AfriClimate AI
AI for a Climate-Resilient Africa

AI for Weather

CCAI Summer School

Shruti Nath | Rendani Mbuvha

Climate Change AI

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

1922 2022

UNIVERSITY OF OXFORD

About AfriClimate AI

Our Mission

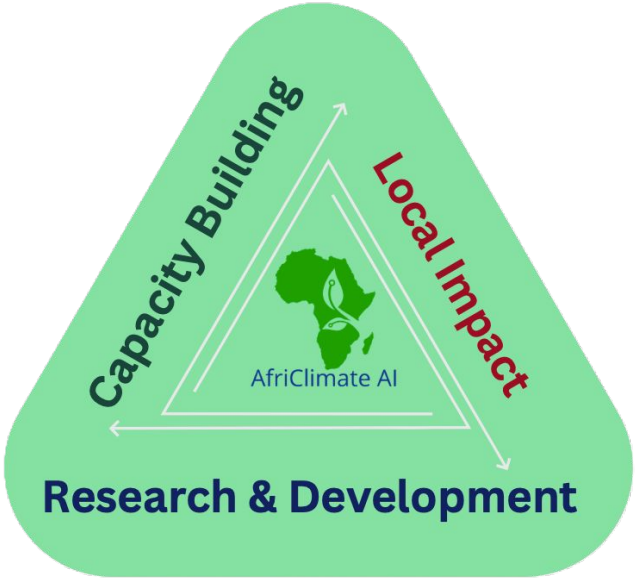
To advance climate and AI research in Africa by:

Building local skills
Training researchers and practitioners across the continent

Context-aware R&D
Research grounded in African climate and socioeconomic realities

AI solution deployment
Moving from prototype to operational climate services

Local datasets
Improving access to high-quality African observational data



Deep Learning Indaba Origins



Deep Learning Indaba, Kigali 2025

Lecture Structure

Part A : Short Range Forecasting (“Weather”)

Part B: Long Range Forecasting (“Seasonal Forecasts”)

Weather as an initial value problem: turn of the 20th Century





Meteorologische Zeitschrift, Januar 1904.

Das Problem der Wettervorhersage, betrachtet vom Standpunkte der Mechanik und der Physik.

Von V. Bjerknes, Universität Stockholm.

Wenn es sich so verhält, wie jeder naturwissenschaftlich denkende Mann glaubt, daß sich die späteren atmosphärischen Zustände gesetzmäßig aus den vorhergehenden entwickeln, so erkennt man, daß die notwendigen und hinreichenden Bedingungen für eine rationelle Lösung des Prognoseproblems der Meteorologie die folgenden sind:

1. Man muß mit hinreichender Genauigkeit den Zustand der Atmosphäre zu einer gewissen Zeit kennen.

2. Man muß mit hinreichender Genauigkeit die Gesetze kennen, nach den sich der eine atmosphärische Zustand aus dem anderen entwickelt.

I.

Die Kenntnis des Zustandes der Atmosphäre zu passenden, übereingekommenen Zeiten zu schaffen, ist die Aufgabe der beobachtenden Meteorologie. In genügendem Umfange für eine rationelle Wetterprognose ist diese

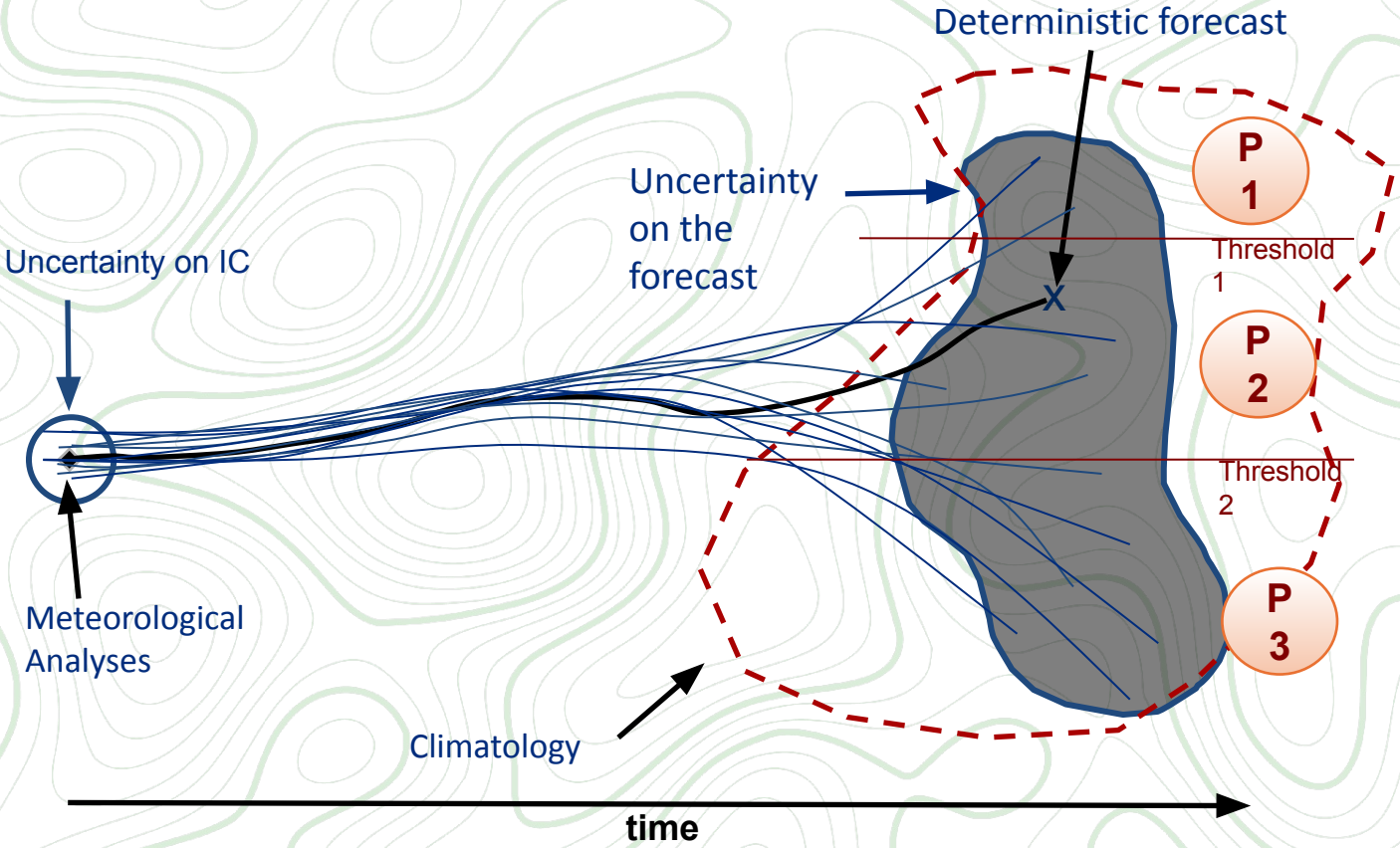
$$\frac{\delta \mathrm{v}}{\delta t} + (\mathrm{v} \cdot \nabla) \cdot \mathrm{v} = \frac{1}{\rho} \cdot \nabla p + \nu \cdot \Delta \mathrm{v} + \boldsymbol{f}$$

$$\Delta \mathrm{U} = \mathrm{Q} - \mathrm{W}$$

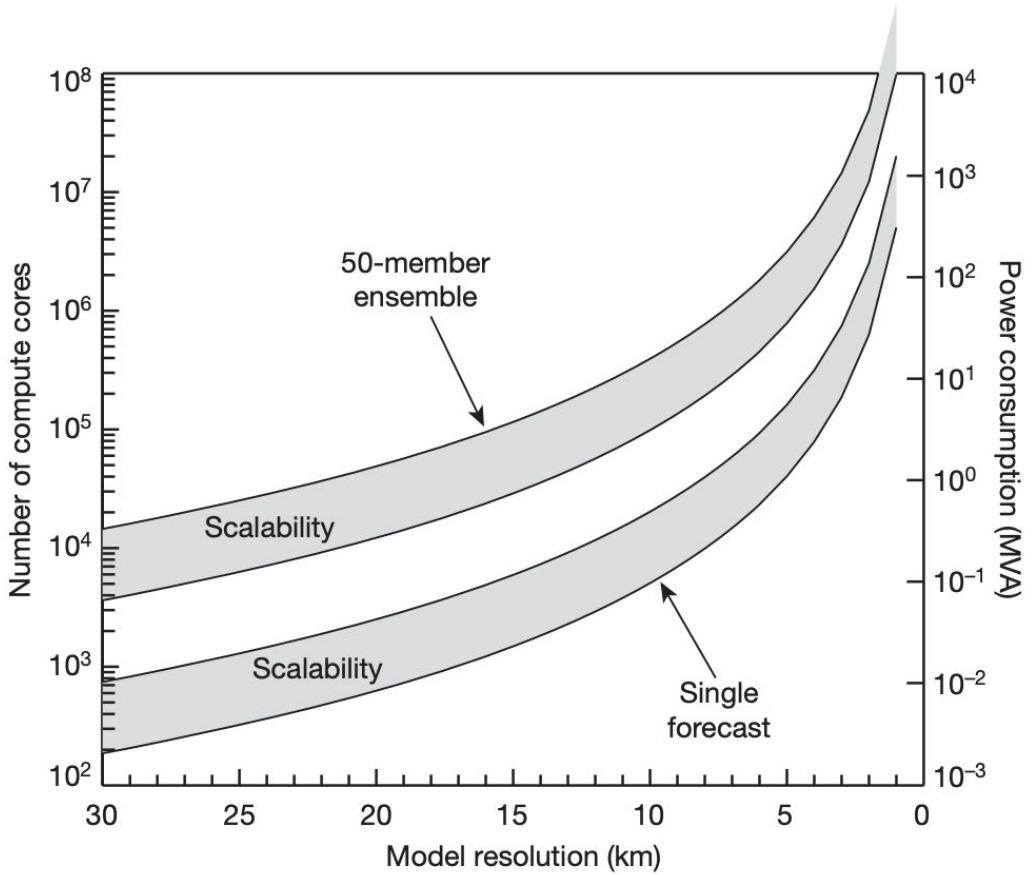
$$PV = nRT$$

What do current forecasts look like?

Probabilistic forecasts



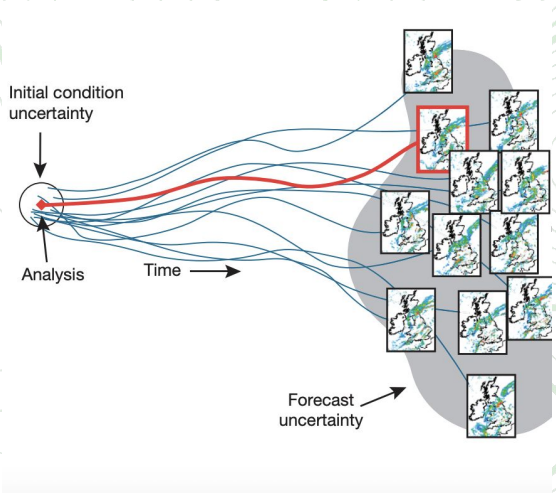
Supercomputing capacity since 1970



REVIEW

The quiet revolution of numerical weather prediction

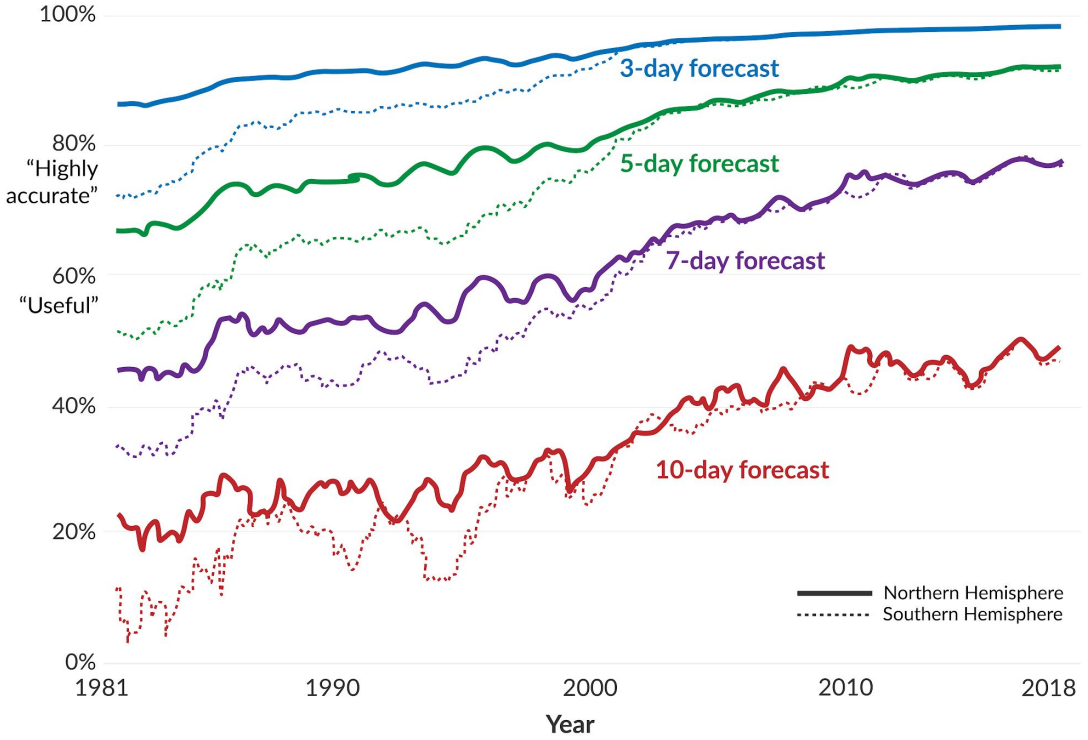
Peter Bauer[1], Alan Thorpe[1] & Gilbert Brunet[2]



The accuracy of weather forecasts has improved

Accuracy is measured as the difference between the forecast and subsequent weather. This is based on the '500 hPa geopotential height' which is a common meteorological metric used to measure air pressure.

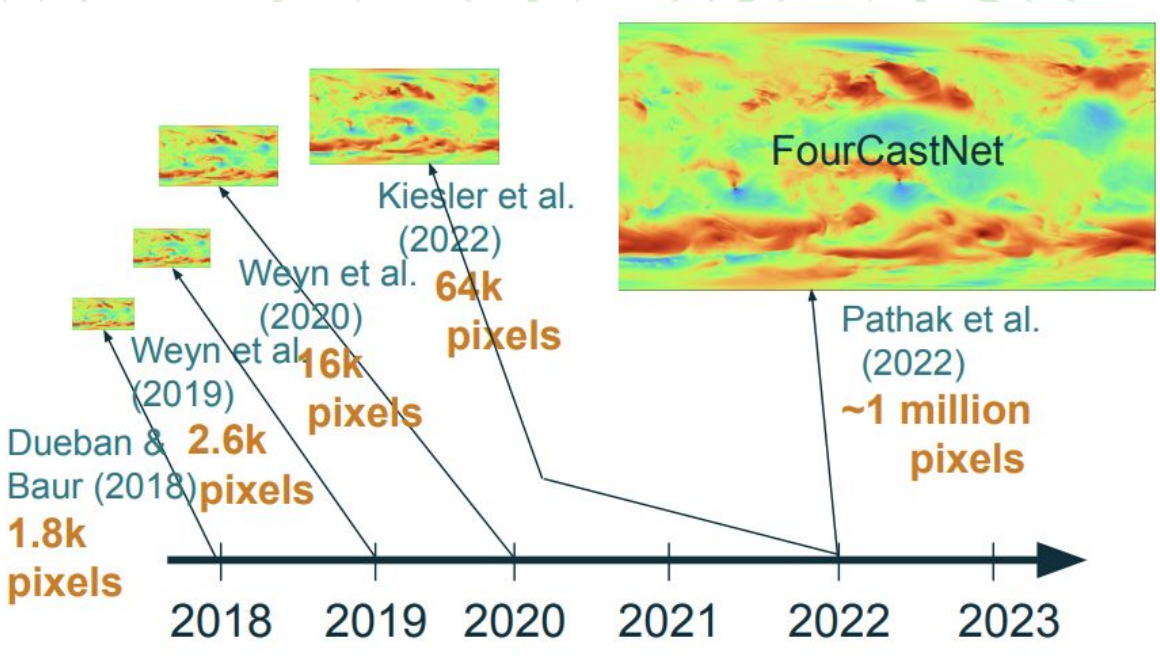
Our World in Data



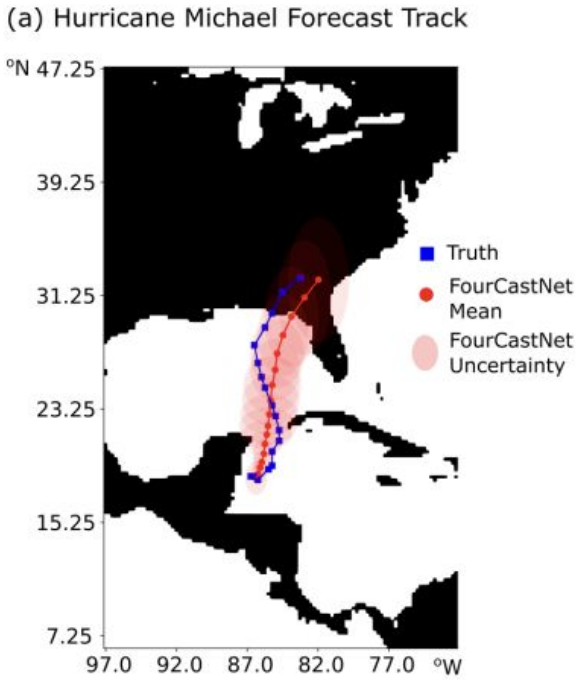
Source: European Centre for Medium-Range Weather Forecasts (ECMWF).

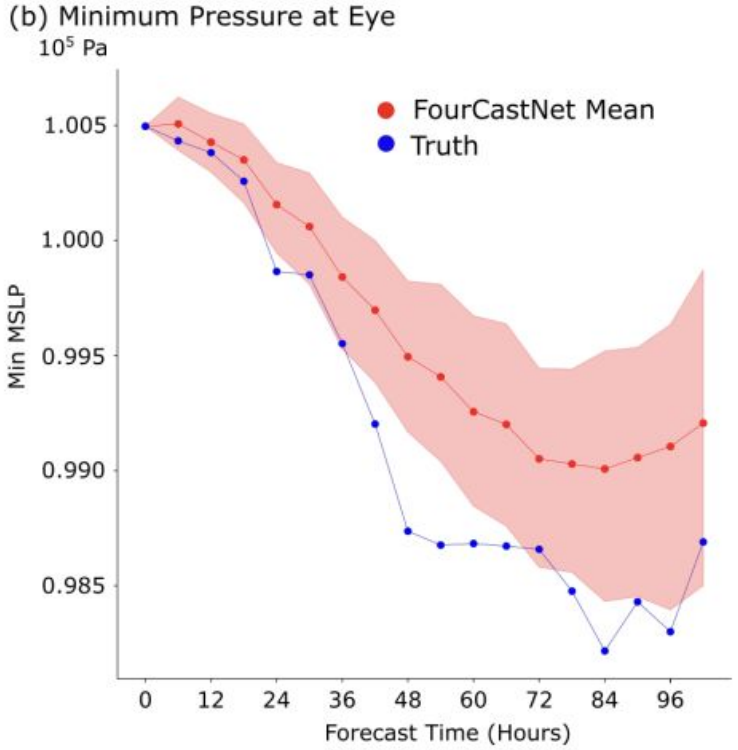
Licensed under CC-BY by the author Hannah Ritchie.

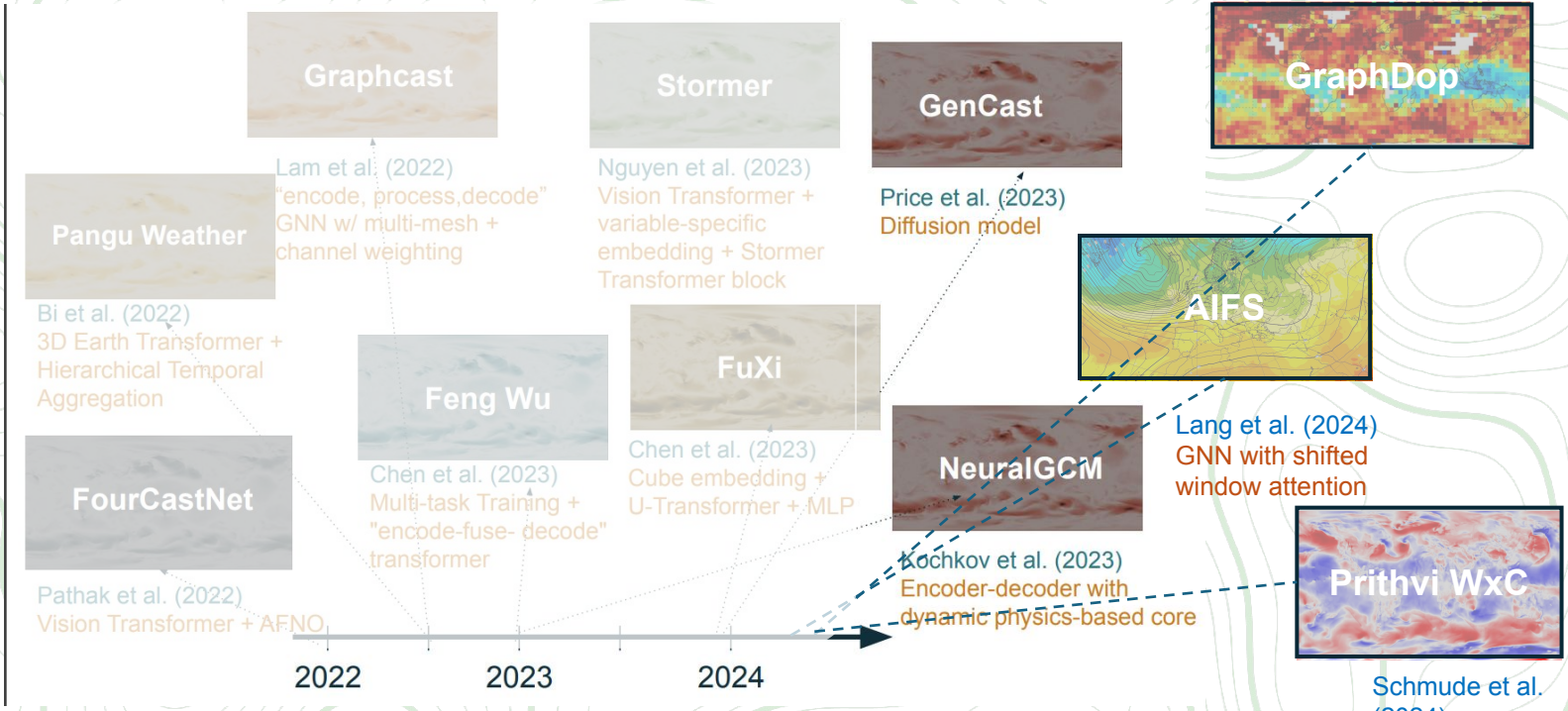
Bringing the noise: AI models



The first PoC:Hurricane Michael forecasts

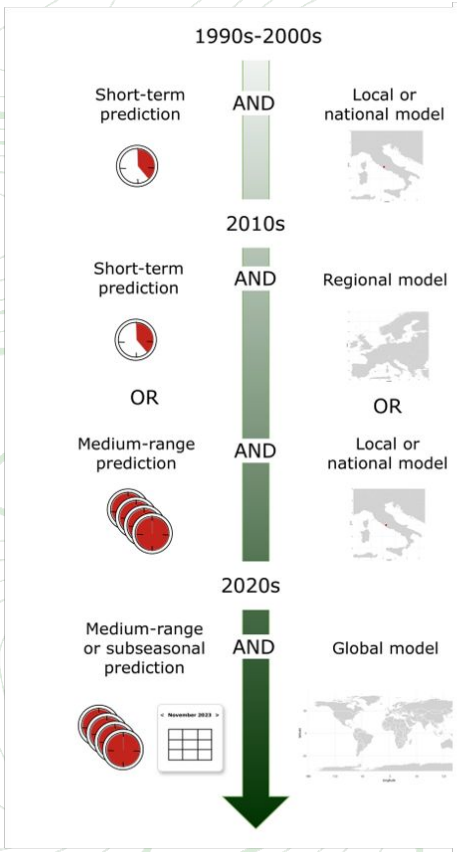


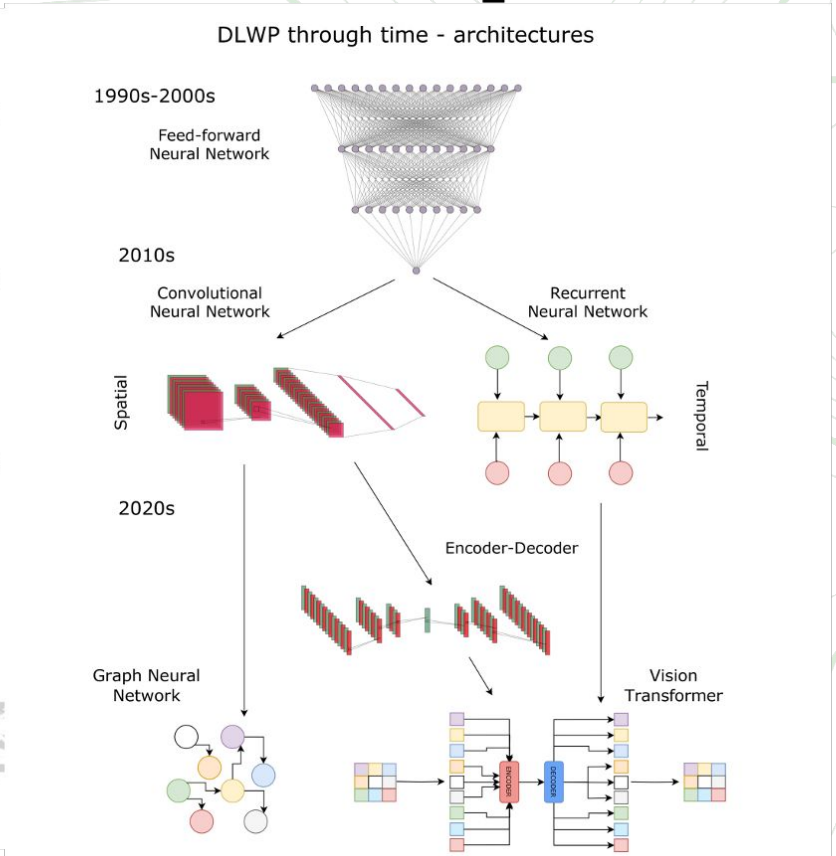




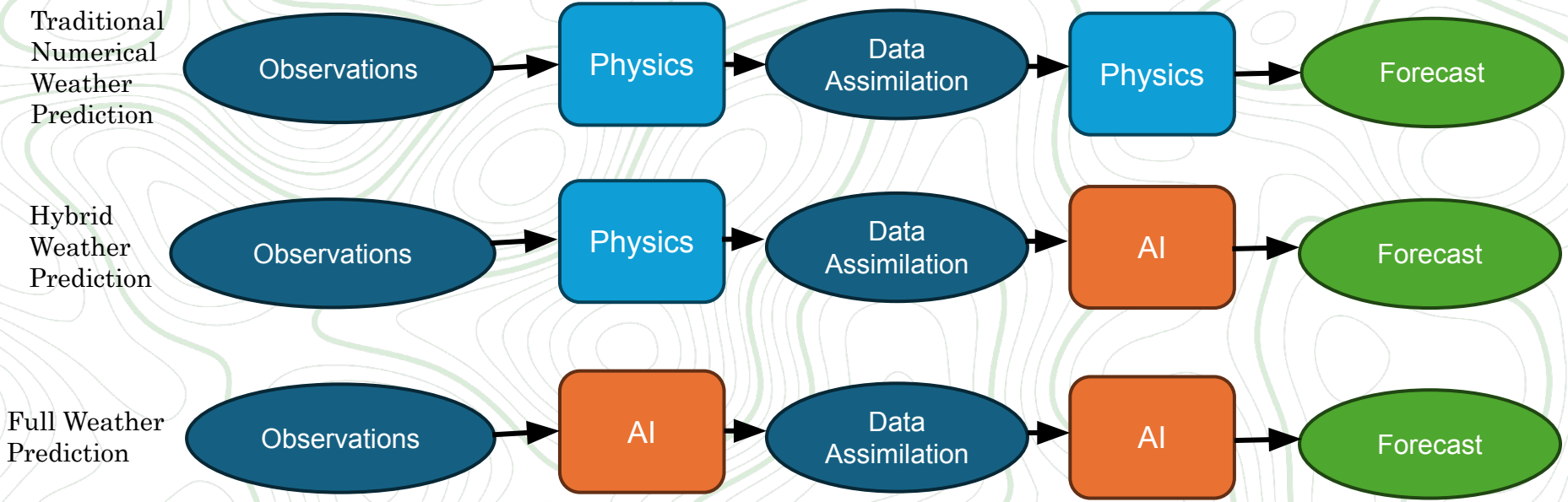
AI model development seems to reinforce itself spurring more and more developments!

General timescales of developments

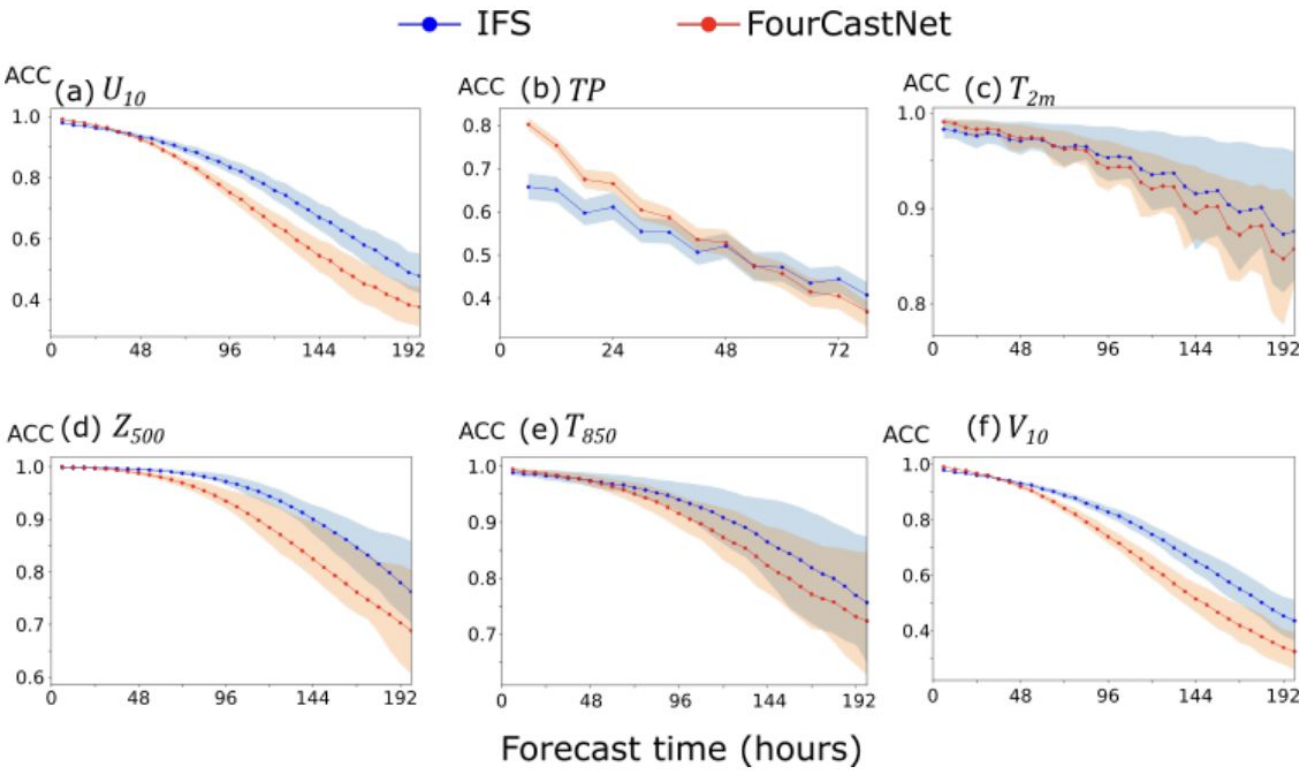




The future of AI?

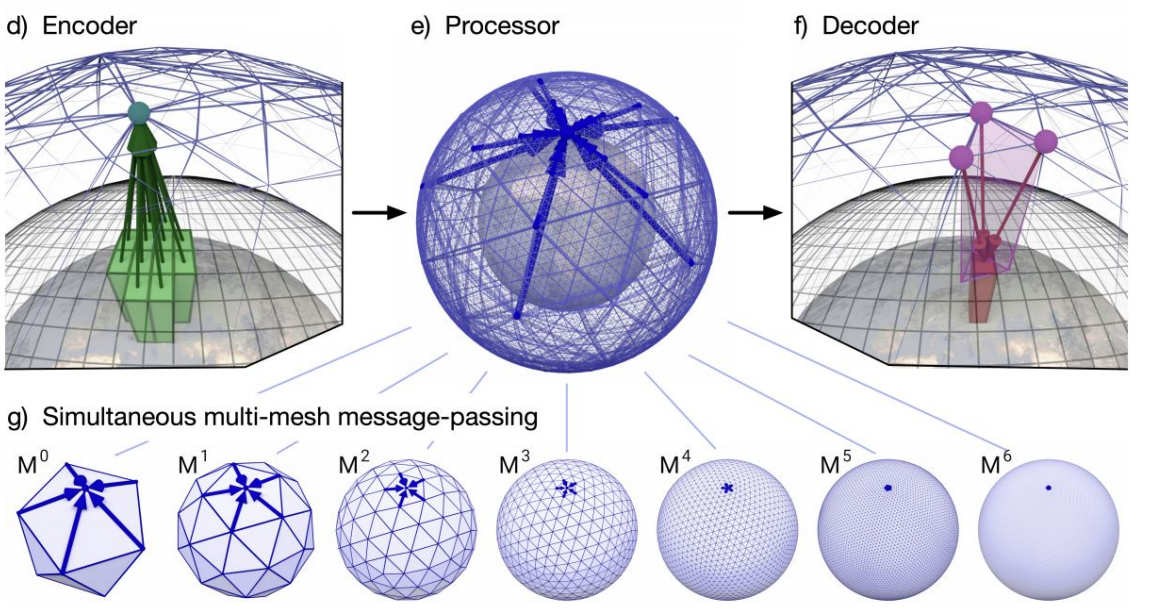


Accuracy was comparable but degraded with time



GraphCast

- Encoder learns node attributed translated from grid to multi-mesh
- Multi-mesh made from iterative refinement: regular icosahedron (12 nodes, 20 faces, 30 edges) iteratively refined six times
- Processor then performs message passing on 16 unshared GNN layers
- Similar roll-out as FourCastNet

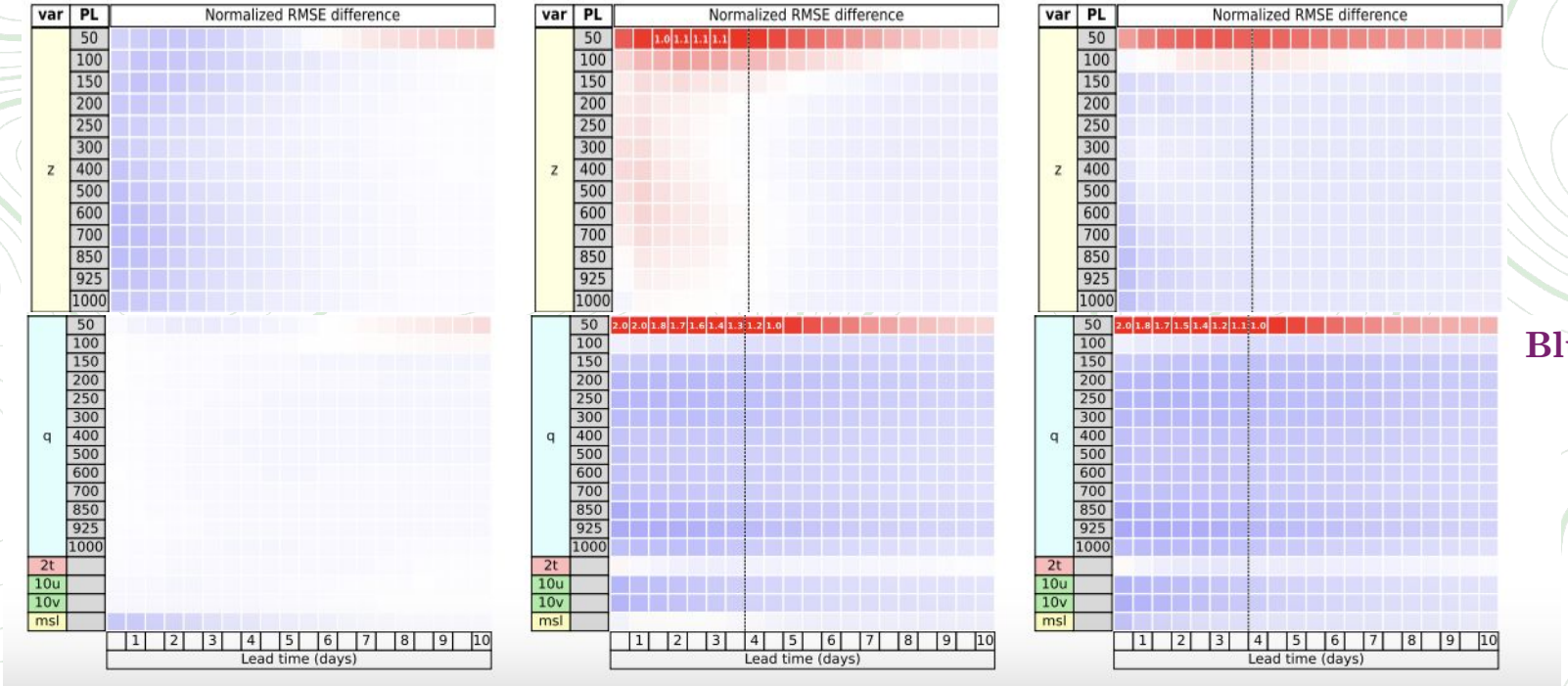


Multi-mesh ablation studies

GraphCast – without multi-mesh

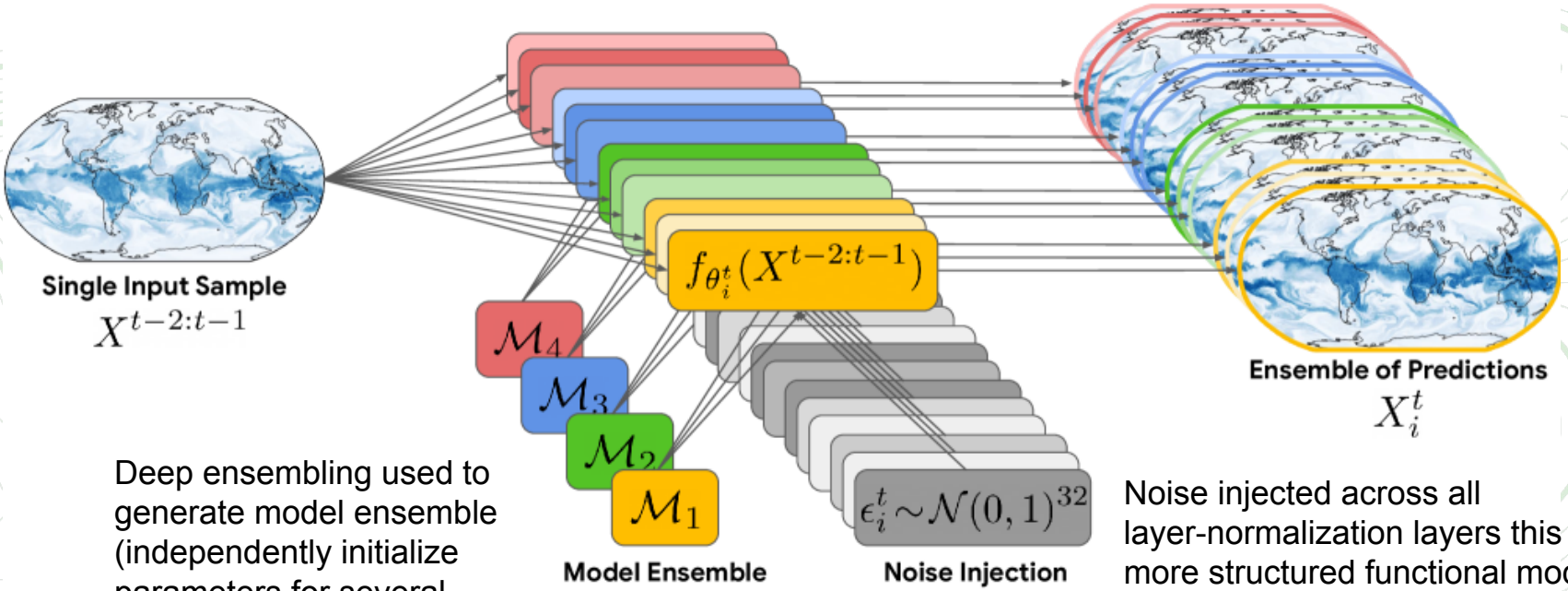
Without multi-mesh – HRES

GraphCast – HRES



Blue is better

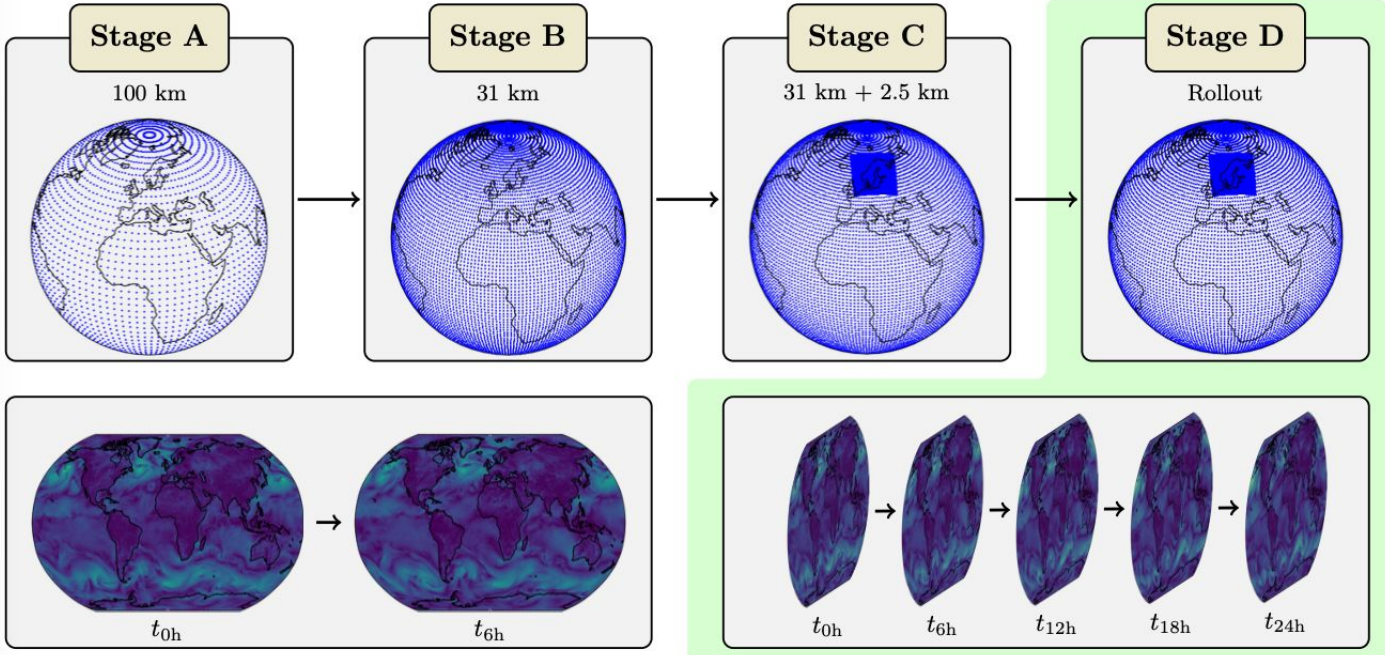
Through deep ensembling and parameter-shared conditional normalisation



Deep ensembling used to generate model ensemble (independently initialize parameters for several models for training)

Noise injected across all layer-normalization layers this allows more structured functional modulation across layers maintaining spatial coherence

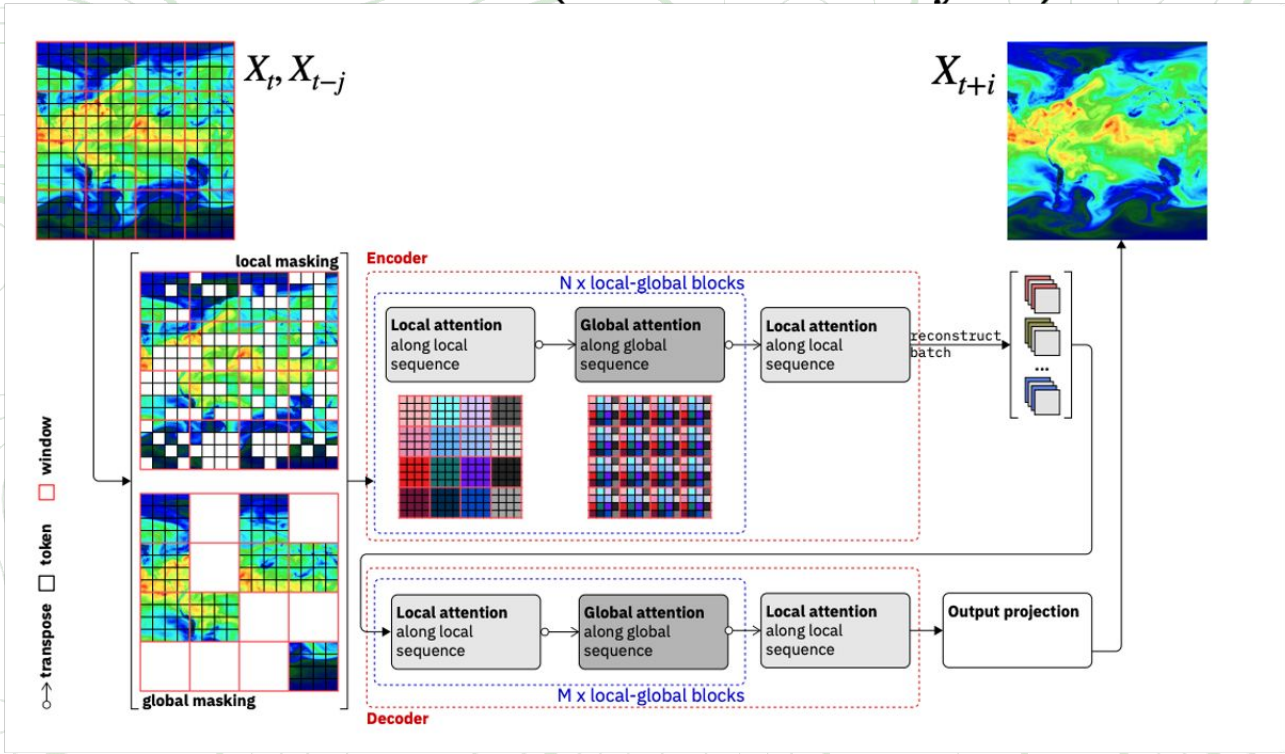
Graph transformers are easy to adapt for variable resolution models



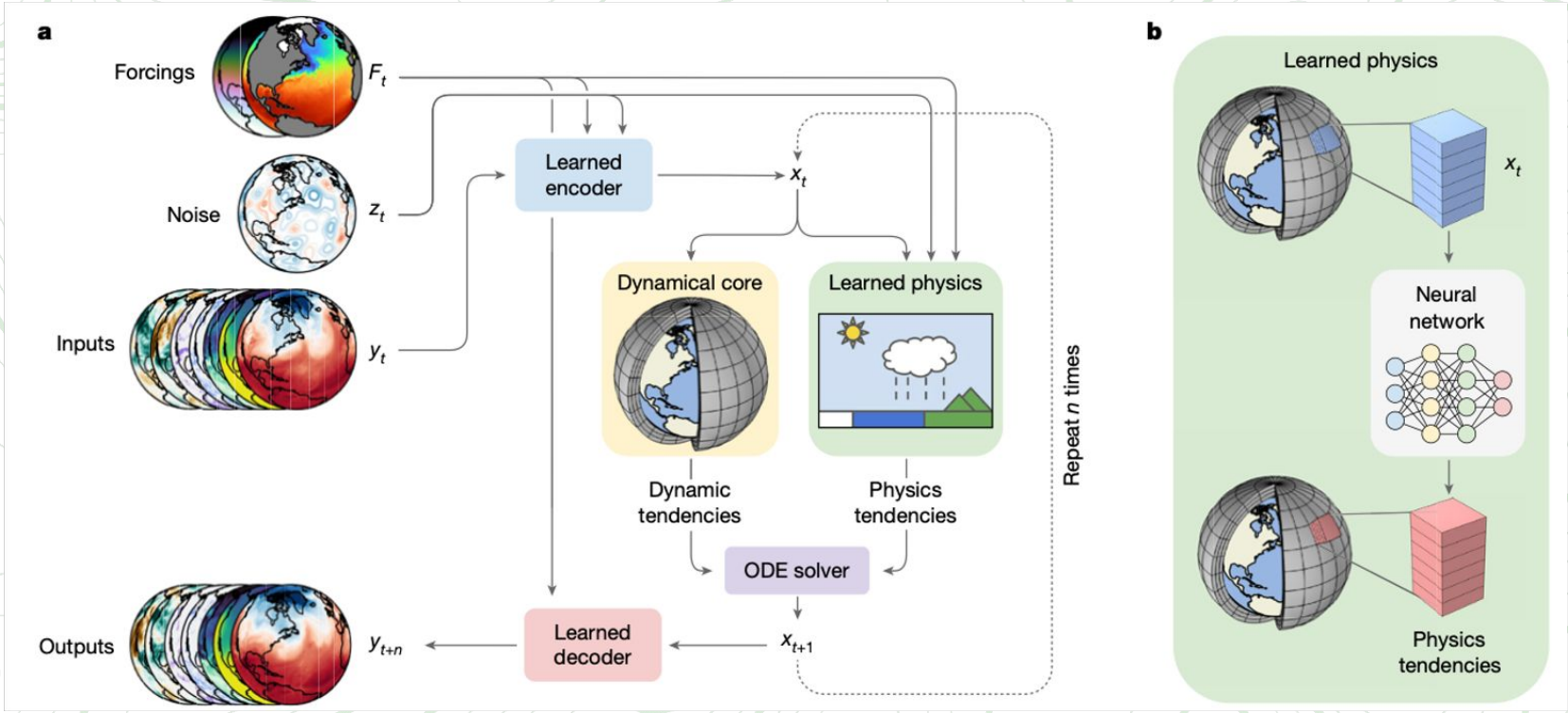
Prithvi WxC

https://arxiv.org/pdf/2409.13598

- Trained on MERRA-2 data (NASA reanalysis)



Hybrid approaches: fully learning all parameterizations



Are these models operationally useful?

Drawing fair comparisons

- For fair comparisons we note that there are different types of forecasts:
  - Deterministic (i.e., one ensemble member)
  - Probabilistic
- AI models are trained on ERA5 data, e.g., for rainfall this is not always realistic and we need to correct towards satellite estimates instead
- Physics-based Numerical Weather Prediction output should also be corrected towards satellite estimates

How do we define forecast utility?

Any scoring rule $S(x, o)$ used to define the skill of a model's predicted outcomes $x \sim X$ against actual observations $o \sim Y$ is proper if

$$E[S(y, o)] \leq E[S(x, o)] \text{ for } y \sim Y$$

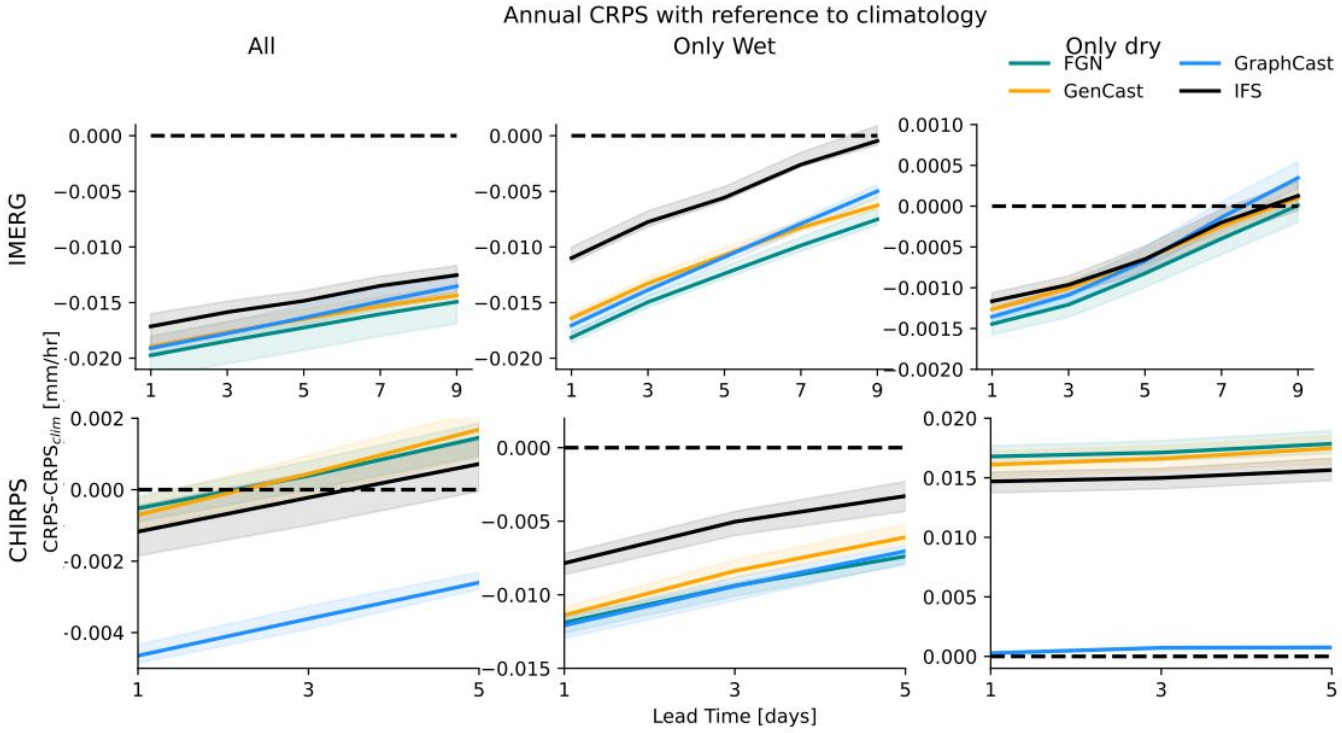
MAE, CRPS, Log-likelihood etc. are all proper scoring rules and can be broken down into

$$S_x = (S_x - S_{cal}) - (S_{ref} - S_{cal}) + S_{ref}$$

Miscalibration | Discrimination | Uncertainty

[1]Gneiting et al. (2025). https://doi.org/10.48550/arXiv.2506.03744

1- and 9- day lead time

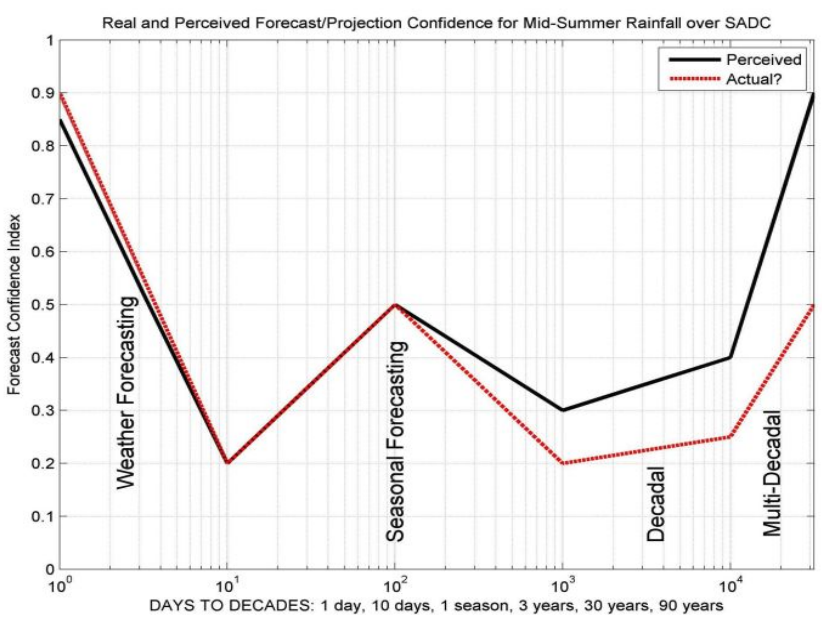


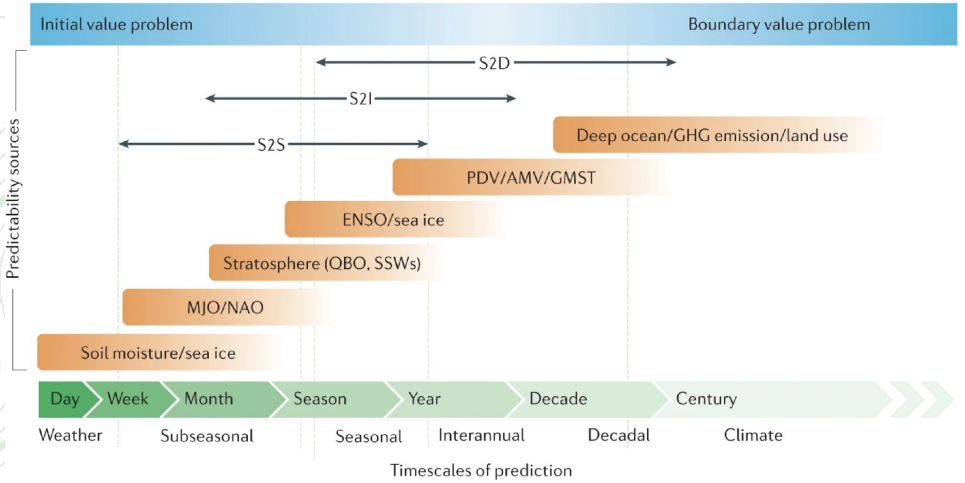
2024-2025 used as test

Forecast 4 Africa
by AfriClimate AI

What is Seasonal Forecasting?

What does the next three 90 days look like?

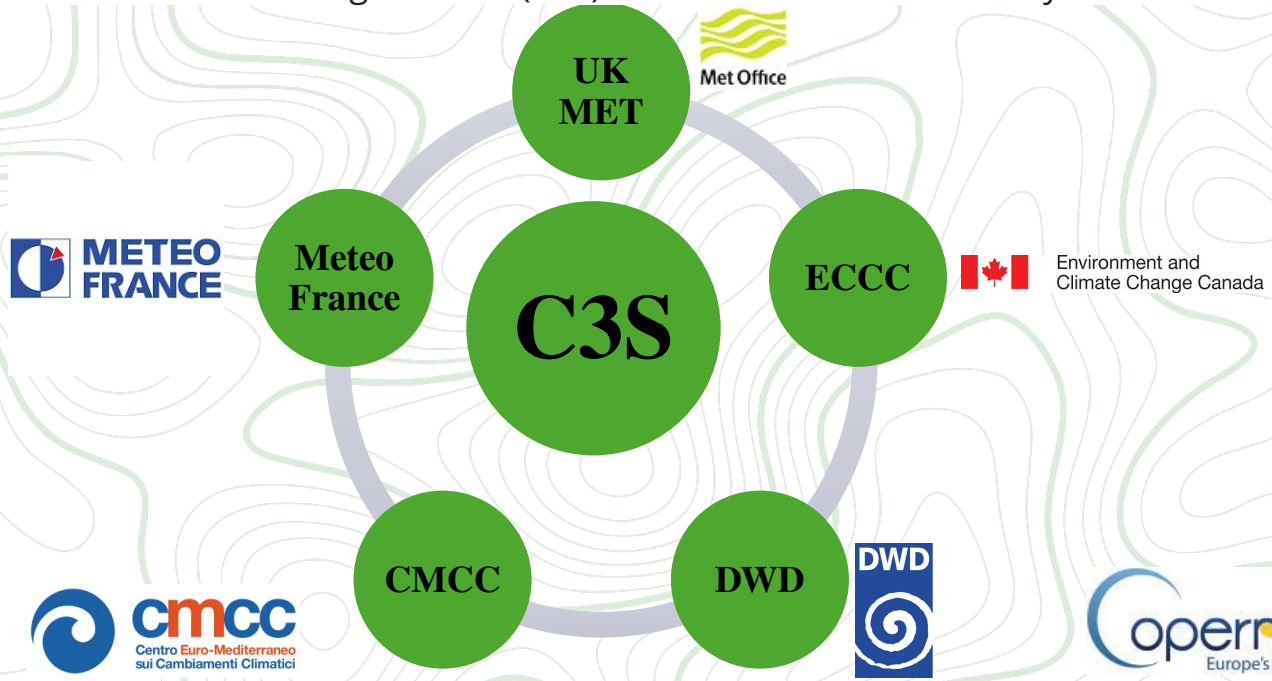




Source: Norwegian Meteorological Institute

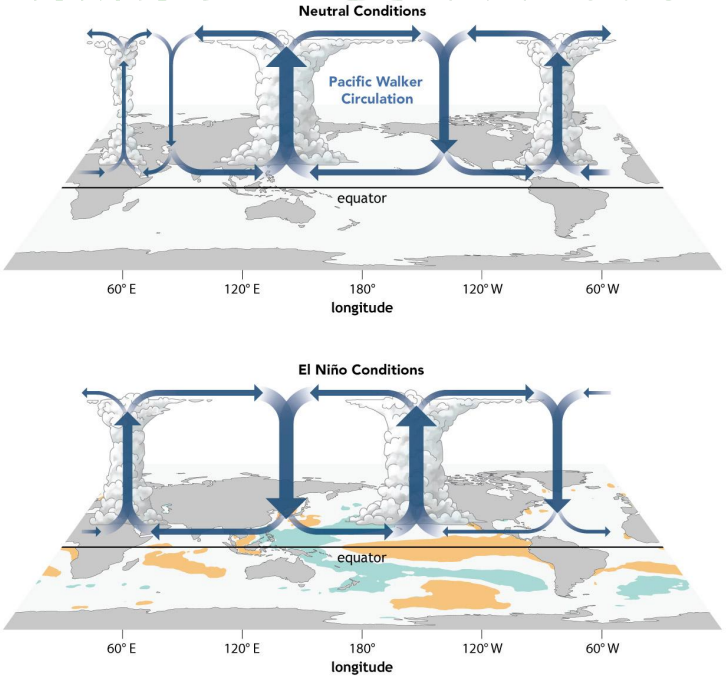
Seasonal Forecasting Models

Copernicus Climate Change Service (C3S) seasonal forecast service by the ECMWF

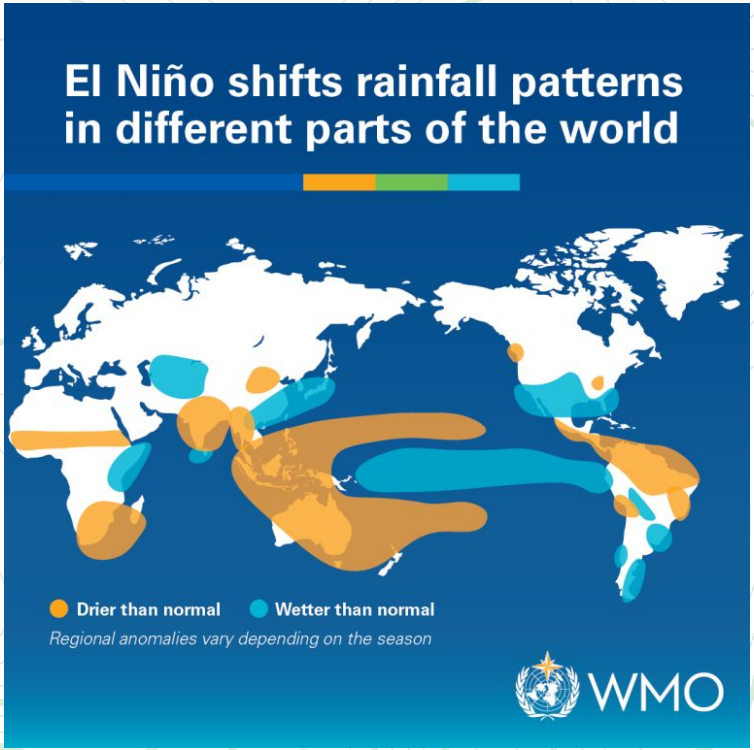




ENSO and Drivers of Seasonal Forecast Skill



The arrival of El Niño. Illustration by NOAA/Climate.gov



Source: WMO and Lenssen, N. J. L., L. Goddard, and S. Mason, 2020: Seasonal Forecast Skill of ENSO Teleconnection Maps. Wea. Forecasting, 35, 2387–2406

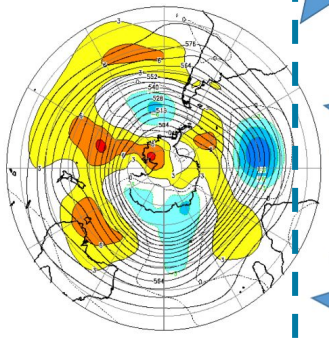
Importance of Seasonal Forecasting

The prediction scheme

1. Phenomena to be predicted should contain a climate signal (e.g. ENSO) in the data; 2. Observed and model time series must be over sufficiently long enough periods so that robust statistical relationship can be developed; 3. and some form of quality control of the observed data had taken place.



ECHAM4.5-MOM3-DC2 ocean-atmosphere GCM



Geopotential heights/ rainfall, etc.

Some transfer function



Riverflow



Malaria



Dry-land crop yield

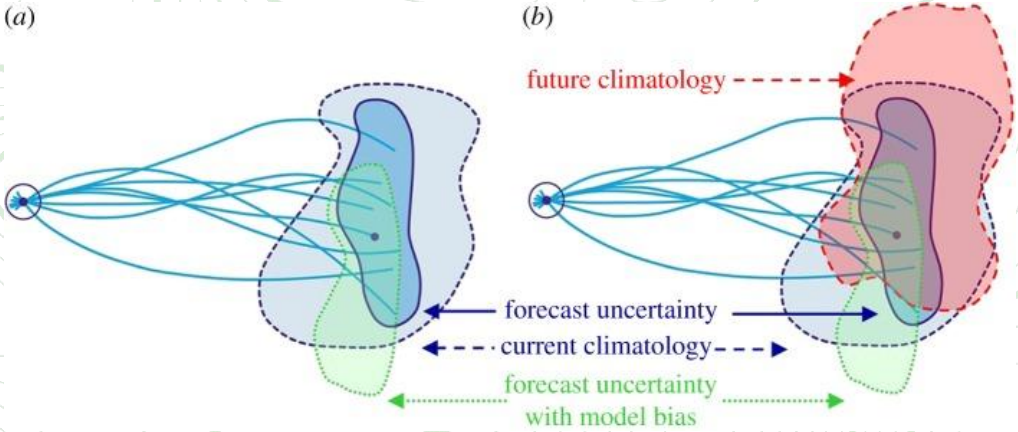
6 CLEAN WATER AND SANITATION

13 CLIMATE ACTION

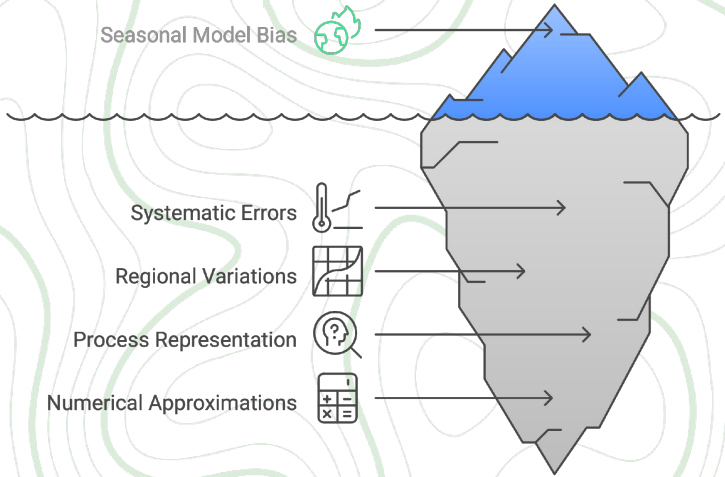
3 GOOD HEALTH AND WELL-BEING

2 ZERO HUNGER

Seasonal Model Bias and Its Sources

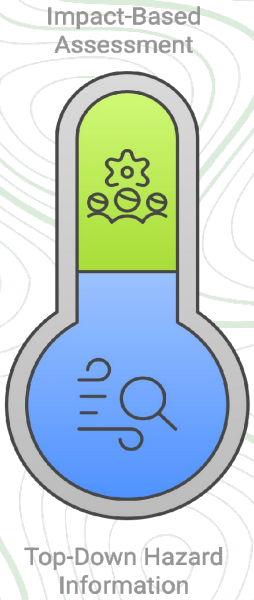


Seasonal Model Bias and Its Sources



Slingo J, Palmer T. Uncertainty in weather and climate prediction. Philos Trans A Math Phys Eng Sci. 2011 Dec 13;369(1956):4751-67. doi: 10.1098/rsta.2011.0161. PMID: 22042896; PMCID: PMC3270390.

Impact-Based Model vs Traditional Evaluation



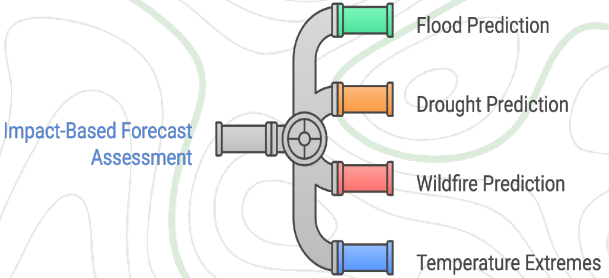
Impact-Based Assessment

Translates forecasts into meaningful information for decision-making

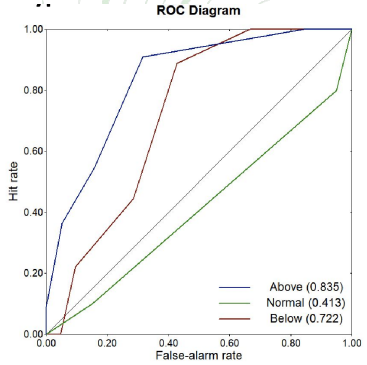
Traditional Assessment

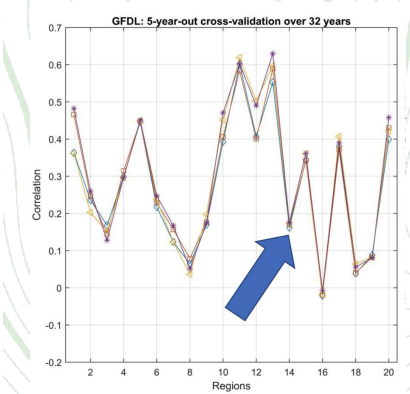
Captures weather patterns compared to reference data

Metrics Based on Indices



Correlations, RMSE and Probabilistic Measures





How good are Seasonal Models out of the “box” on an Impact Basis?

AGU ADVANCING EARTH AND SPACE SCIENCES

Earth's Future

RESEARCH ARTICLE
10.1029/2024EF004936

Key Points:

- C3S models demonstrate notable skill in forecasting seasonal variability of precipitation, particularly in the tropical and sub-tropical regions
- UK-Met and Météo-France consistently outperform other models, demonstrating superior accuracy and reliability in various regions and seasons
- Our Impact-Based Framework offers valuable insights for targeted risk assessments, particularly in wildfire- and flood-prone regions using seasonal forecast models

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

R. Mbuvha,
r.mbuvha@qmul.ac.uk

Citation:

Nikraftar, Z., Mbuvha, R., Sadegh, M., & Landman, W. A. (2024). Impact-based skill evaluation of seasonal precipitation forecasts. *Earth's Future*, *12*, e2024EF004936. https://doi.org/10.1029/2024EF004936

Received 3 JUN 2024
Accepted 16 OCT 2024

Author Contributions:

Conceptualization: Zahir Nikraftar, Rendani Mbuvha, Mojtaba Sadegh, Willem A. Landman
Data curation: Zahir Nikraftar
Formal analysis: Zahir Nikraftar
Investigation: Zahir Nikraftar

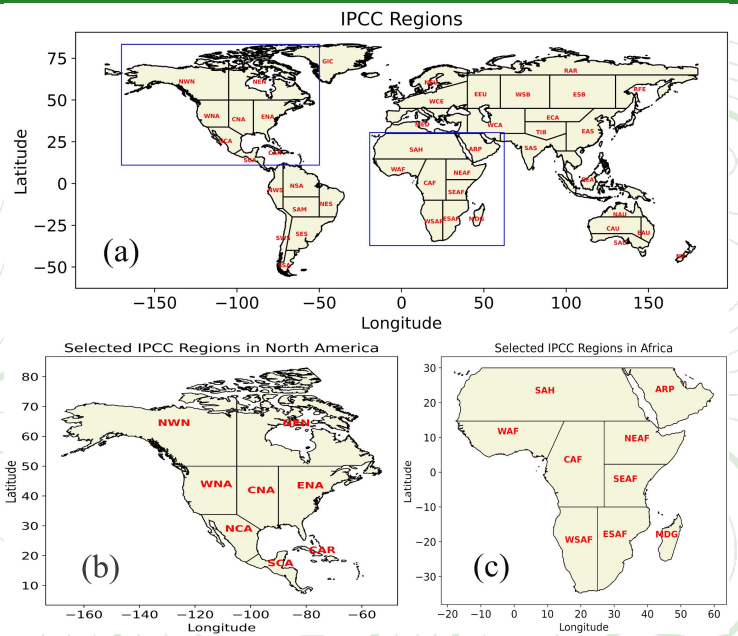
Impact-Based Skill Evaluation of Seasonal Precipitation Forecasts

Zahir Nikraftar[1], Rendani Mbuvha[1,2], Mojtaba Sadegh[3,4], and Willem A. Landman[5]

[1]Machine Intelligence and Decision Systems (MInDS) Research Group, School of Electronic Engineering and Computer Science, Queen Mary University of London (QMUL), London, UK, [2]School of Statistics and Actuarial Science, University of Witwatersrand, Johannesburg, South Africa, [3]Department of Civil Engineering, Boise State University, Boise, ID, USA, [4]United Nations University Institute for Water, Environment and Health, Hamilton, ON, Canada, [5]Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa

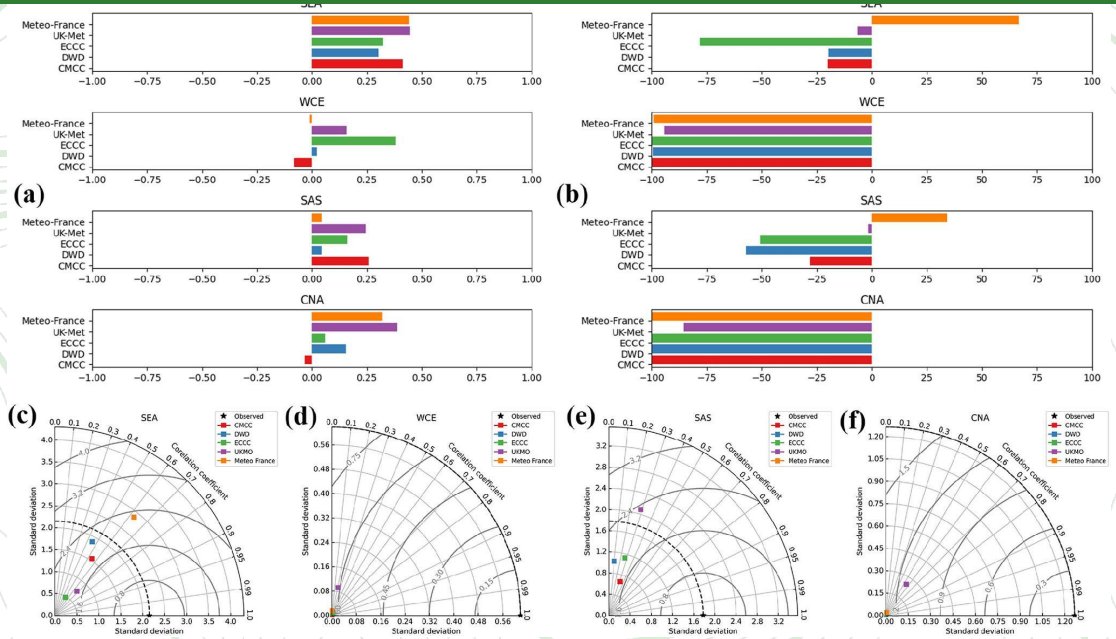
Abstract We introduce an impact-based framework to evaluate seasonal forecast model skill in capturing extreme weather and climate events over regions prone to natural disasters such as floods and wildfires. Forecasting hydroclimatic extremes holds significant importance in an era of increasing hazards such as wildfires, floods, and droughts. We evaluate the performance of five Copernicus Climate Change Service (C3S) seasonal forecast models (CMCC, DWD, ECCC, UK-Met, and Météo-France) in predicting extreme precipitation events from 1993 to 2016 using 14 indices reflecting timing and intensity (using absolute and locally defined thresholds) of precipitation at a seasonal timescale. Performance metrics, including Percent Bias, Kendall Tau Rank Correlation Score, and models' discrimination capacity, are used for skill evaluation. Our findings indicate that the performance of models varies markedly across regions and seasons. While models generally show good skill in the tropical regions, their skill in extra-tropical regions is markedly lower. Elevated precipitation thresholds (i.e., higher intensity indices) correlate with heightened model biases, indicating deficiencies in modeling severe precipitation events. Our analysis using an impact-based framework highlights the superior predictive capabilities of the UK-Met and Météo-France models in capturing the underlying processes that drive precipitation events, or lack thereof, across many regions and seasons. Other models exhibit strong performance in specific regions and/or seasons, but not globally. These results advance our understanding of an impact-based framework in capturing a broad spectrum of extreme weather and climatic events, and inform strategic amalgamation of diverse models across different regions and seasons, thereby offering valuable insights for disaster management and risk analysis.

Plain Language Summary We introduce a novel model assessment framework by investigating the impact of extreme events in areas prone to disasters such as floods and wildfires. We investigate how well five seasonal forecast models predict extreme precipitation, and related events such as floods and droughts. We assess the performance of models through 14 indices that assess the timing and intensity of precipitation during different seasons and in various regions globally. Our results show that some models outperform others in certain regions and seasons. We find that two models, UK-Met and Météo-France, are particularly skillful at predicting extreme events in various regions and seasons. This information is important for improved management and understanding of the risks of natural disasters.



Z., Mbuvha, R., Sadegh, M., & Landman, W. A. (2024). Impact-based skill evaluation of seasonal precipitation forecasts. *Earth's Future*, 12, e2024EF004936. https://doi.org/10.1029/2024EF004936

Impact Based Evaluation – HPD in Flood Prone Regions



Model performance with respect to the number of heavy precipitation days exceeding 10 mm and the 75th percentile of the reference data (nwd10q75): a) Kendall's Tau coefficient, b) Percent bias, and Taylor diagram for c) SEA, d) WCE, e) SAS, and f) CNA regions, respectively.

Bias Correction of Seasonal Forecasting Models

Quantile Mapping

$$F_{\text{obs}}(x_{\text{corrected}}) = F_{\text{model}}(x_{\text{model}})$$

- F_{obs}: Observed cumulative distribution.
- F_{model}: Model cumulative distribution.

Linear Scaling

$$x_{\text{corrected}} = a \cdot x_{\text{model}} + b$$

- $a = \frac{\sigma_{\text{obs}}}{\sigma_{\text{model}}}$: Scaling factor for variability (standard deviation).
- $b = \mu_{\text{obs}} - a \cdot \mu_{\text{model}}$: Bias adjustment for the mean.



AI In Seasonal Forecasting

JGR | Machine Learning and Computation

RESEARCH ARTICLE
10.1029/2025JH000957

Season-Net: A Deep Learning Framework for Bias Correction of Seasonal Forecasting Models

Zahir Nikraftar[1], Rendani Mbuvha[2,3], Mojtaba Sadegh[4,5], and Willem A. Landman[6]

[1]Machine Intelligence and Decision Systems (MInDS) Research Group, School of Electronic Engineering and Computer Science, Queen Mary University of London (QMUL), London, UK, [2]Department of Mathematics, Statistics and Probability Group, University of Manchester, Manchester, UK, [3]School of Statistics and Actuarial Science, University of Witwatersrand, Johannesburg, South Africa, [4]Department of Civil Engineering, Boise State University, Boise, ID, USA, [5]United Nations University Institute for Water, Environment and Health, Richmond Hill, ON, Canada, [6]Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa

Key Points:

- Season-Net combines U-Net and ConvLSTM to correct spatial and temporal biases in seasonal temperature forecasts
- A temporal sliding-window QM loss enforces window-local CDF calibration, boosting extreme accuracy while preserving temporal awareness
- Across regions, Season-Net outperforms QM and LS on MAE and Kendall's tau, with the most notable gains in CRPSS and the Brier skill score (BSS)

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Z. Nikraftar,
z.nikraftar@qmul.ac.uk

Citation:

Nikraftar, Z., Mbuvha, R., Sadegh, M., & Landman, W. A. (2025). Season-net: A deep learning framework for bias correction of seasonal forecasting models. *Journal of Geophysical Research: Machine Learning and Computation*, *2*, e2025JH000957. https://doi.org/10.1029/2025JH000957

Abstract Seasonal climate forecasts play a crucial role in decision-making across sectors like agriculture, energy, and disaster management. However, these forecasts often exhibit spatially structured biases that undermine their reliability, but this structure also enables more effective bias correction, particularly improving performance in predicting temperature extremes. Traditional bias correction methods such as quantile mapping (QM) and linear scaling (LS) are limited by assumptions of stationarity and their inability to capture complex spatiotemporal patterns. To address these challenges, we introduce Season-Net, a hybrid deep learning framework combining U-Net and ConvLSTM architectures. Season-Net is used to perform bias correction on seasonal daily temperature forecasts from the Met Office (GloSea6) and Météo-France (System 8) by learning season-specific spatial and temporal dependencies in a unified architecture. The model is trained with a novel sliding-window quantile mapping loss function that introduces temporal awareness into the quantile mapping process, enhancing its ability to capture temperature distribution and evolution. Evaluations across North America and Africa show that Season-Net consistently outperforms QM and LS in both deterministic (e.g., RMSE and Kendall's Tau) and probabilistic (e.g., Brier skill score and CRPSS) metrics. Furthermore, Season-Net excels in impact-based evaluations, significantly improving the prediction of extreme temperature events. These results highlight the superior capability of deep learning methods in correcting spatially structured seasonal forecast biases and enhancing the utility of climate predictions for climate-sensitive applications. Season-Net offers a promising pathway for advancing seasonal forecast postprocessing with high accuracy and impact relevance.

Climate Dynamics (2026) 64:148
https://doi.org/10.1007/s00382-026-08077-4

ORIGINAL ARTICLE

Seasonal forecasting using the GenCast probabilistic machine learning model

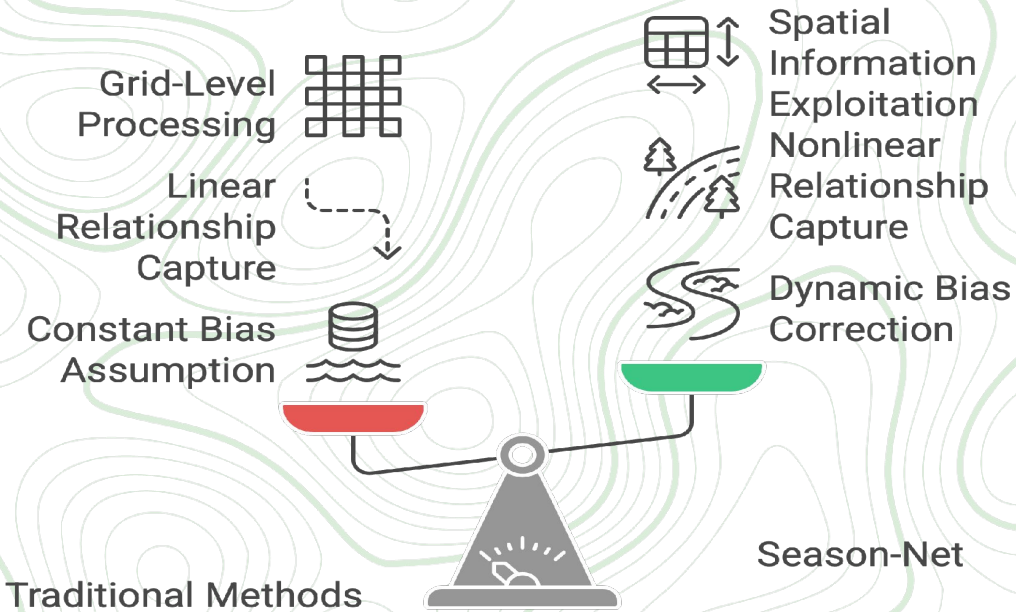
Bobby Antonio[1] · Kristian Strommen[1,2] · Hannah M. Christensen[1]

Received: 12 September 2025 / Accepted: 11 January 2026 / Published online: 16 March 2026
© The Author(s) 2026, modified publication 2026

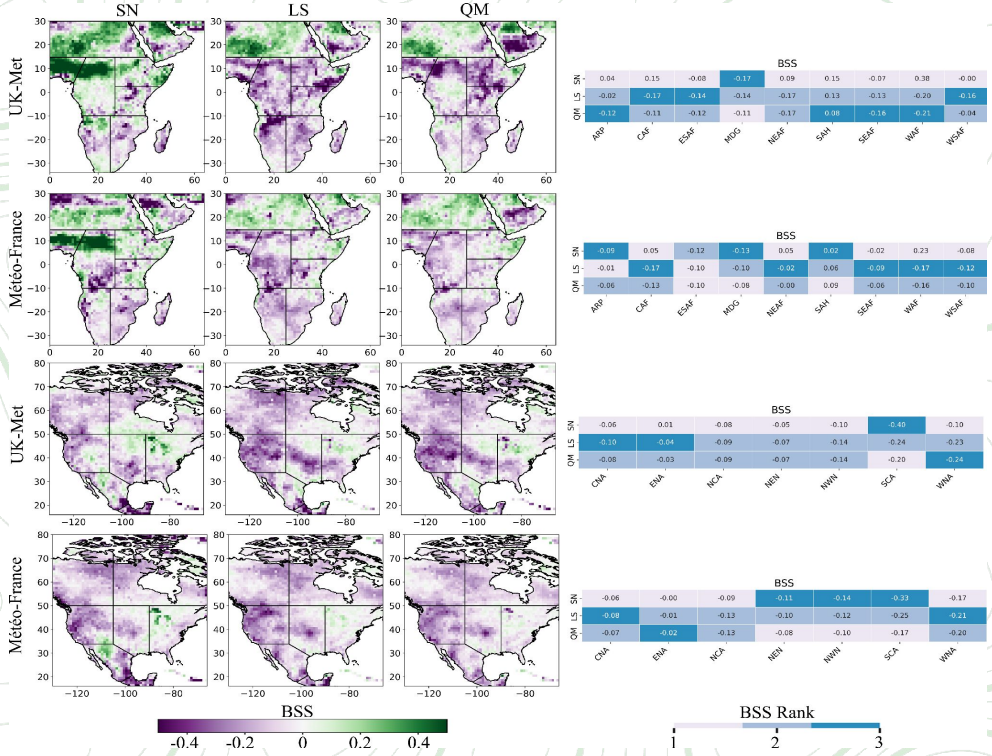
Abstract
Machine-learnt weather prediction (MLWP) models are now well established as being competitive with conventional numerical weather prediction (NWP) models in the medium range. However, there is still much uncertainty as to how this performance extends to longer timescales, where interactions with slower components of the earth system become important. We take GenCast, a state-of-the-art probabilistic MLWP model, and apply it to the task of seasonal forecasting with prescribed sea surface temperature (SST), by providing anomalies persisted over climatology (GenCast-Persisted) or forcing with observed SSTs (GenCast-Forced). The forecasts are compared to the European Centre for Medium-Range Weather Forecasts seasonal forecasting system, SEAS5. Our results indicate that, despite being trained at short timescales, GenCast-Persisted produces much of the correct precipitation patterns in response to El Niño and La Niña events, with several erroneous patterns in GenCast-Persisted corrected with GenCast-Forced. The uncertainty in precipitation response, as represented by the ensemble, compares favourably to SEAS5. Whilst SEAS5 achieves superior skill in the tropics for 2-metre temperature and mean sea level pressure (MSLP), GenCast-Persisted achieves higher skill in some areas in higher latitudes, including mountainous areas, with notable improvements for MSLP in particular; this is reflected in a slightly higher correlation with the observed NAO index. Reliability diagrams indicate that GenCast-Persisted has little skill relative to climatology, whilst GenCast-Forced produces forecasts with reliability comparable to SEAS5. These results provide an indication of the potential of MLWP models similar to GenCast for the 'full' seasonal forecasting problem, where the atmospheric model is coupled to ocean, land and cryosphere models.

Keywords Seasonal forecasting · Machine learning · Numerical weather prediction

Motivation for AI Methods

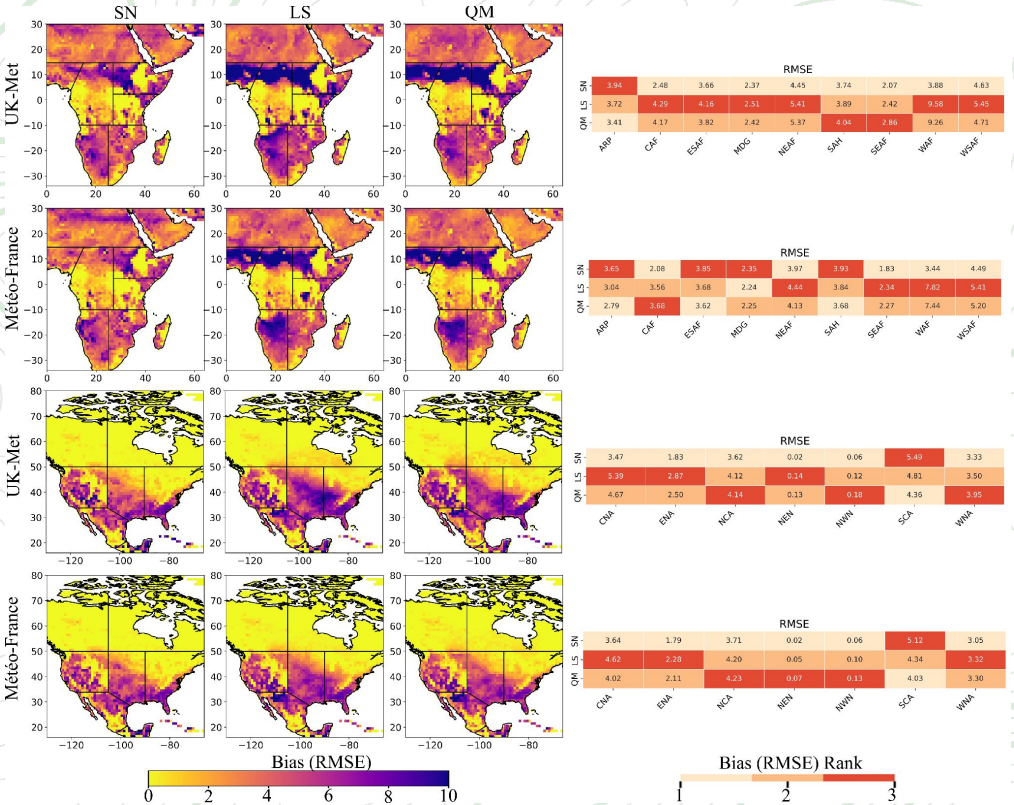


Bias Correction Results – Brier Skill Score



Brier Skill Score (BSS) over Africa (rows 1&2) during SON season and North America (rows 3&4) during JJA season over Season-Net (column 1), LS (column 2), and QM (column 3). The average over IPCC regions is presented in the table ate the right side of each row. The bias colour bar is constrained to the range of [-0.5, 0.5] for consistency in visualization over different maps.

Impact Based Evaluation – Warm Days Frequency (WDF))



Warm Days Frequency (WDF) - counts the number of days exceeding 35°C

RMSE maps for WDF index in the test period over Africa (rows 1&2) during the SON season and North America (rows 3&4) during the JJA season over Season-Net (column 1), LS (column 2), and QM (column 3). The average over IPCC regions is presented in the table ate the right side of each row. The bias colour bar is constrained to the range of [0, 10] for consistency in visualization over different maps

Persisted AI Weather Model Results

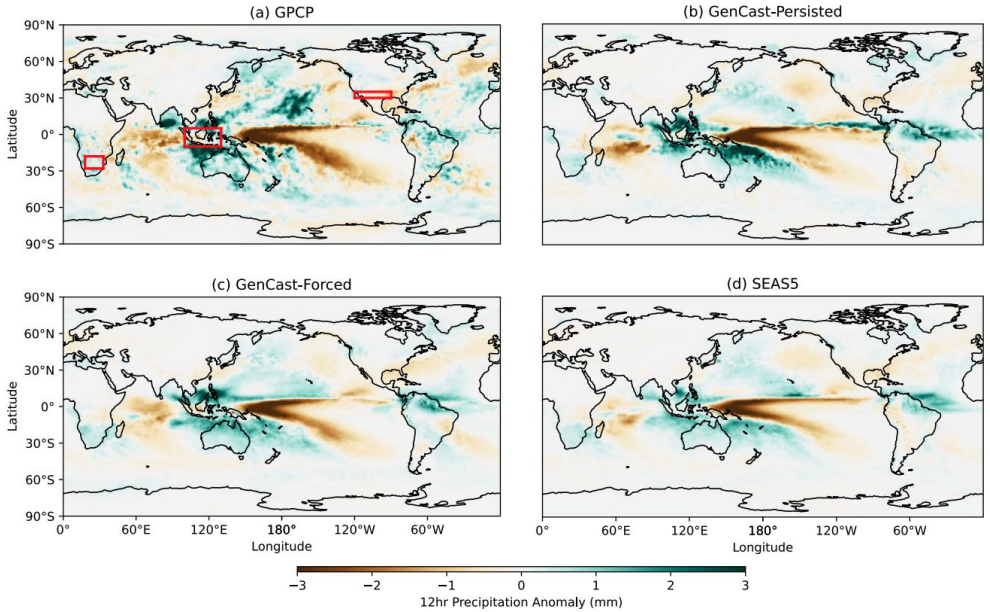


Fig. 1 Seasonal 12hr precipitation anomalies for December 2010–February 2011 (Strong La Niña). Red boxes indicate the subregions used in Fig. 3

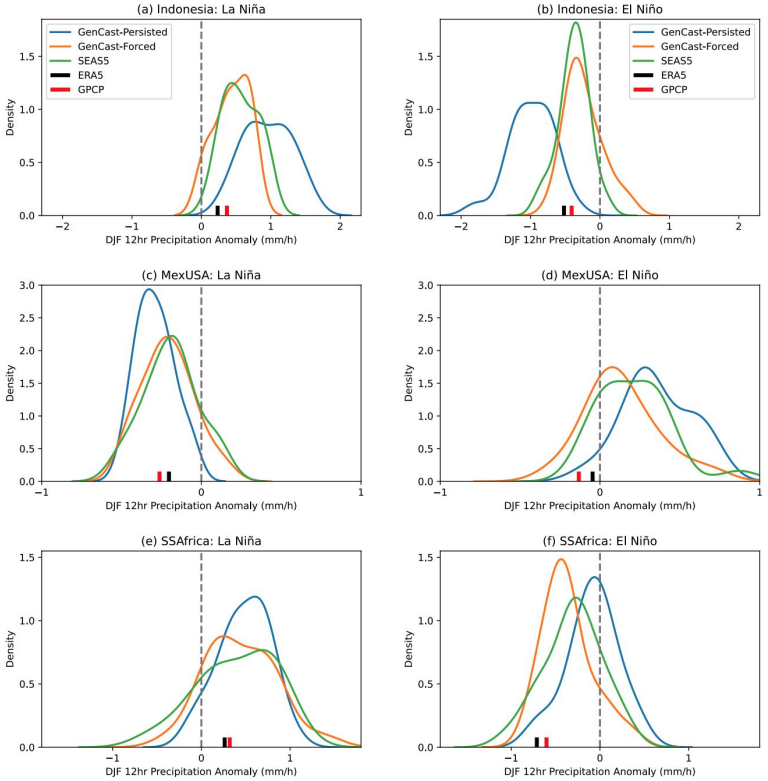


Fig. 3 Kernel density estimate for the distribution of DJF 12hr precipitation anomalies, averaged over each of the subregions shown in Figs. 1 and 2 (rows), and for each of the La Niña and El Niño case studies (columns). In each plot the ERA5 value is shown in black

Antonio, B., Strommen, K. and Christensen, H.M., 2026. Seasonal forecasting using the GenCast probabilistic machine learning model. Climate Dynamics, 64(4), p.148.

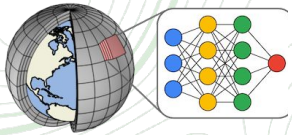
ImvulaNet: Forecast Co-Production for Local Impact

Farmer Sourced Local Rainfall and Yields

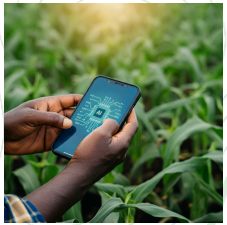








NeuralGCM



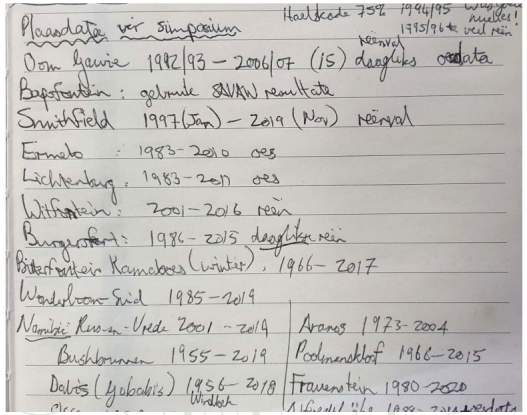
frontiers in Climate

ORIGINAL RESEARCH
published: 08 September 2020
doi: 10.3389/fclim.2020.00005

Citizen Science for the Prediction of Climate Extremes in South Africa and Namibia

Willem A. Landman[1,2]*, *Emma R. M. Archer*[1] and *Mark A. Tadross*[3]

[1] Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa, [2] International Research Institute for Climate and Society, The Earth Institute of Columbia University, New York, NY, United States, [3] Climate Systems Analysis Group, University of Cape Town, Cape Town, South Africa



Source: Landman, W. A., Archer, E. R., & Tadross, M. A. (2020). Citizen science for th prediction of climate extremes in South Africa and Namibia. Frontiers in Climate, 2, 5.

Integrated Decision Support tool for farmers

AfriClimate AI
AI for a Climate-Resilient Africa

IDRC · CRDI
Canada

ARC · LNR

UK International Development
Partnership | Progress | Prosperity

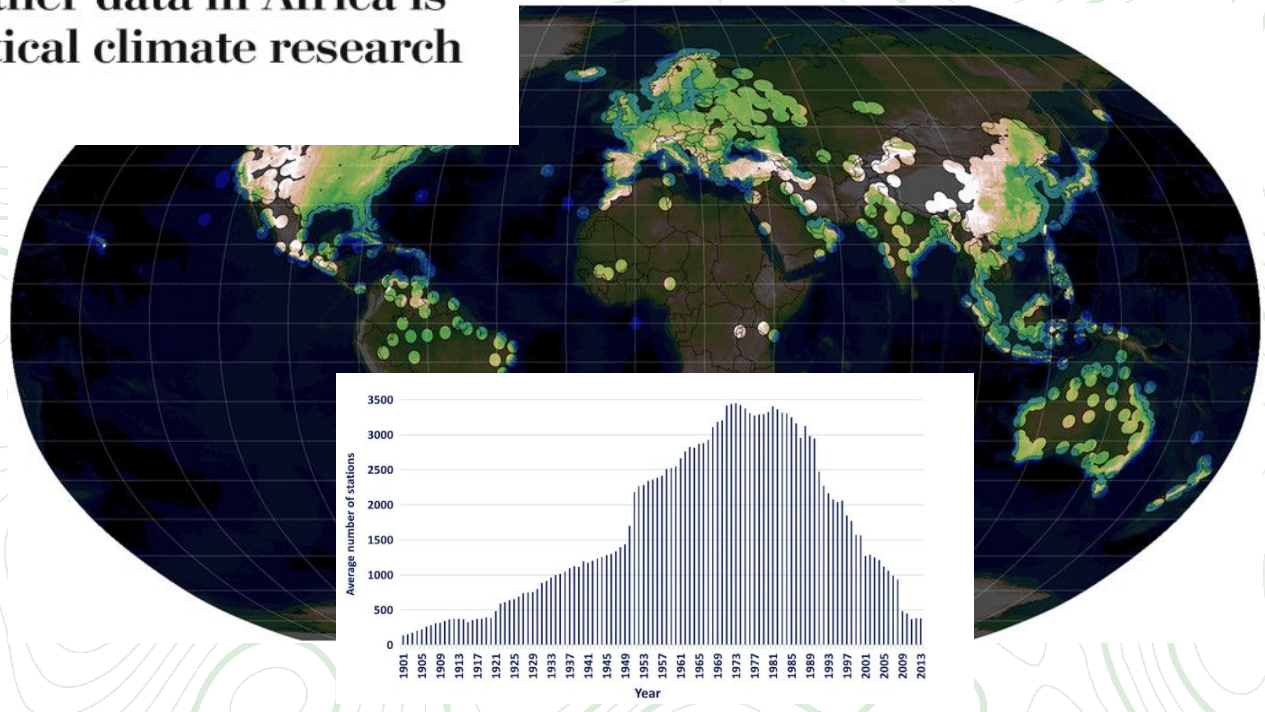
Caveat: The Data Scarcity Curveball

The Washington Post
Democracy Dies in Darkness

AFRICA

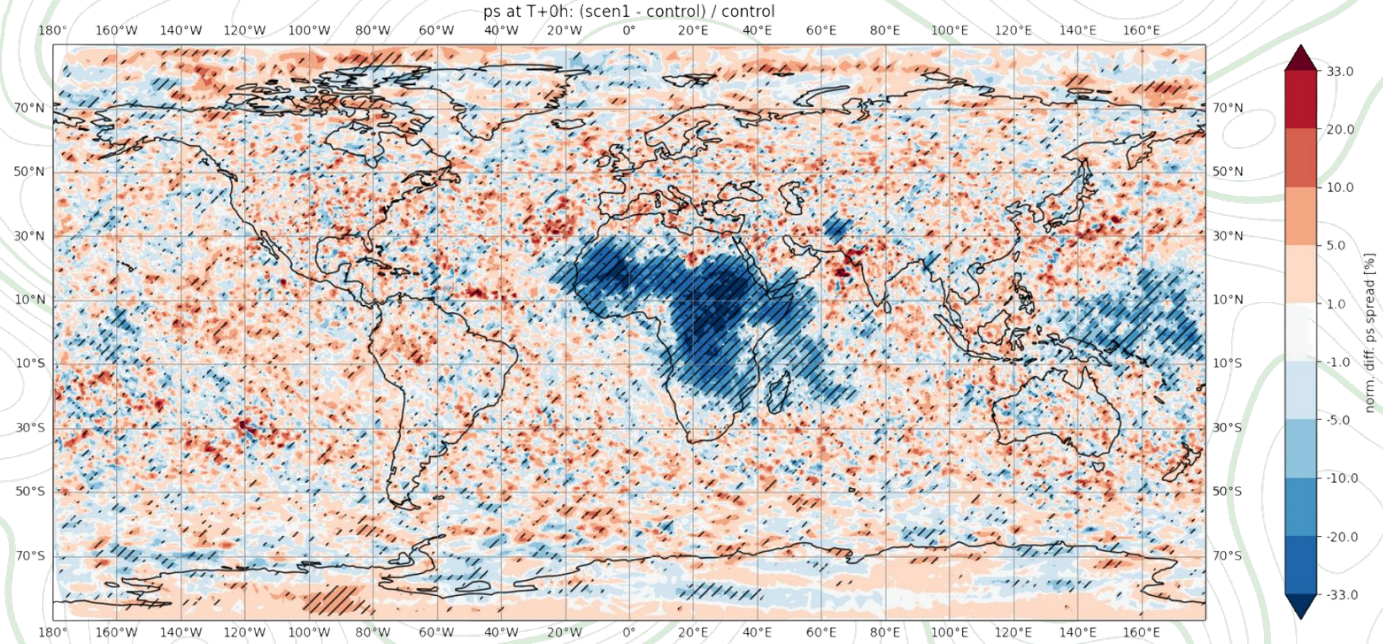
A lack of weather data in Africa is thwarting critical climate research

By Rachel Chason and Rael Ombuor
September 24, 2021 at 5:00 a.m. EDT



Evolution of the number of weather stations used in the GPCC full-data product over Africa (T.DINKU)

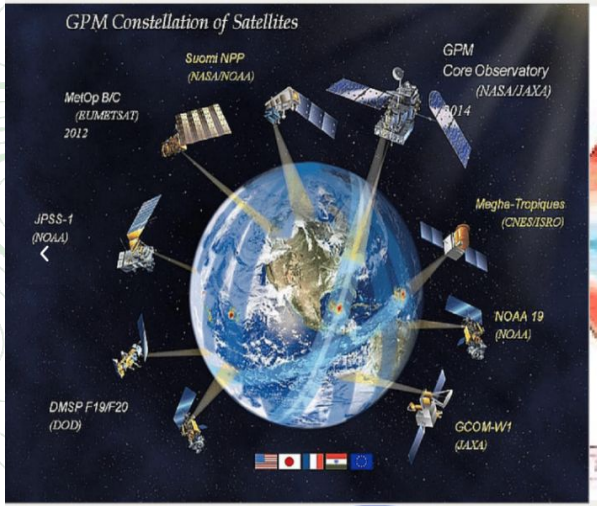
Data Scarcity Hurts Weather Forecasting



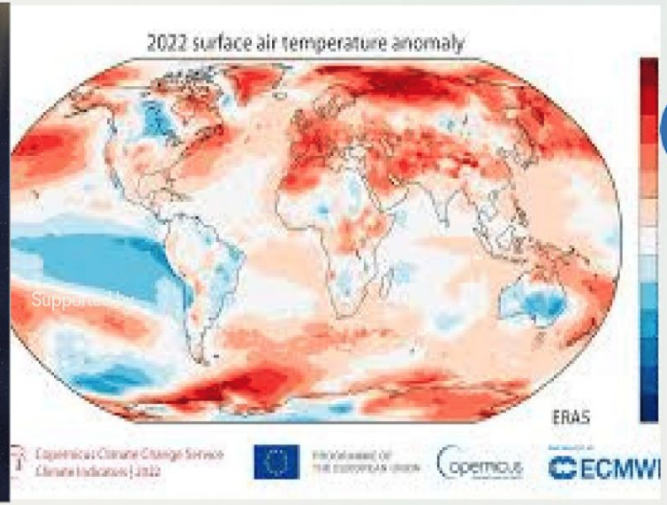
Percentage reduction in the surface pressure analysis uncertainty with simulated additional surface observations over Africa (ECMWF, 2025)

Turn to Global Proxies?

Satellite Products



Renanalysis Products



Merged / Blended Products

MSWEP

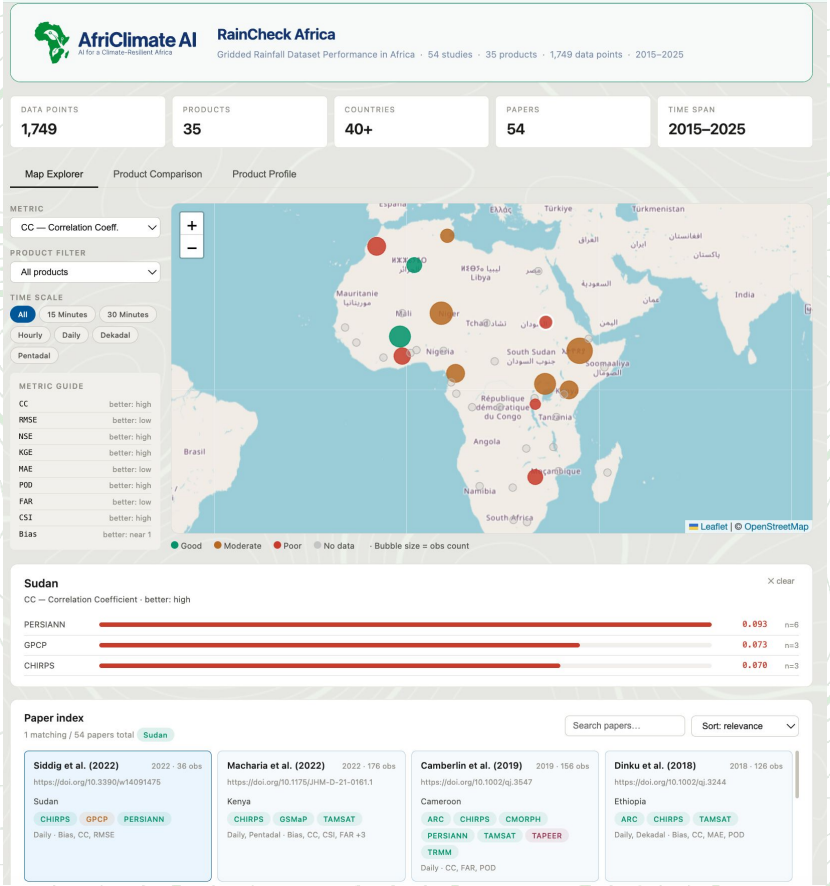
Climate Hazards Center
UC SANTA BAR

NASA

全球降水観測計画
GPM
GLOBAL PRECIPITATION MEASUREMENT

ECMWF

But Which Proxy? …RainCheck Africa



https://africlimate-research.github.io/RainCheckAfrica/

Key Takeaways

Part A — Short-Range Forecasting

- AI models (GraphCast, GenCast, AIFS, Prithvi WxC) now rival physics-based NWP, at a fraction of the compute
- Probabilistic skill — calibration, discrimination & uncertainty (IDR) — matters as much as raw accuracy
- Hybrid AI–physics systems are the near-term operational reality

Part B — Seasonal Forecasting

- Correlation & RMSE can mask poor skill on the impacts that matter heavy rain days, warm spells, drought
- Bias correction (linear scaling, quantile mapping, Season-Net) is essential before forecasts are usable
- Emergence of persisting AIWP for Seasonal and sub seasonal time scales
- AI narrows the gap here too, but gains remain regionally uneven

From Research to Impact

Forecast skill only creates value once it reaches the people who act on it. Co-production with farmers and national met services model output into trusted, decision-relevant information.

Quiz Time: Join on Slido



https://app.sli.do/event/nCP9ZbfHnNTq9N8scZjmjV

Key Readings

1. **Bauer, P., Thorpe, A., & Brunet, G.** (2015). *The quiet revolution of numerical weather prediction.* Nature, 525(7567), 47-55.
2. **Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., ... & Battaglia, P. (2023).** Learning skillful medium-range global weather forecasting. Science, 382(6677), 1416-1421.
3. **Nikraftar, Z., Mbuvha, R., Sadegh, M., & Landman, W. A. (2024).** *Impact-based skill evaluation of seasonal precipitation forecasts.* Earth's Future, 12, e2024EF004936.
4. **Antonio, B., Strommen, K., & Christensen, H.M. (2026).** *Seasonal forecasting using the GenCast probabilistic machine learning model.* Climate Dynamics, 64(4), p.148.

Thank you

shruti@afrilimate.ai

rendani@afrilimate.ai

https://africlimate.ai/

