UncertaintyCat: Accessible UQ with OpenTURNS & Generative AI[1]

Mark Legkovskis

PhD in Engineering — University of Warwick

13 June 2025

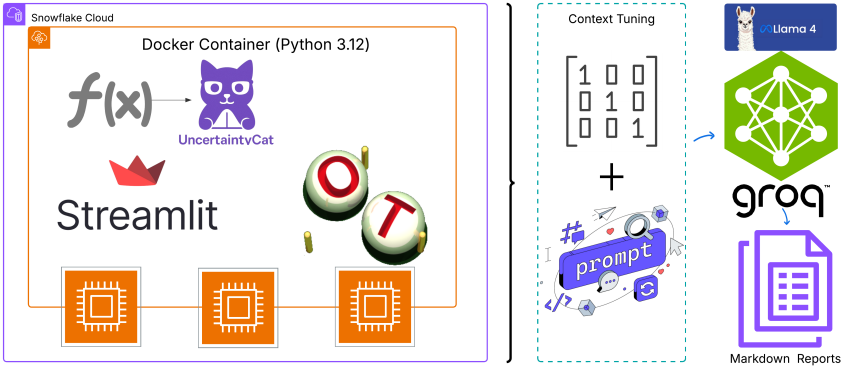
[1] https://uncertaintycat.streamlit.app/

Motivation

- **Born from real industrial need**. During my PhD with **TataSteel**, I built high-dimensional reheating-furnace models. UQ wasn't on the roadmap—until its value became obvious.
- **First Sobol analysis = "wow" moment**. Sobol results pinpointed which inputs truly dominated output variability. Engineers immediately asked for the same insight on more steel-making models.
- **Pain-point**. Each new model meant fresh OpenTURNS code *and* a manual, time-consuming interpretation report.
- **Solution — UncertaintyCat**. One web app that runs UQ with a few clicks and delivers AI-generated, engineer-ready commentary—streamlining the whole cycle from model to actionable insight.

Cloud-Native Architecture (Big Picture)

- **Zero install for users**. Everything runs through a browser over HTTPS.
- **Streamlit Cloud workspace**. Managed Docker (Python 3.12) hosts UncertaintyCat and executes `OpenTURNS` completely in-memory. *User model code is never persisted.*
- **Groq Cloud LLM API**. Numeric results stream to Gemma-2 / Llama-4; the LLM returns compact Markdown insight for the dashboard.
- **End-to-end flow**. Browser → Streamlit (compute) → Groq (LLM) → Browser. (typical full-run wall-clock < 3 min on free tier)



Workflow — From Code / Data to Insight

1. **Pick a Workspace Tab**
   - *UQ Dashboard* — full uncertainty analysis
   - *Dimensionality Reduction* — Morris screening
   - *Distribution Fitting* — fit marginals / copula from data
2. **Select an LLM model** (Llama-4, Gemma-2, DeepSeek 70B) in the sidebar.
3. **Provide Inputs**
   - Dashboard & Dim-Red: paste or write a single-output `OpenTURNS PythonFunction`. Click Validate Model to catch typos early.
   - Distribution Fitting: upload CSV/Excel *or* paste tabular data.
4. **Configure Analyses**
   - *Run Full UQ Suite* — executes Model Understanding, Expectation Convergence, Sobol, FAST, ANCOVA, HSIC, Taylor, Shapley-ML, ... in one go
   - *or* press any individual-module button for a faster, targeted run
   - Dim-Red tab: set *Number of Trajectories*, *Number of Levels*, -threshold
   - Dist-Fit tab: choose distribution families max candidates
5. **Run**. A progress bar tracks each job; results live only in session memory (no model code stored).
6. **Review & Ask**. Interactive Plotly plots + an LLM side-panel chat that already *knows* every analysis you ran — e.g. *"Compare Sobol vs FAST indices and highlight discrepancies."*

Feature Highlights

- **Complete UQ & SA toolkit**. • MC expectation-convergence • Sobol • FAST • ANCOVA • Morris (screening) • HSIC • Taylor • Shapley-ML • Correlation • Exploratory-Data — *11 engines, the modern UQ canon*.
- **Engine-by-engine AI insight**. Each module makes its own Groq LLM call, bundling raw metrics *plus* the relevant OpenTURNS theory. Output: concise, context-aware briefs matched to the plot in view.
- **Dynamic visuals at any scale**. Auto-generated Plotly dashboards adapt from 3-D to $\sim$20-D models; hover tool-tips, filtering and export-to-PNG come standard.
- **Minutes, not days**. A full eleven-module run typically finishes in < 3 min, replacing a week of manual coding, plotting and report-writing.
- **Data-driven inputs**. *Distribution-Fitting Wizard* and *Morris Dim-Red* tabs loop straight back into the main workflow for refined second runs.
- **Built-in benchmarks**. Borehole, Beam-Deflection, Morris, Oscillator, ... ready for demos, teaching and regression tests.

Positioning & Roadmap

Sweet-spot use-cases

- Teaching, quick onboarding, proof-of-concepts
- Rapid "what-if" loops on **10^3–10** samples (runs comfortably on a laptop or single cloud core)
- SMEs & R&D teams that have models but *no UQ pipeline*
- Fully offline deployment with local LLM weights — zero data leaves site

Out of scope

- Replacing billion-sample, HPC-grade brute-force studies on multi-million-cell CFD/FEA models
- Custom pipeline design — engines are pre-wired (Saltelli-Sobol, FAST, ANCOVA, . . .); bring a valid `OpenTURNS` model, click Run, no code-level knob-twiddling
- Enterprise data-governance programmes (handled upstream by the user's own infrastructure)

Roadmap — open for collaborators

- Paste equation $\rightarrow$ auto-generated Python model (no manual coding)
- **Scalable back-ends**: GPU queue / Ray cluster to push MC sweeps to $> 10^6$ samples when needed
- **One-click Docker bundle**: local or private-cloud install with optional on-device LLM weights
- Pilot studies sought with academia & industry (materials, energy, aerospace) — contribute notebooks, datasets, feature PRs

`https://github.com/MLegkovskis/UncertaintyCat`