Report of the 24th Met Office Scientific Advisory Committee Meeting (22-24th January 2020)

Response from the Met Office Chief Scientist in red

We thank MOSAC for their contribution to a very productive meeting, and for producing a detailed and incisive report. This section summarises the main points raised by MOSAC, which are made in more detail within the Annex, and so responses from Met Office are given solely in the Annex in order to avoid duplication.

Summary

The 24th MOSAC meeting was held the 22-24 January 2020. The MOSAC has appreciated that many of the topics included in the meeting agenda were aligned with MOSAC previous suggestions. The Met Office prediction systems, and the Research and Development (R&D) programme is well-aligned with the priorities of the Research and Innovation 2020-30 Strategy. The capability of the Met Office to do fundamental research is significantly greater than that of other national meteorological services and represents a core strength that must be maintained.

In the following we provide the highlights of the detailed MOSAC report in the Annex. The associated recommendations and questions are referred between brackets (e.g., NGMS-R1).

Numerical Environmental and Weather Prediction

Ambitious and impressive efforts: The Supercomputing 2020+ and Next Generation Modelling System. With a six-fold increase in computing capacity in 2022, the Met Office will remain the most compute-capable operational centre for weather and climate in the world, a distinction it has held since 2016 (Top 500). More importantly, vastly increased computing will enable the range of new science and modelling capabilities needed to support the Science and Innovation 2020-30 Strategy, and in turn help clearly establish the Met Office science program as world leading. We are encouraged by the commitment of resources (approximately 50 FTE per year for the next 4 years) to the Next Generation Modelling System (NGMS) project.

Concerns about NGMS computational risks. Preliminary results show poor computational performance and potentially reengineering costs for data assimilation. It is understood that this is a snapshot in time, but the issue on poor computational performance is worrisome (currently reported as being a factor of five to ten slower than the actual Met Office Unified Model). The NGMS programme's efforts at remediation appear to be well-organized and with a level of urgency and attention appropriate to the risk the computational performance deficiencies present. MOSAC provides a series of recommendations that will help mitigate the situation (NGMS-R1, NGMS-R2, NGMS-R3 and NGMS-R4).

NGMS is confronted with significant multi-disciplinary challenges. The data assimilation requirements need to be taken in account sooner than later in the NGMS project plans. MOSAC notes that there is a pressing need to establish formal governance between the Met Office and the Joint Effort for Data assimilation Integration (JEDI) project coordinated by the US Joint Center for Satellite Data Assimilation (JCSDA). MOSAC urges the two initiatives to share their requirements (FS-R2, NGOPS-R1, NGMS-R5). Similarly, the alignment of plans between the atmospheric model and the environmental sciences components, including air quality, atmospheric dispersion, ocean and wave models should not come as an afterthought and the relevant teams should meet early in the design process (FS-R3, OS-R3, AQAD-R1 and AQAD-R4).

Good progress in global Numerical Weather Prediction. The highpoints are the new Numerical Weather Prediction (NWP) experimental suites, the outstanding utilization of neural network for better NWP products and progress towards the coupling of the ocean to the NWP systems with demonstrated benefits. MOSAC applauds efforts to include better data assimilation testing earlier in the model development process, and strongly encourages further moves in this direction. MOSAC suggests a couple of areas in which new results obtained this year merit further investigation beyond immediate operational implementations (WS-R1, WS-R2 and WS-R3).

UK NWP capability continues to develop well. The UK NWP system has been improved with new hourly cycling and land-surface data assimilation. A noteworthy strength of the process is the integration of feedback from forecasters through the O2R program. Numerous remaining challenges were cited in both the report and presentation, showing that there is still effort needed for a better utilization of the observations (UKNWP-R1). We would encourage studies evaluating the benefits of the ocean coupling for weather forecasting, and the proposals to introduce other environmental complexity such as river flows. We would also be interested in having more information on the plans for UK urban modelling, as work to date is focused on London (UKNWP-R2).

The testing and continued development of new physics parameterization in increasingly realistic settings is showing excellent progress. The Met Office team is to be congratulated, but many challenges are still to be dealt with to bring to maturity this new physics parameterization. The added value from the external partners in this research activity was less visible, and it would be helpful to have an update next year (WS-R4, FS-R1, FS-R4).

Very high-resolution NWP over London airports: is it adding value? Plans are to extend the domain size at 300m resolution and to increase horizontal resolution to 100m for improving deep convection forecasting. MOSAC encourages the Applied Science team to make stronger use of the Urban Working Group which brings together Met Office staff and academics working on urban topics to ensure alignment of the spread of activities internally and externally with the UK community (AS-R2).

A long and successful history of ocean forecasting in the UK. Much of the pioneering work on tides, surges, and waves originates from the UK. There are many activities within the Ocean Forecasting section which reflect this long history and the shift of the more traditionally weather focussed prediction to earth system prediction. In view of the emergence of earth prediction systems, MOSAC recommends an increased coordination at the design and planning phase of these prediction systems (e.g., OS-R4), especially within the context of the NGMS and JEDI efforts (OS-R1).

Coastal flooding and erosion: a major hazard in the UK. In Ocean Forecasting R&D provided documents, storm surges and waves are listed as one of the highest causes of natural disaster in the UK. As such and given the long history of the UK on R&D work in coastal flooding, the committee wonders if 1/10 of Ocean Forecasting team is enough to address the challenges of modelling coastal flooding. MOSAC would welcome a presentation on this aspect of the Ocean Forecasting R&D programme at the next year's meeting (OS-R2).

The Atmospheric Dispersion and Air Quality section at an important crossroad. The team has put forward some technical options for model use over the next few years with the aim of maintaining their forecasts and services, while meeting the challenges of the NGMS. MOSAC has provided its preference (AQAD-R1). MOSAC think also it is vital that the long-standing collaborations with UK academia and partners to continue, given the rising importance of air quality and the funding opportunities now coming (e.g. SPF Clean Air). These externally funded projects are providing a terrific opportunity for the section to improve their models and services and we would encourage continued engagement (AQAD-R2 and AQAD-R3).

The Informatics Laboratory: a key success. MOSAC praises the creation of a centre of expertise for the application of data science in meteorology. The applications targeted by the centre are generally regarded as the most promising areas for immediate success. The Met Office is well positioned internationally to lead that emerging aspect of the meteorological data science.

Data Assimilation and Observations Research

The Met Office Earth Observation R&D is unquestionable world-class. Hence the Met Office is very well positioned to challenge statement like the conventional wisdom that more observations are always required. Are current observations all used and well used? Are there numerical experiments that can be done to best know what type and where new observations are needed keeping in mind the Met Office financial plan? This might be a suitable topic for a future MOSAC meeting (WS-R5).

Good news: a 5-year agreement for the Facility for Airborne Atmospheric Measurements. Having this commitment is a good statement to the wider atmospheric science community that it is worth reaching out to the Met Office at the planning stage of large field campaigns. It should ensure the Met Office continues to be there at the inception of key projects and campaigns and is able to shape and influence those. The agreement itself seems flexible and a good deal for the Met Office.

MOSAC recognises the excellent quality of science at the Met Office Cardington. This group has had an excellent reputation for boundary layer science since the classic experiments of the 1970s. The reported work on fog observations clearly demonstrates good collaboration and experimental capability feeding into model development. MOSAC welcomes also the increase activity in the field of UAV instrumentation. Collaboration with the wider academic community on field campaigns is good but could be more strategic and open to more international partners (MOC-R2 and MOC-R3). Benefits would clearly be mutual in terms of optimising use of instrumental resources (including those with NCAS Atmospheric Measurement Service) and providing early awareness of emerging scientific/technological advances or modelling needs (MOC-R1 and MOC-Q1).

Weather-related Hazards and Impacts: Operational and Customized Forecasts

Applied Science activities are continuing growing. MOSAC is impressed to see the fast-growing of the section with 19 new employees. The examples presented (including posters) demonstrate the strong dynamics of Applied Science and multiple future opportunities (e.g., AS-R3): i) the rail leaf adhesion service is an outstanding example of multidisciplinary research work; and, ii) the PhD work on convective nowcasting with machine learning technique is another example of cutting-edge research. MOSAC is pleased to note of the new Strategic Priorities Fund UK Climate Resilience programme in which the UK Climate Resilience team can take a leadership role.

Applied Science scientists and scientific consultants work scope is extremely large. The expertise needed could be unavailable or not easily mobilized within the Met Office research departments. This is clearly a challenge and a concern for the continuation of growth. We encourage stronger engagement with the Met Office Academic Partnership

(MOAP) Universities and wider academic community in terms of providing expertise across a range of disciplines. As an example, there is a clear intention and progress on better scoping and understanding customers and stakeholder's vulnerabilities and exposure to weather and climate hazards. This is an important challenge for Applied Science which requires developing partnership with social scientists. (AS-R1, AS-Q1)

Operations to Research: a genuine cooperation between forecasters and researchers. Evidence of this was seen in how often Operations to Research (O2R) was cited in the counterpart R2O report. Many aspects of this interaction have generated positive progress in nowcasting and in the use of ensemble and probabilistic information. MOSAC provides futures suggestions for the future O2R workplan (O2R-Q1, O2R-Q2, O2R-R1 and O2R-R2). MOSAC suggests that a similar O2R approach would be a good addition to the environmental prediction systems.

The landscape for operational meteorology is rapidly changing. Users want local/personalised information, new technologies becoming available to communicate that information, and new players entering the scene. MOSAC was pleased to see the deep strategic thought that the Future of Operational Meteorology (FoOM) programme is putting into positioning the Met Office to continue to lead in the provision of services "to help you make better decisions to stay safe and thrive". What seems clear is that as routine meteorological forecasting becomes more automated and commoditised the real value of the meteorologist will be as a trusted advisor applying expert knowledge in high impact situations and translating that knowledge to stakeholders in meaningful ways that support their decision making. This will put forward many challenges for operational meteorologists to evolve in this changing technological landscape and service needs (FOM-R1, FOM-R2, FOM-R3, FOM-R4, FOM-R5, FOM-R6 and FOM-Q1).

Climate Research and Services

The Hadley Centre is a world-leading centre for climate science. One of the most striking results of recent years has been the high climate sensitivity (> 5C) from the Met Office CMIP6 climate model. It arises from a combination of model radiation scheme improvements (which increase the CO_2 radiative forcing, and revised aerosol and cloud parametrisations. In a workshop immediately preceding the MOSAC meeting, it was shown that these cloud revisions improved by 6hr weather forecasts, making a very strong case that these changes are physically credible. This is another example of the advantage of the weather-climate seamless unified-model approach that is used at the Met Office where NWP science can inform Climate science and vice-versa.

The seamless perspective opens more and more the prospect of new investigations of climate change at regional and local scales. As an example, the impact of increased resolution was shown, whereby climate-change induced increases in UK average wintertime precipitation were some 27% in a more realistic convection-permitting model, compared with just 16% in a more conventional regional climate model (CS-R1).

The UNSEEN methodology is an important methodology for determining the sorts of extreme events associated with a given return period. It is based on large ensembles of seasonal and decadal integrations that are run to provide counterfactual alternatives of the recent past. In this way, the model may provide better statistical estimates of the real world than the observations themselves do! This study and the two discussed above are of interest to the Weather and Climate scientists. MOSAC proposes that in the future we should try as much possible to present these weather-climate topics in joint MOSAC and SRG sessions (CS-R2).

There is clearly an active and growing area in Climate Services. These activities have several different drivers (e.g., fit to Met Office strategy and capabilities; user needs; funding priorities). With so many projects it was difficult for the committee to get a clear sense of where the ‘centre of mass’ of the portfolio lies. MOSAC recommends that the Met Office attempt to map current activities under some broad headings (HDM-R1 and HDM-R2) and MOSAC strongly recommends further consideration of the Met Office’s place in the complex international landscape, and where it positions itself in the value chain in order to add most value alongside other efforts and actors (HDM-R3).

Concluding Remarks

MOSAC would like to thank the speakers, carousel and poster presenters for their hard work and willingness to discuss all aspects of their work. We would also like to provide many thanks to the Met Office leadership and the administrative staff for this interesting, fruitful and well-organized meeting (including the note taking).

Gilbert Brunet (Chair) on behalf of the Met Office Scientific Advisory Committee (Thomas Auligné, Janet Barlow, Natacha Bernier, Andy Brown, George Craig, Véronique Ducrocq, Elizabeth Ebert, John Michalakes, Tim Palmer and Ian Renfrew)

Annex: Recommendations and Questions

Foundation Science

Foundation Science continues to provide the essential basis for the future excellence of Met Office prediction systems, and the research programme is well-aligned with the priorities of the Research and Innovation Strategy. The capability of the Met Office to do fundamental research is significantly greater than that of other national meteorological services and represents a core strength that must be maintained.

In particular, the Met Office should be commended for the new 5-year agreement for FAAM use. Having this commitment is a good statement to the wider atmospheric science community that it is worth reaching out to the Met Office at the planning stage of large field campaigns. It should ensure the Met Office continues to be there at the inception of key projects and campaigns and is able to shape and influence those. The agreement itself seems flexible and a good deal for the Met Office.

A second key success is the development of the Informatics Laboratory into a centre of expertise for the application of data science in meteorology. The applications targeted by the IL are generally regarded as the most promising areas for immediate success. With this critical mass of data science expertise, there is the potential for the Met Office to become a leading National Meteorological and Hydrological Service in this field.

FS-R1: The introduction of both CoMorph and CASIM into the same model science configuration (GA/RA releases) is a bold plan. It could be problematic. We would recommend that significant testing is carried out beforehand, and this should include testing in the data assimilation system which may perform badly even with a scheme that performs well in model testing. A thorough evaluation will most likely involve working with UK academia and UM Partners and should involve a wide variety of cases. In terms of real cases studies there will be opportunities within some ongoing projects such as Terra Maris and EUREC4A and the new NERC Climate Sensitivity due to Clouds programme.

If and when CASIM and CoMorph pass the initial tests to move from research to the GA release cycle (as part of GA9) they will then undergo rigorous evaluation across multiple resolutions, DA configurations, and regional and global configurations (including coupled). This assessment will include bi-monthly science assurance meetings where the community meets to review performance of each change and their potential interactions across this broad range of tests. The intention is to open these science assurance meetings to our partner community (NERC and UM Partners) via video-conference. We agree that the field experiments mentioned provide a great opportunity to evaluate new model developments at process level.

FS-R2: There is a clear need for data assimilation requirements and testing to be incorporated earlier in the model development process. A particular example of the problems that arise is the negative impact of data assimilation in the UKV model, that may require changes in the model domain or nesting methodology to correct. A broader scale example is the design decisions regarding the degree of integration of the NGMS model in the Joint Effort for Data assimilation Integration (JEDI) system, where DA requirements are critical to the design of the modelling system.

There is a clear recognition from our own recent experience of implementing global model changes at GA7 and GA8 that the potential interactions between model changes and DA need to be assessed much earlier in the development cycle both for global and regional systems. This was a recommendation from the review of global model development undertaken in 2017 after the INTEGRATE project. Revised testing strategies that include much earlier DA tests for GA and RA model development are underway and we can report progress with this at next MOSAC/SRG.

FS-R3: The alignment of plans between the atmospheric model and the environmental sciences components, including ocean and wave models should not come as an afterthought, and the relevant teams should meet early in the design process.

These governance/planning aspects are now being more actively managed through the Global Coupled (GC) Programme Board which involves key Stakeholders (Weather and Climate Science, NERC, and UM Partners). The implementation of coupled NWP into operations in 2021 will also provide a key focus. A similar approach/governance will be followed for the regional coupled (RC) system as it develops. A recent focus discussion promoted by the GC Programme Board has also taken place around improving development of the UKCA components within a seamless coupled prediction system.

FS-R4: The successful application of sub-km modelling is still dependent upon basic research. The development of scale-aware parameterisations or the use of true LES resolution may be required for some applications of this class of models and should be investigated with priority.

We agree and these aspects will be more widely discussed and developed as part of the Research and Innovation themes implementation plans, particularly on “7.Path to high resolution” and “9. Capturing environmental complexity”.

<u>Weather Science</u>

MOSAC is pleased to note good progress across the weather science area, including (non-exhaustively) the parallel suites, the neural network post-processing work (where there was fast progress on an operationally relevant problem) and progress with moves towards coupled NWP. The Cryosat ice thickness assimilation results are encouraging, and it would be interesting to know of plans to look at Soil Moisture and Ocean Salinity observations. It is pleasing to see the progress with getting the Global Atmosphere model developments to operations. MOSAC applauds efforts to include more Data Assimilation testing earlier in the model development process, and strongly encourages further moves in this direction.

ASCAT soil moisture observations have been assimilated for a number of years via the Extended Kalman Filter approach. Regarding ocean salinity, the capability exists in research mode to assimilate SMOS salinity data. Studies at the Met Office with Mercator and others show modest benefits in using SMOS data to constrain salinity. Data handling and operational elements still need to be implemented. There are no firm operational implementation dates at this moment as our priorities are coupled NWP and ORCA12 developments.

MOSAC notes a couple of areas in which new results obtained this year merit further investigation beyond immediate operational implementations and possible tactical fixes, as they raise more strategic questions. It therefore strongly recommends further work in these areas and would be interested to be appraised of progress. These are:

WS-R1: The results that UK DA can lead to worse results than the downscaler approach somewhere around 6-24 hours (rather than as previously assumed adding value for first few hours and then asymptoting back to the same level of performance as a downscaler approach). This is presumably due to some combination of multi scale flow in larger domains, using fewer observations than global within its domain (e.g. satellite), and failure to do large scale flow well because the limited area model can't consider info from outside the domain. Beyond immediate possible practical fixes as described, the issue feels worthy of significant further investigation, both to understand the problem and, in combination with partners, to develop approaches to address it. For example, as domain sizes increase are issues due to the Lateral Boundary Condition constraint/lack of info from outside, or the multi scale flow inside (as global will eventually face) and either way, what are the science options and implications for whole strategy for setting up Limited Area Models (of different sizes) in different parts of the world? Can true nowcasting and day 1 systems be the same or do they diverge?

We agree that the recent UK DA results are worthy of further investigation, especially given they are the opposite impact to similar studies performed several years ago. Our strategy remains to use the same DA approach for NWP-nowcasting (2-12 hours) as longer-range NWP. Our strategy is to assimilate more (not less) observations in high-resolution compared to global DA in the UK domain so we believe this is unlikely to be the primary cause. The recently observed degradation of UK DA hints at the impact of recent operational upgrades (e.g. increasing UK domain, vastly improved global analyses through variational bias correction, hybrid 4DVar, etc). To address this research question, we are planning to merge several projects in the next 1-2 years to tackle this issue directly: a) Impact of domain size, b) Blending of large-scale increments (from global DA) in high-resolution DA, c) Lateral boundary condition specification, and d) Relative performance of ensemble-based and climatological covariances.

WS-R2: MOSAC was presented post-processing results. Beyond the immediate (positive) impact, MOSAC recommends working to understand more fully what parts of the innovation are dominating the improvements (technique, number of predictors, length of training). MOSAC feels this could inform further steps (e.g. if truly get value from 2 years of training). How sensitive is this to model upgrades (more fundamental ones could invalidate results)?

We believe the length of training dataset is one aspect of the improved skill observed in the PS43 tactical implementation of machine learning (ML) in the current 'BestData' system. However, our current focus is on the development of the new, probabilistic-based IMPROVER system. This is a radical redesign, which will likely required the ML to be implemented in a different way hence we do not intend to spend too much further effort diagnosing the impacts in the current, deterministic approach. We will return to these questions in an IMPROVER context.

Overall does this change the thinking on the need for reforecasts?
More generally, MOSAC recommends revisiting the thinking on reforecasts in the light of the presented results, needs and ambitions of the IMPROVER project.

We believe the twin combination of large positive impacts from a) ML-based post-processed and b) The new online, analysis increment-based additive inflation (see below), together with the development of the probabilistic-based IMPROVER system validate the current strategy not to perform reforecasting in its traditional sense. The lack of a consistent, UM-based global, multi-year reanalysis dataset required to initialise the reforecast remains a key blocker (the analysis increment inflation technique appears to be robust for shorter periods of order 1 year or less).

PS43 was clearly a successful implementation, and MOSAC would be pleased to understand in more detail exactly what was implemented (e.g. in terms of the additive inflation, and its relationship to the existing stochastic physics). Furthermore, it notes that the inclusion of a bias correction term in the free running model (not just in the DA cycle or post-processing) is an interesting philosophical development for medium-range NWP. This approach has already shown its potential for seasonal prediction with skill comparable to actual state-of-the-art seasonal prediction systems like the one at the Met Office (Derome, Lin and Brunet, Journal of Climate, 2005: Seasonal forecasting with a Simplified General Circulation Model: predictive skill in the AO and PNA). Attempts to correct moist process biases with a similar SGCM approach has shown to be unsuccessful (Hai Lin, personal communication).

We acknowledge the philosophical development that comes from applying a bias correction term in the ensemble forecast itself. The advantage that this has over a reforecasting or calibration approach is that by correcting the analysed prognostic fields within the forecast, this can reduce the impact of biases in those fields on other prognostic or diagnostic variables. This was demonstrated during the development stage of the project, by applying these corrections to a free-running climate simulation and showing large improvements to the climatologies of fields not directly corrected, such as precipitation and the top-of-atmosphere radiation fields. We have also studied the impact of this correction on the mean errors of the ensemble itself. Its impact is clearly positive in fields with large biases such as lower-level (850hPa) temperatures over continental land during summer, but is marginal in other periods.

WS-R3: It would be interesting to see more detail on how large an impact the bias correction term has (e.g. how do global bias maps late in the forecast compare with and without the correction term)? Particularly if the impact is large, it would be interesting to provide some reflections on implications for the seamless strategy (e.g. would such a change be contemplated for seasonal or climate?) and the testing strategy (will it be harder to upgrade Numerical Weather Prediction model, e.g. with physics change targeting reducing bias if this is already doing it? How will bias terms be kept up to date when model does change)?

In terms of the impact on model development, the current approach is to generate a 1 year archive of analysis increments, that is reused from year to year, which we regenerate following an upgrade to the model physics. This has already been tested once, as the system was developed for use with the GA6-based UM configuration, and implemented with a GA7-based one using upgraded increments. Our proposal is not to use any bias correction during tests to develop new model configurations, and we continue to run a deterministic forecast without these corrections, which allows us to monitor our biases in the operational system.

The testing and continued development of the CoMorph convection scheme in increasingly realistic settings is showing excellent progress, and the Met Office team is to be congratulated. The added value from the external partners in the PARACON project was less visible, and it would be helpful to have an update next year.

We would be happy to provide an update on this next year.

WS-R4: It appears that the introduction of the new modelling framework and dynamical core in NGMS will reduce the speed at which new parameterisations and other components can be introduced in the operational models over the next two years. Since there may be a backlog of improvements to the model physics by the time NGMS is ready, it will be important to prioritise their introduction in future parallel suites to ensure that changes in predictive skill can be traced to changes in the forecast system. Introducing many changes at once will make it difficult to trace changes in predictive skill (positive and negative) to modifications in the forecasting system.

It is true that there will come a point when updates to the physics parametrizations are no longer targeting the current UM but that is not expected to happen at least until well into 2021. The GA9 global model science configuration is targeting the UM and is not due to be frozen until December 2021. GA10, which targets LFRic, will follow. Also, to minimise the disruption to physics developments, we have deliberately adopted the strategy to directly couple the current UM parametrisations into LFRic (the physics within LFRic are built directly from the UM code repository).

Based on a long history of the pitfalls of large model developments, a central tenet of NGMS is to avoid introducing more changes at any one time than are necessary and hence a staged approach will be taken to the operational implementation of NGMS components. For example, NG-OPS is currently planned for implementation in 2023, before introduction of the new model or data assimilation, thus allowing differences in observations pre-processing to be separated.

WS- R5: The conventional wisdom that more observations are always required, for example surface data for high resolution NWP should be examined rigorously. Are current observations all used and well used? Are there experiments that can be done to best know what type and where new observations are needed? This might be a suitable topic for a future MOSAC meeting.

We agree that a presentation on observations use and impact in high resolution NWP could be a topic to consider at a future MOSAC meeting.

Finally, MOSAC notes the progress with the ambitious plans for next generation DA system and its operational environment and encourages very close coordination with the model development projects, particularly in terms of understanding the options for and implications of different levels of JEDI adoption.

Climate Science

The Hadley Centre is a world-leading centre for climate science and Professor Albert Klein Tank gave an overview of future plans for the Hadley Centre, to keep it a world-leading centre.

One of the most striking results of recent years has been the high climate sensitivity (> 5C) from the Met Office CMIP6 climate model. It arises from a combination of model radiation scheme improvements (which increase the CO_2 radiative forcing), and revised cloud-aerosol and cloud microphysics parametrisations. In a workshop immediately preceding the MOSAC meeting, it was shown that these cloud revisions improved 6hr weather forecasts, making a very strong link between MOSAC and SRG activities and reinforcing the importance of the seamless philosophy.

Whilst the ECS is larger than that seen in the previous generation of models, none of the model's forcing or feedback processes are found to be atypical of models, though the cloud feedback is at the high end. The relatively large ECS results from an unusual combination of a typical CO_2 forcing with a relatively small feedback parameter (see SRG-papers). Cloud feedbacks in general, and cloud-aerosol interactions in particular, are found to be the most likely contributors to the high ECS values for several other CMIP6 models too. We have contributed to community papers on this topic (Meehl, Senior et al., 2019) and documented our findings in the JAMS special issue here:
https://agupubs.onlinelibrary.wiley.com/doi/toc/10.1002/(ISSN)1942-2466.UKESM1

Another area where seamless thinking will be important is in the so-called signal-to-noise paradox whereby the correlation-based ensemble-mean skill estimates on sub-seasonal to decadal predictions are significantly larger than correlation-based estimates of ensemble spread. As discussed during the SRG, an expedient to overcome this problem is to run very large ensemble sizes – this effectively filters out all ensemble variability. As an example, skilful decadal predictions of the NAO were shown using 600 member ensembles! Although this problem only becomes manifest explicitly in initial-value problems, it could be manifest in our estimates of regional climate change in the sense that the signal could be much larger than models are indicating. One possible explanation is that the problem is a manifestation of nonlinear model error, whereby persistent quasi-stationary flow regimes over the Euro-Atlantic sector are being modelled poorly. Increasing model resolution (potentially substantially) may be the only cure.

We note that this paradox is one of the drivers for the “ensembles” theme in the new Research & Innovation strategy.

UKCP18-package is now complete and opens the prospect of new investigations of climate change at regional and local scales. The impact of increased resolution was shown, whereby climate-change induced increases in UK average wintertime precipitation were some 27% in a convection-permitting model, compared with just 16% in a more conventional regional climate model.

CS-R1: MOSAC encourages further evaluation (against long-term observations) of the convection-permitting model simulations for various parameters (cloud fractions, screen-level temperature in cities, etc.) at local and regional scales and at sub-daily and daily scales. This may provide evidence for biases in the model that are less visible in NWP and may trigger feedbacks on atmospheric and land surface physical parameterisation development.

Evaluation of CPM simulations are part of the Climate Programme but also included in the Newton/WCSSP programme in which time slices of CPM results for various domains across the world (S-America, Africa, etc) are evaluated against observations.

The UNSEEN methodology, whereby large ensembles of seasonal and decadal integrations are run to provide counterfactual alternatives of the recent past, is an important methodology for determining the sorts of extreme events associated with a given return period. In this way, the model may provide better statistical estimates of the real world than the observations themselves do! Of course, such a result is entirely predicated on the accuracy of the model, and the model error associated with the signal-to-noise paradox could potentially impact on the accuracy of the UNSEEN methodology.

For this reason the model performance is always evaluated against observations for the metric of interest to assure that UNSEEN results are robust.

All these issues raise the question of whether the MOSAC and SRG should continue to meet separately, as there are important scientific issues which are not being communicated well between the two groups.

CS-R2: Alternatively, the agendas can be restructured to focus on making the most critical talks joint talks.

Agreed that communication between MOSAC and SRG on these topics is important. Note that UNSEEN has been the topic of a plenary presentation some years ago. Following recommendations of MOSAC and SRG we selected climate services as the theme for a plenary presentation this year. Papers documenting the progress on UNSEEN have been developed for SRG and shared with MOSAC at the meeting.

An ongoing problem for the Hadley Centre is how to balance the requirements of moving to higher and higher resolutions, with the need to include and develop more and more Earth-System complexity. There is no easy answer to this. Different versions of the model need to be maintained. One crumb of comfort is that studies (e.g. under the EU PRIMAVERA project) have shown that some aspects of high resolution (e.g. tropical cyclone intensity and regime significance) can be partially mimicked using stochastic parametrisation. It was noted that stochastic parametrisation is an area that the Met Office has chosen to delay temporally, for one year or so, because of LFRIC implementation logistic.

This classic dilemma in model development includes ensembles size. Our approach has been described in the R&I strategy (Box…) and is further elaborated on in the Climate Science Roadmap (SRG paper Annex II).

Applied Science

MOSAC is impressed to see that Applied Science and Business Group science activities are continuing growing. This is a fast-growing section with 19 new employees. The examples presented (including posters) demonstrate the strong dynamics of Applied Science. MOSAC is pleased to note of the new Strategic Priorities Fund UK Climate Resilience programme in which the UK Climate Resilience team can take a leadership role.

The variety of work being carried out by applied scientists and scientific consultants in *Applied Science and Scientific Consultancy* and in *Business Group* is extremely large. In order to provide expert added-value of applied scientists on such various topics, beyond classical data post-processing, specialized expertise in each topic may be required. This expertise could be unavailable or not easily mobilized within the Met Office research departments without deviating these scientists from their main longer-term research objectives whereas external hires may present challenges in providing training on significant scientific content relevant to each new employee. This is clearly a challenge and a concern for the continuation of growth.

We recognise the issue flagged, and it something we continue to monitor carefully. When the Met Office receives a request for a service that requires some R&D input, the business team, before any proposal or contract is formulated, engages with Science through what is called “the Green Team” process and ascertains what technical work needs to be done, whether we have the appropriate level of skills within the Met Office and whether that resource can be made available. It is only then that the Met Office will commit to trying to service that work. This process provides safeguards to mitigate against the risks described above. Looking to the future: (i) the SPF climate resilience programme is developing a network of scientists from broader backgrounds to develop new climate services, and (ii) we are currently expanding our Met Office Academic Partners to develop partnerships across a wider range of sciences.

AS-R1: We encourage stronger engagement with the MOAP Universities and wider academic community in terms of providing expertise across a range of disciplines. The addition of two new MOAP partners might help in this regard. The balance between the variety of topics and domains, the investment in training of applied scientists and the resources from other departments must be carefully considered. MOSAC recommends the identification of areas where the Met Office has a strong capability, niches where the return on investment would be highest, as well as areas where too much investment would be needed. A presentation of this analysis for the whole Met Office portfolio would be welcomed at the next MOSAC meeting.

We strongly agree with this recommendation. Applied Science is significantly involved in the call for additional MOAP partners. The competition for new MOAP partners identifies 'hazard to impacts' as one of the three areas for bids. Furthermore, the SPF Climate Resilience programme is building new partnerships, for example in social sciences, as presented by Lowe and Hewitt.

There is a clear intention and progress on better scoping and understanding customers and stakeholder's vulnerabilities and exposure to weather and climate hazards. This is another important challenge for Applied Science which requires developing partnership with social scientists. The SPF UK Climate Resilience brings together UK research councils across disciplines, including the Economic and Social Research Council, and is an opportunity to deepen such interdisciplinary partnerships.

The Dengue fever warning service for Vietnam is an interesting project with clear potential societal benefits. Based on the early career scientist presentation, the connections with the local stakeholders could be improved and might benefit from a different approach. For example, it is not clear that the system was designed with the reaction time for the local government to act. And so, although the work is novel it may not be provided in the most helpful manner.

AS-Q1: For instance, would shorter range forecasts (e.g., a prediction of the risk of a mosquito boom) be provided a week or two before its onset so that local authorities could monitor a few sensitive spots to both validate the information and gain further trust in the early warning system? Is the issue of maintenance and evolution of the service after the development of the system considered?

We are grateful to MOSAC for these specific suggestions. These are important questions that we are discussing with the sponsors and forms part of the requirements gathering exercise being undertaken with the Vietnamese users.

The added value of a sub-kilometre model for forecasting convection over London airports is an open question. Plans are to extend the domain size at 300m resolution and to increase horizontal resolution to 100m for improving deep convection forecasting. These resolutions are inside the grey zone areas for shallow convection and turbulence on which several academic and Met Office physical parameterization teams work.

The possible benefits of moving from 300m to 100m resolution is an important research question, and we agree that the MOAP partners have important skills in this area. We shall look to align our research plans.

AS-R2: MOSAC encourages the applied team to make stronger use of the Urban Working Group which brings together Met Office staff and academics working on urban topics to ensure alignment of the spread of activities internally and externally with the UK community.

We agree that the coordination of urban meteorology could benefit from further coordination, and we shall look to the urban working group to provide a forum for this activity.

Several of the examples shown are clearly at the cutting-edge of the research. The rail adhesion service is an outstanding example of multidisciplinary research work. It also raises the questions of the evaluation of such service and of how to consider the mitigation measures, that are two non-trivial issues. The PhD work on convective nowcasting with machine learning is another example of cutting-edge research.

AS-R3: Regarding the crop statistical modelling, it could be relevant to take into account the change in CO2 too when studying the climate change impacts on the crop varieties.

The change in CO2 was indeed factored into the calculation. In fact, the American Maize variety that was considered in the study is a vegetable that has a saturation threshold of absorption of CO2 of around 400ppm and so future increases in CO2 would not have a significant impact on this particular crop yield.

The role of Operations to Research

MOSAC finds that the new Operations to Research formed in November 2017 and first reported to MOSAC 23 are working well and have fostered a genuine cooperation between forecasters and researchers. Evidence of this was seen in how often O2R was cited in the counterpart R2O report, Strategy for UK NWP (MOSAC 24.8): identifying and prioritized actions on such issues as lack of spread in the MOGREPS-UK, UK screen-level temperatures and fog/low cloud errors with Op. Mets. MOSAC suggests that a similar O2R approach would be a good addition to the environmental systems.

The nowcasting development is proceeding well. Although it is not scheduled to go into service in 2025, it is not too early to begin planning for the next stage.

O2R-Q1: Relevant work is going on in regional ensemble development, blending with other data in IMPROVER, and machine learning methods for post-processing, but what is the strategy to integrate the new nowcasting system with these projects?

We will take a staged approach to the development of a state of the art, Met Office National Capability Nowcast system (NCNow) by 2025 by ensuring full exploitation and integration (where appropriate) of the IMPROVER system and machine learning capabilities. Initially an existing system, Met Office NOWcasting (MONOW), is being considered as a candidate for producing observation based nowcasting services. MONOW could deliver customer bespoke solutions direct to clients or provide data to generate a post processed blended mix of nowcasting information, NWP and machine learning capabilities. We foresee this latter function potentially running on the next generation post processing system (IMPROVER) or via a future data platform

O2R's engagement with NOAA's National Severe Storms Laboratory in running a convection-allowing ensemble in NSSL's Hazardous Weather Testbed (HWT) is an accomplishment towards the need to "learn lessons from other global forecast centres to reduce the need for developing in-house capability." (MOSAC Paper 24.6, Objective 2.4) MOSAC encourages continued initiative and progress in this area.

O2R-Q2: The use of ensemble and probabilistic information is a strategic goal that touches on many of O2R's activities. Is there a coordinated strategy to solicit feedback from operations?

The MOSAC advises that ensemble sub-setting for forecast generation should be approached with caution. While it is common practice in operational tropical cyclone forecasting and has been shown to improve the deterministic track, in general the best ensemble members early in the forecast may not perform well later in the forecast, particularly for a regional ensemble.

The use of ensemble and probabilistic information is a strategic goal that touches on many of O2R's activities. We acknowledge MOSACs caution with respect to forecast generation via ensemble subsetting. We will focus effort on ensemble sensitivity analysis and the opportunity that ensemble datasets offer for information extraction, an extensive and currently untapped resource. The effort will be directed at gaining understanding of the drivers of a forecast's uncertainty and, more importantly, the key factors early in a forecast evolution that are responsible for an outcome of interest. Subsetting may be an option to be investigated as part of the above effort but will be an activity consequential of rather than the sole purpose of the work.

O2R-R1: Science is required to understand when ensemble sub-setting could be appropriate and to what extent automated approaches could be applied, including post-processing to calibrate the ensemble spread.

O2R-R2: Work on radar and nowcasting, integrating other observations and NWP prognostics into ML is impressive and should be continued.

Ocean Forecasting

The UK and UKMO have a long and successful history of ocean forecasting. Much of the pioneering work on tides, surges, and waves originates from the UK. There are many activities within this section which reflect this long history and the shift of the more traditionally weather focussed centres to earth systems centres.

OS-R1: with the emergence of earth systems MOSAC recommends an increased coordination at the design and planning phase of the systems. For example, when considering R&D on components of any of the earth systems model that could feed back into another system early interactions could help align R&D plans for all groups involved (e.g. work in the Planetary Boundary Layer which could affect the ocean surface). Another example is the uncertainty associated with JEDI, the coupler, DA of the wave and ocean systems, or the interface with LFRic.

We agree that broader coordination across Earth System components is required at an early stage in NGMS. The NGMS-Marine project has been in place for over a year, and represents the views of the ocean community in the NGMS programme. We look forward to continuing to discuss this topic with MOSAC as NGMS develops.

OS-R2: In the opening presentation on the science overview, coastal flooding and erosion were explicitly listed in the hazards to decision making. In Ocean Forecasting R&D documents provided, storm surges and waves are listed as one of the highest causes of natural disaster in the UK. As such and given the long history of the UK on R&D work in coastal flooding, the committee wonders if 1/10 of the ocean R&D group is enough to address the challenges of modelling coastal flooding. This includes the capacity to undertake wave-surge coupling and developing overtopping and erosion capacities. MOSAC would welcome a presentation on these issues at the next year's meeting.

We would be happy to present work in this area next year.

OS-R3: Waves are also an important component of the earth system. They are the physical boundary between the ocean and the atmosphere. It is somewhat unclear how the development of the wave system, its DA system, its addition to the coupled ocean-atmosphere-ice system, including the coupler itself, and the transition of all those systems to NGMS is supported and will unfold. Although it is understood that much of this work leans on the wave, ocean, and ice communities at large, internal capacity will be necessary on all fronts. MOSAC would appreciate to see next year plans related to this matter.

We hope that this issue will be mitigated at least in part because the current wave model (WaveWatchIII) is a community model implemented in numerous frameworks. However, we agree that there is a need to ensure the nascent NGMS-wave interface is fit for purpose. Plans are at a relatively early stage in the development of the wave component of NGMS, but we can provide a brief update on developments next year.

OS.R4: Some of the Ocean Forecasting team plans are ambitious. This includes the operationalization of an ORCA12 ocean forecasting system. Work appears well underway on this initiative. The validation against drifter however raised concerns. The committee strongly suggests the use of another dispersion model (e.g., NAME or Arianne) to help identify possible technical issues and the inclusion of the drifter analysis in model verification.

We agree plans are ambitious. The divergence observed in validation against drifters is common and has been acknowledged in several modelling environments. For example, studies comparing met.no, Met Office and other models currents being put through the same drift model with or without the same atmosphere driving giving quickly diverging results (https://odnature.naturalsciences.be/noosdrift/ has some resources which highlight the difficulty).

MOSAC notes the good collaboration with the NWP team on verification of ocean forecasts using the HiRA framework and the MODE method in the MET verification package. MOSAC is also pleased to note that statistical post-processing of ocean forecasts is underway and encourages collaborations with the IMPROVER project. Given the prominence of Data Science in the new strategy, investigations of ML to improve wave forecast is also positively noted.

Strategy for UK NWP

The numerical weather prediction capability for the UK continues to develop well, including the introduction of hourly cycling and land-surface data assimilation. A noteworthy strength of the process is the integration of feedback from forecasters through the O2R program. Numerous examples of this feedback were cited in both the report and presentation, showing that this is a genuine collaborative effort.

UKNWP-R1: The result that data assimilation can perform worse than simple downscaling in the new UKV configuration is surprising and raises important questions about how data assimilation functions in limited area domains. The solution may involve the use of a smaller domain to eliminate synoptic degrees of freedom that are only partially observed in the limited area domain but may also involve innovations in the data assimilation to preserve information on all scales. The impact of observations should also be explored. It is possible that the 4D-var data assimilation of radar reflectivity planned to be operationally tested this year may change the results. A significant research effort may be required to extract the potential benefit of data assimilation in the UKV, as discussed in the recommendations for Weather Science. We look forward to seeing how this interaction between research and operational testing progresses.

See discussion with WS-R1 above. Although the planned implementation of radar reflectivity assimilation in late 2020 is expected to provide a boost to UK NWP skill, we do not believe it will be sufficient by itself to reverse the observed degradation of UK DA. We would be happy to report on the new, combined UK DA improvement project next year.

UKNWP-R2: The strategy for sub-kilometre NWP seems to include two very different aspects: urban modelling (including the 300m London model to support fog forecasting, and higher resolution runs in research mode); and coupling with an estuary-resolving shelf-seas model under the UKEP project. We would encourage studies evaluating the benefits of the coupling for weather forecasting, and the proposals to introduce other environmental complexity such as river flow are to be welcomed. We would also be interested in having more information on the plans for urban modelling over UK, as work to date is focused on London where evaluation data obtained by academic partners is also available.

Extended UKEP evaluation, including impact on weather forecasting, is planned. We can provide more information on our urban modelling activities at a future MOSAC meeting.

Satellite Applications Section

This group is unquestionably world-class in its field. Evidence of this includes the recent introduction of all-sky radiances in operations, as well as the use of tools such as FSOI to estimate the relative contributions of different observation types. There is extensive ongoing work to continuously improve both the amount of data assimilated and its effectiveness in the DA system.

A presentation was given detailing the plans for observation processing in the Next Generation data assimilation system. The plans appear to be clear and achievable. Like all groups that are heavily involved in NGMS development, innovative research has largely been put on hold for about two years, and it is important that priorities are identified for research afterwards, since there may be some catching-up to do. It is possible that the uncertainty over the design and code structure of the new JEDI-based DA system will impact the observation processing work, although in the long term the use of the JEDI framework should reduce such risks.

Future of Operational Meteorology

The landscape for operational meteorology is rapidly changing with users wanting local/personalised information, new technologies becoming available to communicate that information, and new players entering the scene. MOSAC was pleased to see the deep strategic thought that the FoOM programme is putting into positioning the Met Office to continue to lead in the provision of services "to help you make better decisions to stay safe and thrive".

We would like to thank MOSAC for their supportive suggestions and comments on our FoOM programme, which is of high strategic importance to the Met Office.

MOSAC found very interesting the scenario analysis considering the uncertainty and the vision (Figure 2 of the paper). What seems clear is that as routine meteorological forecasting becomes more automated and commoditised the real value of the meteorologist will be as a trusted advisor applying expert knowledge in high impact situations and translating that knowledge to stakeholders in meaningful ways that support their decision making.

FOM-R1: As improved post-processing of model outputs and nowcasts become available rigorous validation of manual and automated forecasts should be conducted to determine which manual tasks can be automated without practical loss of value.

We agree with MOSAC.

FOM-R2: As improved nowcasting capability is developed and implemented MOSAC recommends extending the comparison of manual and automated forecasts to the National Severe Weather Warning Service, Terminal Aerodrome Forecasts, and other high impact forecasts and warnings to understand where forecaster vetting of automated decision support, rather than manual preparation, should become the norm.

We agree with MOSAC. These tests should form part of the assessments of verification and value. We shall need to look at how best to prioritise services for this work though as it could be extensive.

FOM-R3: The use of cognitive science to understand and sharpen forecaster decision making (for example, by addressing cognitive biases) would be valuable, and will likely require outside expertise. Also, in the social science realm, it would be beneficial to be able to quantify in some way the added value provided by meteorologists through their expert advice over and above what could be provided automatically or by centralised providers.

We agree with this recommendation.

FOM-R4: The FoOM programme has identified several obstacles preventing effective use of ensembles and high-resolution NWP in forecasting. MOSAC is very happy to see the R&I theme on "producing and exploiting ensembles" that aims to address these obstacles and would like to hear an update on progress next year.

FOM-Q1: The development of a greater set of diagnostics for convection, fog, and other high impact weather would assist in condensing large data volumes into information that can assist in forecasting. In principle these should be pre-calculated as part of the (multi-?) model post-processing. Is this in scope for IMPROVER?

We have plans to generate additional diagnostics in support of operational meteorology, such as convective diagnostics, as stated in the MOSAC paper. Model-derived convection diagnostics from UK and global NWP system will be used to give a model-based first guess for the likelihood and nature of convection. The current proposal is to develop offline research diagnostics and calculate these based on level 0 (raw model) and level 1 (re-gridded) data.

Studies in the US and elsewhere have highlighted the need for operational meteorologists to become more skilled in mesoscale thinking so they can better interpret the information coming from high resolution NWP. New knowledge on mesoscale meteorological processes from scientific observations and modelling experiments must be effectively pulled through to operations via the training program and development of tools to assist operational meteorologists. MOSAC is pleased to note the embedding of O2R in the Convective Modelling team, which provides a kind of business-as-usual testbed environment.

Careers for operational meteorologists need to evolve to meet the changing service needs. Recruits to the Met Office meteorologist program are of a very high standard; however, it is rare that one person can perform all service functions perfectly well. MOSAC notes the new roles that were defined by the FoOM programme of Met Expert (setting forecast policy, similar to Chief Meteorologist role), Met Advisor (putting weather and climate information into customer context), and Met R&D (developing new products and tools for operational meteorologist).

FOM-R5: It would be worth considering an additional role of "Met Scientific Subject Matter Expert" that works in both scientific research and operations. This may go beyond the proposed 50:50 postings mentioned in the paper – MOSAC would like to know more about the intention of those positions.

This is an interesting suggestion that we shall investigate as part of the review of professions and roles.

To elaborate, people often change careers from the operational side to the research side (and much less often someone starts out in the R&D side and moves to the operational side). Many key users value being able to understand the science behind the weather and climate that they experience and base important decisions on. This understanding can be provided by a subject matter expert who has deep process understanding built on scientific research experience and deep understanding of stakeholders coming from operational experience.

Such individuals would spend a large fraction of their time, perhaps the majority, in a scientific research role but would also be capable of providing additional operational support during extreme events and could provide expert advice to stakeholders that would normally be beyond the knowledge of most operational meteorologists. Their relationships with stakeholders developed in their operational role would give them an enhanced level of trust and effective communication with those stakeholders. They would also share their expert knowledge with the operational cohort through training and mentoring.

We agreed and this supports our position around creating surge capacity.

FOM-R6: Although the demography of the Met Office meteorologist/forecasters is favourable (i.e. young people more able to adapt), enough attention and resource should be committed to meteorologist training to equip people to perform well in these new roles.

Provision of training will be a key part of the FoOM.

Next Generation – Observation Processing System

Joint Effort for Data assimilation Integration (JEDI) infrastructure has the modularity to support the next generation observation processing system. The intent is to keep the actual Observation Processing System (OPS) science like the forward operators. The NG-OPS and NG-DA will be implemented within JEDI infrastructure. We expect that more effort will be needed with the latter. This implementation needs to be compatible with the global and LAM UM and NGMS.

NGOPS-R1: MOSAC would like to see more detailed plans for JEDI (OPS/DA) on this matter next year. The plans should include all the modelling, coupling and DA aspects relevant to weather, earth system and climate applications.

It is correct that the basic strategy for the NG-OPS project is to transfer the science of the current system into JEDI, with an early implementation (around 2022-23). However, there is still time to implement new science between then and the retirement of all UM-based components (2026-2027). Regarding NG-DA, the nascent project includes a coupled DA work package which will begin to look at the requirements for land-surface DA next year. We can report of status next year. Ocean DA is currently weakly-coupled and hence is in scope of the NGMS-Marine project rather than NG-DA. We agree that significant efforts will be required later in the NG-DA project, especially as the development of linear/adjoint models for NGMS-4DVar ramps up in coming years.

Air Quality and Atmospheric Dispersion

The Atmospheric Dispersion and Air Quality team put forward some options for model use over the next few years with the aim of maintaining their forecasts and services, while meeting the challenges of the NGMS and LFRic.

ADAQ-R1: MOSAC was asked to comment on these plans and is broadly supportive of the proposed 'option 3' which we understand will utilise NAME + UKCA for the near term under a modular framework. Given the context and the need for UKCA for global and regional climate purposes this does seem the logical pragmatic choice.

We would add that a benefit to pursuing UKCA in the NGMS is a continuation of the long-standing collaborations with UK academia and partners. We think it is vital for that to continue, given the rise of air quality up the agenda and the funding opportunities now coming (e.g. SPF Clean Air). These externally funded projects are providing a terrific opportunity for the section to improve their models and services and we would encourage continued engagement.

There does also seem to be the plan to continue with NAME into the NGMS, as it should be relatively easy to port and has a distinct purpose and users outside of the Met Office, so this does seem worthwhile.

The adoption of UKCA in NAME strengthens our commitment as it will become the chemistry and aerosol option in all our models. We certainly intend to pursue UKCA in NGMS. Discussions are already underway within the NGMS programme to initiate formal sub-projects to ensure both air quality and NAME are ready for NGMS.

Offline CTM approaches such as NAME offer complimentary options to fully coupled approaches such as NGMS. For example, more limited area studies, use of wider NWP driving data sets, source attribution and also offers greater flexibility for testing of new additions and sensitivities to changes. We therefore agree with MOSAC that there are complimentary roles for both approaches.

ADAQ-R2: Given the effort involved in developing UKCA within a modular framework, engaging with the NAME user group in the testing of NAME + UKCA modular is important.

We agree and thank MOSAC for this suggestion.

How to bring higher resolution into air quality forecasts seems something of a developing research topic. Links through to order 1 km, sub-km or Urban versions of the weather forecasting systems are not in place and it was not entirely clear what the plans are for offline and online chemistry. However, under the SPF Clean Air programme of work we recognise separate tasks on developing high resolution AQ forecasting; high resolution urban NWP; and statistical sub-grid scale post-processing to get to roadside scales.

The limited time available at MOSAC meant that we did not explore aspects of future higher resolution modelling as much as we might have. However, we do have activities/plans in this area. For example, the use of 1.5 km NWP is already routine for NAME non-AQ applications. Within Clean Air, a 2.2 and 1.5 km AQ forecast configuration is being considered for potential future operational forecasting. Sub-km applications, including both ~100m resolution NWP and building-aware approaches are also being considered. We would be happy to present this work in future to MOSAC.

ADAQ-R3: We look forward to hearing how these tasks will be integrated into an urban air quality system over the 2 years of the SPF programme.

The bar is being raised for future volcanic ash and emissions forecasts, with the request to provide 4D forecasts. This is a real challenge and the Met Office will need to collaborate well to meet it successfully. Appropriate use of new satellite products and having the right physics in models is going to be important. There are open questions about how many ash-related processes (such as aggregation, wet and dry deposition, resuspension, and appropriate grain size distributions) will need to be represented for accurate forecasts, as well as the optimal use of NWP output in driving the forecasts (e.g. sometimes orographic flows can be important in the distribution of ash). Moving to ensemble-based ash forecasts in an operational context will be important to represent the uncertainties in the forecasts and working with users will help define the most appropriate products and displays.

ADAQ-R4: MOSAC also encourages engagement with OFRD and the ocean community to explore the feasibility to use NAME numerical algorithms in the ocean. This could yield new collaborations (e.g., inclusion of stochastic processes for trajectory modelling) and strengthen the response capacity (even if the responsibilities are not all held at the Met Office).

This re-use of expertise and capability, with some additional development, from across the Met Office and beyond clearly has the potential to offer considerable gains for limited additional investment. We are aware of similar unifying approaches e.g. as done at CMC with the MLDP model which is very similar to NAME. We will investigate this and what developments might be needed. However, the challenging development programmes in existing priority areas such as air quality and volcanic ash mean that we have very limited resource and scope to explore new activities.

Met Office Cardington

MOSAC recognises excellent quality of science in the key areas of instrument development; significant field campaigns (including international collaborations); upkeep of continuous measurements onsite. This group has had an excellent reputation for boundary layer science since the classic experiments of the 1970s. The reported work on fog observations clearly demonstrates good collaboration and experimental capability feeding into model development.

Collaboration with the wider academic community on field campaigns is good but could be more strategic rather than ad hoc. Benefits would clearly be mutual in terms of optimising use of instrumental resources (including those with NCAS Atmospheric Measurement Service) and providing early awareness of emerging scientific/technological advances or modelling needs.

MOC-R1: The present OBR Advisory Group sits entirely within the Met Office: we encourage the establishment of a UK Partnership around observations research. Its form could include a “light touch” committee (perhaps meeting virtually once or twice a year) leading to more focused workshops. These should straddle both modelling and observational communities, e.g. joint with the UKEP community.

We believe that given the remit of the current OBR Advisory Group – namely identifying and prioritising future observational research requirements and aligning internal effort across groups to ensure pull through – it is appropriate for this group to be internal to the Met Office. However, we appreciate the feedback that there could be a more strategic approach to identifying and engaging external partners to build collaborations around the strategic science and/or technological developments that the OBR SAG helps identify. In collaboration with the Science Partnerships team, we will direct some effort this year to identifying how this external engagement could be facilitated more effectively moving forwards.

Cardington and its staff play a key role in the development of UAV instrumentation as part of national capability, given its staff expertise and status as a CAA endorsed "danger area" permitting flight time. Development of such experimental systems requires institutions with longer term funding and influence such as MO and NCAS to drive the agenda with respect to influencing national guidelines on using such devices for scientific research as distinct from commercial applications and we are pleased to hear of MO efforts in this regard.

MOC-R2: Whilst there is good interaction with the academic community in the UK, we encourage continued collaboration and increased interaction with international groups as there is a lot of activity in this area now.

We agree that continued collaboration, including with international groups, should be an important part of our UAV development activity. In this domain, we have some active involvement and exposure to wider international work, for example through membership of the International Society for Atmospheric Research using Remotely piloted Aircraft (ISARRA). Indeed in 2017, we hosted the flying portion of the 5th annual ISARRA conference at Cardington which involved numerous international UAV systems operating at the site for knowledge sharing purposes. We will look to continue and where possible strengthen these interactions moving forwards.

It was not clear to the committee as to engagement in or contribution of data to international networks focused on surface measurements over a range of land surface types, or the effort required to do this, or how extensively the CEDA dataset is utilised. Generalisability of results from any site in terms of land surface model development is strengthened by intercomparing with other, similar sites.

Cardington data does not currently contribute to international surface monitoring networks such as e.g. Fluxnet. International collaboration regarding use of Cardington-generated surface data, and its contextualisation e.g. in terms of other surface types, is generally facilitated on a project-by-project basis. For example, in the upcoming LIAISE experiment this will be coordinated through the HYMEX project.

A subset of Cardington data is provided for wider community use through CEDA. While greater participation in established surface networks is agreed to be a good thing, we have not pursued this historically on account of prioritising our limited resource (particularly data scientist time) to other activities.

MOC-R3: Given the point above about finding a better mechanism for strategic collaboration the MOSAC committee would appreciate seeing a timeline from OBR of commitments to all activities (long-term observations and short-term campaigns), to identify capacity for wider involvement or resource constraints.

We would be happy to provide this in the future and indeed the table included in the FS overview paper covered just a subset of future activities that the group is committed to. We will also consider wider distribution of this information as part of our response to MOC-R1.

MOC-Q1: Given Cardington's geographic position away from other Met Office centres, does the group see much in the way of secondment, etc. and feel engaged with other directorates or the academic community as much as is needed?

Cardington could play a role in wider outreach efforts – schools respond positively to showing pupils observational work - but also engage in undergraduate/postgraduate student research projects. This could be important in terms of supporting scientist career development and future recruitment across the Met Office.

There are some examples of recent internal secondments to Cardington (e.g. from operations for forecast support during LANFEX, and undergraduate summer students), however it is true to say that Cardington's geographic position has limited appetite for long term placements operating in either direction. Alternative activities are therefore undertaken to try to ensure engagement with other directorates. For example, collaboration with model development teams is supported by Cardington scientists undertaking short research visits to HQ typically several times per year, Cardington research plans (as with other plans in OBR) are jointly developed with APP to ensure strong alignment, and future Cardington activity is also covered within the remit of the OBR SAG. Links with Observations Research and Development are also actively maintained, particularly with a view to use of the Cardington site in future field trials.

Despite these activities, could Cardington be better engaged internally and externally? Yes is doubtlessly the answer. This is an area that we will continue to focus on enhancing in the coming year, noting though that this in itself takes time from staff involved which will need to be balanced with respect to other commitments.

Regarding wider outreach, Cardington staff have been actively involved in STEM activity over the last year (including local school visits) and maintain a keen interest in hosting undergraduate/postgraduate research projects, which have been fruitful in previous years.

Supercomputer 2020 + and Next Generation Modelling System

The MOSAC recognizes the Supercomputing 2020+ and NGMS efforts as ambitious and impressive. With a six-fold increase in computing capacity in 2022, the Met Office will remain the most compute-capable operational centre for weather and climate in the world, a distinction it has held since 2016 (Top 500). More importantly, vastly increased computing will enable the range of new science and modelling capabilities presented during the MOSAC and SRG 2020 reviews, and in turn help clearly establish the Met Office science program as world leading. We are encouraged by the commitment of resources (an average of at least 50 FTEs per year over the next 4 years) to the Next Generation Modelling System (NGMS) project, and by statements during the presentations that talent has being enlisted to fill these positions. While encouraged by these developments, MOSAC has concerns about risks that have been identified, primarily in poor computational performance and in the potential for JEDI-related reengineering costs for data assimilation.

MOSAC appreciates the openness and frankness with which problems relating to numerical stability and deficient computational performance of the current aqua planet LFRic prototype were presented. It is understood that this is a snapshot in time and that, further, solving the instability issues may further aggravate poor computational performance, currently reported as being a factor of five to ten slower than the UM. At best, reaching operational forecast rates with such a performance deficiency would require significantly more computer resources; at worse, if scalability were exhausted, no additional amount of resources would be enough.

The NGMS program's efforts at remediation appear to be well-organized and with a level of urgency and attention appropriate to the risk the computational performance deficiencies present.

NGMS-R1: The MOSAC recommends that NGMS establish and enunciate specific operational performance requirements for LFRic relative to equivalent UM configurations. To cast the requirement in the most relevant way, the requirement should be expressed as the maximum allowable increase in computational resources, compared with an equivalent UM configuration, needed to run LFRic at operational speed. A sufficiently detailed plan for identifying and remedying performance bottlenecks should include detailed performance monitoring and characterization (e.g. Roofline analysis) to identify inefficient code.

We agree and work will be undertaken to draw up a set of requirements. But the challenge of doing so for a system that has as many highly varied configurations and needs as the UM should not be underestimated, particularly when the target architecture(s) for NGMS are not yet known.

NGMS-R2: Since there is a significant increase in data communication costs in LFRiC, further work is needed to study how data transport can be restricted to the bits that describe real and useful information. This can be achieved in part through the use of mixed precision arithmetic (including 16-bit floats where appropriate). However, a more general assessment of how to optimise data transport issues is recommended. To the extent feasible and appropriate, MOSAC would appreciate being apprised of progress in meeting the requirements.

MOSAC commends the strong outreach and partnership with STFC in the development of the PSyclone domain-specific language and framework on which the LFRic software architecture is heavily dependent.

We think that this point refers to the large increase in data exchanges currently observed in LFRic simulations? If so then the leading order issue, as discussed at the meeting, is the pressing need to implement algorithmic changes to reduce this number – it is excessively high at the moment and maintaining that level of communication is **not** the target. Once that number has been reduced as far as possible (by algorithmic improvements) then will be the time to explore how many bits of information are actually required. But as mentioned at MOSAC we are implementing the necessary technical ability to permit mixed precision.

NGMS-R3: We urge the Met Office take steps necessary to ensure that development and support for PSyclone (and any other critical external technical dependencies) will continue uninterrupted through the life of LFRic. This also entails ongoing support for adapting PSyclone to generate efficient code on future computer architectures and other emerging requirements.

MOSAC congratulates the Met Office on commencement of the five-year ExCALIBUR project that will help ready Met Office modelling systems for exascale computing with development in programming models, system co-design, data science and career development.

We most definitely agree – this is an important partnership for us. Additionally, it is essential that the Met Office continues to develop an appropriate level of in-house expertise in this area.

NGMS-R4: We suggest exploring and making connections where appropriate with other forward leaning technology groups within the Met Office such as the Informatics Laboratory. MOSAC also recommends developing strong collaborative links with those partners in Europe who are developing similar dynamical cores, e.g. ECMWF and MPI/DWD.

We agree and would go further: NGMS is about much more than the dynamical core and we need to ensure that we have good connections with as wide a range of appropriate groups as possible.

The NGMS report and presentation to MOSAC expressed concern for "the impact that adoption of the JEDI data assimilation framework will have across the modelling system" and that a system review will be required. MOSAC understands the concern relates only to the use of JEDI to control the entire NWP system. Remaining issues for the original plan to use JEDI only for OBS and VAR involve only governance – raising Met Office requirements to their appropriate priority within JEDI – and contain no technical "show-stoppers."

NGMS-R5: With this in mind, MOSAC suggests proceeding with the original more limited VAR/OBS-only implementation until: 1) an evaluation of the potential problems and benefits of a full-JEDI system is completed; and, 2) until such time as redesign and reengineering efforts for a full-JEDI approach can be accommodated without draining resources from timely implementation of the NGMS system for operational use. Regular checkpoints should be established in the NGMS timeline to ensure that a full-JEDI approach will not be precluded by on-going developments in NGMS software and architecture during the interim. The MOSAC would appreciate an update on the direction and progress for JEDI in NGMS as part of presentations that describe the overarching NGMS architecture at our meeting next year. The update should also include progress using JEDI for ocean-atmosphere and atmosphere-land coupled data assimilation, including developments on moving from weak towards strong-coupling.

We wholeheartedly agree with this recommendation, which is our current strategy to managing this issue.

Hazard to Decision Making – an overview of climate services

There is clearly an active and growing area in climate services. These activities have several different drivers (e.g. fit to Met Office strategy and capabilities; user needs; funding priorities). With so many projects it was difficult for the committee to get a clear sense of where the 'centre of mass' of the portfolio lies.

We thank the committee for their views on this developing area. Pulling through science to climate services is a challenging task and is still in its infancy. In the Met Office there are several different parts of the office that produce climate services. As presented, we are working to provide a clearer description of our combined offering through a new interface group than involves all of these parts.

Organised by use type a significant fraction of the outputs on climate time-scales fit into the categories of:

- International development
- Risk quantification into building resilience and wellbeing
- Evidence to support policy development and use

We could also split our provision into two distinct categories: we provide both data services (which are typically build on by other climate service providers) and end user services.

We also highlight that there is a mix of maturity levels of our services from pre-demonstrator through to operational. For some audiences this is a preferred way of splitting the portfolio.

HDM-R1: MOSAC recommends that the Met Office attempt to map current activities under some broad headings, with particular consideration to where it has (or wishes to develop) particular scientific capabilities that is wishes to form a centrepiece of its offering. For example, seasonal forecasting capability could certainly be part of it, but if so, could it be done more systematically based on knowledge of skill (of multiple models) in different parts of the world?

We will prepare this and would be happy to present it at the next MOSAC meeting or earlier. However, one aspect of climate service development that we believe is particularly important is that they involve a strong user driven component. This means that for a given application we might wish to use a mix of scientific approaches. For instance, for an organisation interested in its exposure to particular hazards (let's say high temperatures) we might recommend using the UNSEEN approach to look at present day risk, UKCP to look at future risks and GLOSEA for early warning. Thus, we might picture the map as a matrix of user categories and the techniques used.

HDM-R2: In terms of RCMs, MOSAC urges care in understanding and communicating how their results should be interpreted, particularly when the regional and host models have very different configurations and characteristics (e.g. relatively to water budget and its closure).

We strongly agree. In fact looking at the water budget in the 2.2km UKCP model simulations was a really important part of the evaluation stage, and of our building up an understanding of how it behaves. In UKCP and EUCP projects we are going further than this and looking at storyline/narrative and large-scale weather typing approaches to better relate driving physics to regional response. We are also trying to understand how to use information from both global models (including CMIP) and regional models (including CORDEX) of different resolutions together to provide a more robust overall statement of the range of uncertainty for users of these models. We would value and specific advice the panel can offer on this topic.

MOSAC supports the Met Office's continued role in the Global Framework for Climate Services, seeking to develop approaches and standards for provision of climate services, building upon, amongst other things, experience gained through many years of international contribution around weather services.

HDM-R3: In terms of its own portfolio, MOSAC strongly recommends further consideration of the Met Office's place in the complex international landscape, and where it positions itself in the value chain in order to add most (sustainable) value alongside other efforts and actors (world bank, private companies, universities, COPERNICUS ...).

We agree this is a very important point, and one we will reflect on further. On a UK scale we will have definite plans to consider this with the UK community within the SPF UKCR programme, which has an aspect starting-up on how to develop the UK framework of climate services. Such a framework needs clear roles for public and private sector organisations, along with views on maturity and quality of services. The role of science in underpinning these services is part of the WCSSP projects and will be further elaborated on in the new Climate Programme 2021-24.