ESMO
Earth System Modelling and Observations

WGNE
Working Group on Numerical Experimentation

WCRP
World Climate Research Programme

39th session of the Working Group on Numerical Experimentation (WGNE)

joint with the 25th session of the Working Group on Subseasonal to Interdecadal Prediction (WGSIP)

Météo-France, Toulouse, France

4-8 November 2024

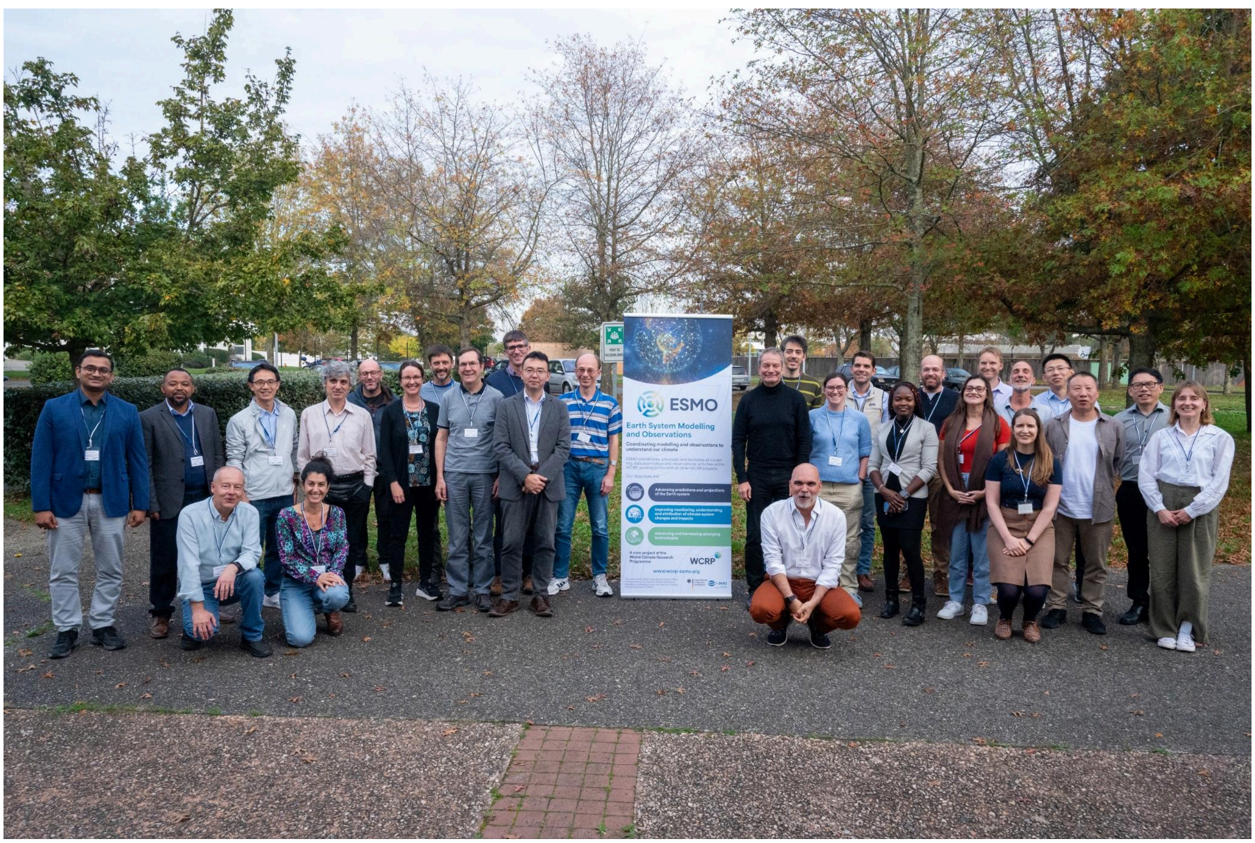

Table of content

ESMO Earth System Modelling and Observations | WGNE Working Group on Numerical Experimentation | WCRP World Climate Research Programme

Joint Session 1 – Introduction

Introduction to the TRACCS programme (David Salas Y Melia, CNRM)

Climate change poses immense challenges for the development conditions of all human societies and their relationship with nature and the environment. The next few decades are crucial for reducing the escalation of climate risks, both in terms of reducing greenhouse gas emissions and adapting to the impacts already present. Distilling climate information from multiple lines of evidence and spatio-temporal scales, and taking the user context into account will increase the relevance of climate information for decision- making. The French climate science community plays a major role at increasing the awareness of climate change issues at the national and international level and has the skills to meet the research challenges required to offer a wide range of knowledge and tools: fitness-for-purpose of climate models, downscaling and bias correction, interdisciplinarity and co-construction with stakeholders of the climate information they need in terms of the type of information and its spatial and temporal scale. Climate sciences are also facing major challenges, linked to technological and scientific advances (high-performance computing, new computing architectures, coupling to additional models such as ice sheets, artificial intelligence techniques, data assimilation) but also to organisational transformations at European and international levels. The exploratory PEPR "TRAnsforming Climate Modelling for Climate Services" (TRACCS) aims at transforming climate modelling to face these challenges and to meet societal expectations, by providing the best possible climate information for climate services. TRACCS will combine several types of activities: i) process modelling for simulating climate from global to local scales so as to provide reliable climate information for assessing both mitigation scenarios and local adaptation actions and their feedback effects on the climate, ii) adaptation of computer codes to new computing architectures, to seize the opportunities offered by exascale in terms of increasing spatial resolution, better representing the complexity of the climate system, and exploring large sets of simulations and obtain better information on model and scenario-related uncertainties, iii) the use of advanced statistical methods and artificial intelligence to characterise climate extremes, accelerate models, and develop emulators (fast statistical models) to better quantify uncertainties, iv) the estimation of the impacts of climate change on different activity sectors and different territories in order to co-construct adaptation, and v) the development of a dialogue between scientists and stakeholders, teaching and communication with all audiences in order to co-construct prototypes of climate services in a transdisciplinary approach. This, together with the training of a new generation of experts in climate change and its impacts, will transform the way scientific advances on climate change are shared with stakeholders, thereby multiplying the capacity for science- based adaptation and mitigation. The joint management by Météo-France and the CNRS of a scientific and research program that is crucial for public action at all levels is a concrete example of an effective cooperation between the academic research community (at the forefront of which is IPSL) and the national operator in the field of meteorology and climate

(Météo-France). This program associates several leading organizations and institutions in this field within the framework of core projects projects, and through the calls for proposals will allow to extend the participation of the whole relevant national scientific community to the program objectives. In order to achieve its objectives, TRACCS consists of 10 Targeted Projects (PCs), 4 of which are relevant to the WGNE and are detailed in the following presentations.

TRACCS-PC5: COMputing PAradigms towards efficient, modular and trainable ClimaTe models (COMPACT) (Julien Le Sommer, IGE)

Climate models are scientific tools for understanding how the climate system works. They are also essential numerical tools that provide quantitative information needed for public policy-making, particularly in the context of climate services. However, their structure will have to evolve significantly in the coming decade in order to respond to the evolution of their applications, and to take advantage of new scientific and technological opportunities. These numerical models will need to provide reliable information at finer spatial resolutions, while quantifying the uncertainties of their projections. They will need to take optimal advantage of new supercomputer architectures, while allowing users to adjust the complexity of their components according to the targeted application. While continuing to be based on robust scientific laws, they will need to make more systematic use of Earth Observation data for their formulation and calibration. For this, they will be able to rely on the development of machine learning and differentiable programming applied to scientific computing. One of the challenges of the decade in the field of climate modeling is therefore to succeed in gradually transforming existing systems to make them more efficient, more modular hence composable, and more compatible with artificial intelligence tools. The primary constraint lies in the fundamental structure of the software systems used in climate modeling. These systems, their components and their interfaces are indeed written in robust programming languages but with low levels of abstraction. In this context, the objective of this project is to guide the evolution of the design and implementation of national climate models to 2030, and to train new scientists to accompany the evolution of our modeling tools beyond 2030. In practice, we propose a combination of exploratory activities and efforts to structurally modify the numerical tools that comprise the national climate modeling systems. We propose to focus primarily on tools that are shared by the two existing systems (IPSL-CM and CNRM-CM). Our activities will be organized around three complementary axes dealing respectively with high performance computing, modularity and the AI hybridization of these tools. The activities proposed on each of these axes will be integrated in the evolution roadmap of each of the tools, ensuring their integration in future versions of national climate models. The project will allow the training and support of a new generation of experts on these topics within our community.

TRACCS-PC6: QUantifying uncertaINties, Tuning and Equilibrating climaTe models (QUINTET) (Aurore Voldoire, CNRM)

Quantifying uncertainties in climate simulations is a long-standing issue that remains largely unsolved despite huge efforts to tackle it, in CMIP6 and several other projects. Today, it has become urgent that the climate modelling community embraces techniques to quantify uncertainties that are less expensive in terms of manpower and computational resources, yet more relevant to societal needs for climate services. The most critical bottleneck of climate model development is the calibration of parameters, as it directly impacts emerging properties from the climate models such as climate sensitivity. So far, most of the calibration of climate models has been done « by hand », based on empirical and pragmatic approaches and on the model developers' expertise. Lately, the French climate modelling community has engaged with mathematicians to develop a new semi-automatic framework for model calibration, which involves uncertainty quantification to reduce the risk of overfitting and machine learning to reduce the computational cost. This promising framework is distinct from optimization as it allows to explore different possible acceptable ('tuned') configurations and thereby different climates. It also allows us to quantify remaining uncertainties once a degree of tolerance is satisfied for a list of given targets. It can also take on all sorts of constraints (global metrics, local observations, statistics from LES…) within a whole model configuration hierarchy. This flexibility is particularly relevant to address the diversity of climate services. However, applying such a framework to components with different typical timescales is still a challenge and questions the methodology of spinning up climate models. There is in particular a need to accelerate spin-up of climate models, so as to deal with the long adjustment time scales. Several other questions also arise, for example how to verify the models over the present period, when observations are available but the climate is transient, and whether or not metrics representing the past transient climate should be included for the calibration. All these questions will be addressed within TRACCS-PC6-QUINTET, following the general objective to facilitate the exploration of parametric uncertainty (perturbed physics ensembles, including assessments of low-likelihood-high-risk scenarios), and to get prepared for higher resolution configurations.

TRACCS-PC7: Improving the physical PRoceSS representatION in Earth System Models (IMPRESSION-ESM) (David Salas Y Melia, CNRM)

The overarching objective of TRACCS PC7 – IMPRESSION-ESM – is to improve the representation of key physical processes in the two French Earth System Models (ESM), possibly adding those that are currently missing, and thereby contribute to some of the main TRACCS objectives: improve our understanding of the Earth system, increase the confidence in ESM output, provide high-quality climate information to downstream relevant applications, and train the next generation of climate professionals. Overall, the improvement of the physical process representation in ESMs will mostly rely on the following approaches:

1. The development of theoretical or conceptual models of physical processes and their connection to the revolved-scale dynamics. Ongoing and solid collaborations between IMPRESSION-ESM partners and the physical Earth System process scientific community will allow IMPRESSION-ESM to benefit from the latest scientific advances.
2. The use of process models, high-resolution simulations, and targeted observations to evaluate and improve the conceptual models at the heart of parameterizations. Some of these models and observations are already well established and used within the IMPRESSION-ESM community, while the development of new reference models or the use of new or upcoming observational datasets will be fostered.
3. The exploration of Artificial-Intelligence-based techniques, applied to observations or more complex models, to design parameterizations that better capture the complexity and diversity of the real world.

More specifically, IMPRESSION-ESM will tackle the following key processes:
- Atmospheric deep convection, its initiation through the cumulus and congestus phases, organisation, variability, and interaction with radiation;
- The role of land surface heterogeneities in driving the water cycle and its interactions with the atmosphere;
- Snow processes at high latitudes and over ice sheets and their interactions with topography heterogeneities or radiation;
- Ocean mixing processes and the interaction of the oceanic circulation with the bathymetry
- The multiphase thermodynamics and rheology of sea ice;
- The air-sea interface and associated numerics, and the couplings across the continuum from continental ice, ice shelves, sea ice to icebergs.

IMPRESSION-ESM proposes a long-term contract framework, with researchers and engineers hired for up to 8 years, which is crucial for the training of highly-expertised scientists, and a small number of shorter-term contracts on targeted development tasks.

TRACCS-PC8: Enhanced treatment of biogeochemical processes - Climate feedbacks, impacts and vulnerabilities (Roland Séférian, CNRM)

Our knowledge on complex and subtle biogeochemical interactions within the Earth system has improved significantly thanks to recent advances in theory, observations and experiments. The current level of integration of these processes in Earth system models (ESMs) is lagging behind the most up-to-date knowledge, mainly due to the competition between increasing model complexity and resolution. This limits their ability to deliver tailored information on climate/Earth system feedbacks, impacts and vulnerabilities.

TRACCS-PC8 aims to explore a wider range of approaches to make a step change in the representation of biogeochemical responses in ESMs and analyze their interactions and feedback with the physical climate. PC8 will focus on components of high level of interest for their impacts on the Earth system, namely: CO2, methane (CH4), exchanges of limiting nutrients across the components such as phosphorus (P), nitrogen (N), carbon (C), and iron (Fe); fires, volatile organic compounds, and the dynamics of semi-arid ecosystems. Machine learning will be used in collaboration with PC5, where relevant, to replace CPU-demanding components of the Earth system models to produce efficient modules for, e.g., radiative transfer in the canopy, soil organic matter mineralisation, process-based description of forest mortality. Most of these developments will be designed jointly by the French climate research community as open-access codes to be integrated into the two ESMs. The climate simulations using these advanced models will provide the basis for studying a wider array of processes and biochemical interactions across the Earth system components, as well as their associated uncertainties. Most of these developments will be done jointly within the French community. This program will bridge the gap between the ESMs developments that have a long tradition, the recent observations and the theory, and will study the societal challenges closely linked to climate changes such as air quality, marine and terrestrial biodiversity, productivity and resilience to climate change, the impacts on the precipitation and drought occurrences.

WCRP & ESMO introduction (Fanny Adloff, ESMO)

Body of the text

WGNE Introduction (Nils Wedi, Ariane Frassoni)

The co-chairs introduced WGNE including the new members, typically global experts vested in enhancing the emerging capacities of operational meteorological modelling centres. The introduction described WGNE's evolution as one of the oldest WMO working groups, and defining its terms of reference and contributions to the new ESMO WCRP core project. It followed a description of WGNE project activities and continued evaluations of systematic errors in weather & climate simulations with examples.

WGSIP Introduction (June-Yi Lee, Bill Merryfield, Ángel Muñoz)

Body of the text

WWRP introduction (Chris Davis, WMO)

WWRP has several projects and working groups that are aligned broadly with objectives of WGNE and, more broadly, with ESMO. The relevant projects include the Paris Olympics RDP that is ending at the end of 2024. This project was concerned with sub-km-scale prediction in urban environments for moist convection, urban heat and air quality. Some of the findings and motivations will progress to WWRP's new urban project starting in 2025. Other projects discussed included PCAPS, looking at prediction and applications for services and communities in polar regions. SAGE is examining applications of sub-seasonal prediction in agriculture, health, energy and disaster risk reduction, in addition to improving our modeling systems to get the most out of the limited predictability on time scales of 2-6 weeks. InPRHA is examining hydrological and atmospheric prediction integration for short-range flood warnings. Finally, the PDEF, DAOS, and Verification (JWGFVR) working groups focus on topics of predictability, data assimilation, observing systems, and forecast verification, all of which are relevant to ESMO broadly. The joint working group on forecast verification research (JWGFVR) is formally joint with WGNE. The talk also advertised open positions on the steering committees for these three working groups and others.

WIPPS & EW4All Introduction (David Richardson, ECMWF)

The WMO Integrated Processing and Prediction System (WIPPS) is a global network of operational centres providing earth-system analyses and predictions to all WMO Members and the wider community. Designated WIPPS centres agree to provide defined sets of mandatory and recommended products to all Members.

In the context of the United Nations Early Warnings for All initiative (EW4All) it is important to develop the WIPPS to make sure that it can provide all necessary forecast information to support WMO Members. This led to a significant increase in the number and resolution of the global and regional products from WIPPS centres. However the increase in the volume of forecast data does bring some big challenges – WMO Members must be able to access and utilize this important new forecast information.

The current rapid development of machine learning (ML) and artificial intelligence (AI) technologies and their use in weather forecasting brings substantial opportunities as well as potential risks. These developments may have significant implications for operational practices and the evolution of the WIPPS.

As users' needs change and the modelling capabilities improve, so WIPPS needs to evolve to meet these new user requirements. Working together with WGNE, WGSIP and ESMO in general will ensure that WIPPS operational activities benefit from the latest developments in earth-system modeling for the benefit of WMO Members.

EPESC (Scott Osprey, NCAS and University of Oxford)

Body of the text

GAW introduction (Sara Basart, WMO)

Body of the text

Modelling Aspects of G3W (Gianpaolo Balsamo, WMO)

The Global Greenhouse Gas Watch reached its implementation and pre-operational phase in summer 2024 with approval of the WMO executive council. The G3W modelling systems are an integral part of the global monitoring support capacity that is under development. The G3W progress and structure was briefly presented and discussed and it also has a dedicated site https://g3w.wmo.int that provides regular updates. In the first months of activity G3W has formed the Advisory Group and Task Teams. Specific needs were presented and discussed and relate to the GHG modelling spatial resolution, and timeliness, since G3W aims at delivering near-real-time net fluxes for 3 major GHGs, Carbon dioxide (CO_2), Methane (CH_4), and Nitrous Oxide (N_2O). The G3W modelling data will be integrated in the WIPPS (see contribution on WIPPS from David Richardson, ECMWF).

Member presentations

JMA (Masashi Ujiie)

The Japan Meteorological Agency (JMA) recently launched two new HPC systems in 2023 and 2024. One system employs an A64FX based architecture. The system is specialized for mesoscale convective system prediction. The other is an intel CPU based system for the other NWP systems. Both systems employ with HBM2 based memory and this high bandwidth memory enabled to achieve 4x effective performance of the JMA's previous HPC system (based on Cray XC50) in the whole two systems.

Using the A64FX based system, JMA is planning to enhance horizontal resolution of LFM (Local Forecast Model), a JMA's regional NWP model, from 2 km to 1km. The recent progress showed that the 1 km LFM can improve representation of convective initiation and transition from shallow to deep convection compared to the 2 km LFM. However, the 1km LFM still tends to represent delayed convective initiation and too rapid transition to deep convection compared to large eddy simulations. Further improvements to vertical transport processes are necessary.

JMA also started investigation of machine learning (ML) based NWP models to catch up rapid progress of ML based models. JMA carried out numerical experiments using existing ML models such as GraphCast, FourCastNet initialized from the JMA's operational global analysis. Through these experiments, JMA has recognized that ML models are more feasible than expected. Experiments of fine-tuning, additional two-years training for GraphCast using the JMA's operational global analysis, have clarified that fine-tuned GraphCast using the high resolution analysis can represent deeper tropical cyclone intensity than normal GraphCast. JMA is also trying to write a low-resolution (1.25 degree) global ML model from scratch to be familiarized with ML models using graph neural network with a hierarchical structure. Even though this low-resolution ML model is simple, the model can show skillful 500hPa height forecasts compared with the persistent forecast. Using existing and toy Ml models JMA accumulates experience on ML model development.

KIAPS (Eun-Hee Lee)

The Korea Institute of Atmospheric Prediction Systems (KIAPS) was established in 2011 with the mission of developing Global Numerical Weather Prediction Systems for South Korea. During the 1st phase of the KIAPS project, the Korean Integrated Model (KIM) was developed and it has been supporting the operational medium-range predictions at the Korea Meteorological Administration since 2020. Now, the 2nd phase of the KIAPS project focuses on expanding KIM applications from short- to extended-ranges. This includes high-resolution modeling targeting ~km-scale, scale adaptive physics and data assimilation, and coupling various Earth system components. Recently, we developed the global variable resolution system and the limited area model based on the original cubed-sphere grid system of KIM. Its km-scale simulations show potential for short-range prediction systems. In addition, KIM coupled systems and coupled data assimilation development is also ongoing. Our recent hindcast experiments with reanalysis initials demonstrate improved performance of the coupled version of KIM in simulating MJO. Besides new applications, we have also supported the upgrade for the operational KIM. The transition to KIM4.0 is nearly complete this year, which includes an upgrade in horizontal resolution of model (8-km) and data assimilation (24-km) with revisions to enhance convective scale-awareness. This upgrade improves overall global skill and prediction performance for high impact weather, showing the capability of high-resolution modeling with appropriate modulation of scale-dependent physics.

CMA (Xingliang Li)

CMA Earth System Modeling and Prediction Center (CEMC) recently undertaken on operational upgrades the main changes including 1km1h cycle's CMA-MESO V6.0, En4Dvar-based CMA-GFS V4.2, physical progress tuning for addressing underestimation of cloudiness and low precipitation level, and more satellite data used in CMA-GFS V4.2 like FY-4B AGRI/GIIRS, FY-3E MWHS 183GHz vapor channal, FY-3E WindRAD scatterometer wind, HY-2B SMR imager,HY2D scatterometer wind and so on. The coming CMA-GFS V4.2 overall performance is improved, especially in the East Asia region.
With a new operational demand for regional hundred-meter scale, global kilometer resolution and the ability of the numerical model to adapt to exascale computing platforms, the regional/global unified nonhydrostatic Multi-moment Constrained finite Volume (MCV) model has been developed for the next generation atmospheric modeling at CEMC. The preliminary case study of heavy rainfall indicates that a global/regional unified next-generation MCV model has great potential to the real weather prediction.

Indian Institute of Tropical Meteorology (Ankur Srivastava)

The Indian Institute of Tropical Meteorology (IITM) is tasked with developing dynamical models under the Monsoon Mission program for short-to-medium range (GFS/GEFS based), extended-range (CFS based) and seasonal forecasts (CFS based). These three models are deployed at the Indian Meteorological Department (IMD) for operations, and the systems are under continuous development at IITM.

The seasonal forecast system (Monsoon Mission Climate Forecast System version 2) has recently been significantly upgraded with a semi-Lagrangian dynamical core in GFS, MOM6, CICE5, and NEMS-based coupler. IITM has also set up a Weakly Coupled Data Assimilation System, and further research is underway to include river routing, COARE-based bulk-flux algorithms, convection schemes, and cloud macro/microphysics. The extended range forecasting system is upgraded to have an 18-member physics ensemble with combinations of convection of microphysics schemes; the system is run at a high resolution (T574) for the first 15 days and a lower resolution for the next 15 days (T382). IITM High-Resolution Global Forecast Model Version 1 has been running in experimental mode at IITM since June 2022. The resolution has doubled to ~6 km from the existing ~12 km using Spectral Cubic Octahedral Grid.

HPC infrastructure has been significantly upgraded: A new HPC system is deployed at IITM, which has a capacity of 11.77 PetaFLOPS (PF) and 33 petabytes of storage, and the facility at the National Centre for Medium-Range Weather Forecasting (NCMRWF) facility features 8.24 PF with 24 petabytes of storage. Additionally, there is a dedicated standalone system for Artificial Intelligence and Machine Learning applications with a capacity of 1.9 PF.

CSIR (Mohau Jacob Mateyisi)

Body of the text

UK Met Office (Tim Graham)

The UK Met Office is now close to getting its new HPC into operational service. The first operational upgrade on the new HPC in 2025 will be GC5 for global modelling and RAL3 for regional modelling. For global modelling, improvements are expected across many aspects including tropical cyclones. Improved model stability has allowed rebalancing of additive inflation and stochastic physics in the ensemble leading to greatly improved CRPS scores. For regional modelling, a key achievement with RAL3 is the ability to use the same science configuration for tropical and mid-latitude regions. This is a significant step on the path to global km-scale modelling.

Further plans for the new HPC included the retirement of deterministic models in favour of ensemble forecast systems. Work is ongoing on the implementation of the LFRic dynamical core with operational implementation planned around 2028.

ECMWF (Inna Polichtchouk)

The European Centre for Medium-Range Weather Forecasts (ECMWF) recently upgraded its operational model cycle to CY49R1, with notable changes including assimilation of 2m temperature observations and activation of the stochastically perturbed parameterisations (SPP) model uncertainty scheme in all ensemble configurations. Since August 2023, ECMWF has also been producing daily deterministic forecasts with the prototype global Extremes Digital Twin of Destination Earth (destinE) at horizontal resolution of 4.4 km that brings better forecast skill in near-surface parameters in comparison to the operational 9 km forecasts.

Preparation is under-way for the ERA6 reanalysis, which will have a finer horizontal resolution (14 km) and a better representation of tropical cyclones and great lakes than its predecessor ERA5. A new OCEAN6/ORAS6 ocean analysis/reanalysis product has been released that has a better representation of SSTs and salinity than its predecessor

OCEAN5/ORAS5, allowing the partial coupling to be switched off in the forecasts in the next operational cycle upgrade.

Other recent achievements include a move toward all-sky, all-surface, coupled data assimilation, the development of the operational AI/ML AIFS ensemble prediction system, and a Data Driven Machine Learning Forecast trained/initialised directly from raw observations.

DWD (Günther Zängl)

The main innovation in DWD's forecasting system was the operationalization of the RUC (rapid update cycle) branch of ICON-D2 on July 10, 2024, which was developed in the "SINFONY" project since 2018. The primary goals are improving the prediction of convective precipitation and related hazards on very short time scales and a seamless blending between nowcasting methods and model forecasts for radar reflectivity and precipitation. In contrast to the standard ICON-D2, the ICON-D2-RUC uses a full 2-moment cloud microphysics scheme (which provides much better radar reflectivities than the standard 1-moment scheme), a shorter data cutoff time (about 15 minutes), and performs hourly forecasts with a lead time of 14 hours.

Additional significant improvements in forecast quality could be achieved by introducing adaptive surface friction in ICON-D2 (February 2024) and by improvements in the ensemble physics perturbations, which were introduced in several steps.

For 2025, DWD prepares 500-m forecasts for Germany, which will be accomplished by using ICON's two-way nesting capability. Starting from the deterministic ICON-D2 analyses, ICON-D05 will provide 48-h forecasts 8 times/day. Most of the expected improvements in forecast quality are related to the better orography resolution, pertaining to 10-m wind speed and gusts, 2-temperature and 2-m humidity.

INPE/CPTEC (Ariane Frassoni)

The presentation highlights INPE/CPTEC's advancements in numerical modeling and future goals. Current operational modeling systems include BRAMS, Eta, WRF regional models and the BAM global model, supporting weather, subseasonal, and seasonal forecasts. The MONAN Earth System Model, designed for regional and global applications, is central to INPE's efforts, emphasizing open-source, community-driven development and should replace the current suite of models. Recent updates focus on the MONAN model, including its atmosphere-land components, with plans to integrate ocean and cryosphere components by 2027. Future deliverables include ensemble forecasts, severe storm nowcasting, and air pollution prediction, leveraging the JEDI data assimilation framework. By 2030, MONAN aims to support climate change scenarios and Earth system modeling for South America. A case study on catastrophic flooding in Rio Grande do Sul (April–May 2024) demonstrated the MONAN model's ability to predict extreme rainfall with accuracy comparable to global models like ECMWF. Key improvements include cold-pool parameterization, updated microphysics, and surface flux evaluations. INPE also hosted its first MONAN training for the South American community in 2024, fostering regional capacity building. INPE plans to enhance MONAN model physics, expand applications, and integrate multidisciplinary components for improved predictions and climate resilience.

NSF NCAR (Peter Hjort Lauritzen)

The NSF National Center for Atmospheric Research (NCAR) is finalizing version 3 of the Community Earth System Model (CESM) in preparation for CMIP7 flagship simulations. An overview of the updates in version 7 of the atmospheric component, CAM (Community Atmosphere Model), was presented.

In summary, the vertical resolution has been increased in the boundary layer, and the model lid has been raised from approximately 42 km to 80 km. The enhanced boundary layer resolution is accompanied by modifications to the deep convection scheme. Other significant changes include updates to microphysics, which substantially reduce the Equilibrium Climate Sensitivity (ECS), the addition of a “moving mountains” gravity wave drag scheme to represent previously missing sources of sub-grid-scale gravity waves (improving zonal mean stratospheric wind biases), and the implementation of a convective gustiness parameterization. Also, the default dynamical core used in CAM versions 4, 5, and 6 has been replaced with the spectral-element dynamical core, which utilizes a Conservative Semi-Lagrangian Multi-tracer scheme (CSLAM) for transport and a separate finite-volume physics grid. Current biases being addressed by the CESM group were also reviewed. Updates on high-resolution CESM efforts were provided, highlighting the coupling of CAM physics with the MPAS dynamical core, targeting a horizontal resolution of 3.75 km.

ECCC (Ron McTaggart-Cowan)

The Canadian Meteorological Centre (CMC) recently completed a major upgrade of its global ensemble prediction suite, with large improvements in both control-member and probabilistic scores obtained through reduction in grid spacing from ~39km to ~25km. Minor adjustments to the stochastically perturbed parameterization (SPP) scheme used to represent model uncertainty were required. In addition to improving ensemble predictions themselves, the higher resolution background error covariance matrix improves predictions in the global deterministic system through the ensemble-variational data assimilation system.

Recent efforts at the CMC have also been focused on the development of hybrid machine-learning (ML)/physically based prediction systems. Predictions from a global model that is spectrally nudged towards ML inferences are significantly more skillful than current operational guidance. From-scratch retraining of the ML model is currently under way with a view towards experimental implementation in 2025.

NCEP (Fanglin Yang)

The NCEP Environmental Modeling Center (EMC) develops and enhances data assimilation systems and environmental models, encompassing atmospheric, oceanic, and land-surface systems. Collaborating with universities, NOAA labs, government agencies, and international scientists, EMC employs advanced techniques to improve operational forecasting. Their Unified Forecast System (UFS), a versatile tool for various applications, spans domains from local to global and predictive timescales from minutes to years, consolidating isolated production suites into a unified framework.

Recent achievements include the implementation of HAFSv1 for the 2023 hurricane season and enhancements to the Rapid Refresh Forecast System (RRFS), including convection scheme updates that improved forecasts. The first interactively coupled atmospheric and air quality model (UFS-AQMv7) was implemented into operation in May 2024, It improved pollutant predictions (ozone and PM2.5) and showed a better agreement with AirNow data for events like Quebec wildfire emissions.

Significant advancements in global model development include transitioning GFSv16 and GEFSv12 to UFS-based GFSv17 and GEFSv13, achieving finer resolution (9 km) for medium-range forecasts and improved Madden-Julian Oscillation (MJO) prediction in sub-seasonal models. Seasonal forecasts benefited from enhancements like the improved Quasi-Biennial Oscillation (QBO) in the SFS prototype. These efforts solidify EMC's role in providing foundational guidance for the National Weather Service.

CNRM (Romain Roehrig)

Météo-France develops and operates a full suite of forecasting systems covering the global (ARPEGE- and regional (AROME) scales and a variety of lead times from near-real time to a few days. The Météo-France regional forecasting system is used over Metropolitan France, French overseas territories and on-demand domains as requested for various applications. In October 2024, the latest version of these systems based on the IFS/ARPEGE cycle 48t1_op1 replaced the older suite. Major upgrades were conducted in the model physics (deep convection and radiative transfer), data assimilation algorithm (Hybrid 4DVar for ARPEGE, 3DEnVar for AROME), and the assimilation of new datasets (esp. all-sky microwave radiances). Technical updates were also included to continuously improve the system efficiency. Overall, the new suite has improved forecasting skills for both its deterministic and ensemble components. Note that, given the recent increased complexity of the Météo-France production line, it took more than a year to fully validate the forecasting system upgrade in its operational context.

In the perspective of the next systems, several research activities are ongoing: revised parameterizations involved in the representation of shallow clouds, improvements of cold and mixed phase microphysics, development of a 4DEnVar data assimilation system for AROME, increased resolution in AROME to reach 750 m or less, inclusion of more Earth system processes (see Cindy Lebeaupin's talk), introduction of a convective organisation parameter in the ARPEGE convection scheme, refactoring of the code for GPUs and other accelerators. The next suite will be based on the IFS/ARPEGE cycle 49t1_op1, will include some of the previously mentioned developments, and should hopefully enter in operations near mid-2026. In parallel, the CNRM is enhancing its AI-based model activities (see Sara Akodad's talk).

On the climate modeling side, the CNRM is preparing its contribution to the AR7 Fast Track and CMIP7. For the former exercise, the CNRM will use the ESM it developed for CMIP6, with slight updates for its use in an emission-driven configuration. The CNRM has more ambition for CMIP7 in terms of model updates. In particular, the strategy is to achieve an atmospheric component as seamlessly as possible with its configuration for NWP applications, both for the technical and physical contents.

Research Board activities (AI TT, Digital Strategy, General) (Nils Wedi)

The RB aims to connect strategic aims of the WMO Executive Council and Congress with the community, e.g. EW4All, G3W, Cryosphere, Hydrology. The RB is also a forum for exchange and alignment of activities across WMO regions, INFCOM & SERCOM, as well as independent science input in alignment with the science advisory panel (SAP) and the research community organised in WWRP, WCRP and GAW. Updates from WWRP included the endorsement of the science project plans for PEOPLE, SAGE, inPRHA, and PCAPS. There is WWRP conference planned in 2027. WCRP presented v11 of the science & implementation plan. There was a discussion on ocean and hydrology activities. GAW highlighted a range of observational activities and G3W, and there was discussion on urban priorities. The RB met this year with the SAP discussing emerging challenges (tipping points, big data, AI), attribution science & youth strategy. Finally, there is a task force team on AI until the end of 2024 to identify & share activities on AI and provide member advice.

Model Uncertainty Model Intercomparison Project (MUMIP) (Hannah Christensen, University of Oxford)

MU-MIP has made significant progress over the last year. Funding from the UK Leverhulme Trust has led to the hiring of three postdoctoral researchers at U. Oxford, U. Exeter and Météo France, to supplement ongoing funding through the NCAR/NOAA Developmental Testbed Centre.. Simulations have been completed with four different models: the ECMWF IFS, Arpege-Climat, Rapid Refresh and GFS, while UK Met Office runs are ongoing. These runs cover a domain in the Indian Ocean at 0.2 degree resolution. Development of analysis scripts is underway.

WGNE Table (Günther Zängl)

The most important upgrades in the HPC systems and NWP configurations of the participating global NWP Centers were summarized. New or upgraded HPC systems have been installed at DWD, NRL, KMA and JMA. Resolution upgrades in global or regional forecasting systems were introduced at UKMO, ECCC, KMA and CMA.

HPC/Exascale Computing Update (Nils Wedi)

The talk provided an overview of trends in preparing for exascale HPC use in weather & climate with input from members. Areas discussed included GPU adaptation successes and pitfalls, reduced or mixed precision approaches, revising I/O pipelines and workflows, in particular in connection with distributed compute and data as well as machine learning. Maintainability and portability of software and data was also discussed. Important trends are digital twin developments, adaptation challenges to AI-function specific hardware, diversity of available accelerator hardware, and challenges with highly complex physical model codes. Improving coupling to and expanding acceleration efforts to other Earth system components is another common theme. Available compiler issues and concerns and separation of concerns concepts with source-to-source translation were mentioned, but stressing the need and substantial effort required to refactor codes nevertheless, and the overall significant time spent by several WGNE members on HPC upgrades and migration. The slides also annex the detailed individual contributions by members.

Physics-dynamics coupling and energy budgets in Earth System Models (Peter Lauritzen)

An overview of the WGNE initiative on Physics-Dynamics Coupling and Energy Budgets in Earth System Models was presented. The motivation behind this initiative is the recognition that no Earth System Model rigorously closes its energy budgets. This is a complex challenge involving unresolved processes, differing assumptions (even within individual model components), and long-standing practices ingrained in modeling systems over decades.

As far as WGNE is aware, there is currently no coordinated effort to discuss or evaluate how Earth System Models address the closure of total energy budgets. To fill this gap, WGNE has initiated virtual discussions held approximately every two months. These discussions focus on specific topics related to physics-dynamics coupling and energy budgets.

The objectives of these discussions are to:

- Illuminate the inner workings of our modeling systems,
- Share insights on what works well and what doesn’t,
- Assess the impact of certain errors, and
- Work towards compiling a WGNE table for physics-dynamics coupling and energy budgets.

Ensemble predictability and forecasting at met centers (Tim Graham)

Many thanks to all contributors for updating the table and providing slides on ensemble developments at operational centres. Several centres have recently upgraded or plan to upgrade their ensemble systems. Developments include increases to resolution, increased number of members, developments in generating initial perturbations and model uncertainty as well as changes to the underlying physical model.

Discussion considered some of the challenges around developing earth system models for use in ensemble systems such as the cost of testing physics developments in an ensemble and the complex interactions of ensemble data assimilation systems with the model physics.

WGNE projects/collaborative activities

AI for Weather (AI4Wx) Research Board Task Team (Catherine de Burgh-Day, BoM)

The AI for Weather Task Team (AI4Wx TT) was formed under the Research Board to provide a rapid response to the needs of members around the rapid advances of AI/ML for weather and climate prediction.

The scope of the AI4Wx TT is to identify existing AI/ML activities within WMO programs and Regional Associations, review activities which can be expanded with limited resources to include AI/ML, identify main science gaps of immediate relevance, identify other key needs of members (e.g. open data, accessibility), and develop a report on these topics including recommendations on immediate actions.

The TT has been meeting fortnightly since mid-August, and has an intended duration of six months. Membership includes representatives from key WMO groups including WWRP, WCRP, GAW, SERCOM, INFCOM and the SAP.

TT members have been surveyed on activities within their WMO area(s) which include AI/ML components, or could easily be extended to include AI/ML activities. They have also contributed to a list of science gaps and priorities within the domain of AI/ML. Information gathering has now been extended to the Regional Associations to explore current AI/ML activities and identified science priorities through that lens.

Immediate-term deliverables which would be of benefit to WMO members have also been canvassed, and include a labeled database of key AI/ML papers (development underway), and a knowledge resource describing current best practice in the development, use and evaluation of AI weather and climate models.

Two key priorities identified by the TT are the importance of maintaining open data and model principles, and ensuring AI/ML model and infrastructure research and development ensure continued autonomy and accessibility for the Global South.

Looking ahead, the TT plans to explore the knowledge building and training needs of members, and will soon begin the process of consolidating its findings and activities into the final report.

TC Verification and initialization overview (Masashi Ujiie)

Tropical Cyclone (TC) verification for 2023 was presented. Position errors in participating centres at T+72 hours were decreased compared to the previous year. This reduction was probably due to year-by-year variation rather than each model's upgrade.

The verification also results showed several new findings on known common biases in the Western North Pacific (WNP) basin. For northward moving bias before recurvature NCEP performed relatively well in both statistical verification and some case studies. For slow bias after recurvature, CMC's bias was relatively smaller than other centres. These characteristics can show implications for better understanding of systematic errors in TC. JMA's forecasts performed higher detection rates of TCs over the several basins. This was presumably due to its model's resolution upgrade from 20 km to 13 km in 2023. Over half of the participating models have weak TC intensity biases in WNP basin over five-day forecast period. However, mean errors of the TC intensity have large variety among the participating models. the Met Office model tends to increase positive bias and the NRL model tend to enlarge negative bias as lead time progresses. These differences can be influenced by model's resolution, physics, earth system components. Out of all, NCEP model performed well in terms of both bias and mean absolute errors.

Considering rapid progress of Machine Learning (ML) based models, JMA, leading the TC verification project, has intention including ML based models to the WGNE TC verification for 2024. For efficient data processing, same protocols (e.g. meteorological parameters, data format and data provision policy etc.) would be adapted to ML based models as for physics based models. Call for participation is to be announced in the beginning of the 2025.

JWGFVR review activities related to WGNE (Caio Coelho, CPTEC/INPE)

The group's mission is to advance the development and application of verification methods to assess and improve the quality of weather and climate forecasts, working closely with WWRP, WGNE, and WCRP. Over the past year, the group has promoted good verification practices, developed new verification tools, and contributed to the revision of standards for operational verification.

Key highlights of 2023-2024 included the successful organization of the 9th International Verification Workshop (IVMW) and associated tutorial hosted by the South African Weather Service (SAWS) in Cape Town, May 2024. This event combined theoretical lectures with practical projects, and significantly enhanced capacity building in the field of forecast verification, particularly for participants from Africa, supported by the Climate Risk and Early Warning Systems (CREWS) initiative. A major outcome of this workshop is a special issue in forecast verification research that is open for submissions until the end of 2024.

A notable area of focus for JWGFVR was its work on verification for several high-impact projects, including the Paris 2024 RDP), the Aviation Research and Development Project (AvRDP2), and the Tropical Cyclone Probabilistic Forecast Project (TC-PFP). The group is also contributing to the verification component of the Implementation Plan of WWRP projects: PCAPS, SAGE, InPRHA.

In collaboration with WGNE, the JWGFVR has developed a Pilot Project for Forecast Verification over South America that focuses on improving forecast verification through two main streams: summary statistical verification across the South American domain and detailed analysis of extreme events, such as heat waves and destructive wind events.

The JWGFVR has also contributed to the revision of the WIPPS verification standards, particularly through the INFCOM Task Team on NWP Standard Verification (TT-NWPSV). This includes proposing new mandatory variables for verification, as well as expanding verification domains, reflecting INFCOM´s efforts to enhance verification practices across global and regional models.

Looking ahead, the JWGFVR plans to further its engagement with international verification projects in WWRP and WGNE/WCRP, continue to enhance verification practices across different domains, unify its web resources for broader accessibility, start planning for the next international verification workshop and tutorial, and keep supporting INFCOM and SERCOM for verification practices.

Discussion points:

- JWGFVR to participate in forthcoming discussions about new WGNE AI model forecasts intercomparison project to be developed jointly with WIPPS
- JWGFVR to consider recommendations of methods for verifying smooth outputs of AI model forecasts
- JWGFVR to keep WGNE informed about plans for next international methods verification workshop and tutorial

The Global Land and Atmospheric System Studies (GLASS) panel: Key research topics and new results (Volker Wulfmeyer, University of Hohenheim)

Body of the text

The AROBASE project: Towards a regional kilometer-scale multi-coupled modelling and forecasting system (Cindy Lebeaupin Brossier, CNRM)

The AROBASE project aims to assemble a fine-scale regional multi-coupled modelling system of the physico-chemical atmosphere, the ocean (including sea-ice and marine biogeochemistry), waves and land surfaces (soil, vegetation, cities, snow, lakes and rivers). The developed numerical system, named AROBASE, that combines AROME/ACLIB/SURFEX/CTRIP/NEMO/MFWAM, is built in the line of the coupled modelling systems developed and used at CNRM. It aims first to improve the understanding and representation of the exchange processes between the compartments of the meteorological and environmental system at the kilometer scale. It is also designed for high-resolution numerical weather prediction, which implies to guarantee its performance in the forecast mode and its relevance for monitoring and anticipating meteorological phenomena and their consequences. AROBASE finally prepares the convection-permitting regional climate model to be coupled to new components.

AI activities for NWP applications at CNRM (Sara Akodad, CNRM)

Machine learning and artificial intelligence techniques applied to numerical weather forecasting (AI-based Numerical Weather Prediction, AI-NWP) are rapidly advancing worldwide, both in large private companies and in national and international research and forecasting centers. AI-NWP offers a highly promising approach for the development of weather forecasts. Its applications are diverse, and AI technologies are gradually being adopted by all teams at CNRM, a pioneer in several areas such as increasing the number of ensemble members through AI and downscaling. The CNRM particularly focuses on developing a high-resolution (kilometric) forecast AI-based emulator in collaboration with partners internal and external to Météo-France.

EW4All survey (Ariane Frassoni)

This presentation outlines the WGNE's contribution to the WMO Early Warnings for All (EW4All) initiative through a systematic error survey. It highlights global and regional modeling centers' capabilities in forecasting the six hazards prioritized by 30 countries, WMO members, like flash floods, droughts, riverine floods, tropical cyclones, thunderstorms, and heatwaves. Forecasting timescales range from hours (e.g., for thunderstorms) to seasonal scales (e.g., for droughts). The survey results show disparities in hazard modeling approaches, with significant efforts focused on improving thunderstorms/squall lines and heatwaves, while few centres make efforts to forecast riverine floods. Examples include DWD's SINFONY system for seamless forecasting and CPTEC's advancements in convection and cold pool parameterizations.

Verification methods include spatial verification of precipitation and water levels, as well as the evaluation of issued warnings. The presentation also identifies region-specific hazards, such as winter-season phenomena (e.g., blizzards) and cold waves, underscoring the importance of tailored hazard forecasting. There is a need for collaboration across modeling centers to improve forecasts, bridge systematic errors, and enhance the utility of predictions for high-impact weather events globally. Participants are invited to contribute further to the initiative.

WGNE – WIPPS collaboration (David Richardson, WIPPS, ECMWF)

Collaboration between WIPPS and WGNE will ensure that WIPPS operational activities benefit from the latest developments in earth-system modeling for the benefit of WMO Members.

A current priority is to review and update WIPPS products to ensure they meet the needs of the Early Warnings for All initiative. The WGNE EW4All survey will help to identify where WIPPS can be extended to meet the prediction needs for the EW4All priority hazards, and where additional model development would be needed. More generally, collaboration with WGNE will help to realize the requirement to extend WIPPS to support a broader range of earth-system applications.

Another priority is to provide guidance to users on the use of AI-based forecasts. Working together with WGNE on an intercomparison of data-driven models and comparison of strengths and weaknesses compared to traditional NWP models would be very useful for WMO Members.

South America - Regional Model Verification Pilot project (Nils Wedi/Ramon de Elia)

This pilot project, initiated during WGNE38 in a joint meeting with members of the Joint Working Group of Forecast Verification and Research (JWGFVR), aims to improve our knowledge of global forecast performance over South America. Although this region is not rich in observations, it contains a large number of reliable weather stations with long records. Work has focused mainly on Argentina and southern Brazil, and has consisted of traditional statistical verification analysis as well as case studies of extreme events. To cover different spatio-temporal scales and weather forecasting challenges, we are analyzing a foehn case over the Andes (Zonda), a severe convective storm, and a record heat wave. Substantial progress has been made through monthly meetings between a good number of Argentinian/Brazilian scientists and forecasters, ECMWF staff, and members of the WGNE and JWGFVR. Results to date have shown that global models over South America perform with comparable skill to other regions of the world, and that the Extreme Forecast Indexes are an excellent tool for developing early warning products in the region. The Zonda case is also an excellent opportunity to test the capabilities of DestinE in steep terrain regions.

WGNE Blue Book (Sara Pasqualetto, ESMO IPO)

This presentation provided an update on the transition of the WGNE Blue Book submission system, including the shift of responsibilities from Elena Astakhova (former WGNE member) to the ESMO IPO. This change has resulted in an adjustment of the submission period from the traditional May timeframe to September/November, a schedule that will be maintained for future editions.

The current status of submissions has been reviewed, with an open invitation to further circulate the call for contributions to broaden participation. Discussions on the future of the WGNE Blue Book have also been highlighted, including a proposal by the IPO to upload future editions to Zenodo for improved archiving and citation capabilities. The group also considered expanding the allowed length of contributions from 2 pages to 4-5 pages to accommodate more comprehensive submissions.

Finally, strategies to enhance WGNE outreach efforts, including the development of newsletters and mailing lists, have been discussed, aiming to engage the broader community more effectively.

Joint Session 2 – Initialization

Pathway to sustained delivery of climate forcing datasets - conclusions of recent scoping workshop (Anca Brookshaw, ECMWF)

Body of the text

Ocean initialization for climate simulations (Gokhan Danabasoglu, NSF NCAR)

Ocean initialization approaches used in climate simulations remain rather ad-hoc with no standard or coordinated (best practice) methods. This is primarily due to practical reasons that include: computational and personnel-related resource limitations; development timelines; and delivery deadlines. Examples of ocean initialization methods include: starting from some observational potential temperature (T) and salinity (S) datasets, usually January-mean, and state of rest; starting from another simulation or a sequence of simulations – this could include starting from Ocean Model Intercomparison Project (OMIP) simulations; running ocean models with surface forcings from coupled simulations; for high-resolution simulations, starting from interpolated T and S fields from a low-resolution simulation with state of rest; and acceleration approaches and using coarser grids for tracers.

It was noted that ocean equilibration times at depth and abyssal ocean are primarily dictated by diffusive timescales which are several thousand years. So, performing long *equilibrium* simulations is not practical, particularly at eddy-resolving / permitting resolutions. We must therefore live with deep / abyssal ocean not being at equilibrium in most simulations, including those participating in Coupled Model Intercomparison Project (CMIP).

Initialization of ocean biogeochemical fields with their rather long equilibration times has its own challenges. Some acceleration techniques, such as Newton-Krylov methods, have been successfully employed to spin-up these passive tracers. Such acceleration methods may not be successfully applied to active tracers.

In coupled simulations, trends in the global-mean potential temperature in the ocean – usually used to gauge oceanic equilibrium – largely reflects that of the energy imbalance at the top of the atmospheric model. However, assessment of oceanic equilibrium just based on this single global-mean metric is rather misleading. This is because there are significant redistributions of heat both across basins and in the vertical even when the global-mean temperature trend is close to zero. It is also important to assess adjustments and trends in other fields of interest. For example, the Atlantic meridional overturning circulation maximum transport can take several centuries to equilibrate when a simulation is initialized from observational T and S.

Different initialization methods and, in particular, different lengths of pre-industrial control simulations in CMIP make the comparison of these simulations rather difficult for oceanic heat uptake and sea level changes (and possibly others). Clearly, a minimum integration length of 500 years for CMIP pre-industrial control simulations is not enough. Furthermore, the rather ad-hoc nature of creation of initial states for prediction simulations is also concerning. These challenges are even more so for high-resolution simulations. To document initialization- and spin-up-related information collectively in a manuscript – noting that such information is not necessarily provided in model description papers, a group of researchers, known as *The Wondering About Ocean Spin-up Team*, has been surveying modeling groups to get details of their ocean initialization and spin-up approaches, including for climate prediction simulations.

Improved initialization of the L-A system? A need to revisit the modeling of surface states and fluxes in coupled model systems (Volker Wulfmeyer, University of Hohenheim)

Body of the text

Initialising subseasonal to decadal predictions (Andrea Molod, NASA/Goddard Space Flight Center, Bill Merryfield, ECCC)

Body of the text

Joint Session 3 – Scale-aware parametrizations & Seamless prediction

Seamless prediction and predictability across weather-subseasonal-seasonal timescales (Priyanka Yadav, NASA/Goddard Space Flight Center)

Body of the text

On causality, sources of predictability and bridging predictions across timescales (Ángel G. Muñoz, BSC)

Body of the text

Aerosol aware / convection aware parametrizations (Georg Grell, NOAA/GSL)

Body of the text

Joint Session 4 – Model Processes Improvements

Teleconnections in CMIP6 (Clementine Dalelane, Deutscher Wetterdienst)

An application of the graph-theoretical analysis tool delta-MAPS was presented, which constructs complex networks on the basis of spatio-temporal gridded data sets, here sea surface temperature and geopotential height at 500 hPa. Complex networks complement more traditional methods in the analysis of climate variability, like the classification of circulation regimes or empirical orthogonal functions, assuming a new non-linear perspective. A number of technical tools and metrics (trend empirical orthogonal functions, distance correlation and distance multicorrelation, and the structural similarity index), borrowed from different fields of data science, are implemented in order to overcome specific challenges posed by our target problem.

Teleconnections and dynamical mechanisms of the 2023 south Brazil austral spring precipitation excess in observations and numerical experiments (Caio Coelho, CPTEC/INPE)

South Brazil experienced in austral spring 2023 a major precipitation excess event. This study aimed to understand what caused this event from a large-scale perspective including teleconnections and dynamical mechanics. The first part of the presentation focused on the diagnosis of the observed atmospheric conditions, and on the investigation of relationships between these conditions with sea surface temperature, atmospheric circulation patterns, dynamical mechanisms, and teleconnections. The second part of the presentation focused on the investigation of the contribution of tropical and extratropical atmospheric thermal forcings to the teleconnections and associated precipitation with linear baroclinic model numerical experiments. These analyses are relevant for advancing knowledge about the underlying mechanisms associated with such extreme precipitation events, and contribute to identifying model virtues and deficiencies in reproducing the involved teleconnections.

Evaluating the impact of Aerosols on the S2S (Ariane Frassoni)

This project, part of the WGNE-S2S-GAW collaboration, evaluates the impact of aerosols (biomass burning, dust, and pollution) on subseasonal numerical weather and air quality predictions. Using multiple models (NASA GEOS-S2S, ECMWF IFS, NOAA, and KMA), re-forecast experiments were conducted with prescribed aerosol emissions. The experiments focused on aerosol direct effects, with some models also exploring indirect effects. Key findings highlight improvements in predictive skill over specific regions (e.g., the USA, Canada, and South America) but mixed results globally. Differences in reanalysis datasets (e.g., ICAP aerosol reanalysis) and ensemble sizes across models impact the robustness of probabilistic analyses. Future steps include addressing these inconsistencies and emphasizing aerosol impacts in key regions like the Amazon. The study underscores the potential of integrating aerosols into subseasonal forecasting to enhance prediction accuracy and societal relevance.

Subseasonal-to-seasonal prediction of the boreal winter stratosphere (Javier García-Serrano, University of Barcelona)

Finally, the talk focused only on seasonal prediction of the winter stratosphere as there was another talk later in the session on subseasonal prediction (by Chaim Garfinkel). The study is mostly derived from the manuscript by Alice Portal et al. [Clim Dyn 58:2109-2130], and centred around three main questions: what is the current status of stratospheric predictability in seasonal forecast systems?; how do they represent upward troposphere-stratosphere coupling?; and, does wave-forcing initialization have an impact on the stratospheric polar vortex skill?. The study employs C3S seasonal hindcasts initialized on November 1st over 1993/94 - 2016/17. By analysing zonal-mean zonal wind at 10hPa for boreal winter (DJF), the results show that forecasts systems tend to be overconfident (underdispersive) at low latitudes, in the QBO region, while they are better calibrated at middle-to-high latitudes, in the polar vortex region. In the tropics, there is high predictability and prediction skill related to the QBO. At middle-subpolar latitudes, there is prediction skill close to the potential predictability. However, at subtropical latitudes there is no prediction skill at all, much lower than potential predictability; an issue worth exploring. The coupling between eddy heat flux at 100hPa [v*T*] and polar vortex strength at 10hPa is reasonably well simulated by the forecast systems, besides model biases in the polar vortex climatology. Troposphere-stratosphere coupling provides prediction skill for the polar vortex; v*T* shows high skill in November but marginal skill in DJF, mainly over the North Pacific (probably related to ENSO) while in reanalysis the key region appears to be Eurasia.

Assessing the impact of accurately representing river freshwater and turbulent fluxes on monsoon variability in a seasonal forecast model (Ankur Srivastava)

This talk focused on major model development efforts at IITM to improve the seasonal forecast skill of the Indian Summer Monsoon (ISM) at seasonal time scales. Major Indian rivers discharge freshwater into the Bay of Bengal (BoB), making the mixed layer shallow and thickening the barrier layers. This stable salinity stratification can impact convective activity at sub-seasonal to seasonal (S2S) time scales during the summer monsoon season. A global river-routing model was coupled (online) to the seasonal forecasting system (CFSv2) at IITM to model this aspect. Representing rivers in coupled models improves the S2S scale variability and the seasonal forecast skill.

The second development focused on improving the surface flux and SST representation in CFSv2 to account for Cool skin and Warm layer Corrections based on the COARE bulk flux algorithm. Diurnal skin temperature parameterization improves the phase (timing of maxima, minima) of SST and the mixed layer depth (MLD). It results in larger diurnal amplitude in SST, MLD, and rainfall. These diurnal scales feedback to intra-seasonal time scales and enhanced SST rectification is seen in the sensitivity run (3-10%) compared to the control run (2-5%) over the Indo-Pacific warm pool region. Reduced seasonal mean biases in rainfall, latent, and net heat flux, and the scale interactions with the diurnal and sub-seasonal processes result in better seasonal prediction skill for ISM.

APARC-SNAPSI study on the role of biases in the stratosphere on S2S predictability (Chaim Garfinkel, Hebrew University of Jerusalem)

Two-way coupling between the stratosphere and troposphere is an important source of subseasonal-to-seasonal (S2S) predictability and can open windows of opportunity for improved forecasts. Model biases can, however, lead to a poor representation of such coupling processes; drifts in a model's circulation related to model biases, resolution, and parameterizations have the potential to feed back on the circulation and affect stratosphere-troposphere coupling. The SNAP project (currently affiliated with APARC, and also the stratosphere sub-project of S2S until December 2023) aims to better understand these processes, with the hope of helping improve model skill.

This talk focused on one of the ongoing SNAP projects (Garfinkel et al., in press). We introduced a set of diagnostics using readily available data that can be used to reveal these biases, and then apply these diagnostics to 22 S2S forecast systems.

In the Northern Hemisphere, nearly all S2S forecast systems underestimate the strength of the observed upward coupling from the troposphere to the stratosphere, downward coupling within the stratosphere, and the persistence of lower stratospheric temperature anomalies. While downward coupling from the lower stratosphere to the near surface is well represented in the multi-model ensemble mean, there is substantial inter-model spread. Models with higher lids and a better representation of tropospheric quasi-stationary waves generally perform better at simulating these coupling processes.

SNAP is very interested in engaging with WGSIP and WGNE communities, and specifically with whatever follows on from the S2S project.

MJO-TF Update (Charlotte DeMott, Colorado State University)

Body of the text

Joint Session 5 – Bias Reduction

Tropical cyclone weak-intensity bias reduction through the semi-Lagrangian dycore (Ron McTaggart-Cowan)

The operational Canadian Global Deterministic Prediction System suffers from a weak-intensity bias for simulated tropical cyclones. The presence of this bias is confirmed in progressively simplified experiments using a hierarchical system development technique. Within a semi-idealized, simplified-physics framework, an unexpected insensitivity to the representation of relevant physical processes leads to investigation of the model's semi-Lagrangian dynamical core. The root cause of the weak-intensity bias is identified as excessive numerical dissipation caused by substantial off-centering in the two time-level time integration scheme used to solve the governing equations. Any (semi)implicit semi-Lagrangian model that employs such off-centering to enhance numerical stability will be afflicted by a misalignment of the pressure gradient force in strong vortices. Although the associated drag is maximized in the tropical cyclone eyewall, the impact on storm intensity can be mitigated through an intercomparison-constrained adjustment of the model's temporal discretization. The revised configuration is more sensitive to changes in physical parameterizations and simulated tropical cyclone intensities are improved at each step of increasing experimental complexity. Although some rebalancing of the operational system may be required to adapt to the increased effective resolution, significant reduction of the weak-intensity bias will improve the quality of Canadian guidance for global tropical cyclone forecasting. (Source: Monthly Weather Review)

Correcting weather and climate models with machine learned nudging tendencies (Oliver Watt-Meyer, Allen Institute for AI)

Body of the text

Adaptive tuning of uncertain parameters in a numerical weather prediction model based upon data assimilation (Günther Zängl)

The adaptive parameter tuning (APT), which was put into operations at DWD in several steps since 2022, uses time-filtered data assimilation increments to improve the estimate of uncertain physical parameter fields from external data. Important examples are physical properties of the soil (heat capacity and conductivity) and vegetation (minimum stomatal resistance, roughness length), which are basically derived from soil type and land use classifications. The main limitation of the latter lies in the fact that the available classifications cover only part of the natural variability of the fields, implying that varying the mapping parameters between the external data and the physical property fields has only limited benefit. Thanks to the APT, the RMS errors of 10-m wind speed, 2-m temperature and 2-m humidity could be improved by more than 10% at short forecast ranges, with gradually decreasing benefit at longer lead times when other error sources (e.g. arising from the synoptic-scale evolution) become increasingly important.

Online bias corrections at ECMWF: what do we gain? (Inna Polichtchouk)

Two approaches to online bias correction that are explored at ECMWF were discussed: 1) coupling a neural network that predicts large-scale flow-dependent model biases (estimated from nudging tendencies) to the IFS, and, 2) spectrally nudging the large scales of the IFS to the predictions of machine learning forecasts produced with AIFS. Approach 2) is a form of online bias correction as AI/ML NWP models are known to have a better large-scale skill than the physics-based models. The main advantage of the hybrid approach 2) over pure AI/ML driven models is that small scales are not damped and a better physical consistency is retained. Both approaches result in a significant improvement in forecast skill on the medium-range time scale, with approach 2) promising a greater skill improvement at a smaller cost than approach 1). Little positive impact was found from implementing approach 1) on sub-seasonal timescales, apart from for the QBO and MJO prediction. During Q&A questions regarding an operational implementation of approach 2) were asked and suggestions were given on how to improve the temporal damping of forecasts arising from the nudging approach used.

Online atmosphere/ocean bias correction in CanESM5 and its impact on seasonal hindcast skill (Bill Merryfield, ECCC)

Body of the text

Updates from partner/collaborative groups

WCRP Digital Earths Lighthouse Activity (DE-LHA) Update (Andrew Gettelman, Pacific Northwest National Laboratory)

Andrew Gettelman, the co-chair of the WCRP Digital Earths Lighthouse activity gave an update on current activities and interactions with WGNE. Specifically there is a joint km-scale (dx < 10km) model development group with WGNE, that meets every quarter or so to discuss current issues. There are also several different specific process teams working on advancing land modeling and convective processes at km-scale. Further discussions of joint interactions with WGNE were presented.

ACE: A fast, skillful learned global atmospheric model for climate prediction (Oliver Watt-Meyer, Allen Institute for AI)

Body of the text

ML/AI for NWP modeling activities at NCEP and the US Community (Fanglin Yang)

Efforts to develop NOAA-native AI models for weather forecasting are advancing, utilizing modifications to Google DeepMind's GraphCast code. The training process involves preparing NOAA datasets, customizing GraphCast for NOAA-specific grids and variables, and implementing multi-GPU training frameworks. Parallel initiatives by NOAA's EMC, NSSL, and PSL/GSL have yielded promising but varied outcomes. While the GFS remains the most consistent model, GraphCast demonstrated comparable performance, excelling in strongly forced environments but underperforming in weaker forcing scenarios.

AI-derived soundings appeared realistic but lacked detail, with models showing inconsistent daily performance compared to GFS. Regional adaptations of GraphCast successfully achieved high-fidelity emulations of WoFS forecasts at 3-km resolution, though deterministic models faced instability challenges. Future efforts are focused on building resilient models that address high-frequency noise and non-physical artifacts.

Key activities include training GraphCast with GDAS data for global deterministic forecasts (ML-GFS) and ensemble predictions (MLGEFSv1.0), emulating air quality models (AQcGAN), and applying machine learning for bias correction. Results are mixed: MLGEFS outperformed GEFSv12 in the southern hemisphere but lagged in the northern hemisphere. Bias correction using UNet effectively reduced temperature and CAPE forecast errors, addressing regional biases. Despite significant progress, additional refinement is needed to operationalize these AI models.

UK Met Office AI/ML activities (Tim Graham)

The Met Office has two large programmes on AI development. AI4NWP is aiming to develop and test ML based models for use in short range forecasting. This includes testing off the shelf models such as Pangu on case studies of particular interest for the UK and carrying out detailed analysis of physical processes such as the West Pacific Subtropical High. A large part of AI4NWP is focused on development of the FastNet model. This is being developed by the Met Office in collaboration with the Turing Institute.

The AI4Climate programme is in its early stages. Currently, there are two main strands to this work. The first is scoping options for ML based climate simulations and is considering questions such as whether it's better to have a single model for all applications or specific models for individual applications. A second activity is looking at the use of ML for regional downscaling of climate projections.

Finally, although not a Met Office activity, an overview of machine learning activities by the NEMO consortium was presented. Current activities included development of emulators of the NEMO model and ML learned physics parameterisations.

AI activities at ECMWF and proposal for AI intercomparison project (David Richardson)

Data-driven weather forecast models are a promising addition to physics-based numerical weather prediction (NWP) models. ECMWF now runs its own medium-range data-driven Artificial Intelligence Forecasting System (AIFS) in an experimental real-time mode. It is run four times daily and is open to the public under ECMWF's open data policy. Two methods have been developed to produce data-driven ensemble forecasts. They have both been found to be similarly skillful and match the skill of the physics-based IFS ensemble in a range of standard performance measures. ECMWF has made forecasts from one of the AIFS ensembles available as open charts under ECMWF's open data policy.

ECMWF is investigating the potential of data-driven ML forecasts trained and initialized directly from observations. Using historical measurements, the network learns correlations between observations from different sources, at different locations and at different times. Then from an input set of real-time observations the network can predict an observation of any type at any required future location and time. Although at an early stage, this Direct Observation Prediction (DOP) shows significant potential.

ECMWF has also developed Anemoi, a Python-based framework for creating machine learning (ML) weather forecasting systems. Named after the Greek gods of the winds, Anemoi is a collaborative, open-source initiative. Anemoi comprises an ecosystem of Python packages, which cover the full life cycle of data-driven modelling.

As well as its own AIFS, ECMWF runs several other ML models in real time and makes the forecasts available on the open charts web platform; verification results are also provided for all the AI forecasts, using the standard WMO verification measures. This provides a useful intercomparison of (some) current AI models and allows forecasters and other users to evaluate the practical usefulness of the forecasts. All these forecasts are run from ECMWF initial conditions and it would be valuable to have a more comprehensive intercomparison that could include e.g. forecasts run from different initial conditions as well as an investigation of appropriate verification measures for ML models. An evaluation of the practical usefulness of the ML models and their performance in forecasting extreme events would provide valuable guidance for forecast users and would help to guide the development and use of AI in WIPPS.