Proposal Panel 1 : 1939547

Back to Proposal

Agency Name: National Science Foundation

Agency Tracking Number: **1939547**

Panel Summary

Panel Summary

This center proposes an ambitious and exciting program in which physics informed machine learning techniques are applied across cryosphere, ocean, atmosphere, and biosphere components of an earth system model. The goal is to target three sources of model uncertainty, physical parameterizations, parameter estimates, and internal stochasticity.

Intellectual Merit:

This was a very well written and exciting proposal with the potential to reduce computational costs and model uncertainty across a range of ESM model components. The panel appreciated the analysis of challenges and felt that working in an existing ESM structure was a benefit.

The panel would like to see more defined timeline. The panel was uncertain about whether ocean biology was a component of the ocean module, and felt this was an area ripe for improvement.

The panel also needed more clarity about how equations might be extracted from the machine learning approaches.

Broader Impacts:

The panel felt that success in any of the modules would be a benefit to society. The broader impacts portfolio was well balanced and aligned with center goals. It was appreciated that the center director would also be director of diversity.

Solicitation Specific Review Criteria:

The proposed work has the potential to be transformative not only for understanding climate, but also potentially in shorter time horizon forecasting. The project is at an excellent scale for an STC.

Rationale for Recommendation:

The timing is excellent for transformative advances in this field and the scale of the center is appropriate to an STC.

Panel Recommendation:
Invite

This summary was read by/to the panel and the panel concurred that the summary accurately reflects the panel discussion.

PANEL RECOMMENDATION:

Proposal All Reviews: 1939547

Back to Proposal

Agency Name:	National Science Foundation
Agency Tracking Number:	**1939547**
Organization:	
NSF Program:	STC Integrative Partnrshps Adm
PI/PD:	Gentine, Pierre
Application Title:	STC: Center for Learning the Earth with Artificial Intelligence and Physics (LEAP)

Review 1

Rating:

Excellent

Review:

Summary

In the context of the five review elements, please
evaluate the strengths and weaknesses of the proposal with respect to intellectual merit.

OVERVIEW & INTELLECTUAL MERIT

The central goal of the proposed STC (Learning the Earth with Artificial Intelligence and Physics, LEAP) is to develop more robust climate prediction models with a focus on reducing existing structural errors, better levering the wealth of observational data that can feed into models, and quantifying the variability, stochasticity and uncertainty of predictive models. The PIs propose to merge physical modeling with the power of machine learning to produce their Community Earth System Model.

Although this is quite far afield from my expertise, I found this to be a well-written and compelling proposal. The PIs nicely summarize the major agents of uncertainty in climate prediction models, including uncertainty in emissions, the large range in equilibrium climate sensitivity (which has not changed since 1979!), and highly imprecise regional climate predictions. The PIs argue that much of our inability to increase predictive accuracy is because of errors caused physical and biological processes operating on scales smaller than the model's grid resolution. Such scale-dependent problems are widely appreciated whether they be in climate change science or basic ecological work.

The PIs propose to target and reduce the impact of three categories of uncertainty in Earth System models: structural errors (which result from imprecision in approximating complex physical and biological processes), model parameter errors, and internal stochasticity. They will do so by leveraging expertise in geoscience, machine learning and advanced computer architecture. Harnessing the power of machine learning is central to their approach and they argue that machine learning can correct structural errors and optimize parameters and, in so doing, facilitate upscaling from fine to coarse scales as well enable extrapolation, which has been challenging for previous efforts employing machine learning. The PIs argue that by integrating physical and biological knowledge with machine learning such extrapolation is possible.

By harvesting multiple data sources to inform their machine learning approach, a strength of this proposal is the breadth of disciplinary expertise and knowledge coverage afforded by the main research thrusts, which focus atmosphere, ocean, biosphere, and the cryosphere to provide a synthetic approach to understanding factors driving climate variability and our capacity to predict it. These thrusts are complemented by four cross-cutting themes: (1) coupling subcomponents of earth system models, (2) machine learning generalization, (3) quantification of uncertainty, and (4) interpretable learning. I found this second theme particularly interesting in its goal to address non-stationarity in climate statistics by injecting physical constraints, accounting for distributional and compositional dynamics into their machine learning models. Also interesting was acknowledgement of the fact that even models informed by observation can have high uncertainty because of, for example, sampling errors.

Although I am not an expert in this areas, I was excited by what the PIs have proposed, which is a compliment to the interesting and well-crafted narrative of this proposal. The work is certainly ambitious and my impression is that is has high potential to produce compelling and transformative insights.

An additional strength of the proposal is that includes an analysis of their competitors. Apparently only the group at Caltech is playing in this space and their approach does not account for structural errors and stochasticity nor does their approach leverage existing community model infrastructure. I was also pleased to see the PIs discuss anticipated bottlenecks associated with the capacity of new hybrid models to generalize and extrapolate and achieving numerical stability. They will address the first bottleneck by integrating physical knowledge and constraints into their machine learning algorithms. To address numerical instability, the PIs will develop diagnostic tools that can be used to reject machine learning models that produce instability. I leave it to the experts to weigh in on the effectiveness of this approach.

One weakness concerns the narrative describing how 'biology' will be incorporated into the models. I found this section somewhat vague and was concerned about the terrestrial bias. Terrestrially-based data from multiple sensing networks will be used to 'optimally define ecosystem functions'. I am not sure what that means and attention to ocean-based processes seems to be a glaring omission. I understand that the PIs will leverage new data sets on the biological removal of carbon to the deep ocean but surely there are other ocean-based ecosystem functions that may be relevant? I realize that one cannot do everything but I recommend the PIs elaborate more on how they will handle processes or ecosystem functions in the terrestrial AND oceanic biosphere.

In the context of the five review elements, please
evaluate the strengths and weaknesses of the proposal with respect to broader impacts.

I thought the broader impacts of this proposal were quite strong. The PIs leverage several approaches to enhance impacts on education (high school thru postdoctoral students) and diversity. These include partnering with the Eagle Academy for Young Men of Harlem (high school students), supporting extensive cross-fertilization between Columbia and NCAR in Boulder (postdoctoral students), and Columbia's Bridge to PhD program (post bac students). Unlike many proposals that are vague in describing the metrics of their education impact, a strength of the proposal was its specific target of supporting 6 postdocs, 10 graduate students, 6 Bridge students, and 12 high school interns. With respect to diversity and inclusion, the above efforts will clearly impact students from URMs. I would have like to have seen some attention to the issue of enhancing the participation of women in science. Other broader impacts include the Climate and Society MA program (facilitated via partnership with Columbia's Program in Climate and Society) and the Center for Science and Society (also at Columbia) which facilitates research and outreach between natural and social scientists, engineers and policy makers. A key feature of this latter effort is the Public Conversations in Climate Prediction, which seeks to engage the public on important climate science issues. The involvement of graduate students in operating this series is a strength. Finally, the Center will hold an annual symposium showcasing the work of Center researchers.

Please evaluate the strengths and
weaknesses of the proposal with respect to any additional solicitation-specific review criteria, if
applicable

Evaluate and comment on the following:

Rationale for an STC, including questions: Is the vision for the project compelling enough to justify the large-scale focus of resources? If so, is an STC the appropriate vehicle? Why is an STC investment warranted at this time? Are the anticipated scientific and societal legacies substantive and transformative?

This was an excellent proposal. It is very ambitious and addresses a key need, which is refining the robustness of climate prediction models. I think the transformative potential of the work is high and is certainly appropriate for an STC.

Research plan, including questions: Are the plans for research and holistic integration appropriately ambitious, leading to significant strategic outcomes? Does the proposal address potential bottlenecks and technical challenges? Does the proposal make a case for the feasibility of significant progress over the next five years?

The research plan is well conceived and described and will likely lead to important advancements. The PIs thoughtfully address bottlenecks and technical challenges and I was convinced that they will be able to overcome them.

Partnerships and Participants, including questions: Is the team of partner organizations and personnel assembled for the proposed Center appropriate and essential? Is the role of each participant clear? Does the team have unique strengths relative to the other groups working in related fields?

This is unquestionably an excellent team will the roles of each participant clearly described. The team has high potential to produce strong results both within and across disciplines. A strength of their approach is that each research thrust is co-led by a geoscientist and a data scientist.

Integration Strategies, including questions: Are the leadership and management strategies promoting a center culture and are the foci of education, knowledge transfer and broadening participation strategically supporting the goals of the center?

The PIs provide an excellent description of the leadership team and their responsibilities for different dimensions of the Center. I was very pleased to see that Center Director Gentice will also serve as the Director of Diversity. This approach, as the PIs note, signals that diversity efforts are taken seriously be the leadership team. The multi-pronged approach to Center oversight and management (2 Advisory Committees, an Executive Committee, a Research Subcommittee, and a Industry Subcommittee) was also a strength.

Summary Statement

For many of the reasons outlined above, I thought this was a excellent proposal and I strongly support the team moving beyond the preliminary proposal stage.

Overall Rating: Excellent

Review 2

Rating:

Excellent

Review:

Summary

In the context of the five review elements, please
evaluate the strengths and weaknesses of the proposal with respect to intellectual merit.

Responses are included in the solicitation specific comments.

In the context of the five review elements, please
evaluate the strengths and weaknesses of the proposal with respect to broader impacts.

Responses are included in the solicitation specific comments.

Please evaluate the strengths and
weaknesses of the proposal with respect to any additional solicitation-specific review criteria, if
applicable

The overarching goal of the proposed center is to greatly improve climate predictions by combing physical models grounded in data with machine learning algorithms through merging and leveraging the best aspects of both approaches. Despite the exponential growth of new, high-resolution, multi-petabyte geophysical data sets from remote sensing and in situ observations, the physical models do not take full advantage of the knowledge from these data sets, but when coupled with machine learning algorithms, they could. The merging of approaches could provide more accurate climate forecasts 20-40 years into the future, allowing decision makers opportunities to help mitigate or facilitate social adaptation to a changing climate. The proposers make a strong case that recent advances in both machine learning and in observational science in the atmosphere, on land, in the ocean and

on/in ice make it timely to propose a major effort to improve climate forecasts. Their main focus is improving climate forecasts and they bring a wide diversity of expertise to do so as an integrated earth science system approach. The breadth of the expertise coupled with the laser beam focus on improved climate models makes this worthy of a center. Their legacy would be a significant improvement in climate forecasting but also in providing a blueprint for future machine learning hybrid approaches.

This proposal makes excellent use of their 12 pages, clearly stating how they will carry out the research, providing much detail. In their physical model approach, they include physics, chemistry and biology and domains of atmosphere, ocean, cryosphere, and biosphere. They correctly address the scale issue, and will focus on the subgrid processes (e.g., submesoscale in the ocean). The organization of the Research Thrusts and Cross Cutting Themes is clearly and logically laid out, and the current road blocks to problems identified. They acknowledge assimilation of data for long term predictions, in contrast to short term forecasts; the path to be taken must improve estimation of the physical model parameters and quantify uncertainty, not as an afterthought, but as a goal. They are cognizant of cost of running models, and look to improved architecture to reduce cost to allow an increased number of simulations. The approach outlined should lead to significant progress û they seem poised to hit the ground running.

The leadership is very experienced and some have worked together. The expertise of the team is clearly and succinctly communicated. Regarding the Research Thrusts and Cross Cutting Themes, it is clear whom about the participants is responsible for what, and who is doing what. The high degree of professionalism in the proposal preparation indicates this team is organized.

The PI is experienced and will be involved in the engagement of underrepresented youth. There will be advisory committees, workshops, etc. There is no description of the anticipated frequency of meetings within the team (and this should be elaborated on in a full proposal), but each subteam leadership includes a geoscientist and a data expert. The proposed course program should enforce good interaction.

The Broader Impacts are clearly outline in terms of benefits to society; advancement of knowledge in earth system models, machine learning, and hybridization; training of a prepared workforce including underrepresented groups. They will include an evaluation expert, who will train the trainees in evaluation criteria. The educational components are well aligned with the goals of the center.

Summary Statement

The proposed center addresses a deficiency in present climate forecast models, highlights advances in observations that could improve the quality of the models as well as advances in machine learning, and proposes combining earth science models with machine learning to reduce uncertainty and improve model parameter. The center has high potential to make significant advances based on the quality of the team and the clear elucidation of the proposed research. The Broader Impacts are at societal and workforce development, and are well aligned with center goals.

Review 3

Rating:

Excellent

Review:

Summary

In the context of the five review elements, please
evaluate the strengths and weaknesses of the proposal with respect to intellectual merit.

This is primarily a Columbia Univ proposal led by Prof. Pierre Gentine (Earth and Env Engineering). This is a large proposal: 36 proposal personnel, of which 10 are directors or co-directors of the Center.

Strength:
This is a hugely ambitious, exciting, transformative proposal. The science and technological merits are self-evident, and will not even bother to repeat the many amazing good things that this proposal can offer to the climate modeling community (and the ML community). Additionally, the timing of this proposal topic (merging big climate model issues with machine learning) could not be better as ML has become the 'buzz word' of the day and is penetrating fast in more and more fields. Proposal is structurally well written, well thought out. The Center Rationale and Center plan are absolutely beautiful. The team is top of the line and is uniquely positioned to try this ambitious project out.

Weeknesses (What to work on further/at the next step):

The proposal is rather vague about (a) the timescales needed for the work proposed, and (b) the precise steps to be taken and their order. Here are

some questions for the authors that I would like to see addressed:
- What will be done first and what will be done second/third, etc.? A flow chart/plan/table would help here.
- How long will it take for each of the steps proposed? How much is feasible right now versus in the future? For example, how long will the integration with ML-based modules with modern GPU and TPU architecture take and how will you proceed?

Also, ML algorithms by definition 'teach themselves' what to do by incorporating real data. This input of data will increase gradually as our observational and satellite datasets grow, meaning that the models can be run and rerun and will get better in time. Because of this, the order in which the various project parts can be done needs to be carefully thought out, e.g. do you proceed with existing obs to infer much needed diffusivity coefficients in the ocean, or do you wait for this step until later on when we will have more Argo float data available, etc.? There are big technical challenges here at every step of the way.

- How much of the mathematics for the ML is known, and how much is not known? The proposal as written seems to suggest that more is known/done than might be the case. My intuition is that in reality much of the first 5 years will be used to build the practical machinery to make ML pertinent to a rather narrow set of first order issues in climate models (and for engineers, computer scientists and climate scientists to be able to even talk to each other) and that the broad scope of the Center will be far from being reached until the next 5 year. Even so, this is stuff that needs to get done and this team seems uniquely positioned to do it!

In the context of the five review elements, please
evaluate the strengths and weaknesses of the proposal with respect to broader impacts.

Strength: great, all good here.
Connection with industry is well thought out. This I imagine can result in potential future funding possibilities. I don't think this center, if given initial funding, will have problems financially to take off the ground.

Weeknesses: none that I can see.

Please evaluate the strengths and
weaknesses of the proposal with respect to any additional solicitation-specific review criteria, if
applicable

Summary Statement

This is an ambitious, exciting, transformative proposal that is forward thinking.
While there are many issues left to be refined and thought through, the breadth of expertise in the Columbia team (and the large size of this impressive team) makes this challenging technical proposal feasible. This proposal in my view absolutely merits funding.

Review 4

Rating:

Excellent

Review:

Summary

In the context of the five review elements, please
evaluate the strengths and weaknesses of the proposal with respect to intellectual merit.

LEAP seeks to take advantage of recent advances in machine learning to develop a whole new class of physical parameterizations for GCMs. There has been a lot of discussion in recent years about the potential uses of machine learning algorithms to fundamentally change how we parameterize GCMs, and the PIs are very well positioned to push this forward. The thing I really like about this proposal is that the lead PI has already demonstrated the potential for these new techniques and has assembled a team to rapidly roll out these new developments across an entire Earth-system modeling framework. This is one of the rare projects that really does explore original, creative, and potentially transformative concepts with rather obvious societal benefits. There is some risk in this project - the use of machine learning techniques in GCM parameterizations is very much in its infancy, and we may well find out in a few years that they don't really offer substantial benefits over traditional physics-based approaches, but the preliminary results look

promising, and I think it's worth investing in this ambitious program. The PIs list some of the key bottlenecks, including numerical stability. It's an honest assessment that I don't think detracts from the merit of the project. I also appreciate the competitive analysis - this project has some similarities to Caltech's CLIMA initiative, but as the PIs indicate, LEAP is focused on using an existing community model while CLIMA is being built from scratch. Both approaches are fine, but it's good that the PIs have identified a complementary approach.

In the context of the five review elements, please
evaluate the strengths and weaknesses of the proposal with respect to broader impacts.

I found the broader impacts educational and diversity program to be quite strong, with components ranging from high school through the postdoctoral level, and including a new MA program in climate & society that will aim to translate some of the new data-driven approaches into a broader policy oriented audience. It's a thoughtful and well-balanced portfolio of educational and outreach activities.

Please evaluate the strengths and
weaknesses of the proposal with respect to any additional solicitation-specific review criteria, if
applicable

While the PIs don't explicitly lay out what they will aim to have accomplished within five years, I think they have the potential to make significant progress in rolling out these new approaches across most of the proposed Earth system modeling components.

Summary Statement

While many groups claim they are ready to make rapid progress in applying new machine learning techniques to GCM parameterizations, the group assembled within LEAP may actually be able to follow through on the claim. This is a credible, well-organized team with potential to make transformative contributions in a new and important area.