EDITOR: Mike Potel, potel@wildcrest.com

DEPARTMENT: APPLICATIONS

Antarctic Water Masses and Ice Shelves: Visualizing the Physics

Greg Abram, *Texas Advanced Computing Center, Austin, TX 78758 USA*
Mark R. Petersen, *Los Alamos National Laboratory, Los Alamos, NM 87545 USA*
Francesca Samsel, *Texas Advanced Computing Center, Austin, TX 78758 USA*
Xylar Asay-Davis, *Los Alamos National Laboratory, Los Alamos, NM 87545 USA*
Darin Comeau, *Los Alamos National Laboratory, Los Alamos, NM 87545 USA*
Stephen F. Price, *Los Alamos National Laboratory, Los Alamos, NM 87545 USA*

High-resolution simulation of global climate physics enables us to model how the climate may change under a variety of future scenarios. Such simulations produce vast amounts of information and dense datasets. If interrogated in tandem, these datasets can provide holistic, vital information on Earth's many integrated systems by revealing the manifold interrelated properties of the atmosphere, ocean, and polar ice, framed by real-world terrain in three-dimensional space as they vary over time. To accomplish this, climate scientists have joined with computer scientists and an artist to develop techniques enabling scientists to see these relationships. The impact of ocean water properties on Antarctic ice shelves illustrates the benefit of this analysis in understanding land ice melt rates and thus sea-level rise.

As rapid climate change begins to affect the Earth's polar regions, scientists are studying the intricacies of ocean, atmosphere, and polar ice to understand how increasing temperatures will affect sea-level rise and coastal communities in the coming century. One area of particular interest for a team at Los Alamos National Laboratory (LANL) is Antarctica's Filchner-Ronne Ice Shelf (Figure 1). Antarctica, a continent larger than the USA and Mexico combined, is covered by an average 2-km-thick ice sheet that, if completely melted, would contribute 60 m of sea-level rise. That would take thousands of years, but 1–2 m of potential sea-level rise in this century depends on the dynamics of floating ice shelves like the Filchner-Ronne. Ice shelves are the product of accumulated snowfall pushing off the continent and contribute to sea-level rise with accelerated melting (in contrast, sea ice freezes and melts seasonally at the ocean's surface and does not affect sea level). Scientists believe that warming ocean waters beneath the ice shelves will increase their melt rates, which may in turn speed up the ice streams that feed the shelves.

As polar currents circulate, they push *water masses*—a classification oceanographers use to separate ocean water into specific ranges of temperature and salinity—beneath the ice shelves, where they become difficult to track. Scientists must rely on models to understand the complex dynamics at play, as real-world observations only come through robotic submersibles or probes drilled through the ice. This creates a difficult visualization challenge in terms of providing clarity, detail, and data interrogation capability in a complex, three-dimensional space. In order to accomplish this, we gathered an interdisciplinary team of expert ocean modelers, computer scientists, visualization professionals, and an artist, who tackled these modeling and design challenges in a two-step process: a data reduction phase, in which the specific requirements of the science drive the transformation of the raw data to higher level representations, and a design phase, in which the use of artistic color theory introduces clarity within the complex multivariate imagery.

Digital Object Identifier 10.1109/MCG.2020.3044228
Date of current version 12 January 2021.

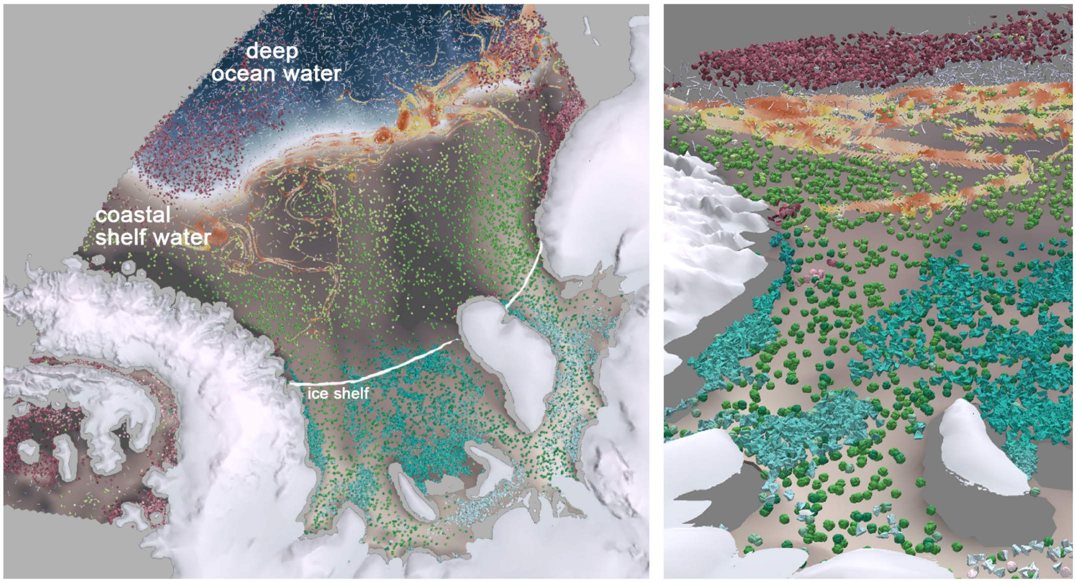

FIGURE 1. This region of Antarctica, south of the Atlantic, contains the Filchner-Ronne Ice Shelf, which extends to the labeled white line. Properties of the water under the ice shelf are represented by the water masses described in the text and Figure 4. The ocean floor is colored by temperature; dark blues indicate the cold deep ocean waters. The yellow–orange streamlines show the currents over the edge of the coastal shelf. On the right is an oblique detail of the region showing the three-dimensional distribution of the water masses.

OCEAN MODELING

Energy Exascale Earth System Model (E3SM)

Climate modeling is a data-rich endeavor. Modern climate simulations cover the Earth with millions of grid cells, run for thousands of simulated years, and can track hundreds of variables. A grand challenge of data science is the analysis and interpretation of such large and varied data sets in order to validate models, test scientific hypotheses, and make projections.

Our ocean simulations used the Model for Prediction Across Scales-Ocean (MPAS-Ocean),[1] developed at the LANL (https://github.com/MPAS-Dev/MPAS-Model). MPAS-Ocean is a component of our E3SM (https://e3sm.org), recently developed by the U.S. Department of Energy (DOE).[2,3] E3SM also includes model components of sea ice, land ice with ice shelves, atmosphere, and land. E3SM is unique in that all components are based on variable-resolution horizontal meshes,[4] so that compute time may be spent on particular areas of interest. It was created to investigate topics at the intersection of climate science and national security. This includes the changing Arctic, coastal flooding, frequencies of storms and droughts, and the causes and effects of sea-level rise.

The ocean, sea ice, and land ice components are modeled using the MPAS framework, where the horizontal meshes are Voronoi Tessellations, and the framework provides unified tools for multicore partitioning, massively parallel I/O, mesh generation, and visualization. Here, we describe the first in a series of collaborative efforts, the visualization of the three-dimensional water masses and their movements under the Filchner-Ronne Ice Sheet.[5] The global ocean and sea ice simulations contain 1.4 million horizontal grid cells that vary in width from 10 to 30 km, and there are 100 layers in the vertical direction. Atmospheric winds, precipitation, and radiation are provided by historical data from 1948 to 2009.[2] Simulations were run on Theta at Argonne National Laboratory, a Cray XC40, 11.7 petaflop system based on the second-generation Intel Xeon Phi processor. Each simulation used 12,292 cores, with a throughput of 2.0 simulated years per wallclock day.

Analysis

In modeling the global ocean, our MPAS team at the U.S. DOE takes a twofold approach to data analysis. First, we run a standard set of automated diagnostics on every simulation, which produces hundreds of plots for our perusal. Second, we explore data interactively with visualization tools and write new analyses to

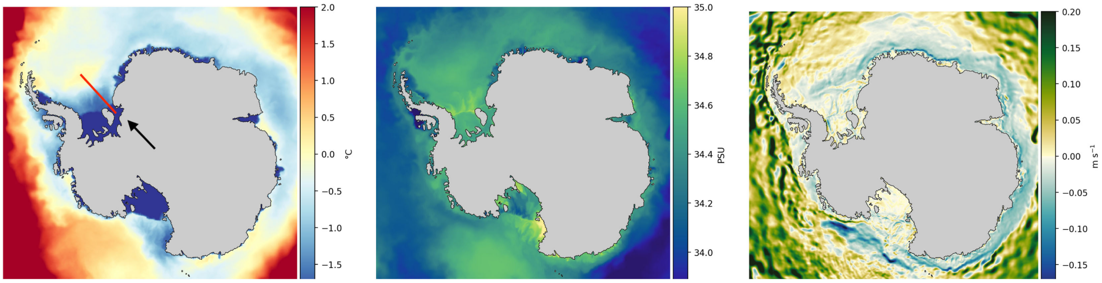

FIGURE 2. Standard analysis output for the ocean surface temperature (left), salinity (middle), and eastward velocity (right), for summer months, averaged over years 21–30 of the simulation. The red line on the temperature plot, left, identifies the location of the slices in Figure 3.

answer particular questions. Both of these methods of standard and exploratory analysis are vital to our evaluation of large data sets—the first quickly provides a broad overview of the simulation, whereas the latter allows for more detailed, customized investigations. Over time, the exploratory analysis that gets used most frequently is formalized and "hardened" to become part of the standard analysis. This is part of the ongoing process of software development we use for scientific data analysis tools.

Standard Analysis

In the typical simulation process, a set of standard analyses are applied to the output data after each simulated decade. MPAS-Analysis (https://mpas-dev.github.io/MPAS-Analysis) is a Python package that averages monthly output over several decades; interpolates to standard grids; computes secondary statistics and differences with observational data; and automatically generates hundreds of plots. Examples of these are shown in Figures 2 and 3. MPAS-Analysis includes horizontal and vertical sections of the prognostic variables (ocean temperature and salinity, velocity, sea ice extent and thickness), heat content, surface fluxes, velocity stream functions, melt rates, and a host of others. Experts in physical oceanography browse these plots to validate the results. A common method of scientific inquiry in climate modeling is to run multiple simulations where only certain parameters are varied. MPAS-Analysis supports this process with direct comparison using difference plots of all analyses.

Water Masses and Ice Shelf Interactions

Tracking water masses helps oceanographers to study how ocean properties evolve over time as they interact with other parts of the climate. This includes studying the impacts of changing ocean currents and their effect on melt rates under Antarctic ice shelves.

Currently, a cold, dense water mass (called "high-salinity shelf water") is found on the continental shelf near many large ice shelves, such as the Filchner-Ronne Ice Shelf shown in Figure 5 (upper left). This cold, dense water and the processes that lead to its formation prevent a warmer water mass (usually called "circumpolar deep water") off the continent from reaching ice shelves. However, in some E3SM simulations, the warm water begins to lap onto the continental shelf and eventually reaches the ice-shelf cavity, dramatically increasing melting. This is not a process observed in the real world so far, but some polar modeling[6] has suggested that such a process could happen under a changing climate. Several other water masses also play a role in this process. LANL scientists felt that visualizing the evolution of these water masses in time and space would help us understand the processes that could lead to such a

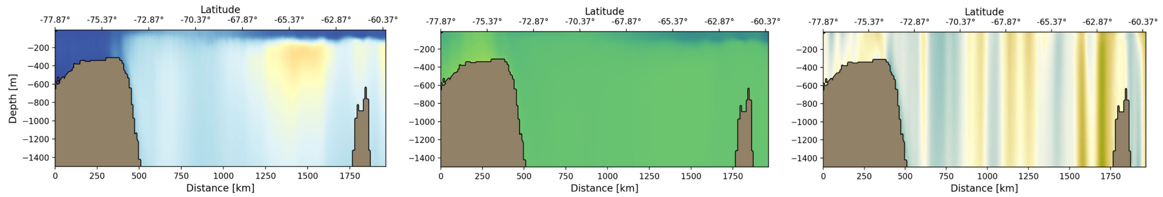

FIGURE 3. Standard output of cross-sections along a longitude line of 42 W, shown in Figure 2 with the continental shelf on the left. Images show surface temperature (left), salinity (middle), and eastward velocity (right), with the same colorbars as the previous figure.

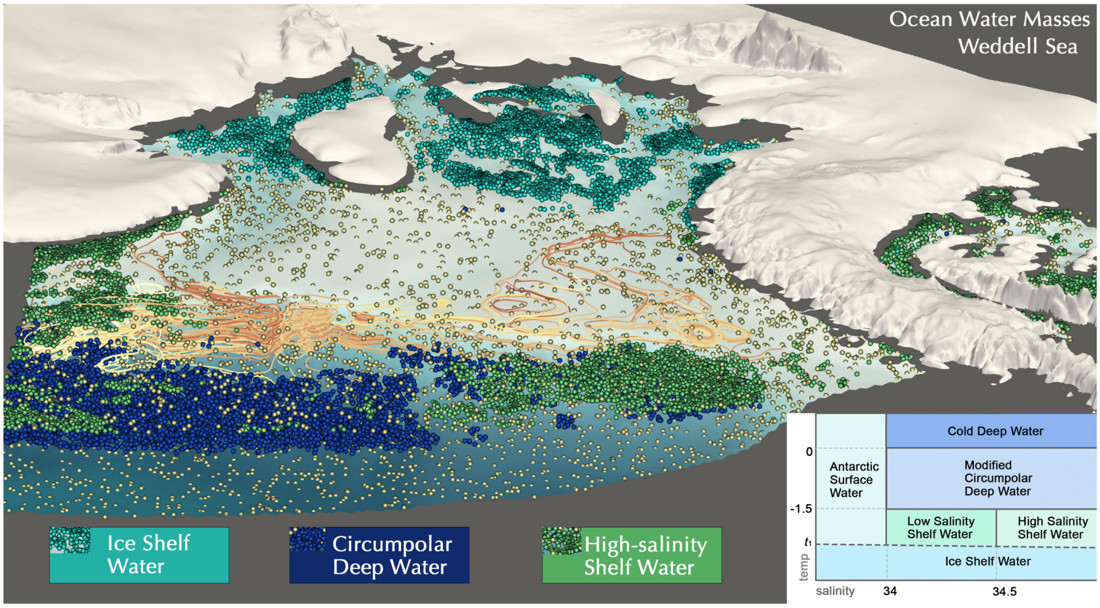

FIGURE 4. Four water masses that circulate below the Filchner-Ronne Ice Shelf and the current movements on the Continential Shelf. The chart on the lower right illustrates where the six categories of Southern Ocean water masses fall in salinity [PSU] and temperature [C].

transformation. For example, what would the intrusions of warm water look like and where would they occur first? At what locations and depths might they occur most frequently? How will water masses on the continental shelf have evolved by the point in time that warm water reaches under ice shelves?

Showing these categories of water masses concurrently is difficult for traditional visualization methods; opaque boundary surfaces obscure one another, whereas translucent methods (both surface and volumetric) result in ambiguous color combinations. Here, we sample the water masses to produce a set of well-distributed points in each water mass, then render these as small spheres. This visualization shows the movement of three water masses through time clearly and without ambiguous color mixing.

When ocean water flows away from Antarctica, it is substantially altered from its interactions with the continent. These altered properties influence the density and current patterns of water far into the Southern Ocean.

TECHNICAL CHALLENGES

There are technical challenges in extracting the water masses from the volumetric data. The water masses are, in effect, a segmentation of the computational space into subsets based on the relationships of temperature, salinity, and depth, as shown in Figure 4. While these segments are nonoverlapping and together fill the computational domain, each is a complex shape consisting of multiple disconnected parts, each with a convoluted boundary. Furthermore, since the data are changing over time, this is actually a four-dimensional problem. Visualizing such data in a movie is problematic. We have three dimensions to utilize to visualize these data: the two-dimensional image plane, and a third dimension of time. Typically, we sequentially slice the four-dimensional space along the axis of time to produce three-dimensional time-step datasets, project each onto the image plane, then play the projections back. However, this leaves us with a single two-dimensional projection of the three-dimensional timestep data to convey the shape of the segments. Simply rendering the boundaries of the segments as opaque surfaces leads to issues of occlusion: the resulting image only reflects the nearest surface to the viewer. Integrative methods, including rendering translucent surfaces or volume rendering the classified volume lead to problems of ambiguity—overlaying different colored surfaces (or volumes) produce colors that do not clearly reflect the data along the viewing ray. Instead, we choose to represent the water masses by producing a set of samples of each, then positioning colored glyphs at the sample points. We arrive at this solution because the science requires understanding the general ebb and flow of the masses rather than their precise boundaries. We can therefore

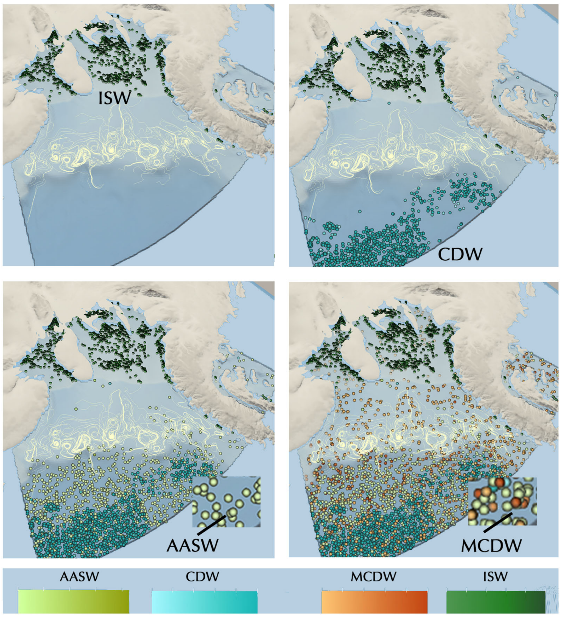

FIGURE 5. Distribution of four categories of ocean water masses are: Ice Shelf Water (ISW), Cold Deep Water (CDW), Antarctic Surface Water (AASW), and Modified Circumpolar Deep Water (MCDW). The yellow–orange streamlines follow the current along the continental shelf. Small spheres are used to show the mixing of different water masses.

reduce the data statistically to a set of pseudorandomly placed sample points and place a colored spherical glyph at each. We can then control the occlusion problem by varying the sampling density and radius of the glyphs. Glyphs for nearby water masses will predominate over those of father away masses, but statistically not occlude them. Further, we can rely on perspective to give additional information about distance; glyphs from distant water masses will appear smaller than those from nearby masses.

Segmenting the Data

The first step of this process—the segmentation of the computational space among the different water masses—is done at the level of the computational grid. Our input data consist of a volume represented by a large set of vertices in three-dimensional space connected into space-filling cells. Values of salinity and temperature are assigned to each vertex by the simulation. Given these and the coordinate location of the vertex, we can evaluate the relationships of Figure 2, bottom right, to create six derived scalar variables, one for each water mass, which are 1 for the vertices that lie inside the water mass and 0 for those outside the water mass. These derived scalar variables allow us to create six probability density functions, one for each water mass. For a particular water mass, we determine a sampling probability by averaging the value of each vertex of the cell, then multiplying the result by the ratio of the volume of the cell to the overall volume of the computational space. This produces a result that, for each cell, reflects the likelihood that a random sample will occur in the corresponding cell.

Sampling the Segmented Data

Given this probability function, we can choose cells for sampling. We do so by using a binary interval search tree based on the running sum of the cell probabilities: the likelihood of a random sample falling in the ith cell is the probability of the sample falling into the $[i,i+1]$ interval of the number line. This results in a sampling algorithm that takes $O(n)$ time to generate the running sum table (for n the number of cells), followed by $O(\log n)$ time to insert each sample. Once we choose a cell for sampling, we simply choose a random point inside the cell and add it to the output sample list. We note that this is imperfect in cells that only partially intersect a water mass; samples should tend toward the vertices of the cells that lie within the mass and away from those outside the mass. However, since the number of cells is far larger than the number of samples, the error introduced has not justified the additional cost of more accurate methods.

Continuity Over Time

The above algorithm can be computed independently for each timestep, but doing so can produce a sparkling effect over time: even if the underlying data changes only very slightly from timestep to timestep, the samples chosen will be completely different. When there are a relatively few samples chosen, this becomes a distracting scintillation. We address this problem by retaining samples from timestep to timestep, re-evaluating the probability of their inclusion in the new timestep, before choosing new samples for the new timestep. A sample from the prior timestep will only be retained in the subsequent timestep if it remains a likely choice in the new timestep. Thus, samples will change relatively smoothly from timestep to timestep. Since the above algorithm samples with replacement, this retention distorts the sampling; areas of overlap between successive timesteps will be oversampled, and new areas will tend to be undersampled. This led us to consider alternative sampling algorithms that have produced good preliminary results. The timestep sequence shown in Figure 6

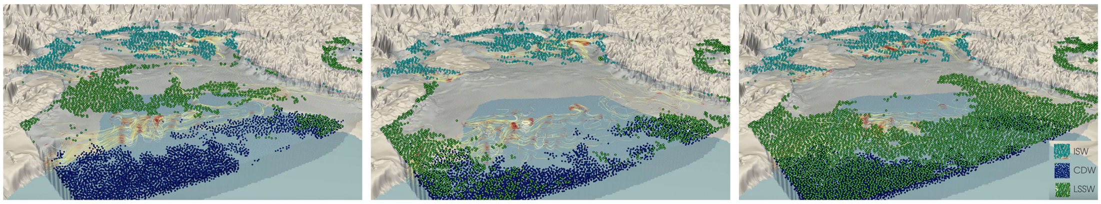

FIGURE 6. Three timesteps showing three water masses ISW, CDW, LSSW and the currents on the continental shelf over the course of a year.

illustrates the ability to follow the changes in location and density of the water masses over time

THE ROLE OF DESIGN PRINCIPLES

Figures 2 and 3 illustrate the traditional format for visualizing the temperature, salinity, and water masses as charts. Figures 4 and 5 in contrast illustrate the four water masses—intersections of temperature and salinity categorized to facilitate the science, as they sit in relation to the terrain and ocean floor depth. Our sampling approach enables us to render the four masses in one image, providing a clear representation of their interrelationships. Figure 6 documents the further value of this method, showing a time sequence of the simulation that enables scientists to track the changes and interactions of the water masses over time. Visualization often faces tradeoffs between clarity and complexity. By employing artistic color theory and design principles,[7] we can enable clear differentiation between the water masses as they move and mingle over the time-varying sequence. Color theory and design principles are central to clarity within multivariate visualization.

Our attention and ability to distinguish between the variables are determined not only by the hues and shapes selected but also by the interaction between these visual elements.[8] As the number of elements in a visualization increases, so does the need to minimize unnecessary contrast; we therefore pay close attention to the relative contrast between the color-encoded variables. Balancing the ability to distinguish between variables while maintaining visual harmony and congruency with the science questions is a delicate balance and a difficult challenge.

Color contrast theory describes the seven types of color contrast, which we employed in this visualization.[7] In Figure 4, seeking to balance the hue selection, need for discrimination, and desire for semantic color association, we selected cool hues to represent the water masses and saturated them for ease of location and distinction. By limiting our hue selection to the cool range of the color wheel for both the water masses and the ocean floor depth, the principal of analogous color contrast enables clear contrast and minimizes the simultaneity effect, a contrast principle defined by the visual vibration that occurs when saturated hues abut one another. We applied a colormap spanning warm hues to the streamlines representing the flow along the continental shelf to provide a clear contrast between the thin streamlines and the water representations. Even the background hue interacts with the color palette of the variables. We applied a warm gray to contrast with the primarily cool light hues encoding the data. Employing Artistic Color Contrast Theory, developed by artists over centuries, to visualization color selection provides guidance toward creating clear and harmonious visualizations, inviting a quiet and easy data exploration.

DISCUSSION

By comparing the standard diagnostics in Figures 2 and 3 to the later figures, we see two different visualization methods. The first is a horizontal or vertical slice of a single variable, averaged in time. The full output includes hundreds of such plots, spanning variables, time, depth, and location. In contrast, Figures 5 and 6 provide a more integrated view of this high-dimensional space. In order to address the science questions asked of the data, we showed the water masses in three-dimensional space, their movement over the course of the year, their interactions with the ocean floor, and mixing caused by the turbulent flows. Visualization researchers bring new ideas and advanced techniques from the visual arts community, which are often missing from standard visualizations. This artistic focus adds valuable engagement and the design and clarity capable of enabling comprehension and communication of these complex scientific concepts both within and outside of the scientific community.

CONCLUSION

Scientific data are growing on size and complexity. Addressing the needs of the scientific community will require new multidisciplinary approaches to visualization. Here we have demonstrated an example of one set of solutions that arose from a close collaboration between the scientists, the modelers, the visualization team, and an artist. While the scientific visualizations discussed here were designed to address our specific need, our research team believes our overall direction rests on fundamentals applicable across multiple scientific domains.

ACKNOWLEDGMENTS

Simulation data are available at https://e3sm.org/data/get-e3sm-data/released-e3sm-data/v1-1-deg-data-cryosphere. This research was supported as part of the E3SM Project, funded by the U.S. DOE, Office of Biological and Environmental Research. Computing resources were provided by the National Energy Research Scientific Computing Center (DOE Contract DE-AC02-05CH11231) and Argonne Leadership Computing Facility (DE-AC02-06CH11357), both awarded under an ASCR Leadership Computing Challenge Award, and by the LANL Institutional Computing Program, supported by the DOE National Nuclear Security Administration under Contract 89233218CNA000001. This work was supported by the National Science Foundation under awards IIS 1704604 and IIS 1704904.

REFERENCES

1. T. Ringler *et al.*, "A multi-resolution approach to global ocean modeling," *Ocean Model.*, vol. 69, pp. 211–232, Sep. 2013.
2. M. R. Petersen *et al.*, "An evaluation of the ocean and sea ice climate of E3SM using MPAS and interannual CORE-II forcing," *J. Adv. Model. Earth Syst.*, vol. 11, no. 5, pp. 1438–1458, 2019.
3. J.-C. Golaz *et al.*, "The DOE E3SM Coupled Model Version 1: Overview and evaluation at standard resolution," *J. Adv. Model. Earth Syst.*, vol. 11, no. 7, pp. 2089–2129, 2019.
4. K. E. Hoch *et al.*, "MPAS-Ocean simulation quality for variable-resolution North American coastal meshes," *J. Adv. Model. Earth Sys.*, vol. 12, no. 3, 2020, Art. no. e2019MS001848, doi: 10.1029/2019MS001848.
5. H. Jeong *et al.*, "Impacts of ice-shelf melting on water-mass transformation in the Southern Ocean from E3SM simulations," *J. Climate*, vol. 33, no. 13, pp. 5787–5807, Jun. 2020.
6. H. H. Hellmer, F. Kauker, R. Timmermann, J. Determann, and J. Rae, "Twenty-first-century warming of a large Antarctic ice-shelf cavity by a redirected coastal current," *Nature*, vol. 485, no. 7397, pp. 225–228, May 2012.
7. L. F. Rathus, *Foundations of Art and Design*. Stamford, CT, USA: Cengage Learn., 2015.
8. J. Albers, *The Interaction of Color*. New Haven, CT, USA: Yale Univ. Press, 2009.

GREG ABRAM is a Research Scientist with the Texas Advanced Computing Center, University of Texas at Austin, Austin, TX, USA, specializing in visualization systems design for use on supercomputers. He received the Ph.D. degree in computer science from the University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. He is the corresponding author of this article. Contact him at gda@tacc.utexas.edu.

MARK R. PETERSEN has been an Ocean Model Developer with the Los Alamos National Laboratory, Los Alamos, NM, USA, since 2005. He is an avid accordionist, a spelling reform hobbyist, and enjoys mentoring young scientists. His biography can be found at http://mark-petersen.github.io. Contact him at mpetersen@lanl.gov.

FRANCESCA SAMSEL is a Research Scientist with the Texas Advanced Computing Center, University of Texas at Austin, Austin, TX, USA. Using her artistic training, she focuses on multidisciplinary collaboration domain scientists to expand the depth of detail and insight that can be exacted from scientific data. Contact her at fsamsel@tacc.utexas.edu.

XYLAR ASAY-DAVIS is a Staff Scientist with the Fluid Dynamics and Solid Mechanics Group, Los Alamos National Laboratory, Los Alamos, NM, USA. His biography can be found at https://xylar.github.io. Contact him at xylar@lanl.gov.

DARIN COMEAU is a Staff Scientist with the Computational Physics and Methods Group, Los Alamos National Laboratory, Los Alamos, NM, USA. He received the Ph.D. degree in applied mathematics from the University of Arizona, Tucson, AZ, USA, in 2013. Contact him at dcomeau@lanl.gov.

STEPHEN F. PRICE has been a Staff Scientist with the Fluid Dynamics and Solid Mechanics Group, Los Alamos National Laboratory, Los Alamos, NM, USA, since 2008. His current research focuses on improving our understanding of ice sheets and their interactions with the Earth system using computational models. He received the Ph.D. degree in geophysics from the University of Washington, Seattle, WA, USA. Contact him at sprice@lanl.gov.

Contact department editor Mike Potel at potel@wildcrest.com.