Acta Geotechnica Slovenica

# PROCESSING ENVELOPES AND AMPLITUDE-RESOLUTION BOUNDS FOR EARTHQUAKE CYCLE-DEMAND MOMENTS

Bence Kovacs[1], Eszter Nagy[2], Balazs Farkas[2], Reka Szabo[2] and Gabor Varga[1,*]

[1] Department of Structural Engineering, Faculty of Civil Engineering, Budapest University of Technology and Economics, Budapest, Hungary

[2] Department of Structural Mechanics, Faculty of Civil Engineering, Budapest University of Technology and Economics, Budapest, Hungary

**Received: 08/12/2025 Revised: 01/03/2026 Accepted: 05/03/2026 Published: 15/3/2026**

***Correspondence:** gabor2varga@gmail.com

DOI: 10.68304/gsej2026.a103

## ABSTRACT

Earthquake cycle demand depends on counting procedure, duration window, and amplitude resolution. Quadratic and quartic amplitude-weighted moments are evaluated for 7,856 NGA-West2 acceleration components with peak acceleration of at least 0.05 g. Peak, mean-crossing, level-crossing, and rainflow counts are compared over complete records and the 5–95% and 5–75% Arias-intensity intervals. Bin-edge inequalities separate finite-amplitude resolution from variation across the twelve processing combinations. Median largest-to-smallest moment ratios are 1.95 for quadratic weighting and 1.48 for quartic weighting; corresponding rainflow amplitude-resolution ratios are 1.76 and 1.89. Prediction uses reference peak acceleration and total cycle count, with 100 event-label groups for fitting, 40 for calibration, and 52 for testing. Ridge regression attains 94.2% simultaneous test-group coverage, with multiplicative interval factors of 4.72 and 6.55. Direct calibration of the quartic-to-quadratic envelope ratio gives a factor of 3.78, compared with 30.90 through rectangular propagation. These differences quantify numerical loading sensitivity. The results support evaluating counting choices and amplitude-bin uncertainty separately before applying cumulative loading measures in geotechnical earthquake analysis.

*Keywords:* earthquake cycle counting; amplitude discretization; duration filtering; cumulative loading; processing envelope; grouped prediction.

## 1. INTRODUCTION

Earthquake loading is neither stationary nor uniform. Its engineering consequences emerge from a sequence of excursions whose amplitudes, directions, frequencies, and temporal spacing evolve throughout the record. Peak acceleration captures an important aspect of that loading, but it


© 2026 The Author(s). Published under the CC BY 4.0 license.

does not specify how often substantial excursions occur or how the loading is distributed around its maximum. For geotechnical materials, this distinction matters because deformation and pore-pressure response can accumulate over successive cycles. A short interval containing a large acceleration pulse and a longer interval containing repeated moderate excursions may have similar maxima while presenting very different loading histories to a soil deposit. Describing this difference rigorously requires separating the properties of the recorded motion from the assumptions used to transform that motion into an engineering demand measure. Equivalent-cycle interpretations have a direct geotechnical foundation [1, 2].

Recorded ground motions permit the loading sequence to be examined directly. NGA-West2 includes recordings from shallow crustal earthquakes in active tectonic regions, together with information supporting their engineering use [3]. Its companion site compilation documents the breadth of recording environments and the importance of site characterization [4]. The broader research programme integrated these resources with ground-motion modeling and uncertainty assessment [5]. These developments provide a rich observational basis, but they do not imply that every derived scalar summary preserves the features needed for a particular geotechnical calculation. The suitability of a transformation depends on the response quantity, observation unit, and intended decision.

The treatment of duration illustrates this difficulty. Significant duration is defined through an accumulated intensity measure and therefore depends on both the selected intensity definition and the selected accumulation interval. Physically parameterized duration equations have been developed for active crustal regions [6]. Such equations help describe systematic duration variation, but duration measured in seconds is not itself a complete account of cyclic loading. A longer record can contain substantial additional low-amplitude motion without proportionately increasing its largest excursions. Conversely, several energetic excursions can occur within a short significant-duration interval. Consequently, replacing a cycle distribution with a duration alone introduces an additional modeling assumption that should be visible when interpreting cumulative demand.

Ground-motion selection research demonstrates why a single scalar rarely determines suitability. Selection procedures can match the mean, variability, and dependence structure of spectral ordinates across a range of periods [7, 8]. The conditional-spectrum formulation explicitly relates the selected representation to the intensity information used for conditioning [9]. These ideas concern a different representation from the cycle moments studied here, yet their methodological lesson is directly relevant: a representation should be judged by the information it retains for the intended analysis. Agreement in a selected intensity measure cannot establish agreement in every other aspect of the motion. The present study therefore avoids treating normalized cycle moments as substitutes for response spectra, duration, or site-specific stress histories.

The same distinction applies to empirical liquefaction assessment. A calibrated triggering procedure relates selected seismic demand descriptors and soil resistance measures to observed case-history outcomes [10]. Procedures for estimating building settlement additionally address deformation mechanisms and the interaction between a structure and liquefiable ground [11, 12]. The acceleration-cycle calculations quantify loading representation before a soil-response or damage model is evaluated. Maintaining this boundary prevents a mathematical comparison of

acceleration-cycle summaries from being mistaken for validation of a triggering equation or a settlement predictor.

Four cycle-counting rules and three duration windows offer a useful setting for investigating the conversion from acceleration excursions to cycle counts. Each rule supplies a particular way of reducing an irregular sequence into cycle counts. Each window changes the portion of the record admitted to that reduction. Their combination produces multiple descriptions of the same measured component. The resulting differences are epistemic in the limited sense that they arise from an analyst's processing choices, rather than from independent realizations of an earthquake. They can reveal sensitivity to a representation, but cannot identify the physically correct representation without information about the material response that the representation is intended to approximate.

Amplitude binning introduces another distinct transformation. A binned cycle-count array records how much count falls within an amplitude interval but does not retain individual amplitudes. Assigning all of that count to the bin midpoint is convenient and explicit. Nevertheless, midpoint assignment is an approximation. Its effect depends on the bin width, the exponent used to weight amplitude, and the distribution of counts within each interval. Increasing the emphasis on larger amplitudes can suppress the contribution of numerous small excursions while simultaneously increasing sensitivity to uncertainty about the positions of larger excursions inside their bins. These competing effects make it inappropriate to infer discretization accuracy from the apparent stability of one processing comparison.

The central research question is therefore whether uncertainty associated with processing choice can be distinguished quantitatively from uncertainty already implied by finite amplitude resolution. A related prediction question asks whether an upper processing envelope can be estimated reliably for excluded motion groups from two descriptors of a reference representation. The first question is deterministic and concerns the information preserved by the cycle-count arrays. The second is statistical and concerns transfer between grouped observations.

Bin-edge summation establishes lower and upper amplitude-weighted moments. Their comparison with the twelve processing combinations separates amplitude resolution from counting and window selection. Ridge regression, a weighted log-mean predictor, and extremely randomized trees estimate the paired upper envelopes from reference acceleration and cycle count. Group-maximum scores calibrate simultaneous prediction intervals, and direct functional scores calibrate the envelope ratio.

Several aspects distinguish the present assessment from a general comparison of counting algorithms. The upper envelope is evaluated for two amplitude weights on exactly matched component identifiers. The finite-bin bounds are calculated separately from processing spread. Components with zero lower processing moments are selected as explicit unbounded cases. Prediction success is assessed over all admitted components within an excluded grouping label, rather than only as a pooled proportion of individual records. Finally, the ratio of the two envelope moments is calibrated directly as well as through a joint rectangle. These choices establish an auditable computational question without claiming that the underlying inequalities or calibration principle

constitute newly invented statistical theory.

The analysis considers recorded peak acceleration, relative-amplitude cycle counts, and event labels. Vertical components are not excluded, and unavailable earthquake metadata are not inferred. Grouped prediction evaluates transfer between the event labels carried by the component identifiers. Its interpretation therefore depends on that grouping, the admitted orientations, and the reference-acceleration threshold.

## 2. ACCELERATION COMPONENTS AND CYCLE COUNTS

NGA-West2 acceleration records have been processed using peak counting, mean crossing, level crossing, and rainflow counting over complete and Arias-intensity-filtered histories [13, 14]. The calculations use these measured-motion cycle counts; the observations were not collected for the present investigation.

Each counting procedure is evaluated over the complete history, the 5–95% accumulated Arias-intensity interval, and the 5–75% interval. The codes pk, mp, lc, and rf identify the counting rules, while suffixes 95 and 75 identify the filtered windows. For each component, peak acceleration is expressed in g and cycle counts occupy ten relative-amplitude intervals, each 0.1 PGA wide. Fractional counts represent partial cycles.

Exact component identifiers pair the twelve processing combinations. Selection requires rf95 PGA $\geq 0.05$ g. This criterion restricts the calculation to stronger recordings; it is not a liquefaction threshold. All recorded orientations remain in the analysis. Multiple orientations and stations associated with an earthquake are dependent observations.

## 3. AMPLITUDE BOUNDS AND PROCESSING ENVELOPES

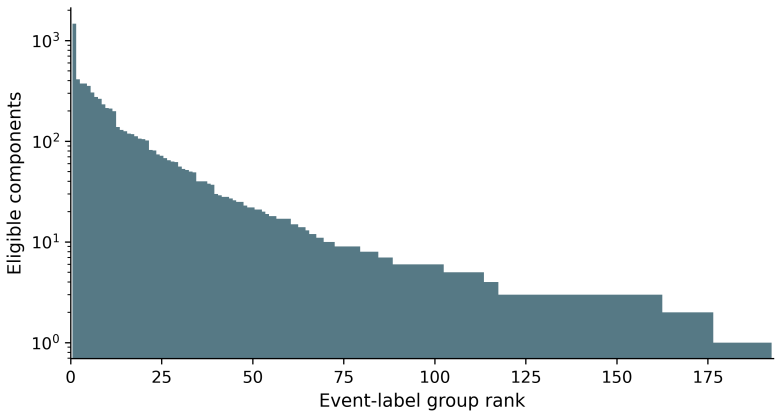



**FIGURE 1.** COMPONENT COUNTS ACROSS EVENT-LABEL GROUPS.

The reference collection contains 64,125 components, of which 7,856 satisfy the matched selection. Component identifiers contain 192 distinct event labels. The event-label token defines exclusion

groups; these labels have not been independently reconciled with an earthquake catalogue. A shared sequence label can combine events, and alternative labels can separate related events. Group-wise prediction statements are conditional on this operational grouping and do not establish independence between earthquakes.

The unequal representation of event labels is shown in Figure 1. Components belonging to one label remain together during fitting, calibration, or testing; the rank distribution describes sampling concentration rather than earthquake frequency.

### *3.1 AMPLITUDE-WEIGHTED MOMENTS*

For component $i$, processing variant $p$, and bin $b$, let $n_{ipb}$ be the supplied cycle count. Relative amplitude lies in $[l_b, u_b]$, with $l_b = (b-1)/10$, $u_b = b/10$, and midpoint $m_b = (l_b + u_b)/2$. For exponent $\beta > 0$, define

$$M_{ip}^{(\beta)} = \sum_{b=1}^{10} n_{ipb} m_b^{\beta}. \tag{1}$$

The two selected exponents, 2 and 4, provide quadratic and quartic weighting of larger relative amplitudes. They are sensitivity choices, not fitted soil fatigue exponents. The moments are dimensionless cycle-weighted summaries. Normalization by PGA removes absolute amplitude from the moment; PGA remains available as a separate descriptor. A high moment therefore does not imply high physical stress demand at a site. The counting procedures produce distinct amplitude distributions for the same Northridge component (Figure 2). The curves use the same Arias-intensity window and therefore isolate the counting-rule comparison for this record; they are not independent motion observations.

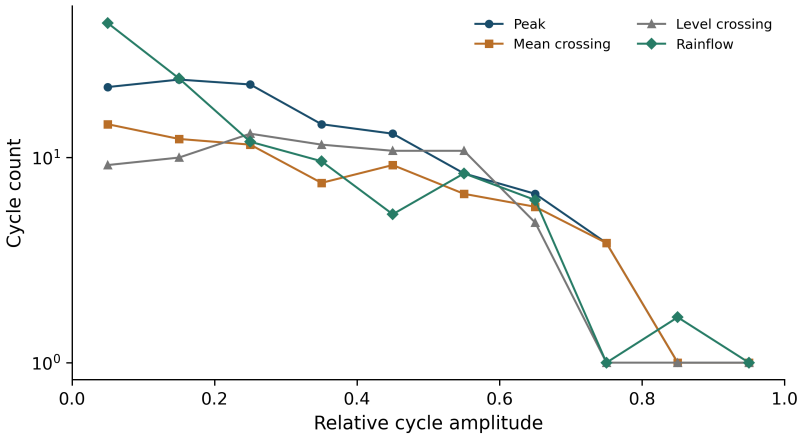



**FIGURE 2.** NORTHRIDGE CYCLE COUNTS OVER THE 5–95% ARIAS INTERVAL.

### 3.2 WITHIN-BIN RESOLUTION BOUNDS

Monotonicity of $a^\beta$ for $a \geq 0$ gives

$$L_{ip}^{(\beta)} = \sum_b n_{ipb} l_b^\beta \;\leq\; M_{ip,\text{unknown}}^{(\beta)} \;\leq\; \sum_b n_{ipb} u_b^\beta = U_{ip}^{(\beta)}. \tag{2}$$

These bounds hold for the amplitude locations compatible with the supplied bin counts. They are not statistical confidence intervals and cannot restore lost loading order. Midpoint weighting supplies a explicit estimate within them. The resolution ratio $U/L$ describes uncertainty from bins alone; when $L = 0$, the upper-to-lower ratio is unbounded and must not be replaced with an arbitrary finite number. The influence of amplitude weighting on the Northridge calculation is illustrated in Figure 3. The continuous exponent range is a deterministic sensitivity calculation, with the two exponents used throughout the paper marked explicitly. The shaded limits follow the recorded bin edges and do not represent confidence probabilities.

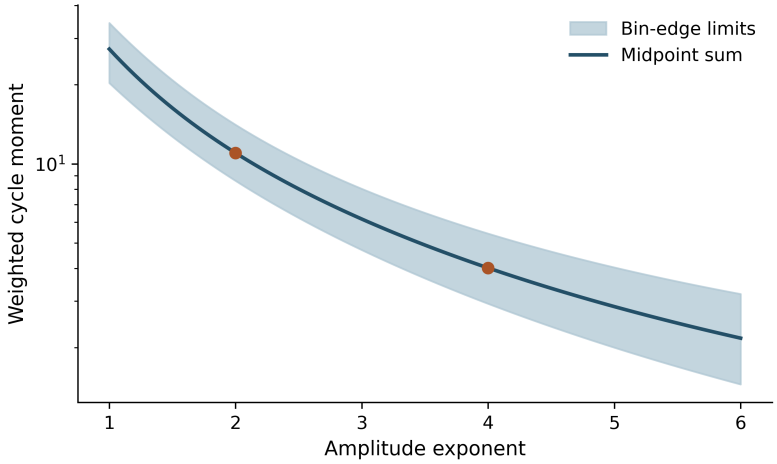



**FIGURE 3.** MOMENT BOUNDS ACROSS AMPLITUDE EXPONENTS.

### 3.3 PROCESSING ENVELOPES

Define the midpoint processing envelope and its spread as

$$E_i^{(\beta)} = \max_p M_{ip}^{(\beta)}, \qquad S_i^{(\beta)} = \frac{\max_p M_{ip}^{(\beta)}}{\min_p M_{ip}^{(\beta)}}. \tag{3}$$

The bounds across both amplitude uncertainty and processing choice are $\min_p L_{ip}^{(\beta)}$ and $\max_p U_{ip}^{(\beta)}$. The twelve procedures are alternatives applied to one component, not twelve independent realizations. The envelope is conservative relative to the included choices and does not imply that every procedure is equally appropriate for every application. Zero denominator moments produce an unbounded spread, selected in the computational record.

Processing spread and within-bin spread answer different questions. The former compares chosen algorithms and windows at bin midpoints. The latter varies the unobserved amplitudes within

one fixed representation. Neither should be interpreted as an estimate of the other, and neither includes site-response or constitutive uncertainty.

### *3.4 MOMENT CALCULATION*

A Northridge component illustrates the calculation. Its rainflow counts over the 5–95% Arias-intensity interval are shown below.

| Relative bin | Count | Contribution to $M^{(2)}$ | Contribution to $M^{(4)}$ |
|---|---|---|---|
| 0.0–0.1 | 66.5 | 0.166250 | 0.000416 |
| 0.1–0.2 | 30.5 | 0.686250 | 0.015441 |
| 0.2–0.3 | 12.5 | 0.781250 | 0.048828 |
| 0.3–0.4 | 9.5 | 1.163750 | 0.142559 |
| 0.4–0.5 | 4.5 | 0.911250 | 0.184528 |
| 0.5–0.6 | 8 | 2.420000 | 0.732050 |
| 0.6–0.7 | 5.5 | 2.323750 | 0.981784 |
| 0.7–0.8 | 1 | 0.562500 | 0.316406 |
| 0.8–0.9 | 1.5 | 1.083750 | 0.783009 |
| 0.9–1.0 | 1 | 0.902500 | 0.814506 |

For exponent 2, midpoint summation gives 11.001250; lower-edge summation gives 8.620000 and upper-edge summation gives 14.085000. For exponent 4, midpoint summation gives 4.019528; lower-edge summation gives 2.938600 and upper-edge summation gives 5.432250. The processing envelope repeats this arithmetic for all twelve matched variants and selects the largest midpoint moment. The two envelope outputs may attain their maxima under different processing choices; their pair is a componentwise upper summary, not necessarily a single realized processing state.

## 4. PREDICTION AND INTERVAL CALIBRATION

### *4.1 TRAINING DESIGN AND PREDICTION*

The response pair is $(\ln E^{(2)}, \ln E^{(4)})$. Predictors are $\ln \mathrm{PGA}_{rf95}$ and the logarithm of total rf95 cycle count. The predictor set deliberately summarizes a single reference processing variant while targets encompass all twelve. This is not a complete site-response predictor and its use requires that the reference processing has already been calculated.

The alphabetically sorted event-label groups are shuffled with NumPy default generator seed 20251101. One hundred groups train the models, forty calibrate intervals, and the remaining fifty-two are tested once. These sets are disjoint. Each training row receives weight proportional to the reciprocal of its group's selected count, rescaled to sum to the training-record count. Predictor means and standard deviations are fitted using training weights only.

Ridge minimizes $\sum_i w_i \|y_i - a - Bz_i\|^2 + \|B\|_F^2$, with an unpenalized intercept. The reference model predicts the weighted training log mean. The tree ensemble uses 200 extremely randomized trees, minimum leaf size five, all predictors, and random seed 42; these settings are fixed before evaluating test responses. Extremely randomized trees have an established methodological origin [15]. No model is selected or tuned using calibration or test responses.

### *4.2 JOINT AND DIRECT-FUNCTIONAL CALIBRATION*

For output $k$, training RMS log residual scale $s_k$ is floored at 0.03. Define $r_i = \max_k |y_{ik} - \widehat{y}_{ik}|/s_k$ and group score $R_g = \max_{i \in g} r_i$. The 90% calibration threshold is the 37th ordered score among forty calibration groups, because $\lceil 41(0.9) \rceil = 37$. Intervals are

$$I_{ik} = [\exp(\widehat{y}_{ik} - qs_k), \exp(\widehat{y}_{ik} + qs_k)]. \tag{4}$$

The finite-sample rank argument requires exchangeability of calibration and future group objects. This remains an assumption, not a result of excluding labels. General conformal methodology supplies the relevant rank principle [16]. Pooled calibration uses the analogous corrected quantile of calibration-record scores; it targets a different event and is selected as a diagnostic comparison.

The direct-functional analysis examines $C_i = E_i^{(4)}/E_i^{(2)}$. Rectangular propagation has log half-width $q(s_1 + s_2)$. Direct scores use $|\ln C_i - (\widehat{y}_{i2} - \widehat{y}_{i1})|/(s_1 + s_2)$ and their maximum within group, followed by the same 37th-order calibration. The direct interval protects the contrast, rather than simultaneous coverage of both envelope moments. Its narrowing follows partly from the triangle inequality.

Point errors are averaged within each test group and then equally across groups. Coverage requires success at every component in a test group. The twelve variants and the two response moments do not increase the number of independent validation groups.

### *4.3 INTERPRETIVE SEPARATION OF THE THREE UNCERTAINTY LAYERS*

The computational design distinguishes information loss, analyst choice, and statistical transfer. Information loss arises because an amplitude interval replaces the exact amplitudes assigned to it. Analyst choice arises because different counting procedures and selected windows describe the same component differently. Statistical transfer arises when a model fitted to one set of grouping labels predicts another. These layers are related through the derived variables, but they are not interchangeable. A narrower prediction interval cannot restore unrecorded amplitude positions, and a small processing spread cannot establish that midpoint approximation is accurate. Keeping the layers separate also clarifies which numerical statement follows from an inequality and which depends on empirical validation.

For a given component and exponent, the midpoint estimate is computed before any machine learning step. The upper processing envelope is likewise a deterministic transformation of the twelve available cycle-count arrays. Fitted models do not define this target, nor are their predictions

used to select the envelope's largest value. This order of operations prevents circular validation in which a predictor determines the quantity against which it is evaluated. It does not eliminate every dependence between predictors and targets: the total reference count is derived from one processing variant included among the twelve. The prediction task is consequently one of summarizing a multi-procedure envelope from a reference representation, rather than forecasting cumulative loading from earthquake magnitude and distance alone.

The input representation is intentionally small. Peak acceleration supplies an absolute-amplitude descriptor, while the reference count supplies a measure of the number of admitted cycle parcels. Neither identifies a frequency spectrum, duration in seconds, site condition, or rupture mechanism. Established ground-motion models use source, path, and site variables to predict intensity measures [17–20]. Those models are not directly comparable to the present regressions because they answer a different question and use a richer conditioning set. Their relevance is to delineate what cannot be inferred from two summary predictors. The fitted relationships should not be used as replacements for earthquake-specific ground-motion models.

### *4.4 DEPENDENCE, WEIGHTING, AND THE PREDICTION POPULATION*

The numerical unit of prediction is a recorded component; the unit used for splitting and simultaneous success is an event-label group. Several components can share a recording environment, a station, or an earthquake sequence, so treating every component as independently sampled would exaggerate the amount of information available for assessing transfer. Grouped exclusion reduces one obvious form of reuse, but its validity depends on the label mapping. Structured cross-validation literature emphasizes that the partition should match the intended prediction task rather than default to random row allocation [21]. Here, the intended task concerns components belonging to excluded labels, and the wording of the calculated coverage follows that definition.

Equal total training weight per group prevents a heavily represented label from dominating the loss solely through its component count. The weighting does not make the groups independent and does not remove regional imbalance. It changes the fitted objective from an observation-weighted approximation toward an average over represented groups. This distinction matters when comparing the training-mean reference with fitted learners: both should use the same weighting convention if their differences are to be attributed to the predictive relationship rather than to different target populations. It also explains why a global component-level error can differ from the equally weighted group error calculated in the main comparison.

Each calibration group contributes one maximum score, regardless of its number of components. A larger group still has more opportunities to produce an extreme score, and this influence is part of the group object being calibrated. The resulting interval has a marginal interpretation over new group objects with comparable sampling characteristics. It does not guarantee the same performance for an arbitrarily large future collection of components. Changes in record selection, instrument coverage, or catalogue labeling can change that object even when the fitted numerical predictors have similar ranges.

### *4.5 MODEL SCOPE AND ALTERNATIVE UNCERTAINTY PROCEDURES*

The transparent ridge predictor provides a controlled comparison with the nonlinear ensemble. Standardization is confined to training observations, the intercept is not penalized, and the fixed penalty avoids selecting a tuning value through the test set. The ensemble supplies a flexible alternative without a broad optimization exercise. Both are implemented through a documented scientific computing library [22]. The comparison is intended to reveal whether modest flexibility changes observed error and interval size, not to identify a universally best learning algorithm. Interpretability is especially useful here because the response is already a multi-stage transformation of the measurements [23].

Other uncertainty procedures could be developed for this task, but they would require their own grouped design. Generalized random forests provide an alternative method for local statistical estimation [24], and conformalized quantile regression can calibrate adaptive intervals based on estimated lower and upper quantiles [25]. Jackknife-based predictive inference offers another route to using limited samples [26]. These methods are discussed as alternatives, not represented as analyses that were performed. The present study retains global residual scales to make the consequences of the calibration unit inspectable. Its wide intervals should motivate an appropriately validated extension rather than an unsupported claim that a different algorithm would necessarily be narrower and reliable.

Finally, the nominal statement remains conditional on a distributional relationship between calibration and future groups. Limits on distribution-free conditional inference prevent interpreting marginal coverage as a guarantee for every individual geological or seismological subgroup [27]. Work on departures from exchangeability and covariate shift makes clear that changes in the observation population require explicit assumptions or additional adjustment [28, 29]. No such correction is fitted here. These qualifications are part of the method's definition and should be selected whenever its numerical results are compared with an engineering reliability target.

## 5. RESULTS

The median processing spread is 1.947 for quadratic weighting and 1.478 for quartic weighting. The median rf95 amplitude-resolution ratios are 1.762 and 1.890, respectively. Two eligible components have zero minimum midpoint moment across processing choices, producing unbounded processing ratios; they are selected explicitly rather than assigned finite values. These are component-weighted descriptive medians, not earthquake-weighted estimates. Figure 4 shows the distributions of processing spread; Figure 5 separates the twelve choices relative to rf95. Different amplitude exponents alter the sensitivity to low-amplitude cycles. Unbounded ratios do not appear at finite horizontal coordinates. Processing-choice map of median moment ratios relative to rainflow counting over the 5–95% Arias interval. Zero on the color scale denotes equal median moment; a value of one denotes a factor of two.

The quartic moment is less sensitive in its median processing spread than the quadratic moment.

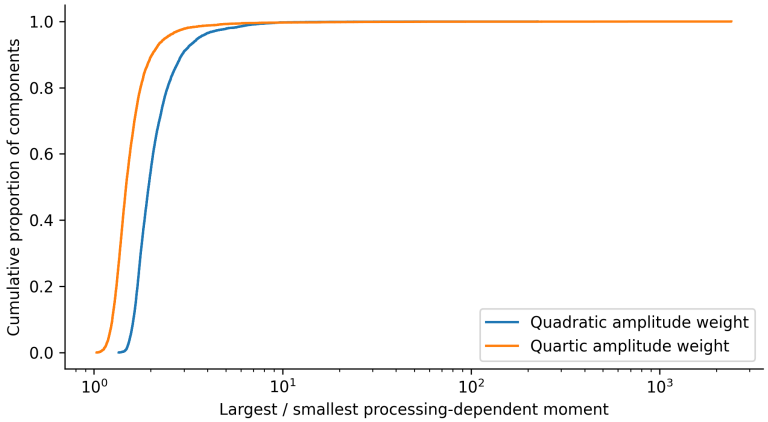



**FIGURE 4.** DISTRIBUTION OF THE LARGEST-TO-SMALLEST MIDPOINT MOMENT ACROSS THE TWELVE PROCESSING CHOICES.

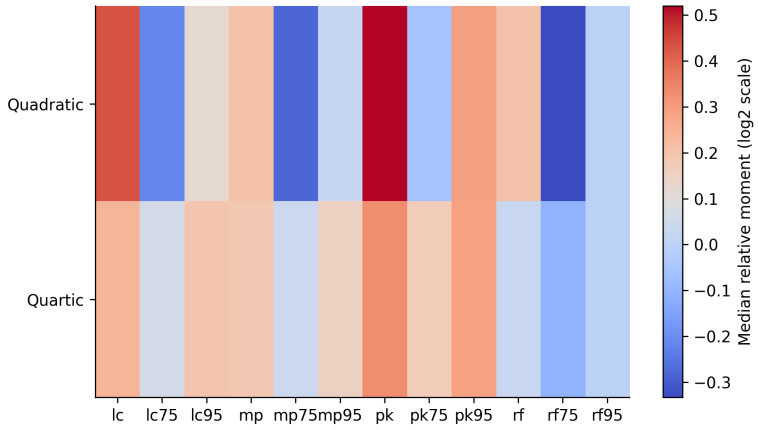



**FIGURE 5.** MEDIAN MOMENT RATIOS ACROSS PROCESSING CHOICES.

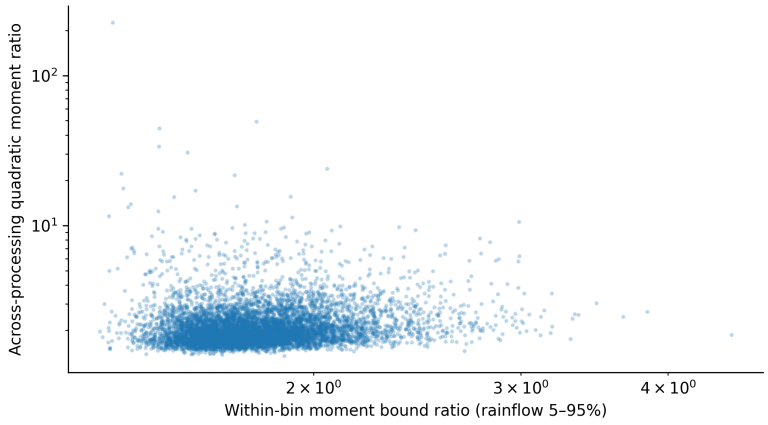



**FIGURE 6.** AMPLITUDE BOUNDS AND PROCESSING-CHOICE VARIATION BY COMPONENT.

This is consistent with reduced influence of numerous low-amplitude cycles, but does not establish greater accuracy for soil damage. Within-bin uncertainty remains material: Figure 6 compares two different sources of ambiguity without assuming they are statistically independent. Component-level relationship between finite-bin amplitude bounds for the reference quadratic moment and variation across processing choices. The two axes measure distinct uncertainty sources; the plot

does not identify a causal relationship.

The fixed test set contains 1,452 components. Table 1 shows similar joint coverage for the two fitted learners but broad intervals. Ridge log errors are 0.367 and 0.436, compared with 0.367 and 0.430 for trees. These close point-error differences do not demonstrate superiority. Pooled-score ridge calibration covers every component in only 57.7% of test groups. This is evidence about the mismatch between record and group targets, not proof that the pooled method fails its intended objective. For ridge, rectangular contrast propagation has factor 30.90; direct functional calibration has factor 3.78, with 94.2% test-group coverage. The improvement reflects the functional target and dependence between the two moments; it must not be interpreted as a reduction in physical ground-motion uncertainty. The quadratic prediction comparison in Figure 7 uses every held-out component. Cell color indicates observation density, and the diagonal denotes equal calculated and predicted envelopes. This display preserves the component-level spread while the numerical performance comparison weights each event-label group equally. Multiplicative factors define intervals $[p/f, pf]$.

**TABLE 1.** PREDICTION AND SIMULTANEOUS COVERAGE FOR THE FIFTY-TWO EXCLUDED EVENT-LABEL GROUPS.

| Model | Log MAE $E^{(2)}$ | Log MAE $E^{(4)}$ | Group coverage | Factor 2 | Factor 4 |
|---|---|---|---|---|---|
| Training mean | 0.534 | 0.536 | 94.2% | 6.57 | 7.68 |
| Ridge | 0.367 | 0.436 | 94.2% | 4.72 | 6.55 |
| Extra trees | 0.367 | 0.430 | 96.2% | 4.70 | 6.78 |

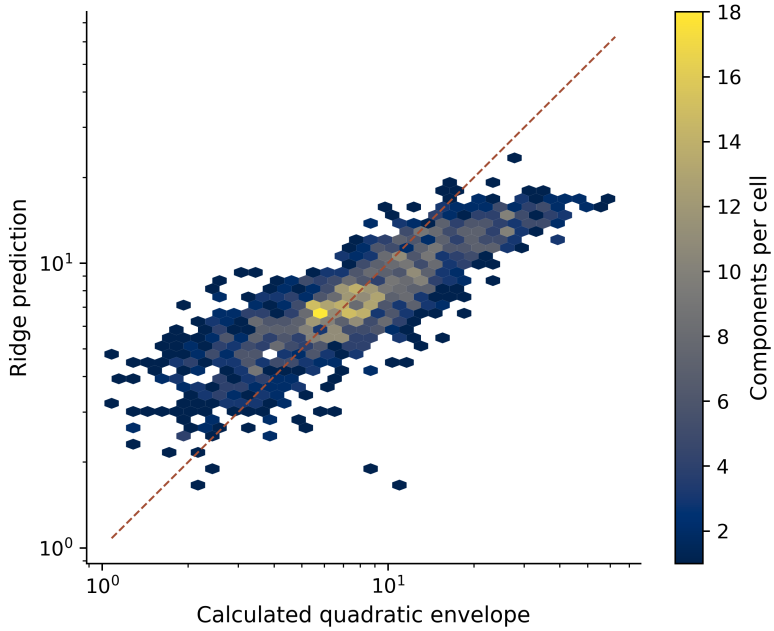



**FIGURE 7.** QUADRATIC ENVELOPES IN EXCLUDED GROUPS.

### 5.1 WHAT THE PROCESSING RATIOS ESTABLISH

The median quadratic spread of approximately 1.95 means that, for the middle component in the selected population, the largest midpoint-based moment among the admitted procedures is nearly twice the smallest. This is a comparison among alternative descriptions of one component. It does not mean that the earthquake has twice the probability of causing damage, that one counting rule is biased by a factor of two, or that a design demand should automatically be doubled. Those interpretations would require a calibrated relationship between the representation and a material response. The result instead establishes that procedural choice is large enough to be calculated explicitly rather than hidden within an otherwise precise-looking scalar.

The smaller quartic median spread, approximately 1.48, indicates that greater emphasis on large relative amplitudes changes the pattern of sensitivity. Numerous small cycles contribute more strongly to the quadratic sum than to the quartic sum. Differences between counting rules that primarily affect those cycles can therefore have less influence after quartic weighting. This explanation follows the form of the moment, but it should not be overextended. The counts assigned to the larger bins can also differ between procedures, and those differences receive greater weight under the quartic transformation. The aggregate result is an empirical balance between these effects, not a monotone rule that higher exponents always reduce processing uncertainty.

The cumulative distributions are useful because they reveal more than a single average ratio. Components near the lower end of a spread distribution are comparatively insensitive to the admitted processing alternatives. Components farther into its upper tail are more sensitive and merit closer examination before a downstream calculation relies on one representation. The distributions do not rank the records by expected liquefaction severity. A weak or short motion can exhibit a large relative discrepancy if its smallest moment is very small. Conversely, a strong motion can have a modest processing ratio while retaining a large absolute loading amplitude. Relative sensitivity and absolute intensity should therefore remain separate descriptors.

Two components have a zero minimum midpoint moment among the processing variants. Their ratios are unbounded by definition. Retaining these cases prevents an artificial impression that every admitted representation supplies a positive cumulative quantity. The computational record preserves them, while a finite logarithmic axis cannot display their horizontal coordinates. A visual distribution should consequently be read together with the stated count of unbounded cases. Replacing zero denominators by a small arbitrary constant would make the apparent upper tail depend on a numerical convention rather than on the stored observations, and is not justified for this descriptive comparison.

### 5.2 HOW TO READ THE PROCESSING-CHOICE MAP

The processing-choice map compares each variant with the rainflow reference over the significant-duration interval. Because the displayed quantity is a median of componentwise ratios, a color difference summarizes a typical relative change within the selected component population. It is not a ratio of two total moments accumulated over the complete component collection. Those two

summaries weight components differently and need not agree. This distinction is important for a collection with many components from some labels and few from others. The map is descriptive at component level, while the later predictive errors deliberately use equal weighting across groups.

The map also should not be interpreted as a factorial experiment with twelve independent treatment samples. Every displayed processing variant derives from the same underlying component identifiers. The comparison benefits from that pairing because source and recording identity are held fixed when the processing label changes. However, the pairing does not create an experimental intervention on the earthquake. Differences across windows and rules are transformations of an observed sequence, and the factors can interact. A change in duration window can remove cycles that one rule counts differently from another, so a simple additive interpretation of rule and window effects would require further evidence.

The reference representation is a normalization choice for visualization and for the predictor set. It is not designated as ground truth. A value near unity relative to the reference only indicates agreement in the selected moment; it cannot establish agreement in the distribution of cycle amplitudes, the temporal loading order, or the time at which an excursion occurred. Two sets of bins may produce the same weighted sum while allocating counts very differently. Likewise, the same total number of cycles can produce different weighted moments when the counts occupy different amplitude bins. These non-uniqueness properties explain why the reference count is an incomplete predictor of the processing envelope.

### *5.3 DISCRETIZATION BOUNDS AND THEIR PRACTICAL MEANING*

The median reference bin-resolution factors, approximately 1.76 for quadratic weighting and 1.89 for quartic weighting, show that the stored bin widths alone admit material variation in the weighted sum. These ratios arise without changing the counting rule or duration window. They therefore cannot be dismissed as disagreement among algorithms. At the same time, they are worst-case bounds over amplitude positions compatible with the bins, not probabilistic estimates of how amplitudes are distributed within those intervals. The unknown moment need not lie near either extreme. Midpoint assignment is a explicit central convention, but the bound calculation makes visible the information that convention suppresses.

The first bin requires particular attention because its lower edge is zero. Counts assigned there contribute nothing to the lower bound, but they contribute positively to the midpoint estimate and upper bound. When most count lies at low relative amplitudes, the ratio of upper to lower bounds can become large even if the absolute midpoint moment is modest. This behavior is mathematically expected. It cautions against using a relative resolution factor alone as a record-quality criterion. A useful audit should inspect both the magnitude of the moment and the absolute difference between its bounds, as well as the ratio calculated here.

The quartic reference resolution factor can exceed the quadratic factor even though the quartic processing spread is smaller in its median. There is no contradiction. Raising the exponent changes sensitivity to where a count sits inside a bin, while the processing spread compares different

count allocations at fixed midpoints. These are different perturbations. The observed contrast demonstrates why one uncertainty layer should not be used as a proxy for the other. In particular, a processing rule that appears stable across duration windows may still have a relatively coarse amplitude representation for a strongly nonlinear weighting function.

The scatter plot should therefore be read as an inspection of co-occurrence, rather than as a calibration curve between uncertainties. A component can have relatively narrow reference amplitude bounds and substantial processing spread, or broad reference bounds with limited spread across the chosen procedures. No independence assumption enters this graphical comparison, and no fitted causal relation is claimed. A later investigation could stratify these patterns by a verified event catalogue, orientation, and site information, but those variables cannot be reconstructed reliably from the present two-descriptor prediction setup alone.

### *5.4 INTERPRETING THE HELD-OUT PREDICTION COMPARISON*

The training-mean predictor supplies a meaningful reference because it uses the same weighted target population without conditioning on the component descriptors. Its larger log errors indicate that the two supplied predictors carry information about the upper processing moments. Ridge and the tree ensemble both reduce those errors in the fixed held-out groups. The total reference count is already related to sums of cycle counts, and the target is another summary of closely related representations. The practical question is how much of the multi-procedure envelope that reference summary can recover and how uncertain the recovery remains.

The small difference between ridge and tree errors suggests that the tested additional flexibility yields limited average benefit on this particular split. It does not establish that the relationships are globally linear or that the learners are statistically equivalent. Their fitted structures, training residual scales, and calibrated upper score thresholds differ. In consequence, similar average point errors need not produce identical predictive interval factors. A claim of superiority would require further evidence across appropriate partitions and a comparison criterion specified independently of test outcomes. The study instead reports both predictors as informative but uncertain approximations to the deterministic envelope.

Observed group coverage is high for both learners, but it is obtained with large multiplicative factors. Such factors should be translated carefully: an interval factor of roughly five places bounds near one fifth and five times the central prediction, rather than a symmetric percentage band around that prediction. Large log-scale ranges are consistent with substantial unmodeled variation between excluded groups. They also reflect the demanding success event, which requires every admitted component and both outputs to be covered. The analysis thus exposes a tradeoff between the strength of the protected event and the precision of the resulting prediction, rather than demonstrating that nominal coverage alone makes a predictor operationally useful.

Pooled calibration produces a much lower proportion of groups with complete success because it uses a record-based threshold. This contrast is informative only when its different target is kept explicit. A pooled interval may cover a large proportion of individual records while missing at

least one component in many groups. That is compatible with its construction. Group-maximum calibration aligns the score with the stated simultaneous event, but its wider intervals should not be presented as a cost-free improvement. The appropriate calibration unit must follow the way the prediction is intended to be used, and the acceptable interval width must follow the engineering decision rather than a statistical convention.

The direct contrast analysis offers another example of target alignment. Because the two envelope moments are derived from related amplitude distributions, their predictive errors may move together. The ratio can therefore be more stable than a rectangle permitting one moment to take its largest admissible value while the other takes its smallest. Direct calibration targets the ratio itself and avoids some of those incompatible corners. Nevertheless, it protects a different event from simultaneous coverage of both moments. The reduction in interval factor cannot be used to claim that uncertainty in each individual moment has disappeared or that both guarantees hold together at the nominal level.

The quartic residuals in Figure 8 are defined as the logarithm of calculated envelope divided by prediction. Positive residuals indicate underprediction. Median residuals are calculated within logarithmic PGA bins for description, without refitting the predictor or adjusting its calibration.

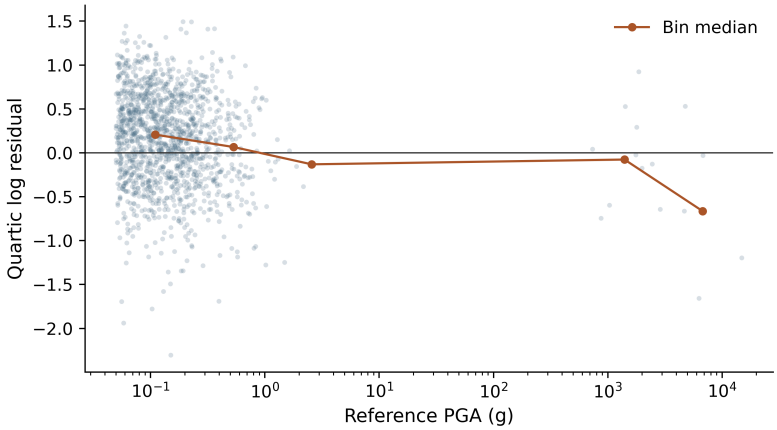



**FIGURE 8.** QUARTIC RESIDUALS AGAINST REFERENCE ACCELERATION.

The interval display in Figure 9 selects twenty components at evenly spaced ranks after ordering the test observations by PGA. Its selection is independent of prediction success. Each horizontal range uses the same calibrated quadratic factor; the calculated envelope and fitted central value are shown separately. These illustrative components do not replace the group-wise coverage assessment.

## 6. DISCUSSION

### *6.1 RELATIONSHIP TO GROUND-MOTION REPRESENTATION*

The findings fit within a broader problem of representing multidimensional earthquake excitation with a small set of intensity descriptors. Standard ground-motion models distinguish amplitude,

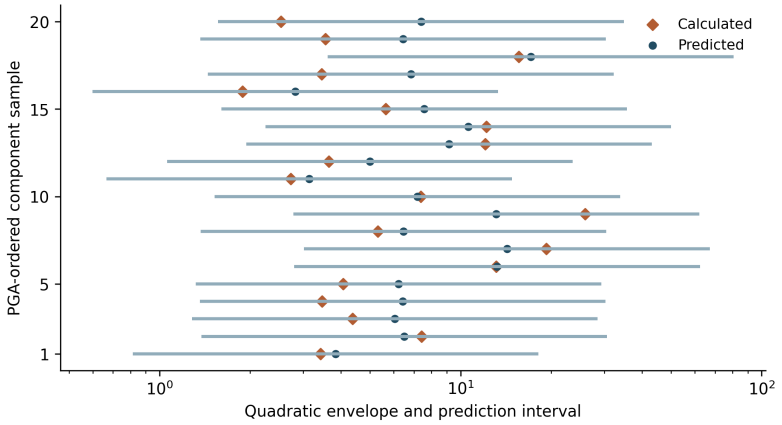



**FIGURE 9.** CALIBRATED INTERVALS ACROSS PGA-ORDERED COMPONENTS.

period dependence, and site effects through their conditioning variables and functional forms. Nonlinear site-amplification studies show why the properties of the surface record cannot be separated automatically from the site response that produced it [30, 31]. The current normalized cycle moments contain neither a complete site model nor the spectral information required to describe that response. They should therefore accompany, rather than replace, the descriptors needed for the subsequent physical calculation.

Orientation adds another dimension. Definitions of seismic intensity based on horizontal components require an explicit treatment of component orientation and combination [32]. The present population retains every admitted recorded orientation because the component identifiers do not provide a separately verified orientation field. This choice avoids an undocumented filtering rule but broadens the interpretation of the results. The processing comparisons apply to recorded components, not to an orientation-independent horizontal intensity measure. Translating them into a horizontal design input would require an additional, documented selection or combination procedure and a renewed assessment of the resulting observation population.

The distinction between acceleration cycles and soil stress cycles is particularly important for geotechnical use. A soil layer responds through wave propagation, impedance contrast, nonlinear stiffness, damping, drainage conditions, and evolving effective stress. Even a perfectly known acceleration-cycle distribution at a recording point does not determine the stress-cycle distribution throughout a deposit. The present bounds quantify an uncertainty in the stored acceleration representation. They do not bound all physically admissible soil responses. This limitation is substantive rather than merely terminological: applying the moments directly as resistance or deformation criteria would omit the mechanism that connects the measured motion with the response of interest.

Research on duration and structural response provides a useful parallel, while remaining outside the present material-response scope. Studies of collapse and inelastic response examine how duration can influence systems whose response depends on loading history [33, 34]. Their relevance here is that peak response and cumulative effects need not be represented equally well by the same descriptor. They do not validate the quadratic or quartic exponents for geotechnical soils. Exponents appropriate for a particular cumulative response would need to be calibrated against

compatible physical observations or constitutive simulations, with uncertainty in that calibration kept separate from the processing comparisons calculated in this paper.

### *6.2 IMPLICATIONS FOR LIQUEFACTION-ORIENTED STUDIES*

Arias intensity and cumulative absolute velocity combine aspects of amplitude and accumulation differently from the moments studied here. Ground-motion prediction research has treated their variability and spatial dependence explicitly [35]. Studies of liquefaction-induced lateral spreading have also examined the information provided by multiple intensity measures [36]. These contributions support considering accumulation alongside peak acceleration, but they do not imply that any cumulative index is automatically sufficient. The definition of the measure, its physical relationship to the response, and the availability of compatible observations remain decisive.

For an analyst selecting input motions for a liquefaction study, the present results support a documented processing audit before the derived cycle summary is adopted. The audit would identify the counting rule, duration interval, amplitude convention, and exponent, then report whether the intended conclusion changes across reasonable alternatives. If a material-response conclusion is stable only under one representation, that dependence should be disclosed. If several representations give similar midpoint demand but their bin bounds are broad, the apparent agreement should not be interpreted as high amplitude precision. The audit provides evidence about representation sensitivity, not an automatic rule for selecting the most damaging alternative.

An upper processing envelope can be useful as a screening device when the cost of underrepresenting cumulative loading is consequential. However, the envelope also has limits. Its two components can be maximized by different processing choices, so the pair need not correspond to one actual processed history. Moreover, a larger normalized moment does not necessarily identify a more damaging physical record when absolute amplitude, frequency, or site conditions differ. The conservative interpretation is restricted to the included numerical summaries. Using the envelope as an input to a nonlinear response model would require a clearly stated mapping and should avoid implying that every upper value is physically realized simultaneously.

### *6.3 GENERALIZATION, CATALOGUE STRUCTURE, AND REPORTING*

The dependence issue cannot be resolved by counting cycle-count arrays or rows. Twelve cycle-count arrays supply repeated processing variants; three components recorded at one station can share much of their event and site context; and records from an earthquake sequence can remain dependent even when their component identifiers differ. The operational label grouping reduces obvious reuse of named groups between training and testing, but it is not a substitute for catalogue validation. Seismological research on clustering and declustering illustrates the importance of explicitly defining relationships between events when interpreting a collection statistically [37]. The present study therefore avoids equating its identifier tokens with an independently verified sample of unrelated earthquakes.

This limitation affects how the observed coverage should be communicated. The excluded event labels do not by themselves establish exchangeability. The labels can differ in record count, geography, instrumentation, and sequence membership. A prediction guarantee derived for a new exchangeable group does not establish uniform reliability across those contexts. The most defensible report identifies the precise group definition, the admission threshold, and the available orientations, then states the empirical test outcome separately from the theoretical rank argument. This separation enables a later catalogue audit without rewriting the numerical meaning of the current results.

The limited input set also restricts transportability. The threshold selects components whose reference PGA exceeds the stated cutoff. Predictions for weaker motions, differently processed records, or representations with other amplitude bins are outside that defined population. Using the model on such records would change either the predictor distribution, the target transformation, or both. Work on predictive uncertainty and distributional shifts provides a reminder that a numerical interval is only as meaningful as its supporting assumptions [28, 38]. No automatic correction for these changes is supplied by the tree ensemble, the ridge penalty, or the residual calibration step.

A practical consequence is that deployment should preserve the reference representation used to construct the predictors. If total rf95 count is unavailable, replacing it with count from another rule is not a harmless relabeling. The predictor would change meaning while the fitted model remained calibrated to the specified definition. Likewise, changing the duration window can change both the reference count and the target envelope. A consistent implementation must either calculate the originally specified inputs or refit and recalibrate the complete procedure. This is a measurement-definition requirement, independent of whether the fitted learner is simple or highly flexible.

### *6.4 AMPLITUDE INFORMATION AND VALIDATION LIMITS*

The calculated moment bounds can be checked directly against the component counts and bin edges. Arithmetic agreement does not establish the correctness of acceleration processing, event-label identity, or a relationship with soil response. Those questions require independent checks of the motion histories, earthquake catalogue, and material behavior.

The bin-edge inequalities hold because amplitude weights increase monotonically and cycle counts are nonnegative. They allow amplitude positions to vary independently within each bin. Constraints from a complete acceleration history or a particular counting algorithm could narrow the admissible range, but those constraints cannot be recovered from aggregate counts alone. The bounds therefore describe the amplitude uncertainty compatible with the recorded intervals; they do not assert that every bound is attained by a physically realizable history.

Verifying earthquake identities would strengthen the interpretation of excluded-group performance. Raw cycle amplitudes would permit direct evaluation of midpoint error. Soil-response observations would be necessary to test whether either moment predicts a material endpoint. The available

acceleration counts support conclusions about processing and amplitude resolution, rather than those unmeasured responses.

## 7. CONCLUSION

The analysis provides an audit of loading representations used before a soil-response calculation. It supports reporting the selected counting procedure, duration window, amplitude binning, and weighting exponent alongside any cumulative-demand summary. A single midpoint moment can otherwise conceal both processing alternatives and the range of amplitudes compatible with the stored bins. The study does not infer liquefaction occurrence, cyclic resistance, pore-pressure generation, residual strength, or settlement. No soil specimens, stress paths, groundwater conditions, or damage observations enter the analysis. The exponents are illustrative sensitivity parameters. Acceleration weighting cannot be assumed equivalent to a site's cyclic shear-stress weighting, and mixing orientations prevents a horizontal-only engineering interpretation. The group labels are reconstructed from component identifiers without an external event catalogue. Therefore nominal group coverage is conditional on the grouping being appropriate; it is not independently established earthquake-level coverage. The PGA restriction affects the target population, the fixed split gives only one validation realization, and no uncertainty estimate treats the 1,452 held-out components as independent earthquakes. An external mapping audit and further grouped validation are necessary before deployment. Amplitude-resolution bounds and algorithm-window envelopes describe different perturbations. Their joint evaluation identifies the numerical uncertainty admitted before a material-response calculation. The calculations show that the median spread across processing choices is nearly a factor of two for the quadratic moment in this selected population. Bin-width uncertainty can be of comparable order and should be calculated separately. Fitted models achieve high observed group coverage only with broad intervals. These conclusions concern acceleration-cycle representation; applying them to liquefaction requires a calibrated soil-response relationship.

## CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest related to this research.

## DATA AVAILABILITY

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

## FUNDING INFORMATION

The authors received no external funding for this research.

## REFERENCES

1. A.H. Liu, J.P. Stewart, N.A. Abrahamson, Y. Moriwaki, Equivalent number of uniform stress cycles for soil liquefaction analysis, Journal of Geotechnical and Geoenvironmental Engineering 127(12) (2001). https://doi.org/10.1061/(ASCE)1090-0241(2001)127:12(1017).
2. R.A. Green, G.A. Terri, Number of equivalent cycles concept for liquefaction evaluations—revisited, Journal of Geotechnical and Geoenvironmental Engineering 131(4) (2005). https://doi.org/10.1061/(ASCE)1090-0241(2005)131:4(477).
3. T.D. Ancheta et al., NGA-West2 Database, Earthquake Spectra 30(3) (2014) 989–1005. https://doi.org/10.1193/070913EQS197M.
4. E. Seyhan, J.P. Stewart, T.D. Ancheta, R.B. Darragh, R.W. Graves, NGA-West2 Site Database, Earthquake Spectra 30(3) (2014) 1007–1024. https://doi.org/10.1193/062913EQS180M.
5. Yousef Bozorgnia, Norman A. Abrahamson, Linda Al Atik, Timothy D. Ancheta, Gail M. Atkinson, Jack W. Baker et al., NGA-West2 Research Project, Earthquake Spectra 30 (2014) 973-987. https://doi.org/10.1193/072113eqs209m.
6. Kioumars Afshari, Jonathan P. Stewart, Physically Parameterized Prediction Equations for Significant Duration in Active Crustal Regions, Earthquake Spectra 32 (2016) 2057-2081. https://doi.org/10.1193/063015eqs106m.
7. Jack W. Baker, Cynthia Lee, An Improved Algorithm for Selecting Ground Motions to Match a Conditional Spectrum, Journal of Earthquake Engineering 22 (2018) 708-723. https://doi.org/10.1080/13632469.2016.1264334.
8. N. Jayaram, T. Lin, J.W. Baker, A Computationally Efficient Ground-Motion Selection Algorithm for Matching a Target Response Spectrum Mean and Variance, Earthquake Spectra 27(3) (2011) 797–815. https://doi.org/10.1193/1.3608002.
9. Jack W. Baker, Conditional Mean Spectrum: Tool for Ground-Motion Selection, Journal of Structural Engineering 137 (2011) 322-331. https://doi.org/10.1061/(asce)st.1943-541x.0000215.
10. R.W. Boulanger, I.M. Idriss, CPT-Based Liquefaction Triggering Procedure, Journal of Geotechnical and Geoenvironmental Engineering 142(2) (2016) 04015065. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001388.
11. Jonathan D. Bray, Jorge Macedo, 6th Ishihara lecture: Simplified procedure for estimating liquefaction-induced building settlement, Soil Dynamics and Earthquake Engineering 102 (2017) 215-231. https://doi.org/10.1016/j.soildyn.2017.08.026.
12. J. Macedo, J.D. Bray, Key Trends in Liquefaction-Induced Building Settlement, Journal of Geotechnical and Geoenvironmental Engineering 144(11) (2018). https://doi.org/10.1061/(ASCE)GT.1943-5606.0001951.

13. B. Carlton, S. Mazzoni, Y. Bozorgnia, EQ Ground-Motion Cycle Count Database: NGA-West2 PGA-Bins Dataset, Zenodo (2025). https://doi.org/10.5281/zenodo.15015061.

14. B. Carlton, S. Mazzoni, Y. Bozorgnia, Earthquake Ground Motion Cycle Count Datasets, Earthquake Spectra (2026). https://doi.org/10.1002/esp4.70115.

15. P. Geurts, D. Ernst, L. Wehenkel, Extremely randomized trees, Machine Learning 63 (2006) 3–42. https://doi.org/10.1007/s10994-006-6226-1.

16. A.N. Angelopoulos, S. Bates, A gentle introduction to conformal prediction and distribution-free uncertainty quantification, arXiv:2107.07511. https://arxiv.org/abs/2107.07511.

17. D.M. Boore, J.P. Stewart, E. Seyhan, G.M. Atkinson, NGA-West2 Equations for Predicting PGA, PGV, and 5% Damped PSA for Shallow Crustal Earthquakes, Earthquake Spectra 30(3) (2014) 1057–1085. https://doi.org/10.1193/070113EQS184M.

18. Brian S.-J. Chiou, Robert R. Youngs, Update of the Chiou and Youngs NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra, Earthquake Spectra 30 (2014) 1117-1153. https://doi.org/10.1193/072813eqs219m.

19. K.W. Campbell, Y. Bozorgnia, NGA-West2 Ground Motion Model for the Average Horizontal Components of PGA, PGV, and 5% Damped Linear Acceleration Response Spectra, Earthquake Spectra 30(3) (2014) 1087–1115. https://doi.org/10.1193/062913EQS175M.

20. Norman A. Abrahamson, Walter J. Silva, Ronnie Kamai, Summary of the ASK14 Ground Motion Relation for Active Crustal Regions, Earthquake Spectra 30 (2014) 1025-1055. https://doi.org/10.1193/070913eqs198m.

21. D.R. Roberts et al., Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure, Ecography 40 (2017) 913–929. https://doi.org/10.1111/ecog.02881.

22. F. Pedregosa et al., Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research 12 (2011) 2825–2830. https://jmlr.org/papers/v12/pedregosa11a.html.

23. Cynthia Rudin, Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead, Nature Machine Intelligence 1 (2019) 206-215. https://doi.org/10.1038/s42256-019-0048-x.

24. S. Athey, J. Tibshirani, S. Wager, Generalized random forests, Annals of Statistics 47(2) (2019) 1148–1178. https://doi.org/10.1214/18-AOS1709.

25. Y. Romano, E. Patterson, E.J. Candès, Conformalized Quantile Regression, Advances in Neural Information Processing Systems 32 (2019). https://arxiv.org/abs/1905.03222.

26. Rina Foygel Barber, Emmanuel J. Candès, Aaditya Ramdas, Ryan J. Tibshirani, Predictive inference with the jackknife+, The Annals of Statistics 49 (2021) 486-507. https://doi.org/10.1214/20-aos1965.

27. Rina Foygel Barber, Emmanuel J Candès, Aaditya Ramdas, Ryan J Tibshirani, The limits of distribution-free conditional predictive inference, Information and Inference: A Journal of the IMA 10 (2021) 455-482. https://doi.org/10.1093/imaiai/iaaa017.

28. R. Foygel Barber, E.J. Candès, A. Ramdas, R.J. Tibshirani, Conformal prediction beyond exchangeability, preprint (2022), arXiv:2202.13415. https://arxiv.org/abs/2202.13415.

29. R.J. Tibshirani, R. Foygel Barber, E.J. Candès, A. Ramdas, Conformal Prediction Under Covariate Shift, Advances in Neural Information Processing Systems 32 (2019). https://arxiv.org/abs/1904.06019.

30. Emel Seyhan, Jonathan P. Stewart, Semi-Empirical Nonlinear Site Amplification from NGA-West2 Data and Simulations, Earthquake Spectra 30 (2014) 1241-1256. https://doi.org/10.1193/063013eqs181m.

31. R. Kamai, N.A. Abrahamson, W.J. Silva, Nonlinear Horizontal Site Amplification for Constraining the NGA-West2 GMPEs, Earthquake Spectra 30(3) (2014) 1223–1240. https://doi.org/10.1193/070113EQS187M.

32. D.M. Boore, Orientation-Independent, Nongeometric-Mean Measures of Seismic Intensity from Two Horizontal Components of Motion, Bulletin of the Seismological Society of America 100(4) (2010) 1830–1835. https://doi.org/10.1785/0120090400.

33. Meera Raghunandan, Abbie B. Liel, Effect of ground motion duration on earthquake-induced structural collapse, Structural Safety 41 (2013) 119-133. https://doi.org/10.1016/j.strusafe.2012.12.002.

34. R. Chandramohan, J.W. Baker, G.G. Deierlein, Quantifying the Influence of Ground Motion Duration on Structural Collapse Capacity Using Spectrally Equivalent Records, Earthquake Spectra 32(2) (2016) 927–950. https://doi.org/10.1193/122813EQS298MR2.

35. Roxane Foulser-Piggott, Katsuichiro Goda, Ground-Motion Prediction Models for Arias Intensity and Cumulative Absolute Velocity for Japanese Earthquakes Considering Single-Station Sigma and Within-Event Spatial Correlation, Bulletin of the Seismological Society of America 105 (2015) 1903-1918. https://doi.org/10.1785/0120140316.

36. William Araujo, Christian Ledezma, Factors That Affect Liquefaction-Induced Lateral Spreading in Large Subduction Earthquakes, Applied Sciences 10 (2020) 6503. https://doi.org/10.3390/app10186503.

37. I. Zaliapin, Y. Ben-Zion, Earthquake Declustering Using the Nearest-Neighbor Approach in Space-Time-Magnitude Domain, Journal of Geophysical Research: Solid Earth 125 (2020). https://doi.org/10.1029/2018JB017120.

38. M. Abdar et al., A review of uncertainty quantification in deep learning: Techniques, applications and challenges, Information Fusion 76 (2021) 243–297. https://doi.org/10.1016/j.inffus.2021.05.008.