



RELIABLE, TRUSTED EVALUATIONS

BEACON

AI TESTBED

Benchmark Evaluations for AI & Conventional NWP

Countless atmospheric emulation models exist, and countless more are innovated at lightning speed to meet the needs of every imaginable forecast system. The Research Applications Laboratory's (RAL) AI/ML testbed, BEACON, provides decision-makers physics-based evaluations that cut through hype and discern bona fide capability. By comparing new model candidates directly against established sources of weather and climate data, we deliver transparent benchmarks of performance, strength, and readiness. With this guidance, sponsors can adopt trustworthy AI faster while ensuring the scientific rigor that makes forecasts reliable.



YOUR TRUSTED PARTNER

RAL, part of the NSF National Center for Atmospheric Research (NCAR), has a long track record of turning weather, climate, and environmental science into solutions for aviation, energy, transportation, national security, and public safety. Building on this foundation, BEACON offers sponsors an independent, side-by-side evaluation of emerging AI models against proven forecasting systems. By applying standardized datasets, transparent metrics, and reproducible workflows, we show where AI adds value, where it still needs work, and when it is truly ready for operations.

Benefits & Impacts

- Authoritative AI model benchmarking
- Industry-standard performance metrics
- Trusted validation protocols
- Definitive AI readiness assessments

NSF NCAR RESEARCH APPLICATIONS LABORATORY
... BEACON AI TESTBED ...



••• Trusted validation protocols • Definitive AI Readiness Assessments •••
 Authoritative AI Model Benchmarking • Industry-Standard Performance Metrics

SMARTER DECISIONS, REAL RESULTS

This approach does more than raise scientific confidence – it reduces financial risk. For renewable energy providers, better forecasts translate into more efficient integration of wind and solar power and fewer imbalance penalties. For agriculture, more reliable seasonal outlooks mean smarter water and fertilizer use, saving input costs while protecting yields. By validating AI tools before deployment, sponsors avoid costly missteps and ensure resources are invested in solutions that deliver measurable impact.

- **Renewable energy** forecasting for grid integration & weather/climate adaptation
- **Transportation** safety through road-weather & aviation hazard tools
- **Air quality** & health risk prediction
- **Wildfire** onset, growth, & remote sensing
- **Extreme weather** risk mapping for infrastructure & real estate
- **Coastal flood** & surge modeling
- **Agricultural** planning through soil-moisture & crop forecasting

TRUSTED AI STANDARDS

AI/ML extends physics-based approaches by rapidly processing data, uncovering patterns, and strengthening decision support for security, resilience, and partner operations. What's missing in today's fast-moving market is a trusted referee—and that's where RAL delivers. As the community's independent testbed, RAL benchmarks AI tools against proven weather and climate systems, using NSF NCAR's internationally adopted METplus framework to establish standards, build credibility, and guide smarter investments.

EXPERTISE YOU CAN COUNT ON

As a Federally Funded Research and Development Center, RAL is a long-standing, impartial partner relied upon by defense, federal agencies, and industry. We bring decades of expertise and trusted tools, including NCAR's Weather Research and Forecasting (WRF) model, the NCAR Model for Prediction Across Scales (MPAS), the emerging Community Research Earth Digital Intelligence Twin (CREDIT) AI framework, a range of leading AI systems, and the enhanced Model Evaluation Tools (METplus), to every evaluation, ensuring that decisions are grounded in science rather than marketing claims.

LEARN MORE!
ral.ucar.edu



OUR WORK IN ACTION

We've applied AI/ML to more than 30 projects across government, industry, and international partners, sharpening the precision off: