AI for Weather Task Team

Chairs: Cat de Burgh-Day and Jørn Kristiansen (and Veronique Bouchet)

Members: Florian Teichert, Stephen Belcher, Ariane Frassoni dos Santos, Chris Davis, David John Gagne, David Richardson, Joshua Fu, Nils Wedi

Ex. Officio: Lauren Stuart, Yuki Honda, Tatsuya Kimura, Estelle de Conig

WORLD METEOROLOGICAL ORGANIZATION

1950-2025 SCIENCE for ACTION

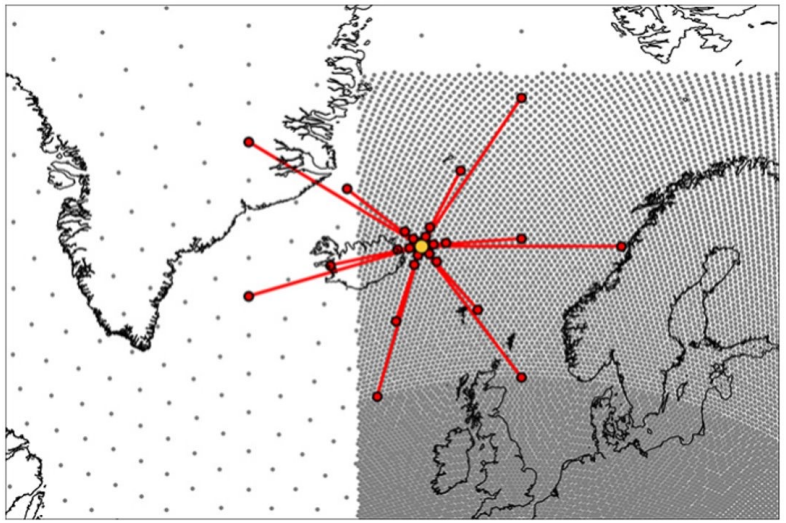
Context

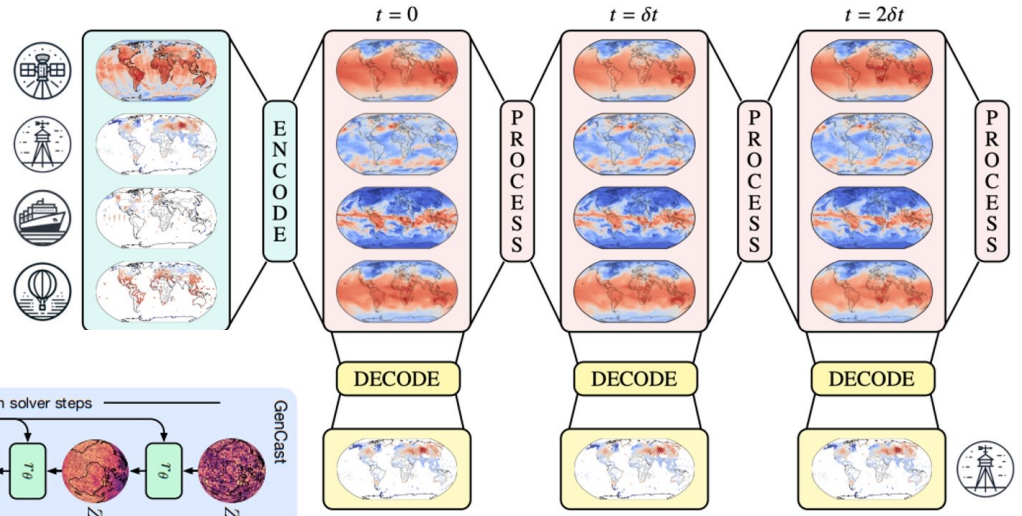
Background

Pre-2022: Two ML atmosphere models in the literature, neither with performance competitive with traditional NWP models

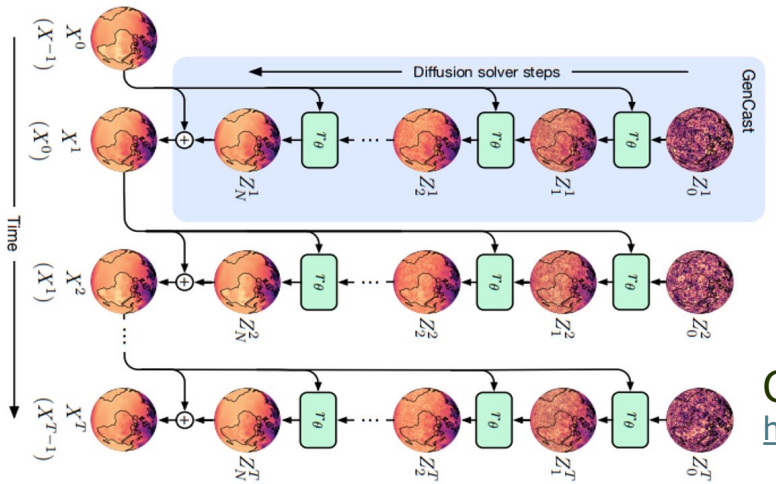
Post 2022 to today: more than 40 AI/ML atmosphere, ocean, land surface, wave, sea ice, coupled, foundation, etc. models. Timescales from DA to hours to weeks (and now months to years…)

From the ECMWF AIFS Blog:
https://www.ecmwf.int/en/about/media-centre/aifs-blog/2024/data-driven-regional-modelling





Aardvark Weather
https://arxiv.org/pdf/2404.00411



GenCast (DeepMind)
https://arxiv.org/pdf/2312.15796

Context

TT-AI4Wx - Overview

- A first response to the AI weather forecasting revolution
- Scope –

  The TT-AI4WX will:

  - *Generate a list of existing activities within WMO programs and regional associations;*
  - *Review activities which can be expanded with limited resources to include AI/ML systems;*
  - *Identify main gaps and science questions of immediate relevance from the community of experts making up the TT; and*
  - *Develop a report on the above including recommendations on immediate actions.*
- Meetings initially fortnightly (~6 months) then monthly (~6 months)
- Membership from RB, SAP, WGNE, WWRP, SERCOM, INFCOM, GAW, and PPE office
- Note lacking RA presence – an issue for LMIC representation
- Concluded at 2025 RB meeting in September – Final Report accepted by RB vote

Context

Activities

- Identified activities underway across the WMO
- Formulated a framework of AI-related work and questions across NMHS activities
- Identified discrepancy between Regions/HICs and LMICs
- Highlighted other focus areas – PPE, training needs, infrastructure needs
- Scientific priorities - other groups have done a great job on specific scientific priorities.TT-AI4Wx sought to fill gaps – technical, infrastructure, data access and needs etc.
- Provided recommendations for the RB, and discussed unaddressed questions for future consideration

Outcomes

Deliverables

- Interim report and open data concept note
- Paper database (~600 and counting…)

Accessibility Benchmark Datasets Climate CNN Coupled
Data Assimilation Data Driven Models Downscaling
Ensembles/Uncertainty Evaluation Explainability Finetuning
Foundation Gappy data interpolation Geometric Deep Learning
GNN Graph neural network Graph Transformer Graphs Group
Hybrid Hydrology/Land Ice/Sea Ice LAM/Regional LLM
Neural Operator Nowcasting Object Identification Ocean
Other Parameterisation Perspective Platforms
Postprocessing Review RNN Seminal Paper
Single variable/process Space Weather Subseasonal/seasonal
Survey Transformer Underpinning Theory/Speculative Research
Visual Transformer Waves

- Final Report and framework of considerations
- Informed guidance on what next
- All outputs are here: https://community.wmo.int/en/task-team-artificial-intelligence-weather-tt-ai4wx

Concept Note: Maintaining the open principles of the WMO in an AI/ML world

Context: With the advances in Machine Learning (ML)/Artificial Intelligence (AI) approaches in observations, predictions and services for weather, climate, and related Earth system information, profound transformations in the development, delivery, and access to hydrometeorological services are underway.

These include:

- AI/ML-based weather and climate models and emulators that could make it feasible to generate forecasts at very low computing costs
- Acceleration in the availability and broad access to data-driven prediction systems with improved quality and accuracy
- Data-driven systems that address the prediction needs at the local to very-local scales

It is the opinion of the Task Team participants that in order to realize these opportunities for all weather services and Members of WMO, the data, code, and intellectual property that

	Fundamental research/ development	Scoping/strategic considerations	NHMS customization/ implementation	Operationalization	Postprocessing/ Downstream applications	Use of Information
Model/ application Development	• Foundation models, • Explainable AI (XAI), • ESM extension (ocean, ice, land, hydrology, etc.), nowcasting • Regional and global AI models, • Ensembles, • Nudging methods	• Evaluation of different AI architectures • Reduction in cost/increase in timeliness/versatility of prediction systems • Product focus (parameter/ global/regional/ensemble)	• Regional AI modelling/fine-tuning • (Global) model fine-tuning • Forecast nudging • Measurement-model fusion by AI	• Path to operations - R2O processes, workflows for ML-ops (repeatable training and results) • Online learning	• Downscaling • AI for forecast augmentation and forecaster support	• Monitoring of forecaster use of AI support and augmentation
Evaluation and Verification	• Verification of AI models • AI based uncertainty quantification • Physical explainability • Attribution to climate change with AI models	• Verification of models • Reduce cost/increase accuracy of constraining prediction uncertainty • Model efficiency/ practicality evaluation (on top of accuracy)	• Verification of models • Add value to forecast products • Add context, e.g. attribution, impact sectors	• Verification of AI against physical models • Building trust in AI forecasts	• Verification of models • AI based uncertainty quantification	• AI based evaluation of forecast outcomes • User response analysis using AI, and of AI models • Measure trust in AI forecasts
Observations (obs), Data Assimilation (DA), Initial Conditions	• AI based observation QA, correction and monitoring • AI for DA and observation operators • AI forecasts directly from observations • AI model performance in data sparse regions and outside training area	• Future obs infrastructure and design How much more of a priority will obs become? • Optimal obs sites for AI • Reduce cost/complexity of global DA • Exploit regional/local knowledge	• AI based observation QA, correction and monitoring • Develop regional AI-based forecast from obs	• AI based observation quality assurance, correction and monitoring • Develop global, regional and local impact AI-based forecast from obs	• AI based impact sector DA/assessment modelling	• Observations of decision making and actions using AI (building datasets for evaluation of user uptake and utility)
AI tooling and accessibility	• Development of tooling for improved accessibility to training data and data/methodology openness • Big data handling, geographically distributed data use	• AI infrastructure and software: - Cost and accessibility - Open-source frameworks - Specific needs of developing countries	• Sharing of open access AI tools and experience • Sharing of data • Access to tooling and models in NHMSs in developing countries	• How to ensure data and models remain open • Ongoing maintenance and support of tooling	• Integration of tooling to include postprocessing and downstream applications	• Routine assessment of state of the art and knowledge of AI tools among stakeholders • Assessment of the uptake of AI tools and technologies in developing countries
Training (especially for developing countries)	• (People) Training in use of AI tools, ML model research and development	• Identifying training/ knowledge sharing needs: - Levels of complexity/expertise - Specific needs of developing countries	• Establish platform for sharing experience and training materials • Training in ML model finetuning, customisation and evaluation	• Training in ML model operational considerations and technical requirements	• Training in ML model postprocessing, and ML models for postprocessing	• Training in communication/user engagement around ML • User education
Data/code access and availability	• Data for training and retraining: collections, access, use guidance	• Data access and sharing guidelines on use • Catalogue of collections • Guidance on public-private partnerships • Ensuring ongoing open data and models as default	• Sharing Data for AI/ML training or retraining • Sharing of inference weights or data	• Data for training or retraining • Licencing of model models and model training data – implications for commercial use of forecasts	• Data for training or retraining	• Understanding obstacles to data access from users of information

WORLD METEOROLOGICAL ORGANIZATION

1950-2025 SCIENCE for ACTION

Outcomes

Scientific Priorities

Interpretability and explainability	Improve AIWP performance on rare and high-impact events while enhancing trust and understanding through physical constraints, case sharing, and targeted benchmarking.
Upstream of the model - observations, DA and initial conditions	Enhance AIWP by integrating AI with data assimilation, optimizing observations, and evaluating model sensitivity to initial conditions, particularly in observation-sparse regions.
Downstream of the model - post-processing	Use AI to improve forecast skill and analyze extreme weather drivers, while managing high data volumes and ensuring careful quality control.
Model complexity and computational cost	Maintain or improve AIWP performance with smaller, efficient, physics-informed models using fine-tuning, regional models, or model distillation to reduce computational demands.
Infrastructure and data design and development	Enable equitable and efficient AI-enabled forecasting through lightweight models, shared platforms, improved data infrastructure, and reproducible standards.

WORLD METEOROLOGICAL ORGANIZATION

1950-2025 SCIENCE for ACTION

**Note – complimentary with scientific priorities identified in the WIPPS AI Roadmap

Outcomes

Recommendations

Key Takeaways

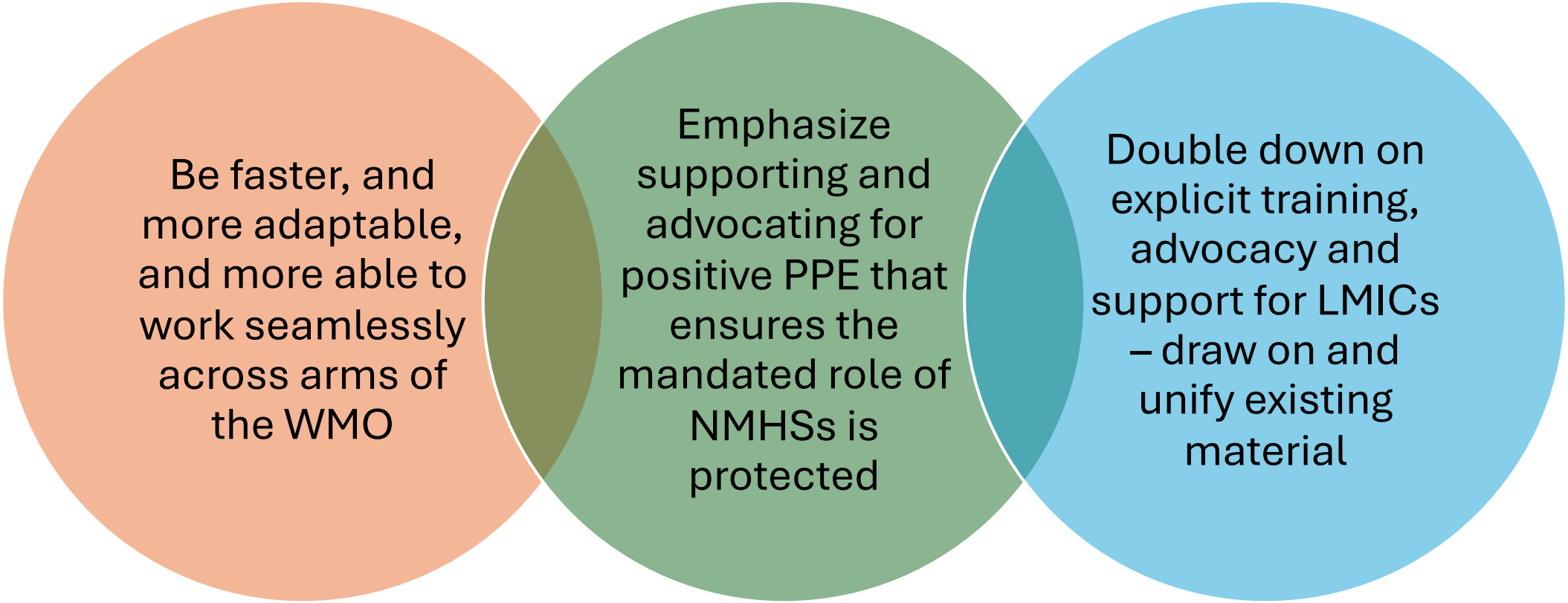
Coordination & Agility	Strengthen cross-body collaboration; enable rapid AI adoption
Equity & Capacity Building	Empower underrepresented regions; provide training, infrastructure, and access to AI tools
Research & Innovation	Focus on AI methods, high-impact events, distributed computing; promote carbon-efficient workflows; address data-sparse regions and LMIC needs
Workforce Development	Build skills in AI implementation, interpretation, and ethics; include societal data expertise
Infrastructure & Standards	Ensure robust observation/reanalysis networks; establish shared platforms, data standards, open AI tools, and trustworthy evaluation

WORLD METEOROLOGICAL ORGANIZATION

1950-2025 SCIENCE for ACTION

Looking forward

What does the WMO need to do differently/more?



WORLD METEOROLOGICAL ORGANIZATION

Looking forward

Future focus areas

ML for sub-seasonal, seasonal and climate

- Coupled models
- Insufficient evaluation data
- Impact of nonstationary climate on training
- Model stability
- Emulators and projections – climate change signal?

Equity and accessibility for LMICs

- Fairness in ML models
- Impact of observations sparsity in LMICs
- Need for evaluation over LMICs
- How to support needs of LMICs in terms of infrastructure?

ML for, and from, observations and DA

- The impact of ML on the design and scale of observation networks
- The future role reanalyses – design considerations for ML
- Progress in ML for DA
- Straight from observations ML models – implications for DA and reanalyses

Social sciences and ML

- The role of ML in modulating engagement how people engage with warnings and impacts
- How ML can be used in social sciences

Training and capacity building - should include academics

Foundation models

- One model for all timescales?

Evaluation and benchmarking

- One-shot and zero-shot models?

Infrastructure, tooling, and standards

- How to ensure open and accessible technologies, code, weights and data?
- How to support the development of accessible and affordable tooling?
- Best practice/standards for evaluation and benchmarking

WORLD METEOROLOGICAL ORGANIZATION

1950-2025 SCIENCE for ACTION

*** Evaluation, verification and benchmarking is essential across all these areas**

Looking forward

Recommended next steps

Ongoing strategic oversight/coordination across WMO AI activities

Coordinating body?

How will it interface with SAP, RB etc?

How will information flow to and from elsewhere?

Targeted consideration of high-level strategic questions

Where can AI contribute?

AI for "last mile" forecast interpretation, communication?

Unique role of NMHSs when the barrier to entry for weather prediction is lower?

Equitable PPE on AI

Impact on LMICs?

What additional roles for WMO in the above?

A nuanced voice for LMICs

Consultative group?

Mandated representation?

Something else?

Coordinated, accessible infrastructure/platforms

Focus on capacity building with existing infrastructure?

Advocating for accessibility?

Supporting portable/accessible infrastructure (e.g. Forecast in a box)?

Targeted actions:

S2S and climate

Observations and data assimilation

Evaluation and benchmarking

Accessibility for LMICS

Training and capacity building needs to include academics

Risk: expectations may be getting too high - How do we balance expectations against ability to deliver?

WORLD METEOROLOGICAL ORGANIZATION

1950-2025 SCIENCE for ACTION

Looking forward

Recommended next steps

Ongoing strategic oversight/coordination across WMO AI activities

Coordinating body?

How will it interface with SAP, RB etc?

How will information flow to and from elsewhere?

JAG-AI

Targeted consideration of high-level strategic questions

Where can AI contribute?

AI for "last mile" forecast interpretation, communication?

Unique role of NMHSs when the barrier to entry for weather prediction is lower?

Equitable PPE on AI

Impact on LMICs?

What additional roles for WMO in the above?

SAP

A nuanced voice for LMICs

Consultative group?

Mandated representation?

Something else?

JAG-AI

Coordinated, accessible infrastructure/platforms

Focus on capacity building with existing infrastructure?

Advocating for accessibility?

Supporting portable/accessible infrastructure (e.g. Forecast in a box)?

INFCOM

Targeted actions:

S2S and climate **(WCRP/WWRP joint effort)**

Observations and data assimilation **(WWRP/GAW)**

Evaluation and benchmarking **(WGNE/WWRP/WIPPS etc.)**

Accessibility for LMICS **(JAG-AI)**

Training and capacity building needs to include academics **(RPs and CDP coordination through JAG-AI)**

Risk: expectations may be getting too high - How do we balance expectations against ability to deliver?

Conclusion: A new task team under the RB is not needed, but RB continues to support and integrate these efforts

Summary

- The TT-AI4Wx has completed an important first step in the WMO's response to the AI revolution
- The WMO will need to continue to respond in full across all elements of NMHSs/member responsibilities
- The rapidly evolving nature of the field will require an equal (and thus arguably unprecedented) level of agility in the WMO
- The RB has noted the large number of activities now underway across the WMO and RB subsidiary bodies, and thus has not formed further AI-related TTs
- WGNE has a critical role to play in guiding and contributing to WMO activities on AI, and supporting WMO members