From Raw Signal to Prediction: What Actually Works (and What Doesn't)

This analysis is based entirely on real-world system data — not synthetic examples.
It demonstrates how raw signals can be transformed into structured insight to determine whether prediction is feasible.

All figures are based on real IBM Quantum calibration datasets. Where necessary, signals are aggregated to provide a system-level representation.

In many real-world systems, raw monitoring data appears complex, noisy, and difficult to interpret.

Despite this, predictive models are often applied directly to such data — leading to unstable results, poor generalization, and months of wasted model development effort.

In many cases, the issue is not the model.

The signal was never suitable for prediction in the first place.

The purpose of this document is to show, using real data,
how signal structure emerges through processing — and how this structure determines whether prediction is feasible before modeling begins.

Although this example is based on quantum calibration data, the same behavior is observed across other domains, including vibration monitoring systems and battery degradation datasets.
The purpose of this example is therefore not domain-specific, but to illustrate a general property of signal-based prediction.

Figure 1 — Pure Raw Data (Industry Input)

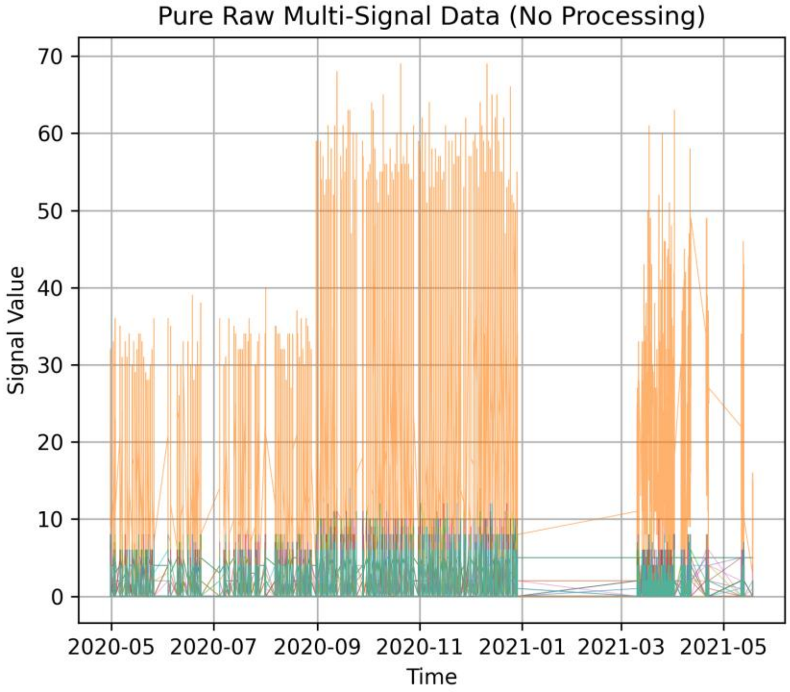

Figure 1 — Pure raw multi-signal data from real quantum hardware. This figure shows unprocessed calibration data obtained from IBM Quantum systems. Multiple overlapping signals (representing different qubits and calibration parameters) are plotted over time without any filtering, aggregation, or transformation. The resulting representation appears complex and difficult to interpret, with no immediately visible instability structure or predictive signal.

This is the stage at which most predictive modeling efforts begin.
At this point, it is not yet known whether the signal contains any structure
that can support reliable prediction.

Figure 2 — Structured Data (First Processing Step)

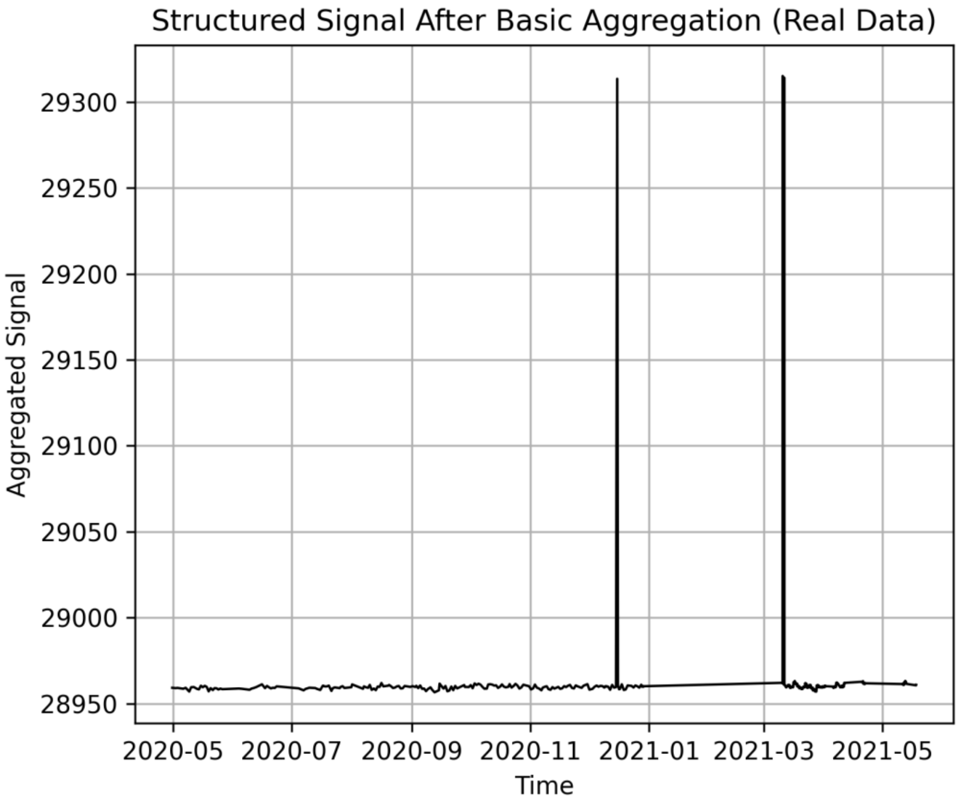

Figure 2 — Structured representation of real-world quantum calibration data. This figure shows the same dataset after basic preprocessing, including time alignment and aggregation across multiple calibration parameters. The resulting signal is more interpretable and reveals overall system behavior, while still containing irregularities, discontinuities, and regime changes typical of real-world systems. No instability detection is applied at this stage.

Although structure becomes visible at this stage,
it is still unclear whether this structure is consistent, reproducible,
or meaningful for prediction.

Figure 3 — Pre-Instability Detection (Analysis Output)

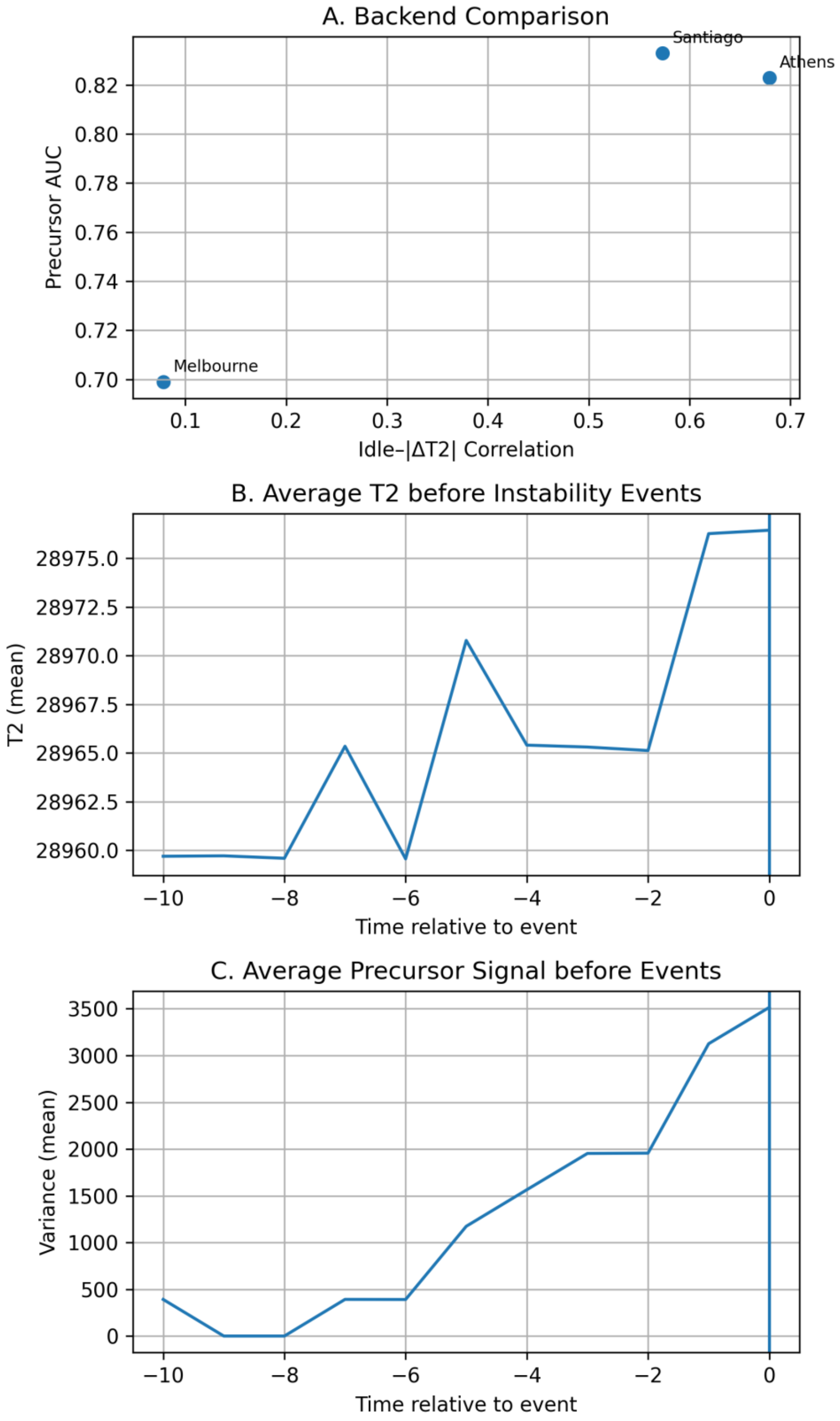

Figure 3 — Pre-instability signal detection using real IBM calibration data. (A) Backend-level comparison of precursor signal strength across IBM Quantum devices. (B) Average signal behavior aligned to instability events (t = 0). (C) Average precursor signal (rolling variance) aligned to the same events, demonstrating that variance increases prior to instability events.

This figure is based on real-world IBM Quantum calibration data. The signal representation is an aggregated proxy derived from multiple calibration parameters to capture system-level behavior.

Interpretation

This sequence of figures illustrates the full transformation pipeline:
Figure 1 shows the raw data as received from industry.
Figure 2 shows the same data after minimal structuring.
Figure 3 shows the extracted analytical insight.

The key question is not how to process the data,
but whether the signal contains structure that is **consistent and reproducible enough to support prediction**.

Without this, predictive modeling is fundamentally unreliable.

Conceptual Interpretation of Pre-Instability Behavior:
The relationship between raw signal behavior, instability events, and precursor signals is not directly visible in real-world data.

The following representation makes this relationship explicit by separating the signal into three components: overall system behavior, event structure, and precursor dynamics.

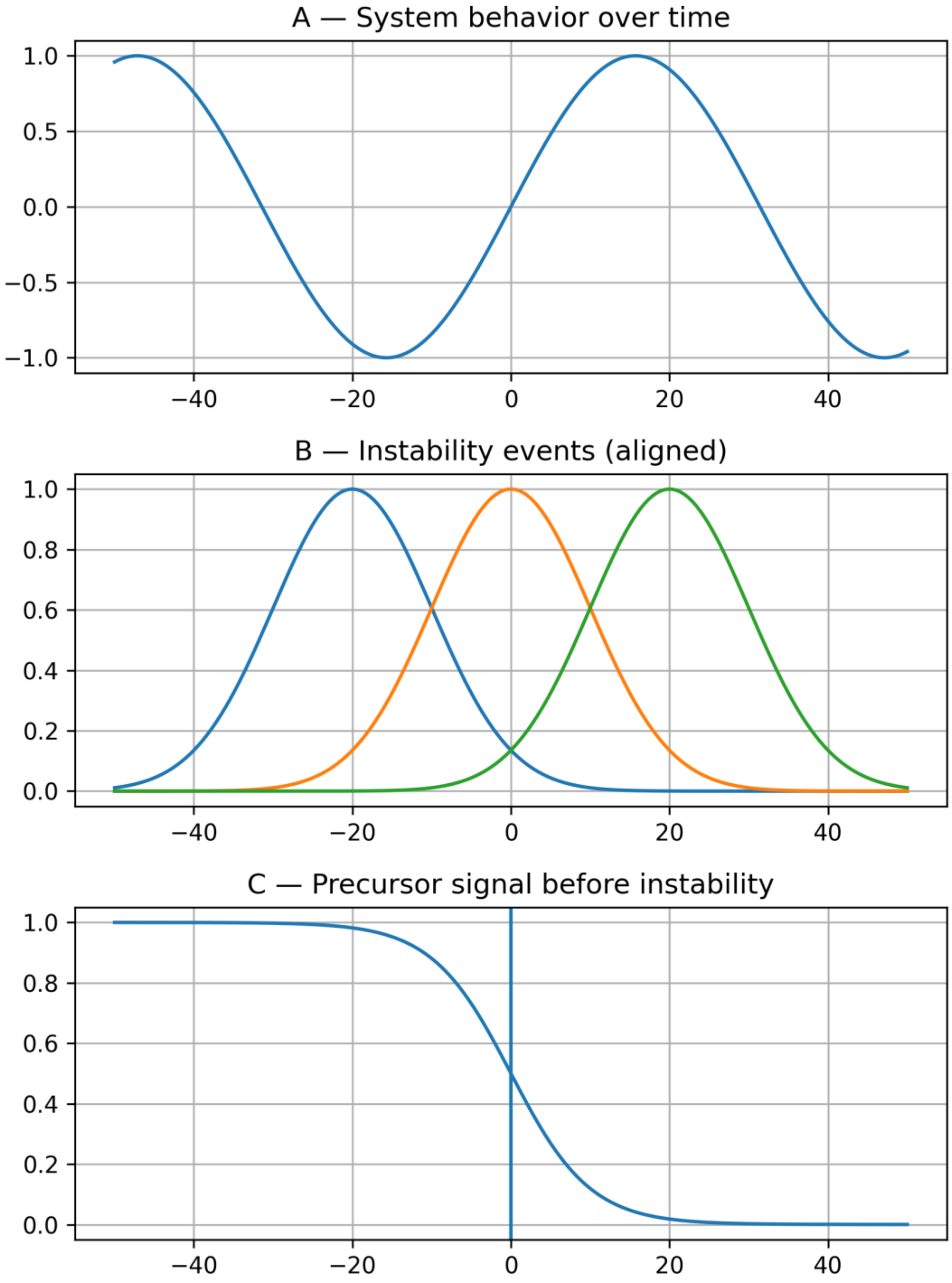

Figure 4 — Conceptual representation of instability detection in quantum hardware.
(A) System behavior over time, showing the raw signal dynamics.
(B) Instability events aligned in time, illustrating that events exhibit structured behavior rather than purely random variation.
(C) Precursor signal behavior prior to instability events, demonstrating that measurable changes occur before the event itself.

This representation makes explicit how instability, structure, and precursor signals are related. This means that detecting structure is not sufficient — the structure must also be consistent and interpretable to be useful. At the same time, it highlights a critical limitation: the presence of detectable structure does not guarantee that this structure is consistent or predictive across observations.

In practice, this occurs frequently:
systems exhibit detectable structural change
that does not translate into usable predictive information.

This leads to a common failure mode:
models are built on signals that appear structured,
but do not support stable or meaningful prediction.

Reproducibility

The analysis presented in this report is fully reproducible using publicly available real-world data.

Dataset source:
IBM Quantum Calibration Dataset (Mendeley)
https://data.mendeley.com/datasets/pmycgb2bt7/1

Data used:
- IBM Athens calibration dataset
- Timestamped calibration parameters (T1, T2, readout error, gate error)

Processing steps:
1. Raw multi-signal calibration data is loaded without modification (Figure 1)
2. Signals are aggregated over time to obtain a structured system-level representation (Figure 2)
3. Instability events are defined as the top 10% largest absolute changes in the signal ($|\Delta T2|$)
4. Precursor features (rolling variance and prior signal changes) are computed
5. Signals are aligned around detected events to evaluate pre-instability behavior (Figure 3)

A complete reproducibility package including raw data, processing scripts, and figures is provided.

Industrial Relevance

This analysis demonstrates that meaningful structure can be extracted from raw time-series data that initially appears unstructured and non-interpretable.

For industrial applications, this has direct implications:
- Raw monitoring data often does not reveal whether prediction is possible
- Basic preprocessing alone does not provide actionable insight
- Only after structured analysis can it be determined whether the signal contains consistent and interpretable behavior

The key outcome is not the detection of instability itself, but the ability to determine whether a signal contains predictive feasibility.
This enables a critical decision step:
GO → proceed with modeling (signal contains consistent and reproducible structure)
LIMITED → modeling may not stabilize (structure present but not reliable)
NO-GO → modeling will not yield reliable results (signal lacks predictive structure)

This is the critical step that should precede any predictive modeling effort:

Can this signal support prediction at all?

If this step is skipped, models are often built on signals
that were never predictive to begin with.

This distinction is essential for reducing wasted effort in model development and focusing resources on signals that contain meaningful information.

In many real-world cases, this step is skipped — leading teams to build models on signals that were never predictive to begin with.

Improving models alone is not sufficient.

Prediction is fundamentally constrained by the information content of the signal.

No model can extract predictive information that is not present.

This is not an edge case —
it is a common failure mode in real-world predictive systems.

Author: Jos Aben
Email: info@tarh-energy.com