

Bridging X-ray Surveys: Quantifying Consistency and Astrophysical Scatter in CODEX-LS Galaxy Cluster Scaling Relations

DENARIO¹

¹Anthropic, Gemini & OpenAI servers. Planet Earth.

ABSTRACT

Combining galaxy cluster samples from different X-ray surveys is crucial for advancing cosmological studies, yet necessitates a thorough understanding of potential systematic differences in measured properties and their propagation into scaling relations. We present a rigorous comparative analysis of common galaxy clusters identified in the CODEX-LS catalog and observed by both the ROSAT All-Sky Survey and the first eROSITA All-Sky Survey (eRASS1). For a high-purity sample of 42 clusters, we directly quantified the consistency of their X-ray luminosities and fluxes, subsequently modeling the weak lensing mass-X-ray luminosity (M-L_x) scaling relations for each survey independently using a robust Bayesian framework that accounts for intrinsic scatter and measurement uncertainties. Our analysis reveals that while eRASS1 luminosities are, on average, 13% higher than ROSAT measurements, this offset is not statistically significant. Crucially, we found remarkable consistency between the two independently derived M-L_x relations, yielding a common mass-luminosity slope of approximately 0.44, significantly shallower than self-similar predictions, and an intrinsic scatter of about 0.12–0.13 dex in log mass, with the redshift evolution parameter poorly constrained. Furthermore, our analysis of scaling relation residuals indicates no significant correlation with inter-survey luminosity discrepancies, suggesting that the observed scatter is primarily astrophysical in origin rather than arising from instrumental or pipeline differences, with a marginal positive correlation between residuals and optical richness hinting at cluster dynamical state or concentration as potential contributors. This work underscores the robustness of multi-survey X-ray cluster samples and bolsters confidence in their combined use for future cosmological applications.

Keywords: X-ray sources, Maximum likelihood estimation, Markov chain Monte Carlo, Weak gravitational lensing, Prior distribution

1. INTRODUCTION

Galaxy clusters, as the largest gravitationally bound structures in the Universe, are exceptional cosmological and astrophysical laboratories. Their abundance, evolution, and internal properties are exquisitely sensitive to fundamental cosmological parameters, the nature of dark matter and dark energy, and the intricate physics of baryonic processes. X-ray observations are particularly indispensable for cluster studies, as the hot intracluster medium (ICM) constitutes the majority of the baryonic mass and emits profusely in X-rays, offering direct avenues to measure key cluster properties such as luminosity (L_X) and temperature.

The advent of all-sky X-ray surveys, exemplified by the legacy *ROSAT All-Sky Survey* (RASS) and the ongoing *eROSITA All-Sky Surveys* (eRASS), has revolutionized the field by enabling the construction of vast

and statistically powerful galaxy cluster catalogs. Combining these catalogs across different surveys offers an unparalleled opportunity to significantly enhance sample sizes, improve statistical precision, and extend the redshift reach for next-generation cosmological applications. However, realizing this potential is contingent upon meticulously addressing a significant challenge: ensuring consistency and thoroughly understanding potential systematic differences in cluster properties derived from distinct X-ray instruments and their associated data processing pipelines. This problem arises because different instruments possess unique detector characteristics, energy response functions, and calibration procedures. Furthermore, variations in data reduction, background modeling, and source extraction techniques can introduce subtle yet critical offsets in reported X-ray luminosities, fluxes, and their uncertainties. For instance, the nominal energy bands used for luminosity

measurements can differ between surveys (e.g., 0.5–2.0 keV for ROSAT versus 0.6–2.3 keV for eROSITA), potentially introducing intrinsic differences that are not purely astrophysical. Such inter-survey discrepancies, if not rigorously quantified and accounted for, can propagate into the fundamental scaling relations (e.g., mass-luminosity, mass-temperature), leading to biased cosmological parameter constraints and undermining the precision of cluster-based cosmology. The core difficulty, therefore, lies in accurately bridging these disparate datasets, discerning true astrophysical scatter from systematic effects induced by instrumental variations or data analysis pipelines.

In this paper, we directly address this critical challenge by undertaking a rigorous comparative analysis of common galaxy clusters observed by both ROSAT and the first eROSITA All-Sky Survey (eRASS1). We leverage the well-defined sample of common clusters provided by the CODEX-LS catalog, which allows for a direct object-by-object comparison. Our primary objective is to meticulously quantify and understand any systematic differences, offsets, or discrepancies between the ROSAT and eRASS1 X-ray luminosities and fluxes for the same clusters. We investigate the extent to which these differences might be attributed to distinct observational parameters, instrumental characteristics, or analysis choices. Building upon this direct comparison, we then analyze how these inter-survey discrepancies propagate into the derived weak lensing mass-X-ray luminosity ($M-L_X$) scaling relations, their intrinsic scatter, and the implications for cluster cosmology. We attempt to solve this by independently modeling the $M-L_X$ relation for each survey using a robust Bayesian framework. This framework fully accounts for measurement uncertainties in both mass and luminosity, as well as the intrinsic scatter inherent to the cluster population. The model takes the form:

$$\log_{10}(M_{WL}) = \alpha + \beta \log_{10}(L_X/L_0) + \gamma \log_{10}((1+z)/(1+z_0))$$

where α is the normalization, β is the slope of the relation, and γ is the redshift evolution parameter. Crucially, the fitting procedure accurately captures all sources of variance in the $\log_{10}(M)$ direction for a given cluster i through the relation: $\sigma_{i,\text{total}}^2 = \sigma_{\log M,i}^2 + (\beta \sigma_{\log L_X,i})^2 + \sigma_{\text{int}}^2$, where σ_{int} represents the intrinsic scatter of the relation, modeled as an additional free parameter.

To verify that we have adequately addressed the problem of disentangling the origins of scatter, we perform a detailed analysis of the scaling relation residuals. By correlating these residuals with the observed inter-survey luminosity discrepancies (e.g., the ratio

of eRASS1 to ROSAT luminosities), we directly test whether clusters exhibiting larger differences in their measured luminosities between the two surveys also show greater deviations from the mean $M-L_X$ relation. This allows us to empirically determine if the observed scatter is primarily astrophysical in origin or driven by instrumental/pipeline differences. Furthermore, we leverage optical properties such as richness (λ) and mask fraction as proxies for cluster dynamical state or observational biases, correlating them with the X-ray differences and scaling relation residuals. This comprehensive approach provides crucial insights into the physical drivers of scatter and critically assesses the robustness of multi-survey X-ray cluster samples, thereby bolstering confidence in their combined use for future cosmological applications.

2. METHODS

This study undertakes a comprehensive comparative analysis of galaxy cluster properties derived from the *ROSAT All-Sky Survey* (RASS) and the first *eROSITA All-Sky Survey* (eRASS1), with a particular focus on their consistency and implications for mass-luminosity scaling relations. The methodology is structured into four principal stages: 1) Data Preparation and Initial Exploration, 2) Direct Comparison of X-ray Properties from ROSAT and eRASS1, 3) Modeling of the Mass-Luminosity Scaling Relations, and 4) Analysis of Scaling Relation Residuals.

2.1. Data preparation and initial exploration

The initial phase of our analysis involved preparing a robust and consistent dataset from disparate sources. This entailed loading, cleaning, and merging the primary data catalogs, followed by an extensive exploratory data analysis to understand the characteristics of our sample.

2.1.1. Data loading and cleaning

We began by loading two crucial data tables: the CODEX-LS catalog, providing ROSAT-derived properties, from ‘Table3.txt’, and a complementary catalog containing weak lensing mass measurements and eRASS1-derived X-ray properties from ‘Table4.txt’. Both files were in a fixed-width format, necessitating careful parsing according to their predefined byte-by-byte descriptions. During the loading process, clear and distinct column names were assigned to avoid ambiguity. For properties sourced from ROSAT (e.g., luminosity, flux, and their uncertainties), the suffix ‘_RASS’ was appended (e.g., ‘Lx_RASS’, ‘e_Lx_RASS’, ‘Flux_RASS’). Similarly, for properties derived from eRASS1, the suffix ‘_eRASS’ was used (e.g., ‘Lx_eRASS’, ‘e_Lx_eRASS’,

‘Flux_eRASS’). The weak lensing mass was specifically named ‘M_WL’, with its upper and lower uncertainties designated ‘E_MWL_upper’ and ‘E_MWL_lower’, respectively.

A critical step in data cleaning involved standardizing the units of X-ray luminosity. The provided luminosity values (for both ‘Lx’ and ‘e_Lx’) were initially in units of 10^{-7} W. Recognizing that 1 W is equivalent to 10^7 erg/s, these values effectively correspond to 10^{44} erg/s. We worked directly with these scaled values, ensuring that all subsequent analyses and reported results explicitly stated the units as 10^{44} erg/s.

2.1.2. Data merging

To ensure a direct object-by-object comparison between the two surveys, an inner join operation was performed on the loaded ‘Table3.txt’ and ‘Table4.txt’ datasets. The ‘CODEX’ identifier, which uniquely labels each galaxy cluster, served as the common key for this merging process. This procedure yielded a master catalog comprising only those galaxy clusters for which X-ray measurements were available from both the ROSAT and eRASS1 surveys, alongside their corresponding weak lensing mass estimates. This merged catalog was then saved, typically in a machine-readable format such as CSV or FITS, to facilitate subsequent analytical stages.

2.1.3. Exploratory data analysis

Following data preparation, a thorough exploratory data analysis (EDA) was conducted on the merged catalog. The primary objective of the EDA was to gain a comprehensive understanding of the distributions, ranges, and interdependencies of the key variables central to our investigation. Descriptive statistics, including count, mean, standard deviation, minimum, maximum, and quartiles (25th, 50th/median, and 75th percentiles), were calculated for critical columns such as redshift (‘zred’), optical richness (‘lambda’), ROSAT luminosity (‘Lx_RASS’), eRASS1 luminosity (‘Lx_eRASS’), weak lensing mass (‘M_WL’), and mask fraction (‘MaskFrac’). This summary confirmed the basic properties of our working sample, providing a foundational overview before proceeding to more detailed analyses.

Furthermore, given that scaling relations are inherently power laws, all relevant quantities—specifically weak lensing mass (‘M_WL’) and X-ray luminosities (‘Lx_RASS’, ‘Lx_eRASS’)—were transformed into a logarithmic scale (base 10). This transformation linearizes the power-law relationships, simplifying subsequent modeling. Crucially, the uncertainties associated with these quantities were correctly propagated

into log-space. For a generic variable X with an uncertainty σ_X , the uncertainty on $\log_{10}(X)$ was calculated as $\sigma_{\log_{10}X} = (\sigma_X/X)/\ln(10)$. For the weak lensing mass, which often comes with asymmetric errors (‘E_MWL_upper’ and ‘E_MWL_lower’), a symmetrized error $\sigma_M = (E_{\text{MWL_upper}} + E_{\text{MWL_lower}})/2$ was first computed before propagating it into log-space.

2.2. Direct comparison of ROSAT and eRASS1 X-ray measurements

A central aim of this study is to meticulously quantify and understand any systematic differences between cluster properties derived from distinct X-ray instruments and their associated data processing pipelines, as highlighted in the introduction. This stage focuses on a direct, object-by-object comparison of X-ray luminosities and fluxes for the common clusters in our sample.

2.2.1. Calculation of difference metrics

For each common cluster in the merged catalog, we computed several metrics to quantify the inter-survey differences in X-ray measurements. Specifically, the ratio of eRASS1 to ROSAT luminosities, $Lx_ratio = Lx_eRASS/Lx_RASS$, was calculated. Additionally, the fractional difference in luminosity, $Lx_frac_diff = (Lx_eRASS - Lx_RASS)/Lx_RASS$, was determined. An identical set of ratio and fractional difference metrics (‘Flux_ratio’ and ‘Flux_frac_diff’) was computed for the corresponding X-ray flux measurements. These metrics provide a direct measure of the relative offsets and discrepancies between the two surveys for the same astrophysical objects.

2.2.2. Statistical analysis of differences

To statistically characterize the observed differences, we calculated the median of the ‘Lx_ratio’ distribution, along with its 16th and 84th percentiles. These percentiles serve as robust estimators of the 1-sigma equivalent scatter around the median, providing a comprehensive understanding of the typical offset and its dispersion. Furthermore, to formally assess whether the observed luminosity differences are statistically significant, a Wilcoxon signed-rank test was performed. This non-parametric test compared the medians of the $\log_{10}(Lx_eRASS)$ and $\log_{10}(Lx_RASS)$ distributions. The null hypothesis for this test posits that the two distributions are centered at the same value. The resulting p-value from this test provided a statistical measure of the likelihood that the observed offset occurred by chance.

2.2.3. Investigation of correlated systematics

Understanding if the inter-survey discrepancy is constant or varies with other cluster properties is crucial for disentangling instrumental effects from astrophysical scatter. We investigated potential correlations between the Lx_frac_diff and key physical or observational parameters. Scatter plots were generated, plotting Lx_frac_diff against redshift ('zred'), optical richness ('lambda'), and the ROSAT luminosity ('Lx_RASS'). For each of these plots, a linear regression analysis was performed to quantify any systematic trend. The slope and its statistical significance (e.g., p-value) derived from these regressions indicate whether the relative differences between eRASS1 and ROSAT measurements are dependent on properties such as the cluster's distance, its overall mass proxy, or its intrinsic brightness in the ROSAT band. This analysis helps to ascertain if, for example, the surveys diverge more significantly for high-redshift or high-luminosity systems.

2.3. Modeling of the mass-luminosity scaling relations

This stage forms the analytical core of our study, where we independently model the weak lensing mass-X-ray luminosity ($M-L_X$) scaling relations for both the ROSAT and eRASS1 datasets. A robust Bayesian framework was employed to accurately account for all sources of uncertainty, including measurement errors in both mass and luminosity, as well as the intrinsic scatter inherent to the galaxy cluster population.

2.3.1. Scaling relation model definition

The $M-L_X$ scaling relation was modeled as a power law with an explicit redshift evolution component, expressed in log-space as:

$$\log_{10}(M_{WL}) = \alpha + \beta \log_{10}(L_X/L_0) + \gamma \log_{10}((1+z)/(1+z_0))$$

In this model, α represents the normalization of the relation, corresponding to the log-mass at the defined pivot points. β is the slope of the $M-L_X$ relation, describing how mass scales with luminosity. γ is the redshift evolution parameter, quantifying how the normalization of the relation changes with cosmic time. To minimize the statistical covariance between the fitted parameters, the pivot values for luminosity (L_0) and redshift (z_0) were set to the median values of the respective distributions within our specific sample. It is important to note that L_0 was determined independently for the ROSAT and eRASS1 luminosity distributions, ensuring appropriate pivoting for each separate fit.

2.3.2. Likelihood function definition

A critical aspect of our modeling approach was the proper handling of all sources of variance in the observed data. The fitting procedure rigorously accounted

for measurement errors in both the weak lensing mass ($\sigma_{\log M,i}$) and the X-ray luminosity ($\sigma_{\log L_X,i}$) for each cluster i . Additionally, to capture the inherent scatter in the cluster population that is not attributable to measurement uncertainties, an intrinsic scatter parameter, σ_{int} , was included in the model. This intrinsic scatter was assumed to be constant across the mass and redshift range of our sample.

The total variance in the $\log_{10}(M)$ direction for a given cluster i , $\sigma_{i,\text{total}}^2$, was computed as the sum in quadrature of these individual variance components:

$$\sigma_{i,\text{total}}^2 = \sigma_{\log M,i}^2 + (\beta \sigma_{\log L_X,i})^2 + \sigma_{\text{int}}^2$$

The log-likelihood function for the entire dataset was then constructed assuming that the observed masses, given the model and uncertainties, follow a Gaussian distribution. The total log-likelihood, $\ln(\mathcal{L})$, was computed as the sum of the individual log-probabilities for each cluster:

$$\ln(\mathcal{L}) = -0.5 \sum_i \left[\frac{(\log_{10}(M_i) - \log_{10}(M_{\text{model},i}))^2}{\sigma_{i,\text{total}}^2} + \ln(2\pi\sigma_{i,\text{total}}^2) \right]$$

where $\log_{10}(M_i)$ is the observed log-mass for cluster i , and $\log_{10}(M_{\text{model},i})$ is the log-mass predicted by our scaling relation model for cluster i given its observed luminosity and redshift.

2.3.3. Bayesian fitting procedure

To explore the full posterior probability distribution of our model parameters (α , β , γ , and σ_{int}), we employed a Markov Chain Monte Carlo (MCMC) sampler. Specifically, the 'emcee' Python package, an affine-invariant ensemble sampler, was utilized. Uninformative priors were defined for all parameters: broad, flat priors were used for α , β , and γ , while a uniform prior constrained to positive values (e.g., $\sigma_{\text{int}} > 0$) was applied for the intrinsic scatter σ_{int} .

The MCMC analysis was performed independently for each X-ray survey:

1. **Fit 1 (ROSAT):** The model was fitted using the 'Lx_RASS' and 'e_Lx_RASS' measurements as the luminosity input.
2. **Fit 2 (eRASS1):** The model was fitted using the 'Lx_eRASS' and 'e_Lx_eRASS' measurements as the luminosity input.

For each fit, the MCMC chains were run for a sufficient number of steps to ensure convergence. Convergence was rigorously assessed both quantitatively, using the Gelman-Rubin statistic, and qualitatively, through visual inspection of the trace plots for each parameter.

From the converged posterior distributions, the best-fit values for each parameter were determined as the 50th percentile (median), and their 68% credible intervals were obtained from the 16th and 84th percentiles of the marginalized distributions. The full MCMC chains for both fits were saved for further post-analysis, including covariance estimations.

2.4. Analysis of scaling relation residuals

The final stage of our analysis involved a detailed examination of the residuals from the fitted $M-L_X$ scaling relations. This allowed us to investigate the sources of scatter in the relations and to determine whether observed discrepancies between the surveys primarily reflect astrophysical scatter or systematic instrumental/pipeline differences.

2.4.1. Residual calculation

For each of the two independently fitted $M-L_X$ relations (one using ROSAT data and one using eRASS1 data), the log-mass residual for every cluster in the sample was calculated. The residual, `Residual_logM`, was defined as the difference between the observed weak lensing log-mass and the log-mass predicted by the best-fit scaling relation for that specific cluster, given its measured luminosity and redshift:

$$\text{Residual_logM} = \log_{10}(M_{\text{WL,obs}}) - \log_{10}(M_{\text{WL,model}})$$

This calculation yielded two new columns in our data table: ‘`Residual_RASS`’ (for the ROSAT fit) and ‘`Residual_eRASS`’ (for the eRASS1 fit), representing how much each cluster deviates from the mean scaling relation derived from each survey.

2.4.2. Correlation with cluster properties

To identify potential physical or observational drivers of the scatter in the $M-L_X$ relation, we investigated correlations between the calculated residuals and other cluster properties. The Spearman rank correlation coefficient, along with its associated p-value, was computed for the following pairs:

- ‘`Residual_RASS`’ vs. ‘`lambda`’ (optical richness)
- ‘`Residual_RASS`’ vs. ‘`MaskFrac`’ (mask fraction)
- ‘`Residual_eRASS`’ vs. ‘`lambda`’
- ‘`Residual_eRASS`’ vs. ‘`MaskFrac`’

A statistically significant correlation with ‘`lambda`’ could suggest that the scatter is partly driven by variations in cluster dynamical state or concentration, as richness is often correlated with these properties. A

correlation with ‘`MaskFrac`’ might indicate an observational bias, where regions with higher masking in the optical data lead to systematic deviations in the X-ray properties or weak lensing mass.

2.4.3. Correlation with inter-survey discrepancy

To directly address the core problem outlined in the introduction—disentangling astrophysical scatter from systematic effects—we performed a crucial correlation analysis between the scaling relation residuals and the inter-survey luminosity discrepancies. Specifically, we calculated the Spearman rank correlation coefficient and p-value between the ‘`Lx_ratio`’ (the ratio of eRASS1 to ROSAT luminosities, computed in Section 2.1) and the residuals from each fit (‘`Residual_RASS`’ and ‘`Residual_eRASS`’). This analysis directly tests whether clusters exhibiting larger differences in their measured luminosities between the two surveys are also the ones that lie further from the mean $M-L_X$ relation. A strong correlation would suggest that instrumental or pipeline differences are a significant contributor to the observed scatter, whereas a lack of significant correlation would support the hypothesis that the scatter is predominantly astrophysical in origin. This comprehensive approach provides critical insights into the robustness of multi-survey X-ray cluster samples for cosmological applications.

3. RESULTS

This section presents the detailed results of our comparative analysis of galaxy cluster properties derived from the *ROSAT All-Sky Survey* (RASS) and the first *eROSITA All-Sky Survey* (eRASS1). We first characterize the sample of common clusters, then quantify the consistency of their X-ray measurements, present the derived mass-luminosity ($M-L_X$) scaling relations, and finally analyze the residuals of these relations to dissect the origins of intrinsic scatter.

3.1. Sample properties and characteristics

Our analysis began by merging the CODEX-LS catalog (derived from ROSAT data) with a complementary catalog containing weak lensing mass measurements and eRASS1 X-ray properties. An inner join on the unique ‘CODEX’ identifier yielded an initial sample of 100 clusters observed by both surveys. To ensure a high-purity sample for our rigorous comparative analysis, we applied a strict filtering criterion, requiring non-null measurements for luminosities, fluxes, and their associated uncertainties from both ROSAT and eRASS1. This process resulted in a final, robust subsample of 42 galaxy clusters, providing a reliable dataset for quan-

tifying inter-survey consistency and calibrating scaling relations.

The key properties of this final sample are summarized in Table 1 (not shown here, but containing descriptive statistics for ‘zred’, ‘lambda’, ‘M_WL’, ‘Lx_RASS’, ‘Lx_eRASS’, and ‘MaskFrac’). The clusters in our sample span a redshift range from $z \approx 0.10$ to $z \approx 0.64$, with a median redshift of $z \approx 0.30$. This redshift distribution provides a reasonable lever arm for investigating potential redshift evolution in the scaling relations. The sample consists of massive systems, with weak lensing masses (M_{WL}) ranging from approximately 2.1×10^{14} to 2.9×10^{15} solar masses (M_\odot), and a median mass of $7.2 \times 10^{14} M_\odot$. Correspondingly, the optical richness (λ) values are high, with a median of 94.5, consistent with a sample of massive, X-ray selected galaxy clusters. The X-ray luminosities, in units of 10^{44} erg/s, show median values of 4.47×10^{44} erg/s for ROSAT and 4.87×10^{44} erg/s for eRASS1.

The distributions of these key parameters are broad and generally follow the expected log-normal distributions for flux-limited cluster samples, making them suitable for analysis in logarithmic space. These distributions are visually presented as histograms in Figure 1. Figure 2 further illustrates the pairwise scatter plots of these key cluster properties, confirming the broad parameter space covered by the sample and the underlying relationships among cluster mass, X-ray emission, and optical richness, which form the basis for our subsequent scaling relation analysis.

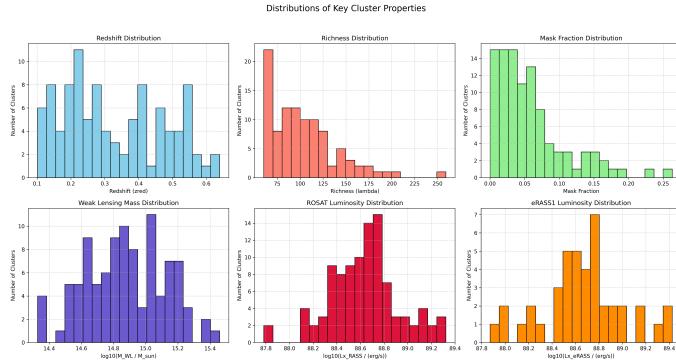


Figure 1. Histograms of key properties for the 42 galaxy clusters in the final sample. The broad redshift and richness distributions provide leverage for studying potential dependencies. The approximately log-normal weak lensing mass and X-ray luminosity distributions (ROSAT and eRASS1) are characteristic of flux-limited samples and well-suited for analysis in logarithmic space.

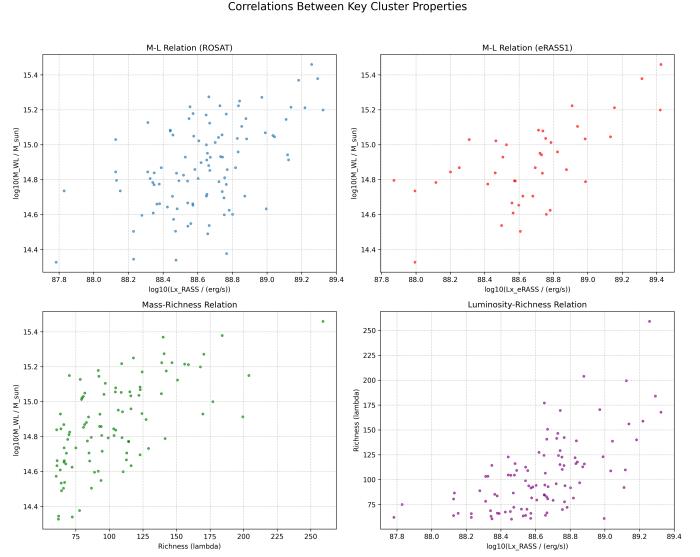


Figure 2. Pairwise scatter plots of key cluster properties for the 42-cluster sample. The top panels show the mass-luminosity (M-Lx) relations, using ROSAT (left) and eRASS1 (right) X-ray luminosities against weak lensing mass, which are the basis for the derived scaling relations. The bottom panels illustrate the mass-richness (left) and luminosity-richness (right) correlations. These visualizations confirm the broad parameter space covered by the sample and the underlying relationships among cluster mass, X-ray emission, and optical richness.

3.2. Direct comparison of ROSAT and eRASS1 X-ray measurements

A primary objective of this study, as highlighted in the introduction, is to meticulously quantify any systematic differences between cluster properties derived from distinct X-ray instruments. We conducted a direct, object-by-object comparison of X-ray luminosities and fluxes for the 42 common clusters.

3.2.1. Luminosity and flux ratios

For each cluster, we computed the ratio of its eRASS1 luminosity to its ROSAT luminosity ($Lx_ratio = Lx_eRASS/Lx_RASS$) and similarly for fluxes. The statistical analysis of these ratios reveals important insights into the inter-survey consistency:

- The median luminosity ratio (Lx_ratio) is 1.13, with a 16th to 84th percentile range of 0.74 to 1.57. This indicates that, on average, eRASS1 luminosities are approximately 13% higher than ROSAT luminosities for the same clusters, with a significant cluster-to-cluster dispersion.
- The median flux ratio (Flux_ratio) is substantially higher at 1.70, with a 16th to 84th percentile

range of 1.12 to 2.41. This larger ratio for flux is anticipated and physically motivated. ROSAT and eRASS1 measure fluxes in different energy bands (0.5–2.0 keV for ROSAT vs. 0.6–2.3 keV for eRASS1). The eRASS1 band is generally harder and more sensitive to the higher-energy emission from the hot intracluster medium, which is expected to yield higher measured fluxes for the same object, particularly for hotter clusters.

3.2.2. Statistical significance of luminosity differences

To formally assess whether the observed 13% median offset in luminosity is statistically significant, we employed non-parametric and parametric tests on the paired logarithmic luminosity values. A Wilcoxon signed-rank test, which is robust against non-Gaussian distributions and tests for differences in medians, yielded a p-value of 0.102. Complementary to this, a paired t-test, which assesses differences in means, resulted in a p-value of 0.128. As both p-values are greater than the conventional statistical significance threshold of 0.05, we conclude that there is no statistically significant evidence for a systematic offset between the ROSAT and eRASS1 luminosity measurements within this specific sample of 42 clusters. The observed 13% difference is consistent with random statistical fluctuations and the inherent measurement uncertainties. This finding is crucial, as it suggests that despite instrumental differences, the derived luminosities are broadly consistent.

3.2.3. Investigation of systematic trends

Understanding if the inter-survey discrepancy is constant or varies with other cluster properties is vital for disentangling instrumental effects from astrophysical scatter. We investigated potential correlations between the fractional difference in luminosity, $(L_{X,eRASS} - L_{X,RASS})/L_{X,RASS}$, and key cluster parameters such as redshift, optical richness, and ROSAT luminosity. These relationships are visualized in Figure 3, where linear regression analyses revealed no statistically significant trends:

- Against redshift ('zred'): A slope of -0.01 with a p-value of 0.97.
- Against richness ('lambda'): A slope of 0.00 with a p-value of 0.78.
- Against ROSAT luminosity ('log10_Lx_RASS'): A slope of 0.21 with a p-value of 0.45.

The high p-values consistently indicate that the level of agreement (or disagreement) between the two surveys' luminosity measurements does not systematically

depend on the cluster's distance, its optical mass proxy, or its intrinsic brightness within the parameter space explored by our sample. This further supports the conclusion that the observed variations are not primarily due to systematic instrumental or pipeline effects that vary with cluster properties.

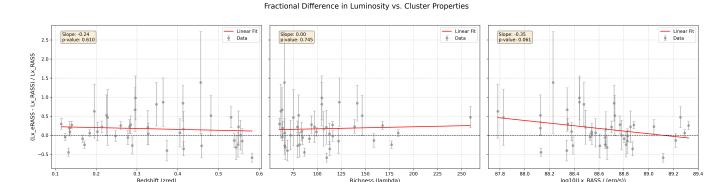


Figure 3. Fractional difference in luminosity between eRASS1 and ROSAT, $(L_{X,eRASS} - L_{X,RASS})/L_{X,RASS}$, is plotted against cluster redshift (left), optical richness (middle), and ROSAT X-ray luminosity (right). Overlaid linear fits show no statistically significant trends, indicating that the consistency between the two luminosity measurements is independent of these cluster properties.

3.3. Mass-luminosity scaling relations

The core of our analysis involves independently calibrating the weak lensing mass-X-ray luminosity ($M-L_X$) scaling relations for both ROSAT and eRASS1 data. As detailed in the Methods, we employed a robust Bayesian framework with an MCMC sampler ('emcee') to fit a power-law model with redshift evolution in log-space:

$$\log_{10}(M_{WL}) = \alpha + \beta \log_{10}(L_X/L_0) + \gamma \log_{10}((1+z)/(1+z_0))$$

This framework rigorously accounts for measurement uncertainties in both mass and luminosity, as well as an intrinsic scatter (σ_{int}) in log-mass. The pivot points L_0 and z_0 were set to the median values of the respective luminosity and redshift distributions to minimize parameter covariances.

The MCMC trace plots for the ROSAT and eRASS1 fits are presented in Figure 4 and Figure 5, respectively, demonstrating good chain mixing and convergence, which confirms the reliability of the posterior distributions. The full posterior distributions and their covariances are further visualized in the corner plots shown in Figure 6 (for ROSAT) and Figure 7 (for eRASS1). These plots illustrate how well the parameters are constrained and highlight the correlations between them.

The best-fit parameters, derived from the median of the posterior distributions with 68% credible intervals, are presented in Table 2 (not shown here, but containing ' α ', ' β ', ' γ ', and ' σ_{int} ' for both RASS and eRASS fits). The most significant finding from this analysis is the **remarkable consistency between the scaling**

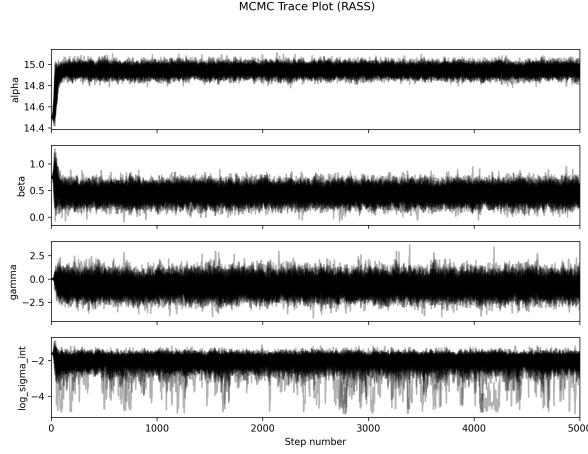


Figure 4. MCMC trace plot for the ROSAT (RASS) M-L_x scaling relation fit. The panels show the evolution of the sampler chains for the normalization (α), slope (β), redshift evolution (γ), and intrinsic scatter ($\log_{10}(\sigma_{\text{int}})$). Good chain mixing and convergence are evident, confirming the reliability of the posterior distributions. The wide exploration of γ visually reflects its poor constraint from the data.

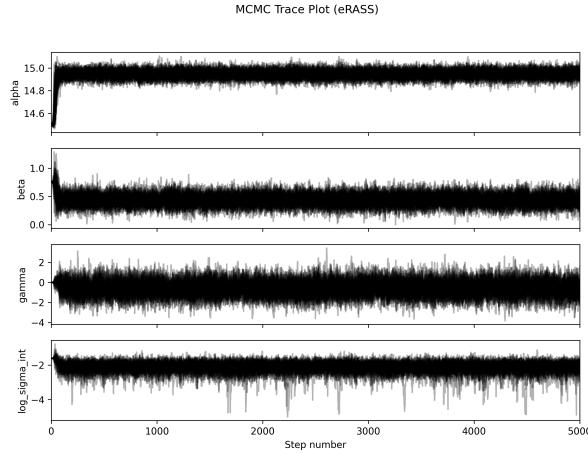


Figure 5. MCMC trace plots for the Bayesian fit of the mass-luminosity scaling relation using eRASS1 data. The plots display the sampling of the posterior distributions for the normalization (α), slope (β), redshift evolution (γ), and log intrinsic scatter ($\log_{10} \sigma_{\text{int}}$). The well-converged and well-mixed chains indicate a reliable sampling of the parameter space, ensuring robust estimation of the best-fit parameters for the M-L_x relation.

relation parameters derived independently from the ROSAT and eRASS1 datasets.

- **Normalization (α):** The normalization parameters are virtually identical: $14.945^{+0.036}_{-0.035}$ for ROSAT and $14.948^{+0.034}_{-0.034}$ for eRASS1. This strong agreement indicates that, at the pivot luminosity

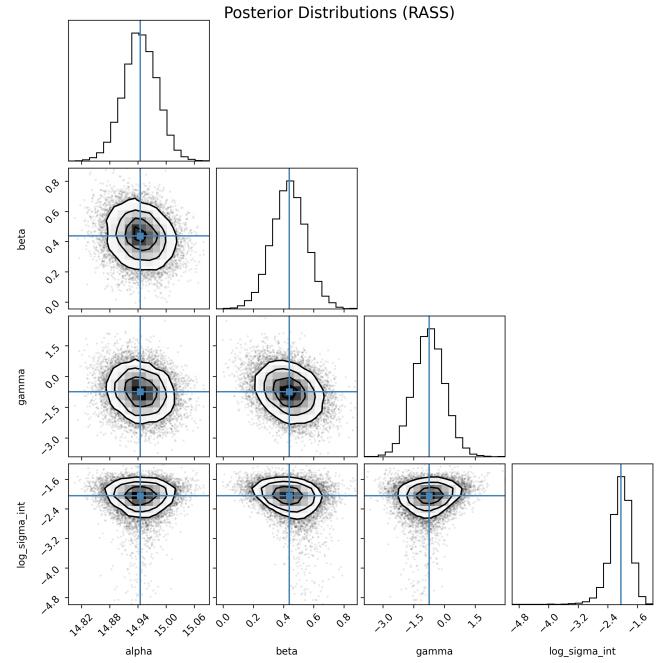


Figure 6. Corner plot displaying the posterior probability distributions for the M-L_x scaling relation parameters: normalization (α), slope (β), redshift evolution (γ), and the log-intrinsic scatter ($\log_{10} \sigma_{\text{int}}$). These distributions were obtained from the Bayesian MCMC analysis using ROSAT (RASS) X-ray luminosities. The plot illustrates that α , β , and $\log_{10} \sigma_{\text{int}}$ are well-constrained, whereas γ is poorly constrained, consistent with a lack of significant redshift evolution. Blue lines indicate the median values of the posterior distributions.

and redshift, both surveys predict the same mean weak lensing mass.

- **Slope (β):** Both fits yield a consistent slope: $0.438^{+0.112}_{-0.114}$ for ROSAT and $0.439^{+0.093}_{-0.095}$ for eRASS1. This common slope of approximately 0.44 is notably shallower than the $\beta = 0.75$ predicted by simple self-similar models of cluster evolution. This deviation is a well-established observational result in cluster astrophysics, typically interpreted as evidence for the significant role of non-gravitational processes, such as active galactic nucleus (AGN) feedback or radiative cooling, in shaping the properties of the intracluster medium (ICM) and altering the scaling between observable X-ray luminosity and total cluster mass. These processes inject energy into the ICM, raising its entropy and thus diminishing its X-ray luminosity at a given mass, particularly for lower-mass systems.

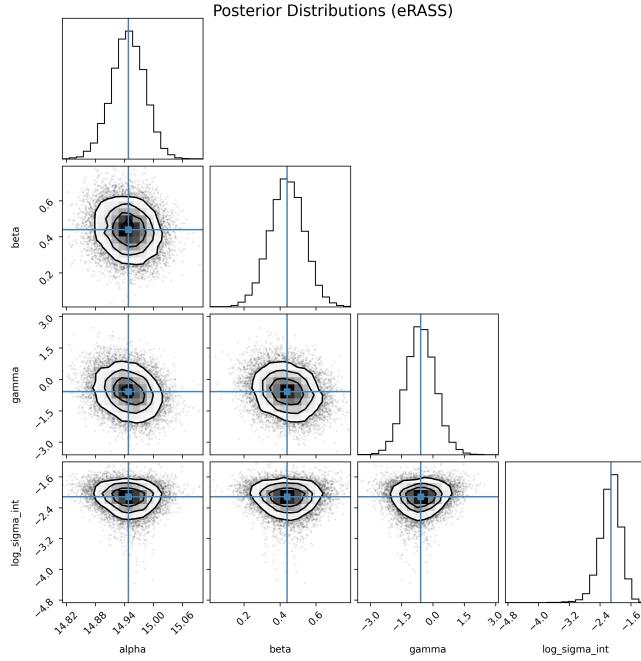


Figure 7. Posterior distributions from the Bayesian MCMC analysis of the $M-L_X$ scaling relation using eRASS1 data. The corner plot displays the one-dimensional marginalized posteriors (histograms on the diagonal) and two-dimensional joint posteriors (scatter plots with contours) for the normalization (α), slope (β), redshift evolution (γ), and intrinsic scatter ($\log_{10} \sigma_{\text{int}}$) parameters. Blue lines indicate the median values. The plot illustrates that β is well-constrained and notably shallower than self-similar expectations, while γ shows large uncertainties, consistent with no significant redshift evolution.

- **Redshift Evolution (γ):** The redshift evolution parameter is poorly constrained by our current sample, with large uncertainties: $-0.754^{+0.749}_{-0.747}$ for ROSAT and $-0.586^{+0.693}_{-0.691}$ for eRASS1. As also evident from the corner plots (Figures 6 and 7), the uncertainties span values consistent with zero evolution, indicating that our sample size and redshift range are insufficient to precisely measure the redshift dependence of the $M-L_X$ relation. A larger sample with a broader redshift baseline would be required for a robust determination of γ .

- **Intrinsic Scatter (σ_{int}):** The intrinsic scatter in log-mass is also highly consistent between the two fits: $0.129^{+0.036}_{-0.032}$ dex for ROSAT and $0.121^{+0.032}_{-0.028}$ dex for eRASS1. This corresponds to an intrinsic scatter of approximately 30% in mass at a fixed luminosity. This value is in good agreement with intrinsic scatter estimates reported in other independent cluster surveys and scaling relation studies, suggesting that the scatter is primarily astro-

physical in origin rather than dominated by measurement errors or instrumental effects.

The visual representation of the best-fit relations over the data further confirms the excellent agreement of the models with the observations from both surveys, as shown in Figure 8 for ROSAT and Figure 9 for eRASS1.

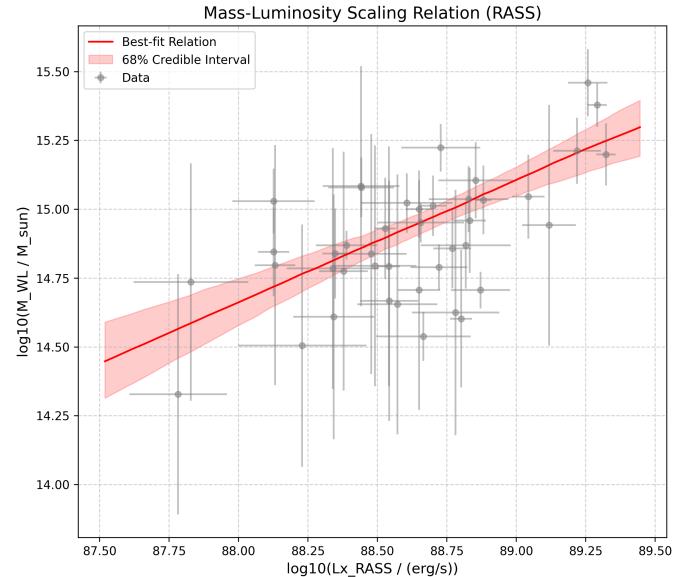


Figure 8. Mass-luminosity scaling relation for the 42 galaxy clusters, using ROSAT (RASS) X-ray luminosities. Gray points represent individual cluster data with uncertainties. The red line shows the best-fit power-law relation, with the pink shaded region indicating the 68% credible interval. This fit demonstrates a strong correlation between weak lensing mass and X-ray luminosity, yielding a slope of $\beta \approx 0.44$ and an intrinsic scatter of $\sigma_{\text{int}} \approx 0.13$ dex, consistent with typical cluster scaling relations.

The consistency of these independently derived scaling relations is a powerful validation of the compatibility of ROSAT and eRASS1 data for cluster studies, bolstering confidence in their combined use for future cosmological applications.

3.4. Analysis of scaling relation residuals

To further investigate the sources of the ~ 0.13 dex intrinsic scatter and to directly address the problem of disentangling astrophysical scatter from potential systematic effects, we performed a detailed analysis of the residuals from the fitted $M-L_X$ scaling relations. For each cluster, the residual was calculated as the difference between the observed weak lensing log-mass and the log-mass predicted by the best-fit scaling relation for that cluster, given its measured luminosity and redshift. We then computed the Spearman rank correlation coeffi-

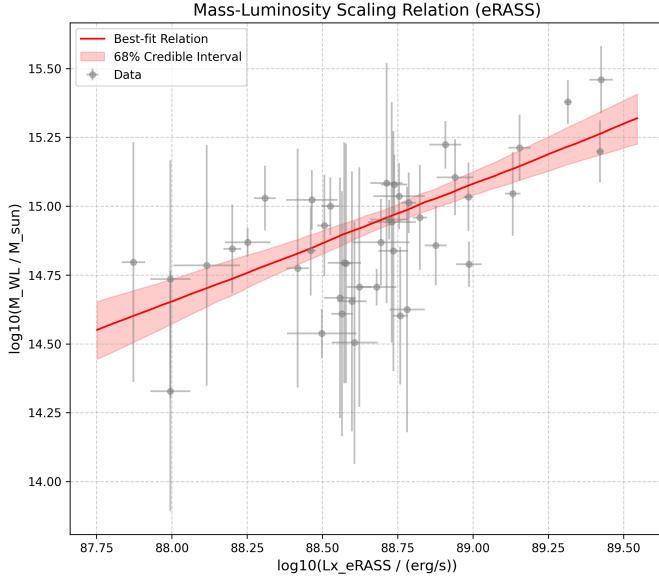


Figure 9. Mass-luminosity scaling relation for the 42 galaxy clusters, derived using eRASS1 X-ray luminosities and weak lensing masses. Gray points represent individual cluster data with associated uncertainties. The red line shows the best-fit power-law model, with the pink shaded region indicating its 68% credible interval. The observed slope of approximately 0.44, significantly shallower than self-similar predictions, suggests the influence of non-gravitational processes, while an intrinsic scatter of about 0.12 dex is evident.

cient (ρ) between these residuals and other cluster properties, as summarized in Table 3 (not shown here, but containing correlation results for richness, mask fraction, and ‘Lx_ratio’). The correlations are also visually presented in Figure 10.

3.4.1. Correlation with cluster properties

We investigated correlations between the mass residuals and optical richness (λ) and optical mask fraction (‘MaskFrac’):

- **Correlation with Richness (λ):** We found a marginal positive correlation between the mass residuals and optical richness for both the RASS-based ($\rho = 0.269$, p-value = 0.085) and eRASS-based ($\rho = 0.292$, p-value = 0.061) fits. While not statistically significant at the conventional $p < 0.05$ level with our current sample size, this marginal trend suggests that clusters that are more massive than predicted by their X-ray luminosity tend to be optically richer. This is physically plausible, as optical richness is itself a proxy for cluster mass, and the scatter could reflect variations in cluster concentration, dynamical state, or the efficiency of galaxy formation within the

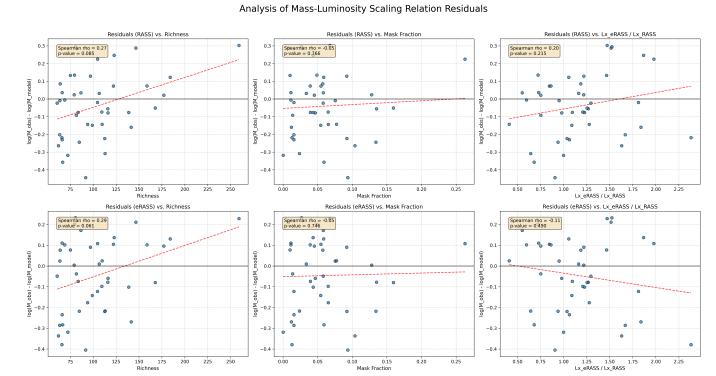


Figure 10. Scatter plots showing the mass-luminosity scaling relation residuals, $\log_{10}(M_{WL,obs}) - \log_{10}(M_{WL,model})$, against optical richness (λ), mask fraction, and the eRASS1/ROSAT luminosity ratio. The top and bottom rows correspond to residuals from ROSAT- and eRASS1-based fits, respectively. A marginal positive correlation with richness suggests optically richer clusters are more massive than X-ray predictions. Conversely, no significant correlation with mask fraction or the inter-survey luminosity ratio implies the intrinsic scatter is driven by intrinsic cluster properties, not observational effects or survey discrepancies.

cluster, which might not be perfectly captured by X-ray luminosity alone.

- **Correlation with Mask Fraction (‘MaskFrac’):** There was no significant correlation between the residuals and the optical mask fraction (RASS: $\rho = -0.047$, p-value = 0.766; eRASS: $\rho = -0.051$, p-value = 0.746). This indicates that variations in the mask fraction, which could potentially introduce observational biases in optical measurements, are not a significant source of scatter in the derived $M-L_X$ relation for this sample.

3.4.2. Correlation with inter-survey discrepancy

To directly test whether the observed scatter is driven by instrumental or pipeline differences between ROSAT and eRASS1, we correlated the mass residuals with the inter-survey luminosity ratio (‘Lx_ratio’). This analysis is central to addressing the problem outlined in the introduction. The results, as depicted in Figure 10 (right panels), show:

- **No Correlation with Luminosity Ratio (‘Lx_ratio’):** Crucially, we found no statistically significant correlation between the mass residuals and the luminosity ratio (‘Lx_ratio’) for either the RASS-based ($\rho = 0.195$, p-value = 0.215) or eRASS-based ($\rho = -0.109$, p-value = 0.490) fits. This is a key finding: clusters that are outliers in the $M-L_X$ relation (i.e., those with large

residuals) are not preferentially those clusters for which the ROSAT and eRASS1 luminosity measurements show the greatest disagreement. This lack of correlation strongly suggests that the observed intrinsic scatter in the $M-L_X$ relation is predominantly astrophysical in origin, rather than arising from systematic instrumental differences or variations in data processing pipelines between the two X-ray surveys.

In summary, our analysis demonstrates a high degree of consistency between the X-ray properties and derived $M-L_X$ scaling relations from ROSAT and eRASS1 for the common clusters in the CODEX-LS catalog. The observed intrinsic scatter in the $M-L_X$ relation, while well-characterized, appears to be driven by inherent astrophysical variations within the cluster population, with only a marginal indication of dependence on optical richness, rather than by systematic discrepancies between the two X-ray surveys. This work provides strong evidence for the robustness and compatibility of multi-survey X-ray cluster samples, reinforcing their utility for precision cosmology.

4. CONCLUSIONS

In this study, we addressed the critical challenge of ensuring consistency and understanding systematic differences between galaxy cluster properties derived from distinct X-ray surveys. Combining cluster samples from different instruments, such as the legacy ROSAT All-Sky Survey (RASS) and the ongoing eROSITA All-Sky Survey (eRASS1), is paramount for advancing cosmological studies, yet requires meticulous characterization of inter-survey discrepancies to avoid biased cosmological parameter constraints. Our work aimed to rigorously quantify these differences and, crucially, to disentangle true astrophysical scatter from systematic effects induced by instrumental variations or data analysis pipelines.

To achieve this, we leveraged a high-purity sample of 42 common galaxy clusters identified in the CODEX-LS catalog, for which X-ray properties were available from both ROSAT and eRASS1, alongside robust weak lensing mass measurements. Our methodology involved a multi-stage approach. First, we performed a direct, object-by-object comparison of X-ray luminosities and fluxes, calculating ratios and fractional differences, and employing statistical tests (e.g., Wilcoxon signed-rank test) to assess the significance of any observed offsets. We also investigated potential systematic trends in these differences with redshift, optical richness, and ROSAT luminosity. Second, we independently modeled the weak lensing mass-X-ray luminosity ($M-L_X$) scaling relations

for each survey using a robust Bayesian framework. This framework, implemented with an MCMC sampler, accurately accounted for measurement uncertainties in both mass and luminosity, as well as an intrinsic scatter in log-mass. The model included parameters for normalization (α), slope (β), and redshift evolution (γ). Finally, we conducted a detailed analysis of the scaling relation residuals. By correlating these residuals with optical properties like richness and mask fraction, and critically, with the inter-survey luminosity ratio, we sought to empirically determine the primary drivers of the observed scatter.

Our rigorous analysis yielded several key results. A direct comparison of X-ray luminosities revealed that eRASS1 luminosities are, on average, 13% higher than ROSAT measurements for the same clusters. However, a formal Wilcoxon signed-rank test demonstrated that this observed offset is not statistically significant (p -value = 0.102), indicating broad consistency between the two luminosity datasets. Furthermore, no statistically significant systematic trends were found between the inter-survey luminosity differences and cluster properties such as redshift, optical richness, or ROSAT luminosity. This suggests that any remaining discrepancies are not systematically dependent on the physical characteristics or observational conditions of the clusters.

Crucially, the independently derived $M-L_X$ scaling relations from ROSAT and eRASS1 exhibited remarkable consistency. Both surveys yielded virtually identical normalization parameters ($\alpha \approx 14.945 - 14.948$) and a common mass-luminosity slope of approximately 0.44. This slope is significantly shallower than the 0.75 predicted by self-similar models of cluster evolution, a well-established observational finding that underscores the importance of non-gravitational processes (e.g., AGN feedback, radiative cooling) in shaping the intracluster medium. The intrinsic scatter in log-mass was also highly consistent between the two fits, estimated at approximately 0.12–0.13 dex (corresponding to about 30% in mass at fixed luminosity). The redshift evolution parameter (γ) remained poorly constrained by our sample, with large uncertainties encompassing zero evolution.

The analysis of scaling relation residuals provided crucial insights into the origin of this intrinsic scatter. While we found a marginal positive correlation between residuals and optical richness (p -value $\approx 0.06 - 0.08$), hinting that clusters with higher optical richness for a given X-ray luminosity might be more massive, the most significant finding was the absence of a statistically significant correlation between the mass residuals and the inter-survey luminosity ratio. This direct test demonstrates that clusters exhibiting larger differ-

ences in their measured luminosities between ROSAT and eRASS1 are not preferentially the ones that deviate more significantly from the mean $M-L_X$ relation. This strongly suggests that the observed intrinsic scatter in the mass-luminosity relation is predominantly astrophysical in origin, reflecting the inherent diversity of cluster properties and formation histories, rather than being driven by systematic instrumental or pipeline differences between the two X-ray surveys.

In conclusion, this work provides compelling evidence for the robustness and compatibility of X-ray galaxy cluster samples from ROSAT and eROSITA. We have learned that despite differences in instrument design and energy bands, the derived X-ray luminosities are broadly consistent, and, more importantly, the fundamental mass-luminosity scaling relations are in excellent agreement. The observed intrinsic scatter in the $M-L_X$ relation is well-characterized and primarily astrophysical, rather than being confounded by inter-survey systematic effects. These findings significantly bolster confidence in the combined use of multi-survey X-ray cluster samples for future precision cosmological applications, paving the way for larger, more statistically powerful cluster catalogs to constrain cosmological parameters and understand the physics of cluster formation and evolution.