Next-Generation Ecosystem Experiments (NGEE Arctic)

Scaling White Paper – Modeling
November 23, 2011

Developing a Hierarchical Scaling Framework for Modeling a Dynamic Arctic Landscape in a Changing Climate

Peter Thornton, Bill Riley, Charles Koven, Forrest Hoffman, Richard Mills, and the NGEE Arctic Science Team

1. Introduction

Our goal for NGEE Arctic is to reduce uncertainty in climate prediction through improved representation of Arctic tundra processes. Our objective with this research program is to achieve a generalization of knowledge and understanding, gained through direct observation and fine-grained simulation of Arctic tundra ecosystems and the mechanisms which regulate their form and function. More specifically, this generalization of knowledge will take the form of improved representation of Arctic tundra states and dynamics in the land model component of a global coupled Earth system model. We begin this process of generalization in Phase 1 of the project and envision that it will be an important and continuing activity throughout Phase 2. Our ultimate deliverable from field, lab and modeling activities will be the development of a process-rich ecosystem model, extending from bedrock to the top of the vegetative canopy, in which the evolution of Arctic ecosystems in a changing climate can be modeled at the scale of a high resolution Earth System Model (ESM) grid cell (i.e., approximately 30x30 km grid size).

One of the most difficult challenges we face is the problem of how to optimally inform process representations in a global-scale model with knowledge and understanding gained through direct observation and process-resolving simulation at smaller scales – we call this the *up-scaling* problem. Of similar importance and just as daunting is the problem of how to provide appropriate large-scale context to guide strategies for direct observations and fine-scale simulation, allowing interpretation of results which can be meaningful at larger scales – what we refer to as the *down-scaling* problem. These two problems are clearly interrelated: appropriate large-scale context provided to guide measurement and process-resolving simulation is fruitless if no mechanism is in place to migrate new fine-scale knowledge to larger scales, while the up-scaled information itself is likely irrelevant if not conditioned in advance by the large-scale context. It is necessary, then, to solve the up-scaling and down-scaling problems together, referred to in tandem as the scaling problem.

To maintain focus in this whitepaper on the Arctic tundra scaling problem and on our intended approach to solving it, we assume here some familiarity with details of Arctic landscape geomorphology, hydrology, vegetation dynamics, and biogeochemistry. Background details on these topics are included in the NGEE Arctic proposal, and will be expanded on as needed in the revised proposal. With concurrence from NGEE Program Management at BER, we also focus this whitepaper primarily on the scaling problem as it relates to up-scale and down-scale migration of process knowledge in a hierarchical modeling framework. We recognize that there are important scaling issues related to measurement methods and sampling strategies, and that these issues are closely related to the structure of the hierarchical modeling framework and to the up-scale and down-scale flows of information within it. For clarity, this document treats the modeling dimensions of the scaling problem comprehensively, while drawing only representative connections to the observational issues. Upon receiving acknowledgement from BER Program Management that the scaling approach described here meets expectations, we will proceed to draw the comprehensive connections between multi-scale measurements and the modeling hierarchy as a major element of the revised proposal.

In the following sections we provide a brief overview of current understanding of Arctic landscape processes relevant to the scaling problem and then outline a scaling philosophy which is consistent with this knowledge (Section 2). We summarize the scaling approach currently employed in the land component of the Community Earth System Model (Section 3) as a demonstration that the community has some relevant experience in solving the scaling problem, and as a way to highlight the deficiencies of the existing approach for process scaling in the Arctic tundra landscape. We next define the physical basis for our proposed scaling approach, which relies on a representation of the landscape as geomorphologically distinct landunits connected by surface drainage networks and subsurface flow paths (Section 4). We describe the nested hierarchical modeling framework proposed to enable NGEE Arctic up-scaling and down-scaling (Section 5), followed by a comprehensive description of how model parameterization information will be derived and passed between scales (Section 6). Additional technical details are provided as an Appendix.

2. Background and Overview of Scaling Approach

Our fundamental scaling approach is to identify processes likely to have the largest influence on climate, based on current knowledge of the Arctic tundra system, and then to define a connected (nested) hierarchy of modeling scales necessary to resolve those processes. This approach allows us to begin immediately to integrate new process knowledge into a climate prediction-scale land model, while establishing a quantitative framework connecting this scale to more process-rich models implemented at finer spatial resolution and over smaller spatial domains.

The scaling problem for NGEE Arctic is bounded at large spatial scales by the need to represent the global pan-Arctic land mass and its interactions with the atmosphere, oceans, and sea ice in coupled Earth system climate prediction simulations. The significance of new process knowledge for global coupled climate prediction depends on the area (and time span) over which the process is relevant, in other words its *representativeness*. Simply stated, the larger the region and longer the duration of influence for a process or phenomenon, the greater its potential impact on the coupled global system. An important challenge for NGEE is to assess the Arctic tundra environment at large spatial scales and over climate-relevant time spans, producing metrics of representativeness, impact, and uncertainty that can direct observation and process-resolving simulation to the most relevant regions in a vast and remote landscape.

A distinguishing characteristic of Arctic tundra landscapes is the existence of recognizable landscape units which are repeated over large domains, and which occur at multiple spatial scales. Previous landscape-scale classification efforts have identified active and drained thaw lakes and polygonal ground as two common landscape units that occur over large parts of the Arctic tundra. Individual thaw lakes range in size from tens of meters to a few kilometers in diameter, and occur in concentrations of one to several lakes per 10 km^2 over hundreds of thousands of square kilometers of Arctic tundra. Individual ice wedge polygons are typically 10-20 meters in diameter and occur in multiple hydrologically connected polygon networks. Polygonal ground commonly co-occurs with thaw lakes, and both are understood as manifestations of permafrost dynamics in low-gradient landscapes. Since our primary focus is on improved climate prediction, we are most concerned with processes that influence the large-scale exchanges of energy between the land surface and the atmosphere, and with processes which affect net fluxes of long-lived greenhouse gases, especially CO_2 and CH_4. As detailed in the NGEE proposal, we expect that the interaction of a warming atmosphere with the interconnected dynamics of thaw lakes and polygonal ground will lead, through the development of thermokarst, to fine-scale changes in surface elevation and hydrologic flow paths, which will be expressed at larger spatial scales as significant change in mean hydrologic state. We expect changing hydrology to lead to shifts in biogeochemical cycles with significant consequences for greenhouse gas flux, and to changes in vegetation dynamics and community structure with significant consequences for surface energy balance and that snowpack dynamics.

Our scaling approach builds on the hypothesis that the transfer of information across spatial scales can be organized around discrete geomorphological units for which processes are represented explicitly at finer scales, with information passed up to coarser scales through subgrid parameterization. By extending an already well-established framework for fractional subgrid area representations to allow dynamic subgrid areas and hydrological and geophysical connections among subgrid units, we expect to be able to characterize permafrost dynamics and the influence of thermokarst at multiple spatial scales in Arctic tundra landscapes. We further hypothesize that hydrologic connectedness controls both structure and dynamics of the Arctic tundra, and our scaling approach accommodates two-way (up-scale and down-scale) hydrologic interactions. One consequence of two-way hydrologic interactions in tundra landscapes is that cross-scale iterative solutions are required to arrive at optimal coarse-scale parameterizations. Our approach implements this iterative solution as a progressive refinement, allowing us to make immediate progress toward process integration at the climate prediction scale before the more advanced process-resolving model scales are fully constructed and tested.

Our assessment of critical Arctic tundra process controls on climate provides strong prioritization of observations and process studies, emphasizing the measurements required to define geomorphological scaling units and to develop model process parameterizations. Different observation methods are possible at different spatial and temporal scales, and the limits of our ability to implement particular scaling strategies depend in part on what system properties can be adequately measured at each scale. We have assessed the current state of the art in measurement technology to set reasonable boundaries for our scaling approach.

Another critical constraint on our scaling approach is that, whatever parameterizations are implemented at each scale, they cannot violate the conservation equations implicit in the numerical solutions. In other words, energy, mass, and momentum must be conserved in the solutions at each scale. One consequence of this requirement is that explicit information on landscape states (e.g. soil temperature or soil ice concentration) and fluxes (e.g. hydrologic flow) can be passed down-scale as boundary condition forcing from coarse-scale to fine-scale models in the hierarchy, but information passed up-scale is limited to diagnostic (non-state) parameters.

More details about these constraints and some worked-through examples are provided in later sections.

In addition to the problem of scaling in space, as framed here, there are different but related problems associated with scaling in the time domain. The modeling hierarchy we propose here has the advantage of using a common (hourly to sub-hourly) time resolution for all model scales, so the granularity aspect of temporal scaling can be ignored. On the other hand, the time domains of interest for climate prediction extend out to centuries, and we *may* find that computational demands make century-scale simulations impractical for our finest resolution process-resolving model components. The up-scaling approach described here can be implemented using temporal subsets of the climate prediction time domain, for example 10-20 year simulations for the finest-resolution model can provide relevant parameterization information for the coarser-scale models. The down-scaling approach is applicable at any point in the coarse-scale model time domain, allowing decadal-scale high-resolution simulations to be initiated at multiple points along the century-scale trajectory.

As a final background note before proceeding to a more detailed explanation of our scaling approach, we acknowledge that the approach described here is one of many possible solutions to the model scaling problem for Arctic tundra systems. This is a new area of inquiry, and we should not ignore the possibility that other, perhaps better, approaches could be proposed or will emerge over time. We believe that our approach is scientifically defensible as well as tractable, but we are anxious to engage a broader research community, bringing our experience and learning from others', and we expect to launch exploratory studies of promising new approaches when and if they emerge.

3. Current Scaling Framework in CLM4

The current land model component of CESM, the Community Land Model v4 (CLM4), already includes a sophisticated spatial scaling framework, the most advanced of any land model component in the current generation of Earth system models. Some explanation of that structure and its intended application is helpful here, since we propose to apply major aspects of it in our more advanced scaling approach for NGEE.

The CLM grid cell is the geographically referenced unit in the model – that is, it has a known geographic center, and fixed geographic extent. Grid cells can vary somewhat in size over the entire simulation grid, but the sizes and areas for each grid cell are fixed for the duration of a simulation. Each grid cell is composed of multiple subgrid fractional areas. These subgrid units are not explicitly geographically referenced: they have a known area, and so a known fractional area representation within a grid cell, but the model has no explicit information about what part of the grid cell each subgrid unit occupies. In this sense the subgrid units are considered to be statistical representations of the subgrid, as opposed to explicit representations.

CLM4's subgrid information is derived from spatial datasets having (generally) a higher spatial resolution than the final model grid resolution. Subgrid fractional units can therefore be prescribed to represent geo-referenced subgrid variability. For example a map of vegetation types (plant functional types, or PFTs) is one of the input layers used in defining a CLM4 grid and its subgrid units, and this map is based on 1 km^2 resolution remote sensing data. Given a

CLM4 grid cell of, say, 0.5° resolution, explicit geographic information regarding the subgrid distribution of PFTs falling within the grid cell is converted to a statistical representation of the subgrid area represented by each PFT. Some state variables in a CLM4 simulation are carried along at the level of the individual PFTs, and so in theory one could extract information from the CLM4 output and apply it to the original geo-referenced subgrid PFT map to generate an output at higher spatial resolution than the CLM4 regular grid cells.

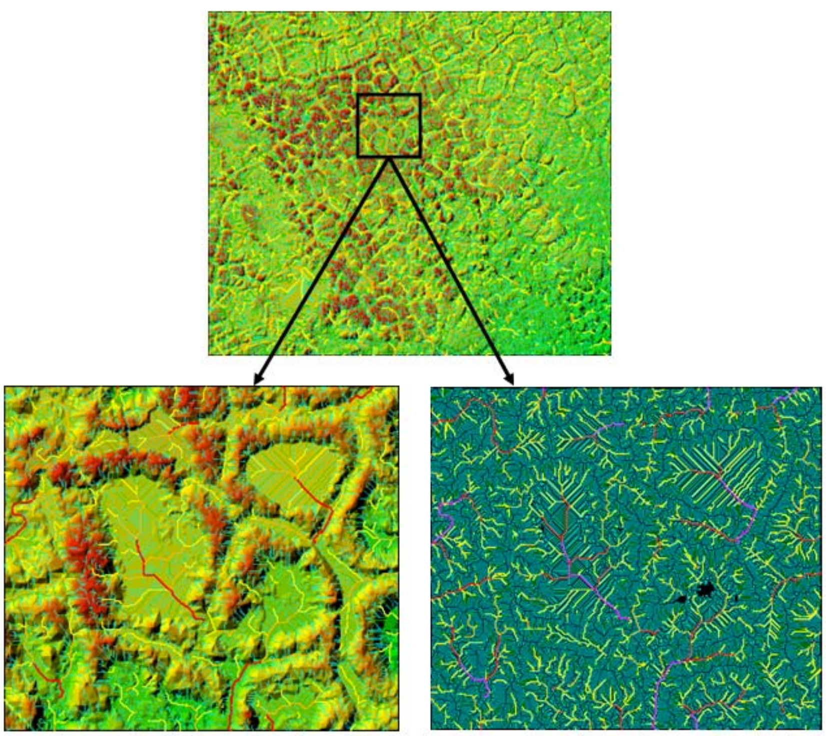
The concept of subgrid units occupying fractional area on a grid cell is actually implemented with one more level of complexity in CLM4, by representing the gridcell as a nested hierarchy of three subgrid types. The first type below the grid cell is the *landunit*, and its intended purpose is to represent subgrid variability that presents itself as geomorphologically distinct regions. For example the current CLM4 subgrid uses the landunit level of the hierarchy to represent the differences between lakes, crops, natural vegetation, wetlands, glaciers, and urban areas. Each grid cell can be composed of one or more landunits, the areas of which completely occupy the grid cell. Each landunit is composed of one or more soil *columns*. The purpose of the column is to represent the state variables and conservation equations for energy, water, carbon, and nitrogen within a multi-layer soil, including the potential for multiple layers of overlying snow, and also including a mean representation of state variables for any vegetation existing on the column. Each column is composed of one or more (usually several) PFTs, each of which has a defined fractional area on the column. The purpose of the PFT level of the hierarchy is to represent the water, carbon, and nitrogen state of physiologically distinct vegetation types and their interactions with each other and with the column state variables such as soil water and temperature distributions.

A final aspect of the existing CLM4 architecture is relevant here: a mechanism is already developed in the model to allow mass and energy conserving changes in subgrid fractional areas represented at the PFT level, and with a few caveats also at the soil/snow column and landunit levels. With a little additional development work to eliminate these caveats at the higher levels in the hierarchy, our NGEE Arctic scaling approach will be able to put the existing scaling architecture in CLM4 to productive use. It is reasonable to say that our scaling approach begins to realize the full potential of the CLM4 nested hierarchy, putting the complexity of its nested subgrid hierarchy into action and exercising it for its intended purpose.

4. NGEE Scaling Framework

A fundamental aspect of our scaling approach is that information derived from high-resolution process-resolving simulation at small spatial scales must have a pathway for up-scale migration in order to properly inform the behavior of a larger-scale and more coarsely-resolved model. To accomplish this we start with a conceptualization at the climate-prediction scale, our coarsest scale, which represents, for each climate model grid cell, multiple distinct subgrid elements based on their geomorphology. These subgrid elements correspond exactly with the CLM4 landunit scaling elements, but we will add a new layer of information for the NGEE scaling approach, maintaining not only subgrid fractional area information, but also a description of the subgrid topology in terms of a surface hydrologic network connecting multiple subgrid units.

CLM4 currently includes a surface hydrologic routing network (the River Transport Model, RTM), but we require a much more sophisticated representation. RTM does not include information on how subgrid areas are connected with a drainage network, nor how they are connected with each other. Based on our hypothesis that warming will promote thermokarst, and that thermokarst will lead to fundamental changes in the hydrologic organization of low-gradient Arctic tundra landscapes, we consider it crucial that our scaling approach accommodate the transfer of information regarding these hydrological reorganizations from the fine scale at which the governing processes are resolved up to the coarser scale where interactions with the climate system can be realized. We will extend the surface hydrologic network information content in CLM4 to include finely-resolved delineations of drainage networks and their associated catchments within individual CLM grid cells. This delineation depends on finely-resolved digital surface elevation maps and automated drainage network and catchment delineation algorithms. We have some example surface elevation datasets for Arctic tundra landscapes, derived from airborne LIDAR measurements, and we have performed tests using existing network and catchment delineation approaches to evaluate the ability of these datasets and tools to rapidly generate meaningful subgrid hydrologic connectivity information. Initial results are encouraging (Figure 1), and we intend to pursue this approach with more extensive LIDAR retrievals and expert interpretation of the results of automated network delineation outputs as an immediate and high priority effort at the start of NGEE Phase 1. We are especially encouraged to see that it is possible to estimate connectedness among individual ice wedge polygons in this very low-gradient landscape, given a high-quality and high-resolution LIDAR elevation dataset.



Example output from automated drainage network and catchment delineation software (RiverTools), applied near Barrow, AK

Figure 1. Automated drainage delineation in a low-gradient landscape. From Hinzman et al., unpublished data.

The use of explicitly resolved subgrid hydrologic connectivity is a critical element of our scaling framework which makes possible the two-way iterative scaling approach (sequential up-scaling and down-scaling) which we hypothesize will lead to improved prediction skill at the climate-modeling scale. Our starting assumption is that at the scale of a high-resolution climate modeling grid cell, say at 30x30 or 10x10 km resolution, it will be possible to delineate subgrid catchments and drainage networks which can be considered fixed on century timescales, as constrained by the large-scale topographic gradients. We note that evidence from the ground suggests that even in these very low-gradient systems, there is enough large scale topographic structure to organize the landscape into catchments which appear to persist on century time scales. Those catchments will form the CLM subgrid landunits. Within those catchments, we further suppose that more finely resolved landscape units such as individual thaw lakes, individual ice wedge polygons, or polygon sub-units such as rims and troughs, can be resolved from LIDAR topography. In contrast with the larger-scale catchments and higher-order network elements, these fine-scale catchments and low order networks are expected to have dynamic topologies under a warming climate, with thermokarst affecting surface elevations, leading to reorganized flow networks. It is precisely this sort of sub-grid reorganization which will be represented explicitly at the finely-resolved modeling scales but implicitly, or through statistical parameterization, at coarser scales. Up-scaling for this property therefore consists of the definition of a suitable parametric expression of the consequences of subgrid flow organization in the coarse-scale model, the explicit representation of this same process in the fine-scale model, and the summarization of model output from fine-scale simulations to optimize parameters in the coarse scale model that allow it to represent the behavior (variance) of the fine-scale process with a quantifiable level of statistical completeness.

Arriving at the appropriate functional form for the parameterized process in the coarse-scale model, and connecting it with explanatory variables available as prognostic outputs or imposed boundary conditions in the coarse-scale model is a topic for focused new model development in the climate-scale model. Fortunately, this new development effort can begin in the absence of parametric input from finer scales, and the new coarse-resolution model can even be exercised by making informed guesses for initial parameter values. The consequence for NGEE implementation is that necessary development and model application efforts can begin immediately and in parallel for multiple modeling scales. This approach avoids the pitfalls of having to wait for completion of fine scale models before commencing development of application of coarser scale models, and vice versa. Of course we can only take full advantage of this approach if there are sufficient human resources deployed to model development and application at each of the necessary scales.

So far this discussion has focused on the up-scaling problem, but organizing the scaling framework around explicit subgrid hydrologic connectivity has consequences also for the down-scaling part of the overall scaling problem. By placing each fine-scale modeling subunit in its proper context within a drainage and catchment network defined at the coarse-scale, it is possible to pass explicit boundary condition information from the coarse scale model to the fine scale in the form of boundary fluxes and states. For example, at the points in the climate-scale drainage network where one catchment landunit communicates with another through surface flow, explicit flow volume, temperature, and biogeochemistry information can be passed from the coarse scale model directly to the boundary condition of a fine-scale model, establishing the large-scale hydrologic and perhaps also energetic and biogeochemistry influence on the processes resolved explicitly within that hydrologic unit by the fine-scale model. Like the up-scaling connections described above, this approach does not require the existence of coarser-scale models to develop and execute finer-scale models. Instead of imposing boundary conditions from the up-scale model, those boundary conditions can instead be estimated from climatologies, as initial guesses.

Before moving on to a more complete technical description of our scaling approach, we summarize the basic steps for the simple case of two models at different scales: a climate-prediction scale model, and a higher resolution process-resolving model (Figure 2). Initial model development is required for both modeling scales to accommodate the information passing and parameterization steps required by the scaling framework. As models are being developed in parallel, early model testing can begin without communication between them. A first scaling step is the up-scale coupling of information from the fine-scale model to inform parameterizations (set parameter values) in the climate-scale model. A second scaling step is to run the climate-scale model with new parameters and pass boundary constraints down-scale to the process-resolving model. This leads to subsequent simulations at fine and coarse scales, in an iterative process through which final parameter estimates for the coarse scale model are expected to converge to their optimal values.

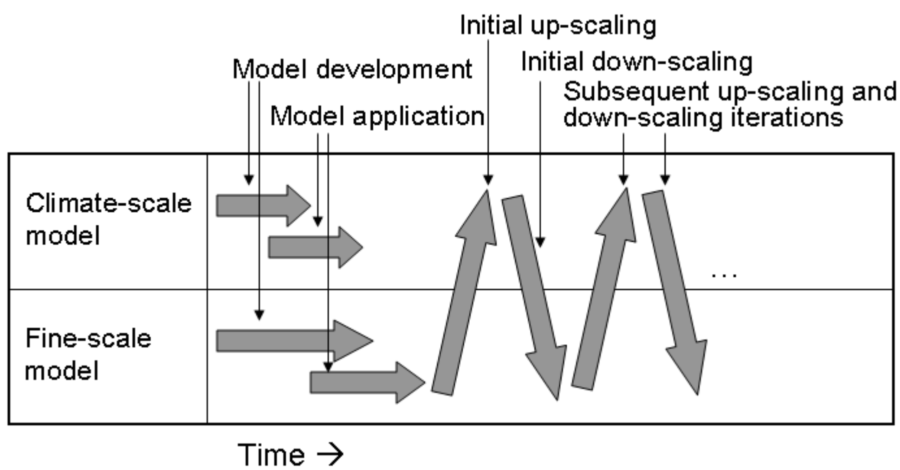


Figure 2. Simplified representation of NGEE scaling approach: two-model case

5. Nested Hierarchy of Models for NGEE Scaling

We did not come to the NGEE scaling problem with a pre-conceived notion of the optimal arrangement of models in a scaling framework, or of the optimal number of models or their domain sizes or resolutions. Our one constraint in this regard has been that our final target is an improved treatment of Arctic tundra landscape processes at the scale of a high-resolution climate prediction model grid cell. As noted earlier, that constraint sets the upper boundary condition for the scaling framework as a land model component that accurately represents tundra landscape processes that most significantly influence climate, that is fully operational within the coupling framework of and Earth system model, and that has a grid resolution on the order of 10x10 to 50x50 km.

In consultation with a broad community of experts on the structure and function of Arctic tundra landscapes, we have concluded that there are two very distinct scales of organization within these landscapes which emerge at resolutions finer than a nominal climate grid resolution of 30x30 km. Moving down in scale from the climate grid cell, the first level of landscape organization which emerges quite distinctly in the Arctic coastal plain is the occurrence of thaw lakes, ranging in size from tens of meters to several kilometers, and, as noted earlier, distributed rather densely over large parts of the pan-Arctic. Our rationale for believing that lakes and ponds have an important impact on pan-Arctic and global climate is detailed elsewhere (in the NGEE proposal). For the sake of our scaling approach it is sufficient to say that these landscape elements occur at scales that are subgrid even for the highest resolution climate prediction simulations. Representing the dynamics of these lakes and their interactions with the surrounding landscape elements requires a higher-resolution model, with grid cell resolution on the order of 100 m, with a simulation domain on the same scale as a single climate-prediction grid cell, say 30x30 km. This scale and resolution permit the explicit representation of hydrologically-connected networks of several to tens of thaw lakes. Figure 3 shows two examples of potential modeling domains at this scale for real landscapes. We refer to this as the intermediate modeling scale.

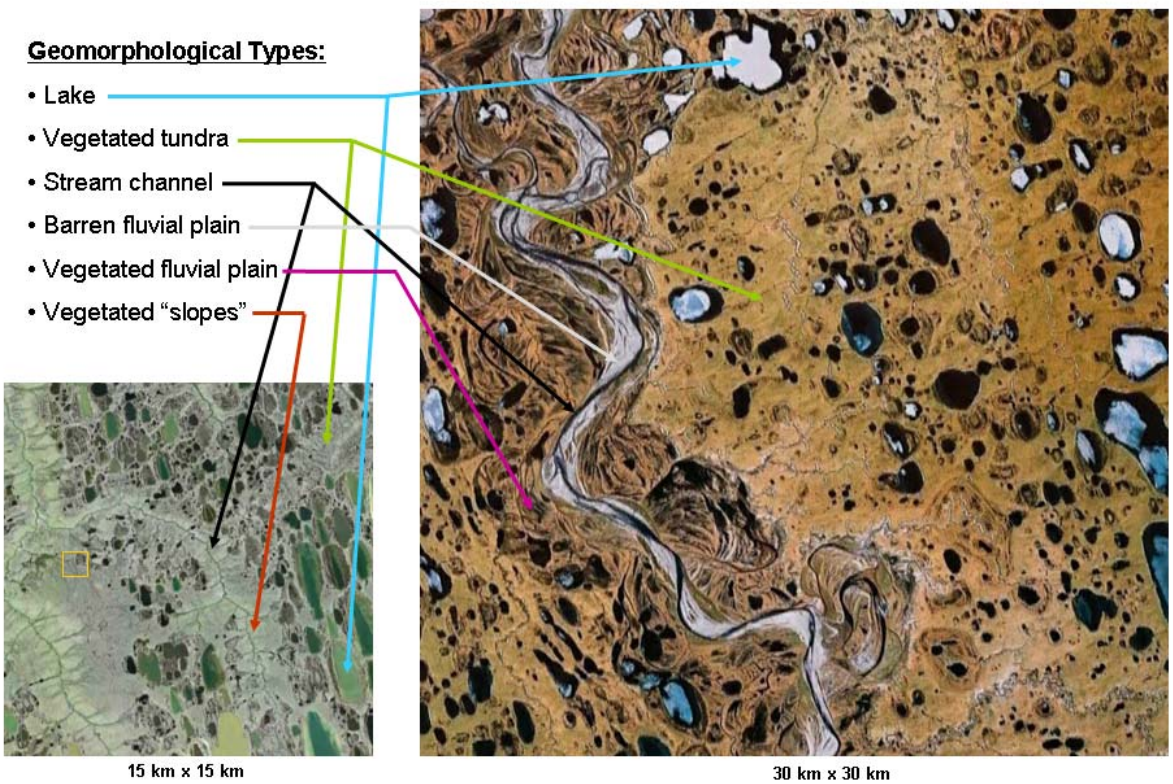


Figure 3. Example intermediate-scale model domains from two different regions on the North Slope of Alaska. Several dominant and recurring geomorphological features are highlighted. The yellow rectangle in the 15 km x 15 km domain marks the location of fine-scale modeling domains shown in Figure 4.

Another level of organization emerges in these landscapes at an even finer scale. The land in between thaw lakes, as well as the remnant beds of drained lakes, is seen very commonly to be composed of polygonal ground, structured by the presence and dynamics of massive wedges of high ice-content soil in the subsurface. The characteristic length scale for these polygons is around 10-20 m, or one to two orders of magnitude smaller than the typical thaw lake. Nonetheless these structures appear to dominate the local hydrologic environment, with follow-on effects for energy balance, biogeochemical dynamics, and vegetation community structure and function. Even more significant for our purposes is the fact that these structures are sensitive to changing thermal conditions, as the microtopography generated by the ice wedges can shift as ground ice accumulates or melts, driving dramatic changes in local hydrology which, when replicated over thousands of similar polygons in extensive networks, is expected to have important effects on hydrologic conditions at larger scales. We therefore consider it essential to capture the dynamics of polygonal ground, especially in response to warming, through explicit simulation at a third and finer modeling scale. Grid or finite element resolution for this scale needs to be on the order of a meter or less in order to explicitly represent individual ice wedges. Domain sizes would need to be on the order of 100-200 m to describe the behavior of multiple self-similar polygons organized in networks, and to describe the interactions of polygon networks with adjacent thaw lakes. Figure 4 shows two examples of potential modeling domains at this scale, nested within the landscape illustrated on the left side of Figure 3. We refer to this as the fine modeling scale.

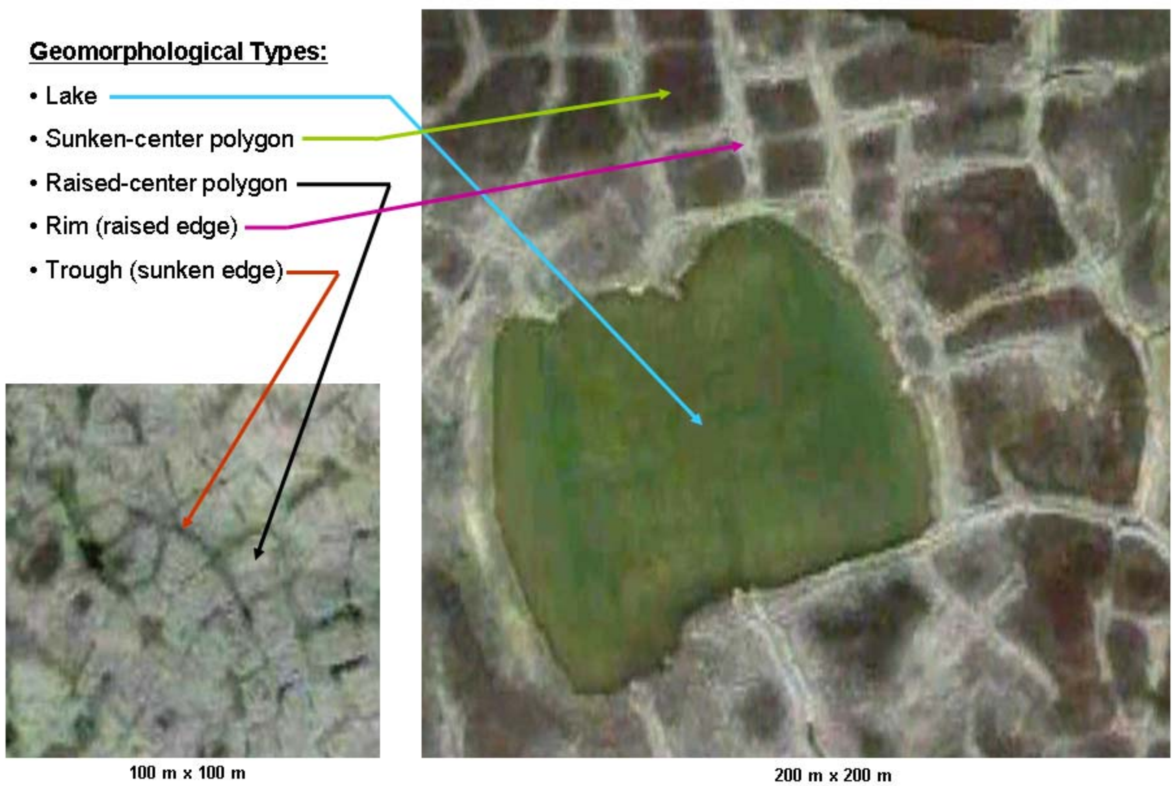


Figure 4. Example fine-scale model domains from a region on the North Slope of Alaska. Several dominant and recurring geomorphological features are highlighted.

There is evidence for an even finer scale of variation within individual ice wedge polygons, in the form of sorted and unsorted circles which can have important consequences for soil vertical structure and vegetation dynamics. These structures appear to have characteristic length scales on the order of a meter, and so it seems possible though perhaps not likely that we would be able to represent these dynamics explicitly in our fine scale model, since this is the same characteristic width of the ice wedges which we do intend to represent explicitly. Seeing no compelling reason to extend our scaling framework any further, we arrive at a three-level nested hierarchical structure: fine-scale model nested within intermediate-scale, and intermediate scale model nested within climate-scale. Both the nature and the extent of model nesting are crucial elements in our scaling framework. In the remainder of this section we describe the nature of how model nesting is accomplished, saving a discussion of nesting extent for the next section.

Our NGEE scaling framework is founded on a simple concept that distinguishes between models at different scales on the basis of which processes are treated explicitly, which processes are treated as subgrid parameterizations, and the model's boundary condition requirements. The ability to perform up-scaling depends on the explicit representation of some process at a finer scale which is treated implicitly, or as a subgrid parameterization, at a coarser scale. The ability to perform useful down-scaling depends on the existence of boundary conditions in a finer-scale model which can be assigned using outputs from a coarser-scale model.

Our climate-scale model will be derived from the current CLM4, and will share all of CLM4's explicit process representations. These include one-dimensional (column-based) mass and energy balance, including permafrost development and active layer dynamics. New development will add thermokarst-driven changes in surface elevation as an explicit one-dimensional process. Mass balance equations include explicit carbon and nitrogen biogeochemistry, and new plant functional types representing real Arctic vegetation will be introduced. The model will include an explicit representation of surface flow in a static drainage network, replacing RTM with a more sophisticated approach that maintains hydrologic connectivity information at the subgrid scale (described earlier). Subgrid parameterized processes added to the model to accommodate the NGEE scaling approach will include overland flow across subgrid units and transfer to the static drainage network, as well as surface and subsurface flow between landunits, including dissolved-phase biogeochemistry and sediment transport. The distribution and dynamics of PFTs are handled as subgrid parameterizations in the current CLM4, and this functionality will be retained. Shifts in subgrid area from one type to another (e.g. polygonal ground to thaw lake, or thaw lake to dry lake bed) will be added as a new subgrid process. Boundary conditions for the climate-scale model include near-surface weather and the imposed structure of the static drainage network and large-scale topography.

Our intermediate-scale model will be based on an existing three-dimensional reactive transport model architecture (either PFLOTRAN or ASCEM/Amanzi), using three-dimensional finite volume computational methods for subsurface calculations, and two-dimensional surface mesh representations for overland flow. All processes treated explicitly in the climate-scale model are also treated explicitly in the intermediate-scale model. Additional explicit processes include thaw lake dynamics, dynamic drainage network organization, and surface and subsurface lateral flow. Processes represented through subgrid parameterization include polygonal ground and ice wedge dynamics, as well as PFT distributions. Boundary conditions for the intermediate-scale model include near surface weather, and surface and subsurface water and energy inflow at domain boundaries. Regardless of the architecture selected, a substantial amount of new model development and integration of existing model components is required to arrive at a fully functional intermediate-scale model. Further details on model technical requirements and computational components are provided in the full NGEE proposal.

Our fine-scale model will use the same computational framweork as the intermediate-scale model, but will be implemented with a smaller domain size and finer horizontal resolution. All the processes treated explicitly for the intermediate-scale model are also treated explicitly in the fine-scale model. Additional explicit representations will include ice wedge polygons and dynamic microtopography. Boundary conditions are formulated in the same way as for the intermediate-scale model. We expect that most of the same process representations employed in the intermediate-scale model can simply be resolved over a finer grid, and initialized with more detailed boundary conditions, to deliver explicit representation of the polygonal ground dynamics, and the interactions among polygonal ground and thaw lakes. In other words we expect that the major distinction between the intermediate and fine-scale models will be in their resolution and initialization data, not in their fundamental representation of physical and biological processes. Finally, we expect that even for the fine-scale model we will not be able to resolve the explicit dynamics of individual plants as they grow, reproduce, and die, and so PFT dynamics will continue to be represented as a subgrid parameterized process at this scale.

6. Model Nesting and Parameterization Across Scales

Given the nature of model nesting described above, we now consider the extent of nesting across the model scales. To help constrain the discussion, we consider the application of our climate-scale model at a resolution of 30km x 30km over a subset of the pan-Arctic tundra, focusing on low-gradient tundra systems of the Alaskan North Slope. The entire region consists of 50-100 climate-scale grid cells at this resolution (Figure 5).

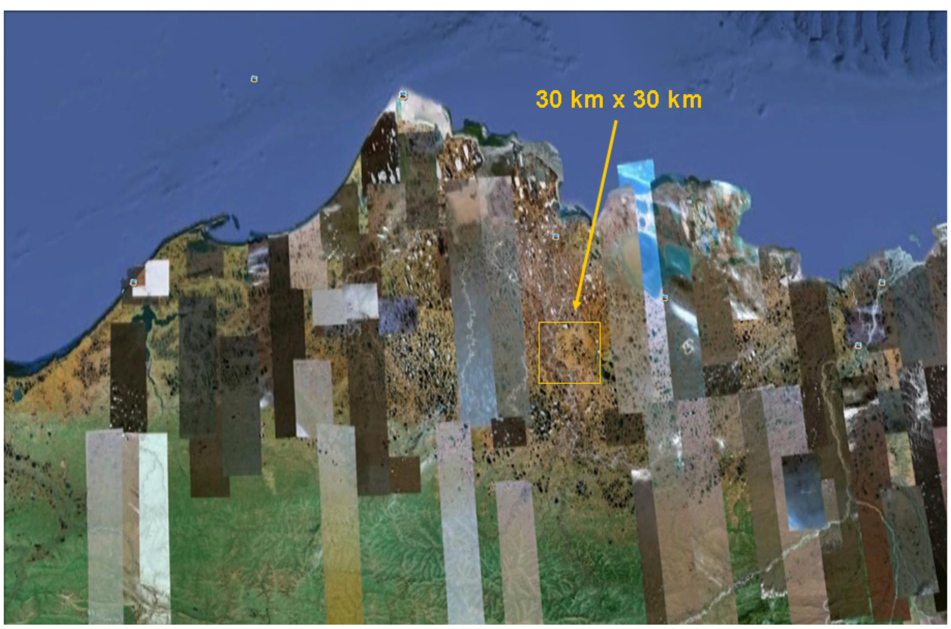


Figure 5. A remote sensing mosaic image of the North Slope of Alaska near Barrow, illustrating the scale of a single climate model grid cell.

An important part of the scaling problem is the characterization of subgrid variance across this large spatial domain. Given a suitable quantification of the type and frequency distributions of subgrid geomorphological units across this complex landscape, it would be possible to select an appropriate number of sub-domains over which an intermediate-scale model would be implemented, for the purpose of generating regionally relevant up-scaling statistics. This is an ambitious goal, but fortunately we find that significant progress in exactly this direction has already been made by several Arctic tundra geomorphology research teams. An example is shown in Figure 6.

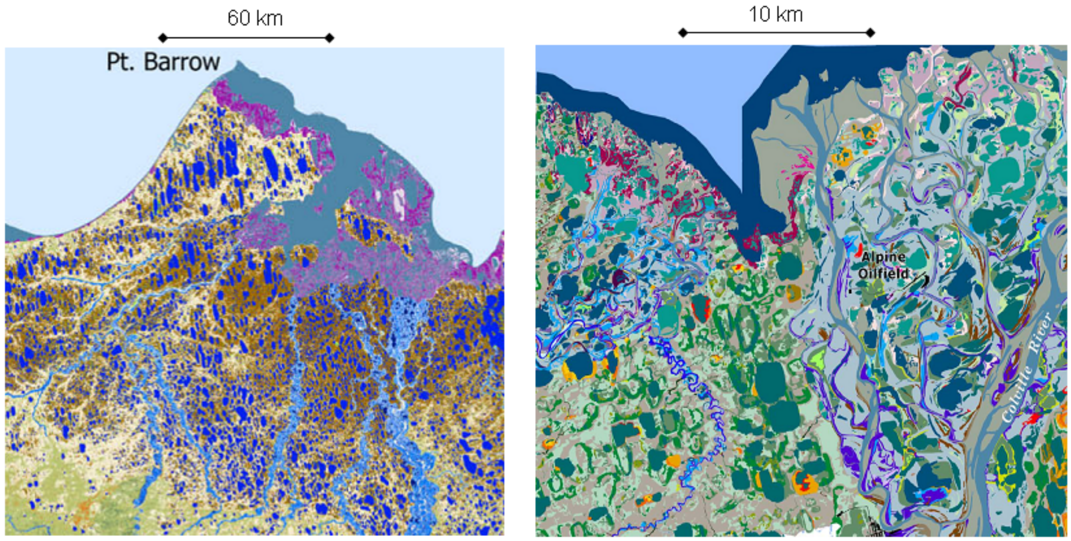


Figure 6. Subsets from two recent remote sensing based efforts to map geomorphological units across the Alaskan North Slope tundra region. Left: from Jorgensen and Heiner, 2004. Right: from Jorgensen et al. 2005.

Related efforts by our own research team indicate that similar approaches can be used to characterize landscape variance in terms of repeated geomorphological units at very fine spatial scales (Figure 1). We intend to implement these approaches as an initial effort in NGEE Phase 1, and to base our decisions regarding the extent of model nesting on the quantitative findings from these early investigations.

For the sake of clarity in the rest of this discussion, we assume that the decisions regarding extent of model nesting have already been made, and we refer to transactions between model scales as aggregated over all finer-scale nested domains in the case of up-scaled parameterizations, or as propagated to all finer-scale nested domains in the case of down-scaled boundary conditions.

The process by which we will arrive at progressively refined parameterizations of Arctic tundra processes, and progressively refined climate predictions, is captured graphically in Figure 7 and described sequentially below. Note that Figure 7 is an expanded version of Figure 2, with the third modeling scale introduced, and with a much more detailed identification of the individual steps required to achieve up-scaling and down-scaling.

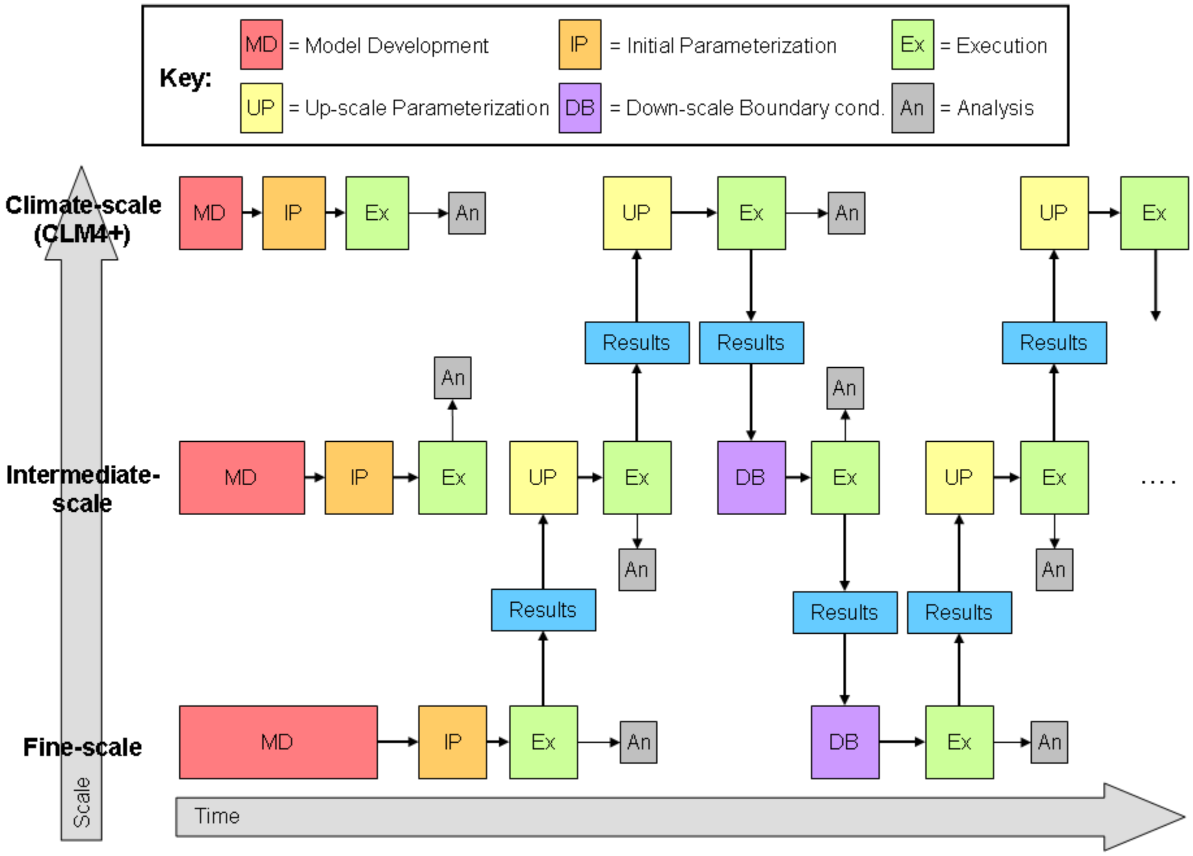


Figure 7. NGEE scaling approach: model parameterization flowchart. See text for explanation.

<u>Model parameterization steps:</u>

1. Based on data for selected domain, define subgrid representations at each modeling scale. Calculate drainage networks and catchments, and distribution of subgrid connectivity and subgrid fractional cover for multiple geomorphological units within each subgrid catchment.
2. Best-guess initial values are assigned to all subgrid parameters at each scale.
3. Simulations are performed at each spatial scale, using best-guess initial parameters.
4. Fine scale simulation results are aggregated to produce parameter values (response surfaces) for subgrid processes in the intermediate scale model.
5. New intermediate-scale simulations are performed, using improved parameters from fine scale results.
6. Intermediate scale simulation results are aggregated to produce parameter values (response surfaces) for subgrid processes in CLM4+.
7. New climate-scale simulations are performed, using improved parameters from intermediate and fine-scale results.
8. Analysis of climate-scale model output: comparison to initial simulations and evaluation against observations.
9. Results of new climate-scale simulations are used to provide improved boundary conditions for intermediate-scale model. New intermediate-scale model execution.
10. Results of new intermediate-scale simulations are used to provide improved boundary conditions for fine-scale model. New fine-scale model execution.

11. Analysis of 3D model output: comparison to initial simulations and evaluation against observations. Repeat from step 4 if solutions have not converged.

A more thorough understanding of our proposed scaling approach requires additional details regarding which processes are parameterized at each modeling scale, how those parameterizations are linked with the nested subgrid scaling hierarchy within CLM4 (or CLM4+, given the NGEE-specific modifications), and how those parameterizations depend on external data sources. We direct readers interested in this level of detail to the Appendix, where the required information is organized in outline form to aid comprehension.

7. Conclusion

The purpose of this whitepaper has been to provide a comprehensive assessment of how our research team will tackle the very challenging problem of cross-scale migration of knowledge to inform a process-rich representation of Arctic tundra landscape dynamics and to improve climate prediction. We would like to stress two points in conclusion. First, we recognize that there are many alternative scaling approaches which might be tried, and we do not expect that our team will adhere strictly to the approach described here. Rather, this description of our approach serves as our best preliminary map of largely uncharted terrain. Second, many science details have been omitted here which will be present in the full NGEE proposal, placing the scaling approach in a much richer context of process observations, laboratory manipulation, and modeling investigations. These omissions are unavoidable, and we hope they do not detract too much from the scaling description.

Appendix: Parameterization Details at each Modeling Scale

The top-level of the following outline refers to each model scale, with sublevels describing the parameterization details at each scale. A final section provides additional details for the up-scale parameterization method.

1. Climate model scale (CLM4+). Fundamental purpose of model at this scale is to represent Arctic tundra processes in climate change prediction.
    a. Existing sub-grid hierarchy used to represent catchments that are fixed on century time scales (CLM4 landunits), as well as fractional area representations of dynamic subgrid components or geomorphological units within each catchment (CLM4 columns). Each column can contain one to many vegetation types (CLM4 PFTs).
    b. Subgrid resolution DEM for pan-Arctic used to define subgrid drainage network and subgrid catchments.
    c. Subgrid catchments are explicitly referenced to the subgrid drainage network, providing a basis for parameterization of subgrid surface and subsurface flow dynamics within each subgrid catchment. The drainage network spans multiple CLM gridcells, providing a large-scale hydrologic integration capability (similar in concept to the existing CLM River Transport Model, but more advanced in implementation).
    d. Geomorphological units within each catchment are characterized on the basis of automated clustering approach (see Section 6), guided by expert knowledge of critical types. Likely categories at this scale include: lake, vegetated flat tundra, vegetated sloping tundra, barren fluvial plain, vegetated fluvial plain, stream channel.
    e. Each column within a catchment landunit is populated by one to several PFTs, with initial specification of PFT fractional cover. Each column is subject to the regular CLM integration of 1D mass and energy balance equations, including vegetation processes and subsurface biogeochemistry. We add the calculation of a 1D dynamic surface elevation to represent thermokarst, providing a critical endogenous link for parameterization from finer scales.
    f. The following parameters link the climate-scale to finer scales, and represent new functionality that must be added to CLM4+ to accomplish the NGEE Arctic modeling and model scaling objectives:
        i. Hydrologic flow among columns.
        ii. Hydrologic flow between columns and drainage network.
        iii. Hydrologic flow between columns and common local groundwater (catchment-based).
        iv. Change in PFT areas within a column
        v. Change in column area within a landunit
    g. Functional forms for these parameterizations are linked to endogenous variables within the CLM4+ simulation, with parameterization constants derived initially from expert judgment and later refined through up-scale parameterization.
2. Intermediate-scale, 3D process-resolving model. Fundamental purpose of model at this scale is to resolve the dynamics of fluvial network, including thaw lake dynamics, and to bridge fine-scale and climate-scale model

a. Explicit 3D (vertical and horizontal) flow representation, with prognostic surface elevation, ponding, surface flow connectivity. Each gridcell also includes all the 1D explicit processes resolved at the climate-scale.
b. High-resolution DEM used to characterize drainage network surface elevations. For at least some applications of the intermediate-scale model, this DEM should be resolved at a finer resolution than the intermediate model, e.g. a sub-meter resolution DEM derived from LIDAR. This allows explicit connection of the intermediate and fine-scale parameterizations.
c. Subgrid elements (geomorphological units) within each intermediate-scale gridcell are characterized on the basis of automated clustering approach (see Section 6), guided by expert knowledge of critical types. Likely categories at this scale include: lake, remnant lakebed, polygonal ground, barren fluvial plain, vegetated fluvial plain, stream channel.
d. Each subgrid geomorphological unit within a gridcell is populated by one to several PFTs, with initial specification of PFT fractional cover.
e. The following parameters link the intermediate-scale to finer and coarser scales:
    i. Hydrologic flow among geomorphological units.
    ii. Hydrologic flow between units and drainage network.
    iii. Hydrologic flow between units and common local groundwater (catchment-based).
    iv. Change in PFT areas within a unit
    v. Change in unit area within a gridcell
f. Functional forms for these parameterizations are linked to endogenous variables within the intermediate-scale simulation, with parameterization constants derived initially from expert judgment and later refined through up-scale parameterization.

3. <u>Fine-scale model (3D process resolving)</u>. Fundamental purpose of model at this scale is to resolve ice wedge dynamics and the influence of thermokarst in reorganization of mass and energy fluxes in polygonal ground.
    a. Explicit 3D (vertical and horizontal) flow representation, with prognostic surface elevation, ponding, surface flow connectivity. Each gridcell also includes all the 1D explicit processes resolved at the climate-scale and intermediate modeling scales.
    b. High-resolution DEM used to characterize drainage network surface elevations. This DEM should be resolved at the resolution of the fine-scale model, e.g. a sub-meter resolution DEM derived from LIDAR. This allows explicit connection of the intermediate and fine-scale parameterizations.
    c. At this scale the resolution of all the major landscape elements will be explicit, and will not require subgrid parameterization. For example, the polygonal ground network will be explicitly resolved as interconnected gridcells representing rims, troughs, and polygon centers.
    d. There may still need to be some subgrid fractional representation of PFT mixtures in the fine resolution model. For example, it is likely that mosses and tundra grasses will be mixed even at the sub-meter resolution. This will require development of automated clustering techniques to diagnose covariation between geomorphological unit and pft types.
4. <u>Up-scale parameterization method</u>

a. In the previous sections we mention the aggregation of finer-resolution model results to arrive at coarser-resolution model parameterizations. This section provides a more detailed discussion of that approach.
b. This approach depends on having an explicit (prognostic) simulation of a system process or state at a finer scale which is represented as a subgrid parameterization (implicitly) at a coarser scale.
c. An example could be the representation of thaw lake size (surface area). This will be represented as an explicit prognostic variable in the intermediate-scale 3D model, but is represented as a subgrid fractional area parameterization in the climate-scale model. By initializing the intermediate-scale model for a region (or subregion) coinciding with a specific climate-scale gridcell, we ensure direct comparability of results from the finer resolution model with parameterizations in the coarser resolution model.
d. Prognostic variable or variables in the finer resolution model which correspond directly with the parameterized quantity in the coarser resolution model are aggregated over each geomorphological unit identified as a subgrid fractional area component of the coarser-scale model, using simple averaging as a first approach, but possibly employing higher order statistics.
e. Multi-variate optimization approach is used to derive parameters at the coarser model scale which return best-fit estimates for the subgrid elements when compared to the explicit results from finer scale. As noted above, we will use endogenous coarse model variables as inputs to the parameterized estimation equations.