Summary Report of the Review of the Geophysical Fluid Dynamics Laboratory May 19 – May 22, 2014

Review Panel

Dr. Anthony D. Del Genio, NASA Goddard Institute for Space Studies, Chair
Dr. Sandrine Bony, Laboratoire de Météorologie Dynamique–Institut Pierre Simon Laplace
Dr. Eric Chassignet, Florida State University
Dr. Kathy Hibbard, Pacific Northwest National Laboratory
Dr. Raleigh Hood, University of Maryland
Dr. Drew Shindell, NASA Goddard Institute for Space Studies
Dr. Rowan Sutton, University of Reading
Dr. Shang-Ping Xie, Scripps Institution of Oceanography

August 18, 2014

Overview

The Science Review of the Geophysical Fluid Dynamics Laboratory (GFDL) was held at GFDL in Princeton, New Jersey, on May 19-22, 2014. The panel reviewed GFDL activities since 2009 in three research areas: (1) Modeling the Earth System; (2) Climate Variability and Change: Understanding and Prediction; (3) Chemistry, Carbon, Ecosystems, and Climate. We note that the review occurred in a very different fiscal and political context for GFDL and NOAA (National Oceanic and Atmospheric Administration) than the 2009 review. This should be kept in mind when considering the panel's recommendations.

The panel is grateful to the GFDL leadership and staff for their willingness to accommodate our requests during the review, and to all GFDL scientists for the obvious care and preparation that went into the presentations and posters that we listened to and viewed. Prior to the review, GFDL also provided a broad range of information to the panel about its history, intellectual achievements, workforce, funding status, educational connections, and stakeholder interactions. The panel recognizes how disruptive reviews such as these are to the life of the laboratory and appreciates the considerable work that was done by GFDL management and staff to provide the panel with a comprehensive picture of the lab and its relationships to NOAA's Office of Oceanic and Atmospheric Research (OAR), the rest of NOAA, the academic community, and local/regional customers of the information it produces.

As per NOAA instructions, the chair did not seek panel consensus in preparing this report. However, the panel agreed on many topics discussed in this review. Where that was not the case, this report notes the differing viewpoints of panel members. The chair wishes to thank the other panel members for the work they put into this review at the expense of their other obligations, their insights in all areas, and especially the highly collaborative character of the panel.

Summary of Laboratory-Wide Findings and Recommendations

GFDL continues to be one of the best, most respected climate modeling laboratories in the world. Hood says that "the quality of the laboratory's R & D ranks very high among other agencies and institutions." GFDL has built on the momentum it first developed a decade ago in the transition to the AM2/CM2 (Atmospheric Model Version 2/Climate Model Version 2) generation of its climate model. AM3/CM3 are among the best-performing global climate model (GCM) systems in the Coupled Model Intercomparison Project (CMIP). This achievement is notable given that GFDL's budget has generally declined since 2009, that its NOAA responsibilities have expanded, and that the latest model development cycle included the creation of its first fully coupled Earth System Model (ESM). Technical expertise at GFDL and the commitment and enthusiasm of its staff to support NOAA's mission are extremely high, and the panel wishes to convey to OAR and to the higher levels of NOAA our belief that GFDL is one of NOAA's highest-performing labs and is central to NOAA's ability to achieve its goals, and thus that its fiscal health should be one of NOAA's highest priorities in the years ahead.

Sutton summarizes GFDL's strengths as "i) world class climate models, being the product of a long term strategic approach to climate model development, focused on a commitment to quality before expediency and founded on robust scientific understanding of climate processes; ii) an outstanding group of scientists who combine expertise in climate model development with expertise in the innovative application of climate models to advance understanding of the climate system. These core strengths are not easy to achieve, nor maintain, and GFDL management should be congratulated for their sustained success." GFDL occupies a crucial niche in the climate change community, as noted by Xie: "GFDL scientists pioneered what I call the dynamical approach to global warming that tackles the challenging problem of circulation change relevant to regional climate change." He goes on to say, " What sets GFDL apart is the top performance of its models, and more importantly its preeminence as a leader in climate change science built on GFDL models." The panel emphasizes that it is critical that GFDL maintain these core strengths as it broadens its research portfolio. For example, Bony recommends that "the hiring of a young expert in geophysical fluid dynamics (especially one who would make a bridge between studies of moist and radiative processes and studies of atmospheric dynamics) should be considered among the top priorities of the lab." GFDL also provides considerable value to NOAA, as noted by Chassignet: "GFDL has done an exemplary job of fulfilling NOAA's mission to transfer research into products and services that can help NOAA meet evolving environmental, social, and economic needs of the nation."

GFDL's commitment to high-resolution, computationally intensive modeling has borne considerable fruit and is central to its future ability to carry out its science and its contributions to NOAA's mission. GFDL's current high-performance computing (HPC) resources are the product of two fortuitous injections of funding associated with national economic and weather disasters rather than a long-term NOAA plan to continuously upgrade GFDL's computational capabilities. The primary computational resource is housed at and operated by the Oak Ridge National Laboratory, whose contract ends next year. Hood expresses concern that "no specific plan was presented… that demonstrated a clear pathway to continued increases in computing capacity. Indeed, only a plan for the opposite contingency was presented, i.e., a plan for how to survive if there are no increases in computing capacity." Shindell states, "It is not clear how long-term planning for HPC is carried out at the highest levels, leaving GFDL in the difficult position of not knowing the future of the HPC systems on which the bulk of its work depends. Xie feels that "this is ever more urgent in light of the needs to build a comprehensive earth system model, for high resolution to better resolve extreme phenomena such as tropical cyclones, and for large initial-condition ensembles to sample natural variability. The computing power GFDL has is already falling behind what Japanese climate modelers have." The panel urges NOAA to assign the highest priority to providing GFDL with the necessary computing resources, as the first part of a decadal plan to periodically upgrade these resources as demands for ever-higher resolution models grow. Failure to do so will compromise GFDL's ability to meet NOAA's mission goals and will eventually erode its standing in the scientific community. Bony suggests that "the lab should anticipate the potential need to adapt to future changes in HPC architecture by including in its 'high-priority hires list' the hiring of an expert in this area."

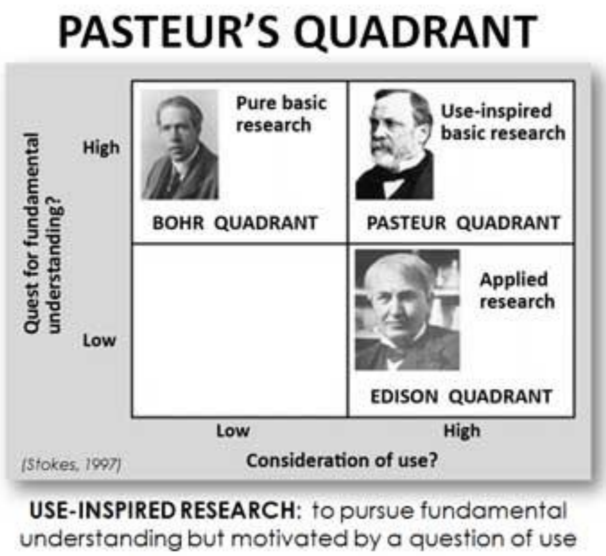

Figure 1

The panel supports the GFDL Director's vision of the lab's position in "Pasteur's Quadrant," at the intersection of fundamental science and applications (Fig. 1). The challenge – and opportunity – for GFDL moving forward will be how to remain there. Given GFDL's capabilities for simulating the current climate, it is perhaps uniquely positioned to address fundamental science questions about physical processes that control climate variability and change, and thereby to lead the community in developing process-based metrics for climate change projections. GFDL has invested considerable resources into its models and CMIP5 simulations. Interesting science is beginning to emerge, but now, in the interim between CMIPs, is the time to use these tools to perform even more exciting science that addresses fundamental questions. GFDL has a history of such science, but the balance may have shifted toward more practical goals such as model bias reduction and forecasts. Sutton recognizes "a burgeoning of opportunities to exploit and apply GFDL capabilities to an ever wider range of science questions and applications. Many of these opportunities are attractive and strongly aligned with NOAA's mission, but they also present a risk that GFDL could be distracted from its core science role of advancing fundamental understanding of the climate system." He goes on to say that "the standard of science presentations in the review was very high, but relatively few addressed fundamental questions about the climate system. It is likely that this is partly because of the heavy focus on model development in recent years, but there is now an opportunity for rebalancing to exploit the new modelling tools." Del Genio worries that "without tending to this, GFDL risks sliding from Pasteur's quadrant into Edison's quadrant – not a terrible thing, but less than what GFDL can potentially be." Xie cites an example from GFDL's past, that sea surface temperature warming relative to the tropical mean determines future changes in tropical cyclones, noting that "GFDL scientists articulated the physical mechanism convincingly and made it widely accepted." Bony says that "GFDL should not only phrase its scientific goals in terms of societal missions and applications but also in terms of more fundamental, long-standing science questions (more than it is currently the case). It will make the lab even more attractive to a wider range of curious and creative minds." The cross-cut of fundamental and applied science will maintain GFDL's presence in Pasteur's quadrant.

GFDL will always have to balance its foundational basic research, obligations to NOAA, and the demands of the Intergovernmental Panel on Climate Change (IPCC). The panel congratulates GFDL for the leadership role played by several of its scientists over time in developing and coordinating the scientific and infrastructure aspects of CMIP. External responsibilities are part of GFDL's value, as Hibbard explains: "GFDL leadership and collaboration with the Program for Climate Model Diagnosis and Intercomparison (PCMDI), CMIP has facilitated the great success of the climate modeling community's development and understanding." However, GFDL as a whole may have been too consumed by CMIP5, devoting 50% of its resources to it.

GFDL intends to reduce its commitment for CMIP6. Shindell supports this and feels GFDL should "create a better balance during that next phase between providing data to the various MIPs and spending time and effort on analysis and interpretation of the simulations to advance understanding of climate processes."

The panel commends GFDL for developing a Strategic Plan, but some panel members felt the Plan was too long and all-encompassing to be of value. Chassignet, for example, suggests that "a significantly shorter version of the Strategic Plan, which highlights the areas that GFDL hopes to emphasize over the coming few years in both its deployment of existing personnel and its hiring/promotion strategies, would be very helpful to the lab's employees. A clear setting of 3-5 research priorities would also help GFDL decide how best to control (rather than be controlled by) the IPCC process and how to deal with the possibility of declining computing resources."

The panel notes the creation of the Science Board and re-organization of the Research Council in response to recommendations for more top-down organization and to ensure that more voices are heard in the development of priorities. This has been partly successful and is necessary, given the need to coordinate GFDL's many activities. It has also allowed several young scientists to be placed in leadership roles, which has increased morale; GFDL should be commended for this. It was not clear, though, how well this structure operates in practice: Whether the increased size of the Research Council has reduced its responsiveness, whether the voices of the rank-and-file and other young scientists are being heard at the highest levels, and whether decisions and the basis for them are being communicated in timely fashion and are understood by the staff. Strategies for recognizing non-supervisory staff should be strengthened, and there should be more frequent input from young scientists to lab science strategy. Del Genio suggests that the current structure may work with better communication, e.g., via prompt posting of Research Council minutes, discussions between group leaders and members about the outcomes of meetings, solicitation of staff input by the Science Board, and clear communication about career advancement criteria. Shindell suggests that "additional efforts to more clearly recognize the contribution of the modeling services personnel and to ensure that non-Federal employees are nominated for those awards for which they are eligible would be beneficial." Bony emphasizes that recognition is important "especially to the people who don't have much visibility in the lab's organizational structure." Several panelists felt that the governing structure itself should be modified. Long-term strategy is primarily the purview of the Science Board, which includes only senior scientists. Hood feels this "is, perhaps, not the best strategy." Sutton comments that "in rapidly changing times, it seems essential to include some younger voices."

GFDL has made some progress since the previous review toward gender balance but still lags behind many other labs. The fairly balanced gender composition of today's geoscience graduate student population is not yet reflected in the GFDL staff, and even less so in leadership. The degree of imbalance at upper levels is such that pure demographic changes over time may not eradicate it. The panel recommends reconsidering hiring and promotion practices and whether these contribute to the problem. Bony encourages GFDL "to organize trainings about the so-called 'unconscious bias', and to set up women's mentoring." Now that GFDL has a few women in leadership positions (though not yet on the Science Board), their perspective is especially important. This is not only a matter of lab morale and retention. It is also a question of whether

GFDL can attract and retain the best scientists to enable it to carry out its NOAA mission and retain its scientific stature if it does not adequately draw from half the pool of available talent.

GFDL's relationship with the Cooperative Institute for Climate Studies (CICS) at Princeton University appears to be healthy and productive. A critical mass of graduate students is advised and taught by GFDL scientists, continuing its impressive history of shepherding young scientists who have become prominent in their own right. CICS provides code used in the GFDL ESM, a successful university-government collaboration that is not replicated at many other labs. Hibbard states that "the CICS relationship has enabled GFDL to grow and develop a high quality land counterpart" to the ocean and marine dynamics activities supported by NOAA. CICS graduate students report that GFDL scientists are "generous with their time" but wish there were more mentors in more areas. The 2009 review panel strongly recommended relocation of GFDL to the Princeton campus to solidify these ties. This sentiment was echoed by the graduate students and the Director of the Princeton Environmental Institute. Several panel members expressed opinions on this, but the topic was not central to the panel's discussions. Xie feels that there are "clear mutual benefits for Princeton and GFDL to move the lab to the main campus" and that a Princeton faculty hire in atmospheric/climate dynamics might strengthen ties and facilitate an eventual move. Del Genio feels that "the relationship as it stands works well and that the GFDL Director and Science Board are in the best position to decide whether a move would be feasible and advantageous to GFDL."

GFDL interacts with stakeholders and with the modeling community through CMIP, but it is more insulated from the outside world than other climate modeling centers. Sutton suggests that GFDL review its partnership strategy: "Many competitor models are used by a much wider user community. Whilst there are undoubtedly costs in supporting external groups or a broader community, there are also potential benefits, especially if partnerships are carefully managed to achieve mutual benefits. The Met Office, for example, has pursued such a strategy effectively." Xie would like to see GFDL "make CM2.1 the standard model of climate dynamics, used by users all over the world, by providing some support for bug fixing and physics upgrading, much as GFDL did for MOM (Modular Ocean Model)." Hibbard suggests "a science plan that outlines near and long-term objectives and relevance not only to NOAA's mission, but to the community writ large." Chassignet wonders about the long-term strategy for GFDL's web, education and outreach activities: "As a flagship, GFDL's web presence needs to be exceptional."

Summary of Individual Ratings

(O = Outstanding; S = Satisfactory; N = Needs Improvement; research areas which a given panel member does not have sufficient expertise to review are left blank; hyphens indicate preferences of some panel members not to break down their rankings into specific categories.)

	Del Genio	Bony	Chassignet	Hibbard	Hood	Shindell	Sutton	Xie
Earth System Modeling	O	O	O		S			O
Quality	S/O	O	O		-			-
Relevance	O	O	O		-			-
Performance	O	O	O		-			-
Climate Var. & Change	O		O				O	O
Quality	O		S+				-	-
Relevance	O		O				-	-
Performance	O		S+				-	-
Chemistry, Carbon, Ecosystems, & Climate				O		O		
Quality				O		O		
Relevance				O		O		
Performance				S+		S/O		

Findings and Recommendations by Research Area

1. Modeling the Earth System

Earth System modeling includes not only the physical climate but also chemistry, carbon, and ecosystems. We restrict discussion in this section to overall Earth System modeling and its physical climate components, and reserve specific comments about the other components for research area #3, although individual panel member comments were given in both sections.

Like all GCM groups, GFDL leads the field in some areas, adopts some elements developed by other groups, and plays catch-up in other areas. Xie notes that "GFDL models are ranked consistently at the top in simulation skill by various metrics. Their model development takes a system approach that combines bottom-up (process-based parameterizations, e.g., atmospheric convection) and top-down (macro-structure performance measures, e.g., ENSO), made possible by strong in-house expertise in both atmospheric/ocean physics and dynamics." Hood says that "GFDL conducts preeminent research in ESM and their advancements represent significant contributions to the scientific community. This statement is particularly true as it relates to the more mature aspects of the ESM research that is being undertaken at GFDL, i.e., ocean and atmospheric modeling." The high-resolution CM2.5/2.6 and HiRAM (High Resolution Atmospheric Model) models perform impressively against standard metrics and for tropical cyclone (TC) proxies, respectively. Del Genio remarks that CM3 is incrementally better than CM2.1 but that this is a considerable achievement since CM3 includes interactive chemistry, aerosol-cloud interactions, and dynamic vegetation that CM2.1 does not: "Coupling to these other aspects of the Earth system often degrades rather than improves model performance because of the additional degrees of freedom and our less complete understanding of other parts of the Earth system. The fact that CM3 maintains the quality of CM2.1 as judged by standard (albeit imperfect) climate metrics while making it a more versatile, comprehensive model of the Earth system, applicable to a wider variety of problems, is a testament to the high quality and dedication of the GFDL model development scientists and evidence that they can work together productively."

The multiple modeling streams at GFDL [CM2.5/2.6, FLOR (Forecast-oriented Low Ocean Resolution), CM3, ESM] were a point of considerable discussion. Xie is mostly positive, calling it "a reflection of the needs for diverse scientific objectives. The downside is the diversion of resources and diffusion of the lab's focus. GFDL recognizes this issue and has set up the lab's effort to develop a single trunk model CM4 that incorporates positive attributes of these stream models. This cycle of divergence into stream models and convergence into a single lab-wide trunk model reflects the exploratory nature of model development. It appears to be an effective strategy to maintain a lab-wide focus while exploring ways to improve models." Bony notes that the multi-stream effort, combined with CMIP5 responsibilities, "had a significant cost in terms of human and computational resources." She therefore supports the consolidation plan for the next-generation CM4 and notes, "The skill of this new model (AM4/CM4) in simulating the current climate already looks very promising." However, tradeoffs are inevitable in combining models developed independently for different science goals into a single modeling system. Chassignet notes, "GFDL would benefit from a scientific discussion of the limitations of current and future Earth Systems models. Given the complexity of the Earth system, GFDL cannot do it all and needs to justify its choices."

GFDL describes their philosophy as one of alternating diversification and consolidation cycles. The current GFDL GCMs use three different cumulus parameterizations, while a new unified cloud and boundary layer turbulence scheme is being tested. Resolving these issues, Del Genio says, "are urgent priorities, because these inordinately influence other aspects of the coupled system and to a large extent control the model's simulation of historical temperature trends and future cloud feedbacks, the cause of much of the spread in climate sensitivity of the CMIP5 models." Shindell supports the trunk concept but notes the upside of model flexibility: "They should nonetheless keep in mind that for the longest climate simulations the optimal balance between complexity and ability to perform adequate ensembles with multiple sets of drivers is not obvious and there might be a role for simulations using a model other than their highest resolution version with the most comprehensive representation of processes. There may also be a role for greater testing of the results' sensitivity to uncertainties in physical processes."

Panel members viewed GFDL's Cloud-Climate Initiative favorably. Bony remarks that it is "a very relevant and opportune opportunity to better connect the different groups of the lab, and to address ambitious science questions. Hopefully this initiative will help GFDL contribute more to the understanding of cloud-climate feedbacks than has been the case over the last years," although she notes GFDL's recent work on convection's role in cloud feedback.

Chassignet says that "GFDL's commitment to high resolution, computationally intensive modeling is central to its future ability to carry out science and support NOAA's mission." He notes, however, “It was unclear from the review how the various modeling groups interact with each other to ensure cohesiveness. Once timelines are set, it may be beneficial to formalize how communication will take place among the groups.” He also points out that “with the increased resolution and additional components, GFDL will soon not be able to fully validate all model outputs. In order to increase external community involvement, particular attention needs to be put on data portals and code availability.” He also suggests that GFDL adopt a more transparent naming convention for its multiple models for the benefit of outside users. Bony notes that the development of the new computationally efficient dynamical core “provides GFDL with a model well suited for high-resolution and seamless prediction studies. Given the development of such a modeling framework expected to be used over a large range of resolutions, it will be important to invest commensurable efforts in the development and testing of scale-aware parameterizations.”

The GFDL MOM is widely used by other modeling groups. GFDL runs MOM at finer and coarser resolutions that roughly resolve and more strongly parameterize mesoscale ocean eddies, respectively. If they can use the former to develop a more capable eddy parameterization for the latter, GFDL can become a community leader in promising fundamental science areas such as the role of the Atlantic Meridional Overturning Circulation and the southern ocean in determining transient climate response. Chassignet notes that “no clear strategy was presented on how ocean overflow parameterizations will be implemented,” and that while this depends on the eventual choices of vertical coordinate for MOM6, it must be planned for in the near future. GFDL has successfully used participation in a Climate Process Team (CPT) to study this topic in the past, as well as more recently to develop an interactive dynamic land ice model that could eventually make it a leader in projecting future sea level rise. The CPT approach has worked less well for atmospheric physics; Bony suggests that “maybe the lab could design and propose CPTs itself that would put atmospheric model development at the forefront of the objectives.”

GFDL has done an outstanding job of performing research that is relevant to NOAA. Xie states that “GFDL develops top-performing models as tools for prediction and projection, thereby contributing to NOAA strategic goals of climate adaptation and mitigation, weather-ready nation, healthy oceans, and resilient coastal communities and economies.” He lists GFDL’s participation in CMIPs and IPCCs, the NOAA North American Multi-Model Ensemble seasonal forecast, the ocean data assimilation system it provides to NOAA’s National Centers for Environmental Prediction (NCEP), and the operational hurricane model it provides to the National Weather Service (NWS) as examples. Further details of NOAA-relevant applications will be discussed in the section on research theme #2. GFDL CMIP/IPCC participation has included model outputs and also the leadership roles played by several GFDL scientists. Bony, however, notes that “GFDL did not participate much in the idealized experiments of CMIP5 focused on science questions such as the understanding of cloud-climate feedbacks or the interpretation of inter-model differences in climate projections.” This may be due to the difficulty in reducing the complexity of GFDL models and to human resource limitations. She suggests that GFDL might address this with the more selective CMIP6 process.

GFDL appears to have weak, at best, connections with several NOAA labs. Bony says that “the connection between GFDL and NCEP's model development efforts is unclear.” Del Genio says, “GFDL is a tremendous model development resource… that is not taken advantage of by NCEP”. He also notes that “GFDL differs from many other global modeling centers in not being part of a larger institution that does process-level and observational science." This limits its ability to balance top-down constraints with bottom-up process-level development. He says, "The true test of whether a GCM has been improved by a new parameterization is not just whether it reduces mean state biases, but also whether it represents the process in question more faithfully." Collaborations with NOAA's Earth System Research Laboratory (ESRL), which does the type of observational and process-level science that GFDL does not, and other NOAA labs might be fruitful in an era of limited resources. Xie worries about continuing GFDL's strength in physical modeling while adding more Earth System components: "Without additional resources being allocated to the GFDL budget, this would force GFDL to make some difficult choices in allocating resources. This, however, represents an opportunity to organize a NOAA-wide earth system modeling effort, where components developed in other NOAA labs are assembled and integrated at GFDL into a coherent and consistent system. Examples include atmospheric chemistry at ESRL, and marine geochemistry at PMEL (Pacific Marine Environmental Laboratory) and AOML (Atlantic Oceanographic and Meteorological Laboratory)."

2. Climate Variability and Change: Understanding and Prediction

GFDL's heritage includes major contributions to the study of climate variability and change, and current lab research continues this tradition. Xie says, "GFDL is one of only two U.S. centers (with NCAR) that conduct comprehensive research into climate variability and change with in-house modeling capability." Sutton points out the challenge of addressing such a large area of research but feels that "the laboratory has made judicious choices about where to focus effort. It is conducting pre-eminent research in many of its focus topics, such as: Seasonal forecasting of regional tropical cyclone risk; ENSO; detection and attribution of regional changes in the hydrological cycle. Much of GFDL's work in these areas is world-leading, and is recognized as such by stakeholders." Chassignet agrees that GFDL "is arguably at the forefront of tackling the problem of how tropical cyclones/hurricanes/typhoons will vary seasonally and in response to anthropogenic climate changes" but feels that other areas are not necessarily of the same caliber.

GFDL leads the world in the climatic understanding of tropical cyclones. Del Genio claims that "with the global HiRAM model, GFDL has simulated the geographical and temporal statistics of TC-like disturbances better than any other modeling group," a result of "the move to 50 km resolution and changes in the cumulus parameterization that were designed to optimize TC simulation." He notes that "since HiRAM is a candidate for the CM4 trunk model, an important question is whether the tuning done to reproduce TCs compromises the mean state, other aspects of climate variability, and features of long-term climate change." At the review, several GFDL studies of the controversial topic of aerosol effects on TCs were presented, with inconsistent conclusions. Del Genio suggests that GFDL coordinate this research among its groups: "GFDL, better than any other institution, is in a position to bring objective science to bear on the question of aerosol effects on the entire Atlantic basin, by looking carefully at the strengths and weaknesses of its aerosol representation vs. observations, by considering how aerosols are interacting with precipitating clouds in their hurricane models, by breaking down the specific role of aerosols vs. that of correlated quantities such as low humidity and warm temperature that accompany aerosol-bearing African air masses, and so on. This could be an area for GFDL to make a large science impact, if it can get its models to converge and make a compelling case for the definitive science."

GFDL also simulates TCs on finer spatial scales, using double-downscaling to make a promising assessment of how Atlantic TCs will respond to anthropogenic warming. At the review, the "fewer but stronger" conclusion that several groups have reached was mentioned, but tentatively. Del Genio suggests that uncertainties in sea surface temperature trends may now be more of an issue than the physics of the TCs, but urges GFDL to "define what they see as the tall pole(s) preventing definitive conclusions about Atlantic TCs and focus on these, because no other group is better positioned to solve this problem." This feeling extends to other aspects of anthropogenic climate change, according to Xie: "Dynamics of radiatively-forced climate change, especially that related to ocean-atmospheric circulation, has been a strength of GFDL research since Manabe and is emerging as a grand challenge that holds the key to reducing uncertainties in regional climate projections. GFDL is home to several leading scientists in this area. With the scientific talents and best-performing models, GFDL is well positioned to address the grand challenge of circulation response to climate change." GFDL provides outstanding service to NOAA by making routine hurricane forecasts for NWS, including a highly successful Sandy forecast, and recent improvements allow GFDL to reduce intensity errors and better predict explosive growth. GFDL also provides track and intensity forecasts to the Joint Typhoon Warning Center, which uses them in an ensemble to initialize their forecasts.

GFDL is also engaging in research on several aspects of seasonal to decadal climate variability. The FLOR model, based on a philosophy that a high-resolution atmosphere can be usefully coupled to a lower resolution ocean for shorter time scale studies, has shown promise. Sutton says, "The development of a seasonal forecasting capability based on the FLOR model is a notable achievement and is yielding valuable results." Xie emphasizes that "this TC-permitting system enables the first dynamic prediction of TC track density at the seasonal and longer leads. It even shows some skills in predicting landfall hurricanes in the Gulf of Mexico." FLOR contributes to the North American Multi-Model Ensemble seasonal forecast. It has also been used to show that El Niño-Southern Oscillation (ENSO) imposes some predictability, up to 9 months, on the strength of the midlatitude storm tracks. Panel members noted that GFDL's success in simulating the dynamics of the coupled system on these shorter time scales should be leveraged in other ways. Sutton suggests "the greater application of atmosphere/ocean dynamical understanding to process-based evaluation of predictions and detection/attribution results. This approach is followed in some areas (e.g. ENSO) but its wider application (e.g. in detection and attribution) would be beneficial." Del Genio notes that GFDL also finds changes in location and strength of the storm tracks on longer time scales, but with some regional similarities and differences relative to the ENSO signal: "If this difference between short- and long-term responses is robust and well-understood…, and especially if it is sensitive to getting a particular aspect of the physics correct, this could be an area for GFDL to establish a useful current climate metric for evaluating one important aspect of projected long-term climate change."

Applications of GFDL decadal predictability research were presented in several areas, including the possibility of skillful multi-year hurricane outlooks, and attempts to understand the recent global warming "hiatus." GFDL's role in this field appears to still be evolving. At the time of the previous review, GFDL's budget was increasing, NOAA had plans for a National Climate

Service (NCS), and the discussion was what GFDL's role would be in providing climate forecasts to the NCS. Five years later, with the NCS concept having been rejected by Congress and GFDL's budget mostly decreasing in recent years, GFDL needs to assess its participation, in both the research and forecast arenas, in decadal-scale climate science. Del Genio notes the great desire by GFDL's regional and local stakeholders for these types of forecasts but feels that GFDL should go no further than experimental outlooks, with caveats made clear: "The important issue for NOAA (and for GFDL staff as well) is to effectively communicate the experimental and highly uncertain nature of such outlooks so that policy at the local level is not driven more strongly by these outlooks than is warranted." On the research side, GFDL should decide what its priorities are, whether extreme event attribution (and understanding, as Sutton emphasizes), Atlantic TCs, or the possible underestimate of decadal natural variability by climate GCMs.

Sutton recommends that GFDL assess its many climate variability activities to strike an appropriate balance and to decide where to participate seriously: "Managing quasi-operational activities is a particular challenge in a cultural environment that has traditionally been focused on basic research. There is a need to clarify the longer term ambitions of GFDL in respect of operational seasonal and decadal forecasting, including GFDL's role in relation to other NOAA labs. The North American Multi-Model Ensemble for Seasonal Prediction (NMME) plays an important role here, but there are many other issues. Seasonal forecasting in particular is developing rapidly internationally. Therefore if GFDL wishes to remain competitive in this field a very clear strategy, and adequate resourcing, will be essential. More attention may be needed to issues such as representation of the stratosphere and initialisation of the atmospheric state. For example, there is evidence that initialisation of the Quasi-Biennial Oscillation is relevant for predictions of the North Atlantic Oscillation, which is itself important for many aspects of climate (AMOC [Atlantic Meridional Overturning Circulation], Arctic, etc). What level of priority will be accorded to the prediction of sub-seasonal variability (which is arguably inseparable from seasonal variability)? In decadal forecasting, the lab has made a valuable contribution to the CMIP5 experiments but there is a need to clarify the longer term strategy in this area." He also argues that GFDL needs to decide "how the needs of seasonal-to-decadal forecasting can and should influence model development priorities." Xie feels that a federal lab is the right place for such forecasts but echoes Sutton's warning about resources in computing - "Securing adequate computing resources is a major challenge for GFDL as a whole but this issue is especially severe for FLOR forecasts" – and human resources - "This seems to be a high priority for federal hirings." Chassignet is concerned that the implementation of data assimilation in CM4 is lagging the coupled model's development: "Data assimilation is strongly model dependent and needs to be developed in parallel."

3. Chemistry, Carbon, Ecosystems, and Climate

At the time of the previous review, GFDL was in the early stages of a transition from a purely atmosphere-ocean Global Climate Model to a more comprehensive Earth System Model (ESM). Not surprisingly, that aspect of GFDL's research was viewed less positively than its more established climate dynamics and change research. Five years later, while the transition is still in progress, GFDL has made significant strides. GFDL CM3, which is based on the latest (AM3) atmospheric physics, is the first operational GFDL model to include interactive chemistry, aerosol-cloud interactions, and dynamic vegetation. GFDL also now has two operational ESMs, both based on the CM2.1 generation climate model, that also incorporate interactive biogeochemistry, including the carbon cycle. The two ESMs (ESM2M, ESM2G) differ from each other primarily in their physical ocean component, each with its strengths and weaknesses. All three GFDL models participated in CMIP5.

Shindell feels that "GFDL's science in the areas of chemistry, carbon, ecosystems and climate is outstanding. Since the last review, they have clearly made great efforts to expand the scope of their carbon and ecosystems modeling. This has clearly been successful, with their modeling in this area now at the leading edge worldwide, and contributing to the status of the overall GFDL AM3/CM3 models as among the very best within the current generation." Hibbard agrees, saying that "GFDL has strongly responded to the 2009 recommendations towards new hires and developed a comprehensive biogeochemistry component to the Earth system modeling framework. What is impressive is that they have not only developed marine biogeochemistry but have also developed concurrent land biogeochemistry modeling components to the traditional atmosphere in a coupled carbon-climate and chemistry framework." Hood acknowledges that the newer Earth System components are still "less mature" than GFDL's core atmosphere-ocean models but recognizes that "the land surface, atmospheric chemistry and the ocean biogeochemical modeling efforts, have also made important contributions more recently." Regarding ocean biogeochemical cycles (BGC) in particular, Hood says, "There is no doubt that the BGC modeling group is doing meritorious work and are making significant contributions to the field." Overall, though, he feels that GFDL is still catching up in these areas: "The land surface, atmospheric chemistry and ocean BGC modeling efforts are not as well known or as well-regarded as some other institutions that have been doing this kind of work for much longer, like, for example, NCAR with their Community Earth System Modeling (CESM) effort."

In atmospheric chemistry, Shindell states, "GFDL performed transient simulations with coupled gas-phase chemistry (tropospheric and stratospheric) and aerosols under CMIP5, again placing them at the cutting edge internationally. They also participated heavily in ACCMIP, leading an analysis and providing input that provided important underpinning to chapters on forcing and future climate in the IPCC AR5." Hibbard notes that "GFDL now has the capability to calculate atmospheric CO_2 from emissions information. This capability will enable comparisons with emissions trajectories provided by the integrated assessment modeling (IAM) communities. Both the chemistry-climate and IAM communities can learn a lot from each other. The simulated emissions trajectories provided by GFDL are a product of the seamless atmospheric chemistry activities and, as such, consider high resolution atmospheric physics, dynamics, ocean and sea ice models." Both Shindell and Hibbard applaud GFDL's participation in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), with Shindell remarking that GFDL led an analysis that "that provided important underpinning to chapters on forcing and future climate in the IPCC AR5," and Hibbard appreciating that GFDL's in-depth study of "the drivers of CH_4 lifetimes will provide insight into the dynamics of CO_2 versus CH_4 emissions and trajectories for many of the global climate and biogeochemistry models."

The land biogeochemistry and ecosystems area includes development of a nitrogen cycle, though Hibbard notes that GFDL recognizes "that the LM3-N is appropriate only as a stand-alone, or offline model as the N limitations in that model distort the seasonal water and CO_2 dynamics." She is impressed by other aspects of the land surface research, e.g., "What is truly innovative is the development of a strategy to incorporate game theory to explore invasives and propagation dynamics," and she suggests that GFDL might usefully consider "some of the old optimal foraging/random walk theories." Hibbard also notes that GFDL is thinking about the representation of vegetation structure and function. She cautions that as model resolution increases, eventually "individual, or patch dynamics will be really really important to represent for example, the spatial co-location of vegetation and overland streams. As time and space resolution increase, a simple statistical representation will not be adequate to capture, e.g., biogeochemical fluxes".

Hood feels that "the GFDL researchers involved in ocean BGC modeling have clearly demonstrated scientific leadership and excellence in their field." He goes on to say, "This group has established a significant national profile in a relatively short time." Hibbard feels that GFDL has strengthened itself with new hires, allowing them to develop "a flexible N:P functionality to update the traditional stoichiometric C:N:P:Z biogeochemical models for marine environments. In addition, the development of the trophic dynamics through phyto- to zooplankton and krill community ecology provide an avenue for applications to NOAA fisheries as well as incorporate the biological components that impact biogeochemistry." GFDL's leadership in the Marine Ecosystem Model Intercomparison Project (MAREMIP) "will provide an avenue to both promote and enable GFDL and NOAA's leadership in marine/fisheries as well as biogeochemical interactions and carbon cycle feedbacks/interactions with the climate system," in Hibbard's view.

GFDL's chemistry, carbon, and ecosystems research appears to the panel to be highly relevant to NOAA. Shindell argues that "GFDL studies of pollution transport from East Asia across the Pacific to the Western US are clearly relevant to the national interest in understanding the local and remote factors contributing to air pollution in that region." He further states, "Their arguably world-leading work on the carbon-cycle with a climate modeling context is clearly of relevance to both NOAA-specific goals related to ocean management and the nation's larger goal of understanding and projecting climate change." For example, GFDL provides the NOAA National Marine Fisheries Service (NMFS) with quantitative seasonal to decadal predictions and projections of interactions between climate and marine ecosystems; NMFS appears to be very enthusiastic about the value of these products. Hibbard suggests that "Coastal resilience and sustainability… are not (yet) a part of the GFDL portfolio, and, perhaps, it is inappropriate for a global climate model to be approaching this pillar, however, the dynamic and highly resolved atmosphere/ocean groups might consider collaborations, or risky science in these arenas, as funding and opportunities present themselves."

Chemistry, carbon, ecosystems and climate (CCEC) is probably the area in which GFDL has had the most success in addressing its gender imbalance problem, but echoing the opinions about the lab-wide issue of women at higher levels, Shindell urges GFDL to focus on the transition to leadership: "The several excellent younger female scientists currently working in carbon and chemistry, as well as newcomers over the coming years, need to feel they have a potential future in the upper ranks at GFDL."

Panel members emphasized the need for GFDL to organize its ESM efforts and make decisions about how model components will be combined and at what level of complexity. Hibbard recommends that the CCEC group develop “a flexible framework or document that outlines current and future planning/strategies for near and long term development. As the biogeochemistry and ecosystem components unfold and develop, it is recommended to also have an understanding of the scope and scale of questions that can and should, versus those that should not be addressed. This will be key for impact, adaptation and vulnerability communities, as well as those involved in scenario development for the international, e.g., IPCC, as well as the development, e.g., UNEP [United Nations Environment Programme] type communities.” Shindell feels that “GFDL’s plans to continue linking atmospheric and biospheric trace species fluxes such as the nitrogen, methane and iron cycles are well thought out and a track they should continue to follow. The leadership in these efforts should continue to evaluate the appropriate level of complexity to incorporate recognizing the tradeoffs between complexity and utility (speed and ease of understanding the model’s behavior) involved.” Hood shares some of these sentiments, stating the need for GFDL “to develop a concrete plan and set of goals especially as it pertains to the scope and aims of its land surface, atmospheric chemistry and marine biogeochemical modeling. This plan should define the scope and scale and scientific objectives of these separate ESM efforts, and it should define a specific strategy and roadmap for ultimately linking these models into a fully coupled ESM in the future. This plan must embrace both scientific challenges (e.g., What are the key state variables that will be dynamically coupled? Which components of the models will not be dynamically coupled and how will the boundary conditions for these uncoupled components be specified?) and methodological issues (e.g., related to how the models will be coded so they can be coupled in a flexible manner). It should also articulate the scientific questions that will be addressed with such a model.” Hibbard recommends that this take place in the context of other U.S. ESMs “to get a sense of where strengths and weaknesses exist and where opportunities and challenges can be met through leveraging or partnership.”

Summary of Recommendations

1. GFDL is one of the world’s best global climate modeling centers and one of NOAA’s most valuable assets. It has strong management, high technical expertise, and a very high commitment and enthusiasm of its staff to support NOAA’s mission. It is imperative that despite the current challenging fiscal environment, OAR and NOAA make every effort to sustain and if possible increase their support for GFDL. In particular, as GFDL expands its purview to broader Earth System Modeling and to more quasi-operational applications for NOAA, it is essential that the core of excellence in physical climate and dynamics that has been GFDL’s hallmark and the source of its stature in the community be maintained.

2. GFDL’s commitment to high resolution, computationally intensive modeling has borne considerable fruit thus far and is central to its future ability to carry out its science and NOAA’s mission. The panel urges OAR and NOAA to assign the highest priority to finding the means to provide GFDL the necessary computing resources to continue its upward trajectory. This should be facilitated by the development of a long-term plan for the evolution of computing capabilities to reduce GFDL’s dependence on serendipitous sources of support for added computing power. Failure to do so will compromise GFDL’s ability to carry out important NOAA mission goals.

3. The panel supports the GFDL Director's vision of the lab's positioning in Pasteur's quadrant. GFDL has done an exemplary job of combining fundamental and applied science, but the challenge going forward is how to remain in that position. Given GFDL's capabilities for simulating the current climate, it may be uniquely positioned to address fundamental science questions about physical processes that control climate variability and change, and thereby to lead the community in developing process-based metrics for climate change. To do so will require a realistic approach, consistent with funding levels, for allocating the time spent on forecast applications, model bias reduction, and CMIP6 participation to allow fundamental science to flourish. A streamlined version of the Strategic Plan, highlighting 3-5 priorities that GFDL wishes to emphasize over the next few years in its deployment of personnel and its hiring priorities, will be very useful to lab staff and help GFDL accomplish their goals. This should be consistent with the lab's plan to balance fundamental and applied research, and short-term tactical decisions should emerge from that.

4. To maintain GFDL's core strengths as it expands the scope of its research, the panel recommends that atmospheric and climate dynamics be a priority area for future hires. This might include a faculty hire at Princeton to further strengthen the ties between GFDL and CICS. The panel supports GFDL's Cloud-Climate Initiative and suggests that one area in which GFDL could try to become a leader is in understanding the interaction between moist and radiative processes and the general circulation, which has been recognized internationally as a grand challenge of climate science. If GFDL decides to become more involved in forecasting and experimental outlooks, a hire in this area will be necessary as well. The panel also suggests that HPC architecture should be another priority for a future hire, given the computational challenges that lie ahead.

5. GFDL's transition to full Earth System modeling has generally proceeded well. The implementation of biogeochemistry into the modeling system is viewed by the panel as being at an appropriate level of complexity. Ecosystem science has gotten off to a good start, and good leaders and young scientists are in place. GFDL now needs to develop a concrete plan for near- and long-term Earth System model development in the areas of land surface modeling, atmospheric chemistry, and marine biogeochemistry. Tradeoffs between complexity and utility should continually be re-evaluated as knowledge and resources evolve. Decisions should be made in the context of the scientific questions that GFDL feels they can address themselves and those challenges that are best met by leveraging or partnerships with other ESM groups.

6. The panel fully endorses the consolidation of GFDL's various modeling streams into a single CM4 trunk model and urges that decisions about the individual parameterizations of this model be made as soon as possible. The panel recognizes, though, that some flexibility in resolution and level of complexity will be useful to retain to allow GFDL to address questions on a variety of time scales from seasonal to centennial and that testing of sensitivity to parameterization assumptions can be a valuable component of its research.

7. GFDL cannot do everything in a field as broad as climate variability and change. The panel recommends that given constrained resources, it focus on the areas in which it can have the most scientific impact. GFDL leads the field in tropical cyclone research and clearly should continue

to emphasize this. In other areas, GFDL must make choices about where it can and wishes to be most competitive and allocate resources appropriately.

8. GFDL has been exemplary in fulfilling NOAA's mission to transfer research into products and services that meet evolving environmental, social, and economic needs of the nation. It provides great value to the U.S. in areas such as tropical cyclone forecasting, seasonal outlooks, and downscaling for regional and local stakeholder use. These are appropriate activities for a federal lab, but it is very important that GFDL decide on the level of its participation in operational or quasi-operational forecasting based on a very clear long-term strategy that considers resource, cultural, and competitive issues. This might result in GFDL taking the lead in one or more areas, or something less ambitious. Part of the strategy should involve clear plans about how prediction system performance will be used to inform model development priorities. It will be crucial for GFDL to clearly communicate the limitations of any regular outlooks to stakeholders, who are prone to over-reliance on such information in making policy decisions.

9. GFDL's cooperative agreement with Princeton University (CICS) appears to be a great success, perhaps one of the best examples of a government-university partnership in the U.S. The two institutions appear to be well integrated, have attracted a critical mass of graduate students, and have together produced a large number of Ph.D.s who have gone on to be successful in the field. There is enthusiasm for moving GFDL to the main Princeton campus from both students and faculty, but the logistical challenges and the success of the current arrangement need to be considered in deciding what course of action would be best for GFDL.

10. GFDL should consider the potential benefits of broadening its partnership strategy beyond that which already exists with Princeton. In particular, GFDL models other than MOM are not as widely used by the community as many competitor models. Other centers have found that the costs of interacting with a large community of users are more than offset by the wider range of applications and analysis of model strengths and weaknesses than is possible using only in-house resources. This could be one component of a more general strategy to expand the lab's web presence and its education-outreach activities. The new Cloud-Climate Initiative might be a proving ground for such an effort.

11. GFDL should also consider broadening its research partnerships within NOAA, where its interactions with other labs appear to be weak. NCEP model development does not appear to benefit from GFDL's more active model development efforts as much as it might. GFDL's relative lack of observational science and process modeling research suggests the potential for beneficial interactions with ESRL, PMEL, and AOML.

12. The panel recognizes the creation of the GFDL Science Board and the re-organization of its Research Council to ensure that many parties' voices are heard in the development of lab strategic priorities. It is not clear how well this process works in practice. GFDL should explore ways to increase communication between leadership and the rank-and-file. This includes prompt communication and discussion forums about decisions, and a means for younger scientists to have input into lab strategic planning, possibly via representation on the Science Board.

13. GFDL has made some improvement since the previous review toward the goal of gender balance but still lags behind other labs. This is especially true at the level of leadership and may not be strictly a demographics issue. The panel recommends a reconsideration of hiring and promotion practices, training about possible unconscious biases, and a concerted effort to engage its young female leaders to accelerate progress in this area.

Comments on the Review Process

The panel generally found it difficult to work within the format of the review worksheet. The definition of "Outstanding" in the Review Panel Guide ("outstanding in all areas") is an unrealistic expectation for any organization when applied to broad areas that encompass many different aspects of research. Taken literally, no lab should ever receive such a ranking. On the other hand, "Satisfactory" seemed to some panel members an unduly lukewarm assessment that might have unintended negative consequences in an area in which the lab is actually doing a good job. It is unclear what "curve" OAR uses to interpret such grades. For example, science proposals to the panel chair's agency are reviewed on a 5-tier rating system for which the highest three grades ("Excellent," "Very Good," and "Good") allow for more shades of evaluation than the panel had available to them, yet experienced reviewers know that a proposal that ranks only "Very Good" – seemingly a strong positive assessment – has little chance of being funded. With nothing in between "Outstanding" and "Satisfactory," the panel was presented with something close to a Hobson's choice. Consequently some panelists chose instead to give "Outstanding/Satisfactory" or "Satisfactory +" rankings, while others did not but expressed a similar sentiment. Hibbard, for example, notes that being between outstanding and satisfactory is where "all groups should strive to be. There is always room for improvement." The intended interpretation of the panel's rankings is thus that GFDL was not judged to be sub-standard in any area. "Outstanding" is differentiated from "Satisfactory" according to whether a given reviewer feels that GFDL is a leader in the field vs. being comparable to a typical modeling center.

Some panel members gave only a single overall rating for a given research area, finding it difficult to separate Quality, Relevance, and Performance. (Can the quality of the models be judged without evaluating their success in simulating the modes of variability that NOAA deems relevant? Must all GFDL research hew to a Strategic Plan written years earlier to be high-performing, or is there room for new ideas?) Note also that the review materials were unclear on this point, with the Review Panel Guide requesting separate ratings for the three themes for each research area, while the Review Evaluation Worksheet only offers a single overall rating for each research area. Some panelists likewise found it difficult to separate the second or third research areas from the first. Climate variability research at GFDL depends on its Earth System Models and is likewise one way in which these models are evaluated. Chemistry, carbon, and ecosystem science at GFDL depend on its Earth System models, yet these areas are what differentiate an "Earth system model" from a "climate model."

There was broad agreement among panel members that the review itself was very well-organized and we appreciated the responsiveness of GFDL staff to inquiries and the pre-review telecon information. Several panel members felt that the review agenda itself might have been organized somewhat differently, in the following ways: (1) Fewer presentations overall, and presentation of not only high-level issues but scientific challenges for the future; (2) Less emphasis on demonstrating how good GFDL models are and more on the fundamental science that is being done with these models, including a summary of GFDL's most important and interesting science findings in the past few years; (3) A greater opportunity to engage the panel in scientific discussion; (4) More time with staff and students to enable a better understanding of the "pulse" of GFDL; (5) Private discussion time with the Science Board and Research Council to better understand how it operates and to discuss strategic science issues. (6) Some discussion of education and outreach activities, web presence, and GFDL's overall relationship to the outside world beyond just its formal stakeholders, e.g., plans for community use of its models.