34th SESSION OF THE

WORKING GROUP ON NUMERICAL EXPERIMENTATION (WGNE-34)

Deutscher Wetterdienst, Offenbach, Germany

24-27 September 2019

Prepared by

Carolyn Reynolds and Keith Williams (WGNE Co-Chairs)

Michel Rixen (WCRP JPS)

WCRP REPORT 16/2019

WGNE34 attendees

TABLE OF CONTENT

WGNE MEMBERS: Carolyn Reynolds (Co-chair), Keith Williams (Co-chair), Oscar Alves, Elena Astakhova, Julio Bacmeister, François Bouyssel, Francois Engelbrecht, Ariane Frassoni, Ron Mctaggart-Cowan, Jian Sun, Masashi Ujiie, Nils Wedi, Fanglin Yang, Günther Zängl

WGNE EX-OFFICIOS: Caio Coelho (remotely), Michael Ek, Peter Gleckler (remotely), Daehyun Kim, Daryl Kleist, Ulrich Loehnert, Daniel Klocke, Marion Mittermeier, Peter Steinle

INVITED EXPERTS: Gerhard Adrian, Angela Benedetti, Simon Dadson, Peter Dueben, Greg Carmichael, Jens Christensen, Tatiana Ilyina, Sarah Jones, Jan Polcher, Michael Riemer, Cath Senior, Detlef Stammer

WCRP JPS: Michel Rixen

WMO: Eunha Lim, Paolo Ruti

1. Introduction

1.1 Objectives of the meeting – Carolyn Reynolds and Keith Williams

WGNE Co-chairs Carolyn Reynolds and Keith Williams welcomed all participants and thanked DWD warmly for hosting the meeting and their organizational support. Keith welcomed Carolyn in her new role as WGNE Co-chair and also 2 newly appointed WGNE members, namely Ron McTaggart-Cowan and Fanglin Yang, as well as Masashi Ujiie. A round table of introductions allowed all attendees to introduce themselves. Co-chairs recalled that this session was being held in the context of a recently approved WMO Constituent Body Reform and a recently approved new WCRP Strategic plan, offering WGNE a great opportunity to contribute to broader and more integrated Earth system numerical model developments around the weather-climate enterprise. The agenda of the meeting was designed to tackle these important challenges across our WCRP, WWRP and GAW research programmes.

The expected outcomes of the meeting were:

- to develop a roadmap for an expanded WGNE;

- to develop a 'Fast-Track' initiative in the context of the WCRP Implementation Plan;

- to propose concrete projects to bridge to new partners outside the typical WGNE remit;

- to discuss governance options in light of the new WMO Research Board.

The agenda was approved without further changes except some re-arrangements due to availability of some remote speakers.

1.2 Welcome by DWD – Gerhard Adrian and Sarah Jones

Prof Gerhard Adrian, WMO President and DWD President welcomed all participants, highlighting the important role of WGNE regarding atmospheric model development and quality assurance of weather forecast and climate predictions to serve WMO Members.

He noted the one to one correspondence between the WMO Strategic Plan 2020-2023 and the WMO Governance Reform towards a more efficient, effective and agile organization to ensure the overall coordination role of WMO across Members. Science has a special place in the new structure, with the newly established Science Advisory Panel and the Research Board to which WGNE will report, both bodies being now non-governmental in nature and open to experts beyond NMHSs in support of an integrated Earth system and science to service value cycle approach.

Sarah Jones, Chair of the WWRP Steering Committee, and Director for Research and Development at DWD highlighted the institutional seamless strategy over the last 10 years with the development of the ICON model, covering global and regional modeling, convection permitting ensembles where scores speak for themselves. She recalled the WGNE long history of achievements and contributions that build on critical model development and collaborative work, and includ earl career scientists and academia. Its future evolution is fitting in the new world of seamless Earth system prediction.

1.3 WMO reform – Paolo Ruti

The WMO Constituent Body Reform has recently approved the establishment of a new Research Board to replace and expand the scope of the former WMO Technical Commission on Atmospheric Sciences.

The Key priorities for the Research Board are:

- Science for Services exploiting a Value Cycle Approach
- Innovation in regions, strengthening capacity in less developed countries and SIDS
- Fostering an integrated and multidisciplinary research approach

To address those challenging topics, the non-governmental nature of the Research Board will include representation from a.o. NMHSs, WMO co-sponsored progs, World Meteorological Centers, WGNE, IOC, ISC, UNEP and maybe IAMAS, Developing Countries and Early Career Scientists.

Historically, single groups were dealing with single problems and the challenge in front of the WMO Strategic Plan is to bring more integration (value cycle, seamlessness, Earth system elements) and relevance through a regional lens. A seamless Earth system approach will require dedicated coordination on model development and associated process understanding.

Whilst the current WGNE DNA resides mainly in the atmospheric domain, the Earth system ambition of the new WMO Strategic Plan 2020-2023 requires stronger engagement with the ocean, hydrology, cryosphere, atmospheric composition and bio-geosphere communities.

The proposed expanded scope for WGNE will probably require some additional members or ex-officio members to fully represent this complexity whilst maintaining the technical expertise, which has characterized the working group since it was created more than 30 years ago. A gradual approach incorporating new members or additional ex-officios could be envisaged as members rotate off.

Boundaries with other modeling groups (e.g. S2S, WGCM, WGSIP) will have to be revisited as well.

1.4 WWRP overview – Sarah Jones

The World Weather Research Programme puts a strong emphasis on catalizing innovation in weather science to improve forecasts from minutes to seasons with a user perspective. Connections with operational partners and enhanced role of Early Career Scientists are critical in this context. Programme priorities emerged from the World Weather Open Science Conference held in 2014 with four societal challenges on High Impact Weather, Water, Urbanization and New Technologies and three core projects on resp. High Impact Weather, Subseasonal to Seasonal Predictions and Polar Prediction supported by a portfolio of regional projects. Amongst various Action Areas presented, Uncertainties and Extreme-scale computing and data handling were emphasized.

The discussion noted the importance of data handling, both from a research-operations GDPFS perspective and in support of smaller scale 'WGNE-like' projects. Machine learning is expected to support numerical modeling in a number of areas, including parameterization, data assimilation and post-processing.

1.5 GAW overview – Greg Carmichael

Greg Carmichael, Chair of the Global Atmospheric Watch (GAW) presented the elements of the programme value chain from observations to environmental services and its overall objective to improve predictions capabilities by incorporating atmospheric composition. Stronger collaboration between GAW, WCRP and WWRP is required to improve microphysics and representation of direct and indirect effects of aerosols to enhance model skill. Model verification represents a significant challenge because of representativeness of environmental observations, especially in urban areas. Air quality forecasting requires robust information on emissions at surface level.

1.6 WCRP overview – Detlef Stammer

Detlef Stammer, Chair of the WCRP Joint Scientific Committee recalled the co-sponsored nature of the Programme by WMO, ISC and IOC, which was recently reviewed to initiate a new Strategic Plan, now approved by all 3 sponsors, and which will be complemented by an Implementation Plan currently under development. The new Strategy will focus on the following goals: 1) understanding; 2) near-term prediction 3) long-term response; and 4) climate and society, all supported by critical infrastructures, including high-end computing and data management, as well as hierarchies of modeling capabilities and sustained observations and reference data sets.

The development of the implementation plan will involve revisiting current programmatic structures. This process will take about 2 years from now via a careful approach to preserve community engagement and current activities for the time being. Two workshops are planned for early next year to further develop science questions and identify the future elements of the programme.

It was noted that a number of activities can be carried out through close partnerships, in particular on modeling, observations and data (analysis) issues. Two Task Teams are currently being set up to advise the JSC on those matters.

Members suggested a broader modeling coordination mechanism across all WCRP, WWRP and GAW following the model of WMAC, which currently performs this role across modeling components of WCRP only.

CMIP will likely have to evolve to ensure the critical elements such as the generation of forcings, the infrastructure and data requests efforts are somewhat institutionalized to match the expectation on the service and policy side.

1.7 WGNE in light of WMO reforms – Michel Rixen

There is no logical scientific argument for separating the physical climate system from full Earth system science. We have now improved tools in the toolbox: seamless analysis and predictions across spatio-temporal scales and the Earth system. Despite tremendous progress in climate modelling, our ability to provide robust estimates of the risk to society is still constrained by numerous barriers to progress. Moore's law does not hold anymore and will require new approaches to exascale computing required for high resolution and larger ensembles to capture extreme events and uncertainties. Machine learning may play a growing role in this context, in particular to emulate complex process which we still do not fully understand. WGNE's future research will have to interface closely with operational infrastructures (observations, global processing and forecasting systems, etc) to rise to this challenge. This is in essence the aim of the WMO Constituent Body Reform: increase value for society through a seamless and integrated search system science and science for services approach, particularly relevant for a WGNE evolution in the context of the newly established non-governmental WMO Research Board.

1.8 WGNE in the context of GDPFS – Eunha Lim

Global Data-processing and Forecasting System (GDPFS) is a network of operational centres operated by WMO Members to make operationally available numerical weather and climate forecast, related products and services to WMO Members and relevant organizations. It is to evolve seamless GDPFS across all time scale, all disciplines and all Earth components through Earth System Modelling and Prediction. Establishment of the working mechanism to coordinate the development of Earth System Model and the linkage between research and operations was recommended at the World Meteorological Centres' (WMCs) workshop (Beijing China, 26-29 March 2019) and adapted the eighteenth World Meteorological Congress (June 2019) . WGNE was suggested as a basis to build the coordination mechanism at the WMCs workshop, so the status and plan of GDPFS along with suggestions for the coordination mechanism was introduced to WGNE members in order to allow them to help shape it.

The Earth system thinking does not imply that all models necessarily have fully interactive Earth system components. The GDPFS concept applies to all global, regional and national levels, including developing countries that may not have the capacity to run global models but can nevertheless contribute to the system in many ways, such as post-processing, verification, etc. There is an obvious benefit in also developing the GDPFS around a hierarchy of modeling suites.

2. Updates on WGNE projects

2.1 Systematic errors survey – Carolyn Reynolds

WGNE conducted a survey among modelling centers to determine the priority of the 17 major systematic errors produced as an outcome of the Fifth Workshop on Systematic Errors in Weather and Climate Models. Fourteen centers responded to the survey, contributing 35 surveys in total. While priorities naturally reflected the type of system being developed, several issues (convection, surface fields and fluxes, and microphysics) where of almost uniform concern and will serve as a guide as WGNE and other groups organize community efforts to address these pressing problems. There are already ongoing and proposed efforts to look at some of these biases and issues, including the WGNE surface flux inter-comparison project and the WGNE MJO Task Force examination of the ability of models to capture aspects of convective organization. Other groups are tackling related issues (e.g., the GLASS LIAISE project examining surface temperature over land, and the SOCRATES project examining microphysics over the Southern Ocean). A follow-up survey will be conducted in coordination with GASS to identify processes of major concern that likely lead to these systematic errors. WGNE's focus on the reduction of systematic errors is closely aligned with the WCRP strategic plan priorities as well as several WWRP implementation plan action areas. The WGNE systematic error survey summary may be found on the WGNE web site.

Members noted that the survey was atmospheric-centric and could deserve an update to shape the new WGNE boundaries in light of its future evolution. That said, it is the experience of many modelling centres that a good number of errors in other Earth system components can be traced back to the physical atmosphere.

Some WGNE projects already project on emerging priorities as well as onto the WCRP Strategic Plan and WWRP Implementation Plan.

It remains easier to identify systematic errors than their causes, which deserves an increased focus in WGNE. Coupling may address some of them, but may possibly create new ones.

CMIP6 includes a specific focus on identifying systematic errors. Early results suggest higher climate sensitivities than the ones estimated in CMIP5.

A recent workshop was held at ECMWF on developing observational campaigns to complement long-term operational observing systems to aid model development and improve forecasts. The collaboration with the UK Met Office to address systematic errors on the Maritime Continent was highlighted in that context.

2.2 Surface fluxes project – Carolyn Reynolds and François Bouyssel

The Surface Flux Phase 1 intercomparison is motivated by the fact that surface flux bias has been identified as an important and widespread issue for the modeling community (in, e.g., the WGNE systematic error survey). The protocol, disseminated in February 2019, asked for instantaneous and accumulated surface and near-surface fields and fluxes every 6 hours from daily 120-h global forecasts for July 2018 and January 2019 to be submitted by July 2019. Ten centers have participated (CMA, CPTEC, DWD, ECCC, ECMWF, MF, NCEP, NRL, RU, UKMO) thus far, with about 1.5 Tb of data in total being archived at Météo France. Météo France has also performed quality control, identifying issues such as missing fields and fluxes with the wrong sign convention, and has begun preliminary analysis. Initial analysis of zonally-averaged fields shows strong agreement among the centers for 2-m temperature, but substantial differences in surface evaporation, and surface short-wave and long-wave radiation. The next steps are to continue the analysis of the data set and engage with other groups (e.g., GASS, WCRP Surface Flux Task Team) on surface verification.

Members noted the absence of climate models in this initiative. Initial comparisons show some large discrepancies across models, which may find their origin in using different maps/ancillary files, as the drag project suggested. The importance of such initiatives in an Earth system context was highlighted and could be part of some proposal to e.g. the Schmidt Foundation.

2.3 WGNE-S2S-GAW Aerosols project – Ariane Frassoni and François Engelbrecht

The evolution of the offline to the online integrated modeling system was mentioned as the way forward to better represent the meteorology/climate/chemistry processes in numerical models. The improvement of the modelling capability with the inclusion of the atmospheric composition fully coupled with the meteorology/climate was cited as a common research goal of WCRP, WWRP and GAW programs. Therefore, the phase II of the WGNE-Aerosol project considers a joint-collaboration between WGNE, S2S and GAW to coordinate and implement the WGNE-S2S-GAW Aerosol project. The final version of the protocol was presented. The protocol includes the regional experiments focusing in the short-range forecasts for South America and South Africa, Egypt and East Asia considering the years 2016-2018. For the S2S timescale, global experiments with at least 32-day simulations, an ensemble approach, and a 2003-2018 hindcast period are considered. The forecast verification methodology follows the recommendation of the JWGFVR and includes deterministic and probabilistic statistical scores. The deadline for modelling groups to deliver the regional experiments is October-December 2019, while for S2S experiment the deadline is August 2021.

The focus of this project is on the aerosol direct effect but it was noted that the indirect effect via precipitation might also be important and is currently also one of the Action Areas for WWRP. Overall, members noted the benefit of having multiple centers doing off-line and interactive aerosols. It was recommended to possibly extend the simulation period to include 2019 so as to cover the massive biomass burning events which occurred in Amazonia this past year.

2.4 GASS report incl. joint WGNE-GASS drag and grey-zone projects – Daniel Klocke

GEWEX, the Global Energy and Water Cycle Exchanges Project of WCRP focuses on water and energy in the Earth System, with the mission to measure and predict energy and water flows in the earth system. The Global Atmospheric System Studies (GASS) panel is one of four GEWEX panels, with the mission to develop and improve the representation of the atmosphere in weather and climate models and contribute to the development of atmospheric models. Following a re-initiation of GASS in 2018, the panel currently consists of two chairs, six members and is still looking to grow. Four active projects, COORDE, Demistify, LS4P and diurnal cycle of precipitation project are all relevant to numerical weather prediction and are in the phase of producing first results. COORDE (COnstraining Orographic DRag Effects) is a joint project with WGNE focusing on understanding impacts of differences in orographic drag parametrizations for modelled circulation and tries to quantify drag from small-scale orography with the use of high-resolution simulations. Demistify is an LES and NWP fog modeling intercomparison to establish the ability of atmospheric models to simulate radiation fog. The goal is to identify key processes and the relevant level of complexity to adequately predict fog and its development. The Impact of initialized land temperature and snowpack on S2S prediction (LS4P) project investigates the impact of surface and sub-surface temperature anomalies in the high altitudes of the third pole (TP) on downstream precipitation predictions on seasonal time-scales and to quantify their impact relative to SST anomalies. First results indicate global impact on precipitation anomalies for months following temperature anomalies in the TP. For the diurnal cycle of precipitation project the experimental protocols are defined and first data submissions are expected soon. Further projects are in the setup phase. Specifically, a second phase of the 'Gray Zone' project (together with WGNE) with a deep and a shallow component and a project targeting physics-dynamics coupling are close to being finalized. GASS greatly profits from the collaboration with WGNE and is looking forward to continued exchange.

New GASS projects are selected after community consultation, white paper developments and volunteer contributions. It was suggested to associate WGNE, WWRP and WCRP closely in those consultations. It was clarified that LS4P is focused on how snow cover contributes to predictability, rather than on process understanding.

2.5 GLASS report – Mike Ek

The Global Energy and Water Cycle Exchanges Project (GEWEX) focuses on understanding, measuring and predicting water and energy in the Earth System. The Global Land/Atmosphere System Study (GLASS) role within GEWEX is to provide a better representation of the Earth System by understanding the role of land, and to improve modeling of land-processes and land-atmosphere (L-A) interactions. The PLUMBER effort is a land model (LM) benchmarking project to gauge whether models are performing as well as they can, given the information they're provided with, and if not, to identify practical avenues for LM improvement. Additional PLUMBER efforts will include the use of several hundred surface flux sites for PLUMBER-2 (versus 20 sites in PLUMBER-1), an Urban-PLUMBER, increased focus on soil processes, and adding a hydrological component to future PLUMBER studies. The SoilWat initiative focuses on improving soil and subsurface processes in models. The goal of the Local Land-Atmosphere Coupling project (LoCo) is to understand, model, and predict the role of local L-A coupling in the evolution of L-A fluxes and state variables, and the respective water and energy cycles, including clouds. Under LoCo, a number of L-A coupling metrics have been developed that should be used to assess whether weather and climate models are representing L-A coupling properly. Similarly, the diurnal land/atmosphere coupling experiment (DICE) examines L-A coupling in a single column model (SCM) setting via stand-alone LM and stand-alone SCM tests, followed by coupled LM+SCM tests, then examines the sensitivity of LMs and SCMs to variations in forcing; additional DICE case studies are being identified. In summary, there are several process-level studies in GLASS on LM benchmarking and L-A interaction that are critical for improving weather and climate models and are highly relevant to the WGNE community.

Machine learning could become increasingly useful in the GASS context to overcome lack of process understanding although members were cautioned on the risk of using such a 'black box' approach. Model hierarchies can support the gradual model complexity to transition a Single Column Model (SCM) into fully coupled model. GLASS may benefit directly from the WGNE Surface Flux project outcomes. The NASA decadal survey offers opportunities to convey observations requirements for the Planetary Boundary Layer (PBL).

2.6 Stochastic physics project – Keith Williams on behalf of Hannah Christensen

A protocol for a project to evaluate and inform the development of stochastic physics schemes has been circulated by Hannah Christensen. The project involved running km scale simulations, coarse graining them and applying to single column models and comparing their evolution. Hannah requested feedback on the proposal and this was collected at the meeting. It was also suggested that this might be combined with a project proposed to investigate the upscaling of forecast perturbations. It was thought that both might make use of the DYMOND global km scale simulations, which involves 9 centers running models at high resolution and for which the EU project ESYWACE supports the corresponding infrastructure.

2.7 Initial tendency project – Keith Williams

A protocol for a project to intercompare forecast initial tendencies, including the contribution from different physics schemes along with the dynamics, was outlined. Feedback on the proposal was received. There was also a suggestion for a project investigating forecast evolution from a common analysis and it was agreed that consideration should be given to combining these two projects.

ECMWF, UKMO, JMA, MétéoFrance, CMA, and possibly NCEP, CPTEC, DWD, NRL could contribute to this effort. Discussions included the use of specified vs original model levels, time steps.

2.8 MJO-TF report incl. progress with YMC exploitation – Daehyun Kim

Participants were briefed on recent research outcomes of the MJO Task Force. Currently, the research activity of the group is organized around four research themes: MJO simulations and mechanisms, MJO prediction, MJO and extratropics, and MJO-MC interactions. The group has found a robust relationship of the mean state moisture pattern with MJO simulation fidelity in climate model simulations and with MJO prediction skill in hindcast datasets. The group is collaborating with a team at PCMDI to implement several MJO metrics into PMP. MJO prediction skill in S2S and SubX models have been assessed. A standardized set of diagnostics and metrics that are designed to evaluate the model representations of MJO teleconnection is being developed and applied to model simulations. Several members of the group are using or have plans to use observations collected during YMC to better understand the MJO-MC interaction. Climate model simulations in which the land-sea contrast, topography, or properties of parameterized convection suggest that MC land masses affect the propagation of the MJO by altering the mean state moisture pattern near them.

Members suggested investigating the impact of initial conditions on MJO prediction skill. There is so far no evidence that going to higher resolution would improve skill further. Chi Dong is working on making The Year of Maritime Continent (YMC) data sets available for further exploitation.

2.9 Exascale review update (atmosphere and ocean model scalability & dwarf components) – Nils Wedi

Following the survey from last year on ocean modelling scalability efforts, this presentation summarized worldwide efforts, collected from WGNE members, in preparation for more scalability, for adaptation to emerging hardware architectures, for code infrastructure changes including I/O and for increasing Earth system complexity, and in preparation for exascale computing capability. This showed different levels of advances, ranging from more traditional MPI/OpenMPI/OpenACC code acceleration, to exploring domain-specific language (DSL) toolchains. Expectations towards effective acceleration of existing NWP and climate model codes range from 2-8 in the next 5 years.

2.10 Reduced precision tests – Nils Wedi

The presentation summarised the adoption of single or mixed precision in NWP operations or research among different members. While many had started efforts on mixed precision, the effort involving the model physics is substantial and several adapted double precision for this part, except one where the dynamics was treated in double but not the physics. Only a few have adopted reduced precision in operations but this is likely to change in the next two years with the majority of centres operating at least part of their code in reduced precision. Speed-up gains range from 25-40%, often resulting from better memory access/cache use and improved communication bandwidth.

2.11 Global model intercomparison with same initial conditions – Nils Wedi

A short presentation introduced the idea of a global model intercomparison project with initialized simulations starting from the same initial conditions to gain robust statistics and to better understand systematic errors in fully parameterized operational models. Linus Magnusson is leading this effort at ECMWF, providing initial conditions and a brief protocol. The subsequent discussion suggested a merge with the proposal on investigating tendencies from Keith Williams and Mark Rodwell, possibly selecting a period that overlaps with the Dyamond II protocol of initialised storm-resolving simulations (some coupled) covering the period of the EUREKA field campaign.

The question was raised as to possible biases related to own vs other’s initial conditions. Models are expected to cope with those initial – in principle well balanced – conditions. It would be up to each group to implement what they think is best system configuration (such as the land forcing). Output sampling through the diurnal cycle was recommended.

2.12 Developments in numerical methods – Michael Baldauf and Nils Wedi

A short summary about the this year’s “PDEs on the sphere” workshop in Montreal was given. As the most prominent topics were identified: the mixed or compatible finite element approach; now used by the UK MetOffice for their new LFRic project. Furthermore the spectral element method (continuous Galerkin) used e.g. by the US Navy for their planned operationalization of NEPTUNE around 2025 and for the Energy Exascale Earth System Model used by several US labs. Several talks have been given about exponential integrators, which is a topic since several years but seems not yet used/planned for operational models.

In the second part of the first talk, investigations at DWD about the Discontinuous Galerkin method as a future possible alternative dynamical core for the ICON model have been described. First results with a toy model for either 2D vertical slice Euler equations and for shallow water equations on the sphere have been shown.

ECMWF's progress in different areas of numerical methods development was summarized. Highlights included the improved wind-pressure relationship related to wave drag over oceans at high wind-speeds, the impact of higher-order interpolation and vertical resolution on lower stratospheric temperature biases, and the first real-weather comparison of IFS-FVM and IFS-spectral on the same grid with the same physics.

The need for sufficient vertical resolution for simulating the orographic drag was noted.

2.13 WGNE computing table – Günther Zängl

DWD will acquire a NEC Aurora vector computer during the next couple of years. As a result of the bidding process, preference was given to this machine with respect to CPU-based architectures because of a better power efficiency and because NEC offered a larger speedup factor for the given budget. New computer systems were installed at JMA, HMC (Russia), ECCC and NCEP in the last year, three of which are being delivered by Cray. A major upgrade of the operational model was made at NCEP by moving from the spectral GFS to the new 'NGGPS' based on the FV3 Dynamical Core. At ECCC, a major revision of the physics package was put into operation. A general observation is that several centers postponed or reduced their plans for model resolution upgrades compared to previous years.

Members thanked Günther for his continued support of this activity.

2.14 WGNE Blue book – Elena Astakhova

The Blue book is a nickname for the WMO edition “Research activities in atmospheric and oceanic modeling” published since 1992. The nickname is explained by the blue covers of its paper issues in 1992-2006. Only online versions are available since 2006. All electronic issues of the WGNE Blue book are on the WGNE Web site. The number of contributions was about 60-70 during the last years. Mostly contributions came from Japan, USA, and Russia. This year it is worth mentioning a variety of papers from USA and Japan dealing with upgrades of their prognostic systems, a nice description of the WGNE aerosol project, and a paper about new Canadian meteorological supersites in the Arctic. In the discussion it was suggested to change the title of the edition and its corresponding description on the WGNE Web site to reflect the WGNE focus on ESM. Publishing summaries of of WGNE projects in the Blue book was also suggested.

Members thanked Elena for her great work on the Blue Book and supported the idea to maintain this grey literature as a convenient format to report on emerging (including more controversial) topics, new model developments and system upgrades. Encouragement of ‘Earth system’ model development contributions was suggested.

3. Knowledge expansion areas for WGNE

3.1 Machine learning – Peter Dueben

Challenges of state-of-the-art weather and climate models were outlined and motivate the use of machine learning to improve these tools. To prove the great potential for machine learning in our domain, results from a couple of recent studies were presented that use deep learning to learn the equations of motion of the atmosphere, to emulate model components of weather forecast models and to enhance usability of weather forecasts. Main challenges for the application of deep learning to improve weather and climate predictions were reviewed, suggesting approaches to improve usability in the future.

Machine learning can help developing emulators to overcome limitations in parameterization and to speed up numerical codes. Members cautioned on the limitation of Artificial Intelligence to extrapolate beyond the training set. The Copernicus Climate Change Service is also looking into these developments, also in the area of model calibration and post-processing.

3.2 Review of machine learning activities in centres – Keith Williams

Contributions for machine learning activities specifically on model development and/or evaluation (rather than, for example, post processing or DA) were requested ahead of the meeting. Five centers provided contributions on this ahead of the meeting with a further 2 indicating work in this area at the meeting. There is significant activity in this area in some centers/countries. For model development it can mainly be split into parametrization emulation and optimization of model parameters. For the former, putting the trained neural net back into the full model seems to be universally a challenge because of the difficulties of calling Python from Fortran. Emulation of radiation seems to be particularly popular, however it is questionable whether this is the best candidate as current parametrizations are reasonably well optimized between cost and performance (at least on CPU chips).

Radiation code is usually computed on reduced grids, with 'all sky' Machine Learning. Large Eddy Simulations can be used if no observational reference exists, an approach adopted by NOAA. CPTEC is also using Machine Learning to emulate precipitation for seasonal forecasting. There may be broader lessons learned as to why Machine Learning research cannot always be easily transitioned into operations. Collaborations with AI institutes (e.g. in Japan) and with the private sector (e.g. ECMWF and NVIDIA) are developing. Other areas of interest for Machine Learning include statistical downscaling, where attention needs to be paid to stationarity assumptions.

3.3 Coupled Atmospheric Composition - Meteorology/Climate Modelling – Angela Benedetti

Atmospheric composition (AC) has been recognized as an important source of predictability at different time scales. Several meteorological centers are now incorporating increasingly complex physical/chemical processes and interactions between the different Earth components into the models.

At ECMWF for example, development activities have built a capacity to simulate and assimilate a variety of AC species in the Integrated Forecasting System.

Focusing on aerosols in particular, recent work has shown the potential of aerosols to enhance model skill at the subseasonal-to-seasonal (S2S) scales via the aerosol modulation by the Madden-Julian oscillation. The prediction of the aerosol fields at the monthly scales was also shown to be skillful.

Moreover, using prescribed observed fire emissions, the ECMWF S2S system was able to re-forecast accurately the surface temperature anomalies associated with the massive Indonesian wildfires of 2015 six months in advance. Some open questions remain related to the degree of complexity needed for NWP applications and cost/benefit considerations.

Some centers indicated an interest to run simulations to confirm results in Benedetti's recent paper. S2S groups have already been contacted to participate in those experiments. There is however no evidence of a direct benefit on MJO scores from including aerosols. Members encouraged experiments including the aerosol indirect effect.

3.4 Longer timescale ESM components (carbon cycle, chemistry, etc.) – Tatiana Ilyina

The objective of the Grand Challenge on Carbon Feedbacks in the Climate System is to understand how biogeochemical cycles and feedbacks control CO2 concentrations and impact on the climate system, including drivers of land and ocean carbon sinks, possible amplification mechanism and role of highly vulnerable reservoirs. The uncertainty in carbon cycle projections (>300 ppm) is comparable to differences across socio-economic scenarios. Southern Ocean is responsible for about half of the ocean carbon sink and dominates its multi-year variability which is still poorly understood, e.g. due to uncertain relative contribution of biological vs. physical processes. There is a fair global agreement between land carbon models and estimate from global carbon budget but large uncertainties remain at the process level, e.g. plant response to CO2 increase. Future projections show decreasing buoyancy driven carbon storage and increasing biology driven in the Southern Ocean. Future land sinks in RCP scenarios remain very uncertain. Near-term changes in emissions may not be detectable in atmospheric CO2 observations over several years due to natural variability and process uncertainty. Predictability of 2-3 years is expected in initialized ESM-based prediction systems for CO2 fluxes and variations in atmospheric CO2 growth rate. Improving coupling between existing components will enable the full spectrum of climate-relevant interactions within the Earth system.

3.5 Hydrology modelling – Simon Dadson

In order to advance our ability to predict the future availability of water resources and the risk of water related disasters under a changing climate, it is necessary to bring together the land surface and hydrological modelling communities. Simon Dadson gave a presentation to describe Hydro-JULES, a new research programme funded by the UK's Natural Environment Research Council. The Hydro-JULES programme will build a three-dimensional, open source, community model of the terrestrial water cycle to support and enable collaborative work across the research and academic communities in hydrology and land-surface science. An advanced terrestrial hydrological model will be generated that couples to the Joint UK Land Environment Simulator (JULES) and related models, including the UK Met Office Unified Model. This new five-year programme, is supported by NERC National Capability funding and will be delivered by the Centre for Ecology and Hydrology (CEH) in partnership with the British Geological Survey (BGS) and National Centre for Atmospheric Science (NCAS).

Hydro-JULES aims to address critical research questions in the fields of hydrology, land-atmosphere feedbacks, carbon and nutrient cycles, data science and integration with novel instrumentation and Earth observation technologies; quantify the risks of hydro-climatic extremes (e.g., floods and drought) in a changing environment to support long-range planning and policy decisions; improve hydrological forecasting using new sensors and modelling technology. The Hydro-JULES project covers topics in land-surface science and hydrology including: quantification of hydro-meteorological risks, using high-resolution climate predictions for hydrological applications, calculation the impacts of environmental change on evaporation, transpiration, and soil moisture, modelling flood inundation over large areas, representing anthropogenic interventions in the water cycle, and application of new techniques including Earth observation and data assimilation. The presentation will explore the scientific drivers for integrated land surface and hydrological modelling, outline the main elements of the structure of the work programme, and describe mechanisms for community engagement and collaboration.

Members suggested GLASS as a likely useful interface to the hydrology community. The land surface, as part of a fully coupled system, includes hydrology and plays a key role for both NWP and climate, and there are many opportunities to use new Earth Observations records from e.g. SMOS, SMAP, SWATH, etc. ECMWF is currently experimenting with an on-line land surface scheme and there is a need to include scoring for hydrology.

4. Other WGNE activities and collaborations

4.1 Report from DAOS – Daryl Kleist

The Data Assimilation and Observing System (DAOS) working group has experienced significant turnover in the past year and is looking to establish several new initiatives to contribute to the WWRP implementation plan. Recent highlights from the working group include the completion of three WWRP publications including: 1) coupled data assimilation recommendation paper, 2) guidance for performing observing system simulation experiments, and 3) international data assimilation symposium summary. Building on previous intercomparison efforts, the DAOS group is endorsing and participating in a Forecast Sensitivity to Observation Impacts (FSOI) study with the aim to expand to a near-real time effort and pilot project. The project is being led by the Joint Center for Satellite Data Assimilation (JCSDA) in the United States, and is leveraging cloud computing and private partnerships. The goal is to further expand the effort to allow for more general exchange of assimilation-based diagnostics. The working group is also in the early stages of planning for the next international data assimilation symposium, hopefully with connections to the WCRP-led reanalysis conference. Planning for a new effort focused on km-/convective-allowing resolution data assimilation experimentation is underway. Finally, DAOS has agreed to a joint effort to survey current practices and make recommendations on tropical cyclone initialization for global models.

It was suggested to explore a possible joint conference around data assimilation and reanalysis around 2021/2022. Opportunities for possible synergies exist between DAOS and the WGSIP work on initial shocks (unstable boundary conditions) and drifts. Other possible areas of collaboration include the initialization of Tropical cyclones and land coupled data assimilations.

4.2 YOPP and the Polar Prediction Project – Gunilla Svensson

A presentation on the activities planned for the consolidation phase of Year of Polar Prediction within the WWRP Polar Prediction Project was presented. The YOPPsiteMIP, which coordinates model output and observational data for supersites in the Arctic, Antarctic and third pole areas, is facilitating process evaluation of models to aid parameterization developments. A considerable effort has been made on developing file formats and semantics. Planned activities to benefit from the year-long Arctic expedition MOSAiC were also discussed. Modeling centers are already engaged in the activities and they are interesting for the GLASS and GASS communities. There will be possibilities for continued model intercomparisons after the conclusion of YOPP and PPP and WGNE could possibly be the host for these activities then.

WGNE could possibly contribute to YOPP, in particular in model verification, including on some extreme events during the Special Observing Periods (SOP).

4.3 HIWeather Project – Michael Riemer

This presentation introduced the WWRP core project HIWeather. After a brief general introduction and motivation, a HIWeather key concept – the warning value chain - was introduced. The warning value chain illustrates the inter-disciplinary and inter-agency challenges that arise when linking final decisions to initial observations and forecasts of the atmospheric state. These challenges motivate the inter-disciplinary approach taken in HIWeather. HIWeather activities performed by HIWeather task team members and deemed relevant to the WGNE was presented. Similarly, links of HIWeather to larger projects that may cover topics relevant to the WGNE were presented. In particular, latest results from the German Waves to Weather consortium on upscale error growth were presented. Finally, the potential for upscale-error-growth experiments at different operational centers to elucidate and inter-compare the generation of ensemble spread in the respective models was discussed.

Members suggested considering some collaboration with WGNE around convective permitting models on small domain, perhaps in the framework of DYAMOND.

4.4 Nowcasting and Mesoscale WG – Peter Steinle

WGNMR's primary modelling related activities are associated with rapid update, high resolution, limited area modelling, and merging the output from these systems with both nowcasting systems at one end, and larger scale global ensembles at the other end. We are also very active in building capability and capacity to use nowcast and short range prediction systems away from major numerical modelling centres, via demonstration projects and the like.

A current high priority area for WGNMR is advancing modelling of urban areas at resolution of around 100m. A key activity in exploring the current capabilities of such models is the 2024 Paris (and Marseille) Olympics Research Demonstration Project. This incorporates exploring the use of high density, 3rd party (citizen) observations.

There are a number of areas of mutual interest with WGNE, including

- The use of AI/ML both within modelling systems and to post-process data
- The generation of stochastic perturbations for ensembles that have appropriate relationships to both large and small scale forcing, large and small scale uncertainty (particularly in the boundary layer), and hence containing physically realistic high resolution relationships in space and time.

Given the value of regular or routine comparison of global modelling systems, WGNMR is also keen to explore mechanisms for comparing limited area models. With the protocols for standard comparisons with observations developed for YOPP and EUREC4A, and the move to consortia based modelling systems, each being used in many areas of the world, intercomparisons may becoming more feasible.

Members noted the challenges to get suitable observations at the city scale and to develop suitable surface schemes.

4.5 Report from SPARC focus on errors in the representation of the stratosphere and coupling with predictability of the troposphere - John McCormack

A description of recent research activities organized by SPARC (Stratospheric-tropospheric Processes and their Role in Climate) was presented to the workshop attendees. This description focused on work related to WGNE goals of improving model representation of the stratosphere and its coupling with predictability of the troposphere. Five specific SPARC activities were discussed: the Data Assimilation Working Group (DAWG), the Stratospheric Network for Assessment of Predictability (SNAP), Dynamics and Variability of the Stratosphere and Troposphere (DynVar), the SPARC Reanalysis Intercomparison Project (S-RIP), and the SPARC Gravity Wave Activity. Specific areas of collaboration between SPARC and WGNE were discussed, including: (1) Stratospheric data assimilation to produce analyses for model initialization and verification; (2) Treatment of both resolved and unresolved gravity wave forcing in models due to orographic and non-orographic sources. Group discussion pointed out that SPARC/WGNE collaborations could benefit from future joint workshops on subject areas relevant to both groups. It was noted that the leaders of the SPARC Gravity Wave activity are tentatively planning to hold a workshop on the subject of orographic gravity wave drag from the surface to the stratopause in late 2020.

The discussion highlighted the WWRP DAOS effort on coupled data assimilation and OSSE experiments. It was also suggested to investigate Gravity Wave Drag with high resolution simulation on 'hot spots' such as the Antarctic Peninsula involving NCAR, ECMWF, UKMO, etc.

4.6 WGCM update and CMIP6 status (incl. model development, HighResMIP and systematic errors) status - Cath Senior

Model simulations for CMIP6 are now progressing with rapid activity over the next few months as the AR6 timelines approach (December 2019 for submitted papers). Model outputs are now being served by ESGF from 21 institutions (45 models) and more will be made available over the coming months. The ESMValTool is now routinely applied to CMIP6 data as it is uploaded to ESGF. Results from HighResMIP show a significant reduction in some long-standing regional model errors. The recent CMIP6 Analysis Workshop held in Barcelona, 24-28 March 2019 brought together 250 people from 26 countries and involved representation from all MIPs and 25 modeling groups. Early results from CMIP6 indicate a higher climate sensitivity than in previous CMIP rounds. A perspective paper is currently being develop to report on these emergent results.

WGCM coordinates model intercomparisons to understand and predict natural variability and forced changes. Future plans for CMIP will aim at a distributed organization with separation of timescales between DECK and MIPs science questions. The CMIP essential infrastructure is currently delivered by volunteer efforts by the WGCM members, the CMIP Panel, the WIP (WGCM Infrastructure panel) and individual scientists in often partly/un-funded efforts. A recently approved WMO resolution will look into institutionalizing critical components of the CMIP enterprise.

The Grand Challenge on Clouds, circulation and climate sensitivity is currently finalizing two assessments on resp. Climate Sensitivity and Aerosol Radiative Forcing. The Grand Challenge on Carbon Cycle is making solid progress on decadal predictions.

The discussion noted the need for process oriented metrics on top of the standard ESMValTool.

There was a general agreement that WGNE could focus on Earth system model development across weather-climate, whilst WGCM would focus on the longer-term feedbacks and climate sensitivity (i.e. model user side mainly).

Systematic errors develop at early stages of models runs but have implications on all time scales. The next challenge but also opportunity for WGNE will be to move into the Earth system approach bridging current disciplinary boundaries. Component process studies will still be fundamentally required.

4.7 Timescales on which systematic errors develop – Julio Bacmeister

A series of Transpose-AMIP (TAMIP) experiments with the high-top version of CAM were described. These experiments were aimed at examining the growth of climate model biases. They were initialized once-per-day with ERA-interim reanalyses and run for 5-days. Precipitation, wind, wind stress and surface pressure biases were targeted. CAM6 was observed to have a significant spin-up in precipitation – with the global means during the first 24 hours 20% below the climatological mean for the model. After this spin-up phase (from Day 2 onwards) both global means and global RMS for precipitation exhibit little interesting evolution in CAM6 with TAMIP biases closely resembling climate biases. By contrast wind, stress, and pressure biases exhibit evolving magnitudes and patterns in TAMIP up to Day 5. The TAMIP runs will be extended to 15-30 days to fully characterize their evolution.

It was noted that systematic errors are comparable to observational uncertainties in many cases. Members suggested perhaps running simulations from various reanalyses.

5. WGNE member presentations

Presentations in this part of the agenda were requested to focus on fixing model deficiencies and/or coupling to other ESM components.

5.1 Reducing systematic biases in NOAA's next generation global prediction system – Fanglin Yang

In June 2019 the US National Centers for Environmental Prediction (NCEP) implemented its first version of the next generation Global Forecast System (GFS) into operation, which is based on the finite volume cubed-sphere (FV3) dynamical core and an advanced microphysics scheme. While this model showed superior performance to its predecessor in many aspects of the medium-range weather forecast, a few systematic model biases remain to be resolved. NCEP is currently developing its next version of the GFS for operation in 2021, which will have a doubled vertical resolution and extended model top to the mesopause. A few major physics components will be updated as well. Reducing systematic model biases has been one of the foci in the development process of this model. The use of a new scheme that parameterizes model unresolved non-stationary gravity wave drag largely reduced temperature and wind biases in the upper atmosphere and improved the simulation of the Quai-Biennial Oscillation. Previous versions of the GFS often failed to capture the planetary boundary layer (PBL) inversion and missed the forecast of frozen rain in cold seasons. Model background diffusion coefficients, which were prescribed values and vary with height, are now made to be a stability dependent function. This modification greatly improved the simulation of the PBL inversion. An updated cloud-radiation interaction scheme also helped to reduce the cold biases found in the lower troposphere in winter. A scheme is being tested to spin up the land surface model with observed precipitation before extended model forecast starts to provide better land initial conditions. This is aimed to reduce forecast biases in land and near surface fields such as soil moisture and 2-m temperature.

It was noted that precipitation might be too smooth over Europe, but not related to orographic processes in the model. Regarding tropical cyclones and hurricanes, resolution and coupling to ocean/waves play an important role.

5.2 Climate sensitivity in CESM – Julio Bacmeister

Equilibrium climate sensitivity (ECS) in CESM2 is close to 5.5K, a substantial increase over the value of 4.2K that existed in CESM1. This increase in ECS was accompanied by changes in the time-evolution of fully-coupled abrupt 4xCO2 increase experiments. Sequestration of heat into the deep was found to be stronger in CESM1 leading to a slower approach to equilibrium in 4xCO2 runs. A consistent analysis of regional contributions revealed a dominant contribution from Southern Ocean shortwave forcing to both the global climate sensitivity and in the time evolution of the 4xCO2 simulations. Finally, an analysis of CESM2 4xCO2 experiments with re-tuned or modified physics was presented. ECS as low as 4.2K is attainable with physics essentially as in CESM2. Our analysis suggests that tuning of low clouds to obtain radiative balance in pre-industrial control runs may be an important factor in determining ECS/TCS.

Members noted that climate sensitivity is not an emergent constraint *per se*, but can be derived from a suite of simulations as an emergent pattern from which we can possibly derive some conclusions or at least hypothesis, about the climate model response, which cannot be easily supported by any observational record.

5.3 Reducing systematic errors in DWD models – Günther Zängl

Several improvements at the level of the physical process description were made to the cloud cover scheme in order to improve the prediction of cloud cover and radiation. In addition, a simple parameterization for solar radiation reflected by cumulus clouds was introduced to better describe the diffuse fraction of radiation.

- In the convection scheme, a limiter for convection overshooting over the tropopause was introduced in order to reduce the moist (and cold) bias in the lowermost stratosphere. This change was also successfully tested in IFS (by Peter Bechtold).

- Ongoing work comprises the usage of higher-resolved and more accurate external parameter datasets for, e.g., soil and vegetation properties, the implementation of a canopy-layer scheme, a scheme for turbulent orographic form drag, an upgrade from the RRTM radiation scheme to ECRad (in collaboration with ECMWF), and coupling ICON to a wave model

The discussion covered the cold bias in the tropical region, the use of several aerosol schemes of increasing complexity and the importance of using appropriate soil maps.

5.4 JMA activities for reducing systematic errors – Masashi Ujiie

Three topics on JMA's recent activities for reducing systematic errors were presented. In the presentation, refinement of subgrid orographic drag and its impact on synoptic-scale circulation over the East Asia were shown. The drag refinement reduces a long-standing "shallow trough" bias over Japan in the JMA's global model and is planned for introduction into the operational system in FY2019. In the second topic, the importance of recognizing systematic differences in analyses and careful interpretation in verification against self analysis were shown through JMA's recent experience of data assimilation experiments. In the third topic, the usefulness of precipitation verification scores, particularly bias scores, plotted in the WMO LC-DNV website was shown. The monitoring and comparison of the score help us detect impacts of model improvement to physics parameterizations.

Members cautioned about drawing conclusions from verification using one's own reanalysis when investigating systematic errors.

5.5 Commodore Ocean – Nils Wedi

A short introduction to the COMMODORE ocean modeller group, and the plans to initiate a process-level comparison of different ocean models, benchmarking against analytic, laboratory and observational data, was provided. There is a workshop taking place during 28-31 January 2020 in Hamburg to discuss further details. WGNE seeks a group like GASS or GLASS that brings together a similarly focused group for process-level ocean model development, and COMMODORE may be such a group, but further discussion may be required to establish such a link, also in view of other ocean groups (e.g. GODAE OceanPredict, the CLIVAR OMDP, the CMIP/OMIP).

Members discussed the possibility to endorse such initiatives and suggested a more coordinated approach to model verification across Earth system components. Consulting with IOC/GOOS to identify suitable ocean model development groups with which WGNE could engage was suggested.

5.6 Ocean numerics on the Uniform Jacobian cubed sphere – François Engelbrecht

Francois Engelbrecht, Professor of Climatology at the Global Change Institute (GCI) of the University of the Witwatersrand (Wits) in South Africa, reviewed the progress made towards the development of the first African-based Earth System Model. The model is formulated on the Uniform Jacobian cube-based grid (UJ-grid), and uses a cube-based global atmospheric model of the Commonwealth Scientific and Industrial Research Organisation (CSIRO) as the atmospheric component. The land-surface component is the CSIRO Atmosphere-Biosphere Land Exchange model (CABLE). The model uses as the ocean component a new cube-based ocean model developed at the GCI in South Africa. The ocean biochemistry used in the model is being developed by the Council for Scientific and Industrial Research (CSIR) in South Africa. Engelbrecht demonstrated that use of the UJ-grid significantly reduces grid-imprinting problems associated with alternative forms of cube-based grids. He further demonstrated that the stretched-grid capabilities of cube-based grid systems facilitate the flexible development of regional Earth System Models. He emphasized the need for an increased focus not only on global, but also on regional Earth System Modelling, and pointed out that a flagship project along these lines is currently underway in South Africa under the auspices of CORDEX (Coordinated Regional Downscaling Experiment). The project uses the Africa-based Earth System model as regional model to simulate coupled ocean-atmosphere-land-chemistry processes over the western parts of southern Africa. Engelbrecht finally showed some regional paleo climate modelling experiments using this regional modelling system, and pointed out that the simulation of paleoclimate poses new tests of the robustness of Earth System Models.

Members commented on the nice alignment of this hierarchical modeling approach with the WCRP Strategy and how this can also contribute to regional science.

5.7 Numerical modelling at the Hydrometcenter of Russia – Elena Astakhova

Several results were briefly presented.

(a) For objective estimation and stochastic modeling of model tendency errors due to unresolved scales, tendencies of the (low-res) model in question were compared with tendencies of a high-res model started from the same initial conditions. When the low-res model starts from coarse-grained high-res data, imbalances arise that mask the model error. To mitigate the problem and get stable model-error estimates, it was suggested to split the model error into two components, which are forecast differences between 1) two high-res runs starting from high-res and smoothed high-res data and 2) a high-resolution and a low-resolution runs started from the low-res initial conditions. The resulting estimates of tendency errors for COSMO with 2.2 km grid spacing compared with COSMO with 220 m grid spacing are presented and will next be used to train a spatio-temporal model-error model (to be used in convective-scale ensemble prediction and ensemble data assimilation);

(b) A model-error perturbation scheme named AMPT (Additive model-error perturbations scaled by Physical Tendency) for limited-area models was developed based on the previously developed stochastic pattern generator (SPG). AMPT was successfully tested with a COSMO-based EPS, showing better results as compared with SPPT.

(c) A new MACv2 aerosol climatology was implemented into the COSMO model. A four-month comparison of 24-h forecast with Tanre, Tegen, MACv2 and CAMS reanalysis demonstrated a strong effect of aerosol climatologies on 2m temperature and the good skill of MACv2. (d) Application of a new cloud-aerosol-radiation scheme (mostly changes in microphysics) in COSMO resulted in a substantial improvement of T2m forecasts in the Moscow region. (e) Upgrades of medium-range and long-range prognostic systems based on the SLAV model are under testing.

Members noted the challenge regarding model physics and associated vertical-horizontal resolution dependencies when moving to higher vertical levels. The model is currently uncoupled.

5.8 Addressing the cause of large-scale circulation errors in the Met Office global model – Keith Williams

Work to investigate a particular circulation bias was presented, namely a large high pressure bias over the Arctic. It was found that an overall reduction in low-level drag was required, however to achieve positive performance on all timescales a redistribution of the drag was needed with an increase in plant effective roughness lengths and a reduction in drag from sub-grid orography. The existence of a cross-timescale modelling system, perturbed physics ensemble and comparison with another model displaying an opposite error all proved useful. It was pointed out that developing solutions to systematic errors is rarely straightforward and often involves dealing with a number of systematic errors. The impact on performance is normally detrimental until a package of changes to address these compensating errors can be developed.

Members noted the need to interpret carefully the Greenland bias and the importance of validating a change after being at least neutral on both NWP and climate runs.

6. Verification

6.1 Tropical cyclone verification – Masashi Ujiie

TC verification for 2017 was presented. Position errors in participating global models over the Western North Pacific area tended to be smaller than those in 2016. However, this seems to be due to year-by-year variation rather than model upgrades. In the long term trend, all the centres certainly have reduced position errors in all areas. The intensity verification showed that the global models still tend to represent shallower TCs than those of the best tracks. As exceptions, recently, several centres such as ECMWF, Met Office, and NCEP can represent deep TCs and sometimes over-deepen them. This is presumably due to recent model upgrades. The TC verification of regional models was summarized as follows; (i) TC positions in regional models is constrained by their driving global models, and (ii) regional models have the potential of providing additional information relevant to TC intensity.

Members noted that results may depend on resolution and data assimilation methods, where some collaboration with DAOS on Tropical Cyclone initialization could be pursued. They suggested also complementing the verification with false alarm rates, which is something that the WMO Tropical Cyclone programme could pick up or at least support, and which would require centers to derive tracks in the same way. Difference in model behavior may have some origin in the data assimilation methods.

6.2 JWGFVR report including progress with process-orientated and surface verification – Marion Mittermaier

The Joint Working Group is involved in a number of verification projects and activities. YOPP provides a framework for analyzing current practices in the Polar Regions, reveal issues and investigate solutions, in particular on sea-ice prediction and NWP process evaluation during the Special Observing Periods and associated supersites. Likewise, the S2S database is being used for hindcast quality assessment and the plan is to transition that into operations via the Inter-Programme Expert Team on Operational Prediction on Sub-seasonal to Longer Time Scales. In cooperation with the WWRP HIWeather project, new forecast verification metrics are being developed using non-traditional observations. Examples of specific efforts were given related to HIGHWAY training, surface temperature scores and precipitation verification (SEEPS).

6.3 CMDP – Climate precipitation panel report – Peter Gleckler

Peter Gleckler (PCMDI, USA) reported on a recent workshop striving to establish an objective framework for benchmarking a diverse suite of precipitation characteristics in ESMs. The workshop, supported by the US Department of Energy, was inspired by discussions that Christian Jakob (U. Monash, Au; ex WGNE co-chair) and Peter Gleckler had led during previous sessions of both WGCM and WGNE. Attendees of the precipitation workshop included a diverse group of modelers and analysts, including, for example, members of the GEWEX Data Assessment Panel (GDAP) that are currently working on a report evaluating available gridded observationally-based precipitation products. Two teams were formed as a result of the precipitation workshop, one focusing on establishing a limited set of characteristics for benchmarking, and a second to identify precipitation-related characteristics and processes that are actively being researched but to date lack widely adopted measures for established benchmarking. The benchmarking team is beginning to implement the objective performance tests into a common analysis framework, and both teams are working to document simulated precipitation in CMIP6 and earlier generations. By the next WGNE session it is expected that the ground-work will have been completed to consider possible next steps, namely, how to challenge the modeling community to use this benchmarking resource as a guide to improve models.

In a general discussion, potential synergies between the JWGFVR and ESM precipitation benchmarking were considered. Attendees welcomed the proposed metrics. They also noted the sensitivity to resolution when it comes to verification, and it was agreed that this should be an issue to stay in the foreground of the benchmarking effort. It was noted that previous efforts to establish linkages between NWP and climate benchmarking were challenging, but that the focus on precipitation was a solid target that could be beneficial to both efforts. Marion Mittermaier and Peter Gleckler agreed that the ongoing precipitation assessment of CMIP6 models was a good opportunity, and that the distribution of precipitation was one example identified for potential collaboration.

7. Strategic directions for WGNE

7.1 High ranking systematic errors and projects required to address these - Ron Mctaggart-Cowan

Background: following the WGNE Systematic Error Workshop and a subsequent survey, WGNE has identified a number of key priorities to address systematic errors, in particular:

- *Convection: Joint WGNE-GASS grey zone phase II project and EUREC4A and GATE III field campaigns, WGNE MJO-TF focus on exploiting YMC observations*
- *Surface fluxes new project with initial focus on the ocean*
- *…*

For additional information, see:

- *https://www.wcrp-climate.org/JSC40/12.7.%20WGNE_for_JSC40_0519.pdf*
- *https://www.wcrp-climate.org/JSC40/12.7b.%20WGNE_Systematic_Error_Survey_Results_20190211.pdf*

Based on the results of a 2018/19 WGNE survey, the perceived leading source of uncertainty in NWP is the representation deep convection either through a parameterization or via poorly resolved processes in higher resolution models. The Eurec4a and GATE MIPs will address problems related to shallow and deep convection, respectively.

The representation of surface processes is also a major challenge for NWP, particularly in relation to the accurate depiction of the diurnal cycle, energy budget, momentum budget and hydrology. Surface/subsurface schemes are complex because of the multiplicity of processes that they must represent, a problem compounded by the current aggregation/blending techniques that are used to convey boundary information to atmospheric models. The WGNE Surface Flux and GASS Drag MIPs are actively addressing some of these issues.

The impact of aerosols on NWP is considered to arise primarily through their direct effects on radiative transfer. Complete aerosol/chemistry schemes are complex and computationally expensive, making it challenging to implement them operationally. It remains unclear what level of complexity is required for NWP, and whether a more complete treatment of aerosols is required in the S2S context. Some of these questions will be considered by the Aerosol MIP.

Identifying the sources of errors in NWP systems is a difficult task, particularly as more model components are coupled together. The elimination of compensating errors is an important objective, one that may be addressed using a hierarchy of models in which elements of the problem can be constrained by observations. Although there remains no systematic methodology for identifying error sources in models, the use of conditional and processed-oriented verification techniques may be of significant benefit in this regard.

Representing model error in a physically consistent way remains another important challenge for NWP developers. In particular, relying on the natural upscale growth of uncertainties that appear at the convective scale – as represented by stochastic parameterizations – appears to lead to an under-estimate of model error. A number of MIPs to address related issues have been proposed (SCM-based, common-initialization and upscale error growth), but there was no clear WGNE consensus on how these efforts should be combined to reduce the load on participants.

The WGNE survey was an extremely useful tool for discussions on systematic errors. A follow-up survey consisting of additional questions and a request for free-form input about systematic errors and centre/institutional priorities for model development will be drafted. The responses from the new survey will help WGNE to track progress on these issues and to identify common themes in the field that may be addressed by collaborations or MIPs.

The expected evolution of WGNE suggests revisiting the survey as a number of systematic errors have their origin in poor coupling/interfacing across components. The new Earth system approach will offer a number of opportunities to address some long standing systematic errors.

Some examples were highlighted in that context:

- how to translate uncertainty in hydrology, CO2 etc back on the model
- indirect effect in aerosols and microphysics (measurements from Australia/New Zeeland available soon)
- drag project: need for a reference across the PBL, SSO/TOFD could be part of it but may require higher resolution to resolve blocking

7.2 WGNE's role in Exascale/stochastic physics/machine learning – Daniel Klocke

Background: see dedicated sections in this report and also https://www.wcrp-climate.org/JSC40/10.4_FastEmergingTopics-JSC40_PKabat.pdf

Weather and climate model codes are about 5% efficient on current computing architectures and new computing technologies need changes to the current codes to use them efficiently. The challenge is to exploit competing architectures (CPU, GPU, vector) without losing flexibility, portability, efficiency of the code and efficiency in further development of the mode. One approach is to separate the concerns of machine specific requirements and the other model parts, but it is unclear, if this approach is practical and who will take responsibility to maintain the code parts which are machine specific. Most centres use precision reduction in some parts of their models to increase efficiency. Efficiency optimisations in terms I/O were little discussed and in terms of data assimilation many centres rely on the development of JEDI.

One possible way to exploit new technologies could be machine learning. Most groups have a machine learning project, or are at least involved in one. There is a clear role of machine learning for data assimilation, parameter estimation and verification. With all the activities and excitement, the usefulness for still needs to be demonstrated in terms of weather and climate modelling and concerns remain, that rare events, or extremes (very valuable information from forecasts) will be hard to handle. To fully exploit machine learning techniques, the data storage and computing infrastructure would need to change.

Stochastic physics are important to represent model error and in modelling correct error growth with forecast time. They play an important role in the spread of ensemble prediction systems. At higher resolutions, e.g. in the convective grey-zone, stochastic physics can be important for the correct representation of physical processes.

WGNE's role is to monitor and communicate relevant developments in HPC technologies, help to set new standards and coordinate the weather and climate community to optimally exploit future

HPC technologies. Members discussed process understanding in the context of Machine Learning which could also be trained on observations. WGNE could develop some recommendations on this. It was suggested to develop a MIP around machine learning. Connecting with industry could also offer some opportunities.

7.3 Ideas for "Fast Track" initiative(s) on model development- Ariane Frassoni

Background: some emerging WGNE topics on Earth system modelling, exascale modelling, machine learning, etc could be attractive to a number of WMO members, donors, sponsors, foundations, agencies. What would be WGNE's priorities, recommendations, and role?

The presentation around the "Fast Track" initiative(s) on model development key specific question in the dedicated discussion highlighted some important topics addressed during the previous days of the WGNE meeting. For the topics related to computer architectures, a suggestion to run chemistry models on emerging computer architectures like FPGA was mentioned as chemistry models are computationally expensive. One important subject mentioned during the discussions was the practical transfer of Machine learning (ML) research to operations. Centers that can apply ML for seasonal forecasting in parallel with standard system and compare results will add value in the knowledge of ML practical application. Seasonal forecasting is a possible way to benefit from ML due its intrinsic low predictability specially over tropics. Experiments performed with coupled models versus prescribed SST in addition to a set of the same initial conditions (IC, in an ensemble approach) can help identify anomalies in subseasonal and seasonal forecasts. High impact weather is a challenge for numerical prediction. To benefit from the WGNE experiment prescribing the initial conditions (experiment to be led by ECMWF), a model intercomparison focusing in a set of important strong weather events identified in the aforementioned experiment can be an avenue to better understand one of the sources of model errors.

Members suggested a typical time limited effort of 3-5 years for those initiative, which shall focus on emerging topics attractive to donors with an appropriate societal/sales pitch. Ideas could range from a reasonable level of maturity towards a clear outcome to exploratory 'blue-sky' research with a suitable high risk/high return ratio. The area of seamless Earth system modeling combined with Machine Learning with a focus on extremes and uncertainty (verification, stochastic physics) was proposed as a good candidate.

7.4 WGNE role in promoting large international modelling efforts and associated partnerships - Peter Steinle

Background: WGNE, as the international body to coordinate numerical experimentation across weather and climate communities, could play an active role in strategizing, leading and/or supporting major flagship projects like Extreme Earth (https://www.extremeearth.eu/), not only at the implementation level, but also in co-designing calls for agencies to support them or elements thereof (see for example https://schmidtfutures.com/our-work/scientific-knowledge/vesri/).

There have been two recent potential funding opportunities to mobilize resources for key research, and such opportunities are likely to rise again. There has also been a recent statement from WMO Congress on the need for Public-Private Partnerships: https://public.wmo.int/en/our-mandate/how-we-do-it/partnerships/Public-Private-Academic%20Sectors%20Engagement.

It is also expected that the WMO Research Board will have the role of promoting and/or endorsing these funding programs, and will be discussing how to mobilize resources. While the group was aware of several issues such as ethical considerations and the ability of government agencies to receive private funding, there are other fields where this is successfully managed, such as medical research.

WGNE is well positioned to

1. Be a primary point of contact for funding agencies for advice on high priority research areas that can be addressed on the order of a few years
    a. To enable a quick response to such calls for funding, WGNE should have groups of major topics to align with likely interests of these new major funding agencies
    b. Target areas should range from addressing major challenges for current modelling techniques, to enabling emerging research areas to become better established.
2. Setting standards for assessing the value of major developments
3. Assisting with protocols to better enable the wider analysis and use of the research.

Allied with this issue is how major commercial data providers can be included in the research activities of interest to WGNE: They are welcome to attend as observers provided research and results are subject to independent review in a consistent manner to the information from current WGNE affiliates.

Some members cautioned about being funded by foundations, some of which are closely associated with the private sector and not risk-free. This question will also be addressed by the Research Board, as a follow up to the corresponding Statement which came out of the recent World Meteorological Congress. It was commented that industry is now issuing operational forecasts but yet, they do not want to be part of – or are not interested in - any international evaluation exercise. WGNE has an important role to play in setting the verification standards in that context, as well as advising on project proposals when they have a large multi-national footprint.

7.5 Roadmap for a WGNE evolution on Earth system modelling – Nils Wedi

Background: following the WGNE33 session discussion and in preparation for JSC40, WGNE had developed a proposal for an evolution of its scope/mandate in the context of the WCRP New Strategy and WMO Constituent Body Reform

See https://www.wcrp-climate.org/JSC40/12.7c.%20WGNE_WCRP-CASreform_positional.pdf

The range of discussions and comments made during the week on the WGNE evolution on Earth system modeling were summarized, and how this roadmap would fit into GDPFS (blue part of the WWRP "squid"). A vision statement to widen the perspective of WGNE was presented to include a wider range expertise of Earth System Model components in the Working Group, in order to address systematic errors and the emerging complex interactions between Earth system components used in weather and climate research as well as operational service applications.

For the record, the draft Terms of Reference as presented during the WGNE34 session are reproduce below:

- *Foster and leverage the worldwide, long-term development of complex coupled Earth System Models (ESMs) with a cross-disciplinary panel of process-level modelling expertise in atmosphere, land, ocean, sea-ice, waves, atmospheric composition, hydrology and scientific computing to simulate the weather-climate-water-carbon cycle for application in research and operations*

   - *Possible action: Add core model development expertise to WGNE as appropriate to cover all of above at different time-scales of prediction/projection*
- *Identify, share, prioritise, link, understand and communicate systematic errors and their solutions across different time-scales in coupled ESMs*
- *Coordinate/co-design within a seamless GDPFS to appropriately balance (tiered) research and service (operational) requirements*
   - *Advice on the right level of complexity required in coupled ESMs given a particular target application, research or service priority*
   - *Maintain close links to specialised modelling, user, working or focus groups (redefined or current ones)*
- *Advise WMC-GDPFS (via Research Board) on technology and science trends in Earth System Modelling, steer development with links to applications, users and services*
- *Quality assurance, ensuring an internationally accepted standard for ESM performance and their efficient and accurate use in operational weather & climate services*

All Members indicated their strong support for proposed new WGNE mission statement, which provides a strong, timely, relevant and ambitious direction for the Working Group. They commented on the need for Terms of Reference to be broad enough to remain valid for the foreseeable future and on the importance to also develop and maintain model development expertise and efforts in respective Earth system component disciplinary areas, including continued efforts on atmospheric model development. For this vision to be successful, Earth system approaches would also need to transpire in NHMSs. They also indicated that what makes WGNE special is its technical work and expertise, which shall be maintained, including its traditional model development. In many ways, WCRP has contributed to Earth system model development for a long time and can only benefit from a strong collaboration across weather and climate communities.

7.6 WGNE's role in developing the seamless GDPFS – Eunha Lim

Background: model development should be an integral part of a data processing and forecasting system, so that gradual improvements on the research side can be transitioned as appropriate into operations. Vice-versa, feedback from operational systems shall inform research priorities *and investments. How can WGNE strengthen this data-processing-forecasting value cycle across NMHs, academia and agencies?*

See [*https://www.wmo.int/pages/prog/www/DPS/documents/DraftofSEAMLESSGDPFSIMPLEMENTATIONPLANver4.0.pdf*](https://www.wmo.int/pages/prog/www/DPS/documents/DraftofSEAMLESSGDPFSIMPLEMENTATIONPLANver4.0.pdf)

Working areas for Earth System Models, building the coordination mechanism and relevant modification of the ToR was already discussed as part of the previous section.

It was proposed to start with a pilot project to shape the coordination mechanism and to find the gaps to build seamless GDPFS with support from the Secretariat. The interaction between aerosols and cloud/radiation could be an appropriate pilot initiative because it covers quite wide range of time scales and bring experts from other areas. This could start as soon as the formal coordination between WGNE and GDPFS is established.

Data from different domains (marine, hydrology and air quality) anddifferent owners (institutes, academia, and private sectors) are crucial to develop and validate ESMs. Although those data do not follow the quality control or standard format provided by WMO as the procedure to handle operational data, sharing them is still worthwhile. Ways to increase discoverability and accessibility was discussed. Each WGNE member will consider the mechanisms to enhance the discoverability of data from other domains and other sectors.

Members noted the challenges in briningboth observations and model simulations into a seamless system, as a number of different infrastructures, protocols, standards and policies currently exist (ESGF vs WIS, Grib vs NetCDF, etc).

7.7 Governance options in light of new WMO Research Board – Michel Rixen

Background: Terms of Reference for WGNE will have to be revisited. The relationship to the WMO Research Board will have to be defined.

See

- *WGNE : http://wgne.meteoinfo.ru/sample-page/terms-of-reference/*
- *Research Board: https://library.wmo.int/doc_num.php?explnum_id=9866 (page 45)*

Following the WMO Reform, the Research Board becomes a natural oversight body for WGNE and its expanded version.

Terms of Reference could focus on the following elements in particular:

- Innovation in Earth System Model development across weather - climate
- Systematic errors, coupling, parameterization, exascale computing, ML, process-oriented, metrics/verification

It is suggested to root the revised membership across operational centers, research and academic sectors, with a multi-disciplinary focus: Atmosphere/Land-hydrology/Ocean/Cryosphere and Carbon/biogeochemistry

It will also be important to define what WGNE will not do to ensure this complements other crucial research activities falling under the research board. There was a general agreement that data assimilation/reanalysis/observations impact could form a suitable bundle and critical mass for another group under the Research Board which whom WGNE could also collaborate for specific projects.

Members noted that the ocean community for example has its own model development networks and WGNE could reach out to e.g. OceanPredict, CLIVAR/OMDP, etc to start with.

8. WGNE Business

8.1 Actions

Actions from this meeting have been consolidated with actions from previous sessions and are reported in APPENDIX A.

8.2 Date and location of next meeting

Julio Bacmeister offered to host the WGNE35 meeting at NCAR sometime in the mid-September mid-November 2020 period. Ariane Frassoni re-iterated the CPTEC offer to host the WGNE36 session in 2021.

APPENDIX A - List of Actions

WGNE-34 actions

Recommendation – Centres encouraged to submit surface weather scores to WMO-LC.

Action 1: Ariane, JWGFVR & GAW to consider verification of aerosol/air quality forecasts and propose a way forward.

Action 2: Ariane to look into redefining period 2017-2019 for aerosol project (cf Amazonia)?

Action 3: Carolyn to develop a follow-up systematic errors survey in discussion with Daniel Klocke (GASS) & WGNE members.

Action 4: co-chairs to provide names of WGNE members (or people in their centres) who could be involved with TC initialization review to DAOS.

Action 5: François and Carolyn to see if they can find someone to work on surface fluxes project. François to send email to WGNE members to see if it would be possible to include funding for the surface flux analysis in a Schmidt proposal 2nd phase.

Action 6: Keith to liaise with Mark Rodwell and Linus Magnusson about combining the initial tendency and initial condition proposals, taking into account WGNE feedback.

Action 7: Daniel Klocke, Michael Riemer and John Methven to consider upscaling proposal requirements for DYAMOND and how to interact with Stochastic physics project.

Action 8: Keith to send feedback on stochastic physics to Hannah Christensen.

Action 9: Julio to present longer timescale (day 5-30) bias evolution at WGNE-35.

Action 10: Masashi / Eunha to follow up with WMO TC programme to sustain effort and contributions

Action 11: Co-chairs to identify WGNE member(s) to contribute to review of DAOS/Trop Cyclone initialization problem.

Action 12: Elena to consider renaming Blue Book to be broader than “Atmosphere and Ocean”.

Action 13: Try to get a member of JWGFVR to be on the Climate precipitation panel and vice versa. In the mean time exchange information as work progresses.

Action 14: Identify potential ocean equivalent of GASS/GLASS (perhaps COMMODORE Ocean?) and in other disciplines (hydrology, etc)

Action 15: WGNE Co-chairs to recommend extended scope of future WGNE around Earth System model development and to propose corresponding draft ToRs and Membership.

Action 16: Mich to send WMO Cg18 Congress statement on Public Private Partnership to WGNE Members

Action 17: WGNE Members to promote accessibility to observational data sets

Actions from WGNE-33

Action 15 – **All members** to alert their seasonal forecasting groups to S2S proposal for simulations with and without SPPT.

Ongoing: Awaiting update on protocol. Some centers interested, others will not participate. Berner and Pegion have a pilot study approved in NCAR and NOAA models which can serve to guide a wider comparison. Carolyn will forward more information when it becomes available.

Action 19 – **Keith Williams and Mark Rodwell** to draft protocol for initial tendency comparison.

Ongoing.