The Data-Ingestion Bottleneck for Quantum Weather Models

Compression, Circuit Depth, and Load ability of Real ERA5 Fields

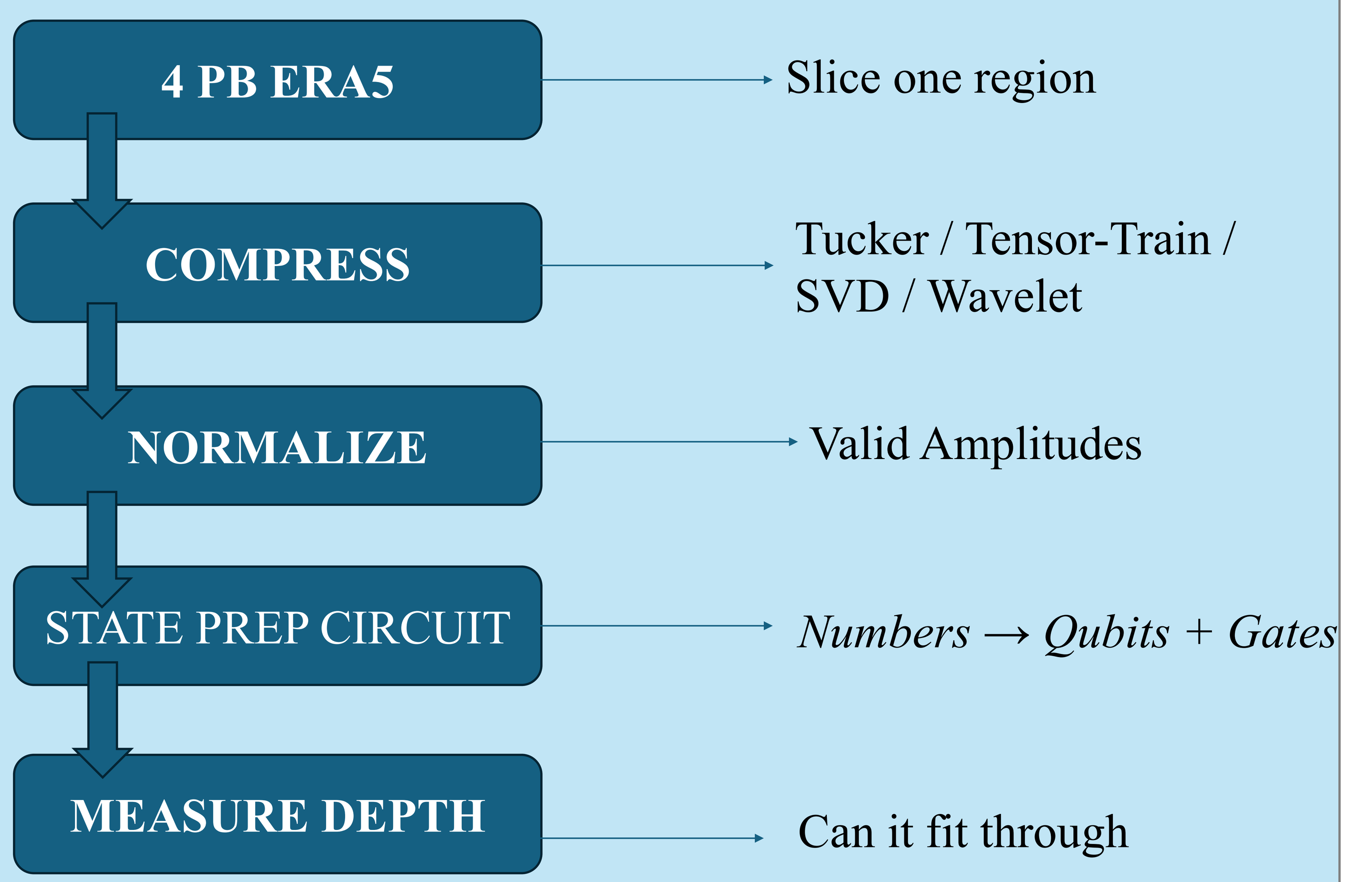
Srikanth Pavuluri
+19402994234
srikanthp2807@gmail.com

1 · OBJECTIVE

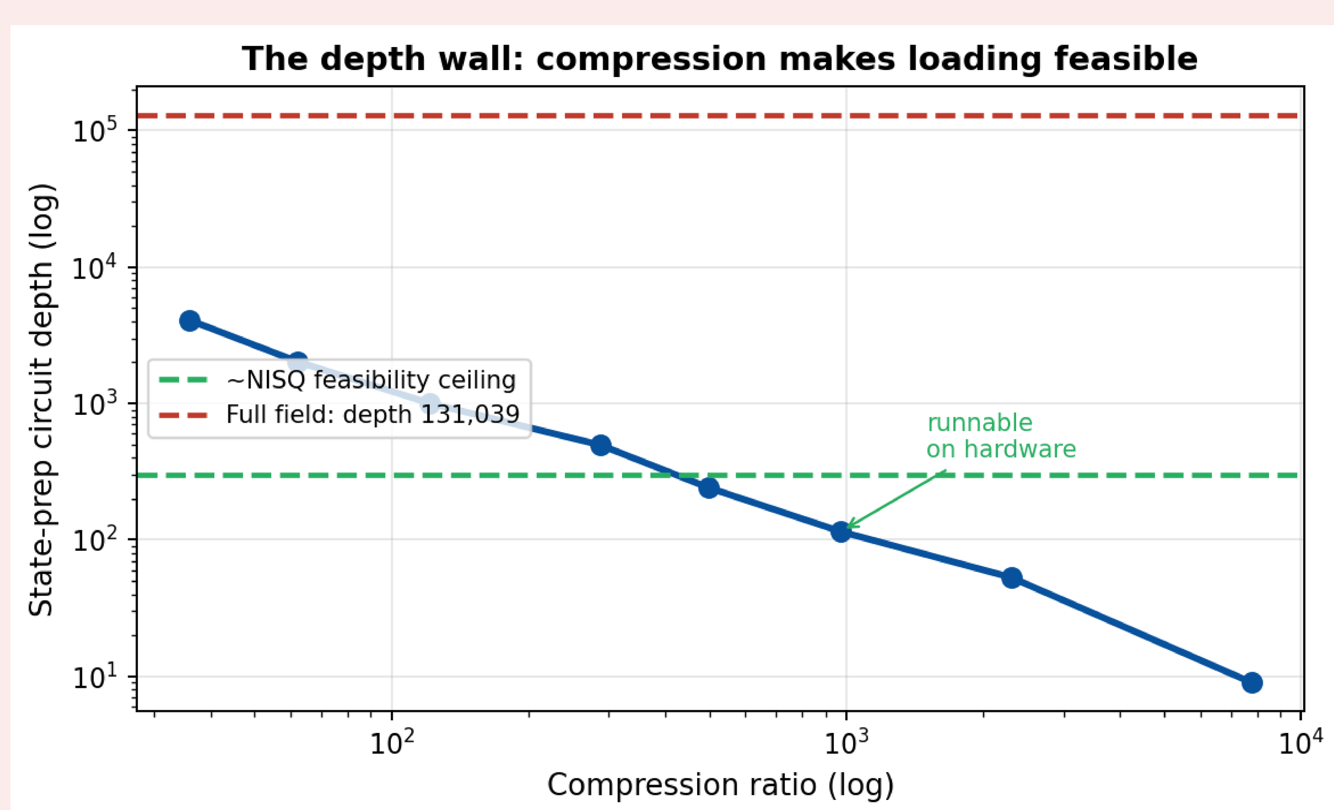
Quantum weather and climate models will need real data loaded into them, but can it fit? This work measures the classical-to-quantum "on-ramp" for real ERA5 climate data: the practical bottleneck between petabyte-scale observations and noisy quantum hardware, where the real limit is circuit depth, not the qubit count..

Data: ERA5 reanalysis (ECMWF), via Google's public ARCO-ERA5 cloud store. Hurricane Ida, 2021-08-29, Gulf of Mexico region, 37 pressure levels.

2 · METHOD — THE PIPELINE



3 · THE DEPTH WALL



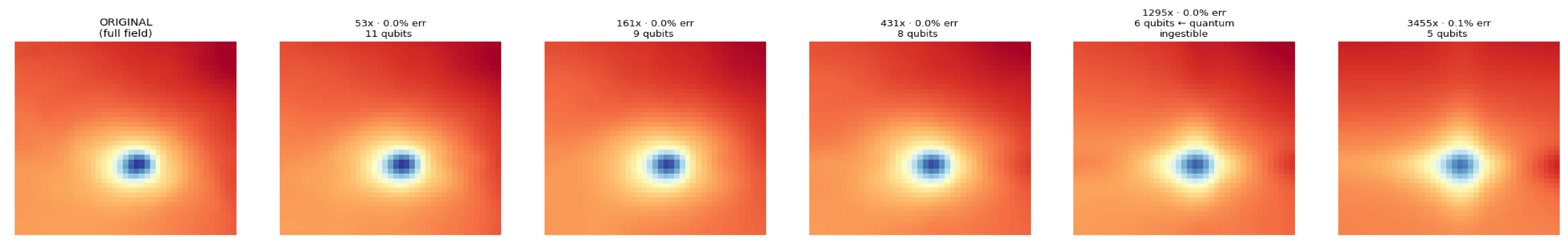
Loading the full hurricane field needs circuit depth 131,039 — about 1000× past what today's noisy hardware can run. Compression collapses that depth into the runnable regime. The bottleneck is Circuit depth, not qubit count.

4 · VARIABLE CONSIDERATION

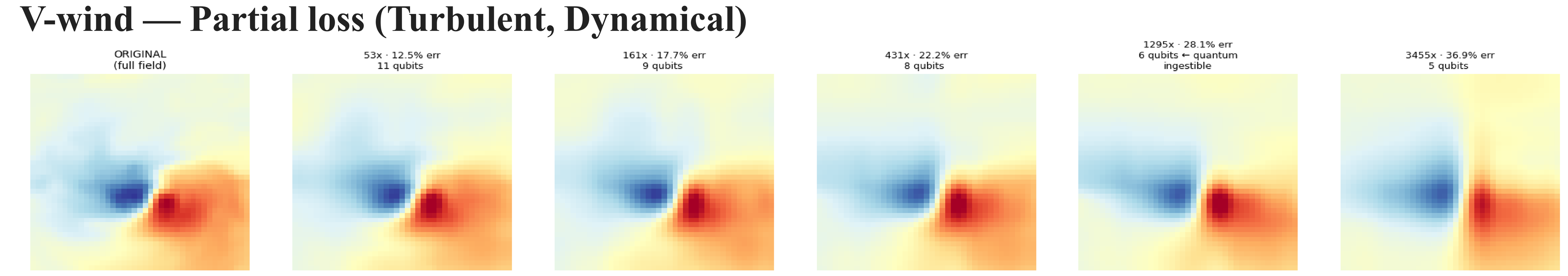
Variables were chosen to span the atmosphere's compressibility spectrum from smooth, large-scale thermodynamic fields to fine-scale turbulent dynamical fields because compressibility (and therefore load ability) is governed by a field's scale structure. All are 3D pressure-level state variables that weather models and data assimilation actually use. Geopotential, Temperature, U,V-wind, Vertical Velocity, Potential Vetrocity, Specific Humidity are considered

5 · COMPRESSION (5 levels) ON EACH VARIABLE (Original to Ingestion Stage)

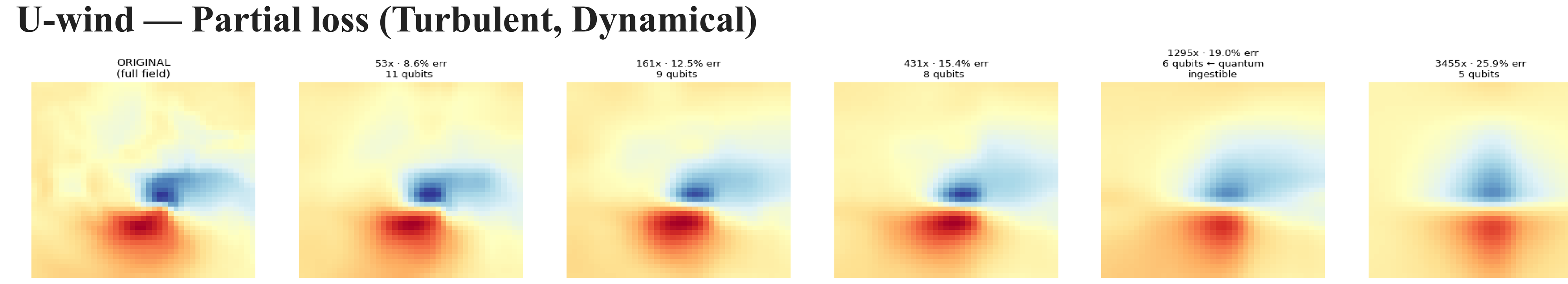
Geopotential — Survives (Smooth, Thermodynamic)



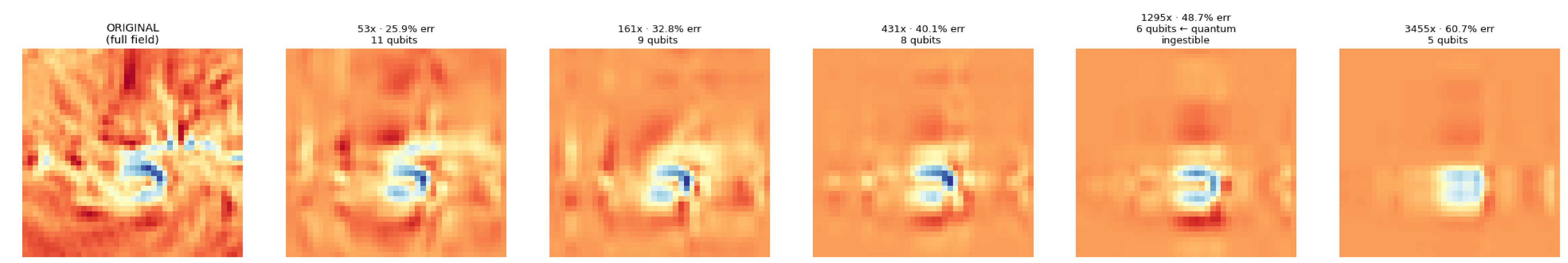
V-wind — Partial loss (Turbulent, Dynamical)



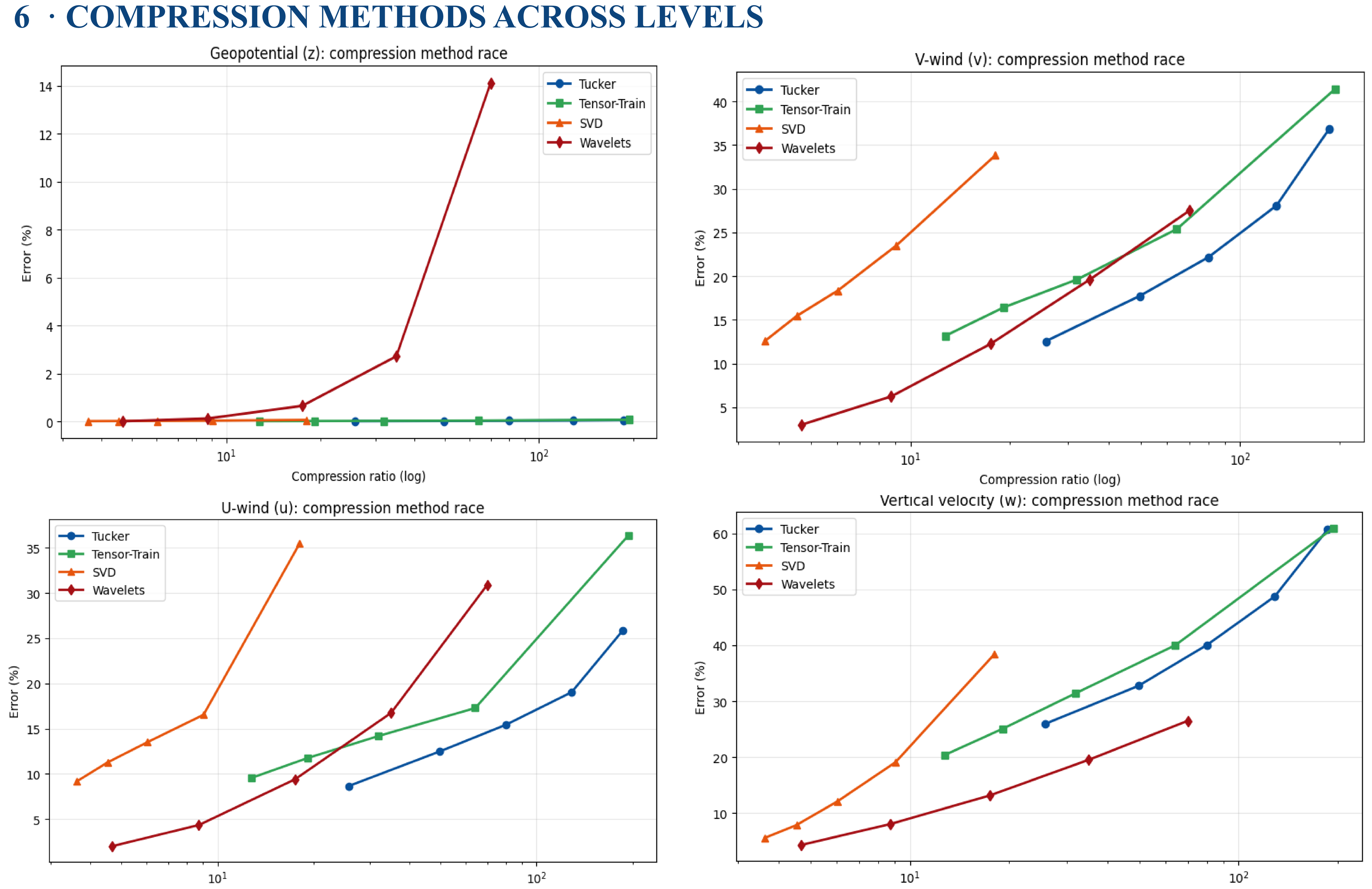
U-wind — Partial loss (Turbulent, Dynamical)



Vertical velocity — Destroyed (Most Turbulent)

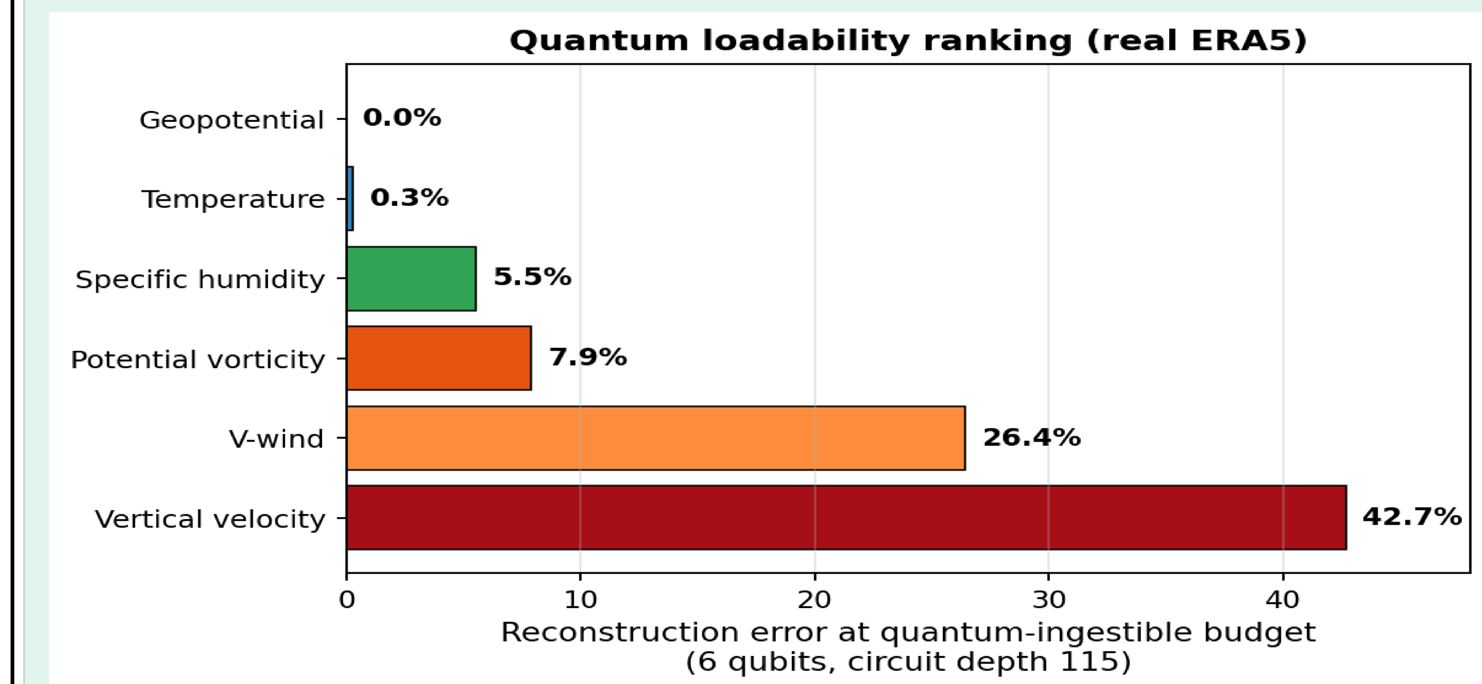


6 · COMPRESSION METHODS ACROSS LEVELS



The crossover: wavelets win at gentle compression (they capture sharp local features), but the low-rank tensor methods (Tucker, tensor-train) catch up or win at aggressive compression precisely the regime quantum ingestion requires. A few global patterns describe a whole field better than a few local bumps.

7 · THE FINDING



The on-ramp is variable-dependent. Smooth thermodynamic fields (geopotential, temperature) load at near-zero cost. Turbulent dynamical fields (wind, vertical velocity) resist and those are exactly the fields forecasts and extreme-event prediction depend on most.
NOTE : Decimal point rounded so 0.1%-0.4% error rounded to 0%

8 · LIMITATIONS ON COMPRESSED DATA

- **Compression deletes fine detail.** It keeps big smooth patterns and throws away small ones so storm cores, fronts, and sharp features are the first to go. It is great loss on physics for weather data.
- **The average error can lie.** A field can show "0.4% error" overall while the one feature you care about (a hurricane eye) is locally destroyed. Always check *where* the error is, not just the number.
- **Turbulent fields have a hard floor.** For wind and vertical velocity, real structure exists at every scale so *no* method can compress them small without heavy loss. It's the physics, not the tool.
- **Results are for one case.** Measured on one hurricane snapshot; other storms, seasons, or resolutions may shift the numbers.
- **Real hardware adds noise on top.** Even a runnable-depth circuit accumulates errors from noisy gates on real NISQ hardware. These depth measurements are for an *ideal* machine a real one degrades the loaded state further, beyond the compression loss.

7 · IDEAS & FUTURE WORK

- **Better methods for turbulent variables.** Wavelets beat the tensor methods at *gentle* compression on turbulent fields, because they capture sharp local features well. But they lose at *aggressive* compression the regime quantum needs because their many localized "bumps" can't describe a whole field from just a few pieces. Finding a method that handles turbulent fields at high compression is open.

- **Domain decomposition (tiling).** Slice the field into tiles and load each one separately fewer qubits, shallower circuits per tile then reassemble. This attacks the depth wall structurally rather than by compression. Limit: atmospheric physics is connected across tile edges (the boundary/halo problem), so isolated tiles have wrong edge physics.

- **Feature extraction on-ramp.** For fields with a coherent feature (a storm, a front), extract and load only the physically meaningful region instead of compressing the whole field. A better on-ramp than blind compression — but only when the physics is localized, not spread across the whole field..