perhaps suggesting even that this structure was highlighted precisely because it exhibits anomalous behaviour.[34] The database entry appears to originate from a systematic study exploring compositional extremes rather than representing a typical K–Rb–Cl solid solution. This case illustrates that structures can enter databases with explicit warnings in text fields yet still be retrieved by

automated workflows as if representative, emphasising the need for machine-readable quality flags rather than human-readable comments.

These structures represent legitimate crystallographic work from their eras and provide valuable historical documentation of phase identification and lattice parameter determination. However, their measurement precision and refinement completeness fall below standards required for computational reference structures. The bond valence sum failures do not indicate these structures are "wrong" but rather signal that their quality metrics differ from modern determinations, enabling stratification for fitness-for-purpose: appropriate for historical phase diagram compilation, inappropriate for machine learning training data requiring sub-0.01 Å precision in bond lengths.

Type 6B: Thin film and non-bulk determinations (1 structure, 0.2% of dataset, though many other similar cases were excluded prior to final dataset collection). PdO (CollCode 77650, 1989, −83.0% deviation) represents a vacuum-deposited thin film characterised by electron diffraction, with structure determination reported as having lattice parameters to two decimal places and no R-values or thermal parameters.[35] Thin films exhibit strain, preferred orientation, substrate effects, and potential nonstoichiometry absent in bulk materials. The extreme deviation (−83%) flags this structure as unsuitable for bulk reference applications, yet the database entry lacks searchable metadata distinguishing thin film from bulk synthesis – in fact, the entry is listed as powder data. The CIF download does not have sample form as a tagged data field either.

Thin film crystallography serves important technological purposes for characterising coatings, interfaces, and device structures, and such data merit preservation in crystallographic databases. However, database practices could be enhanced by systematic flagging indicating synthesis method, enabling automated distinction between bulk reference structures appropriate for computational training and thin film structures appropriate for device-oriented applications – somewhat achieved in the ICSD, and not recorded within the CIFs users download. The extreme BVS deviation effectively provides such a flag through indirect means: any structure deviating by >50% likely reflects either experimental failure or fundamental inappropriateness for bulk structure applications.

Type 6C: Multi-phase refinement contamination (4 structures, 0.7% of dataset). Structures refined as secondary phases in multi-phase patterns to rightly improve fit quality showed systematic deviations when deposited in databases as if representing pure bulk phases:

• **NiO (CollCode 241420, 2016, −32.6%):** Refined as <1% impurity phase in an excellent battery electrode study of $LiNi_{0.5}Mn_{1.5}O_4$.[36] The primary focus rightly was cathode characterisation; NiO refined without estimated standard deviations solely to account for minor diffraction peaks from trace impurity.

• **NiO (CollCode 131090, 2019, −8.0%):** Secondary phase in birnessite (layered manganese oxide) study.[37] The study dis-

cusses "NiO-type $Ni_{0.26}Mn_{0.74}O_x$" with unconfirmed oxygen content and notes NiO appearing as a second phase. Refined parameters lack ESDs and lattice parameter reported as 4.198 Å without uncertainty.

• **MgS (CollCodes 64415, 64422, 2023, −27.8% and −11.2%):** Most instructive example, representing MgS formed during Mg–S battery discharge and analysed via Rietveld refinement of electrode materials. SI shows 'Sample 1' contained C, SiC, S, MgS (7.5 wt%), Mg metal, and "MgCS" phase; 'Sample 2' after recharge still contained 9.3% MgS alongside multiple magnesium-containing phases from electrolyte decomposition (MgO, $MgSO_4$, MgF_2).[38] The study employed positron annihilation spectroscopy to study defect evolution, with XRD providing phase identification. The MgS represents non-equilibrium electrochemical product with nanoscale morphology, incomplete reversibility, and multi-phase contamination, fundamentally distinct from bulk-synthesised reference MgS. The authors never claimed to synthesise or characterise pure bulk MgS – that was not their objective. However, database deposition practices treat all structures equivalently regardless of synthesis context, leading computational workflows to retrieve electrochemically formed MgS in multi-phase battery electrodes alongside high-purity single-crystal MgS determinations.

This category represents the most concerning database quality issue identified because it reflects neither historical measurement limitations (Type 6A) nor specialised application contexts (Type 6B) but rather inappropriate deposition of subsidiary structural information from studies whose primary focus lay elsewhere. Multi-phase Rietveld refinements routinely determine lattice parameters and sometimes atomic coordinates for all phases present to improve overall fit quality, but such parameters represent optimised values within multi-phase constraints rather than independent structure determinations. Database practices depositing each phase from multi-phase refinements as separate entries without contextual metadata enable retrieval of inadequately determined structures by users assuming single-phase bulk characterisation.

Type 6D: Miscategorized measurement conditions (1 structure, 0.2% of dataset). YbO (CollCode 77710) synthesised at 4–5 GPa and 873–1373 K lacks pressure/temperature metadata in database entries, leading to classification as "ambient" despite extreme synthesis conditions.[39] The structure derives from a 1978 Comptes Rendus note titled "Synthèse de YbO et YbO sous haute pression" (Synthesis of YbO and YbO under high pressure), which suggests that the phase and data may be of high-pressure origin. The +13.8% deviation likely reflects either pressure-induced geometric changes or nonstoichiometry stabilised under extreme conditions rather than parameter inadequacy for ambient YbO.

This case may exemplify metadata loss during database deposition: synthesis conditions stated in source papers failing to propagate to database entries as machine-readable flags. A user retrieving this structure programmatically receives lattice parameters and atomic coordinates but no indication that these represent a high-pressure phase potentially exhibit-

ing metastable geometry. Systematic BVS screening identifies such cases through deviations inconsistent with chemical expectations for ambient structures, enabling *post-hoc* literature examination to recover lost synthesis context.

3.4.1 Composite assessment. The fourteen database quality cases (Types 6A–6D) represent 2.4% of the ambient dataset and suggest different remediation approaches for each category. Pre-modern crystallography (Type 6A) requires metadata enhancement flagging measurement precision and refinement completeness to enable fitness-for-purpose assessment without removing historically valuable data. Thin films and specialised syntheses (Types 6B, 6D) require systematic flagging of synthesis method and conditions. Multi-phase contamination (Type 6C) represents the most actionable category: database deposition policies could require explicit indication when structures derive from multi-phase rather than single-phase refinements, enabling users to make informed decisions about appropriateness for reference applications.

No unexplained outliers remained after systematic literature examination. Every structure failing validation criteria without falling into Types 1–5 (parameter inadequacy, methodological limitations) proved traceable to documented quality issues discoverable through source paper examination. This demonstrates the intended function of database-scale systematic quality assessment: structures need not be excluded from databases, but they must be sufficiently characterised for computational tools to make informed decisions about fitness-for-purpose. A structure appropriate for phase identification may be inappropriate for machine learning training. A structure measured under extreme conditions is not defective but requires condition metadata to avoid classification as an ambient-condition parameter failure. A structure refined as 7.5% minority phase in a multi-phase battery electrode provides valuable information about phase evolution during electrochemical cycling but should not serve as a bulk reference structure. Systematic BVS analysis enables automated identification of such fitness-for-purpose distinctions that human curation at deposition time cannot anticipate, as appropriateness depends on intended application rather than inherent structure quality.

4 Discussion

4.1 Practical guidance for database users: red flags and decision trees

Researchers building computational datasets or training machine learning models should apply systematic quality filters before accepting database entries. The following red flags warrant scrutiny:

4.1.1 Pre-1980 structures lacking *R*-factors or thermal parameters. Film-based measurements provided adequate lattice parameters for phase identification but insufficient precision for bond length-dependent applications. A case in point being the example of SmO discussed previously – appropriate for

confirming NaCl structure type, inappropriate for Rietveld use, BVS validation or ML training.

4.1.2 Multi-phase refinement products. Check source papers for synthesis context. The 2023 MgS structures (CollCodes 64415, 64422) represent 7–9 wt% phases in Mg–S battery electrodes alongside C, SiC, MgO, $MgSO_4$, MgF_2. Appropriate for electrochemical phase evolution studies; inappropriate as bulk MgS references.

4.1.3 Structures with site disorder $\Delta V \geq 2$. $LiCoO_2$-type materials report diffraction-averaged M–O distances (~2.05 Å) satisfying neither Li^+ (requires ~2.15 Å) nor Co^{3+} (requires ~1.95 Å) bonding requirements. Look for spectroscopic validation (XANES, Mössbauer) confirming local oxidation states before use in bonding-sensitive applications.

4.1.4 Solid solutions at intermediate compositions. $Na_{0.5}K_{0.5}Cl$ exhibits 64% BVS deviation not because parameters are wrong but because diffraction averages over two distinct local environments. Use end-member compositions (NaCl, KCl) for parameter validation; flag intermediate compositions for phase diagram work and general information only, or use with appropriate caution.

Users could consider an example decision tree for structure selection when interacting with a database such as the ICSD:

1. Check measurement decade and technique: pre-1980 film-based determinations often lack refinement quality indicators required for bonding-sensitive applications, whilst modern CCD and synchrotron measurements typically provide adequate precision.

2. Verify *R*-factor < 0.05 and presence of complete anisotropic thermal parameters.

3. Examine composition for mixed occupancy exceeding 10% at any site, as substantial disorder introduces geometric averaging artefacts.

4. Confirm single-phase synthesis by consulting source papers for terms indicating multi-phase refinements ("Rietveld", "multi-phase", "secondary phase").

5. For large-cation oxides ($r > 1.0$ Å: Ca, Sr, Ba, Eu), treat BVS validation failures as expected parameter inadequacy rather than automatic disqualification, provided other quality metrics are acceptable.

4.2 Database curation: implementable metadata enhancements

Specific machine-readable flags would make for much easier automated fitness-for-purpose assessment, particularly if these could also be included in crystallographic information files when downloaded:

4.2.1 Refinement context field. • *primary_structure*: true/false (was this the main subject?)

• *phase_fraction*: 0.00–1.00 (in multi-phase refinements)

• *refinement_constraints*: free/constrained/fixed (were parameters independently refined?)

4.2.2 Measurement provenance field. • *technique*: film/scintillation/CCD/synchrotron/electron_diffraction

• *sample_type*: bulk_single_crystal/bulk_powder/thin_film/nanoparticle/electrode

• *synthesis_context*: equilibrium/metastable/in_situ/high_throughput

4.2.3 Validation confidence field. • *composition_confirmed*: [XPS/XANES/EDX/none]

• *independent_measurements*: N (how many different labs?)

• *literature_agreement*: within_1_percent/within_5_percent/novel

The implementation burden for these changes would be modest: most information exists in source papers but fails to propagate to CIF files. Database deposition forms could require explicit answers for refinement context and sample type, with automated parsing extracting technique from diffractometer metadata.

4.3 Parameter revision priorities and critical experiments

Alkaline earth oxides (CaO, SrO, BaO) alongside europium oxide represent the highest priority for parameter revision. The complete failure of Ca^{2+}–O^{2-} parameters across 15 structures spanning a century (1922–2023) with exceptionally low inter-structure variance (<1%) indicates genuine parameter inadequacy rather than measurement scatter or experimental artefacts. The oxide–chalcogenide validation inversion provides definitive evidence: calcium sulphide validates in 94% of structures whilst calcium oxide fails in 100%, demonstrating that the inadequacy is anion-specific rather than reflecting errors in cation radii or coordination environment detection. The same inversion appears for barium and europium, establishing a systematic pattern that demands parameter revision rather than acceptance of current values.

However, a critical ambiguity remains. No CaS, SrS, or BaS structure has been measured since 2006, creating an 18–44 years gap during which diffractometer technology evolved substantially. The MgS case demonstrates that early measurements can mask inadequacies revealed by improved precision: structures from the 1990s validated consistently, whilst 2023 measurements using modern techniques failed systematically. Modern high-precision synchrotron measurements of alkaline earth sulphides using refinement protocols matching recent MgS work represent the critical experiment distinguishing genuine anion-specific parameter adequacy from measurement precision evolution masking inadequacies in historical data. If 2024 CaS measurements validate whilst contemporaneous CaO measurements fail, anion-specific parameter revision is justified. If both fail, the oxide-sulphide inversion reflects historical measurement limitations rather than fundamental anion-dependent chemistry.

The systematic linear relationship between cation radius and BVS deviation ($R^2 = 0.996$, slope -30% Å^{-1}, $p = 0.002$) is remarkable given that Gagné & Hawthorne fitted each M^{2+}–O^{2-} pair independently using their GRG optimization method rather than imposing systematic scaling relationships. This correlation suggests either that their alkaline earth oxide reference structures contained correlated systematic errors, or that the exponential functional form itself inadequately captures bonding in large polarizable cations paired with small electronegative O^{2-}. The oxide–chalcogenide validation inversion provides the critical diagnostic: calcium, strontium, barium, and europium validate reliably in sulphides but fail systematically in oxides despite identical rock salt coordination. This anion-dependence rules out cation-specific issues (radius errors, polarizability treatments) and indicates anion-specific parameter inadequacy. Whether this reflects inadequacies in the M^{2+}–O^{2-} reference structure set employed during parameterisation or fundamental limitations of the Brown–Altermatt functional form for this specific bonding regime cannot be resolved from rock salt data alone; examination of alkaline earth oxide parameter performance in perovskite and fluorite structures would distinguish structure-type-specific from general functional form failures.

Lanthanide monoxides ($Ln^{2+}O$, where Ln = La, Ce, Pr, Nd, Sm, Eu, Gd, Yb, Lu) merit systematic investigation but present substantial practical challenges. These exotic phases require extreme reducing conditions to stabilise unusual Ln^{2+} oxidation states, where this is even possible, with several of the reported structures originating from DFT-based structure prediction rather than direct experimental synthesis. The current dataset is too sparse ($n = 1$ for most lanthanides) and heterogeneous (mixing experimental determinations with computationally predicted structures) for definitive parameter assessment. However, the observed crossover from negative deviations (early lanthanides) to positive deviations (late lanthanides, particularly YbO and SmO) suggests parameter behaviour transitions within the series that warrant investigation. The lanthanide contraction provides a chemically similar series spanning substantial ionic radius variation (La^{2+} 1.17 Å to Lu^{2+} 0.977 Å in octahedral coordination), potentially enabling systematic probing of size-dependent parameter adequacy. Future work synthesising high-quality $Ln^{2+}O$ reference structures across the complete series using comparable synthetic protocols and characterisation techniques would constrain whether single M^{2+}–O^{2-} parameterisations capture behaviour across the full radius range or whether size-dependent functional forms are required.

Chloride and iodide validation rates (40% and 56% respectively) warrant investigation but represent lower priority relative to alkaline earth oxides. The smaller sample sizes and greater chemical diversity spanning alkali halides, transition metal halides, and mixed-occupancy solid solutions complicate interpretation. Systematic failures in specific systems (LiI, K–Na–Cl solid solutions) may reflect parameter inadequacy, but distinguishing these from size-mismatch effects in compositionally averaged structures requires more detailed analysis than the systematic, reproducible failures observed for alkaline earth oxides.

4.4 Broader implications: data integrity in AI-Era materials science

Recent AI-driven materials discovery initiatives highlight the critical need for systematic database quality assessment. GNoME predicted 2.2 million structures with 0.03% experimental validation, whilst the A-Lab autonomous synthesis platform claimed 71% success (41 novel materials from 58 targets)

yet independent analysis revealed likely zero genuine discoveries, attributed to inadequate structural characterization alongside computational prediction errors.[10,40,41] These failures underscore that training data quality and human oversight, not merely prediction algorithms, determines validation success. The present work demonstrates that even for simple rock salt structures under ambient conditions, only 49.5% validate through bond valence sum analysis, with failures stratifying into systematic parameter inadequacy (17.1%), methodological limitations (10.5%), and database quality issues (2.4%). Every structure exhibiting 20% BVS deviation without documented cause potentially introduces systematic error into ML training data. As autonomous laboratories and high-throughput computation accelerate structure deposition into databases, automated validation tools providing rapid quality assessment become essential infrastructure for maintaining data integrity in AI-driven workflows.

Tools like Eir can enable non-experts to access quality assessment previously requiring specialist knowledge, but uncritical tool application risks replacing one problem (insufficient quality checking) with another (mechanical filtering lacking chemical understanding). The failure mode classification (parameter inadequacy/missing parameters/vacancy disorder/heterovalent disorder/solid solutions/database contamination) represents an educational framework, not just automated sorting. Users must understand WHY structures 'fail' to make informed decisions about appropriateness.

Automated quality assessment scales to datasets spanning hundreds of thousands of structures – manual curation demonstrably cannot. The complete traceability achieved here (no unexplained outliers) demonstrates that systematic computational screening successfully identifies issues requiring attention without generating false positives, provided failure modes are properly characterized rather than treated as pass/fail thresholds.

5 Conclusions

Automated bond valence sum analysis of 840 rock salt structures achieved complete traceability: every structure either validated (49.5%) or exhibited deviations attributable to documented parameter inadequacy (17.1%), methodological inappropriateness (10.5%), database quality issues (2.4%), or borderline measurement precision (14.0%). This transforms BVS from a pass–fail filter into a diagnostic tool that successfully identifies what requires attention and why, while also showing some of the limitations of BVS analysis.

Three critical findings challenge assumptions about crystallographic database quality. First, alkaline earth oxides exhibit systematic parameter inadequacy (CaO, SrO, BaO, EuO: 100% failure, <1.5% inter-structure variance across measurements spanning 1922–2023), demonstrating that conventional parameters for seemingly simple binary oxides require revision. Second, the oxide–chalcogenide validation inversion – where Ba, Ca, and Eu validate in 94–100% of sulphide structures

whilst failing in 100% of oxide structures – provides definitive evidence that parameter adequacy depends on both cation and anion identity. An 18–44 years gap in alkaline earth sulphide measurements prevents distinguishing genuine anion-specific adequacy from measurement precision evolution, making modern high-precision sulphide determinations the critical experiment. Third, database contamination from multi-phase refinements (MgS refined as 7.5% minority phase in battery electrodes, NiO as <1% impurity in cathode materials) demonstrates that structures passing peer review may prove inappropriate as computational reference structures without enhanced fitness-for-purpose metadata.

The convergence of high-throughput crystallography and AI-driven materials discovery makes such quality assessment critically urgent. GNoME predicted 2.2 million compositions with 0.03% experimental validation; whilst prediction algorithms contribute substantially to this low rate, training data quality represents one correctable factor. Every structure exhibiting 20% BVS deviation without documented cause potentially introduces systematic error into ML models predicting bond lengths, formation energies, or ionic conductivities. Bond valence sum analysis coupled with intelligent confidence classification provides one essential component of the multi-method validation frameworks required to maintain database integrity as crystallographic repositories serve AI-driven materials discovery at scale.

The open-source Eir tool assists this assessment at database scale (1–2 structures per second), with extensions to spinel, perovskite, and layered structures underway to test whether rock salt patterns – particularly the oxide parameter inadequacies and heterovalent disorder limitations – generalise across diverse coordination environments. The demonstrated ability to classify structures according to failure mode, combined with implementable metadata enhancements for database curation, establishes systematic quality assessment at AI-relevant scales.

Author contributions

NRM produced all contributions to this research paper: coding, analyses, interpretations and writing.

Conflicts of interest

There are no conflicts to declare.

Data availability

The Eir 1.2 software for automated bond valence sum analysis is freely available under an open-source MIT licence at https://github.com/SuperBladesman/Eir with comprehensive documentation and installation instructions. Processed validation results for all 686 assessable rock salt structures are provided as six SI CSV data files: (S1) complete structure listings with

validation status and failure mode classifications (612 structures); (S2) structures excluded due to missing BVS parameters (154 structures); (S3) Type 1 systematic parameter inadequacy cases (107 structures); (S4) Types 3–5 methodological limitations arising from structural disorder (95 structures); (S5) Type 6 database quality issues (13 structures); and (S6) borderline structures requiring manual review (79 structures). These machine-readable files enable complete reproduction of manuscript statistics and support computational workflows for database quality assessment. Source crystallographic data (CIF files) were obtained from the Inorganic Crystal Structure Database (ICSD, https://icsd.fiz-karlsruhe.de, access requires institutional subscription); ICSD collection codes for all analysed structures are provided in the SI data files. Complete analysis workflows, Python processing scripts, and extended documentation are available in the GitHub repository.

Supplementary information (SI) is available. See DOI: https://doi.org/10.1039/d5dt03003j.

References

1 A. K. Cheetham and R. Seshadri, Artificial intelligence driving materials discovery? Perspective on the article 'Scaling deep learning for materials discovery', *Nat. Mater.*, 2024, **23**(1), 15–16.

2 N. Reeves-McLaren and S. Moth-Lund Christensen, Data integrity in materials science in the era of AI: balancing accelerated discovery with responsible science and innovation, *J. Mater. Chem. A*, 2026, **14**, 276–283, DOI: 10.1039/D5TA05512A.

3 A. Belsky, *et al.*, New developments in the Inorganic Crystal Structure Database (ICSD): accessibility in support of materials research and design, *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.*, 2002, **58**(3), 364–369.

4 D. Bimler, in *Cryptocrystallography: When X-ray diffraction suggests imaginary molecules and protein structures*, 2022, DOI: 10.21203/rs.3.rs-1537438/v1.

5 U. Christmann, The State of Accurate Mass Data in the Supporting Information Files of Organic Letters Publications, *Org. Lett.*, 2024, **27**(1), 4–7.

6 M. C. Kozlowski, *et al.*, Editorial: Importance of Accurate Mass Data in Supporting Information, *Org. Lett.*, 2025, **27**(1), 1–3.

7 G. H. Major, *Detection and Mitigation of Faulty Data, Data Analysis Errors, and Reporting Practices Within X-ray Photoelectron Spectroscopy Literature*, Brigham Young University, 2023.

8 G. H. Major, *et al.*, Assessment of the Frequency and Nature of Erroneous X-ray Photoelectron Spectroscopy Analyses in the Scientific Literature, *J. Vac. Sci. Technol., A*, 2020, **38**(6), 003.

9 D. R. Baer, *et al.*, XPS guide: Charge neutralization and binding energy referencing for insulating samples, *J. Vac. Sci. Technol., A*, 2021, **39**(1), 013204.

10 N. J. Szymanski, *et al.*, An autonomous laboratory for the accelerated synthesis of novel materials, *Nature*, 2023, **624**, 86–91.

11 I. D. Brown, Recent Developments in the Methods and Applications of the Bond Valence Model, *Chem. Rev.*, 2009, **109**(12), 6858–6919.

12 I. D. Brown, *The Chemical Bond in Inorganic Chemistry: The Bond Valence Model*, Oxford University Press, Oxford, 2016.

13 I. D. Brown and D. Altermatt, Bond-valence parameters obtained from a systematic analysis of the Inorganic Crystal Structure Database, *Acta Crystallogr., Sect. B:Struct. Sci.*, 1985, **41**(4), 244–247.

14 A. Prodi, *et al.*, Charge, orbital and spin ordering phenomena in the mixed valence manganite $(NaMn^{3+}_3)(Mn^{3+}_2Mn^{4+}_2)O_{12}$, *Nat. Mater.*, 2004, **3**(1), 48–52.

15 J. L. Tallon, *et al.*, Generic superconducting phase behavior in high-Tc cuprates: Tc variation with hole concentration in $YBa_2Cu_3O_{7-x}$, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1995, **51**(18), 12911–12914.

16 J. M. Goodman, *et al.*, Experimental Data Checker: Better information for organic chemists, *Org. Biomol. Chem.*, 2004, **2**(21), 3067–3070.

17 C. J. Williams, *et al.*, MolProbity: More and better reference data for improved all-atom structure validation, *Protein Sci.*, 2018, **27**(1), 293–315.

18 A. Spek, Single-crystal structure validation with the program PLATON, *J. Appl. Crystallogr.*, 2003, **36**(1), 7–13.

19 R. W. W. Hooft, *et al.*, Errors in protein structures, *Nature*, 1996, **381**, 272.

20 A. Jain, *et al.*, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation, *APL Mater.*, 2013, **1**(1), 011002.

21 O. C. Gagné and F. C. Hawthorne, Comprehensive derivation of bond-valence parameters for ion pairs involving oxygen, *Acta Crystallogr., Sect. B:Struct. Sci., Cryst. Eng. Mater.*, 2015, **71**(5), 562–578.

22 O. C. Gagné and F. C. Hawthorne, Bond-length distributions for ions bonded to oxygen: Results for the transition metals and quantification of the factors underlying bond-length variation in inorganic solids, *Acta Crystallogr., Sect. B:Struct. Sci., Cryst. Eng. Mater.*, 2016, **72**(4), 602–625.

23 I. D. Brown, in *Accumulated List of Bond Valence Parameters*, 2020. Available from: https://www.iucr.org/resources/data/datasets/bond-valence-parameters.

24 F. Bosi, Bond valence at mixed occupancy sites. I. Regular polyhedra, *Acta Crystallogr., Sect. B:Struct. Sci., Cryst. Eng. Mater.*, 2014, **70**(5), 864–870.

25 I. Brown, On measuring the size of distortions in coordination polyhedra, *Acta Crystallogr., Sect. B:Struct. Sci.*, 2006, **62**(4), 692–694.

26 L. Pauling, The principles determining the structure of complex ionic crystals, *J. Am. Chem. Soc.*, 1929, **51**(4), 1010–1026.

27 J. R. Hester, A validating CIF parser: PyCIFRW, *J. Appl. Crystallogr.*, 2006, **39**(4), 621–625.

28 M. Wojdyr, GEMMI: A library for structural biology, *J. Open Source Softw.*, 2022, **7**(73), 4200.

29 N. Rego and D. Koes, 3Dmol.js: molecular visualization with WebGL, *Bioinformatics*, 2015, **31**(8), 1322–1324.

30 W. P. Davey and E. O. Hoffmann, The cubic form of some ions, *Phys. Rev.*, 1920, **15**, 333.

31 S. Holgersson, Structure of the sulfides of Mg, Ca, Se, and Ba, *Z. Anorg. Allg. Chem.*, 1923, **126**, 179.

32 H. A. Eick, N. C. Baenziger and L. Eyring, Lower oxides of samarium and europium. The preparation and crystal structure of SmO0.4–0.6, SmO and EuO, *J. Am. Chem. Soc.*, 1956, **78**(20), 5147–5149.

33 V. S. Urusov and V. V. Blinov, Defect characteristics in solid solutions NaCl-AgCl, *Izv. Akad. Nauk SSSR, Ser. Khim.*, 1970, **2**, 278–282.

34 M. Ahtee, Lattice constants of some binary alkali halide solid solutions, *Ann. Acad. Sci. Fenn., Ser. A6*, 1969, **313**, 1–11.

35 J. Kumar and R. Saxena, Formation of NaCl- and Cu_2O-type oxides of platinum and palladium of carbon and alumina support films, *J. Less-Common Met.*, 1989, **147**, 59–71.

36 L. Boulet-Roblin, *et al.*, Operando Neutron Powder Diffraction Using Cylindrical Cell Design: The Case of $LiNi_{0.5}Mn_{1.5}O_4$ vs Graphite, *J. Phys. Chem. C*, 2016, **120**(31), 17268–17273.

37 X. Shan, *et al.*, Framework Doping of Ni Enhances Pseudocapacitive Na-Ion Storage of $(Ni)MnO_2$ Layered Birnessite, *Chem. Mater.*, 2019, **31**(21), 8774–8786.

38 E. M. Sheha, *et al.*, Positron Annihilation Spectroscopy as a Diagnostic Tool for Probing the First-Cycle Defect Evolution in Magnesium–Sulfur Battery Electrodes, *Phys. Status Solidi A*, 2023, **220**(3), 2200661.

39 J. M. Leger, *et al.*, *Chimie minérale. Synthèse de YbO et YbO sous haute pression*, in *Comptes Rendus des Seances de l'Academie des Sciences, Serie C: Sciences Chimiques*, 1978, vol. 286.

40 A. Merchant, *et al.*, Scaling deep learning for materials discovery, *Nature*, 2023, **624**, 80–85.

41 J. Leeman, *et al.*, Challenges in High-Throughput Inorganic Materials Prediction and Autonomous Synthesis, *PRX Energy*, 2024, **3**(1), 011002.