33rd SESSION OF THE

WORKING GROUP ON NUMERICAL EXPERIMENTATION (WGNE-33)

Japan Meteorological Agency, Tokyo, Japan

9-12 October 2018

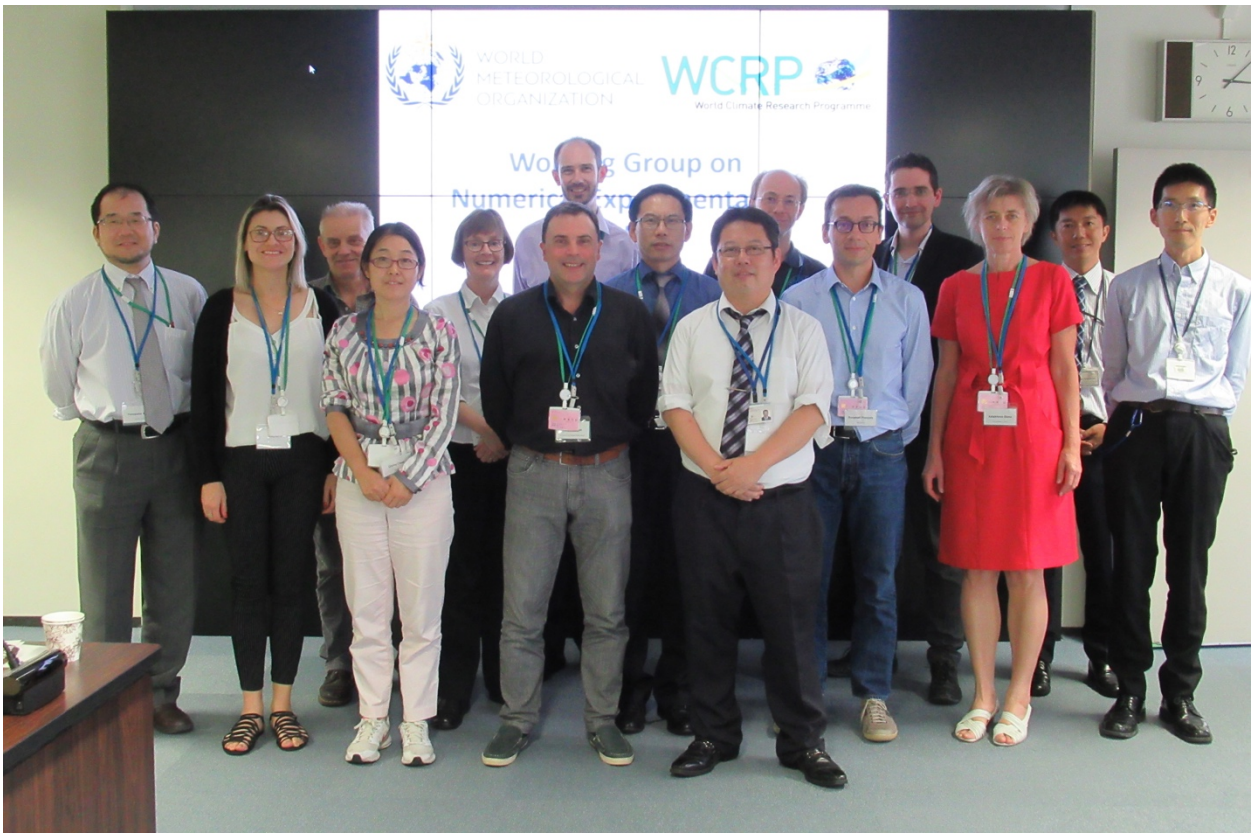
Prepared by

Ayrton Zadra and Keith Williams (Co-Chairs of WGNE)

Carolyn Reynolds (WGNE Member, incoming Co-chair)

Michel Rixen (WCRP JPS)

WCRP REPORT 19/2018



WGNE33 attendees

WGNE33 – PDEF4 Joint session attendees

TABLE OF CONTENT

WGNE MEMBERS: Ayrton Zadra (Co-chair), Keith Williams (Co-chair), Oscar Alves, Elena Astakhova, Julio Bacmeister, Francois Bouyssel, Francois Engelbrecht, Ariane Frassoni, Junichi Ishida, Carolyn Reynolds, Nils Wedi (remotely), Gunther Zaengl

WGNE EX-OFFICIOS: Caio Coelho (remotely), Michael Ek (remotely), Daniel Klocke (remotely), Steve Woolnough, Xubin Zeng (remotely)

INVITED EXPERTS: Hannah Christensen, Daniela Domiesen, Peter Gleckler (remotely), Christian Jakob (remotely), Karl Taylor (remotely), Fanglin Yang, Kunio Yoneyama

PDEF: Oscar Alves, Judith Berner, Craig Bishop, Manuel Fuentes, Mio Matsueda, Zhiyong Meng, John Methven, Mark Rodwell, Olivia Rompainnen-Martius, Susanne Theis, Munehiko Yamaguchi

WCRP JPS: Michel Rixen (remotely)

WMO: Estelle de Coning

Tuesday 9 October: WGNE session

Introduction

Welcoming remarks were given by the Director-General of Forecast Department at JMA. He pointed out the urgent need to extend lead times and provide smaller warning areas because the aging population is becoming harder to evacuate. The JMA new strategic plan focuses on four priorities: 1) threshold rainfall; 2) typhoons; 3) impact on socioeconomic activities and 4) impact of climate change. These were to be discussed further at the JMA seminar on the Friday afternoon.

Review of outstanding action items (Keith Williams and Ayrton Zadra)

Outstanding action items were discussed. A new action item list, including actions that arose during WGNE33 and outstanding items from previous years, was discussed at the end of the meeting on Friday.

WGNE-systematic errors survey (Ayrton Zadra)

The draft survey was discussed and suggestions for modifications provided. Suggested modifications included 1) having the flexibility to note if the answers are for single or multiple systems or timescales. The question of how best to rank the systems was discussed, and it was decided to ask for a ranking of only the issues/errors that were deemed important. The users should note if the system is coupled, and it was decided to broaden the surface heterogeneity question. The survey will be modified accordingly and will be distributed through WMO mailing list with a request to forward as appropriate.

Update on WGNE blue book (Elena Astakhova)

The "Research activities in atmospheric and ocean modeling" is called the blue book because it originally had a blue cover. The call for contributions is usually in February and the deadline is in May. A new issue used to be published at the beginning of July. The number of contributions have decreased since the 1990s, but have held steady at around 60-77 for last four years. The section with largest number of contributions is data assimilation (DA). There were very few contributions in computing, resolution effects and verification, although it was felt that these sections are still needed. Despite the trend to publish results in peer-reviewed journals, it was decided that the blue book is still useful for early exchanges of information (especially description of system upgrades). Methods to promote a more flexible exchange of information without proving too burdensome on Elena were discussed. Elena noted that the "news" section on the WGNE web site could serve this purpose (people could, e.g., provide links to tech reports), and the group was encouraged to contact Elena with relevant news items.

GLASS (Mike Ek)

Land surface has a dominant impact on predictability on intermediate (multi-week) timescales (e.g., snow melt, vegetation change). The role of GLASS is to focus on the better representation of the earth system through understanding the role of the land surface. GLASS promotes many community activities that improve best estimates and model representation of state variables, understanding of land - atmospheric feedbacks, and understanding role of the surface on predictability. Benchmarking efforts include PLUMBER, PALS, ILAMB, GSWP3, and SoilWat. Land-atmosphere interaction efforts include LoCo, GABLS/DICE, LS3MIP, and LUMIP. Several of these efforts are cross-cutting with other projects. Results from different projects were briefly discussed. One example was from DICE, which is looking at land-atmosphere coupling for the diurnal cycle under several different environments. Findings so far show that there are large errors in evaporation and other surface fluxes. The key science questions for next 5-10 years include going "Back to basics" to understand processes and observe and model them and hierarchical model development to improve earth system models. Projects would benefit from closer connection between GASS and GLASS.

Years of the Maritime Continent (Kunio Yoneyama)

The scientific motivation for YMC is primarily related to the Madden Julian Oscillation (MJO) maritime continent (MC) barrier. This is in contrast to the previous project, DYNAMO, which focused on initiation of the MJO. The purpose of YMC is to expedite progress in improving understanding and prediction skill of local multi-scale variability of the MC weather-climate and its global impact. Time period is July2017-February 2020, and it involves 70 institutions from multiple countries. YMC web sites are www.bmkg.go.id/ymc and www.jamstec.go.jp/ymc. YMC has placed an emphasis on data sharing, field campaigns, modeling, prediction and applications, and outreach and capacity building. Local agencies have agreed to provide routine observations in addition to Intensive observing period data. A big issue has been getting approval from local countries for platform deployment. JAMSTEC has provided near-real-time NICAM 14-day (7 km) and 30-day (14 km) forecasts in support of YMC. The 7-km runs better represent the diurnal cycle phase peak in precipitation. Model diurnal precipitation peak in NICAM is later than observed which is probably because the model isn't able to deal with steep and complicated topographic effects. Part of YMC-Sumatra is a radiosonde intercomparision and quality-control of observations (radiosonde humidity biases have an impact on data assimilation). Observations

from the intensive observation periods (IOPs) will be released to the general public 1 year after collection.

TC Verification Project (Junichi Ishida)

A review of model performance for 2016 was provided. The best TC track models in the western North Pacific were ECMWF, Met Office and KMA. Errors were larger in 2016 then 2015, probably reflecting sampling issues and intrinsic predictability, rather than system changes. Most models have a slow bias after recurvature, reflecting a common systematic error. Most models can't capture deep TC intensity (not unexpected at global model resolutions), but NCEP seems an exception. The current NCEP model has tendency to overdeepen TCs, but this is corrected in next generation system. All models have shallow bias east of Philippines. The TC intercomparison website will be available soon.

Surface Flux Intercomparison (François Bouyssel and Carolyn Reynolds)

The intercomparison idea was prompted by the fact that biases in surface fluxes (especially over oceans) were identified as an important and widespread issue during the WGNE Workshop on Systematic Error in 2017, and the PAN-WCRP and WGNE32 meeting in Exeter in October 2017. A draft protocol had been circulated to WGNE members and GEWEX (D. Klocke) and CLIVAR/GSOP (A. Kumar) and comments subsequently incorporated. The 2nd draft protocol was presented for further comment. Comments specific to instantaneous variables and soil moisture along with other suggestions will be incorporated into the next (hopefully final) protocol to be sent out shortly. Preliminary results for January 2018 comparing ARPEGE (Météo-France) and IFS (ECMWF) global operational NWP systems for surface water, heat, and momentum fluxes were presented. While the momentum fluxes were comparable between the two systems, substantial differences were seen in the net shortwave radiation, particularly over the southern oceans, and latent heat fluxes and evaporation, particular over tropical ocean regions. Substantial differences were also seen in the amount of snow at high latitudes, and in the partition between stratiform and convective precipitation.

Aerosol Project (Ariane Frassoni and Francois Engelbrecht)

A recap of phase 1 results was presented and an extension on the first phase (phase 2) was proposed. Phase 1 results found the direct effect is important for NWP skill and that atmospheric model quality is very important for air quality forecasting. Phase 2 (WGNE-AER2) proposes to extend the first phase to include regional models as well as S2S experiments. The S2S extension is supported by the results of Benedetti and Vitart (2018), who found that accounting for direct effect was important for surface temperature as well as upper-air variables in extended-range forecasts. The current S2S draft protocol would include 5-member 32-day forecasts starting 1 March, 1 April, 1 May 2003-2018 for dust impact, and 32-day forecasts starting 1 August, 1 September, 1 October 2003-2018 for biomass focus. For the limited area domain, 72-h forecasts for 2016-2018 are proposed for different regions and time periods (for dust and biomass burning). These regions are chosen such that advection into the domain should not be a major issue. Several aspects of the protocol remain to be defined. This is a great opportunity to collaborate with WGSIP and S2S.

Climatological Look at Precipitation in NWP Systems (Julio Bacmeister)

Precipitation from three centers (DWD, ECMWF, Meteo-France) was compared with TRMM observations to evaluate the similarities between the model output and TRMM. Pattern correlations of the monthly-mean model precipitation for each 24-h accumulation out to 5 day and TRMM were computed. Correlations between the model precipitation and TRMM drop off more sharply over ocean than land (perhaps due to orographic forcing of precipitation patterns). Correlation in the differences between the forecasts and TRMM are larger for day 5 than for day 1, indicating that the model biases become more alike as forecast lead time increases. ICON and IFS have different biases over Andes and S. America in general. The lack of orographic details in summer western US in TRMM as compared to the models may suggest errors in TRMM. Intensity PDFs of accumulated precipitation were examined and found to be very similar between first and fifth day. Interestingly, the highest values in extreme precipitation in the IFS and CAM were not produced by the convective scheme, perhaps related to orographic enhancement. Future work will include extended validation, especially in complex terrain. It was suggested that some of the precipitation information may be available in TIGGE and Julio is going to investigate TIGGE as a source of model data.

WGNE Computer Table (Gunther Zangl)

An update on trends from the WGNE computer table was presented. There is a clear trend in the growing importance of ensembles at both global and regional scales. In addition, applications including aspects of atmospheric composition are increasing. Changes to the table were proposed and discussed. Gunther prosed that the sustained/peak performance question be discontinued in favor of something like processor type, #nodes, #cores/node? He also suggested that the table include more specific information about partition of resources between, e.g., DA vs. deterministic vs. ensemble forecast, or between operations and research. It was agreed upon that the aerosol/composition tables be more "free-form" to allow for more specific information on aerosol DA and forecasting to be provided

GASS Report (Daniel Klocke)

GEWEX and GASS are very relevant to WGNE as they focus on process-level understanding and model improvement. The meeting "Understanding and Modeling Atmospheric Processes", held in Australia in February 2018 has been summarized in BAMS. There are many projects on topics that include 1) surface drag and momentum; 2) initialized land temp and snow on S2S; 3) fog; 4) diurnal and sub-diurnal precipitation; and 5) 2nd phase of grey zone (drag and grey zone follow on projects of the WGNE projects). Participation in the following projects is encouraged: 1) follow-on drag project (COORDE, irina.sandu@ecmwf.int); 2) initialization of land temperature and snow-pack (ILSTSS2S) in partnership with TPEMIP (yxue@geog.ucla.edu); 3) Fog modeling intercomparision, based on LANFEX (ian.boutle@metoffice.gov.uk). A phase 1 protocol for a project on the diurnal and sub-diurnal precipitation in different climate regimes is being finalized. The 2nd phase of the grey zone project is based on EUREC4A and GATE Phase III and focuses on scale awareness, stochasticity, and convective organization (Lorenzo.tomassini@metoffice.gov.uk.

MJO-TF Update (Steve Woolnough)

The MJO-TF has five subprojects (process-oriented diagnostics, forecast evaluation, air-sea interaction, interactions with MC, and interactions with the extratropics). A recap of major findings was provided. Process-based diagnostics highlight the importance of background mean moisture state (horizontal advection of moisture key). Models with a good representation of mean moisture gradients tend to have better skill for phase 6-7 initialization. Models with a good cloud-longwave feedback tend to have better skill for phase 2-3 initializations. Coupling improves eastward propagation. Models with a too-dry boundary layer can have exaggerated coupling strength. MJO propagation is sensitive to the representation of convection over the maritime continent (emphasis of YMC). Future plans are to keep the sub-projects on the MC, MJO-extratropics, and MJO prediction, and to subsume the process-oriented diagnostics and air-sea interactions into "mechanisms and simulations". Future plans also include consider the role of interannual modes of variability (ENSO, QBO) and applying existing and new process-oriented diagnostics and air-sea interaction diagnostics to CMIP6.

Wednesday October 10th, joint WGNE session with Predictability, Dynamics, and Ensemble Forecasting:

WCRP/CAS/WWRP/GAW matters (Estelle de Coning, Michel Rixen)

Estelle provided an overview and update on both WWRP and WCRP (for Michel Rixen, who could not attend). WWRP mission is focused on research for more accurate and reliable forecasts from minutes to seasons, aiming for seamless prediction. Impact-based forecasts, strengthening academic-operational partnerships, and enhancing the role of early-career scientists are all priorities. The three core projects are HIWeather, S2S and PPP. WWRP has 4 challenges with interconnectivity (extremes, water, urban, and new technology). WCRP is developing a new strategic plan (WCRP is 40 years old now, WWRP is only 20 years old). The WCRP external review panel noted the importance of bedrock science, the seamless approach (in terms of time, space, earth system modeling, and research to operations). Links to services and policy and the need for an implementation plan were emphasized. An overarching theme is closing the energy, water and carbon cycles, and connecting climate science to decision making. The strategic plan will be introduced at the town hall meeting at the Fall AGU meeting in Washington DC in December 2018. The review panel suggested three capability themes (earth system processes across scales, climate variability, predictability and prediction, and climate change and earth system feedbacks). They also suggested forming a WCRP working group on climate model development,

which may provide an opportunity for WGNE to move into a new era of earth system model development.

Summary of the Workshop on Systematic Errors (Ayrton Zadra)

Common systematic errors remain a major focus for WGNE and modeling centers around the world, prompting the Systematic Error Workshop held in Montreal. The meeting summary was recently published in BAMS. Recommendations include: 1) WGNE-WGCM to survey field to prioritize errors (survey in preparation); 2) Extend/generalize the drag project (new WGNE/GASS project COORDE); 3) Examine surface flux errors (new WGNE surface flux project protocol being drafted); 4) Make use of S2S databases; 5) discuss with S2S/WGSIP an extension of the aerosol project to seasonal timescales (see presentation by A. Frassoni and F. Engelbrecht, above); 6) Consider a cross-cutting weather-climate group looking at initial tendency analysis of common biases; and 7) Have the next systematic error workshop focus on solutions.

Flow-dependent Error Growth in TIGGE models (Mark Rodwell)

Examination of ensemble performance should focus on both reliability and sharpness. Reliability in ECMWF ensemble is improved through SPPT and ensemble of 4DVAR analyses. There is some evidence that day to day variability in spread is correlated with day to day variability in error (i.e., sharpness). The Lagrangian error growth rate for the TIGGE model (with advection of error removed) is examined. Cluster analysis indicates a strong growth rate for the warm-conveyor-belt (WCB) cluster. The WCB cluster also shows larger per case contribution to unreliability. Results suggest a future focus on accounting for uncertainty in convective momentum transport. The removal of tapering of stochastic physics near the boundary layer increases ensemble variance near surface and increases reliability. A budget approach is suggested as a diagnostic that will help improve proper scores. Applying this approach to multiple ensemble systems is suggested as future work.

NAWDEX (John Methven)

The North Atlantic Waveguide and Downstream Impact Experiment (NAWDEX) field project occurred during September-October 2018 and a review article has appeared in BAMS. The goal was to quantify effects of diabatic processes on disturbances to the jet stream and influence on downstream forecasts. Dropsondes from the field program provide evidence of an underestimation of wind on the flanks of ridges and above the tropopause and overestimation of winds below tropopause (i.e., the observed jet max is sharper in observations than in model). These differences increase with forecast time. This is relevant as biases in the model Potential Vorticity PV gradient impact Rossby Wave propagation. Very high resolution is required to represent the observed sharpness. This is an example where sub-grid-scales have an effect but cannot be treated with stochastic forcing (although it was noted that stochastic physics could sharpen gradients). A NAWDEX-TAMIP experiment (running NAWDEX cases with climate models in NWP mode) has begun and participation is encouraged. This experiment will look at simulation sensitivity to both vertical and horizontal resolution.

Improving Stochastic Parameterization Schemes using High-resolution Model Simulations (Hannah Christensen)

A lack of scale representation motivates stochastic parameterizations. The framework for studying this problem is to identify uncertain processes in forecast models, characterize this uncertainty, and design stochastic parameterizations to explore this uncertainty. Coarse-graining is used to measure error (Christensen et al, 2018, JAMES). This study uses the ECMWF IFC single column model (SCM), available through open IFS. It builds on an existing high-resolution dataset: Cascade (from UK Met Office), at 4km resolution. A case study is being used to explore the physical basis for SPPT and to evaluate if some SPPT assumptions may be addressed using coarse-graining. SCM simulations are run for two hours, and the first hour is disregarded due to spin-up issues. The “optimal e” parameter (the random component of the tendency magnitude in SPPT) is calculated for the SCM at every grid point based on the high-resolution run. The optimal “e” values and statistics can be used to improve parameters in SPPT. The results suggest that skewness should be incorporated into the SPPT scheme. This technique is proposed as a new method for assessing model error. The use of existing high-resolution simulations and SCM result in a relatively low entry bar.

Building Stochastic Representation of Uncertainty into Physical Parameterizations (Christian Jakob)

The clear distinction between model error and model uncertainty should be made. This talk focuses on model uncertainty. Parameterizations should be informed by observations and contain statistical elements where equations are not known, and physical elements where they are known. Grey-zone resolutions require new techniques for developing new convective parameterizations. We learn about the relationship between large and small-scale states using observations (e.g., CPOL density plot of convective area fraction vs. domain average convective rain). In observations, the degree of stochasticity can be a strong function of the predictors chosen. E.g., CAPE looks a lot more stochastic than vertical motion. This information and a Markov-chain model are used to predict area fractions. The model can deal with the stochastic behavior of convection, co-existence of different types of convection, organization (cold pools, coast lines, other), resolution-awareness and dynamic switching of convection. Early results in the MPI model are provided in Peters et al., JAMES. The scheme results in much smaller mass fluxes, strong improvement in autocorrelation of daily rainfall, and smoother autocorrelation of time step rainfall. The effects impact the large scale, improving Kelvin waves and Outgoing Longwave Radiation (OLR), and the RH-rainfall relationship.

Discussion on Ways Forward with the Representation of Intrinsic Model Error (C. Reynolds, C. Bishop, J. Berner, and O. Alves):

The discussion began with a description of a coarse-graining experiment proposed by PDEF at WGNE32, including the potential pros and cons (such as the issue of initial shock or spin-up) and ways to reduce entry barriers. Discussion was devoted to what centers could provide very-high-resolution runs in a global or regional sense (regional would be more computationally feasible). It was proposed that potential ways forward on this topic include the methods described by H. Christensen and M. Rodwell, above, and discussion will continue on these topics. C. Bishop has volunteered to put together a draft proposal on focusing efforts around cloud data assimilation. A proposed S2S stochastic physics experiment was also discussed. In this experiment, S2S forecasts would be performed and made available with and without stochastic forcing (particularly SPPT) to examine the impact of stochastic forcing on sub-seasonal to seasonal forecast timescales. This project is proposed under the WWRP/S2S project.

SPARC Update (Daniela Domiesen)

SNAP (stratospheric network for the assessment of predictability) is now an official subproject of S2S. The initial phase of QBOi included 17 models contributing to coordinated experiments. Currently QBO amplitude in many models is too weak in the lower stratosphere and may vanish from lowermost stratosphere. DynVarMIP is an ongoing targeted model intercomparison, currently evaluating CMIP6 models. A new focus is on dynamics, predictability, and extremes across timescales (a workshop is planned for Madrid in October 2019). S-RIP (SPARC reanalysis intercomparison project) evaluates biases in atmospheric parameters such as temperature, wind, and ozone. It was noted that the S2S database has much better representation of stratospheric fields than SubX and that the 2nd phase of S2S will include even more stratospheric data. Good stratospheric simulations should help with S2S predictive skill. There is a positive correlation between tropospheric and stratospheric skill. While stratospheric events are not very predictable, once you have an event, tropospheric impacts exist for several weeks.

TIGGE Status Report (Manuel Fuentes)

THORPEX Interactive Grand Global Ensemble (TIGGE) was started in October 2006 as a key component of THORPEX. THORPEX finished in 2014 while TIGGE has continued for another five years, and its further continuation is being reviewed in 2019. NCAR no longer an archive centre (now only ECMW and CMA are archive centers). There are 3 Petabytes of data in the archive, and it grows at 4 TB/day (2 million fields/day). There are currently 2680 users. Recent improvements include an enhanced web infrastructure, technical wiki, new tools including a history page, and usage statistics. NCAR still hosts the tropical cyclone tracks archive (rda.ucar.edu/datasets/ds330.3). TIGGE-LAM (Limited Area Model) has data from 1 Jan 2013, from European models, and far fewer users than global TIGGE. Given uncertain funding last year, many actions were delayed as time was spent on re-establishing contacts with data providers, filling gaps, and migrating from unsupported hardware. Use of TIGGE in TC Forecasting (Munehiko Yamaguchi): TCs are the most common focus of TIGGE papers, followed by heavy precipitation. RSMC Tokyo currently provides multi-model TC ensemble track predictions to Typhoon committee members as an example of R2O. TC forecast center survey results indicate that ensemble forecasts are used by nearly all forecasters, and are particularly important in track and genesis. Ensembles are less well-used in intensity forecasting and used the least in forecast TC size/structure and extratropical transition. Less than half of the forecasters use calibrated ensemble forecast information. While ensembles are used by forecasters, ensemble-based

probabilistic products are not commonly used. In the future, survey respondents were keen to see probabilistic products for impacts. A future challenge would be to use deterministic forecasts, TIGGE and S2S together effectively. There are some discrepancies in information provided to TIGGE CXML, which would benefit from standardization.

TIGGE General Discussion

Participation in the TIGGE S2S workshop-ECMWF April 2019 is encouraged, and will be useful to promote the continuation of TIGGE and get operational centers' commitment. Reproducibility of published results is another argument for WMO to continue TIGGE. Is the CMIP multiple center archive an alternative data archive model given the huge amount of data? This was the original TIGGE idea, but didn't go that route because of complications. PDEF is going to establish a TIGGE panel.

Review of Center Ensemble Systems (Junichi Ichida and Carolyn Reynolds)

The ensemble system recent and planed upgrades were first presented for both global and regional models. For the global system, in addition to model upgrades, plans are to increase ensemble size at Meteo France and RHMC, introduce 50 ensemble data assimilations at ECMWF, and increase horizontal resolution at CPTEC and RHMC. For the regional systems, several centers have plans to increase ensemble size and ensemble frequency. Given the joint meeting with PDEF, the report on research focused on incorporation of model design. Both ECMWF and Meteo France are testing stochastic parameter perturbations, and several groups are considering stochastic forcing based on analysis increments for their ensemble DA and or ensemble forecasts (Met Office, DWD, and NRL). Other developments include testing of the development of a stochastic deep convection scheme (ECCC), dynamical core departure point uncertainty (ECMWF), stochastic boundary layer perturbations (Met Office), and additive model-error perturbations scaled by physical tendencies (RHMC).

Thursday 11 October

CMDP Report (Peter Gleckler)

The CMIP diagnosis, evaluation, characterization of Klima (DECK) is an entry card for CMIP, and includes AMIP simulations, pre-industrial control simulations, 1%/yr CO2 increase, and abrupt 4xCO2 runs. Historical runs are also needed for CMIP. WGNE encouraged CMDP to identify a limited set of metrics for CMIP class model evaluation. CMIP DECK was defined, in part, to inspire ongoing model evaluation. Weighting model projections (depending on model skill, or model dependencies) remains an active area of research. Model weighting can have a big impact on, e.g., Arctic sea-ice projections. WGNE/WGCM metrics and diagnostics panel's broad scope has made focusing challenging. There are several community-based tools that may be relevant for CMIP DECK evaluation. WCRP strategic and developing implementation plans provide an opportunity to rethink CMDP. Progress has been lacking on advancing specific scientific methods and topics. It has been suggested to consider a pan-WCRP pane which inspires targeted expert teams to define and implement a limited set of model metrics (e.g., MJO diagnostics task force, CFMIP community diagnostics code, benchmarking of simulated precipitation).

JWGFVR report (Caio Coelho)

The Joint Working Group on Forecast Verification Research has many activities including SWFDPs, PyeongChang Olympics verification planning, HIGHWAY/L. Victoria, AvRDP (aviation), TLFDP (typhoon landfall), CBS (flash flood). The working group continues to actively contribute to the Polar Prediction Project (PPP), Sub-seasonal to Seasonal prediction project (S2S) and High Impact Weather project (HIW). A chapter on forecast verification time scales, including precipitation verification metrics, has been included in upcoming S2S book. A linear association between MJO and precipitation prediction skill has been found and reported in a recently published article. Atmosphere-ocean coupling likely leads to better subseasonal MJO and precipitation prediction. There is also ongoing work on quantifying observation uncertainty in verification measures (e.g., MesoVict). Following an action from WGNE32 the working group made important progress on process-oriented verification. This verification work includes, for example, T2m error stratified by total cover, and total cloud cover error stratified by cloud top height. New WMO guidelines are in place for exchanging surface forecast scores, including precipitation scores. Global NWP centers are encouraged to exchange surface scores via the WMO Lead Centre for Deterministic NWP verification (LC-DNV).

Fixing Systematic Errors at CPTEC (Ariane Frassoni)

At CPTEC, the operational regional model is now WRF at a 5 km resolution. BRAMS will remain in operations for air quality until WRF-CHEM comes on line. Current work includes improvements to biomass burning emissions forecasts to improve air quality forecasts in the regional domain. The implementation of a new parameterization to better represent cloud cover over Amazon based on probability density function using a climatological run of the Brazilian Atmospheric global Model (BAM) was implemented. The results compared with the UK Met Office global model in the context of The Climate Science for Service Partnership Brazil were encouraging showing improvements of the cloud cover distribution. An optimization method has been applied to improve convective parameterizations using by optimizing the weighting of the five different choices of mass flux. Preliminary results indicate that applying the optimal weights determined by this method improved forecasts in the regional model. Future activity will include the implementation of the method in the global model.

Reducing Systematic Errors in ICON (Gunther Zangl)

DWD focused on biases in the diurnal cycle of T2m and RH2M, particularly a cold and moist bias in late afternoon/early evening, which was most pronounced during summer months. A hypothesis was formed based on a recent paper that taking into account the asymmetric diurnal cycle of plant evaporation could reduce the early evening bias. In Terra (a simple model), introducing a prognostic variable for integrated plant evaporation that varied during the day reduced the bias (but did not eliminate it). Moisture and cloud biases in snow-covered regions with high vegetation was addressed through artificial reduction of snow-cover fraction of melting snow, reduction of snow temperature used for calculating potential evaporation, and other modifications. A temperature bias dipole in tropical troposphere was reduced by a factor of two through tuning changes in the convection scheme and a correction in the physics-dynamics coupling. Unresolved bias issues include 1) a North African bias (too little cloud cover, too dry, and too cold), and 2) a cold bias in 200 hPa in NH summer. The task force ResQME was established in April 2018 to investigate bias problems (for ICON in both global and regional configurations).

Reduction of Systematic Errors in GFS FV3 DYCORE and Advanced Microphysics (Fanglin Yang)

FV3 implementation will be at the same resolution, but the resolution of the ensemble DA will increase. Physics changes include replacing the Zhao-Carr/Sundqvist microphysics with GFDL microphysics, updated ozone photochemistry, new parameterization for middle atmosphere water vapor photochemistry, revised bare soil evaporation, modified convective schemes to reduce excessive cloud top cooling, and improved terrain (resulting in big difference in Greenland and the Andes). The physics will remain at 64-bit precession, but the dynamics precision will be reduced to 32 bit. In DA, TC relocation and full field digital filtering will be eliminated. FV3 does better job at conserving stratospheric ozone than old model and the stratospheric temperature bias is reduced. Precipitation diurnal cycle over the CONUS is improved. FV3 represents well the observed relation between surface pressure at storm center and surface maximum wind. A large cold bias over the North Pacific and Alaska was fixed with changes to the foundation temperature analysis in the near sea-surface temperature model. Remaining concerns include too progressive synoptic patterns, precipitation dry biases for moderate rainfall, and poor representation of inversions, which may be tied to errors in stratus decks.

Activities at CMC to Reduce Systematic Errors in GEM (Ayrton Zadra)

CMC has implemented a new and successful approach to improve forecast skill, which includes challenging the model across resolutions/scales, developing parameterizations that are less sensitive to vertical resolution, imposing conservation principles, participating in international projects, and expanding R&D strategies. Coding issues were discovered through testing physics at different resolutions. The use of a hierarchy of models, often moving back and forth between the models, proved useful. Initial tests were performed at reduced resolution to conserve limited computational resources. Forecasters were included in expert evaluations of case studies and assessment of parameterization tendencies. MHEEP: Model hydrological and Energy budget Evaluation Project was employed. This consisted of running four 13-month 24km free runs using prescribed SST and sea-ice fraction. Most improvements came from physics changes, rather than increase in resolution.

20th Century Temperature Biases in CESM2.1 (Julio Bacmeister)

CESM2.0, frozen in June, included big changes in physics. CMIP6 DECK simulations using CESM2.0 are currently underway. The initial problems with CMIP6, detailed last year, were possibly due to the 2nd aerosol effect being overestimated. There was a large sensitivity to CMIP6 emissions compared to CMIP5 emissions. While nothing systematic was identified, fixes for other problems led to gradual improvements. Equilibrium climate sensitivity is considerably larger in CESM2 (the stronger aerosol interaction might be responsible), but there is a fair amount of uncertainty in how that is computed. CESM2 doesn't stabilize after abrupt 4xCO2 like CESM 1 does. Models that behave more like CESM1 have brighter low clouds. Big impacts can be made just by tuning a parameter in CLUBB (as impactful as primary physics modifications). Future work includes establishing a connection of 4xCO2 behavior and 20th Century simulations, comparisons with other methods, and investigating the impacts of land and ocean.

Challenges in Reducing Systematic Biases in the Indian Ocean (Oscar Alves)

The Indian Ocean (IO) east-west dipole is important for Australia climate. The UKMO GC2 in seasonal prediction mode has an overactive IO dipole as well as DJF biases. Various experiments suggest that the atmosphere is not the basic problem. Examining the sensitivity to the ocean data assimilation scheme (UK vs. BOM) showed big differences in the IO subsurface currents. BOM system is multi-variate and produces current increments where NEMOVAR does not. BOM does not assimilate altimetry data while NEMOVAR does (and assimilating altimetry on the equator can be very tricky). The bias in year 1 is very different from years 2-5. Future work includes measuring how much ocean initialization errors impact model bias, and conducting a large hindcast set with different ICs.

Reducing Systematic Errors in Seasonal Forecasts over Southern Africa (Francois Engelbrecht)

Seasonal CCAM AMIP-style simulations are performed at a 200-km resolution. DJF simulations for 850-hPa height anomalies are skillful (resulting from a strong ENSO signal), but other seasons

are not. Using observed rather than climatological stratospheric ozone increases the fraction of ensemble members that are able to reproduce the "normal" rainfall for the 97-98 El Nino. Experiments indicate that skill can be obtained by including observed stratospheric ozone forcing (only for the spring-summer season).

Improvement of SL-AV Model Climate and the Impact on Medium-range Weather Forecasts (Elena Astakhova)

The model dynamical core is a finite difference SI/SL (semi-implicit semi-Lagrangian) formulation. Parameterizations are taken from ALADIN/ALARO consortium, with differences in 1) radiation (CLIRAD SW + RRTMG LW), 2) INM RAS SRCC MSU multilayer soil model, and 3) marine stratocumulus parameterizations. 10-day weather forecasts plus 60-day seasonal forecasts are performed (the center participates in the S2S multi-model archive). There was a big reduction in the cloud bias through the introduction of the marine stratocumulus cloud parameterization. With the increased vertical resolution (100 vs 28 levels), the higher upper level (0.05 vs 5 hPa), and parameterized convective gravity wave drag the model is able to reproduce the QBO rather well. Improvements in model climate accompany medium-range forecast improvements.

HIGH-TUNE, High-resolution Simulations to Improve and Tune BL Cloud Parameterizations (François Bouyssel)

The objective of HIGH-TUNE is to improve parameterizations involved in the representation of low-level clouds. SOA statistical tools are used to propose values for parameters based on SCM/LES for an ensemble case. The tools are used to both tune parameterization free parameters, and to disentangle structural errors from tuning issues. The effort is using several campaign observations for 1D and LES cases. It is used to determine parameter space that cannot yet be ruled out. The objective is to identify how plausible parameter ranges vary over all the 1D cases? The method would be useful for defining ranges of stochastic parameter value perturbations.

HPC Trends and the Scalability of Atmosphere and Ocean Models (Nils Wedi)

At ECMWF, the increase in degrees of freedom has followed a modified version of Moore's law, doubling every 24 months. The current computational resource allocations are approximately 40% dynamics and 30% physics, with larger and larger fractions taken up by other components (e.g., ocean, wave, ice). New HPC trend include 1) manycore CPUs combined with accelerators; 2) arrival of open instruction set architectures; 3) exascale race driven by a concern for energy footprint and physical distances between processors (low-power processors, memory hierarchies, liquid cooling …); and 4) machine learning driving both processor and algorithm development. Many programs, including Extreme Earth, have been developed to address these issues. The ECMWF ESCAPE project has several objectives including defining and encapsulating the fundamental algorithmic building block (weather and climate dwarfs). ATLAS is a library for NWP and climate modeling, available on github (https://github.com/ecmwf). Deep learning may influence algorithmic choices for weather and climate. Future architectures will probably work best with lower precision and efforts are underway to determine what components work well with reduced precision. The results of a questionnaire were summarized. There is a trend away from spectral models, but there might be a reverse trend due to machine learning. Scalability challenges include 1) large-time-step models; 2) ocean/sea-ice/waves; 3) strong coupling in DA ; 4) chemistry and biogeochemistry and 6) unstructured grids. In the future, perhaps the questionnaire should be extended to sea-ice, chemistry/aerosol, biogeochemistry, and regional modeling to establish closer connection and sharing of key algorithms across ESM components.

Friday 12 October: WGNE session

CMIP/HiRes MIP status (Karl Taylor and Peter Gleckler)

CMIP is a huge effort and an order of magnitude more output is expected for CMIP6 than CMIP5. Currently only some model outputs are available, while most output will be made available over the next year. CMIP6 infrastructure improvements include specifications developed and documented in series of WIP position papers, and a reference set of controlled vocabularies defined to enable independently-developed infrastructure components to communicate. Input4MIPS is collecting and controlling forcing data. CMIP6 guidelines can be found at pcmdi.llnl.gov/CMIP6/Guide/dataUsers.html. Data is available from ESGF data portals at IPSL and PCMDI/LLNL. Obs4MIPS data specifications are more closely aligned with CMIP6 and many new data sets are expected soon. A CMIP6 model analysis workshop will be held in Barcelona, Spain, 25-28 March 2019.

The HighResMIP purpose is to investigate robustness and convergence across multi-model ensembles as resolution is increased (only resolution is increased, physics are not tuned). There are 3 tiers: atmosphere –only, a coupled control, and future simulation to 2050 (both atmosphere only and coupled). 6 European groups have completed simulations as part of EU-PRIMAVERA project. Model data are currently available (3Pb) on single CEDA-JASMIN platform with analysis and processing tools and a compute cluster (enabling coordinated analysis without the need to download data). The plan is to start the upload to ESGF in about two months. TC tracks and climate extreme indices will be calculated and made available. NWP centers participating include Meteo-France, UK Met office, ECMWF, and perhaps JMA. Work is in progress to produce additional time-averaged fields.

APPENDIX A Actions from WGNE-33

Action 1 - **co-chairs** to update survey (only rank relevant issues, one survey for each timescale. possibly add whether coupled, consider asking how center investigates systematic errors) and circulate to WGNE/WGCM/WGSIP mailing lists.

Action 2 – **All members** to consider sending more 'news' items to Elena for website (including model system upgrades).

Action 3 – **Elena Astakhova** to email members requesting short presentations on model systems.

Action 4 - **Co-chairs** to refer bias for tropical cyclones to be too weak in analysis to DAOS.

Action 5 – **Junichi Ishida** to liaise with CMA to address missing data issues and possibly remove from published plot until this has been done.

Action 6 - **Francois Bouyssel and Carolyn Reynolds** to complete protocol for surface flux project (collect metadata e.g. on convective precipitation, detail of surface soil moisture request) and circulate around WGNE members

Action 7 - **Ariane Frassoni, Francois Engelbrecht and Qiying Chen** determine where aerosol project data should be stored. Finish protocol (especially considering data volumes of regional experiment) and circulate around WGNE members.

Action 8 - **Julio Bacmeister** to look if required precipitation fields are available from TIGGE or S2S in order to extend the NWP 'climate style' precipitation evaluation.

Action 9 - **Gunther Zangl** when the computer table is next updated to request fewer lead times for the table, request only fill in information where known, aerosols page more free-form, consider clarification on performance requests, resource usage by suite components.

Action 10 - **co-chairs** to circulate GASS white papers. **All members** to pass on to relevant people in their centers and those people to make contact with the project leads. Especially consider participating in COORDE.

Action 11 - **co-chairs** to liaise with Michel Rixen to extend MJO-TF for at least 3 years, maybe 4 years.

Action 12 – **All members in climate centers** to consider submitting transpose-AMIP experiments to NAWDEX.

Action 13 - **Hannah Christensen** to draft protocol for limited area coarse graining runs.

Action 14 – **Craig Bishop** to draft white paper to lever funding for data assimilation of unused datasets requiring high resolution simulations.

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

Action 16 – **All members** to encourage use of TIGGE and encourage contributing centers to attend ECMWF meeting 2-5 April 2019.

Action 17 – **All members in NWP centers** to consider their center submitting high resolution control to TIGGE.

Action 18 – **All members** to be aware of Non-Hydrostatic modelling meeting at JMA in 16th Nov 18.

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

Action 20 - **Caio Coelho** to circulate document on process-orientated verification to WGNE when complete (expected early 2019).

Action 21 - **Caio Coelho** to consider how best for all center's to participate in process-orientated verification (is it for the code to be shared or for center's to send data to one person e.g. Thomas Haiden) for analysis.

Action 22 – **All members** in NWP centers to consider submitting to CBS surface scores.

Action 23 - **Peter Gleckler and Caio Coelho** to make contact after March meeting on future of WCRP metrics. Identify a person/people who could sit on both JWGFVR and Climate precipitation panel and develop a plan on how to work together on precipitation evaluation.

Action 24 – **All members** to provide feedback to Nils Wedi regarding national efforts on scalability.

Action 25 – **All members** whose centers have undertaken precision (e.g. double vs single precision) studies to provide a presentation at WGNE34.

Action 26 – **All members** to consider development of dwarf components suitable for exascale testing.

Action 27 – **Julio Bacmeister** to report at WGNE34 on NCAR's investigations into the timescales on which systematic errors develop.

Action 28 – **François Engelbrecht** to represent WGNE at CMIP6 meeting in March 2019

Action 29 – **All members** to check whether publicly available verification results (esp. precipitation) are available at their center and send link to Elena.

Action 30 – **Co-chairs** to ensure developments in numerical methods features as part of WGNE34.

Action 31 – **Co-chairs** to draft white paper on how WGNE might fit into the proposed WCRP structure and circulate to members for comment.

Action 32 – **Co-chairs** to discuss with Michel Rixen membership of exofficios mailing list. **All** to use both members and exofficios mailings lists when circulating information.

Action 33 – **Gunther Zangl** to identify potential dates for WGNE34, avoiding w/b Oct 7 2019.

Ongoing actions from WGNE-32

Action 12 – **WGNE members** to provide Elena Astakhova with email addresses to extend/update mailing list for blue book. *Updates still required, especially where the format of email addresses has changed. Ongoing.*

Action 19 – **Michel Rixen** to close WMO WGNE website at beginning of January 2018 with a redirect to new website. *Will be done once transfer of archive is complete. Ongoing – co-chairs to highlight to Michel that old website still exists.*

Ongoing actions from WGNE-31

Action 27 - **Ariane Frassoni** to request Saulo Freitas to include estimates of statistical significance of the impacts of aerosols in the case studies. Ariane to continue working on this until the end of the year. *Ongoing.*