The Cost of Writing Down a Number

Michael P. Laing
mlaing@post.harvard.edu
Written with Anthropic Claude Fable 5

August 2026

Companion essay to "Measure-Relational Algebra: The Core" (arXiv, August 2026). Canonical reference: realql.org.

A forty-year ceiling that wasn't there

For four decades, geophysicists believed Earth's magnetosphere saturates. Push the solar wind harder and harder, and the planet's magnetic response bends over and flattens — a ceiling. The curve was smooth, reproducible, and physically plausible. At least nine mechanisms were proposed to explain it. It was built into the models used to estimate what an extreme solar storm would do to power grids and satellites.

In 2026, a paper in *Nature* (Sivadas et al., **655**, 1143–1147) showed that the observations give no evidence of a ceiling. The driver of the response — an electric field carried by the solar wind — cannot be measured where it acts. It is measured by spacecraft 1.5 million kilometers upstream and propagated forward through an uncertain transit. Each recorded value was therefore not the true driver but a noisy observation of it. Everyone knew this. It was written in the data documentation. And then every analysis for forty years did the same thing: it took the recorded number as the value, and computed onward.

That single act — recording a measurement as if it were exact — is enough, all by itself, to bend a straight line. Extreme recorded values are, on average, less extreme in truth: a reading in the far tail is more likely to be a moderate truth plus a large error than an extreme truth measured cleanly. Regress a response against such records and the top of the curve droops. The droop looks exactly like saturation. Correct for the known measurement uncertainty and the relationship is linear across the whole observed range — which means that, extrapolated without a ceiling, the geomagnetic impact of a Carrington-class storm is roughly **twice** what the saturating models predicted.

Forty years of theory addressed an artifact of the data model.

The mechanism is general, and it has a direction

Nothing in that story is specific to space physics. The ingredients are three, and they are everywhere:

1. **A quantity known imperfectly.** Every measurement delivers not a value but a state of knowledge about a value — an error bar at minimum, a distribution in honest practice. This is the founding premise of metrology.

2. **A record that keeps the number and discards the knowledge.** Databases, spreadsheets, data files, and the pipelines between them store points. The error model, if it survives at all, survives in a PDF nobody joins against.

3. **A nonlinear conclusion.** A square, a product, an exponential, a threshold, a ratio, a maximum. Almost everything interesting is nonlinear.

Combine the three and the result is not noise. It is **bias with a sign** — fixed by the curvature of the calculation, growing with the uncertainty that was discarded, and completely immune to averaging. Take a million more measurements through the same instruments and the artifact does not shrink; it sharpens. That is what makes it dangerous: it presents as a clean, reproducible empirical law. There is nothing in the recorded data to mark it as spurious, because the mark was discarded at the moment of recording.

The mathematics behind this is old — Jensen's inequality, regression dilution, errors-in-variables. Statisticians have owned the corrective tools for a century. The point of the accompanying paper is not new mathematics. It is a proof of exactly where the boundary lies: collapsing a distribution to a point commutes with a calculation **if and only if** the calculation is affine — sums, scalings, averages — or the value was exact to begin with. Inside that boundary, dropping the error bar is provably free. Outside it, the collapsed computation is not an approximation of the correct one. It is a *different* computation, and it can manufacture structure — ceilings, thresholds, power laws — that the underlying system does not possess.

Why this is urgent now

The failure mode has existed as long as measurement. Three things have changed.

We now live at the extremes. Climate records are not occasionally broken; record-breaking is the operating regime. And extremes are precisely where the mechanism bites hardest, twice over. Measurements are least reliable at the edge of an instrument's range, at the edge of a sensor network's coverage, in the reconstructed past of a proxy record. And the conclusions drawn from extremes are the most nonlinear we compute: damage scales as a high power of wind speed; flood loss is a threshold function of water level; mortality curves hockey-stick in heat. Convex conclusion, uncertain input, point-valued record: the estimated consequences of the events that matter most are systematically distorted in ways the recorded data cannot reveal. To be precise about the claim: this does not mean any particular climate result is wrong. It means that under point-valued data practice, *the audit that would tell you is prohibitively manual* — it must be reconstructed by hand, per analysis, by someone who already suspects the problem. The Sivadas result is what such an audit finds when someone finally performs one, forty years late.

Risk analysis is nonlinearity applied to uncertainty, industrially. Value-at-risk, tail expectations, stress tests, insurance pricing, reinsurance layering: these are convex functionals evaluated on estimated quantities — correlations, volatilities, exposures — every one of which is a measurement with error, stored as a point. An unstable world moves these estimates fast and widens their uncertainty exactly when their tails are being interrogated. A risk number computed from collapsed inputs is not conservative or anticonservative by accident; the direction of its error is determined by curvature, and it is knowable — but only if the uncertainty is still there to propagate.

The pipeline no longer has a human in the middle. The classical defense against all of this was the vigilant analyst — the statistician who suspects error-in-variables and reaches for the correction. That defense never scaled, and it is now being removed entirely. Data flows from sensor to database to model to decision, increasingly with machine learning systems both consuming and producing the numbers. A model trained on collapsed data learns the artifacts as if they were physics — it will faithfully learn the saturation ceiling that isn't there. Its outputs, stored again as points, seed the next round. Automation does not merely fail to catch the bias; it launders and compounds it.

Vigilance does not scale. Structure does.

The instinctive response — *analysts should be more careful* — misdiagnoses the problem. In the geomagnetic case, no one was careless. The uncertainty was documented. The corrective statistics sat in textbooks. What was missing was any structural connection between the two: the data model in which everyone worked had no place to put a distribution, so the distribution went nowhere, and forty years of careful people computed carefully on numbers that had already discarded what mattered.

This is a data-model problem, and data-model problems have data-model solutions. Database systems already embody one of the great examples: relational normalization moved consistency from a burden on every programmer's vigilance into a property of the schema — you cannot commit the update anomaly, because the structure forbids it. The same move is available here. Let the stored object be the state of knowledge itself — the distribution, with its dependencies — and let "collapse to a single number" be an explicit, recorded operation rather than a silent default at the moment of ingestion. The mathematics then says exactly when the system may collapse early for free (the affine parts — which is why such systems need not be slow), and prices the error everywhere else. Correctness stops depending on whether this year's analyst happens to suspect this dataset's error structure.

The accompanying paper works this out as mathematics: cells as probability measures, dependence carried explicitly, operations as pushforward, the classical model recovered exactly as the special case of exact values, and one theorem marking the boundary where point-valued practice stops being right. Seven open problems — representation, language, optimization, decision operators, a benchmark — mark the distance from mathematics to working systems, and they are offered as work, openly. Everything is published under CC BY; there are no patents and nothing to license. The framework is called Measure-Relational Algebra; the reference point is realql.org.

The claim, plainly

A measurement is a state of knowledge, not a number. Recording the number and discarding the rest is a choice. For seventy years of data processing that choice was invisible, because nothing in our systems could represent the alternative. It is now a proven theorem where the choice is free and where it is not; a *Nature* result showing the price can be forty years of misdirected science and a two-fold error in the risk estimate that mattered; and a century-old body of corrective statistics waiting to be moved from the analyst's vigilance into the data model's structure, where it can finally be applied by default.

We are entering the decades in which humanity's most consequential decisions — about a heating planet, about cascading financial and geopolitical risk, about medicine practiced at machine speed — will be computed from measurements taken at the edge of what our instruments can see, by pipelines with no statistician in the loop. Those decisions inherit the curvature of their calculations and the uncertainty of their inputs whether we record that uncertainty or not. The only question is whether the record keeps what we actually know — or keeps a number, and lets the difference masquerade as a law of nature.

---

*The paper: M. P. Laing, "Measure-Relational Algebra: The Core," arXiv, August 2026. The motivating case: N. Sivadas et al., "Regression to the mean can explain saturation of geomagnetic storms," Nature **655**, 1143–1147 (2026).*