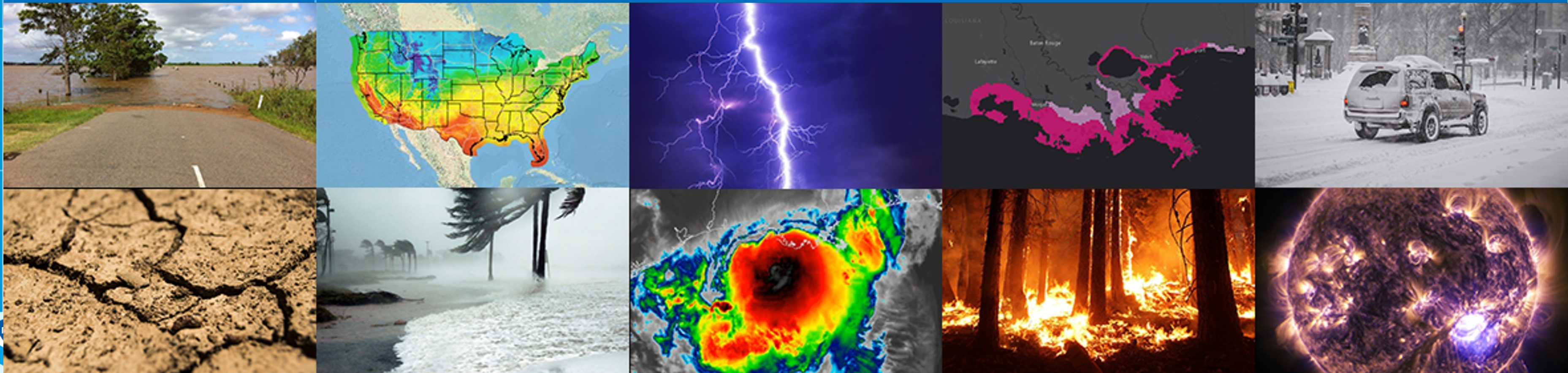
Leveraging AI for the Hurricane Forecast Improvement Program

Kevin Garrett, Office of Science and Technology Integration
Modeling Program

Contributions from: Aaron Poyer, Will Komarami, Jason Anderson, Jack Kain, Deepthi Achuthavarier, HFIP Community

NATIONAL WEATHER SERVICE

November 29, 2023



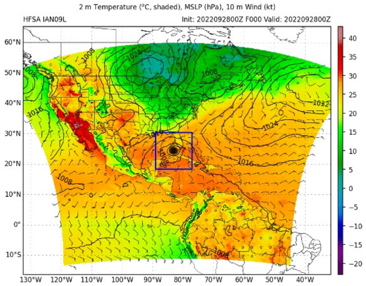
(Soon to be updated)
HFIP Strategic Objectives

- **Improve guidance**
  - Extend forecast guidance from 5-day to 7-day with no loss of skill
  - Halve forecast guidance errors from 2017
  - Develop capabilities for enhanced products based on probabilistic guidance
- **Improved forecasts of TC hazards**
- **Enhance communication of hazard risk and uncertainty**
  - Incorporation of social, behavioral, and economic sciences (SBES) research for more effective Tropical Cyclone (TC) hazard product suite

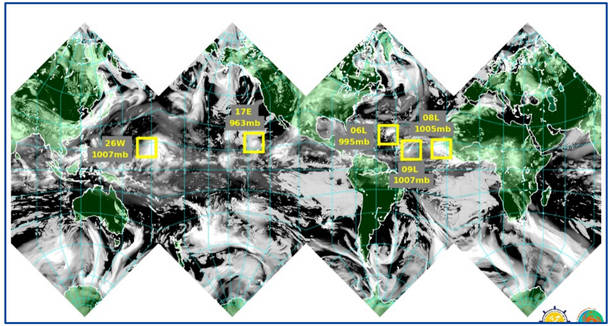
HFIP VISION

Organize the hurricane community to dramatically improve numerical forecast guidance to the National Hurricane Center in 5-10 years.

Get from here…..



to here



FY23
HAFS v1

FY24
HAFS v2

FY25
HAFS v3

FY26
HAFS v4

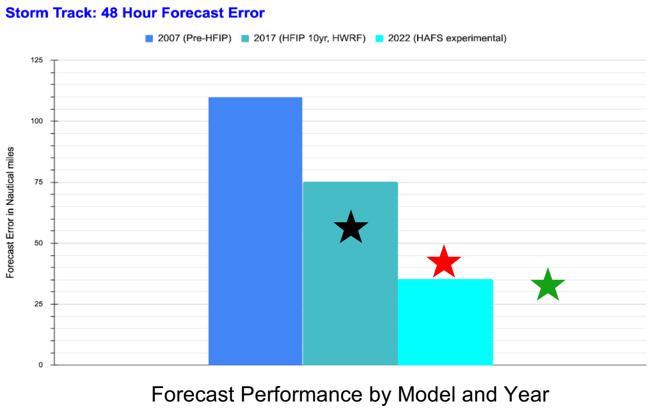
FY27 HAFS
Multiple moving nests in all tropical basins

FY23
SBES project triangulation

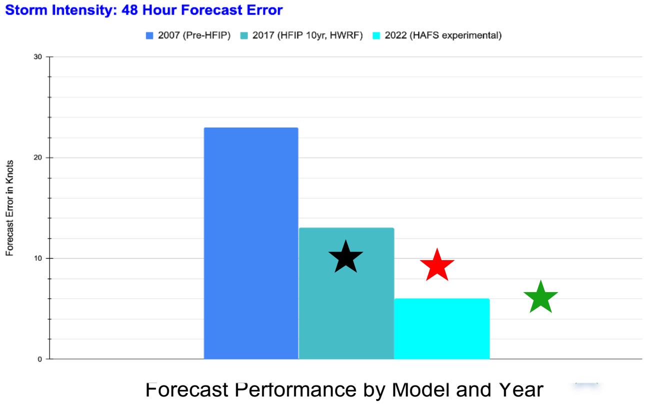
FY24
SBES recommendations mapped to operations

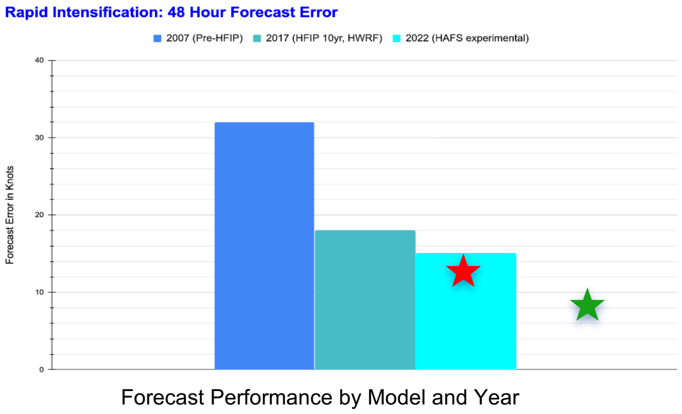
FY25+
New product assessment

HFIP Strategic Objectives (performance)



- **Model Track Error:** Meeting 5-year HFIP goals from 2019
- **Model Intensity Error:** Exceeding HFIP 10-year goals from 2019
- **Model Rapid Intensification:** Close to achieving 5-year HFIP goals from 2019.



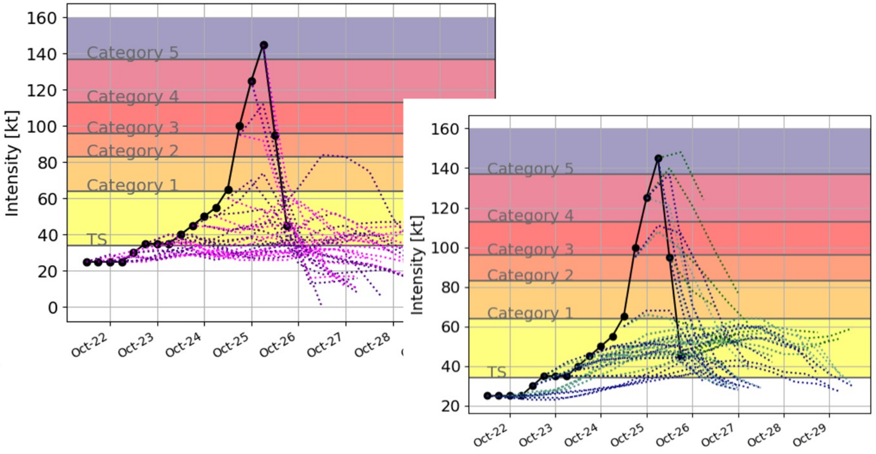


★ Weather Act 10-yr goal

Congress & HEOB

Science Priorities for HAFS

- Improved TC track, intensity, storm size/structure, maximum wind radii, rapid intensification, TC genesis
  - Multiple moving nests in single (global) domain
  - Increased spatial resolution
  - Improved/self-cycled data assimilation
  - Vortex initialization
  - TC physics
  - Atm-Ocean/Atm-land coupling
  - Probabilistic guidance/ensemble configuration



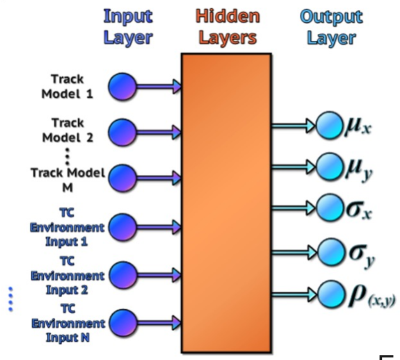
Forecast and estimated intensity for Hurricane Otis. HAFS (left) and statistical models (right) missed Otis RI event (courtesy L. Bucci)

A Machine Learning Model for Estimating Tropical Cyclone Track and Intensity Forecast Uncertainty

DeMaria et al., HFIP Annual Meeting 2023

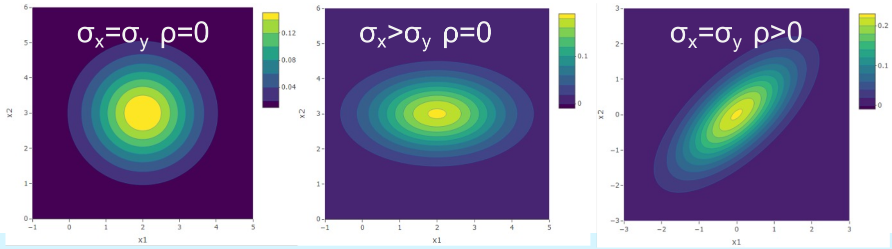
The TCANE Model

- A Machine-Learning model being developed at CIRA under NOAA Hurricane and Ocean Testbed (HOT) support
- Predicts the track and intensity error distributions of the *NHC official forecast*
- Projects ensemble forecast information onto NHC official forecast error distributions

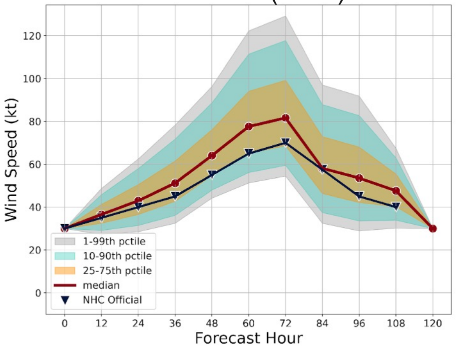


μ_x , μ_y = forecast biases in the x,y directions
σ_x ,σ_y = standard deviations of x and y errors
$\rho(x,y)$ = correlation of x and y errors

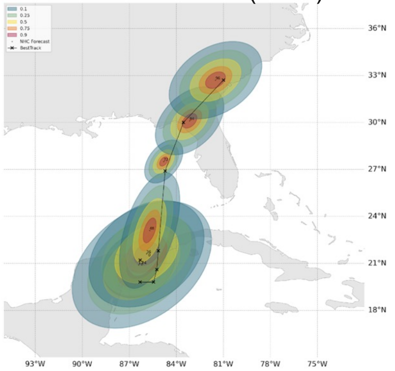
Bivariate normal distributions for various input parameters



Example TCANE intensity forecast for Hurricane Idalia (2023)



Example TCANE track forecast for Hurricane Idalia (2023)



Review: Using Machine Learning for Data Assimilation, Model Physics, and Post-Processing model outputs

Vladimir Krasnopolsky - NCEP/EMC

ML for Initialization/DA

- Fast forward models for direct assimilation of radiances
- Fast observation operators
- Fast ML Models and Adjoints
- Data Pre-processing and Quality Control

ML for post-processing NWP model outputs

- Nonlinear Bias Correction
- Nonlinear MOS
- Nonlinear ensembles
- Nonlinear multi-model ensembles
- Model error prediction
- Model output downscaling
- Tropical cyclones detecting and tracking

ML for model physics

- Fast radiation
- Fast and better microphysics
- New ML parameterizations
- Full suite of model physics
- Fast ML models
- Stochastic physics

ML Application	Methodology developed at EMC	Prototype developed at EMC	Estimated time for development
Forward model and Jacobian	yes	yes	<1 year
Observation operator	yes	yes	1 year
Fast models and adjoints for DAS	no	no	a few years
Data QC and pre-processing	no	no	1 year
Fast radiation	yes	yes	1 year
Fast microphysics	yes	no	1 to several years
New ML parameterizations	yes	no	several years
Full ML physics	yes	no	1 to several years
Fast ML models	no	no	several years
Stochastic physics	yes	no	1 - 3 years
Bias corrections	yes	yes	1 year
Nonlinear MOS	no	no	several years
Nonlinear ensembles	yes	yes	a few years
Nonlinear multi-model ensembles	yes	yes	a few years
Prediction of model errors	no	no	1 to several years
Model output downscaling	no	no	several years
Tropical cyclones detecting and tracking	no	no	several years

Leveraging Machine Learning/AI

- Preprocessing
  - Vortex initialization
- Data Assimilation
  - Fully data-driven analysis
  - Obs operators, QC, bias correction
  - Background errors, handling non-linearities
- TC dynamical model
  - physics
- TC model emulation
  - Just key parameters? (track and intensity, RI)
  - Global or regional (high res/needs structure)
  - Ensembles
  - Downstream applications (surge, QPF)
- Post-processing
  - Model error correction
  - Optimize probabilistic guidance

AI/ML Added Value

Improved science
- Meet HFIP objectives
- Focus on improving specific cases, structure, etc

Increased efficiency
- Increased model cadence
- Reduced latency
- Higher resolution
- More obs assimilated
- Improved uncertainty est.

Drivers
- pIDSS
- Agile/mobile workforce

BACKUP

Can Artificial Intelligence-Based Weather Prediction Models Simulate the Butterfly Effect?

Selz and Craig, GRL 2023

"Current artificial-intelligence-based models cannot simulate the butterfly effect and incorrectly suggest unlimited atmospheric predictability."

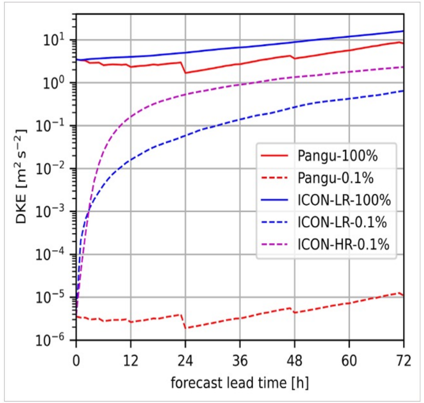


Figure 1 Open in figure viewer PowerPoint

Globally-averaged difference kinetic energy (DKE) as defined by Equation 1 on 300 hPa over time for the different experiments (hourly output time step).

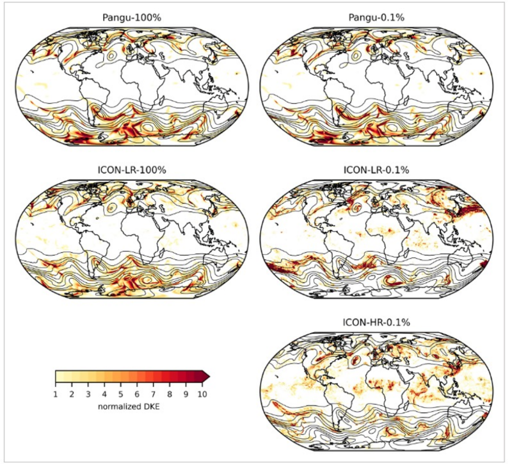


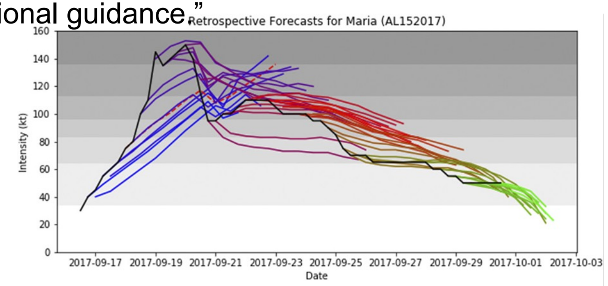
Figure 2 Open in figure viewer PowerPoint

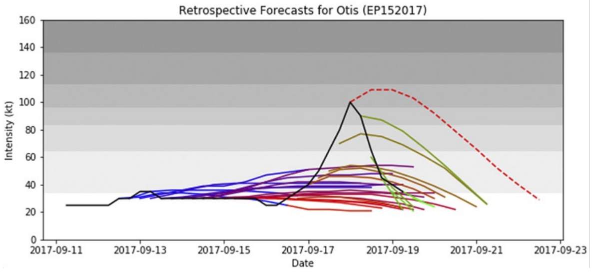
Global maps of normalized DKE on 300 hPa after 72-hr lead time. The thin black lines show the 300 hPa geopotential of the ensemble mean for reference (linespacing 1,500 m^2 s^{-2}).

Development and Evaluation of an Evolutionary Programming-Based Tropical Cyclone Intensity Model

Schaffer et al. 2020, AMS/MWR

"A statistical–dynamical tropical cyclone (TC) intensity model is developed from a large ensemble of algorithms through evolutionary programming (EP). Deterministic performance, as defined by MAE"...(in the Atlantic)..."is competitive with the operational Statistical Hurricane Intensity Prediction Scheme and Logistic Growth Equation Model at these times. In the eastern and central North Pacific"..."it is generally less skillful than OCD5 and all operational guidance."

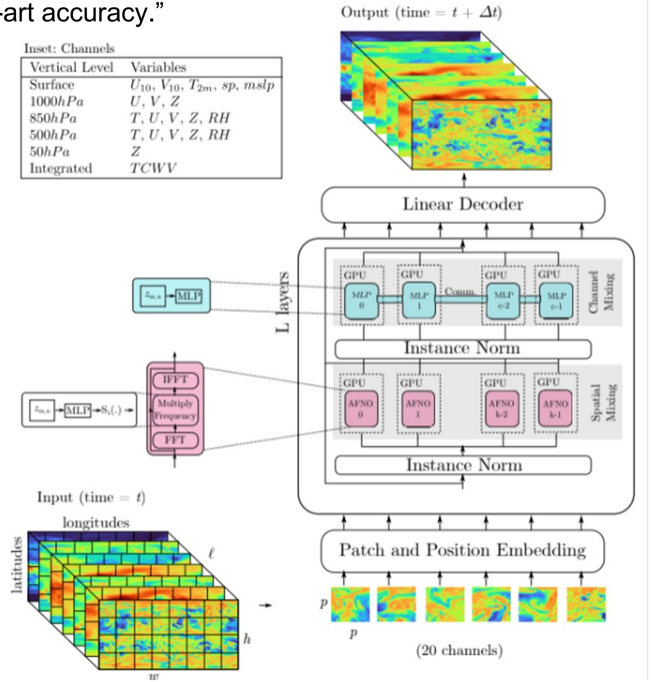




FourCastNet: Accelerating Global High-Resolution Weather Forecasting using Adaptive Fourier Neural Operators

Kurth et al. 2023, PASC '23, Davos, Switzerland
NVIDIA Corporation

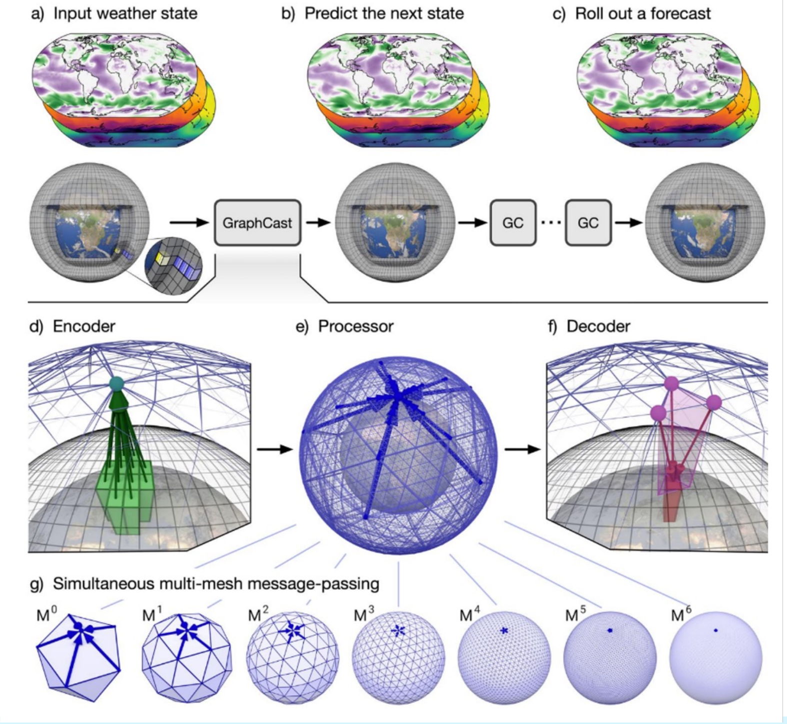
"We report that a data-driven deep learning Earth system emulator, FourCastNet, can predict global weather and generate medium-range forecasts five orders-of-magnitude faster than NWP while approaching state-of-the-art accuracy."



Learning Skillful Medium-Range Global Weather Forecasting

Lam et al. 2023, Science
Google DeepMind

"GraphCast significantly outperforms the most accurate operational deterministic systems on 90% of 1380 verification targets, and its forecasts support better severe event prediction, including tropical cyclones tracking, atmospheric rivers, and extreme temperatures."

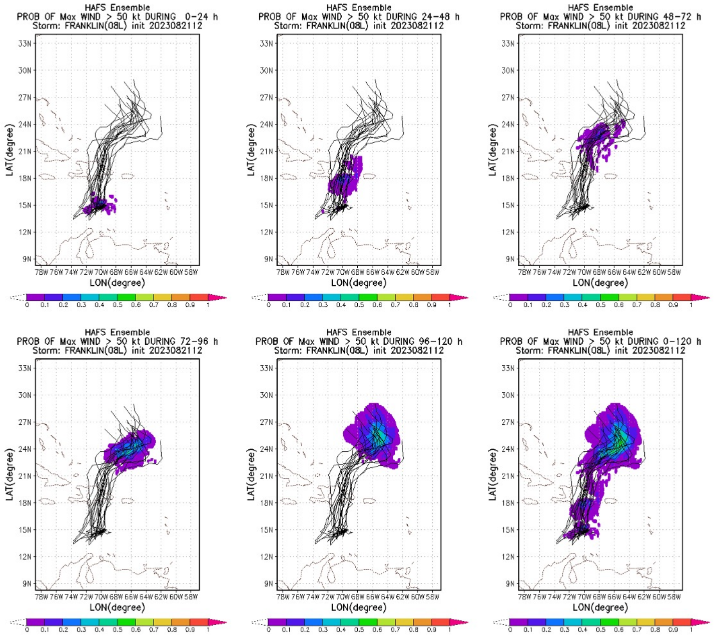


Modeling/supplemental program activities

Development of high resolution HAFS Ensemble based on operational HAFS

HFIP 2017 Strategic Goal: Incorporate risk communication research to create more effective watch and warning products.

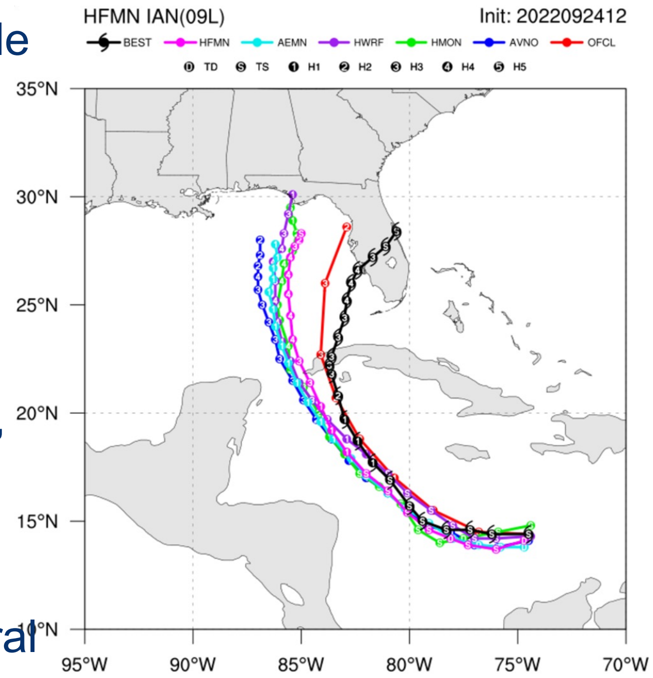
- 21-member ensemble with two-way ocean coupling,
- 120h forecast length 4x per day (00Z/06Z/12Z/18Z),
- Physics perturbations chosen for ability to project onto TC track, intensity, and/or structural diversity.
- Testing in 2023 near-real time experiments, alternate config (static 6km domain)

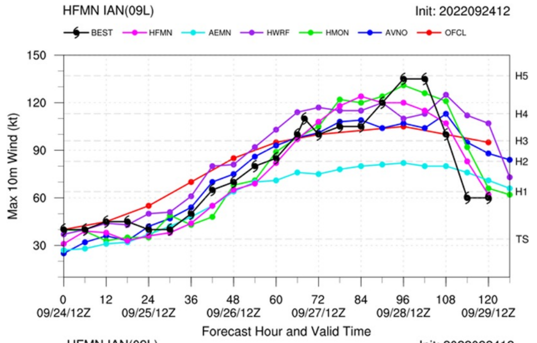


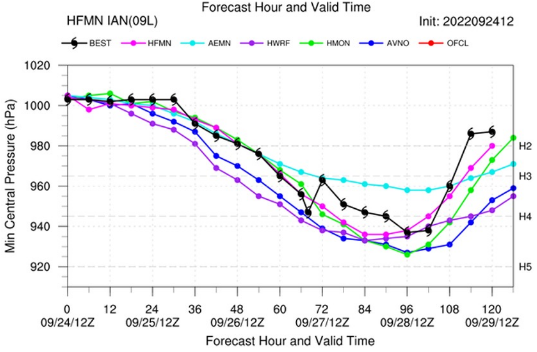
Tropical Storm Franklin HAFS Ensemble probability of wind speed > 50 kts 0-120 hrs https://www.emc.ncep.noaa.gov/HAFS/HAFSEPS/tcall.php

Hurricane Analysis and Forecast System (HAFS)
HAFS Ensemble Real-time on Cloud (HERC)

- High resolution ensemble developed from operational HAFS,
- 21-member ensemble with two-way ocean coupling,
- 120h forecast length 4x per day (00Z/06Z/12Z/18Z),
- Physics perturbations chosen for ability to project onto TC track, intensity, and/or structural diversity.

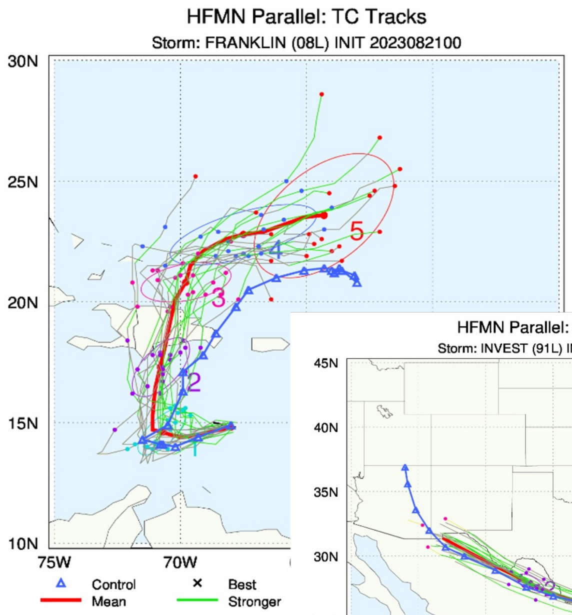


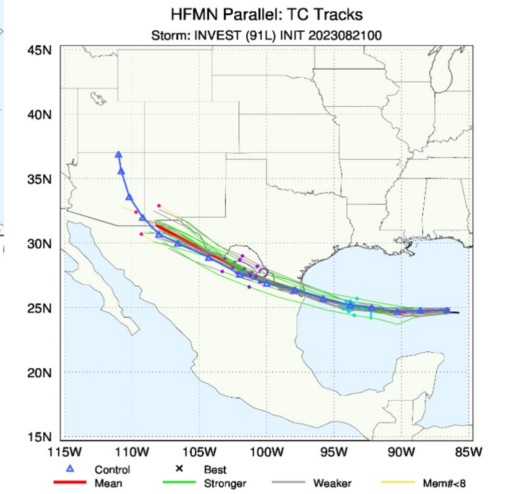




Hurricane Analysis and Forecast System (HAFS) Real-time DA experiment on Jet

- Plan: Self-cycled DA system on a static 6-km domain
- 20 ensemble members, 10 running 6h forecasts for EnKF only, 10 run 5-day forecasts.
- Possible pivot due to limited compute resources on jet real-time:
  - Run the DA system only (6 h forecasts for 20 members) within jet reservation, and run 5-day, 10-member ensemble forecasts option on other rdhpcs.
- If resources insufficient, will evaluate the 5-day ensembles after season, rather than real-time.





Themes from NCEP 10 year strategy

- ***JEDI as foundational infrastructure to enable (faster) innovation***
- Blur the lines “across/throughout the funnel” – these activities are iterative and not sequential.
  - **Leverage partnerships**, research / cooperative agreements
- Embrace change
  - **Reimagine** *how* we do assimilation (example: global hourly-updating system as pathway toward more continuous DA)
- New Technologies – proactive instead of reactive
  - CI/CD & Automation
  - Modern Programming
  - HPC, cloud, toward exascale
  - AI/ML
- Workforce development, recruitment, and retention

HPC Cloud Market Drivers

Rapid adoption of LLMs not fully reflected in market numbers…yet

- **AI & LLMs, including availability of GPUs**
- **Investments by cloud service providers (CSPs) to ease migrations to and integration with the cloud**
- **Users' maturing understanding of an expanding number of cloud-appropriate workloads**
- **Other recurring drivers with shifting priority order**
  - Cost-effectiveness relative to same job on-premises
  - Flexibility with surge workloads
  - Scale of available resources relative to on-premises infrastructure
  - Access to new technologies

Current JTTI Project Status

Chandra Kondragunta, JTTI Program Manager

Total number of R2O projects funded to date : 155

External = 119

- **Academic sector = 109**
- **Private sector = 10**

Internal (NOAA) = 36

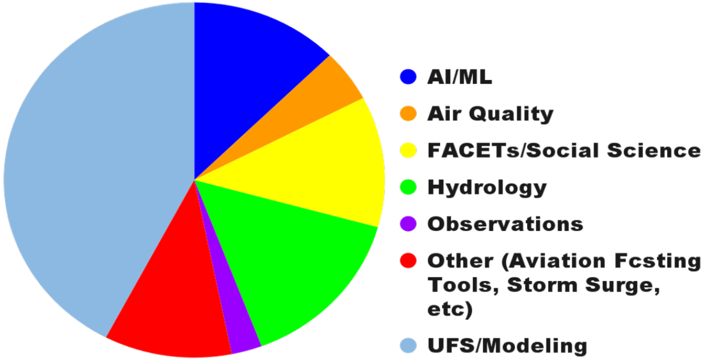
Total number of transitions = 20

(Includes SBES=2, AI/ML=2 an

Private Sector = 2)

JTTI funds and transitions interdisciplinary R2O projects from the American Weather Enterprise to the NWS operations

JTTI Funding by Topic Area (FY16-FY22)



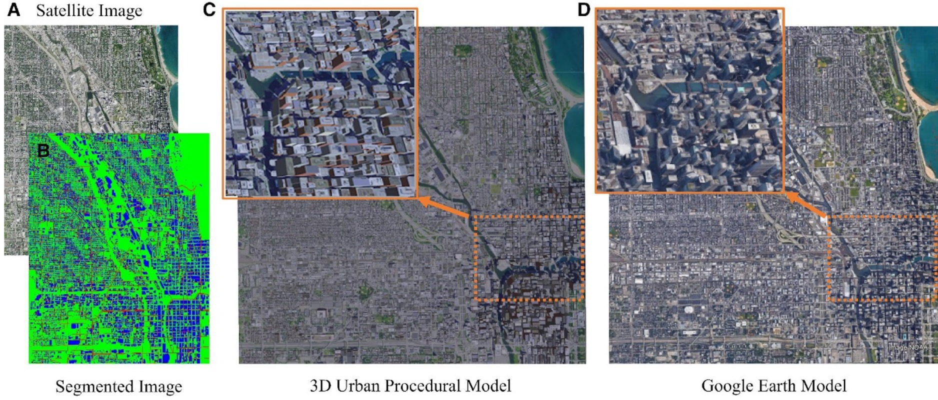
UT GLOBUS

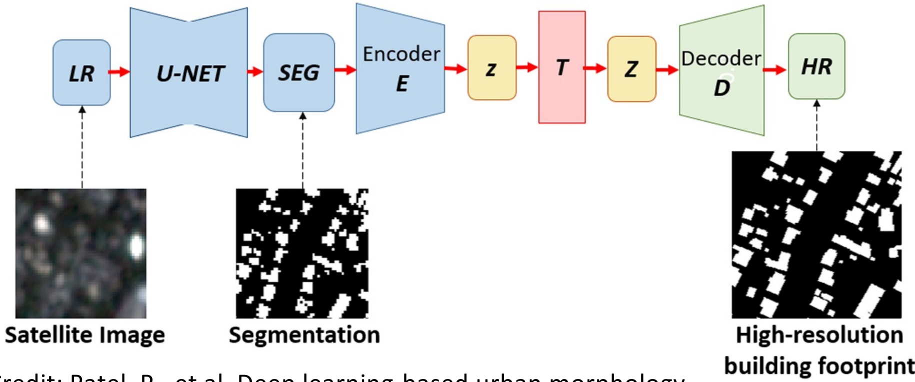
Dev Niyogi, UT Austin

An Open-Source Dataset that utilizes AI to develop land data including urban building heights information

- Uses open-source spaceborne satellite data.
- Employs machine learning approaches to predict building level information.
- Easily ingested in weather models.
- Used to calculate thermal comfort.

Synthetic dataset from AI and Gaming Environments for real world weather applications - Deep learning-based urban morphology





(Credit: Patel, P., et al. Deep learning-based urban morphology for city-scale environmental modeling)

A Review of Recent and Emerging Machine Learning Applications for Climate Variability and Weather Phenomena

Molina et al. 2023, Artificial Intelligence for the Earth Systems

Sources of predictability for modes of climate variability

- Representation of climate modes in ESMs is biased or uncertain.
- Results from XAI can be inconsistent (e.g., collinearity among inputs).
- Need to further the use and development of causal methods.
- Human biases can leak into XAI interpretation.

Feature detection

- Transferability of pretrained ML models across ESMs is unclear.
- Lack of explainability for ML-based feature detection.
- Ambiguities in feature definitions (as provided by domain experts).
- Certain phenomena lack standardized datasets (e.g., global monsoon).

Extreme weather & climate prediction & precursors

- Limited observational record (particularly for climate extremes and cascading or compounding extremes).
- Class imbalance of extremes.
- Characterization of extreme event precursors (e.g., genesis) needed.

Observation-model integration

- Inhomogeneous data coverage (spatially and temporally).
- More communication between observationalists and modelers needed.
- Need for more physics-informed ML and uncertainty quantification.
- Methods that further quantify observation-model agreement.

Downscaling & bias correction

- Open-source microscale benchmarking data are limited or lacking.
- Reasonable priors for uncertainty quantification are unclear.
- Both data driven and physics-based approaches are needed.
- Lack of downscaling in both space and time.

AI/ML in NWP

Studies have shown mixed results in terms of AI/ML model performance versus more traditional NWP models.

However, once developed and trained, AI/ML models consistently run considerably faster.

Two potential avenues for AI/ML in NWP, both of which need to continue to be explored

“In-line” applications (integrated into model cycle or timestep):

- Replace or speed up dynamical core
- Represent or replace model physics parameterizations (cumulus, PBL, microphysics, radiation)
- Direct integration into data assimilation cycle

“Offline” applications (run after the model run completes):

- Ensembles: reproducing the existing error and spread characteristics of an existing ensemble, upscale to 100s-1000s of members
- Post-processing and bias correction