UFS R2O Project Review Criteria

Review Panel Meeting March 12-13, 2020

Reviewer 1

Please provide responses to the following questions, based on the UFS R2O Project written proposal; these responses may be updated based on the panel review materials and discussions.

1. Does the proposed Project include a set of significant and achievable forecast goals for 2 or more NOAA UFS-based numerical weather forecast application systems, with at least half expected to be ready and available for NWS transition to operations by the end of the project period?

   The proposal is clear about the five achievable goals in the 1-2 years with a longer-term vision over the 3-5 years. The five goals include i) a fully coupled ensemble MER prediction system ready for pre-operational testing, ii) a public release of the RRFS operating on the CAM scales, iii) retirement of the existing mesoscale models, iv) prototype 1-km forecasting system (WoF), and v) a fully coupled forecast model for research use.

   Among them, two of the goals i) and v) depend on the development of the fully coupled prediction MER system. This new system represents a major upgrade from the current operational CFS in a few significant ways including the model infrastructure, data assimilation, model physics and accessibility to the public. The technical plan on delivering a MER/S2S coupled prediction system is realistic to a large extent and addresses several key challenges identified, which warrants the likelihood of success.

   The accomplishments of the goals ii) and iii) are also important, representing a breakthrough in realization of the high-resolution and ensemble regional CAM forecasting over the North America. The efforts of unifying the NWS regional prediction systems by phasing out the legacy mesoscale models pave the path for future unification of global and regional prediction systems. While the RRFS may not be ready yet for transitioning into operations by the end of the 2 years, it is a critical step forward.

   Compared to other primary goals, the WoF development is still at the prototype stage, making it less likely an item close to operations in the near term. It is of great interest for sectors interested in severe and localized weather events; on the other hand, there are many foreseen challenges in resolving physics and computing issues.

2. Please comment on the overall scientific and technical merit of the proposed project.

Overall, the proposers have successfully identified the major modeling issues that exist in the current prediction systems such as near-surface temperature and precipitation biases, middle and upper troposphere flow, convection initiation and boundary moist processes. The proposed work is organized with a total of 12 sub-projects, each of which has clear goals in the 1-2 years. Given the scope of the project, the team has demonstrated great teamwork in coordination and integration of individual tasks. Specifically, the proposed project merits in the following perspectives:

- It couples atmospheric and land models unprecedented with the sea ice and ocean models in the NWS predictions on the S2S scales. According to the review panelists, the benefits of using such a fully coupled and ensemble system are evident in forecasting surface temperature and precipitation for a few days out and a longer time period. The fully coupled system ensures improvement in forecasting capability from the current CFS, which has been around for more than 10 years. The sub-project "Coupled Model Development'' discusses in detail about the coupling strategy of benchmarking and testing of the sea ice and ocean models.

- The new coupling system is interfaced with the CCPP (common community physics package). This facilitates leverage of the community modeling developments of model physics and aligns with the purpose of making the UFS as the community weather prediction systems.

- The proposed model physics development work including the sub-projects "Atmospheric physics" and "Atmospheric composition" are well-targeted on the known and prioritized issues in the atmosphere model. For example, preliminary work on the model cumulus convection parameterization with the Tiedtke scheme shows evident improvement in boundary layer convection initiation that justifies the needs for upgrades in model convection and moist process. Large improvement is also seen in aerosol predictions compared to the operational GFS model, and these improvements will be made available in the coupled GEFS system via the CCPP interface for transitioning into operations.

- Data assimilation is a critical part of the MER forecast for initialization and concomitant conditions. One of the main outcomes from the 2-year project is to deliver the next-generation data assimilation infrastructure called JEDI, with the aim of replacing the current GSI used for atmospheric data assimilation. The development of JEDI is closely coordinated with the coupled system development interacting with each of the underlying component models, not just for the atmospheric model. Thus, the JEDI system would provide a unified infrastructure for performing data assimilation across the coupled prediction

system. In particular, this project includes two new DA capabilities of: a JEDI-based marine assimilation system for ocean and sea ice application and initializing soil and snow moisture for the land model of the UFS system.

- The development of the RRFS enhances the CAM-scale forecasting capability over the North America with the FV3 dynamical core, an interface to the CCPP for model physics, and JEDI data assimilation same as the global system. It is a critical step toward unifying the global and regional UFS systems down the road and executing the transition from the existing mesoscale models. In addition, efforts will also be made in advancing the hurricane predictions by the HAFS and evaluating the 1-km forecasting skills by the WoF system, both of which may be incorporated into the UFS systems in the long-term vision.

- Lastly, the coupled forecast system will include land, ocean, sea ice, atmosphere physics and constituents, which shares the same model infrastructure including the CCPP physics package and the JEDI data assimilation system. These model developments extend the current GFS/GEFS to a full-scale Earth System modeling system with advanced weather forecasting capabilities ranging from several days to the sub-seasonal to seasonal long. The developed UFS coupling system offers the flexibility for the weather research community to contribute to the future development of the NWS operational weather prediction system.

3. Does the proposed project include an effective strategy to bring research-to-operations and operations-to-research (R2O2R)? Does it include engagement with the research community to achieve forecast improvements and engagement with forecast community to establish forecast goals?

   Yes, the proposed project includes an effective strategy to ensure the research-to operations and operations-to-research. The model physics developments will be integrated into the UFS prediction systems through the CCPP, which is accessible and contributed by the research community. The MER global system and CAM regional system teams will select and evaluate the model physics schemes from the CCPP to prepare for transition to operations. The project team is also engaged with the weather forecasters and stakeholders to prioritize the modeling improvements and developments. In addition, the UFS systems are released to the forecasting and research community through GitHub. This facilitates the community-based evaluation and testing of the system and feedback to the UFS development team.

4. Does the proposed project adequately foster code unification, contributing towards the reduction of the number of code versions in use by the UFS and NCEP?

Yes, the proposed RRFS includes the implementation of FV dynamical core and the JEDI data assimilation at the CAM scales. The use of the common model infrastructures represents a step towards the unification of global and regional prediction systems. The realization of the RRFS will also pave the pathway to retire the existing mesoscale modeling suite and simplify the code versions for regional forecasting.

5. Does the Project have a clear management and organization plan?

   The project has a reasonable management plan, although it is noted that the project leads do not seem to have specific FTEs allocated for project management. There are some concerns over lack of coordination about the computing resources and keeping track of sub-project progress on the high level. Some sub-projects such as "RRFS and retirement of legacy models" have more detailed timeline and milestones than others.

   The present format in which the total of 12 sub-projects is organized includes some overlaps in the tasks. For example, both 1.1 "Coupled RA and R&R" and 1.5 "Data Assimilation" talks about the JEDI development. Also, the stochastic physics is included in both 1.2 "Coupled Model development" and 1.6 "atmospheric physics". Additionally, marine model (MOM) developments appear in both 1.2 "Coupled Model development" and 1.4 "Model Components", although they may not address the same issues. It would be great to eliminate the duplications and streamline the project structure.

6. Is the proposed project appropriately aligned with other modeling efforts, including the broader UFS community, other programs/projects at NOAA, and with the broader national/international NWP and modeling community?

   The upgrade of the GFS/GEFS system to the fully coupled system including ocean and sea ice is definitely aligned with the efforts by the major weather forecasting centers for the MER/S2S predictions. The proposed work on data assimilation and atmospheric composition in this project also complement with the OWAQ-supported programs on "Marine/land/atmospheric composition observation", "Aerosol data assimilation" and "Aerosol and constituent modeling".

7. Has the project established appropriate success metrics?

   The project has clear goals for deliverables in the 1-2 years. The overall success of the project depends upon the readiness of the coupled global prediction system for MER/S2S, and the public release of the RRFS regional system. For each sub-

project, the success is measured largely by improvement from the new model physics or system against the current versions. However, the coupled global system is not really comparable against the CFS, as the latter is not a fully coupled system and by nature inferior to the new model. It would be desirable to establish a set of defined metrics for evaluating the coupled global system as a whole, i.e., comparing with the ECMWF or the UK met office model, which belong to the same generation of the modeling system.

8. Does the proposed project effectively leverage internal resources, including e.g. time commitment of federally funded NOAA investigators, computing resources, in-kind contributions, and other funded efforts?

   Yes, the proposed project does a great job of leveraging the resources from other NOAA funded projects.

9. Does the proposed project have adequate and appropriate resources, including sufficient time-commitment of team members?

   As mentioned previously, the sub-project 1.1 “Coupled DA and R&R” seems to be short on staffing with 6.25 FTE TBD. Otherwise, the project resources are reasonable.

10. Does the project team have adequate and appropriate expertise to achieve the proposed goals?

    Yes, the project team has appropriate expertise for the proposed work.

11. Please provide a brief overall assessment of the proposed project and an overall score for the proposed project, where 1 is poor and 10 is outstanding. You may also provide ranking for sub-areas or aspects if you like.

    Redacted

    The overall score is ## out of 10. The proposed project is structured with clear goals for the first 2 years and the 5-year vision in purpose. The primary outcomes

include a ready-for-pre-operational testing version of the coupled ensemble global prediction system for MER/S2S and a public release of the regional CAM system. These two developments constitute major milestones towards the unification and simplification of the UFS systems with advanced forecasting skills on global and regional scales. In particular, the development of the coupling prediction system (L-O-SI-A-Ae) will bring the UFS up to the level same as other state-of-art MER weather forecasting systems. The project team did a great job of assembling and integrating the sub-projects to address the model biases identified in the current prediction systems. There are some overlap in tasks between the sub-projects that may be better streamlined. It seems appropriate to prioritize some sub-projects or tasks in the sub-projects to the year 3 and after. It is highly desirable to establish a set of defined metrics for evaluating the coupled global MER system with observations or other fully coupled modeling systems.

UFS R2O Project Review Criteria

Review Panel Meeting March 12-13, 2020

Reviewer 2

There is an overall need for better coordination and, in some cases, consolidation of efforts. The main examples pertain to the development of coupled data assimilation systems, and specific CAM applications, in which individual subprojects working in similar or complementary areas should have more cohesion. These are noted in some of the specific comments below.

NOAA's commitment to making UFS development open and accessible to the community through these collaborative projects is highly commended. University partners have the potential to make significant contributions to this effort, with the key to success being their commitment to conduct research and development *constrained by NOAA's operational imperatives*. Nevertheless, there are obvious and inevitable tensions between operational and academic priorities, including their respective metrics of success. This, in my view, presents a significant risk to maintaining UFS project coherence, and one that program managers and project leaders will need to monitor and, likely, mitigate, continually. It also bolsters the argument for articulating more explicit metrics and targets for forecast improvement wherever possible.

There are other overarching issues that will affect UFS success but lie outside the scope and funding limits of the current proposal. Access to computational resources is paramount here. The recent NOAA procurement of a 12-Pflop (initial configuration)

machine will be a big boost, but even this resource will likely be stressed in a relatively short time. And as the sophistication of the operational version of the UFS increases, the computing needed to run representative test systems, even at reduced resolution, will increase dramatically. As a reference, historically, ECMWF devotes 70-80% of their compute capacity to development. Several of the UFS subprojects in this proposal point to as-yet-undefined cloud computing resources to address these needs.

In a related vein, NOAA must invest in research and development for advanced I/O handling, including data compression algorithms, increasing communications bandwidth, etc. so as to ensure that post-processing and validation activities can keep pace with the increased rate of science development that the UFS project will likely enable. This has direct implications for the planning and utility of the critical verification and post-processing functions proposed under subproject 3.2.

While absent from the present proposal, it is strongly recommended that UFS-R2O project work include preparation for advanced computing environments, including exascale capability. Effective use of UFS resources will require additional investment in this area, including R&D into co-design software development, the reworking of major elements of the FV3 code base, and effective management of future data sharing and archiving. Again, the science demands of the UFS (in particular the increasing emphasis on ensemble-based, coupled systems) will put great pressure on NOAA's compute resources, notwithstanding the recent good news about its next operational upgrade. While code optimization is clearly a part of the UFS-R2O project, flexibility in modeling infrastructure also means that science capability and development can proceed quickly and exceed the optimization capacity to support these enhancements. Just as community development of the UFS will be beneficial for developments in modeling and data assimilation, similar joint efforts are required to address the scalability and exascale capability of the system. While it was mentioned at the panel review that such work is being undertaken in other parts of NOAA, it should be noted that centers like ECMWF and the MetOffice have well-established efforts in this area, which are well coordinated with the development plans for their next generation operational systems. Within the context of the current proposal, this work would be a preferred focus of the FV3-target component of subproject 1.4 (Model Components).

Subproject-specific comments

Subprojects 1.1 and 1.2

Development of a coupled Earth system (initially atmosphere-ocean-sea ice) data assimilation capability is an important strategic direction for NOAA and should be a

priority for the UFS. I view subprojects 1.1 and 1.2 as complementary efforts in this regard (along with subproject 1.5, to be discussed later), with 1.1 focused on data assimilation and the delivery of a reanalysis for S2S applications, and 1.2 focused on the development of the requisite coupled model. However, there is also a clear and necessary intent to include data assimilation experimentation in 1.2. It was confirmed in the presentation to the panel that DA integration and testing are indeed integral to this subproject. However, the two projects make no reference to one another and there is some disconnect in the timelines. A concern is with the timelines in subproject 1.1, including a likely underestimate of the time and effort required to produce even a weakly coupled (WCDA) system that performs well, both technically and scientifically. There are many aspects to the exact configuration of a weakly coupled system that need to be determined, including the choice of windowing in each medium, the choice of DA algorithm that works best for each medium and in conjunction with one another, etc. This may take considerably longer than a year, with extensive testing required at low resolution before moving on to high resolution as proposed in year 2. In the same vein, the proposed move to strongly coupled DA in year 2, even as preliminary research, may be overambitious and a distraction from the development of the much more practical and obtainable WCDA system. This is especially of concern given the plans to conduct a scientifically viable weakly coupled reanalysis in year three (since even uncoupled reanalysis usually take a few years of system testing, tuning, bias mitigation, observation preparation, etc.) This is also where I think there may be a disconnect with the timeline in subproject 1.2, which proposes to be conducting extensive development of the coupled model through years 3 and 4. That timeline seems realistic.

The other critical component for the coupled system development is the JEDI data assimilation infrastructure proposed in subproject 1.5. Given the reduction in the program budget, it might be recommended that subprojects 1.2 and 1.5 be viewed as the primary vehicles for developing the UFS coupled data assimilation system, and that the reanalysis proposed in subproject 1.1 be folded in as a specific application, or deliverable, of this development. As stated in 1.1, additional future funding will be required to maintain an ongoing R&R capability, but it's not clear the current UFS budget should support 1.1 as an entirely separate project at the current time.

Subproject 1.3

The proposed work is well thought out and of high quality, however much of the underlying capability being proposed here in terms of aerosol DA has been developed to a high level of maturity at NASA GSFC. According to the subproject leads, there is an existing collaboration with GSFC, not funded by the UFS-R2O program. With this in mind, and given the reduced UFS-R2O budget, the immediacy for a UFS-R2O-funded project as proposed here is not clear.

Subproject 1.4

This subproject describes proposed R&D for three components of the UFS: the FV3 dynamical core, the Noah-MP land surface model, and the MOM6 ocean model. On the one hand, the grouping of these elements into a single subproject seems *ad-hoc*, while on the other hand, there are no obvious places within the current structure of the proposal where the individual elements might better fit. The adoption of Noah-MP is important, as it should provide significantly improved representation of land surface processes compared with the current Noah model. For MOM6, the developments targeted for *regional* applications should be prioritized, given the planned applications under CAM. In both cases, these efforts should be well coordinated with other foundational subprojects in the proposal, namely 1.6 (physics) in the case of the Noah-MP development, and 1.2 (coupled model development) and 2.4 (hurricanes, HAFS) in the case of the MOM6 regional development. The proposed developments for FV3, while relevant for climate applications and whole-atmosphere modeling, seem less critical, especially if funding for the UFS-R2O project is to be reduced. A redirection of effort toward the preparation of FV3 for exascale computing environments might be considered.

Subproject 1.5

Modernization and unification of the data assimilation infrastructure for all operational system components should be a top priority for NOAA and the UFS. This subproject lays out a reasonable, systematic plan for reaching this goal within five years, with likely achievable goals along the way. It seems that this subproject in combination with subproject 1.2 are critical for achieving NOAA's objectives in coupled Earth system DA. In the short term, the investigators emphasize the challenge of ensuring the conversion to JEDI of all currently used observation types, especially for atmospheric DA. In this context they briefly mention the additional investment needed for ensuring the conversion and convergence of capabilities for pre-processing, QC and other front-end data handling procedures. This is critical, especially for conventional data types, which are diverse and often require specific, non-uniform procedures for QC and data selection. Addressing this challenge will be vital to the success of JEDI, irrespective of all its other obvious benefits.

Regarding marine DA and hybrid GODAS, phase 2 calls for a 30-yr reanalysis by 2021, but phase 3 states that operational implementation of the system won't occur until 2023. Given the latter, will there actually be a reanalysis-worthy system available as early as 2021? If so, is there a concern that it will be of significantly lower quality than the 2023 system? Does this matter? Finally, is there any connection between the work proposed here and the R&R work proposed in subproject 1.1 on coupled DA? It seems natural that

there should be, but there is no mention of this here or in subproject 1.1. Am I confused about the nature/goals of these subprojects?

Subproject 1.6

This is a critical and generally well thought out section of the proposal which, if successful, could likely have the largest potential impact on GFS/GEFS forecast skill in the next 2-4 years. While I have remaining concerns about whether the proposed multi-agency task will work efficiently to develop the next generation moist physics for the UFS (in addition to existing WGs and other committees), I am sympathetic to the subproject lead's motivation: to get the existing physics team members to leave behind their individual long-time favorite schemes and begin working on something new and best suited for operational implementation.

Subprojects 2.1 and 2.2

These two subprojects, and especially the development of the RRFS, are keys to the important unification/simplification goals of the UFS (which are most sorely needed for the regional application suite) and should be top priorities for CAM and the UFS. In both cases the transitions from the existing systems to the proposed ones appear to be complex and arduous. In the case of 2.1, and specifically the development of the 3D-RTMA/URMA (including improving the analysis products and their efficiency and latency), the proposal text is short on detail. Another area of concern is the adaptation to JEDI. JEDI is clearly mentioned in both projects. However, both projects, and especially the RRFS development, appear to propose continuing significant DA development within the GSI framework. Is this necessary and strategic? If so, how can significant delays in migration to JEDI-based DA be avoided or ameliorated, given the proposed timelines for operational implementation?

Subproject 2.3

The utility of the WoF system appears to depend heavily on the feasibility of running at 1-km resolution, which is yet to be established. From a science and strategic point of view, might the RRFS itself provide adequate capability, with its planned 15-minute update cycle and high resolution? From a technical point of view, the investigators also make clear their concern about the difficulty and required time for the transitions from WRF-ARW to FV3 SAR, and then from GSI-EnKF to JEDI. Overall, there appear to be non-negligible scientific and technical risks.

Subsection 2.4

The work included in this subproject, and especially the development of HAFS, is an understandable priority for NHC. Much of the text in this subproject section describes the existing system components for HWRF and HMON, and general statements about which program funds and existing efforts will be leveraged to develop HAFS, but not as much about exactly how this will be done. Also, as in other subprojects in the CAM section, the need to transition to JEDI is mentioned but the plans for accomplishing this are not. Again, there appears to be significant DA work planned first in the current GSI-based system. I realize this may be necessary to some degree, but I am concerned about the redundancy of efforts.

Subsection 3.1

Strong support for modeling infrastructure is important and this subsection seems comprehensive and well thought out, overall. In particular, it seems that development and coordination of (a hopefully limited array of) workflows for UFS applications should be a near-term priority for this subproject. This is an absolute requirement to achieve the UFS vision of extensive community engagement, as people will only use and develop the UFS if they can easily run and adapt it.

Section 3.2

Similar to the cross-cutting subproject 3.1, support for a unified tool set for verification and postprocessing, with agreed-upon core metrics, is essential for UFS success both across NOAA and for community contributors. The list of proposed diagnostics given in the technical approach section of this project is lengthy and rather specific to current GFS and GEFS shortcomings. This is laudable on the one hand, but perhaps a bit too targeted on the other. Some descoping may be necessary, especially in light of the fact that I/O bottlenecking already is a major concern, which will grow as UFS applications move to higher resolution and more ensemble-based approaches.

Subproject Ranking

While all the subprojects have merit for advancing NOAA's forecast systems, and most were well written and well presented at the panel meeting, the following prioritization is offered in response to the program managers' request for such guidance.

Redacted

Reviewer 3

The team has done the tough job of identifying the most necessary subprojects for bringing near-ready research into operations of UFS. Since the proposal is primarily organized by these subprojects, with separate milestones and effort tables for each, this review and answers to the questions below will also be primarily by sub-project.

1. Does the proposed Project include a set of significant and achievable forecast goals for 2 or more NOAA UFS-based numerical weather forecast application systems, with at least half expected to be ready and available for NWS transition to operations by the end of the project period?

High level goals for improving the forecast were given in the executive summary and in the presentation. In both, the emphasis for MER/S2S was to reduce biases and improve variability (the proposal specified which fields and modes), deliver a fully coupled ensemble prediction system with coupled DA, improve quantification of uncertainty in coupled ensemble forecasts especially at interfaces and advance the initialization through improved use of obs and advances in DA. For CAM, the high-level goals are to deliver a 3D RTMA, unify the regional CAM models, develop a Warn on Forecast system and establish connections between CAM and other systems (HAFS, NWM, CMAQ). Each subproject had its own specific goals.

1.1 Goals for UFS-based systems ready in 2 years: WCDA for GEFSv13
Other goals: a prototype SCDA, a low-resolution WCDA R&R.

1.2 All goals focus on the coupled GEFSv13/GFSv17 and they plan to have them ready at the end: Reduced biases in forecast, stochastic physics and tested coupling strategies for the fully coupled GEFSv13/GFSv17. Reducing biases may not be achievable since these are long-standing.

1.3 Plan to have aerosol data files, aerosol processing tools and AOD DA algorithms ready for both GFSv17 and GEFSv13. Reducing biases from including aerosols is not guaranteed.

1.4: Have improved land model and more consistent FV3 ready for GFSv17/GEFSv13. Other goals are FV3 ready for space weather apps and high-res MOM6 ready for coupled hurricane forecasts.

1.5: Have JEDI ready to replace GSI in GFSv17/GEFSv13 for the atmosphere with new obs. Also have JEDI-based marine and land (snow and soil moisture) DA. Other goals: 30-year ocean and sea-ice analysis, hybrid inter-comparison study for deterministic DA.

1.6: New moist physics and new gravity wave scheme ready for GFSv17/GEFSv13. Additional goals of identifying root causes of long standing errors is not realistic.

2.1: Deliver a 3D-URMA and a more efficient 3D-RTMA with better analysis products. Additional goal is to transition this system to JEDI.

2.2: Deliver UFS-based system for Rapid Refresh Forecast System for use in regional forecasts in GFSv17/GEFSv13. Other goals include moving the DA for this system to JEDI and expanding assimilation to GOES-16 all-sky radiance.

2.3: Have an FV3-based WoF system ready for use at the same time as GFSv17/GEFSv13. Other deliverables include moving the WoF DA to JEDI and assessing the cost/benefit of 1km WoF.

2.4: Deliver DA for HAFS in time to be used operationally. Other deliverable is HFIP experiment with both the old and new systems.

3.1: CMEPS ready for use in GEFSv13/GFSv17 and general SE support of GFSv17/GEFSv13 and SFS v1. Also new releases of CCPP and SCM. Other deliverables include user support for MR and Hurricane applications and the start of workflow consolidation.

3.2: New UEPP system for use with GEFSv13 and improvements to precip in NBM. Other goals include CAMPS development and a metrics workshop with a goal to simplify the visual presentation of multiple metrics.

2. Please comment on the overall scientific and technical merit of the proposed project.

Overall the proposal has a high level of scientific and technical merit. All of the work would, if carried to completion, result in a greatly improved UFS. The work on stochastic physics and coupled data assimilation are exciting areas of research. Warn on Forecast and high-resolution coupled forecasts are all leading-edge problems. The most admirable component is the bringing together of people from the research and operations community to work towards a common goal of improving the next generation of GFS/GEFS (in particular GSFv17/GEFSv13 scheduled for operations in Q2 FY24). Bringing research and operations together is important for NOAA to advance its mission and a jointly-proposed, jointly-funded proposal like this is a great way to start that transformation.

3. Does the proposed project include an effective strategy to bring research-to-operations and operations-to-research (R2O2R)? Does it include engagement with the research community to achieve forecast improvements and engagement with the forecast community to establish forecast goals?

The main strategy is to staff each sub-project with scientists from operations centers (primarily EMC) and research entities (NOAA research labs, Cooperative Institutes and universities). R2O2R happens organically as researchers from different communities work together over time. It's a good strategy provided the project leads can provide ways for all the parties (at separate institutions) to work together (See management plan comments below). The project will end roughly 18 months before GFSv17/GEFSv13 goes operational so it will be essential that each sub-project has a person in EMC who is familiar with its work and can carry it to the finish line.

The staffing plan as it relates to R2O2R and any specific R2O or O2R actions is commented on here. Overall effort is addressed later.

1.1: Of 9.35 FTE total, only 3.75 (40%) of effort is from EMC with no named collaborators. The SCDA is a more research-oriented part of this sub-project. While it is important to get WCDA in to the GEFS, this sub-project does not seem to have the necessary engagement from the operations side.

1.2: Of 18.1 FTE total, 13.1 (70%) is at EMC as are the sub-project leads. Their amount of time is not specified but this implies a mostly operational effort. The stochastic physics is more research-oriented.

1.3: Of 6.1 FTE, only 0.05 FTE is at EMC but Stajner is a co-lead. Data sets and tools for operations will be provided but not clear the case can be made to operations to take them forward.

1.4: Of 9.6 FTE total, 4.6 (48%) is at EMC however the ocean component has no EMC effort.

1.5: Of 19.7 FTE total, 11.25 (57%) is at EMC. This is one of the more balanced sub-projects and in addition has a specific O2R goal of doing documented releases of JEDI for the research community to use.

1.6: Of 7.12 FTE total, 1.8 (25%) is at EMC but Kain is one of the co-leads.

2.1: Of 17.6 FTE total, 6.5 (37%) is at EMC.

2.2: Of 37.6 FTE total, 11.85 (30%) is at EMC.

2.3: Of 6.5 FTE total, there is no EMC effort although Jacob Carley was listed as a co-lead in the presentation. But if NSSL also runs this model operationally, EMC effort may not be needed.

2.4: Of 27 FTE total, 8 (30%) is at EMC.

3.1: Of 45.3 FTE total, 30 (66%) is at EMC. 25 of that is identified as “Engineering Team” at EMC and it is not clear if all 25 are 100% on this sub-project. Specific O2R activity includes user support for MR and HAFS.

3.2: Of 23.55 FTE, 12.6 (53%) is at EMC.

3.1 and 3.2 supports both O and R since they provide infrastructure used in both.

4. Does the proposed project adequately foster code unification, contributing towards the reduction of the number of code versions in use by the UFS and NCEP?

To this reviewer, the key word in this question is “reduction”. While many of the sub-projects replace existing narrowly-scoped and “privately” developed codes with community codes, which is good, that does not always result in a reduction of code (with a notable exception).

1.1: Little direct code unification. Replaces CFS CDA. The reliance on JEDI should reduce the amount of custom CDA code that has to be maintained and I assume the JEDI software is used for both coupled and un-coupled data assimilation.

1.2: No code unification. Part of the larger plan to replace CFS with GEFSv13

1.3: No code unification. Adds emissions processing code and algorithms necessary to include aerosols in AD.

1.4: No code unification.

1.5: No code unification but working to replace GSI with JEDI.

1.6: No code unification.

2.1: No code unification. The JEDI transition should help reduce the overall code base but that will just get started here.

2.2: A main goal of this sub-project is to replace legacy systems with FV3-GFS/GEFS based systems. It won’t get there at the end but will enable that goal. There is also the transition of the CAM DA to JEDI.

2.3: Unifies the WoF system around FV3, reducing code (by removing WRF). Also removes a custom DA system and replaces it with JEDI.

2.4: Will help replace another non-FV3 system with FV3.

3.1: Will mostly replace code (CMEPS for NEMS) but the workflow effort could result in code unification and reduction.

3.2: UEPP is new but can be built on existing code. Not much code unification/reduction.

5. Does the Project have a clear management and organization plan?

The org chart for the project is reasonable with the sub-project leads reporting to team leads and then to the overall project leads. Confluence is a good choice for creating/hosting project-specific content but the leads should define a structure before rolling it out to help keep things organized. JIRA is a popular choice for issue tracking but does not have the friendliest user interface.

The telecom schedule might need adjustment. The Teams are too big to have weekly telecoms. The larger sub-projects should instead call weekly, with team leads sitting in, and not be left on their own. A team-level telecom could be monthly. The project should invest in a web conference system that allows video and screen sharing instead of just an audio call.

The presentation mentioned a project manager although one is not in the text or its budget. A project manager will be essential to a success especially if the leads cannot each spend 50% or more on this for the full 2 years. Besides tracking progress, the PM should be viewed by the project as a PI-proxy. Their phone calls and emails should not be ignored. The PM should work for all the project leads even if they technically direct-report to just one or to the program office.

The leads should add at least 2 face-to-face meetings (a kickoff and a mid-project) for all staff working on the project (funded directly and leveraged). This is essential for establishing collaborations within a new team. The overall and sub-project leads should meet 2 additional times to make any adjustments to the work plan. The project leads should also use a chat application such as Slack or Microsoft Teams for all project staff. Chat is a great way for geographically distributed teams to stay in touch and "hang out" virtually, building team comradery.

6. Is the proposed project appropriately aligned with other modeling efforts, including the broader UFS community, other programs/projects at NOAA, and with the broader national/international NWP and modeling community?

1.1: Aligned with JEDI and individual model DA efforts as well as development of the coupled model for GEFSv13.

1.2: Aligned with EMC coupled model development, Aligned with GFDL research (using MOM6, FV3). Helping NOAA catch up with international modeling centers using coupled models for forecasting. Stochastic physics is a research topic throughout the earth science community.

1.3 Catching up with international centers that have interactive aerosols. Aligned with NASA and Harvard GEOS-Chem and ongoing development of FV3.

1.4 Aligned with NOAA goals in space weather forecasting. Aligned with GFDL research in MOM6.

1.5 Aligned with multi-agency JEDI program and multi-agency JCSDA. Catching up with international centers that ingest land data.

1.6: Aligned with CCPP effort.

2.1: Aligned with nearly-complete 3D-RTMA effort from JTTI.

2.2: Aligned with JEDI and CCPP efforts and Hurricane forecast efforts.

2.3: Aligned with JEDI and CCPP efforts.

2.4: Aligned with various HSUP efforts.

3.1: Aligned with work at NCAR. Also the JEDI and CCPP efforts.

3.2: Aligned with users/developers of METplus and associated tools.

7. Has the project established appropriate success metrics?

Nearly all sub-projects identified appropriate success metrics. Nearly all the metrics were for judging success at the end.

1.1: functional and cycling WCDA system for UFS. Any improvement in OMF error for WCDA over regular initialization. SCDA is better than WCDA.

1.2: Compare forecast skill to uncoupled GFS v16 in many fields including ability to predict low frequency phenomena (QBO, MJO). Compare stochastic skill to GEFSv.12 (medium range) and UFS without stochastic (S2S).

1.3: Compare aerosols with observations for hindcasts. Better S2S skill then GEFS v12.

1.4: The metrics were not specific enough. In each model area, they are compared against current systems but what exactly will be improved? The FV3 work can check conservation. The land work is delivering a test system which when applied will hopefully show how to improve the land model.

1.5: Measure difference between new UFO and GSI observer. Improved skill when using JEDI and more efficient operation. Compare marine reanalysis with legacy GODAS products.

1.6: Compare forecast skill with old and new schemes.

2.1: Fit to obs and consistency compared to 2D-RTMA.

2.2: Performance of deterministic forecast skill relative to baseline using legacy models.

2.3: Completion of report, compare 2 WoFs with the same analysis and show improvement.

2.4: Successful delivery of tested and evaluated DA for HAFS.

3.1: New or newly-introduced infrastructure does not change answers in the UFS application or does for known reasons. New releases of ESMF, CCPP. User engagement measures such as issues opened/closed, github PRs and forks. Amount of infrastructure successfully transferred to operations.

3.2: Successful use of METplus and UEPP by UFS researchers. Addition of new features.

8. Does the proposed project effectively leverage internal resources, including e.g. time commitment of federally funded NOAA investigators, computing resources, in-kind contributions, and other funded efforts?

The project does not provide any of its own compute resources and will be leveraging both in-house NOAA facilities and, when possible, non-NOAA compute facilities. The TBD cloud-based capability is too unknown to judge if it will be able to provide the indicated hours. If it does not, that will mostly affect the data assimilation efforts (1.1 and 1.5).

There is very little time in the effort tables for the project leads. They should have enough support to spend at least 30% each managing this project.

1.1: 14% Leveraged effort: 0.6 FTE from JCSDA, 0.5 from PSD, 2.75 from EMC. (Non-leveraged is 5.5 from PSD, 2.0 from JCSDA.). Also leverages single-model JEDI-based DA efforts. R&R relies heavily on cloud computing. Overall 67% is TBD

1.2: 60% Leveraged effort: 6.6 FTE from EMC, 0.2 from ESRL-PSD. Non-leveraged is 2.1 at GMU, 2.7 at ESRL-PSD and 6.5 at EMC. Need additional computer time on NCEP operational computers. 90% (16.7 FTE) is TBD but that is broken down in to 9.6 (53%) existing staff and 7.1 (37%) TBH.

1.3: 40% Leveraged effort: 1.6 FTE from ESRL, 0.1 from ARL, 0.4 from JCSDA, .1 other. Non -leveraged is 1.3 from ESRL, .8 from ARL, 1 from JCSDA and .5 from NESDIS/STAR. 32% is TBD.

1.4: Only 1 FTE leveraged from EMC out of 9.6 total from EMC, UM, GMU, FSU and Rutgers. 41% is TBD.

1.5: Only leverages 1.2 FTE out of 19.7: .25 EMC, .7 JCSDA, .2 GMAO and .05 PSD. Non-leveraged is 11 EMC, 7.5 JCSDA. 14 FTE (71%) is TBD but that is broken down in to 8 (41%) current, but un-named, staff and 6 (30%) TBH.

1.6: This sub-project has the most confusing effort table. Leverages 0.9 (12%) out of 7.12. 4.5 (63%) is TBD but that is not clearly broken down. There is 1 new hire at PSD and a current staff at PSD. Other organizations just say TBD.

2.1: 60% leveraged effort: 5.5 from EMC and 7.1 from GSD. The leads have time on the proposal but everyone else (17.4 of 17.6) is un-named current staff (so no new hires necessary).

2.2: 70% leveraged effort: 7.25 from EMC and 18.1 from GSD. The leads have only 0.1 each. 86% is TBD but they are all identified as current staff. No new hires.

2.3: 54% leveraged effort all from NSSL. Non-leveraged is also NNSL. Also leveraging NSSL compute.

2.4: 66% leveraged from EMC and HRD mostly from HSUP funding. From budget table, all new effort (9 FTEs) will be in EMC.

3.1: 73% leveraged from EMC, NCAR/CGD and GSD. Non-leveraged is for NCAR/CGD, EMC, DTC-NCAR and DTC-GSD. EMC is all "current staff" including the TBD. Not clear if DTC TBD effort is current staff or new hires.

3.2: 33% leveraged from EMC, DTC and others. At least 9.25 is TBD including 2 FTE of new hires. There may be more TBD but some cells are totals of a named person and "other".

9. Does the proposed project have adequate and appropriate resources, including sufficient time-commitment of team members?

I'm assuming the computer time requests are sufficient as there is not enough information to tell if they are asking for too much or too little.

1.1: The leads and experts have sufficient time for the WCDA and SCDA efforts. Too much of the R&R effort is TBD. No EMC staff are identified by name to help with transition to operations.

1.2: A table provided by the program managers provided the PSD per-person effort and no one is more than 20% for the stochastic physics work there. Prof. Stan is mentioned in the table text and it is assumed Stan's effort is the .1 in 2.1 for GMU. The leads are both at EMC but we don't know how much time they are committing. It is not clearly explained what all the TBD effort will do but the 3 new hires at EMC are for sea-ice coupling, coupling strategy and bias reduction.

1.3: The experts have sufficient time. The EMC lead is only 0.05 FTE.

1.4: Most of the experts have sufficient time. 3 senior people are on for a token amount. Alan Wallcraft and a postdoc will do most of the ocean work.

1.5: The 4 core at JCDSA are full time on this sub-project with one more at 50%. The co-lead at EMC (Fox) has an appropriate 0.25 while other in-kinds are at 0.2. Whitaker has a token amount on this sub-project but overall has significant effort.

1.6: The leads each have about 0.2 FTE. Olson is the only named staff with sufficient time.

2.1: Lots of people will be on this full time but none of them are named. Leads are each 0.1 FTE

2.2: The leads only have 0.1 FTE each. The un-named staff will mostly be working 100%.

2.3: The lead has .33 FTE. The un-named staff may have sufficient time but it is hard to tell. If "Research Associate III" is listed under 2 different milestones, is that the same person or 2 different people?

2.4: Very hard to tell. Sub-project leads not listed in table. 2 people listed don't have an effort level.

3.1: Effort is not specified per named staff member but if it is evenly distributed, then it's an appropriate level.

3.2 Did not see any effort listed for the 2 sub-project leads. The MDL CAMPS staff may be stretched thin on this sub-project. Other named staff are grouped with "other" in the effort count. The total effort should be sufficient if staffed appropriately.

10. Does the project team have adequate and appropriate expertise to achieve the proposed goals?

The leads, sub-project leads and other named staff are all highly qualified to lead the proposed work. But most of the proposed effort, including where most of the real work is done, is from un-named individuals so a confident answer is not possible.

1.1: Good expertise in CDA and in DA for the specific components (land, ocean, sea ice). Most of the work for the R&R is TBD. Amount of operations expertise is not clear.

1.2: Good expertise in stochastic physics, land, atmosphere and ocean modeling and in operations. 80% of the effort is un-named so cannot tell if these is enough expertise.

1.3: This sub-project is commended for having the lowest amount of TBD effort. The team is well qualified and several have substantial effort.

1.4: A well-staffed subproject with many ocean, land and atmosphere experts that would benefit the overall project.

1.5: Has several experts in JEDI software and DA.

1.6: Experts in microphysics and gravity waves are present.

2.1: The 2 leads have the appropriate experience to lead this sub-project. Nearly all the effort is from unnamed "current developers" so cannot judge expertise.

2.2: The name investigators on this sub-project all have adequate and appropriate expertise. But it is hard to just the sub-project as a whole since 86% of the effort is un-named.

2.3: The sub-project lead is highly qualified. Other staff are not named.

2.4: The sub-project leads have experience in the current hurricane forecast system. Most of the effort is not named.

3.1: The leads and all named staff are experts in the software packages being developed. Most of the effort is not named.

3.2: The named staff with provided CVs all have adequate expertise. Most of the effort is not named.

11. Please indicate which areas of the proposed work are most meritorious and crucial for improving NOAA NWS forecast systems. Also indicate those areas that could be descoped under reduced budget conditions.

Work chosen to be descoped typically was too research oriented. If operations indicates they really want it, could be brought back up. Any of the "crucial and meritorious" work that cannot identify their substantial to-be-hired or TBD effort could be justifiably descoped down to existing staff.

12. Please provide a brief overall assessment of the proposed project and an overall score for the proposed project, where 1 is poor and 10 is outstanding. You may also provide ranking for sub-areas or aspects if you like.

Score: ##. Overall, this is an ambitious and important proposal. The deduction is for proposal organization problems which made review difficult: in-consistent content in effort tables, significant numbers of un-named staff and a lack of clear linkage between the sub-projects.

UFS R2O Project Review Criteria

Review Panel Meeting March 12-13, 2020

Reviewer 4

1. Does the proposed Project include a set of significant and achievable forecast goals for 2 or more NOAA UFS-based numerical weather forecast application systems, with at least half expected to be ready and available for NWS transition to operations by the end of the project period?

Based on the Operational Models Consolidation Timeline (Slide 13 from the UFS Project Overview), MER/S2S operational transitions: RTOFSv3(Q1FY23) and GODASv3 (Q1FY23) will occur soon after the end of the two-year project, while GFSv17 and GEFSv13 will become operational in FY24. CAM operational transitions: HAFSv1 (Q3FY22) and RRFSv1 (Q2FY23) are scheduled to occur either during or just after the end of the two-year period, while 3D RTMA/URMAv1 will be operational in Q3FY23.

The proposed MER/S2S project consists of a broad range of activities that were categorized under six sub-projects. The key activity that stands out as supportive of the near-term MER/S2S operational transition timetable is the weakly coupled UFS/JEDI data assimilation (DA) scheme using MOM6/CICE6/FV3/WWIII components based on efforts in Subsections: 1.5 (Data Assimilation), 1.2 (Coupled Model Development) and 1.1 (Coupled DA and R&R): Milestone 1. It was explained that the land model and WWIII are likely to represent serious challenges, along with the observational quality control/handling. Other subprojects mainly support transitions in FY24 and onwards. Subproject 1.1 (Coupled DA and R&R) contains two other tasks: a multi-decadal R&R using the WCDA and the development of a strongly coupled DA (SCDA) system. The WCDA-based R&R effort could overwhelm the WCDA development/testing timeline, reducing the quality of the WCDA system. The SCDA development effort will be demanding and it is questionable whether the CDA would be transitioned to operations in years 3-5 as is proposed on Page 16 or whether a successful R&R could be carried out using the CDA in these same years.

Most of the regional UFS CAM applications: HAFS, RTMA/URM-3D, and RRFS have reasonable pathways forward to operational transition, although only HAFSv1 will be ready in late FY22. However, the WoFS part of the CAM proposal states that operational transition cannot occur before FY24-25 and that the first two years of the project will be spent transitioning from WRF to FV3 and transitioning to JEDI. The proposers of this subproject are very clear about the challenges and are not over-promising.

The CCI Modeling Infrastructure sub-project has a vast but vital set of sub-goals to produce the infra-structure for the global and regional UFS applications. In the two-year time-frame of the project, they propose releases of ESMF/NUOPC in Q2FY21 and Q3FY22, CCPP and SCM releases in Q3FY21 and Q3FY22, and JEDI/NUOPC three-dimensional data exchanges and an updated pre-processing tool chain for MOM6, CICE5 and WWIII in Q3FY22. A clearer prioritization of the subtasks in the first two years would have been valuable with so many different activities. The Verification and Post-Processing subtask is focused on the addition of diagnostics and verification algorithms to

the METplus software along with new releases of CAMPS and UEPP in the project timeframe.

2. Please comment on the overall scientific and technical merit of the proposed project.

Scientifically, the NWS needs to consider processes on time scales from sub-hourly through seasonal both regionally and globally. The proposal did separate out the CAM and MER/S2S timescales and proposed modeling approaches to adequately represent these processes. Some sub-tasks did a better job than others articulating the scientific goals and how they would be addressed. However, this proposal is motivated by an improved operational capability not the exploration of new dynamics.

The technical focus of the proposal is to reduce the number of regional and global applications currently being used operationally into a framework where the same model components are used throughout applications. The models used are FV3, MOM6, CICE6 and WWIII, while the DA systems are under the JEDI umbrella. The model components are consistent with those used in the broader community and the development approach of the JEDI framework should hasten progress. The success of the DA for the ocean model will depend on the availability and ingestion of ocean observations. Coordination between ODA and CCI/V&PP is needed to ensure that certain observations are "with-held" for subsequent veracity tests. The regional CAM applications are stand-alone; the lack of an active ocean, especially in the hurricane applications, is a concern. The implementation of open boundary conditions in MOM6 early in the project by Rutgers along with the MER/S2S coupling component work should allow for its use in some of the regional CAM applications.

3. Does the proposed project include an effective strategy to bring research-to-operations and operations-to-research (R2O2R)? Does it include engagement with the research community to achieve forecast improvements and engagement with the forecast community to establish forecast goals?

The proposed project is strongly motivated by operational transitions and generally includes well-thought through strategies to forward their sub-projects towards transition (R2O). However, the links between sub-projects could be further explored for missing opportunities such as using MOM6 in the CAM applications. Other linkages such as enhancing the relationship between the CCI Verification and Post-processing (CCI/V&PP) sub-project and the ocean data assimilation effort should be encouraged.

The O2R strategy, on the other hand, was more *ad-hoc*. It was stated that model output would be made available to the community for analysis, especially for graduate students, and that there would be "data" repositories. However, no specifics regarding the release

of model output were provided (or maybe I missed it in the CCI section). The CCI/V&PP sub-project plans to enhance METplus for model evaluation/diagnosis. The MER/S2S Coupled Model Development sub-project included the sentence “Engage the user community for evaluation”. The MER/S2S sub-project Atmospheric Composition plan to participate in an aerosol impact experiment (WGNE).

The proposed project engages with the research community. Examples include the university NOFO sub-projects in support of the MER/S2S sub-projects on model components and coupled model development and the CCI project led by Dunlap (NCAR). A major goal of the CCI/V&PP sub-project is a model evaluation system to be used by the UFS community; a meeting will be organized to decide on metrics.

4. Does the proposed project adequately foster code unification, contributing towards the reduction of the number of code versions in use by the UFS and NCEP?

Yes. Both the MER/S2S and CAM applications propose a time-line to reduce the number of applications by using the UFS and retiring many current applications. Most of the MER/S2S current applications will be replaced by GFSv17 and GEFSv13, while the CAM applications are subsumed into RRFSv1/v2.

5. Does the Project have a clear management and organization plan?

The proposed project has a clear management plan. The scope of the proposed project is very broad and has many dependencies. Project engineers will monitor the project and its dependencies, while senior leadership and the program office will adjust milestones as needed.

It is proposed to use confluence and JIRA (both web tools), conference calls and in-person meetings to enhance communication. Quarterly reports will be provided to the program office. I did not see a clear plan for all-hands meetings.

6. Is the proposed project appropriately aligned with other modeling efforts, including the broader UFS community, other programs/projects at NOAA, and with the broader national/international NWP and modeling community?

In terms of model component choices, the use of MOM6, developed at GFDL, is aligned with NCAR’s adoption of this model for use in future coupled models. NRL is also considering using it but a decision has not been reached. CICE6 is broadly used, both nationally and internationally. The ocean data assimilation sub-project plans to use variational and ensemble Kalman Filter DA schemes. GFDL is using an ensemble Kalman Filter (KF) scheme for ODA with MOM6, so the proposed ODA is unsurprisingly aligned

with GFDL. However, the FSU NOFO plans to largely use the Navy Coupled Ocean DA scheme (based on multi-variate optimal interpolation - MVOI) with global 1/12° MOM6. This ODA activity seems divergent with the ensemble KF approach currently being tested for the GODAS/SOCA effort.

The ocean data assimilation effort and the CCI/V&PP sub-project may consider reaching out to the NOAA/OAR- sponsored ocean observation networks when they are considering ocean data streams for the ODA and model evaluation. These networks include NOAA laboratories (PMEL) and university partners. Alignment with these ocean experts could enhance the ocean simulation and analysis part of the proposed project.

7. Has the project established appropriate success metrics?

Overall, they are not well-articulated but tend to be comparisons with existing applications. For example, the performance of 3DRTMA and RRFS will be measured relative to the performances of their predecessors: 2DRTMA and HREFV3, respectively. Sub-project 1.1 plans to use observation-minus-forecasts of the WCDA system versus the uncoupled UFS forecast. Sub-project 1.5 regards the completion of a multi-decadal ocean/sea-ice reanalysis as a benchmark of success. CCI/V&PP will base its success on its user base.

Regards the success of the ocean and sea-ice data assimilation: “with-held” ocean observations are needed to gauge success. As mentioned above, coordination with CCI/V&PP is needed to ensure that certain observations are “with-held” from the ODA or sea-ice DA for subsequent veracity tests.

8. Does the proposed project effectively leverage internal resources, including e.g. time commitment of federally funded NOAA investigators, computing resources, in-kind contributions, and other funded efforts?

All of the MER/S2S sub-projects leverage internal time commitments. For example: Sub-project 1.1 plans for 9.35 FTEs and 1.35 of those are leveraged. Sub-project 1.2 asks for 11.3 FTEs and 6.8 are leveraged. Sub-project 1.3 asks for 4.26 FTEs and 2.2 are leveraged. Sub-project 1.4 lists in-kind support but is not specific about leveraged FTEs except for 1 FTE at EMC. Sub-projects 1.5 and 1.6 have modest amounts of in-kind salary support. Sub-projects 1-5 also leverage office space, some travel and some computing facilities including existing high-performance computing allocations. All CAM sub-projects have in-kind salary leveraging, office space, and list leveraged computing frequently. CCI Modeling and Infra-structure: 12.3 FTEs are requested and 33 FTEs are leveraged; they also leverage office space and computing resources. V&PP: 15.8 FTEs and 8.8 FTEs leveraged, plus office space and computing infrastructure.

9. Does the proposed project have adequate and appropriate resources, including sufficient time-commitment of team members?

Apart from Milestone 1, Sub-project 1.5 is largely relying on new EMC hires or unnamed current EMC staff. For example: Milestone 4 is to be led by Guillaume Vernieres who has 0.1 in-kind FTE; his team consists of 4 FTEs of unnamed current staff. Other milestones will be staffed by new hires with very modest in-kind FTE efforts by Daryl Kleist, Andy Fox, Jeff Whitaker, and Rahual Mahajan. I have singled this out as it is a key activity to UFS success. Other sub-projects had better identified individuals working on the projects.

The budget for the V&PP seems very high; it includes 8.25 FTEs for EMC evaluations and metrics.

10. Does the project team have adequate and appropriate expertise to achieve the proposed goals?

The NOAA proposal team and their university partners are competent to achieve the proposed goals. Is a physical oceanographer involved in the ocean metric design as part of the CCI/V&PP sub-project? The CVs of the project participants did not indicate an ocean expert. Maybe they will rely on the metrics workshop to provide ocean and sea-ice guidance?

11. Please indicate which areas of the proposed work are most meritorious and crucial for improving NOAA NWS forecast systems. Also indicate those areas that could be descoped under reduced budget conditions.

Redacted

12. Please provide a brief overall assessment of the proposed project and an overall score for the proposed project, where 1 is poor and 10 is outstanding. You may also provide ranking for sub-areas or aspects if you like.

My overall ranking of the proposal is xx out of 10. The proposal is impressive in terms of scope and its plans to advance forecasting on time scales from sub-hourly to seasonal. However, the plans could benefit from another round of internal iteration to find linkages and to emphasis the most important aspects of the forecast systems. The lack of an ocean model in the hurricane forecasting system is a concern, especially since an ocean model was used in an earlier system. The state of readiness of experienced staff to work on the data assimilation sub-project is another major concern.

UFS R2O Project Review Criteria

Review Panel Meeting March 12-13, 2020

Reviewer 5

Please provide responses to the following questions, based on the UFS R2O Project written proposal; these responses may be updated based on the panel review materials and discussions.

1. Does the proposed Project include a set of significant and achievable forecast goals for 2 or more NOAA UFS-based numerical weather forecast application systems, with at least half expected to be ready and available for NWS transition to operations by the end of the project period?

   There are three main objectives to this proposal: The development of common model infrastructure as well as post processing software; the development of a UFS-based MER/S2S forecasting system, and the development of a UFS-based Convection Allowing Mesoscale (CAM) Model. At a high level, it seems that the probability to make significant progress in all three areas seems high. Of these, the CAM has a high probability to be ready for testing by the end of the two years, while the MER/S2S will require more time. The infrastructure (JEDI and common infrastructure) must accompany these developments and is also likely to be ready in two years given that it has multiple funding streams that are pointing towards this objective already. Some of the smaller sub-projects may also be done within two years, but as detailed later, the three items above are really the core of this project and what needs to be prioritized if funding is not sufficient to cover all aspects of this project.

2. Please comment on the overall scientific and technical merit of the proposed project.

   This is a very large and complex proposal and it would be easy to go through it to find and challenge some of the individual assumptions. Instead, I think it is important to note that this proposal is unique in that it represents a truly innovative community-based framework for transitioning Research to Operations. This framework can act as a blueprint for how NOAA moves forward not only in R2O but eventually O2R. That said, and because this project is really focused on demonstrating a community-based approach for R2O, it may not be the most novel scientific proposal (except for a few efforts to improve the model itself), but rather should be viewed on its technical strength of integrating a community-based framework for improvement that lays the foundation for future development. That is worth doing.

3. Does the proposed project include an effective strategy to bring research-to-operations and operations-to-research (R2O2R)? Does it include engagement with the research community to achieve forecast improvements and engagement with the forecast community to establish forecast goals?

   The proposal does very well in starting from community and forecaster-based needs that are translated into a collaboration between academia, OAR and EMC to eliminate the valley of death. This project starts with a clear path forward by involving people form all the organizations that constitute the normal chain of maturation. Whether they succeed will of course depend on the quality of the forecasts that the new FV3 core and data assimilation can provide but the path is correct. Thus, there is a clear path from Operations to Research back to Operations (O2R2O) laid out in this proposal. It is premature in my view to cycle back to Research in this project as the first two years, and possibly 4, will be needed just to get a community-based model into Operations.

4. Does the proposed project adequately foster code unification, contributing towards the reduction of the number of code versions in use by the UFS and NCEP?

   Code unification is a very strong part of this proposal. On the High Resolution side, the RRFS is intended to replace NAMnest, HRRR, HiResWindows and HREF in 2023 on its way to more broadly replace the NAM, RAP, HRRR and HREF. This is a significant consolidation of models if the users see the new systems as at least equivalent. It would have been nice to see early plans for the HWRF to also become part of this package but my understanding is that the Weather Forecast Act, and funding through the Hurricane Supplemental would make this exceedingly complex. Nonetheless, it might be something that can be encouraged now so that future transitions are enabled rather than made more difficult through parallel development. The Project Leads stated during the review that this would happen naturally through JEDI etc., but planned and continued communication between HWRF and RRFS might be desirable.

5. Does the Project have a clear management and organization plan?

   I found this part of the project lacking. It seemed to be structured like most academic science proposals, with only nominal time for the PIs and most of the funding going to people doing the actual work. While this works well for small grants, I believe this project needs more help with coordination and resource allocation. The project will need significant help coordinating tasks and ensuring that groups get software they need in time to work efficiently. Computing resource allocations also need to be coordinated. I therefore strongly suggest that NOAA find a way to allocate at least one, if not two project engineers to the

PIs to act as a resource to them that is not currently budgeted. Funding some additional time for the University PI might also help.

6. Is the proposed project appropriately aligned with other modeling efforts, including the broader UFS community, other programs/projects at NOAA, and with the broader national/international NWP and modeling community?

   I found the project to be specifically guided by the UFS community questions and aligns really well with NOAA's desire to have a community based framework to improve forecast models. The alignment with EPIC might be made stronger once EPIC is stood up and begins to develop capabilities.

7. Has the project established appropriate success metrics?

   While each sub-project certainly attempted to define its own metrics, and those should be preserved, the project lacked a set of "overall" metrics that would define success of the project as opposed to the success of the sub-projects. This becomes particularly important given the visibility of this project as a pathfinder for the community-based model improvement. As such, I recommend that the Project elevate the priorities they identified during the review process (abridged slightly) as the Project Success Metrics that they should report at the highest level.

   - Fully coupled (from day 0) ensemble prediction system, including coupled DA, ready for pre-operational testing
     - Addressing science priorities and leading to improvements in forecast priority areas
     - Including reanalysis/reforecast capability for calibration/bias correction (production in year 3)
     - Public release of coupled MER/S2S application
     - Public release of JEDI (including observational data store)
   - Regional rapid refresh (1-hour cadence) ensemble forecast system for CAM scales implemented
     - Public release of regional RRFS system
   - Start sunset of existing mesoscale prediction systems

8. Does the proposed project effectively leverage internal resources, including e.g. time commitment of federally funded NOAA investigators, computing resources, in-kind contributions, and other funded efforts?

I found that the proposal struck a reasonable balance between EMC, OAR Labs, and academia. That said, it struck me that EMC will need 60 new employees (listed as TBD) in the proposal, and in some of the material distributed after the review. I wondered if they were already at EMC but the budget request (8.75M/year) is consistent with hiring 60 new people. This represents a significant risk if that many new people are needed. In fact, I would suggest that this may be the real bottleneck in the project that should be watched very carefully – especially if the project is only for a few years.

9. Does the proposed project have adequate and appropriate resources, including sufficient time-commitment of team members?

   I found that team members seemed to be spending enough time. My concern, repeated from above, is for the leadership that needs to coordinate the larger picture as well as their own teams. The project is understaffed at the PI and at the Project Lead level. The Sub-project is probably staffed correctly as the people in those positions are naturally managing those activities anyway.

10. Does the project team have adequate and appropriate expertise to achieve the proposed goals?

    The team that put this proposal together is as good a team as one can hope for. The concern, and this goes back to item 8 above, is the large number of TBD's, particularly at EMC.

11. Please indicate which areas of the proposed work are most meritorious and crucial for improving NOAA NWS forecast systems. Also indicate those areas that could be descoped under reduced budget conditions.

    Redacted

12. Please provide a brief overall assessment of the proposed project and an overall score for the proposed project, where 1 is poor and 10 is outstanding. You may also provide ranking for sub-areas or aspects if you like.

    I thought the proposal is really important for the community as it lays the foundation for how forecast models may be developed in a more open and agile manner going forward. That alone, merits a xx or a xx. Of the subprojects, I listed the ones I thought were crucial for the success of this project above. From only a sub-project perspective, I'd like to give kudos to the air quality sub-project. While I don't think it is the most important, I thought it was the best written.

The 3D-RTMA and RRFS were also very well written, while the Coupled Model sub-project was perhaps the least convincing if not for the importance that the coupled models take in the overall project.

UFS R2O Project Review Criteria

Review Panel Meeting March 12-13, 2020

Reviewer 6

1. Does the proposed Project include a set of significant and achievable forecast goals for 2 or more NOAA UFS-based numerical weather forecast application systems, with at least half expected to be ready and available for NWS transition to operations by the end of the project period?

The proposed project includes a set of significant deliverables, but not necessarily significant forecast goals. This seems to be by design: the team mentioned that they have deliberately avoided assigning metrics to assess forecast skills improvement goals. That being said, there are a number of deliverables listed in the proposal that are significant and with at least half of those promised for transition to operations (by the project's end): the coupled data assimilation system, the 3DRTMA, the coupled forecast model, are just a few examples. However, it is not evident that the amount of work promised to transition to operations is realistic for the 2 years time frame for this project.

2. Please comment on the overall scientific and technical merit of the proposed project.

The proposed project is a scientifically high-quality effort. All the team members are highly qualified as well. That being said, I see a recoverable flaw in the project that could also be considered an opportunity to streamline and realize cost savings. The project, with the exception of the air composition sub-group, seems to be essentially ignoring a large body of work related to observations (in particular satellite data) exploitation (quality control, bias monitoring and correction, product generation, etc). And we know that satellite data/observations is one of the most important factors in NWP and with the biggest potential to enhance forecast skills if done properly. This is particularly true for the sub project MER/S2S data assimilation that presumably will rely a lot on global DA and on satellite data. Assimilation of Aeolus wind profiles from space is an example (now being assimilated by most major NWP centers). But there are many other examples, including assimilation from non-traditional sources such as Loon stratospheric balloons, etc). Sometimes it is simply a matter of involving the right teams (already funded) and bringing them into the fold (to address certain aspects). This is envisioned to reduce the cost requirements for the data assimilation sub-group, to free up funds (either to be saved, or to be used for other sub-groups that might be under-funded).

Another potential risk to the project is the apparent disconnect of this team from the players who are actively working on similar issues and who may be non-traditional

partners (in academia, private sector and Govt agencies including NOAA). Some disruptive technologies (such as AI and ML) have matured recently and are advancing rapidly (beyond post-processing that the project has alluded to as planning to use ML), which presents a real opportunity to improve efficiency and realize cost savings.Perhaps this is not the role of the project and should be dealt with under EPIC, but if the UFS is locked in for the next couple years, with a rigid design, it will be hard to adapt it in future versions, for new technologies.

3. Does the proposed project include an effective strategy to bring research-to-operations and operations-to-research (R2O2R)? Does it include engagement with the research community to achieve forecast improvements and engagement with the forecast community to establish forecast goals?

   The project is by design well suited for R2O2R. The overall project as well as the sub projects are all led by operational and research representatives (both NOAA and academia), ensuring that the operational objective is the major driver and ensuring that the research community is using the operational versions of the model. The project has well defined timelines for the execution of the effort. These timelines account for operational constraints including freeze periods.

   Not all private sector partners are driven exclusively by profit. Some actually are looking to engage with NOAA and willing to help us improve our prediction skills (they ultimately benefit indirectly from this).

4. Does the proposed project adequately foster code unification, contributing towards the reduction of the number of code versions in use by the UFS and NCEP?

   Yes. That is one of the major objectives of this project (code unification). Especially for the regional modeling aspect. Although some of these target goals are scheduled to be achieved after the period of the project ends.

5. Does the Project have a clear management and organization plan?

   Yes. The project has a clear management and organization plan.

6. Is the proposed project appropriately aligned with other modeling efforts, including the broader UFS community, other programs/projects at NOAA, and with the broader national/international NWP and modeling community?

The project is quite well aligned within the UFS community. But the NWP community is wider. The satellite observations community is a glaring gap in this project which is unfortunate because there is a lot of data assimilation / optimization / simulation being performed in this community and claimed results that seem to improve the forecast skills. Freeing up energy and resources for UFS implementation.

7. Has the project established appropriate success metrics?
   The project has listed in its individual sub-projects, a list of deliverables and metrics for success. What is perhaps lacking is an overall UFS set of high-level deliverables (a coupled modeling system, a coupled DA system, etc.) and list of specific metrics in terms of actual prediction skills enhancement (improvement of Anomaly correlation, QPF, etc. by a certain amount or percentage), similar to what HFIP program has.

8. Does the proposed project effectively leverage internal resources, including e.g. time commitment of federally funded NOAA investigators, computing resources, in-kind contributions, and other funded efforts?

   See my comments above in section 1.

9. Does the proposed project have adequate and appropriate resources, including sufficient time-commitment of team members?

   Yes, the proposed project has adequate and appropriate resources planned and dedicated to the effort being proposed. Some commitments are at a 0.05 level (2 hours a week), which represents a slight risk to the project. Comment to the program managers: 14 FTE approximately seem to be requested to fund the modeling infrastructure. Wouldn't this aspect at least be partially outsourced to the EPIC baseline activities? (and not be part of the UFS).

10. Does the project team have adequate and appropriate expertise to achieve the proposed goals?

    Yes. The team has definitive and adequate expertise to achieve the proposed goals.

11. Please indicate which areas of the proposed work are most meritorious and crucial for improving NOAA NWS forecast systems. Also indicate those areas that could be descoped under reduced budget conditions.

It is my opinion that the most crucial components of the UFS that are necessary to improve the NOAA NWS forecast systems in the future, are (1) coupled modeling systems, (2) coupled data assimilation systems, (3) improved and flexible atmospheric physics package, (4) User friendly visualization and post-processing package, (5) the consolidation of the regional forecasting into the RRFS and an overall (6) O2R infrastructure for an efficient engagement with the wider community.

Given its focus, it is assumed that the *modeling infrastructure* will be covered under the EPIC base program (at least partially).

In case of reduced budget conditions, the following recommendations could be considered:

Redacted

12. Please provide a brief overall assessment of the proposed project and an overall score for the proposed project, where 1 is poor and 10 is outstanding. You may also provide ranking for sub-areas or aspects if you like.

The UFS project is highly needed for NOAA to regain its leadership role in the NWP community and to remain up to the state of the art standards. The engagement with the community, defining clear and transparent protocols for O2R and R2O, supporting the wider community to help NOAA, are all extremely important goals. The approach of consolidating projects at this stage between these groups (mainly OAR and NWS), is a cornerstone of the project, and should be highly commended, and should provide added value in the near future.

Despite this OAR-NWS cohesiveness, the project would have been significantly better received if it was more inclusive of other key NOAA partners (within NOAA) engaged in observations exploitation (for NWP, for improving forecast skills).

The project seems to perhaps try to achieve both infrastructure setup (for the community) and at the same time, improvement in forecast skills. It is likely not realistic to do both in the time frame allotted to the project. The project should engage stakeholders and decide on what is the most important to achieve and then stagger the project over longer periods of time.

In addition, the project runs the risk of putting ‘all eggs in one basket’ for OAR and NWS and be locked down for the next 4-5 years, if the UFS is not open to potential changes in directions in the field of NWP or ESP in general (Earth System Prediction). It is recommended that the project leaves the door open to changes as it moves forward. Changes that could be due to new techniques, emergence of new players in the community. This should be reflected in the management structure and Terms of reference.

<u>The overall project is rated as xx.</u>

The individual projects are rated as follows: Redacted

UFS R2O Project Review Criteria

Review Panel Meeting March 12-13, 2020

Reviewer 7

I will note that I have not separated the sub-projects and combined general impressions of the proposal as a whole under each answer. Sub-projects are mentioned where needed and briefly broken down in the section 11 priority summary.

1. Does the proposed Project include a set of significant and achievable forecast goals for 2 or more NOAA UFS-based numerical weather forecast application systems, with at least half expected to be ready and available for NWS transition to operations by the end of the project period?

   The proposal includes 12 sub-projects aiming to replace and consolidate a large number of existing forecast systems over a period that extends beyond the two-year proposal.

   The 6 sub-projects under MER-S2S are aimed at the coupled global model, atmospheric composition, data assimilation and physics. The target date for GFSv17/GEFSv13 is Q2FY24, two years after the second year. Given that this is transitioning from an atmosphere-only GFSv16 to a coupled GFSv17, the time to achieve this is reasonable. While it was not mentioned in the proposal, the presenters did note that initial tests of a fully coupled system have already produced reasonable results compared to CFSv2, especially benefiting from improved sea ice. Without this assurance, the proposal may have seemed to lack sufficient tasks related to coupled testing within the above-mentioned task areas, but it appears this is being done. SFSv1 for seasonal forecasts is targeted for Q4FY24. Upgrading reforecasts and reanalysis is an important dependent product.

The 4 sub-projects in the CAM area are a 3D real-time analysis, the RRFS regional forecast system, Warn on Forecast and Hurricane forecasting. This area has some short-term targets within the 2-year period which are having HAFSv1 ready for operational testing and having a stand-alone-regional (SAR) FV3 member in HREFv3. Longer term targets are the 3DRTMA/URMA replacing the existing RTMA/URMA system, RRFSv1 targeted for Q2FY23 replacing the regional NAM, hi-res windows, RAP/HRRR and HREF systems with an ensemble DA and forecast system at around 3 km grid size nationally. RRFSv2 for Q2FY24 would also subsume air-quality model CMAQ. Furthermore, the regional SREF is to be replaced by GEFSv13 if possible. WoF is a longer-term target beyond FY24.

The 2 sub-projects in the CCI area include infrastructure for the modeling system and its support and verification and post-processing. These are ongoing tasks in the proposal and are critical to the success of the community aspect of the UFS while also improving the interface for outside developers and for running and verification of the UFS system. The MR Weather application has just been released as a first public version. Graduate Student tests have been done to make sure the system is easily usable. These are viewed as a good foundation for these sub-projects.

To answer the question, the inclusion of a SAR member in the HREFv3 3 km ensemble for Q4FY20 is an early operational target. Perhaps for cost reasons it is done as a replacement of a member rather than as an addition which can be viewed as a risk given the short preparation time. This gives a means for evaluating the FV3 core at convection-allowing resolutions early in the project. The HAFSv1 implementation at the end of the proposal period is the next operational target. It is understood that this sub-project is parallel with a larger funded HFIP task, and it requires substantial effort because the hurricane models have a complex system of initialization, ocean coupling, and moving nest capabilities that also would need to be replaced. However, in the presentations, early tests of the FV3-based system with modified global physics have looked promising, so this gives confidence in the model part of this project.

The bulk of the operational readiness is beyond the end of this first phase, but given the wholesale shift in the dynamical core, the physics, ocean model, data assimilation system and infrastructure, this is not surprising. By the end of 2 years both the MER-S2S and CAM sub-areas will be advancing as rapidly as seems possible towards operations and replacement of multiple current systems.

2. Please comment on the overall scientific and technical merit of the proposed project.

The scientific merit of this proposal is overall high. It is understood that the goals are results-oriented in terms of the need to provide the best possible forecasts at all scales. Given the uncertainties in initialization, it is important to have state-of-the-science data assimilation capabilities including ensemble methods that should at least in part come through the international JEDI effort. The GEFSv13 and RRFSv1 ensembles also recognize the value of gauging uncertainties for forecasting. GFSv17/GEFSv13 also will use a weakly coupled data assimilation system and a strongly coupled system also has proposed work. These represent capabilities at the forefront of science and it will be a complex area requiring expertise in atmosphere and ocean modeling as well as data assimilation. Strong coupling has some risks, but there is no dependence on it working soon.

In terms of physics, the presentation also showed room for improvement in the cumulus parameterization with regard to tropical waves, for example. The physics sub-project includes an early task force to identify root causes of errors and priority is rightly given to moist physics. While double-moment microphysics may be a good target that also will help CAM applications, since the results-based priority appears to be for the MR application, the task force should look objectively at whether improving the sub-grid convective scheme should take priority over the microphysics.

The technical merit is also overall high. Aspects included in this are the data assimilation infrastructure, the new common physics infrastructure, the adopted coupling infrastructure, and the general goal of making the system easily usable to the UFS community. The project is fully adopting ideas that make adding and swapping components easier that will help streamline future development. For example, adding physics becomes easier when the interface to the model is well defined and obeys some rules rather than having to know much about the dycore, parallelism, and memory structure. This is definitely a positive direction, and JEDI offers similar benefits.

3. Does the proposed project include an effective strategy to bring research-to-operations and operations-to-research (R2O2R)? Does it include engagement with the research community to achieve forecast improvements and engagement with the forecast community to establish forecast goals?

Streamlining R2O has clearly been the underlying goal of this proposal that promises to have long-term benefits. It has responded to a real need that was identified in recent reviews for the national weather models to open up their scientific development in a way that can leverage the rather strong US academic community and other agencies with similar modeling interests. This would work well especially if

NOAA funds the community with parallel projects to this on particular areas of need that would be identified through forecasting or technical problems in the existing operational systems. Identifying the needs and funding relevant research would form the O2R component.

The proposal itself has a strong research component in the areas of physics, coupling, data assimilation, and extensions to deep atmosphere and atmospheric composition. It includes researchers at NOAA labs, ESRL and NSSL, and at several universities.

Being a little familiar with the UFS effort, I know that community engagement is a serious priority. The release of public versions of operational codes with support and offering tutorials and workshops are important aspects of the broader effort that exist beyond this proposal.

4. Does the proposed project adequately foster code unification, contributing towards the reduction of the number of code versions in use by the UFS and NCEP?

Yes, this is clearly the long-term goal. The first two years take the necessary steps to make it possible to have a global MER-S2S and limited-area CAM capability in the future, both sharing the same dynamics and capable of sharing the same physics. These two systems would replace many separate systems used for reanalysis/reforecasts, S2S prediction, global MR prediction, national mesoscale multi-day forecasts, and CAM rapid-refresh forecasts, including ensembles for global scales and CAM scales.

There are built-in safeguards to continue existing systems if the new system degrades the performance. The main example of this is the SREF that is to be replaced by the GEFS where the risk is that the global physics may not be adequate for the North American region to replace the SREF. The proposal recognizes this and includes specific areas where such deficiencies are already known such as the planetary boundary layer. Another example is in the Warn-on-Forecast system planned for multiple years later where some differences are noted in the simulation of resolved convective updrafts between the ARW and FV3 dynamics. These issues probably need to be resolved before acceptance and progress will likely be looked at annually in the NSSL's HWT program.

5. Does the Project have a clear management and organization plan?

The three topic areas MER-S2S, CAM and CCI have co-leaders and the 12 sub-project areas also have designated leaders. In addition to this internal structure is the more external UFS team structure organized along applications and subject-area working groups. The internal structure is needed to manage timelines and deliverables specific to this proposal. It is clear that some sub-projects such as Physics and Data Assimilation also have cross-cutting aspects and should not be stove-piped within the topic area. I believe this is recognized. How the internal structure works with the external UFS structure remains to be seen, but as long as the proposal managers are represented in all the working groups and application areas, there should be no disconnects.

6. Is the proposed project appropriately aligned with other modeling efforts, including the broader UFS community, other programs/projects at NOAA, and with the broader national/international NWP and modeling community?

   It is not explicit in the proposal, but I am aware of connections with other modeling projects. The UFS working groups have representatives from other national modeling groups, such as NCAR and NRL, possibly NASA and DOE. The UFS work is also regularly presented in the ESPC telecoms that focus on physics interoperability across modeling groups. The CCPP is being adopted and developed across several centers including NCAR and NRL, and this is highly collaborative. Although I am less familiar with it, JEDI is another example of collaboration with other modeling centers on unifying data assimilation software. Coupling is also in collaboration with NCAR's CESM work. The MET verification package has been adopted for broader use in the UFS having initially been developed at NOAA's DTC. NCEP EMC have committed to using outside developments in verification and the physics framework which was an essential step that is already going into operations with GFSv16 and existed through its test phase. Candidate physics for GFSv16 included the RAP suite developed at NOAA's ESRL and other externally provided packages via the CCPP. The CCPP continues to be a focal point for physics collaboration.

7. Has the project established appropriate success metrics?
   While not all the sub-projects outlined metrics for success in the proposal, it is recognized that the main metric for success relates the forecasts with the new system and the current system. The reviewers had some discussion on this because even if the benefits in the forecast systems are small at first, there is great value in adopting a more open and community approach as a goal of the UFS which serves NOAA well for future development and collaboration with external partners. The other clear benefit is streamlining to support fewer models that are more closely related to each other

compared to the current rather diverse set. The increased public availability of operational codes and reduced set of models provide other measures of success.

8. Does the proposed project effectively leverage internal resources, including e.g. time commitment of federally funded NOAA investigators, computing resources, in-kind contributions, and other funded efforts?

   This is hard to judge, but the proposal relies on substantial in-kind contributions in all areas, especially DA, 3DRTMA, HAFS, RRFS and verification. This is likely because it is considered a base-funded activity. Several full FTE in-kind contributions exist in these listed areas. With large leveraged reliance, project management may have to compete with other time needs unless project timelines and priorities are explicitly shared with those FTEs not directly funded.

   In the discussion it was noted that the computing estimates also challenge the computing resources. Since a new computer is coming, some may be available in pre-testing there.

   For the planned operating system, the extra computing resources required by this system are accounted for by putting reforecasting and reanalysis and the future WoF system on cloud computers.

9. Does the proposed project have adequate and appropriate resources, including sufficient time-commitment of team members?

   Including the substantial leveraged in-kind FTEs, I believe the resources have been adequately estimated, and key personnel have sufficiently funded and in-kind involvement.

10. Does the project team have adequate and appropriate expertise to achieve the proposed goals?

    The project team has experience with physics and data assimilation for most of the required new forecasting systems. An area that has not had much recent development is coupled modeling so this presents new challenges and needs specific expertise. The coupled model evaluation is being done in collaboration with GMU. This is an area that may show up biases in the coupled system that do not appear in an atmosphere-only system. The expertise has to include areas such as radiation/cloud interaction and surface stress in the atmosphere and ocean.

11. Please indicate which areas of the proposed work are most meritorious and crucial for improving NOAA NWS forecast systems. Also indicate those areas that could be descoped under reduced budget conditions.

    The top priorities are those leading to GFSv17/GEFS and RRFSv1 that are the core of the new system in my opinion.

    Redacted

12. Please provide a brief overall assessment of the proposed project and an overall score for the proposed project, where 1 is poor and 10 is outstanding. You may also provide ranking for sub-areas or aspects if you like.

    Overall assessment is xx/10. The scope includes everything that is needed for a next generation modeling system that replaces outdated multiple unrelated models, and streamlines support, research and development while making the new system more open to the research community enabling both R2O and O2R. In section 11, I have outlined which sub-projects are closer to the core needs and rated them accordingly.

UFS R2O Project Review Criteria

Review Panel Meeting March 12-13, 2020

Reviewer 8

1. Does the proposed Project include a set of significant and achievable forecast goals for 2 or more NOAA UFS-based numerical weather forecast application systems, with at least half expected to be ready and available for NWS transition to operations by the end of the project period?

The proposed Project includes several significant and achievable goals for several NOAA UFS-based numerical weather forecast applications systems, including GFSv17&GEFSv13, RRFS, HWRF & HMON and WoF.

For GFSv17 and GEFSv13, several important goals are likely to be achieved by the end of the project period. They include:

- Use of JEDI for atmospheric data assimilation
- Stochastic physics for atmosphere/land/ocean
- Weakly coupled-data assimilation re-analysis system
- Improved physics (deep atmosphere dynamics, upgraded microphysics …)

Several other goals should be ready and available for NWS transition to operations by the end of the project period:

- Maintenance and upgrade of HWRF & HMON
- Transition of the FV3 SAR in the WoF system
- Replacement of 2D-RTMA by 3D-RTMA
- Development of RRFS and retirement of several mesoscale modeling systems

2. Please comment on the overall scientific and technical merit of the proposed project.

This is a very ambitious proposal. It is organized around 12 sub-projects, each of them proposing significant modeling improvement to the Unified Forecast System. Overall, this is an excellent proposal which should address the ultimate goal of improving the skill of the NOAA's operational forecast models. This proposal would also help unify the various forecasting systems at NOAA as well as broaden the access and usage of UFS by the research community.

This project has high scientific merit. The proposed research should lead to significant improvements to UFS and address the high priorities listed in the project Executive Summary. In particular, some proposed research activities such as coupled data assimilation, new microphysics, ocean-atmosphere coupling, stochastic physics changes to the gravity wave drag and interactive aerosols should lead to improved performance of the UFS. I was less convinced by the potential benefit of other research topics such as deep atmosphere dynamics and improved conservation properties.

This project covers a large number of research topics. However, some important topics do not seem to have sufficient attention: initial uncertainty while model uncertainty (e.g. stochastic physics) is well covered, stratospheric chemistry (particularly ozone) which can have an important impact on the lower stratosphere biases, sea-ice assimilation, scalability of the different model components (although this is more technical, optimizing the cost of the UFS is of crucial importance for reducing the cost of testing). In addition, the project focuses mostly on increased complexity of the UFS. Therefore, the proposed innovations may not all be implemented in the next version of GEFS and GEFS.

The technical merit of this project is also very high. This project will lead to the implementation of JEDI and FV3 in several key forecasting systems. This will represent a very important step forward for the UFS. The use of the cloud will also be a very important step to engage the outside research community. This could represent a game changer for model development in NOAA. For me, these technical aspects are the strongest part of this project and, alone, justify its endorsement.

I think it was a very good idea to include a third area “Cross-Cutting activities”. This area covers the development of common diagnostics (MetPlus). It could be extended to define a strategy for model testing. Establishing the cost/benefit of each proposed model change will be crucial for this project. A common strategy for the evaluation of model improvements will be needed.

3. Does the proposed project include an effective strategy to bring research-to-operations and operations-to-research (R2O2R)? Does it include engagement with the research community to achieve forecast improvements and engagement with the forecast community to establish forecast goals?

In my opinion, one of the strongest aspects of this project is the development of a clear strategy to bring research to operation. As mentioned above, it will create a paradigm where the research community can be engaged more easily to achieve forecast improvements thanks to the unification of codes and use of cloud. This will make it easier for research scientists outside of NOAA to develop and test new changes to the UFS. This will accelerate the implementation of innovations into UFS. This new paradigm should also make it easier for research scientists to test and verify operational configurations which will provide them inputs for possible future improvements. Therefore, this strategy should be effective in bringing research to operations as well as operations to research. Another positive aspect of this project is the fact that it involves people from many different communities.

4. Does the proposed project adequately foster code unification, contributing towards the reduction of the number of code versions in use by the UFS and NCEP?

If successful, this project should lead to code unification and the reduction of the number of code versions. This is particularly the case in the CAM sub-projects which will lead to the retirements of several regional mesoscale systems and to the reduction in the number of forecasting systems. The use of JEDI and FV3 in several systems will also represent an important step forward towards code unification.

5. Does the Project have a clear management and organization plan?

The organization of this proposal is complex, with many interdependencies and sometimes overlaps between sub-projects, particularly in MER-S2S. A major challenge for this project will be the management of these interactions. This could be helped by reducing the number of sub-projects in MER-S2S (6 sub-projects are probably too many) by merging some of them (e.g. Coupled DA and RR could merge with Data assimilation. Atmospheric physics could be part of Model components). The number of MER-S2S sub-projects could be reduced to a maximum of 4.

The proposed management structure includes project leaders, team leaders, sub-project leaders as well as project engineers. The role of the project engineers will be crucial since their role will be to facilitate the interaction between the various sub-projects. This structure seems adequate although the 3 layers of management (project, team, sub-project) seems a bit heavy. Maybe having a leaner 2-layer structure (no team leaders) instead would not be more efficient.

6. Is the proposed project appropriately aligned with other modeling efforts, including the broader UFS community, other programs/projects at NOAA, and with the broader national/international NWP and modeling community?

The proposed project is well aligned with other modelling efforts in the US. It includes interactions with other entities part of UFS. The fact that it includes people from all communities will help its alignment with other modelling communities. The alignment with the international community is less obvious, except for a few sub-projects such as Atmospheric composition which plans are aligned with WGNE experimentation

7. Has the project established appropriate success metrics?

Although each sub-project has established its own success metrics, the success metrics for the whole Project need to be better established. They need to include a clear target for model skill improvement as well as the number of model retirements.

8. Does the proposed project effectively leverage internal resources, including e.g. time commitment of federally funded NOAA investigators, computing resources, in-kind contributions, and other funded efforts?

The proposed project leverages internal resources, including in-kind contributions and other funded efforts. The use of cloud computing will help leverage some computing resources. More efforts may be needed to leverage computing resources.

9. Does the proposed project have adequate and appropriate resources, including sufficient time-commitment of team members?

I think that the level of resources is adequate and appropriate, except maybe for computing resources as mentioned earlier. The time-commitment of team members seems adequate.

10. Does the project team have adequate and appropriate expertise to achieve the proposed goals?

All the project teams have adequate and appropriate expertise to achieve the proposed goals.

11. Please indicate which areas of the proposed work are most meritorious and crucial for improving NOAA NWS forecast systems. Also indicate those areas that could be descoped under reduced budget conditions.

All 12 sub-projects have merit and would deserve funding. However, under reduced-budget conditions, priorities have to be established.

Redacted

12. Please provide a brief overall assessment of the proposed project and an overall score for the proposed project, where 1 is poor and 10 is outstanding. You may also provide ranking for sub-areas or aspects if you like.

Overall, this is an excellent project which should address very well the ultimate goal of improving the skill of the NOAA's operational forecast models. This project would also help unify the various forecasting systems at NOAA as well as broaden the access and usage of UFS by the research community. Score: xx

Sub-projects:

UFS R2O Project Review Criteria
Review Panel Meeting March 12-13, 2020

Reviewer 9

Please provide responses to the following questions, based on the UFS R2O Project written proposal; these responses may be updated based on the panel review materials and discussions.

1. Does the proposed Project include a set of significant and achievable forecast goals for 2 or more NOAA UFS-based numerical weather forecast application systems, with at least half expected to be ready and available for NWS transition to operations by the end of the project period?

Yes. The project is very clearly focused on delivery through to operations, and where there is not a plan for operational delivery, this is mostly to bring lower TRL developments up to higher TRLs, ready for a potential future implementation.

2. Please comment on the overall scientific and technical merit of the proposed project.

   Both scientifically and technically, the top-level goals of the UFS R2O project are challenging, but achievable. The developers appear to be asking the right scientific questions, and what became clear during the review is that they are willing to adapt their specific plans in response to the answer to those questions. It was encouraging to hear the term "evidence-based decision making" used during our discussions; I would encourage all sub-project leads to keep this in mind during the project.

   The cross-cutting technical projects will provide the right technical framework with which to develop, deliver and maintain the UFS (modelling infrastructure) and with which to assess and monitor this delivery and inform future development requirements (verification and post-processing).

3. Does the proposed project include an effective strategy to bring research-to-operations and operations-to-research (R2O2R)? Does it include engagement with the research community to achieve forecast improvements and engagement with the forecast community to establish forecast goals?

   Mostly. The R2O approach is very clear, particularly for relatively mature science, as this is obviously the focus of the project. I would advise the developers of the UFS releases to work with the model/system developers and the developers of the VPP sub-project to develop not only standard packages and test cases, but standard test/assessment suites and scorecards that can be used (alongside the researcher's more process-based diagnostics) to assess the impact of their work on the wider behavior of the model.

   It's less clear how the projects will deliver to the O2R process. One suggestion would be that when the O2R process identifies an issue with such a high priority that NOAA would like some significant efforts/development from the wider research community, that you provide test cases and assessment metrics as part of a release that is "tailored" to measure impact on that specific issue.

4. Does the proposed project adequately foster code unification, contributing towards the reduction of the number of code versions in use by the UFS and NCEP?

   Yes. Overall, the project appears to have the right balance between ambition and pragmatism in this goal. The global sub-projects focus on pulling together different components of the modelling suite (e.g. atmosphere, ocean, waves, etc.) whilst keeping options open for limited or one-way coupling. The CAM applications focus much more clearly on the consolidation of modelling systems, which appears to be

a major source of duplicated effort and resource, whilst accepting that the "sunsetting" of some of the older systems may take time and/or rely on specific mitigation efforts.

In some of the material, the pragmatic/mitigating approaches (i.e. plan B) could have been made more explicit, but the proposers were open and honest about these during the presentation and discussions, so there was no sign of them trying to gloss over these.

5. Does the Project have a clear management and organization plan?

   Yes. On paper the project plan looks sound, although there are some details agreed in how the project will be monitored (in both software and practical coordination activities like meetings etc.). The main concerns that were raised by other members of the panel, and which I share, are how the coordination between different aspects of the project can be achieved, particularly in monitoring inter-dependencies etc. Additionally, I was slightly concerned by the messaging that the project leads see this as one of 2-3 full time jobs. I obviously understand the sentiment behind this, and have dealt with similar situations myself, but would like to emphasize the importance of visible and proactive sponsorship for any project in its success.

   Finally, there was some discussion within the panel on a potential re-organization of sub-projects 1.1, 1.2 and 1.5. I don't have strong opinions about this, so long as the sub-projects are coordinated appropriately. This coordination is particularly important, given that contributions are coming from different parts of NOAA.

6. Is the proposed project appropriately aligned with other modeling efforts, including the broader UFS community, other programs/projects at NOAA, and with the broader national/international NWP and modeling community?

   Yes. From what I know of the UFS and the wider NOAA activities, this project is mostly focusing on the things that will be needed to successfully develop and deliver the UFS for operational implementation. In terms of a general direction of travel, the consolidation and unification of software in the NCEP operational suite is an absolute necessity, as otherwise, the US efforts in NWP will fall further behind; this would be unacceptable, given the intellectual and financial resource going into meteorological science, technology and industry in the US. The wider aims of the consolidations, i.e. increased coupling, a consolidated global to regional (and possibly to potentially even more local scale) forecasting suite, are also aligned with the wider NWP community.

   Obviously, the main focus of the UFS and UFS R2O is improvements to the systems and the model development process. I would advise those leading these implementation activities, however, to also consider improvements to their operational implementation processes, including sufficiently early engagement with stakeholders to bring them on board with the developments being made. I

would also advise starting discussions with stakeholders about the possibility of a slightly more agile development process (with smaller, incremental upgrades being made more regularly). This might add to the workload in this implementation activity, but it should speed up the rate of improvement in the systems.

7. Has the project established appropriate success metrics?

   Mostly. The logical comparison for most of these systems will be to use the standard assessment metric/scorecards, because the project is building upon or replacing a good quality baseline system. It is clearly important that on its initial implementation (application-by-application), the UFS is able to match the current levels of performance, and possibly address some of the areas of concern in the current systems. The long-term success of the UFS, however, will be measured not on its initial impact, but on the rate of improvement in operational systems following this. That will be much harder to measure, so again, the project has correctly identified measures like community engagement etc., that it believes will be necessary to deliver that longer-term goal.

8. Does the proposed project effectively leverage internal resources, including e.g. time commitment of federally funded NOAA investigators, computing resources, in-kind contributions, and other funded efforts?

   Yes. This appears to be well spread across the different sub-projects. I assume that some of this is double-counted from/contributing to other work (e.g. 25 FTE EMC engineering team!) but assuming that the most of the other estimates are right, then this seems appropriate.

   There is far less contribution in kind directly noted in sub-project 1.5, and unlike other panel members, I initially found it difficult to disentangle the work in this project from any wider work from NOAA contributing to the development of JEDI. I think the final clarification was that this sub-project was basically the EMC contributions to JEDI development and the subsequent implementation of a JEDI-based DA system. I think this means that we can take as read a wider “contribution in kind” toward the underpinning development and support of the JEDI framework from other JCSDA partners.

9. Does the proposed project have adequate and appropriate resources, including sufficient time-commitment of team members?

   Yes, subject to my comment above about the project sponsors/leads. As I am unaware of the NOAA compute capability and the process for bidding for resources, I will leave other panel members to comment on that.

   Finally, for the project to succeed, I would encourage the project leadership to develop a joint responsibility for success to the different sub-projects, so that if one project is well on track and another important project is not, that the project

manager/sponsor can encourage some “in kind” contribution between the different sub-projects, where expertise allows

10. Does the project team have adequate and appropriate expertise to achieve the proposed goals?

    Yes. It seems fairly clear that the right people are involved in the right sub-projects.

11. Please indicate which areas of the proposed work are most meritorious and crucial for improving NOAA NWS forecast systems. Also indicate those areas that could be descoped under reduced budget conditions.

    Given that you’re trying to work to relatively fixed timescales and you have limited resources, I’ve tried to apply an agile MoSCoW approach to assessing this (i.e. to split my recommendations for the sub-projects into Must have, Should have, Could have and Won’t have this time). Note that there is no ordering of sub-projects within each of these categories; they are in sub-project number order.

    Redacted

12. Please provide a brief overall assessment of the proposed project and an overall score for the proposed project, where 1 is poor and 10 is outstanding. You may also provide ranking for sub-areas or aspects if you like.

    The UFS R2O project is an important part of the wider UFS initiative as it underpins the timely delivery of operational UFS-based systems. The project is well planned and the activities selected for inclusion map on the user requirements and future development priorities. For an overall assessment, I would assess the score for this proposal an xx. It represents the start of an ambitious push by NOAA to better harness the wealth of NWP development taking place in the USA and to apply this more efficiently to the development of operational systems, and hence user applications. Whilst there are some aspects of the project that are less critical than others, there are no activities that I would recommend not pursuing, and I would encourage the Weather Program Office to reconsider the budget available to allow more of the sub-projects to be funded.

    Alternatively, you could delay the start of some of the sub-projects/activities. In that case, you should expect the delivery dates to be later than in the proposal, however, rather than introducing the requirement for even more investment in the later years of the project.

    Final/additional comments.
    I would like to commend the project leads for pulling together the proposals and presentation material at what sounded like very short notice. They engaged with the panel review process openly and contributed to constructive discussions on several topics. The visual representation of timelines and interdependencies in the

CAM presentations, and in some parts of the cross-cutting activities presentations, were particularly appreciated.

UFS R2O Project Review Criteria
Review Panel Meeting March 12-13, 2020

Reviewer 10

Overall Reply (R): This proposal is quite different from regular proposals, as it is a compilation of selected subprojects that have gone through some community efforts (particularly SIPs). The inperson panel review and discussion with the proposal team substantially clarified many of the issues I had.

I appreciate the efforts of NWS/OSTI and OAR/OWAQ in jointly supporting this UFO-R2O project, substantially advancing the collaboration between NWS and OAR, as pursued by the two Assistant Administrators (Louis Uccellini and Craig McLean).

I also appreciate the efforts of the project team for bringing all subprojects together in an overall coherent way in a short time. This project opens a new era in the community effort (NWS, OAR, NCAR/UCAR, and some universities) in developing UFS and accelerating R2O.

1. Does the proposed Project include a set of significant and achievable forecast goals for 2 or more NOAA UFS-based numerical weather forecast application systems, with at least half expected to be ready and available for NWS transition to operations by the end of the project period?

R: The project has a set of significant and achievable forecast goals for each subproject. Overall, the goals of using RRFS to replace several legacy regional models, development of the crosscutting infrastructure, and the delivery of the new coupled modeling system and its coupled DA are excellent.

While deliverables of the two subprojects of CCI are likely to be ready for NWS transition to operations in 2 years, the readiness for subprojects in MER/S2S and CAM (particularly in delivering the coupled modeling and DA systems and RRFS/HAFS) depends on the computing resources available to this project.

If computing resources are limited, the top priority would be the delivery of the coupled modeling and DA systems, as their output drives RRFS and HAFS in forecasting and the model dynamics (FV3), DA, and some of its physics will be shared with regional models.

2. Please comment on the overall scientific and technical merit of the proposed project.

   R: Overall the project has both scientific and technical merits. It opens a new era in the community effort (NWS, OAR, NCAR/UCAR, and some universities) in developing UFS and accelerating R2O. The modeling and DA systems and the infrastructure will also help entrain more involvement of the academic community and possibly the private sector, further accelerating the UFS development in the future.

   The challenge is that the project needs to combine the long-term goal (in delivering the coupled modeling and DA systems and infrastructure) with the short-term goal (in showing some success in improving weather forecasting). The latter is needed to ensure sustainable support of the project after the first two years, and it is achievable as some of the subprojects focus on forecasting bias reduction.

3. Does the proposed project include an effective strategy to bring research-to-operations and operations-to-research (R2O2R)? Does it include engagement with the research community to achieve forecast improvements and engagement with forecast community to establish forecast goals?

   R: R2O strategy is clear (because of substantial EMC involvement and the partnership of OAR, NCAR/UCAR and several universities). O2R (primarily after the first two years) will be enabled by the coupled modeling and DA systems and infrastructure developed under this project.

4. Does the proposed project adequately foster code unification, contributing towards the reduction of the number of code versions in use by the UFS and NCEP?

R: Yes for the regional modeling over U.S. and for hurricane forecasting.

There are four land models (Noah, Noah-MP, RUC, LM4) used at NOAA. The default option for

UFS is Noah-MP. This project should finalize the unification of Noah, Noah-MP, and RUC, with Noah-MP (with some possible revisions during this process) for both global and regional models in UFS.

5. Does the Project have a clear management and organization plan?

   R: Yes, there is a clear management and organization plan, as discussed at the panel review meeting.

   As subprojects are more or less independent projects, it is unclear how much control the project leaders have on individual subprojects (particularly when progress is slower than projected). It is also desirable for project leaders (at least those for overall project and three teams (MER/S2S, CAM, CCI)) to commit at least 2 months/year to the project.

6. Is the proposed project appropriately aligned with other modeling efforts, including the broader UFS community, other programs/projects at NOAA, and with the broader national/international NWP and modeling community?

   R: The proposed project aligns well with other modeling efforts at NOAA. It is also consistent with the overall national and international model and DA development strategy for the future.

   Closer collaboration with NASA/GMAO (e.g., in DA) and OAR/GFDL (e.g., in dynamic core and Earth system modeling) is desirable.

7. Has the project established appropriate success metrics?

   R: Yes, each subproject has established success metrics. The overall success metrics were also discussed at the panel review meeting. Most of the metrics are appropriate.

   The challenge is that the success metrics for the coupled model and DA development are largely just comparison with existing NOAA systems, and the

bar may be regarded too low for this relatively large investment. A more appropriate bar would be comparison with the current ECMWF modeling system and with future ECMWF systems (in terms of reducing the gap between NCEP system versus ECMWF system).

8. Does the proposed project effectively leverage internal resources, including e.g. time commitment of federally funded NOAA investigators, computing resources, in-kind contributions, and other funded efforts?

   R: Yes.

   One challenge is the availability of requested large computing resources that would strongly affect the progress of this project.

9. Does the proposed project have adequate and appropriate resources, including sufficient time commitment of team members?

   R: yes, there are sufficient time-commitment of team members.

   It is desirable for project leaders (at least those for overall project and three teams (MER/S2S, CAM, CCI)) to commit at least 2 months/year to the project.

10. Does the project team have adequate and appropriate expertise to achieve the proposed goals?

    R: Yes.

    It is unclear how these participants were selected.

11. Please indicate which areas of the proposed work are most meritorious and crucial for improving NOAA NWS forecast systems. Also indicate those areas that could be descoped under reduced budget conditions.

R: All subprojects are meritorious.

Identification of subprojects or specific areas for descoping or reduced budget depends on the goals emphasized by the program management.

12. Please provide a brief overall assessment of the proposed project and an overall score for the proposed project, where 1 is poor and 10 is outstanding. You may also provide ranking for subareas or aspects if you like.

R: Overall, this is an exciting project involving a large number of entities in NWS and OAR within NOAA, UCAR/NCAR, and a few universities. The main ideas in subprojects went through prior community discussions (e.g., SIPs). There will be three major outcomes: the global forecasting system will include, for the first time, fully coupled modeling and DA system; the regional forecasting system will lead to the retirement of several legacy systems; the infrastructure will accelerate the model development and help entrain the broader community in the model/DA evaluation and development. Together they represent a remarkable modernization of the NWS suite of forecasting systems.

Key weaknesses are: The primary goals (delivering the new modeling system and DA) will depend on the availability of requested large computing resources, leading to a possible delay. It is unclear how much control the project leaders have on individual subprojects (particularly when progress is slower than projected). It is desirable for project leaders (at least those for the overall project and three teams (MER/S2S, CAM, CCI)) to commit at least 2 months/year to the project. Besides comparison with existing NOAA systems, the success metrics for model and DA development should include comparison with current ECMWF modeling system and with future ECMWF systems.

Overall score: xx

UFS R2O Project Review Criteria

Review Panel Meeting March 12-13, 2020

Reviewer 11

1. Does the proposed Project include a set of significant and achievable forecast goals for 2 or more NOAA UFS-based numerical weather forecast application systems, with at least half expected to be ready and available for NWS transition to operations by the end of the project period?

I would say yes. The proposed deliverables are seemingly aggressive given the length of the project period and the fact that most project areas have significant interdependencies on either other projects or implementations. Priorities and plans to mitigate the risks should be developed for some key sub-projects. Flexibility to descope would be reasonable. In addition, a well-coordinated testing plan across sub-projects should be designed to ensure that most deliverables can be ready as proposed by the end of the project.

2. Please comment on the overall scientific and technical merit of the proposed project.

   The proposed work provides a well-thought-out path forward to advance NOAA's operational forecast models. The R2O2R focus is beneficial to both NOAA and the broader community, and hopefully to pave a way to more efficient model development cycle and yield to more accurate systems with contribution from community. The use of cloud service for computation and data distribution is also an exciting technical innovation that can attract community engagement.

3. Does the proposed project include an effective strategy to bring research-to-operations and operations-to-research (R2O2R)? Does it include engagement with the research community to achieve forecast improvements and engagement with the forecast community to establish forecast goals?

   Yes, mostly. The code management plan for engaging community involvement and the tutorial and documentation activities to support community use is proposed. It is worthwhile to elaborate on the plans to engage and support the community in using the UFS system and datasets derived from this system. For instance, the level of support will be available to the community users on HPC and in Cloud computing environments. Data support/management plan on cloud-based observation data store and reanalysis data product.

4. Does the proposed project adequately foster code unification, contributing towards the reduction of the number of code versions in use by the UFS and NCEP?

   Yes.

5. Does the Project have a clear management and organization plan?

   Yes, mostly. To have a project engineer in place soon to overlook the whole project is desired. Web tools other than Confluence and Jira, to use an instant messaging platform (i.e. Slack) will facilitate communication between multiple institutions.

6. Is the proposed project appropriately aligned with other modeling efforts, including the broader UFS community, other programs/projects at NOAA, and with the broader national/international NWP and modeling community?

   Yes. Just to add to the communities mentioned in the proposal: The proposed activities here are also align with many model development and infrastructure tasks being proposed and conducted at E3SM (Energy Exascale Earth System Model) modeling center, i.e., component model development, modeling infrastructure, verification and diagnostics tools, infrastructure to create initial conditions for doing weather forecast runs (E3SM-SCREAM and ECP).

7. Has the project established appropriate success metrics?

   I focused on success metrics of coupled model, atmospheric physics and model components within MER/S2S applications, for familiarity. The success metrics for each of these components has close ties. Individual success metrics are presented for each sub-area, but it will be good to track the progress and tackle interdependencies with integration of all the model components. It will be helpful to establish metrics by listing the configuration of each version of the model and presenting the improvement of each component with version advancement. (Atmospheric composition needs to be included if aerosol-atmosphere-cloud interaction is an area of interest)

   In addition to success metrics of each sub-area, it would be good to set success metrics and measure progress for each Application Team as a whole.

8. Does the proposed project effectively leverage internal resources, including e.g. time commitment of federally funded NOAA investigators, computing resources, in-kind contributions, and other funded efforts?

   Availability of cloud resources for data storage and distribution is not discussed.

   Is community support planned for using cloud computing services?

9. Does the proposed project have adequate and appropriate resources, including sufficient time-commitment of team members?

   Computing resources are lacking based on current request

10. Does the project team have adequate and appropriate expertise to achieve the proposed goals?

Some model development tasks could be understaffed (need new hire or allocate employees from across institutions)- i.g. Sea ice, atmospherics physics developers, and coupled DA experts.

11. Please indicate which areas of the proposed work are most meritorious and crucial for improving NOAA NWS forecast systems. Also indicate those areas that could be descoped under reduced budget conditions.

    Most of the proposed areas are all meritorious and crucial for NOAA NWS forecast systems. One area that we can consider to descope might be the redacted. The infrastructure is important, but it is not obvious that the software packages proposed to create and develop in the session is well justified. A unification of these packages seems desirable. Alternatively, to prioritize these tools and come up with a strategy to co-operate on these packages in a unified environment will be useful.

12. Please provide a brief overall assessment of the proposed project and an overall score for the proposed project, where 1 is poor and 10 is outstanding. You may also provide ranking for sub-areas or aspects if you like.

    I would give xx to the overall project based on the proposal.