The Abdus Salam
International Centre
for Theoretical Physics

6th Summer School on Theory, Mechanisms and Hierarchical Modelling of Climate Dynamics: Artificial Intelligence and Climate Modelling | (SMR 4067)

05 May 2025 - 16 May 2025
ICTP, Trieste, Italy

T01 - ASHFAQ Moetasim

Towards Development of Generative AI-based Sub-Seasonal to Seasonal Forecasting System

T02 - BADURU Balaji

A Probabilistic Forecast for Multi-year ENSO Using Bayesian Convolutional Neural Network

T03 - CASALLAS GARCIA Alejandro

From observations to realistic simulations, the mechanisms that develop and prevent convective organization

T04 - COLFESCU Ioana

A machine learning method is used to identify sources of long-term ENSO predictability in the ocean (sea surface temperature (SST) and heat content) and the atmosphere (near-surface zonal wind (U10)). Tropical 10m wind sseem to represents an inreresting source of ENSO predictability.

T05 - CORNER Jeremy Malcom

Investigating Extreme Precipitation Event Attribution Through Deep-Learning Approaches

T06 - CORON Domitille Tihaine

AI-Driven reconstruction of oxygen profiles in the North Atlantic Oxygen Minimum Zone

T07 - GALVAN FRAILE Victor

Enhancing Seasonal Predictions with Machine Learning: A Global Perspective on SST Influence in Early Winter

T08 - GORENSTEIN Iuri

Towards a Physics-Informed Neural Network for Ocean Circulation Modelling

T09 - GRANDE Davide

Learning the Ocean: Leveraging AI to Improve Ocean Data Assimilation

T10 - KAYODE Florence Abiodun

Improving Global Wildfire Simulations Using Machine Learning

T11 - KIEU Chanh Quoc

Retrieving tropical cyclone intensity from climate output: AI approaches and challenges

T12 - MANIMPIRE Gasana Elysee

RECONSTRUCTION OF THE SURFACE GEOSTROPHIC CURRENT IN THE MEDITERRANEAN SEA FROM ABSOLUTE DYNAMIC TOPOGRAPHY GENERATED BY GENERATIVE ADVERSARIAL NETWORK (ADTGAN)

T13 - OLMO Matias Ezequiel

Climate projections in the Euro-Mediterranean region: atmospheric circulation patterns and temperature and rainfall future changes

T14 - PALMA I GARCIA Lluis

On the application of generative modelling for seasonal climate predictions

T15 - PECHACOVA Alzbeta

Data-Driven Equation Discovery of In-Cloud Vertical Velocity

T16 - PEREZ ZANON Nuria

Exploring windows of opportunity at sub-seasonal time scales for extreme precipitation events: the Valencia DANA case study

T17 - POLESELLO Andrea

Investigating deep convective systems with Machine Learning and causal inference across a hierarchy of datasets

T18 - PROVENZANO Elena

CNN-based forecasting of early winter NAO using sea surface temperature

T19 - SANSINE Vateanui Francois Ernest

Fine-tuning a global weather model for superior subseasonal forecasting

T20 - SOTO LOPEZ Carlos Enmanuel

Data-Informed Inversion Model (DIIM): a framework to retrieve marine optical constituents in the BOUSSOLE site using a three-stream irradiance model

T21 - YUNITA Rezky

Local Climate Zone Classification in Jakarta Using AI: Machine Learning Approach with Landsat 9 and Land Surface Temperature Integration

T01

Towards Development of Generative AI-based Sub-Seasonal to Seasonal Forecasting System

Authors: Moetasim Ashfaq[1], Takuya Kurihana[1,2], Nathaniel Johnson[3], and Katherine J Evans[1]
1: Oak Ridge National Laboratory, 2: Fujitsu Research of America, Inc.,
3: NOAA Geophysical Fluid Dynamics Laboratory

While internal atmospheric variability plays a crucial role in sub-seasonal to seasonal (S2S) forecasting, its predictability remains a significant limitation in physical seasonal forecasting systems. Specifically, our previous work underscores its importance in accurately predicting winter precipitation variability in the Northern Hemisphere. To address this limitation, we are developing a Generative Artificial Intelligence (AI)-based S2S system.
Our initial development focuses on adapting a conditional denoising diffusion probabilistic model—one of the leading generative AI algorithms—to generate a 50-member ensemble for S2S predictions over the Northern Hemisphere (20°N-90°N). The model is designed to forecast two-week averaged geopotential height at both 200 hPa and 500 hPa in advance, for the Northern Hemisphere winter. The model is trained on seasonal ensemble forecasting datasets from the Copernicus Climate Data (1981–2021) and further fine-tuned with ERA5 reanalysis data from the same period. This training was performed on 16 AMD MI250X GPUs on the Frontier supercomputer at the Oak Ridge Leadership Computing Facility.
We evaluated the model's performance during the winter months (November to February) from 2022 to 2024, comparing it against ERA5 reanalysis data. Figure 1 presents snapshots of two-week averaged geopotential height at 200 hPa between January 15 and January 28, 2022, as predicted by our diffusion model (Diff.) and six seasonal ensemble forecasting models. Our model not only generates qualitatively realistic spatial patterns but also quantitatively outperforms physical seasonal forecasting systems on average (Figure 2). Furthermore, the model effectively captures key atmospheric modes of variability, such as the North Atlantic Oscillation (NAO), and other extratropical forcing patterns critical for precipitation variability over Central and Southwest Asia. This suggests that generative AI models could offer significant improvements in predicting atmospheric variability.

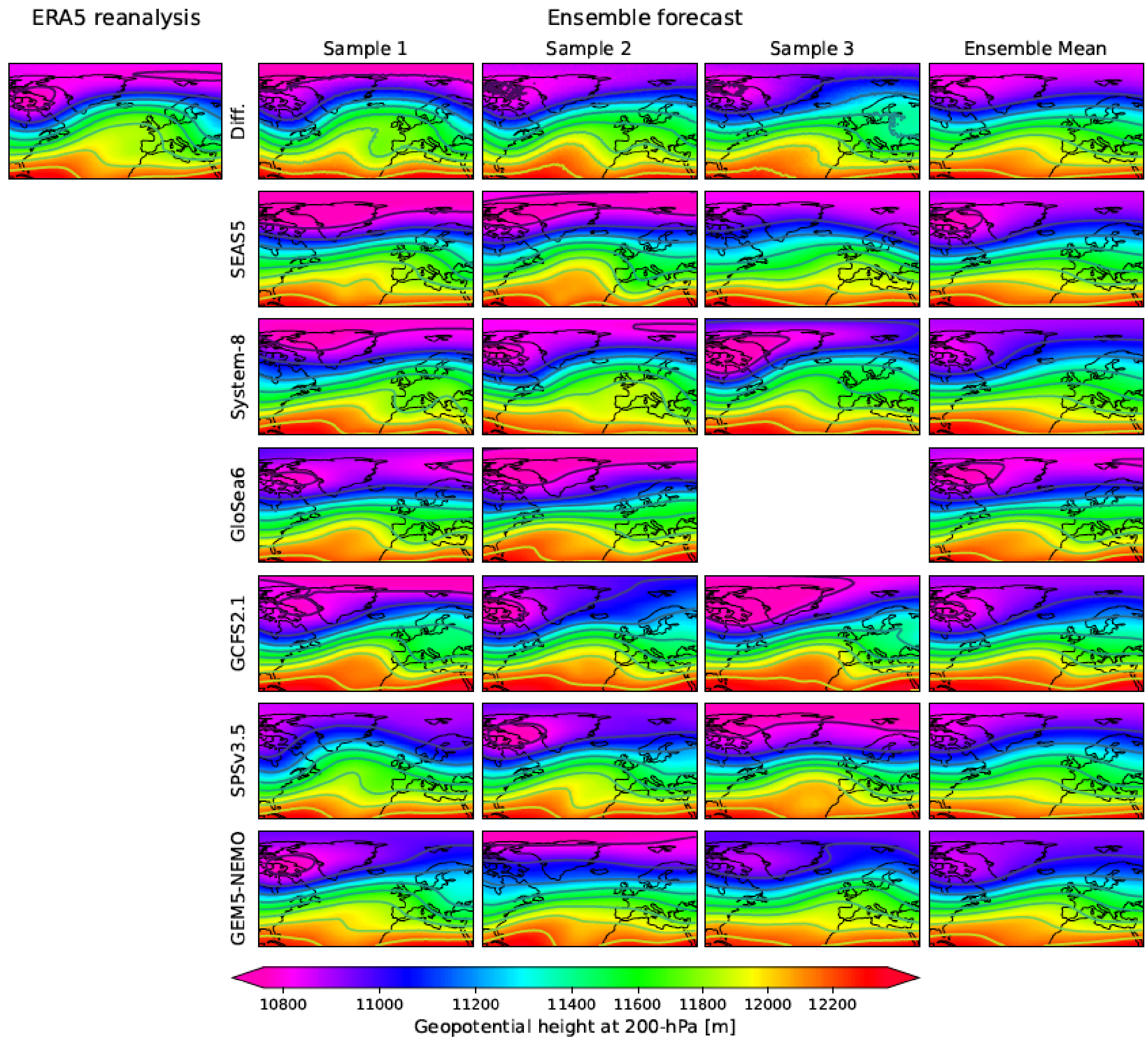

Figure 1. Comparison of snapshots of two-weeks averaged geopotential height at 200 hPa over North Atlantic and surrounding areas (90°W - 40°E, 20°N - 90°N) from January 15, 2022 to January 28, 2022. We display three selected ensembles from each forecasting system, along with the ensemble means, to compare ERA5 reanalysis (left column top row). Note that GloSea6 has only two ensemble members available, making the third column empty.

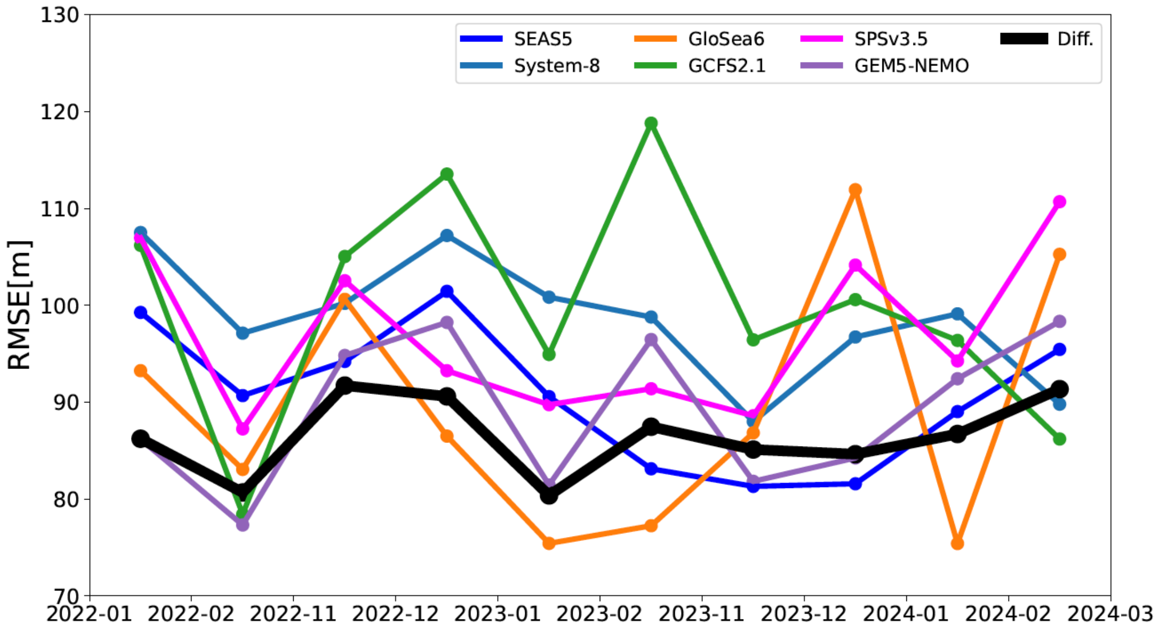

Figure 2. Median of the latitude-weighted root mean squared errors among ensemble members from 2022 January to 2024 February (winter months November to March) over the Northern Hemisphere (20°N-90°N). Our diffusion model consistently demonstrates good prediction performance.

T02

A Probabilistic Forecast for Multi-year ENSO Using Bayesian Convolutional Neural Network

P Sreeraj [1] , B Balaji [1,2] , Arya Paul [1] , and P A Francis [1]

[1] *(Presenting author underlined) Indian National Centre for Ocean Information Services (INCOIS), Ministry of Earth Science (MoES), Hyderabad, India*

[2] *Indian Institute of Tropical Meteorology, Ministry of Earth Science (MoES), Pune, India*

A robust El Niño Southern Oscillation (ENSO) prediction is essential for monitoring the global climate, regional monsoons, and weather extremes. Despite dedicated efforts spanning decades, the precise prediction of ENSO events through numerical modeling beyond a couple of seasonal lead times remains a daunting challenge. The advent of deep learning-based approaches marks a transformative era in climate and weather prediction. However, many machine learning-based studies attempting ENSO prediction are confined to singular estimates, lacking adequate quantification of uncertainty in learned parameters and overlooking the crucial need for a nuanced understanding of ENSO prediction confidence. Here, we introduce a deep learning-based Bayesian convolutional neural network model that provides robust probabilistic predictions for ENSO with a lead time of up to 9–10 months across all seasons. The Bayesian layers within the convolutional neural network maintain the capability to predict a distribution of learned parameters. Augmented with bias correction, our model reproduces the amplitude of the Niño 3.4 index with fidelity for lead up to 9–10 months. The inherent capacity for uncertainty modeling enhances the reliability of bayesian neural networks (BNNs), making them particularly valuable in operational services. This research holds substantial socio-economic implications as it enhances our forecasting capabilities and rigorously quantifies forecast uncertainties, providing valuable insights for planning and policymaking.

[1] P. Sreeraj, B. Balaji, A. Paul, and P. A. Francis, Environ. Res. Lett. **19**, 124023 (2024).

T03

From observations to realistic simulations, the mechanisms that develop and prevent convective organization

Alejandro Casallas [1], **Caroline Muller** [1], **and Adrian Tompkins** [2]

[1] *Institute of Science and Technology Austria*
[2] *Abdus Salam International Centre for Theoretical Physics*

Over 30 years ago, it was discovered that clouds cluster in idealized models, significantly influencing climate sensitivities and the hydrological cycle [1]. More recently, studies have shifted to exploring the consequences of convective organization in observational contexts. Research has shown that the spatial arrangement of convection affects extreme precipitation, the tropical radiative budget, and hydrological dynamics [2].

Convection in the Pacific warm pool (2°–9°N, 135°–145°E) clusters during boreal winter and spring when the region lies along the moist boundary of the warm pool. However, westward-propagating waves disrupt this clustering by altering wind patterns and moisture distribution, creating break-up episodes. In boreal summer and autumn, convection becomes random despite favorable diabatic feedbacks. This study investigates mechanisms driving these changes, focusing on the breakup episodes during boreal winter and spring [3,4].

Using satellite observations and reanalysis data, we identified variables influencing clustering and disorganization [3,4]. WRF model simulations tested the roles of moisture advection and wind patterns under scenarios isolating advection, moisture alone, and winds alone. Results showed that moisture advection drives organization-disorganization cycle, with the moisture alone being critical for convective organization, while low-level winds, especially meridional winds, primarily drive the disbandment of clusters.

To better understand the role of the meridional wind, we applied a sensitivity analysis using a Random Forest algorithm to isolate key variables influencing clustering, because modifying specific levels in WRF simulations can cause momentum imbalances. The analysis revealed that high-level winds and vertical shear have minimal impact on convective clustering. Instead, low-level and surface winds emerged as critical factors. Southerly flows were found to import moisture from the southern boundary, shifting convection northward and promoting random convection. Conversely, weak northerly flows retained moisture and convection in the southern region, supporting clustered convection.

These findings emphasize that while large-scale humidity is essential for clustering, it cannot entirely prevent breakup episodes. Southerly winds play a dominant role, driving moisture redistribution and convection disbandment. Surprisingly, zonal winds had little effect on clustering or disorganization. This study highlights the interplay between large-scale moisture, wind dynamics, and convection, advancing our understanding of the factors controlling organized and random convection in the Pacific warm pool.

[1] Held, I. M., R. S. Hemler, and V. Ramaswamy, 1993: Radiative-Convective Equilibrium with Explicit Two-Dimensional Moist Convection. *J. Atmos. Sci.*, **50**, 3909–3927

[2] Muller C, Yang D, Craig G, Cronin T, Fildier B, Haerter JO, Hohenegger C, Mapes B, Randall D, Shamekh S, Sherwood SC (2022) Spontaneous aggregation of convective storms. Annu Rev Fluid Mech 54(1):133–157.

[3] Tompkins, A. M., Casallas, A. and De Vera, M. (2025) Drivers of mesoscale convective aggregation and spatial humidity variability in the tropical western Pacific. npj Climate and Atmospheric Sciences, 8, 69.

[4] Casallas, A. (2024) Understanding Convective Organization and Seeking it Through Observations. Ph.D. thesis, Università degli Studi di Trieste. https://hdl.handle.net/11368/3076620.

T04

A Machine Learning-Based Approach to Quantify ENSO Sources of Predictability

Ioana Colfescu, Hannah Christensen, David John Gagne

A machine learning method is used to identify sources of long-term ENSO predictability in the ocean (sea surface temperature (SST) and heat content) and the atmosphere (near-surface zonal wind (U10)). Tropical SST represents the primary source of predictability skill. While U10 does not increase the skill when associated with SST, our analysis suggests U10 alone has apredictive skill comparable to that of SST between 11 and 21 months in advance, from late fall up to late spring. The long-lead signal originates from coupled wind-SST interactions across the Indian Ocean (IO) and propagates across the Pacific via an atmospheric bridge mechanism. A linear correlation analysis supports this mechanism, suggesting a precursor link between anomalies in SST in the western and wind in the eastern IO. Our results have important implications for ENSO predictions beyond 1 year ahead and identify the key role of U10 over the IO.

T05

Investigating Extreme Precipitation Event Attribution Through Deep-Learning Approaches

Jeremy M. Corner [1] , **Antonios Mamalakis** [1] , **and Kathleen A. Schiro** [1]

[1] *University of Virginia, Department of Environmental Science, Charlottesville, VA, USA*

Abstract

Understanding the drivers and variability of extreme precipitation events remains a significant challenge in climate science, but deep-learning models offer a new and promising potential. Deep-learning models are particularly useful in attributing climate drivers to atmospheric phenomena, especially when the underlying relationships are *nonlinear*. Herein, we examine the capability of deep learning models to capture maximum precipitation accumulations over the Southeast US during the extended Winter (i.e., November–March). The model inputs are standardized anomalies of column-water vapor, 500-hPa geopotential heights, and surface temperature with zero lead time, retrieved from the second version of Community Earth System Model Large Ensemble. Three distinct training approaches with convolutional neural networks (CNNs) are systematically evaluated against a baseline linear model: 1) a traditional approach with 14 members used for training data and 2 members of validation and testing data each, 2) an approach that utilizes customized loss functions that pair various statistics (e.g., absolute error, Anderson-Darling, etc.) with mean square error to more efficiently capture extremes , and 3) a "transfer learning" approach that involves initial broad training of the CNNs, followed by fine-tuning them to predict high-percentile precipitation events.

The top CNNs performance is achieved when a costumed loss is used, corresponding to an R^2 value of 0.86 on an independent testing data, outperforming the linear baseline model. This provides robust evidence for nonlinear dynamics at play driving precipitation that are captured by the CNNs. Furthermore, the CNNs can better temporally capture extreme precipitation events relative to the linear models. We then "interrogate" the best-performing CNN by implementing explainable artificial intelligence (XAI) techniques. XAI techniques increase model interpretability and provide important context about which input variables/patterns are most relevant to the deep-learning model output. The investigation of the CNN allows us to highlight key features that drive extreme events. Lastly, by comparing the patterns that the linear model is using to the ones used by the CNN, we isolate the critical nonlinear relationships of extreme precipitation events.

T06

AI-Driven reconstruction of oxygen profiles in the North Atlantic Oxygen Minimum Zone

Domitille Coron[1], **Etienne Pauthenet** [1], **Florian Sévellec** [1], **and Esther Portela** [1]

[1] *Univ Brest, CNRS, IRD, Ifremer, Laboratoire d'Océanographie Physique et Spatiale, Plouzané, France*

The oxygen content in the ocean has strong impacts on the sustainability of marine ecosystems [1]. In eastern tropical oceans, the combination of weak ventilation and enhanced oxygen consumption due to the proximity of productive neighboring upwelling areas creates the conditions for Oxygen Minimum Zone (OMZ) to exist [2, 3]. My research focuses on the study of the North Atlantic OMZ (NAOMZ). This OMZ is characterized by two oxygen minima: a deep minimum around 400-m depth and a shallow minimum around 100-m depth [2, 4]. Particularly, we will investigate the regional, and long-term oxygen variability, as well as its drivers, and its coupling with the West African upwelling system. The end-goal is to be able to analyze the contribution of different physical drivers of sub-regional variations of dissolved oxygen on seasonal and long-term time scales. In particular, we are interested in distinguishing anthropogenic and natural signals in the long-term evolution of the North Atlantic OMZ. However, analyzing multidecadal oxygen variability on a sub-regional basis has been hindered due to lack of data with adequate spatial-temporal cover. Furthermore, numerical models used in these studies struggle to reproduce the observed oxygen structure and its changes in OMZs areas [3, 6]. To circumvent this issue, as the first step or my PhD research, we develop a reconstruction of oxygen profiles based on all available *in situ* measurements of oxygen, temperature and salinity data from Argo, bottle and CTD profiles, in the tropical North-Atlantic ocean. All profiles are linearly interpolated on regular pressure levels and span the period 1906-2025. Using a Multi Layer Perceptron model (MLP), missing oxygen data are reconstructed from temperature and salinity data. The result is assessed by comparing the profiles obtained with profiles from the database (for an unused test period), with a particular focus on the shallow and deep minima prediction throughout the years. In order to carry out this comparison, a detection method is developed for profiles with a shallow minimum, or with both a shallow and a deep minima. It is based in part on criteria of oxygen saturation and the presence of an oxygen minimum around critical depths.

[1] L. Stramma, E. D. Prince et al, Nature Climate Change. **2**, 33-37 (2012).
[2] L. Stramma, P. Brandt et al, Journal of Geophysical Research: Oceans **17**, C4 (2008).
[3] J. Hahn, P. Brandt et al, Ocean Science **13**, 551-576 (2017).
[4] P. Brandt, H. W. Bange et al, Biogeosciences **12**, 489-512 (2015).
[5] L. Stramma, V. Visbeck et al, Geophysical Research Letters **36**, 20 (2009).
[6] L. Stramma, A. Oschilies et al, Biogeosciences **9**, 4045-4057 (2012).

T07

Enhancing Seasonal Predictions with Machine Learning: A Global Perspective on SST Influence in Early Winter

Víctor Galván Fraile[1], Marta Martín-Rey [1], Irene Polo [1] and Belén Rodríguez-Fonseca[1,2]

[1] *University Complutense of Madrid (UCM), Madrid, Spain.*
[2] *Institute of Geosciences (IGEO-CSIC), Madrid, Spain.*

The seasonal predictability of atmospheric patterns during early winter (November–December) is strongly influenced by anomalous sea surface temperature (SST) conditions. These oceanic anomalies play a central role in shaping winter atmospheric variability at a global scale, although their influence differs depending on the region and forecast lead time. In the extratropics, particularly, forecasting remains challenging due to the complexity of atmospheric dynamics and the interplay of various sources of predictability. In regions like the North Atlantic, the SST anomalies have been shown to significantly affect early winter atmospheric behaviour. However, current seasonal forecast systems—largely dependent on the El Niño–Southern Oscillation (ENSO)—tend to exhibit substantial biases in extratropical SSTs. These biases hinder the predictive skill of key climate variables in areas such as the Euro-Atlantic region (EAR), underscoring the need for alternative modelling strategies.

This work explores the ability of global SST anomalies to anticipate November–December sea level pressure (SLP) anomalies, using lead times from 1 up to 10 months. We apply three statistical techniques: Maximum Covariance Analysis (MCA) to detect major modes of co-variability between SSTs and atmospheric fields; a neural network (NN) approach tailored to capture non-linear teleconnection patterns; and a hybrid method that integrates both techniques.

Our findings reveal high predictability in several oceanic regions, depending on initialization timing. By contrasting traditional linear (MCA) with more complex non-linear methods (NN and hybrid), we provide a comprehensive assessment of early winter predictability. Notably, the NN-based models demonstrate substantial skill in the EAR, particularly at long lead times (7–10 months). The study also examines the temporal variability of these teleconnections over the 1940–2019 period, identifying non-stationary behaviour and highlighting specific time windows with enhanced forecast potential. These insights contribute to a better understanding of ocean-driven atmospheric teleconnections and offer valuable guidance for the development of more accurate seasonal prediction models—crucial for managing climate-related risks across the globe.

T08

Towards a Physics-Informed Neural Network for Ocean Circulation Modelling

6th Summer School on Theory, Mechanisms and Hierarchical Modelling of Climate Dynamics: Artificial Intelligence and Climate Modelling

Iuri Gorenstein [1] , **Ilana Wainer** [1]

[1] *Laboratório de Oceanografia Física, Clima e Criosfera (OC2) - Instituto Oceanográfico USP - São Paulo, Brasil.*

Numerical ocean models rely on discretizing primitive equations to simulate ocean dynamics, but their computational cost increases exponentially with resolution [1]. As an alternative, recent advances in neural networks (NNs) have demonstrated promising applications in climate science, offering orders-of-magnitude faster simulations than traditional models through data-driven algorithms that can be continuously improved with observational data training [2,3]. However, NN-based approaches for ocean modeling face unique critical challenges, including stability, physical consistency, and generalizability [4]. This study presents a novel approach to embedding the barotropic dynamics of the Coastal and Regional Ocean COmmunity (CROCO) model (a.k.a the AB3-AM4 algorithm) [5,6] into a physics-informed neural network. By leveraging a variational Koopman autoencoder, we transform the nonlinear barotropic dynamics into a latent space where they can be evolved linearly. The network is trained on an Equatorial Rossby Wave simulation, learning to represent barotropic free-surface height and momentum dynamics. We test the trained model on a barotropic tsunami simulation in a previously unseen domain and initial conditions to assess its generalizability. The results demonstrate a stable free-surface height evolution, preserving key physical structures despite the domain shift. Future developments will incorporate parameterizations of subgrid-scale processes and bulk forcing mechanisms, enabling a fully generalizable, data-driven ocean model. This approach bridges the gap between traditional numerical models and NN algorithms, paving the way for hybrid modeling techniques that improve ocean simulation beyond computational limitations.

[1] SOUFFLET, Yves et al. On effective resolution in ocean models. **Ocean Modelling**, v. 98, p. 36-50, 2016.

[2] HAYKIN, Simon S. **Neural networks and learning machines: International Version**. Pearson Education (US), 2009.

[3] RUSSKOVA, Tatiana. Performance analysis of a parallel Monte Carlo code for simulating solar radiative transfer in cloudy atmospheres using CUDA-enabled NVIDIA GPU. In: **23rd International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics**. SPIE, 2017. p. 311-321.

[4] SONNEWALD, Maike et al. Bridging observations, theory and numerical simulation of the ocean using machine learning. **Environmental Research Letters**, v. 16, n. 7, p. 073008, 2021.

[5] Francis Auclair, et al. Coastal and Regional Ocean Community Model. 2.0.1, **Zenodo**, 7 Oct. 2024, doi:10.5281/zenodo.13898339.

[6] SHCHEPETKIN, Alexander F.; MCWILLIAMS, James C. The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. **Ocean modelling**, v. 9, n. 4, p. 347-404, 2005.

T09

Learning the Ocean: Leveraging AI to Improve Ocean Data Assimilation

Davide Grande[1,2,3], **Andrea Storto**[2], **and Roberto Buizza**[3]

[1]*(Presenting author underlined) Scuola Universitaria Superiore IUSS, Pavia, Italy*
[2]*CNR-ISMAR, Rome, Italy*
[2] *Scuola Superiore Sant'Anna, Pisa, Italy*

The assimilation of ocean observations into numerical models remains a central challenge in Earth system science. Ocean data assimilation (DA) must cope with sparse and irregular observations, complex boundary conditions, and nonlinear multiscale dynamics—features that often limit the effectiveness of traditional variational and ensemble-based methods.
My PhD project explores the potential of machine learning (ML) to complement and enhance ocean DA frameworks. As a first step, I am conducting a systematic review of the recent literature (2020–2025), aimed at identifying how ML techniques are being used to address eight key challenges that are particularly critical in the ocean context. These include observation integration, boundary treatment, sparse data reconstruction, uncertainty quantification, and the embedding of physical constraints in data-driven schemes.
The review organizes and evaluates a broad selection of recent studies covering diverse ML approaches -such as neural surrogates, latent-space DA, physics-informed networks, and hybrid coupling architectures— and assesses their level of maturity, generalization capabilities, and potential applicability to both regional and large-scale ocean prediction systems.
This work lays the foundation for the second phase of the PhD, where selected ML-based components will be implemented and tested within tailored numerical experiments, in order to evaluate their contribution to improving the accuracy, consistency, and efficiency of ocean data assimilation.
Ultimately, the project seeks to clarify when and how ML can be reliably integrated into ocean DA pipelines, and to contribute design guidelines for their future inclusion in coupled Earth System Models.

T10

Improving Global Wildfire Simulations Using Machine Learning

Florence Kayode[1], Christian Seiler[1], and David McLagan[1,2]

1 School of Environmental Studies, Queen's University

2 Dept. Geological Sciences & Geological Engineering, Queen's University

Wildfires significantly impact atmospheric carbon levels, with about 20% of global fire emissions being irreversible sources of CO2 that cannot be offset by forest regrowth or soil carbon recovery over decadal to centennial timescales[1]. The complex interplay between fire weather, fuels, and ignition complicates the differentiation between anthropogenic and climatic influences on fire dynamics both globally and regionally[2]. In the short term, rainfall during the dry season suppresses fire activity, while over longer periods, fuel accumulation during wet years in arid ecosystems can lead to increased burned areas in subsequent years. The CanESM land surface component, CLASSIC, captures global wildfire emissions well but underestimates the magnitude of wildfires in boreal regions[3,4]. Recent developments suggest that incorporating the Fire Weather Index (FWI) into the model's parameterization significantly improves performance. However, this approach has only been tested in limited domains, and incorporating FWI necessitates retuning parameter values; a process traditionally guided by experience, expert judgement, and intuition.

We propose using Genetic Algorithms (GA), a machine learning technique that mimics natural evolutionary processes to solve optimization problems. In this method, each parameter is encoded as a string of fixed binary length and assigned a fitness value related to an objective function. Unlike other optimization methods, GA searches for improvements using probabilistic transition rules, aiming to explore multiple peaks in parallel. We anticipate that GA will provide optimized parameter values for use in CLASSIC, enhancing the representation of wildfire emissions globally. This approach will also offer insights into the interactions between certain model parameters, demonstrating the effectiveness of machine learning algorithms in environmental research. While we expect significant improvements, careful consideration of preprocessing, parameter settings, and scalability is essential for successful implementation and accurate predictions.

Citations:

[1] Bo Zheng *et al.*, Increasing Forest fire emissions despite the decline in global burned area. *Sci. Adv.***7, eabh**2646(2021). DOI:10.1126/sciadv.abh2646
[2] Stijn Hantson, Douglas S. Hamilton, Chantelle Burton, Changing fire regimes: Ecosystem impacts in a shifting climate, One Earth, Volume 7, Issue 6, 2024, Pages 942-945, ISSN 2590-3322, https://doi.org/10.1016/j.oneear.2024.05.021.
[3] Melton, J. R., V. K. Arora, E. Wisernig-Cojoc, C. Seiler, M. Fortier, E. Chan, and L. Teckentrup (2020). Classic v1.0: the open-source community successor to the Canadian land surface scheme (class) and the Canadian terrestrial ecosystem model (CTEM) – part 1: Model framework and site-level performance. Geoscientific Model Development 13(6), 2825–2850.
[4] Seiler, C., J. R. Melton, V. K. Arora, and L. Wang (2021, May). CLASSIC v1.0: the open-source community successor to the Canadian land surface scheme (CLASS) and the Canadian terrestrial ecosystem model (CTEM) – part 2: Global benchmarking. Geoscientific Model Development 14(5), 2371–2417.

Title: Retrieving tropical cyclone intensity from climate output: AI approaches and challenges

Abstract: In this presentation, I will discuss the challenges and potential strategies for retrieving tropical cyclone (TC) intensity from gridded climate reanalysis data. Using two widely utilized deep learning architectures—convolutional neural networks (CNNs) and vision transformers (ViTs)—we will show that these models can estimate TC intensity with a mean absolute error of ±7.5 kts. Further analysis of model performance under different data sampling methods reveals however that TC intensity retrieval is highly sensitive to data sampling strategies, particularly for convolutional architectures. In contrast, the transformer-based approach shows greater resilience to these data sampling variations, making it a robust solution for TC intensity retrieval or TC climate downscaling applications in the future.

T12

RECONSTRUCTION OF THE SURFACE GEOSTROPHIC CURRENT IN THE MEDITERRANEAN SEA FROM ABSOLUTE DYNAMIC TOPOGRAPHY GENERATED BY GENERATIVE ADVERSARIAL NETWORK (ADTGAN)

M. G. Elysee[1], A. Apirro[2], E. Mauri[2], and M. Menna[2]

[1] *University of Trieste, Italy*
[2] *National Institute for Oceanography and Applied Geophysics (OGS), Italy*

Abstract

Sparse and noisy observational data introduce biases into model predictions, leading to inaccurate forecasts, reduced data reliability, and errors in decision-making [1], [2]. These challenges are critical in ocean dynamics and temperature modeling, which directly influence storm formation and other natural disasters [3]. Traditional methods, including numerical simulations and data assimilation techniques such as optimum interpolation, have been widely used by organizations like the Met Office and CMEMS [4], [5]. However, these methods often smooth out mesoscale features, impacting weather prediction accuracy [6], [7].

To address these limitations, we propose ADTGAN (Absolute Dynamic Topography Generative Adversarial Network), a model designed to reduce data sparsity and noise by generating realistic missing data and denoising uncertain observations [8], [9]. ADTGAN achieves lower RMSE, reduced correlation errors, and higher skill scores, demonstrating its robustness in improving oceanographic data reliability and weather forecasting accuracy [10].

References

[1] J. Baxter, L. Smith, and R. Davies, "The challenge of sustaining ocean observations," *Front. Mar. Sci.*, vol. 6, p. 105, 2019.

[2] X. Liu, Y. Zhang, and H. Wang, "Big data challenges in ocean observation: A survey," *Res. Gate*, 2015.

[3] L. Gregor, A. Fay, and C. Schmid, "Spatiotemporal upscaling of sparse air-sea pCO_2 data via physics," *Nat. Commun.*, vol. 15, p. 97, 2024.

[4] J. A. Cummings and O. M. Smedstad, "An optimal data assimilation method and its application to the ocean," *J. Mar. Syst.*, vol. 165, pp. 133-142, 2017.

[5] P. Y. Le Traon, M. J. Bell, and B. Hackett, "A global-ocean-data assimilation for operational oceanography," *J. Mar. Sci. Eng.*, vol. 11, no. 12, p. 2255, 2023.

6] M. Beauchamp, P. Roux, and C. Ouellet, "OceanBench: The sea surface height edition," *arXiv preprint*, arXiv:2309.15599, 2023.

[7] Y. Fang, X. Lin, and J. Gao, "Three-dimensional ocean sound speed field reconstruction using tensor neural networks," *arXiv preprint*, arXiv:2308.04930, 2023.

[8] W. Sun, B. Tan, and D. Xu, "Direct simulations of wind-driven breaking ocean waves with data assimilation," *arXiv preprint*, arXiv:1410.0418, 2014.

[9] J. Wang and L. Zhang, "Investigating ocean circulation dynamics through data assimilation," *arXiv preprint*, arXiv:2404.07134, 2024.

[10] S. Levitus, J. Antonov, and T. Boyer, "Ocean reanalysis," *Wikipedia*, 2022.

T13

Climate projections in the Euro-Mediterranean region: atmospheric circulation patterns and temperature and rainfall future changes

Matias Olmo[1] , **Cos P.**[1], **Campos D.**[1], **Soret A.**[1], **Doblas-Reyes F.**[1]

[1]*Barcelona Supercomputing Center, Barcelona, Spain*

The Western Mediterranean (WMed) has been pointed out as a hotspot region for both warming and drying signals. However, there is still large uncertainty in future projections due to model uncertainty and misrepresentation of specific processes. Thus, the need for a better understanding of the future climate of the WMed becomes evident. Improved climate indicators for decision-making can benefit from a deeper insight on future climate extremes and their related atmospheric circulation, taking into account the spread in model performances over the WMed region.

The present work is based on the analysis of future projections of rainfall and temperature extremes from a set of CMIP6 global climate models (GCMs) during 2070-2100, according to their representation of a classification of atmospheric circulation patterns (CPs). CPs are defined using daily mean sea level pressure (SLP) using hierarchical clustering and data reduction through empirical orthogonal functions. The ERA5 reanalysis during 1950-2022 was considered as the reference to evaluate the historical GCMs simulations, constructing the CPs with SLP and analyzing their link to surface variables including precipitation, maximum and minimum temperatures. To assess the future atmospheric circulation, the clustering algorithm is replicated and future CPs are compared to the historical ones in terms of frequency, shape and intensity changes.

Based on the historical CPs, a model ranking is generated using a combination of spatial and temporal reproduction metrics for the SLP patterns and the associated surface conditions. GCMs manage to reproduce the annual cycle of the CPs frequency, with a dominant summer CP enhancing warm and dry conditions. However, the correct timing of this pattern and the transitional CPs still need to be more accurate. The analysis of the associated surface patterns shows good model performance, better for temperature than for rainfall, particularly in the transition seasons, for which the GCMs spread in their skill score becomes larger. This process-based evaluation leads to a performance ranking that is used for model subselection based on cost function minimization considering model performance, spread and independence.

In terms of climate extremes, the uncertainty in future projections of the indices —including the expected increases in the frequency of warm days and dry spells— can be quantified by selecting specific subsets of GCMs, according to the process-based ranking. In particular, the warming and drying signals over areas such as the northeastern Iberian Peninsula are clearer in the best-performing GCMs ensemble. This constraining procedure shows more clear results in summer than in winter, when natural variability has a larger role in modulating the WMed changing climate.

These changes in temperature and rainfall were associated with a changing frequency of the CPs driving the specific extremes. CPs present some differences in their seasonal distribution for the late 21st century compared to their historical records, while the centroids of the CPs often present changes, evidencing modifications in the intra-pattern variability.

T14

On the application of generative modelling for seasonal climate predictions

<u>Lluís Palma</u>[1,2], **Alejandro Peraza**[1], **Amanda Duarte**[1], **Albert Soret**[1] **and Markus Donat**[1,3]

[1]*Barcelona Supercomputing Center, Earth Sciences Department, Spain*
[2]*Facultat de Física, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona*
[3] *Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain*

Reliable probabilistic predictions at the seasonal time scale are critical for key societal sectors such as agriculture, energy, and water management. Current operational approaches face significant challenges: General Circulation Models (GCMs) are computationally expensive and often limited by low spatial resolution and model deficiencies. At the same time, traditional statistical methods struggle due to big modelling assumptions (i.e. linearity or Homoscedasticity) and short historical records. Generative models emerge as a cost-effective, promising alternative, offering the potential to model complex nonlinear climate dynamics inherently probabilistically and at a reduced computational cost.

Thus, the present study evaluates the effectiveness of different generative methodologies in predicting global or regional gridded fields of temperature and rainfall seasonal anomalies. The predictions cover all four seasons and are initialised one month before the start of the season, aligning with most climate services providers. We employ climate model output from CMIP6 and CEMS-lens2 during training and ERA5 reanalysis data during testing to circumvent the short span of current reanalysis and observational datasets. We analyse the method's performance in predicting interannual anomalies beyond the climate change-induced trend. Additionally, we test the proposed methodology in a regional context with a use case focused on Europe. We show that the model's ensemble generation capabilities allow it to provide diverse ensemble members, allowing the derivation of relevant probabilistic information and potentially reliable predictions. While climate change trends dominate the skill of temperature predictions, additional skill over the climatological forecast in regions influenced by known teleconnections is found. We reach similar conclusions based on the validation of precipitation predictions.

This work further demonstrates the effectiveness of training generative models on climate model output for seasonal predictions, providing skilful seasonal climate predictions beyond the induced climate change trend at time scales and lead times relevant for user applications, motivating further research.

[1] Palma, Ll., Peraza, A., Civantos, D., Duarte, A., Materia, S., Muñoz, Á. G., Peña, J., Romero, L., Soret, A., Donat, M. G. (2025). Data-driven seasonal climate predictions via variational inference and transformers. In arXiv. http://arxiv.org/abs/2503.20466 [In review]

T15

Data-Driven Equation Discovery of In-Cloud Vertical Velocity

Alzbeta Pechacova[1], **Alejandro Casallas**[1], **Lokahith Agasthya**[1] **and Caroline Muller**[1]

[1]*Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria*

Vertical velocities of updrafts within clouds are known to directly impact the precipitation rate and lightning frequency while playing a central role in other cloud processes [7]. However, they remain poorly understood. Existing theories to predict in-cloud vertical velocity involve a variety of plume/thermal models (e.g. [3, 7]) and the terminal velocity of raindrops (e.g. [6]) to develop and apply scaling hypotheses to cloud-resolving simulations.

In this project, we propose an alternate method using equation discovery to better understand what determines the vertical velocity (w) within an individual cloud. We run idealized simulations using a cloud-resolving model (SAM) [4], and subsequently adapt and apply a tracking algorithm from [2] to follow clouds as Lagrangian objects and obtain data per cloud. Using SINDy [1], we then acquire equations for w and its derivatives, and compare these to existing theories. To avoid including unnecessary non-informative terms and bias from correlated variables, we perform conditional independence testing [5].

Initial results indicate that the process above can extract an equation for w with a high coefficient of determination ($R^2 \approx 0.8$). The variables with high predictive power include the cloud condensate mixing ratio as well as the terminal velocity, and improvements are also seen when including the temperature anomaly and relative humidity. While promising, several steps remain to tune the data selection as well as the machine learning method, to correctly evaluate and interpret the results.

1. S. L. Brunton, J. L. Proctor, J. N. Kutz, *Proceedings of the National Academy of Sciences* **113**, 3932–3937 (2016).
2. A. Casallas, A. M. Tompkins, C. Muller, G. Thompson, *Journal of Advances in Modeling Earth Systems* **17** (2025).
3. D. Hernandez-Deckers, T. Matsui, A. M. Fridlind, *Atmospheric Chemistry and Physics* **22**, 711–724 (2022).
4. M. F. Khairoutdinov, D. A. Randall, *Journal of the Atmospheric Sciences* **60**, 607 –625 (2003).
5. L. Kook, A. R. Lundborg, *Briefings in Bioinformatics* **25** (2024).
6. A. Parodi, K. Emanuel, *Journal of the Atmospheric Sciences* **66**, 3449–3463 (2009).
7. M. S. Singh, P. A. O'Gorman, *Quarterly Journal of the Royal Meteorological Society* **141**, 2828–2838 (2015).

T16

Exploring windows of opportunity at sub-seasonal time scales for extreme precipitation events: the Valencia DANA case study

Núria Pérez-Zanón[1]

[1] *Barcelona Supercomputing Center*

Accurately predicting extreme weather events at sub-seasonal timescales (2–4 weeks) remains a major challenge in climate services. Identifying windows of opportunity—periods when forecasts exhibit enhanced predictive skill—can significantly improve early warning systems and decision-making. This study focuses on a high-impact isolated upper-level low-pressure system (DANA) event that affected Valencia, Spain, in October 2024, analysing its predictability and potential for skilful forecasts within a sub-seasonal framework.

An exploratory analysis using the NCEP CFSv2 sub-seasonal climate forecast system [1] indicates that standard calibration [2][3] of total accumulated precipitation fails to highlight a clear signal of above-normal rainfall over the Iberian Peninsula for the week of 28 October to 3 November 2024. However, some ensemble members predicted intense precipitation up to four weeks in advance, while one- and two-week lead forecasts misrepresented the spatial distribution of the precipitation terciles, suggesting potential predictability despite misplaced signals.

To improve forecast reliability, we aim to further analyse key dynamical variables, including air temperature and geopotential height at 500 hPa, and sea level pressure, to establish a methodology for extreme precipitation event warnings. Additionally, we aim at exploring the potential of AI-based calibration methods to enhance predictive skill. Random Forest and Gaussian Process Regression could improve point-based bias correction, while Self-Organising Maps (SOMs) may help to refine spatial calibration. However, the choice of methodology depends on data availability and the inherent limitations of sub-seasonal climate forecasts for training AI models.

[1] Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H.-Y., Iredell, M., Ek, M., Meng, J., Yang, R., As, M., Mendez, P. A., Dool, H., Zhang, Q., Wang, W., Chen, M., & Becker, E. The NCEP Climate Forecast System Version 2. J. Clim., 27, 2185–2208. https://doi.org/10.1175/JCLI-D-12-00823.1 (2014).
[2] Doblas-Reyes, F. J., Hagedorn, R., & Palmer, T. N. The rationale behind the success of multi-model ensembles in seasonal forecasting – II. Calibration and combination. Tellus A: Dyn. Meteorol. Oceanogr., 57, 234–252, https://doi.org/10.3402/tellusa.v57i3.14658. (2005).
[3] Manzanas, R., Gutiérrez, J. M., Bhend, J., Hemri, S., Doblas-Reyes, F. J., Torralba, V., Penabad, E., & Brookshaw, A.. Bias adjustment and ensemble recalibration methods for seasonal forecasting: a comprehensive intercomparison using the C3S dataset. Clim. Dyn. https://doi.org/10.1007/s00382-019-04640-4 (2019).

T17

6th Summer School on Theory, Mechanisms and Hierarchical Modelling of Climate Dynamics: Artificial Intelligence and Climate Modelling

Andrea Polesello[1], **Alejandro Casalla**[1], **Remy Roca** [2], **Francesco Locatello** [1], **and Caroline J. Muller** [1]

[1] *Institute of Science and Technology Austria, Klosterneuburg, Austria*
[2] *Laboratoire d'Etudes en Geophysique et Oceanographie Spatiales, Toulouse, France*

Deep convective systems (DCSs) play a crucial role in the tropical hydrological cycle and radiative budget (1; 2). In particular, the largest and longest-lived of those cloud systems contribute to a high fraction of the extreme precipitation in the Tropics (3). Therefore understanding what drives these types of systems is crucial.

To that end, Abramian et al. (4) developed a new method to predict the maximum area of DCSs using the DYAMOND-Summer simulation with the cloud-resolving global model SAM, and the TOOCAN algorithm to track cloud systems. The method uses simple machine learning models, trained on information on the early stage of the systems and their surrounding environment, including dynamical and thermodynamical variables, morphological features of the systems and the characteristics of their neighbors. We first extended this method to other datasets: in particular, we used the DCS tracks identified by TOOCAN in satellite data (5), combined with ERA5 data for the physical variables, and also the TOOCAN tracks in the SAM run of the RCEMIP project (SST=300 K) (6). We used different neural network models, achieving good prediction accuracy even using just information from the first two hours of life of the DCSs. The results are also consistent across the three datasets, showing the universality of the prediction power of such a method.

We are currently improving this method by upgrading our machine learning models to more interpretable ones, such as the recently invented B-cos networks (7), that allow the user to readily understand which features contributed the most to the prediction, without the need of running any offline calculation after the training process.
The final stage of the investigation would be to run causal discovery and causal inference algorithms, to identify causal relations between the predictors and the maximum area, to be then able to identify physical drivers of the DCS evolution, distinguishing them from other variables that might only be stronly correlated with the maximum area, without a causal link.

[1] Stephens, G., and Coauthors, 2023: The First 30 Years of GEWEX. Bull. Amer. Meteor. Soc., 104, E126–E157, https://doi.org/10.1175/BAMS-D-22-0061.1.

[2] Roca, R., J. Aublanc, P. Chambon, T. Fiolleau, and N. Viltard, 2014: Robust Observational Quantification of the Contribution of Mesoscale Convective Systems to Rainfall in the Tropics. J. Climate, 27, 4952–4958,.

[3] Roca, R., Fiolleau, T. Extreme precipitation in the tropics is closely associated with long-lived convective systems. Commun Earth Environ 1, 18 (2020). https://doi.org/10.1038/s43247-020-00015-4

[4] Sophie Abramian. Physical origins of the properties of mesoscale convective systems and implications for high impact events. Climatology. Université Paris sciences et lettres, 2023. English. ⟨NNT : 2023UPSLE012⟩. ⟨tel-04809470⟩

[5] Fiolleau, T. and Roca, R.: A database of deep convective systems derived from the intercalibrated meteorological geostationary satellite fleet and the TOOCAN algorithm (2012–2020), Earth Syst. Sci. Data, 16, 4021–4050, https://doi.org/10.5194/essd-16-4021-2024, 2024

[6] Wing, A. A., Reed, K. A., Satoh, M., Stevens, B., Bony, S., and Ohno, T.: Radiative–convective equilibrium model intercomparison project, Geosci. Model Dev., 11, 793–813, https://doi.org/10.5194/gmd-11-793-2018, 2018

[7] M. Böhle, N. Singh, M. Fritz and B. Schiele, ”B-Cos Alignment for Inherently Interpretable CNNs and Vision Transformers,” in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 6, pp. 4504-4518, June 2024, doi: 10.1109/TPAMI.2024.3355155.,

T18

CNN-based forecasting of early winter NAO using sea surface temperature

E. Provenzano[1], **G. Gastineau**[1], **C. Mejia**[1], **D. Swingedouw**[2], **and S. Thiria**[1]

[1]*UMR LOCEAN (IPSL), Sorbonne University, Paris, France*
[2]*UMR EPOC, University of Bordeaux, France*

March 2025

The North Atlantic Oscillation (NAO) is a primary mode of atmospheric variability over the North Atlantic, characterized by a dipole pattern of sea-level pressure anomalies that significantly influence temperature and precipitation anomalies over Europe [1]. While the NAO's impact on North Atlantic sea surface temperature (SST) is well understood, the atmosphere's response to SST is weak and nonlinear.
During early winter (ND), El Niño and La Niña events influence NAO behaviour, with differing effects in late winter (JF) [2]. Indian Ocean SSTs, as well as the North Atlantic Horseshoe SST anomaly, have also been identified as important drivers of early winter NAO variability [3][4].
Climate models used for seasonal forecasting, however, often struggle to capture the NAO response to SSTs, particularly during early winter, where they typically exhibit low signal-to-noise ratios [5]. To address this gap, this study adopts a statistical modelling approach, employing convolutional neural networks (CNNs) to predict the NAO in early winter using 1-month leading SST fields as predictors. This approach aims to enhance seasonal forecasting capabilities and advance our understanding of ocean-atmosphere interactions.

ERA5 Hourly Data reanalysis (1940-2023) was used and re-gridded to a 2°×2° grid for the SLP (20°N–70°N, 100°W–40°E) and to a 1°×1° grid for the SST (15°S–70°N) using bilinear interpolation. Running means of 10 and 30 days were applied to the SLP and SST fields, respectively. The NAO index is calculated with the first principal component from the Principal Component Analysis of (ND) SLP anomalies, which accounts for 22% of the total variance.
The CNN architecture consists of three convolutional layers, totalling 261.745 trainable parameters. A linear regression model, consisting of 30.601 parameters, is used as a benchmark.
The dataset was divided into training (1940-2012) and test (2013-2023) periods. Following model training, the CNN achieved superior results compared to the baseline model, with a normalized RMSE of 0.97 and a correlation coefficient of 0.24, significant at the 95% confidence level. In contrast, the baseline model had a nRMSE of 1.03 and a correlation of 0.02.
To better understand the CNN's predictive insights, Grad-CAM analysis was employed. This technique revealed that the model effectively focused on key SST regions, particularly the ENSO region, North Atlantic, and Indian Ocean, aligning with established physical connections in NAO variability.

This study highlights the potential of CNNs to enhance seasonal climate forecasting by capturing nonlinear relationships often missed by traditional models, while emphasizing the need for further refinement through additional predictors such as stratospheric conditions and sea ice concentration [6].

[1] A. A. Scaife *et al.*, J. Climate **21**, 72 (2008).
[2] W. Zhang and F. Jiang, Climate **11**, 47 (2023).
[3] G. Gastineau and C. Frankignoul, J. Climate **28**, 1396 (2015).
[4] S. C. Hardiman *et al.*, Atmos. Sci. Lett. **21**, e1005 (2020).
[5] F. Molteni and A. Brookshaw, Clim. Dyn. **61**, 2673 (2023).
[6] R. Hall *et al.*, Int. J. Climatol. **35**, 1697 (2015).

T19

Abstract template for SMR4067 – Contributed talk
Fine-tuning a global weather model for superior subseasonal forecasting

V.Sansine[1], T.Izumo[1], Marania Hopuare[2], Damien Specq[3], Sophie Martinoni-Lapierre[3]

[1]French National Institute of Research for Sustainable Development, UMR-SECOPOL
[2]Gepasud Laboratory – University of French Polynesia,
[3]Météo-France, Toulouse, France.

Accurate subseasonal forecasting is socio-economically critical [1, 2] yet remains a great scientific challenge [3, 4, 5, 6]. Recent advances in machine-learning based global weather forecasting demonstrate superior skill on medium-range [7, 8, 9, 10] (1 to 15 days ahead) and subseasonal-range [11] (15 to 42 days ahead) than the best traditional weather forecasting system. These data-driven models require immense computational resources for training, which are not widely available. Here we show, by using medium-range Graphcast model as pre-trained model and focusing on reducing iterative error accumulation, that fine-tuning is an efficient strategy to achieve improved results for subseasonal forecasting. Our fine-tuned model GraphFT rapidly converges (trained on just three years of data), and significantly outperforms Graphcast and the leading deterministic traditional subseasonal forecasting system at 3-4 week leads [12]. Demonstrating the potential of fine-tuning for improving, with low computational costs, possibly both atmosphere and ocean forecasts at long leads.

[1] J. H. Richter *et al.*, « Quantifying sources of subseasonal prediction skill in CESM2 », *Npj Clim. Atmospheric Sci.*, vol. 7, nº 1, p. 1-9, mars 2024, doi: 10.1038/s41612-024-00595-4.

[2] E. Coughlan de Perez *et al.*, « Action-based flood forecasting for triggering humanitarian action », *Hydrol. Earth Syst. Sci.*, vol. 20, nº 9, p. 3549-3560, sept. 2016, doi: 10.5194/hess-20-3549-2016.

[3] C. J. White *et al.*, « Potential applications of subseasonal-to-seasonal (S2S) predictions », *Meteorol. Appl.*, vol. 24, nº 3, p. 315-325, 2017, doi: 10.1002/met.1654.

[4] M. Chen, W. Wang, et A. Kumar, « Prediction of Monthly-Mean Temperature: The Roles of Atmospheric and Land Initial Conditions and Sea Surface Temperature », févr. 2010, doi: 10.1175/2009JCLI3090.1.

[5] F. Doblas, J. García-Serrano, F. Lienert, A. Biescas, et L. Rodrigues, « Seasonal climate predictability and forecasting: Status and prospects », *Wiley Interdiscip. Rev. Clim. Change*, vol. 4, juill. 2013, doi: 10.1002/wcc.217.

[6] F. Vitart, A. Robertson, et D. Anderson, « Subseasonal to Seasonal Prediction Project: Bridging the gap between weather and climate », *WMO Bull.*, vol. 61, janv. 2012.

[7] R. Lam *et al.*, « Learning skillful medium-range global weather forecasting », *Science*, vol. 382, nº 6677, p. 1416-1421, déc. 2023, doi: 10.1126/science.adi2336.

[8] S. Rasp *et al.*, « WeatherBench 2: A Benchmark for the Next Generation of Data-Driven Global Weather Models », *J. Adv. Model. Earth Syst.*, vol. 16, nº 6, p. e2023MS004019, 2024, doi: 10.1029/2023MS004019.

[9] K. Bi, L. Xie, H. Zhang, X. Chen, X. Gu, et Q. Tian, « Accurate medium-range global weather forecasting with 3D neural networks », *Nature*, vol. 619, nº 7970, p. 533-538, juill. 2023, doi: 10.1038/s41586-023-06185-3.

[10] J. Pathak *et al.*, « FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators », 22 février 2022, *arXiv*: arXiv:2202.11214. Consulté le: 25 juillet 2024. [En ligne]. Disponible sur: http://arxiv.org/abs/2202.11214

[11] L. Chen *et al.*, « A machine learning model that outperforms conventional global subseasonal forecast models », *Nat. Commun.*, vol. 15, nº 1, p. 6425, juill. 2024, doi: 10.1038/s41467-024-50714-1.

[12] V. Sansine, T. Izumo, M. Hopuare, D. Specq, et S. M.-L. Pierre, « Fine-tuning a global weather model for superior subseasonal forecasting », 10 mars 2025, *Research Square*. doi: 10.21203/rs.3.rs-5619528/v1.

T20

Data-Informed Inversion Model (DIIM): a framework to retrieve marine optical constituents in the BOUSSOLE site using a three-stream irradiance model

Carlos Enmanuel Soto López[1,2], **Fabio Anselmi**[1], **Mirna Gharbi Dit Kacem**[1,2], **and Paolo Lazzari**[2]

[1]*Università degli Studi di Trieste, 31127, Italy*
[2]*Istituto nazionale di oceanografia e di geofisica sperimentale - OGS, Trieste, 34010, Italy*

In recent years, detailed underwater light propagation models have been used in a number of marine biogeochemical models to investigate biooptical interactions that regulate marine ecosystem dynamics. Our work focuses on the inversion of a three-stream light model [1], determining biogeochemical properties of the ocean color observed by satellites. Simultaneously, the three-stream light model, used as a forward model, enables the simulation of the ocean color by reconstructing the remote sensing reflectance according to the biogeochemical properties of seawater. A coherent inversion model, as presented here, can be used as an observation operator to assimilate the remote sensing reflectance.

Due to the nature of inversion problems, it is a challenge to optimize the parameters of such models, with the aim that the inversion procedure determines properties that are coherent with real-world measurements. For this reason, we explore a Bayesian approach, allowing the retrieval of quantities together with their respective uncertainty. Uncertainty estimation is of crucial importance in data assimilation procedures as well as in many other applications. As part of the approaches explored, with the attempt of finding a computationally optimal procedure, which allows for fast computations and efficient parameter calibration, we also explored a neural network based inversion method, using the so-called Stochastic Gradient Variational Bayes (SGVB) framework [2], the same framework used to train Variational Auto Encoders (VAE). We called the method that uses the SGVB framework for parameter calibration, and as a method of optimizing the inversion, Data Inform Inversion Method, or DIIM [3].

[1] Dutkiewicz, S., Hickman, A., Jahn, O., Gregg, W., Mouw, C., and Follows, M.: Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model, Biogeosciences, 12, 4447–4481, 2015.

[2] Kingma, D. P. and Welling, M.: Auto-encoding variational bayes, arXiv preprint arXiv:1312.6114, 2013.

[3] Soto López, C. E., Anselmi, F., Gharbi Dit Kacem, M., and Lazzari, P.: Data-Informed Inversion Model (DIIM): a framework to retrieve marine optical constituents in the BOUSSOLE site using a three-stream irradiance model, Geosci. Model Dev. Discuss. [preprint], https://doi.org/10.5194/gmd-2024-174, in review, 2024.

T21

Local Climate Zone Classification in Jakarta Using AI: Machine Learning Approach with Landsat 9 and Land Surface Temperature Integration

Rezky Yunita [1], Muhammad Rifki Taufik [1], Erwin E. S. Makmur [1], and Donaldi Sukma Permana [1]

[1]*Indonesian Agency for Meteorology, Climatology, and Geophysics*

Local Climate Zone (LCZ) classification [1] is essential for urban climate studies, land use planning, and environmental monitoring. Many studies of LCZ have been conducted in big cities around the world especially for urban heat island assessment [2-7]. However, studies of LCZ in Indonesia, especially Jakarta as the capital city, are still limited, and previous studies comparing the performance of AI classification methods for LCZ is hardly found. This study explores the utilization of Artificial Intelligence (AI) techniques to classify LCZs in Jakarta using Landsat 9 remote sensing data in 2023. Three machine learning algorithms—Random Forest (RF), Support Vector Machine (SVM), and Gradient Tree Boosting (GTB)—are implemented and compared based on spectral (band 1-7) and spatial features. Additionally, Land Surface Temperature (LST) data is incorporated to enhance classification accuracy and analyze temperature variations within each LCZ. The performance of the models is evaluated using overall accuracy metrics. The results indicate that RF achieves the highest accuracy, outperforming both SVM and GTB, with GTB ranking second. The integration of LST further enhanced LCZ classification by providing insights into temperature distribution across urban and natural landscapes. These findings emphasized on the utilization of machine learning methods and the thermal data in LCZ classification for urban climate analysis.

References

[1] Stewart, Ian D., and Tim R. Oke. "Local climate zones for urban temperature studies." Bulletin of the American Meteorological Society 93.12 (2012)

[2] Leconte, Francois, Julien Bouyer, Rémy Claverie, and Mathieu Pétrissans. "Using Local Climate Zone scheme for UHI assessment: Evaluation of the method using mobile measurements." Building and Environment 83. 39-49. (2015)

[3] Geletič, Jan, Michal Lehnert, Stevan Savić, and Dragan Milošević. "Inter-/intra-zonal seasonal variability of the surface urban heat island based on local climate zones in three central European cities." Building and Environment 156. 21-32. (2019).

[4] Yang, Xiaoshan, Lilliana LH Peng, Zhidian Jiang, Yuan Chen, Lingye Yao, Yunfei He, and Tianjing Xu. "Impact of urban heat island on energy demand in buildings: Local climate zones in Nanjing." Applied Energy 260. 114279. (2020)

[5] Chang, Yue, Jingfeng Xiao, Xuxiang Li, Ariane Middel, Yunwei Zhang, Zhaolin Gu, Yiping Wu, and Shan He. "Exploring diurnal thermal variations in urban local climate zones with ECOSTRESS land surface temperature data." Remote Sensing of Environment 263. 112544. (2021)

[6] Ma, Lei, Guoan Huang, Brian Alan Johnson, Zhenjie Chen, Manchun Li, Ziyun Yan, Wenfeng Zhan, Heng Lu, Weiqiang He, and Dongjie Lian. "Investigating urban heat-related health risks based on local climate zones: A case study of Changzhou in China." Sustainable Cities and Society 91. 104402. (2023)

[7] Zhang, Rui, Jun Yang, Xinyue Ma, Xiangming Xiao, and Jianhong Cecilia Xia. "Optimal allocation of local climate zones based on heat vulnerability perspective." Sustainable Cities and Society 99. 104981. (2023) on local climate zones: A case study of Changzhou in China." Sustainable Cities and Society 91. 104402. (2023)