Visually Communicating Future Climate in a Web Environment

COREY DAVIS AND HEATHER ALDRIDGE

State Climate Office of North Carolina, North Carolina State University, Raleigh, North Carolina

RYAN BOYLES

Southeast Climate Adaptation Science Center, U.S. Geological Survey, Raleigh, North Carolina

KAREN S. MCNEAL AND LINDSAY MAUDLIN

Department of Geosciences, Auburn University, Auburn, Alabama

RACHEL ATKINS

Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, North Carolina

(Manuscript received 9 December 2019, in final form 13 July 2020)

ABSTRACT

While there is growing demand for use of climate model projections to understand the potential impacts of future climate on resources, there is a lack of effective visuals that convey the range of possible climates across spatial scales and with uncertainties that potential users need to inform their impact assessments and studies. We use usability testing including eye tracking to explore how a group of resource professionals (foresters) interpret and understand a series of graphical representations of future climate change, housed within a web-based decision support system (DSS), that address limitations identified in other tools. We find that a three-map layout effectively communicates the spread of future climate projections spatially, that location-specific information is effectively communicated if depicted both spatially on a map and temporally on a time series plot, and that model error metrics may be useful for communicating uncertainty and in demonstrating the utility of these future climate datasets.

1. Introduction

Decision-makers attempting to address the impact of future changes in climate are generally interested in four basic pieces of information for a local area [National Research Council 2010a; U.S. Government Accountability Office (GAO) 2009]: 1) observations of current climate conditions, 2) observations of climate impacts and vulnerabilities, 3) projections of climate

changes, and 4) projections of climate change impacts (costs and benefits).

There is a wealth of climate data available including observations and modeling output. However, the National Research Council (2010a) and GAO (2009, 2015) state that stakeholders face challenges in accessing, analyzing, and interpreting historical climate data and future climate change projections. The authors' experience with climate services in the southeastern United States confirms these challenges. In particular, stakeholders from a range of sectors pose similar questions about using future climate change projections, including, "Can you give me the best global climate model projection? What is downscaling, and do I need to use it? Which downscaling method is best? Can I just use the multi-model average? Or, can I use the best two models to see a range of future impacts?" These kinds of questions suggest a need for technical expertise to access such data; but, importantly, there is a greater

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/WCAS-D-19-0152.s1>.

Corresponding author: Ryan Boyles, rboyles@usgs.gov

DOI: 10.1175/WCAS-D-19-0152.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

need for expertise on how to interpret climate model projections.

These questions are not easily explained in simple terms for data users. There generally is no single best climate simulation or downscaling method to use (Pierce et al. 2009; Knutti et al. 2010). While using a multimodel mean can provide lower historical error than any single simulation, a range of projections is often recommended to better understand the range of possible future climates and the range of possible impacts. The answers to such questions can overwhelm stakeholders including those who desire to incorporate climate projections into their decisions and planning.

While one might expect nonscientific users to ask such questions, our experience suggests that these same questions are also commonly asked by nonclimate scientists and researchers who generally understand data analysis and modeling but who lack expertise in climate data and climate models. Engineers, hydrologists, plant breeders, and foresters have all posed similar questions to the authors in recent years. Overall, communicating future climate projections to stakeholders—whether they have scientific background or not—is a fundamental challenge for climate scientists and climate service providers (Moser 2010; Somerville and Hassol 2011; McNie 2013).

From our experience working with stakeholders, there are five resources that are widely used by nonclimate scientists and resource managers in the United States to explore future climate projections, though this list is not exhaustive:

- 1) The National Climate Assessment (U.S. Global Change Research Program 2017) is the U.S. government's authoritative product on climate change and its impacts to the country. It provides extensive narrative, regional maps of average projected change for temperature and precipitation, and summaries of regional impacts of projected climate changes.
- 2) The USGS GeoData Portal (Blodgett et al. 2011) provides a wealth of high-resolution downscaled climate projections from a variety of producers. It provides an interactive map for individual global climate models (GCMs), multimodel averages, and provides users a way download grid-specific time series data.
- 3) Climate Wizard (The Nature Conservancy 2009) was developed by The Nature Conservancy and partners and was widely used by natural resource professionals prior to 2019. It allows users to view maps of both historical and projected climate variables from individual GCMs and multimodel ensemble averages. Users can also download time series of the projections for regions of interest. As of 2019, Climate Wizard is no longer available.
- 4) The USGS Climate Change Viewer (Alder and Hostetler 2013) provides statistically downscaled projections and allows users to view climate change maps, regional time series, and the distribution of GCMs for a selected region. Summary reports can be generated that include additional charts and graphics averaged over predetermined state/county areas or hydrologic unit (HU) HUC2–HUC8 watersheds.
- 5) NOAA Climate Explorer (<https://crt-climate-explorer.nemac.org/>) allows users to explore maps and time series plots of historical data and downscaled future climate projections. Users can search for data by entering a location of interest, multiple variables of interest (e.g., mean daily precipitation, heating degree-days), or a topical area of interest (e.g., coastal, water, ecosystems).

Each of these resources has been developed to specifically support the improved understanding and use of future climate change projections by stakeholders, but each takes a different approach in style and detail, and little is available on the development methods or intended audiences for most of these resources. Based on our experiences, these products provide very important utility. Each one attempts to address the challenges stated by the National Research Council (2010a) and GAO (2009, 2015)—user needs to access, analyze, and interpret climate change projections—through the use of visually engaging and intuitive map displays that depict downscaled future climate projections with different greenhouse gas (GHG) emissions and/or representative concentration pathways (RCPs). However, these resources have three primary limitations that we find are critical to communicating future climate projections:

- 1) Single maps are used, which suggests erroneously that a single future climate projection might be appropriately used for decision-making instead of a spread of possible future climates across the downscaled GCMs. While this specific example is focused on climate projections, others have found similar issues in visualizing uncertainty in geospatial data such as Kinkeldey et al. (2014), Deitrick (2012), Deitrick and Wentz (2015), and Kinkeldey et al. (2015).
- 2) City- and location-specific values are not easily communicated. The National Research Council (2010a) and GAO (2009, 2015) suggest that making such localized information available is key for stakeholders to be able to use future climate data. While the USGS Climate Change Viewer and NOAA Climate Explorer do contain spatial scales down to county and HUC8 levels, this is insufficient

according to National Research Council and GAO recommendations. Ideally, location-specific data would be included for a wide range of future time periods averaged over 20 years or more, as well as for different GHG emissions/RCP scenarios.

- 3) Model error, or other characterization of uncertainty, is not included or not easily interpreted. This is critical for many resource managers as their ability to effectively assess risk is often tied to an understanding of uncertainty in the model projections.

To help non-climate-science stakeholders better understand and use future climate projections, we need tools to assist the visual communication of climate change over space and time including the spread of possible future climates at meaningful spatial scales. The goal of this manuscript is to tell the story of the authors attempt to build a visualization tool, designed for natural resource professionals (especially foresters), that addresses the above limitations through a somewhat intensive and iterative user engagement process. We attempt to explore which combination of text and graphics leads to the most consistently correct and efficient interpretation of climate change projections. Such visuals should improve on the efforts of others by addressing the above limitations. Here, we describe our methods and experience to develop such tools and the lessons learned.

2. Methods

Methods presented here were intended to support development of a system of web-based tools (including layout, narrative, and graphical visuals) specifically to foster consistent and efficient interpretation of climate projections by users. These methods were not designed to address specific research questions. For brevity, more detail on the approaches (including tasks and questions from usability testing and eye-tracking methods) is provided in the supplemental material.

a. Audience and climate projection data

We explore a variety of visual communication products using an audience of professional resource managers. A study of Southeast forestry professionals' climate change attitudes revealed a substantial need and potential appeal for weather- and climate-related tools to assist with forest management while considering likely future climates (Boby et al. 2016). In addition, the study highlighted that effective climate change outreach to this audience should focus on ways to reduce risks, incorporate uncertainty into decision-making processes, and manage for resilience to future climate change

impacts in this region (Morris et al. 2014; Boby et al. 2016). A web-browser-based climate futures visualization platform was developed according to this recommended model for communicating with the target audience with an emphasis on the three limitations above. This platform, also called a decision support system (DSS) is designed to inform the planning and decision-making for this target audience of professional foresters. This DSS was developed as part of the Pine Integrated Network: Education, Mitigation, and Adaptation Project (PINEMAP), which was a large regional USDA project to inform the management of southern pine trees under climate change. PINEMAP provided research, extension, and educational engagement with forest researchers and professional foresters across the southeastern United States. (The final DSS is publicly available at <https://climate.ncsu.edu/pinemap/>.)

At the core of the visualization is the climate data upon which all graphics are based. Based on engagement with the community of professional foresters, data needs were prioritized to focus on sufficiently fine spatial resolutions and daily temporal resolution, and they allow for the option of variables beyond just temperature and precipitation. We chose to use the multivariate adaptive constructed analogs (MACA; Abatzoglou and Brown 2012) as it met each of these priorities. MACA uses a statistical downscaling method that provides localized guidance about future climate conditions across the United States at a resolution of 6 km, which was sufficient for this group of decision-makers. This dataset was assembled by researchers at the University of Idaho using 20 GCMs from phase 5 of the Coupled Model Intercomparison Project (CMIP5; WCRP 2011) and IPCC's Fifth Assessment Report (AR; Pachauri et al. 2014).

While no technique is perfect, MACA is a method that meets the needs of PINEMAP researchers by including features such as a daily time step and the meteorological variables (e.g., wind speed, specific humidity, solar radiation, minimum/maximum temperature, and precipitation) important for ecological applications and ecological process modeling. While four different Representative Concentration Pathways (RCPs) were developed for IPCC's Fifth AR (Moss et al. 2010), only RCP4.5 (stabilization pathway) and RCP8.5 (high greenhouse gas pathway) were used for the MACA downscaling and thus, within the visualization. Similarly, more than 40 GCMs were included in CMIP5, yet MACA downscaling method was only applied to 20 GCMs since daily data for all parameters was required for this technique (MACA; Abatzoglou and Brown 2012).

Although daily time steps were needed as inputs for various ecological process models and to analyze the

change in frequency of specific events (e.g., the number of days with minimum temperature below freezing), the graphical display of daily projections could be both overwhelming and misleading if they were interpreted as actual weather forecasts for a given future date. Instead, to simplify the available options and ensure that the future time periods visualized in graphics are on climate (not weather) time scales, model output for visualization was averaged across four 20-yr time periods: 2020–39, 2040–59, 2060–79, and 2080–99.

b. Design of the communication elements of the DSS

For the initial graphics to communicate climate futures, we focus on four measures, identified through conversations with professional foresters, for assessing climate risks and opportunities related to loblolly pine planting and growth. These include tools that would communicate 1) changes in summer average temperature and precipitation, 2) occurrences of extreme winter minimum temperature events at a variety of different intensity thresholds, 3) shifts in hardiness zones that can assist with seedling deployment based on annual extreme minimum temperature, and 4) changes in summer dryness index or the ratio of summertime growing degree-days to summer precipitation—the latter of which can be a useful measure of relative drought stress on loblolly pine trees (Sabatia and Burkhart 2014).

Analysis of MACA climate model output identified several challenges for visualization. First, the interactive maps needed to provide climate projections at spatial extents ranging from the entire southeastern United States to individual grid points. Effective visuals across such spatial scales would require more than just a map interface. In addition, each time period and RCP included three measures of spread from the model ensemble: a mean value and values two standard deviations above and below the mean. Incorporating all three of these values in the results required an interface more complex than previously developed by others (e.g., the five example resources described above). To show detailed results at both regional and local spatial scales, we designed and tested two separate displays. A map provides a regional view of conditions for a snapshot of one time period and RCP. For example, users viewing changes in summer temperatures could look at the average change, in degrees Fahrenheit, for the 2040–59 time period and RCP8.5. To view local conditions, a user could select a location by clicking the map or entering a latitude and longitude, which displays a second time series visualization below the maps. This time series plot provides an overview of conditions at the selected location including the historical average and projected change or average data for all future time periods and RCPs.

With that broad vision in place, determining the specifics for the map and time series displays required a balance between the need for accessibility and the desire for detailed output for three types of climate analysis: historical average, projected change, and projected average. We anticipate that users would follow a progression of steps to explore the climate data as shown in Fig. 1.

The *historical average* is calculated using meteorological data (METDATA; Abatzoglou 2013), which has spatial resolutions similar to MACA with the same daily meteorological variables. This historical data was averaged over the reference period 1986–2005, which represents a recent 20-yr rotation period in pine tree production. The *projected change* is calculated by computing the projected future outcomes (e.g., the average summer precipitation) for each downscaled GCM and emissions scenario, then subtracting the climate model baseline covering the period from 1986–2005. This communicates the importance of the model change, not the raw model output, as being useful for assessing future climate change risk. The *projected average* is calculated by adding the projected change of each variable to the historical average. Projected change and projected average are presented as three separate but connected map displays: the multimodel mean (the center map) and the values two standard deviations above and below the mean (the right and left maps, respectively). This three-map layout was proposed and tested as a possible framework to more effectively communicate model spread and interpretation of uncertainty based on the authors experience and the finding by Kinkeldey et al. (2014), Deitrick (2012), Deitrick and Wentz (2015), and Kinkeldey et al. (2015).

Model error was also generated to demonstrate how well the downscaled GCMs performed. Few existing displays of climate model output include this measure, but it is an important one with implications for decisions made based on the data. To calculate model error, the climate model baseline data (1986–2005) for all 20 downscaled GCMs was compared against METDATA historical observations (1986–2005) using mean absolute error. If the climate model baseline data matched reasonably well with what happened historically, the error should be low and users would infer more confidence regarding the utility of the future projections. If the error exceeds the future projected changes, the data are not as useful for decision-making. Thus, the values on the projected change and projected average map displays are masked out with a light gray color. For these same two data displays, model error is shown on the time series plot as an orange bar with a light gray zone overlaid on the future values. [Model error visuals are

FIG. 1. Schematic of expected user flow through key components of the climate DSS. Users are first presented with menus to select climate variables, and then are shown a map of historical average conditions. Users then select to view climate for a specific period in the future and emissions scenario. Finally, users can click on the map to view a time series for the selected location of interest.

described in more detail in [section 3a\(3\)](#) with examples shown in [Figs. 7 and 8](#).] Future values that fall within this zone of model error can be interpreted as not-as-useful guidance. Knowing when values begin falling outside the model error can determine, for example, whether professional foresters plan for conditions similar to those in the past or for changes from those past conditions.

The iterative process of development also incorporated feedback from stakeholders outside of the professional forester community, namely experts in communication. In an attempt to anticipate and overcome challenges related to effectively communicating climate science jargon and navigational features in the DSS, we engaged in discussions with a technical communications graduate student during the spring 2015 semester. Changes were proposed to the interface layout, wording, and graphics to promote ease of use, clarity, consistency, aesthetic appeal, functionality, and relevance to the target audience ([Misenheimer 2015](#)). Graphic designers at Southern Regional Extension Forestry (SREF) also assisted with enhancing navigational features in the DSS. During winter 2016, a new menu system was deployed and graphics were enhanced on the DSS introductory content to more easily guide first-time users through the interface.

c. User beta testing

Development was informed by an iterative design process that obtained feedback from the end users. Usability tests were performed to help determine the effectiveness of the map and time series displays with regard to communicating future climate change projections. Previous research by [Breuer et al. \(2009\)](#) demonstrates that an iterative process of development, which includes feedback from stakeholders, is beneficial during the development of a decision support system. Several other studies emphasize that this iterative approach is effective for communicating risk ([Fischhoff](#)

[1996](#); [National Research Council 1996](#); [Patt and Dessai 2005](#); [National Research Council 2010b](#)).

To support the iterative design of these visualizations, forestry researchers and practitioners served as beta testing users with input in several ways (details on the tester experience and feedback is provided in the online supplemental material). Initial map layout designs used conceptual mock-ups featuring “dummy” datasets such as gridded real-time fire risk estimates as a stand-in while climate model output was processed. These conceptual mock-ups were presented to testers for an informal evaluation. Next, users had a 2-week independent study period in which they were emailed basic information about the tool(s) they tested, a task or two to work through using the DSS, and a survey instrument related to the task(s) containing questions such as, “What did you expect *this feature* to do?” and “How did you interpret *this*?” This independent exploration period allowed beta testers to work through the task(s) on their own time since that best simulates the environment of actual DSS users. At the end of this 2-week period, users were separated into small groups of three to five individuals who generally shared similar responses to the survey. Several weeks later, these small groups of users spoke with DSS developers by phone to discuss their responses, including layout and color options for maps and legends, wording choices, and any comments, concerns, or suggestions they had related to the tool(s). After these discussions were completed, DSS developers documented necessary adjustments, using the GitHub issue tracking software to identify progress and completion (<https://github.com/>). Results from the user testing were documented in a similar manner on GitHub. A total of 21 researchers and practitioners participated in beta testing and provided feedback.

d. Eye tracking

To gain additional insights into usability and user interpretation of DSS features, we conducted an eye-tracking

study of the DSS with non-beta testers. Eye tracking is a noninvasive way of collecting data about gaze duration and location without interfering with normal viewing patterns. Tobii eye-tracking hardware and software were utilized to determine user attention, for example, where they look, when they look, and for how long. Our eye-tracking analysis was exploratory, and results include heat maps, gaze plots, and statistical measures of attention to various areas of interest. Detailed methods and approaches are provided in the supplemental material, but also provided in a companion paper Maudlin et al. (2020) and in those by Bojko (2006), Holmqvist et al. (2011), and Fiedler and Glockner (2012). This DSS eye-tracking study was performed at the Appalachian Society of American Foresters meeting in Durham, North Carolina, on 28–29 January 2016. A total of 30 volunteers ranging from students to professional foresters to researchers participated in the study, which took approximately 30 min and involved completing tasks and answering questions related to three tools in the DSS. Institutional Review Board (IRB) human subjects research approval was obtained before any usability research commenced. Using a design-based research approach (Brown 1992; Collins et al. 2004), the findings from the eye-tracking study were used to improve the design of the PINEMAP DSS.

3. Results and discussion

Using the methods described above, analysis of user testing is provided for the key design elements to gain insight into the efficacy of the graphical communication. Below we describe the results of the beta testers that were used to implement changes for the subsequent eye-tracking assessment of the DSS.

a. Results from beta testers

1) MAP DISPLAY

Initial map layout designs used conceptual mock-ups were presented to testers for an informal evaluation. First, as depicted in Fig. 2, a single-map layout was created that showed only the multimodel mean value with no options to display various percentiles. Although a single map is simple and familiar to many users, it does not communicate the possible spread of future outcomes. We presented this layout during a facilitated meeting of natural resource researchers, extension specialists, and stakeholders. Feedback indicated that a single map did not provide enough information about the range of likely future outcomes to make informed planning decisions. In response, a three-map environment was developed to

include the multimodel mean and model spread, which used values two standard deviations above and below the multimodel mean. These were chosen to represent the model spread because they were straightforward to calculate and explain to an audience that would likely be unfamiliar with other statistical measures of uncertainty. Two additional mock-ups included these map display options (Figs. 3 and 4). One version, shown in Fig. 3, kept a single map but added tabs for all three options so the user could toggle between them. Testers suggested that the three outcomes should be visible at the same time to be most effective. Another version (Fig. 4) depicted all three outcomes simultaneously by displaying the multimodel mean map in full-size above two half-sized maps showing the values two standard deviations above and below the mean. This version was difficult for reviewers to view in its entirety without scrolling, and they suggested that comparing between maps was difficult in such a stacked layout.

As DSS development continued, a layout was implemented with the three maps side by side and the multimodel mean in the center (Figs. 5 and 6). To accommodate this interface, the DSS web page content area was widened from the PINEMAP project standard of 960 pixels to 1260 pixels. This width proved to be one of the main limitations during the DSS development process. If the page were too wide, users with older or lower-resolution computer monitors could be forced to scroll horizontally to see the three-map layout in its entirety.

Google Analytics data showed that for the calendar year 2015, 58% of visitors to the main State Climate Office of North Carolina website (which hosts the DSS) had screen resolutions at least 1280 pixels wide. Keeping the DSS width less than 1280 pixels ensured that the full layout would fit in a full-screen browser window for most users, although our survey of beta testers found that not all users viewed their browsers in full-screen mode. The wider page layout did make the DSS impractical for use in a mobile environment, and that user base is not trivial. In 2015, 27.8% of visits to DSS host website were from mobile devices. Discussions with professional foresters and extension agents suggested that they would primarily view the DSS on a desktop computer, so the compromise was made to develop for that environment in which the wider layout is more accessible.

One other important decision concerning this three-map layout was the relative size of each map. As shown in Fig. 5, initial designs made all three maps the same size, which was appropriate to convey the equal likelihood of each among the spread of outcomes. However, the limited horizontal screen space meant all three maps were relatively small and lacking the desired level of

FIG. 2. An early design mock-up of what eventually became the Minimum Temperature Thresholds tool included a single map showing the multimodel mean. In this mock-up, dummy data were displayed instead of a calculated mean model projection.

detail on a regional scale. Instead, the decision was made to have one map—by default, the middle map showing the multimodel mean—appear twice as large as the other two. This design is shown in Fig. 6. Although only one map at a time is enlarged and shown with this higher level of detail, users can click either of the side map titles to enlarge them.

A final decision regarding the map layout was the wording for the map titles themselves. While the calculation for the side maps—two standard deviations above and below the multimodel mean—was relatively straightforward, explaining this in short, simple terminology was not as easy. The early mock-ups and functional DSS designs used the titles “minimum projected change,” “mean projected change,” and “maximum projected change” for the three maps, with the term “change” substituted for “average” on the Projected Average display. This wording caused some confusion because beta testers interpreted the side maps as

showing the single models with the highest or lowest values, when in fact they represented the 2.5th and 97.5th percentiles from the model output. A beta tester familiar with climate model data helped evaluate several other wording options. The final choices were “lowest likely change” or “lowest likely outcome” for the leftmost-side map, “multi-model mean” for the middle map, and “highest likely change” or “highest likely outcome” for the rightmost-side map. Although a term such as “likely” includes an implication about the probability, which Budescu et al. (2014) found are subject to a wide range of interpretations among users, the map titles are supplemented with tool tips and a more detailed FAQ page that explain the exact calculations involved in generating the map data.

2) TIME SERIES DISPLAY

A time series plot was needed to show the range of future outcomes for a selected location. These plots

FIG. 3. A second-iteration mock-up for the Minimum Temperature Thresholds tool including tabs to toggle the multimodel mean and extremes, and the first version of a time series plot showing historical data and future projections. All information displayed was dummy data instead of actual model simulations.

needed to load quickly, update dynamically, and be understood by the target audience of professional foresters.

The USGS National Climate Change Viewer (NCCV) tool (Alder and Hostetler 2013) uses a time series plot with a continuous line through the historical and future time periods. However, this plot does not include the model spread, which is a limitation we tried to address. In addition, because we were averaging historical and future data into 20-yr time periods, connecting these

single values with a line seemed inappropriate and could have led to the erroneous extrapolation of single-year values from the line. Instead of using this single-line approach, each future period was represented by two discrete bars: one for RCP4.5 and one for RCP8.5. Designing these bars to best represent the spread of model solutions was a major challenge, and six designs were ultimately created before they were evaluated with beta testers climate modeling experts.

FIG. 4. An alternative mock-up of the Minimum Temperature Thresholds tool showing the (top) mean projected change and (bottom) half-sized (left) minimum and (right) maximum projected change maps. All information displayed was dummy data instead of actual model simulations.

The time series graphics tested are provided in Figs. 7a–f. Results from beta testers is provided in Tables 1–3 with more detailed information on testing and evaluation in the supplemental materials. The initial “box and whiskers” design (Fig. 7a) had two components: filled bars spanning plus and minus one standard deviation around the multimodel mean and thin error bars spanning plus and minus two standard deviations. Although reviewers gave this design high scores for its

clarity in displaying future ranges (Table 2), ease of determining specific values (Table 3), and success in concisely presenting information (Table 1), they also identified several fundamental barriers to correctly interpreting the information. This design used error bars to represent the spread of model solutions, so this was interpreted by some to be showing the model error. Even in our own explanations of the time series, we often struggled to reference this feature without

FIG. 5. An initial version of the Minimum Temperature Thresholds tool implemented in the web-based PINEMAP DSS showing three equal-sized maps.

referring to it as “error bars.” In addition, the values plus and minus one standard deviation were deemed extraneous information by 9 of the 14 evaluators, and because these values were not on the maps, it created an inconsistency between the two displays.

A derivative of this design, called the “column and whiskers” (see Fig. 7b), did not include the values plus

and minus one standard deviation and instead had the filled bar span from zero to the multimodel mean. This design closely matched one used in the IPCC’s AR5 Summary Report to show assessed likely ranges of historical warming trends (Pachauri et al. 2014). While this design received high scores for clarity in showing future ranges (Table 2), beta testers were split on its ease for

FIG. 6. The final design of the Minimum Temperature Thresholds tool, including a three-map layout.

determining specific values (Table 3). Evaluators noted confusion about why the full range of values from zero to the multimodel mean were spanned by the filled bar if not all of those values were actually part of model solutions. Likewise, a similar design called the “column and thick bars” (Fig. 7c) replaced the error bars with a wider, shaded bar, but this design was deemed as one of the least effective by the evaluators; it received the lowest average score for the ease of determining specific values. Despite removing confusion related to the inclusion of error bars, this design created additional confusion in part because the data labels for the multimodel mean value overlapped with the thick bars, making it unclear as to which bar that numeric value described.

Two other designs, called “scatter range” plots, removed the bars and instead represented the multimodel mean value with a small circle and the values plus and minus two standard deviations with triangles pointing toward the mean, as shown in Fig. 7d. A similar design, displayed in Fig. 7e, added diamonds to represent values plus and minus one standard deviation. Evaluators were also unreceptive to these designs; Figs. 7d and 7e received the lowest two average scores for the clarity of displaying future ranges, and Fig. 7e received the lowest average score for the success of clearly and concisely presenting information.

With simplicity at a premium in the eyes of the reviewers, one final design used just one bar to span the values plus and minus two standard deviations about the multimodel mean, with the opacity of the bar highest near the mean and lowest (most transparent) at the top and bottom edges. This “faded bars” design (shown in Fig. 7f) received the highest average scores for how clearly it displayed future ranges, highest for ease to determine individual values, and the second-highest average score for its overall success and clarity. Testers noted more comfort with this version compared to the box and whiskers since they deemed it less likely that the model spread would be interpreted as model error. With that feedback in mind, the faded bars design was chosen for use in the DSS. Although questions did arise about assuming a normal distribution of model output and representing that distribution with a basic linear fading, the evaluators agreed that trying to show small variations in the model distribution on already-small bars in this time series plot would compromise the clarity that they valued in this display.

3) MODEL ERROR

Displaying model error on the maps and time series proved quite challenging. The model spread of future outcomes displayed in the three-map layout was sometimes

erroneously referred to as “model error” by beta testers, so adding an extra layer that actually depicted model error had the potential to add even more confusion. To avoid these misinterpretations, it was important to find a way to present the data intuitively so that users could clearly see whether values were within the range of model error and correctly interpret the meaning of model error as relevant to the magnitude of change.

Showing a single, raw model error data layer as another map display was found to not be meaningful to the DSS target audience (if it was even viewed). Instead, locations where the absolute value of the projected change exceeded the mean absolute error value were displayed in gray, as depicted in Fig. 8. This masked out the raw Projected Change or Projected Average values at those points for all future time period and RCPs. This approach is somewhat similar to the Melillo et al. (2014), which uses white shading on maps for areas where the projected changes do not exceed the magnitude of natural variability. The National Climate Assessment also uses hatching to highlight areas where projected changes are “significant and consistent among models” (Melillo et al. 2014). On the time series plot, the model error was displayed as a pair of dashed lines spanning the entire plot horizontally with transparent shading between them, which allowed for easy comparisons with the future bars. An example of this display is shown in Fig. 9. Model error metrics were tested, but ultimately not implemented due to challenges in interpretation [see section 3a(3) challenges below].

b. Results from eye tracking—Layout and design

Feedback from beta testers during the DSS development process helped shape the wording, color palettes, time series displays, and other features that were implemented. However, this feedback could not show how users actually interacted with the DSS, especially when they first visited the site. Instead, results from the eye-tracking study using 30 non-beta tester participants helped identify several key usability issues with the DSS. This included analysis of heat maps showing where users looked for the longest amount of time as in Fig. 10. Notably, when asked to identify the spread of values for a particular location, only 6 respondents used the three-map layout, showing an overall underutilization of the side maps in contrast to the feedback from beta testers. Other users preferred the time series plots to determine their answers. When asked to compare the spread of outcomes across all future time periods, 16 users looked only at the size of the bars and/or the values at the top and bottom of each bar to answer the question. Seven users also consulted the legends explaining what each bar represented.

FIG. 7. Variety of time series plots tested. (a) A box-and-whiskers-style time series plot with bars spanning plus and minus one standard deviation around the multimodel mean and whiskers spanning plus and minus two standard deviations around the mean. (b) A column-and-whiskers-style time series plot with bars spanning from zero to the multimodel mean and whiskers spanning plus and minus two standard deviations around the mean. (c) A column and thick bars time series plot with wider bars spanning from zero to the multimodel mean and narrower bars spanning plus and minus two standard deviations around the mean. (d) A scatter-range-style time series plot with circles representing the multimodel mean and triangles at plus and minus two standard deviations around the mean. (e) Scatter-range-style time series plot with circles representing the multimodel mean, diamonds at plus and minus one standard deviation around the mean, and triangles at plus and minus two standard deviations around the mean. (f) A faded bars time series plot with bars centered on the multimodel mean and faded out at plus and minus two standard deviations around the mean.

Other findings from the eye-tracking study showed where users were not looking: at embedded help options; at the Layer Options menu on the map, which was not expanded by default; and at the options to view 5° ranges on the seedling deployment tools, which were arguably the most important features of those tools.

These 5° ranges (described by Schmidtling 2001) have been used by southern pine foresters for seedling selection for nearly two decades as guidance to maximize growth and balance cold risk. With that information in mind, adjustments were made to each of these features to make them easier to find and use. The length of the

FIG. 7. (Continued)

text was reduced, and the most important information was placed as the first bullet point. The structure and layout of tool tips was reworked, shortening the text and migrating some content to a “Frequently Asked Questions” web page that provides supplemental information when users require assistance and a tutorial for new users. Figure 6 depicts the final three-map layout selected. Figure 7f shows the final times series design selected for conveying location-specific projections.

c. Challenges

During the DSS development process, challenges arose due to conflicting views of the beta testers—probably

related to differences between level of expertise and their interpretation of the data. For instance, some users wanted to see detailed explanations of how the data were generated yet others struggled to understand what was being presented on the maps and time series plot due to the use of unfamiliar terminology (e.g., climate modeling jargon). Many studies have explored the differences between discipline experts and novices (Barfield 1986; Simmons and Lunetta 1993; Chen et al. 2006; Hmelo-Silver et al. 2007; Jarodzka et al. 2010; Gegenfurtner et al. 2011; Kastens et al. 2016). Experts are usually scientists or college-level instructors with formal training and experience in a particular domain, while novices are typically

TABLE 1. Aggregated responses ($n = 16$) from beta testers to the question, “Based on how clearly they display future ranges, rank the six time series plots from 1 (the worst) to 6 (the best).”

Chart type	Number of responses per rating						Average score
	1 (worst)	2	3	4	5	6 (best)	
Box and whiskers	0	1	3	10	2	0	3.81
Column and whiskers	3	1	1	4	2	5	4.00
Column and thick bars	3	4	5	0	3	1	2.94
Scatter range: Mean plus and minus two std dev	1	7	5	2	1	0	2.69
Scatter range: Mean plus and minus one and two std dev	8	2	1	3	0	2	2.44
Faded bars	1	2	3	4	0	6	4.13

students or the general public with little or no training in the same domain (Simmons and Lunetta 1993). Furthermore, what experts know within a specific domain is not just quantitatively different but also qualitatively different (LaFrance 1989). When looking at complex and dynamic visualizations, experts tend to focus their attention on the more relevant information surrounded by irrelevant information than do novices, demonstrating perceptual/attentional skills (Jarodzka et al. 2010). Experts are also often able to complete tasks faster than novices (Jarodzka et al. 2010). A companion manuscript by L. Maudlin et al. (2020, unpublished manuscript) details the role of expertise on usability of this DSS.

Attempts were also made to maintain a balance on the DSS between user flexibility and speed. For example, beta testers of the dynamic hardiness zone tools desired the option to define additional 20-yr periods. Time and processing costs prevented the ability to choose custom periods, so instead beta testers recommended prepopulation of two additional near-term 20-yr periods (2010–29 and 2030–49), which were deemed important since current foresters will be making decisions during those times.

Beta testing confirmed what others have found about color: different audiences may interpret the same color ramp in different ways depending on their background and familiarity with the data (Allen et al. 2006; Canham and Hegarty 2010; Hegarty et al. 2010). As an example,

the color ramp on the Minimum Temperature Thresholds tool initially varied from green, representing a smaller number of extremely cold days, to red, which represented more days with minimum temperatures below the selected threshold. Some beta testers suggested this was confusing because red is often associated with warmer temperatures, yet on this tool, it represented areas with a greater frequency of cold temperatures. However, simply reversing the color ramp was also misleading since other testers said they interpret green as being good, and more cold days might not be good for trees. To avoid these misconceptions, the color ramp for this tool was changed to span blue–white–yellow–orange–red–pink colors to mimic the USDA Hardiness Zones map (USDA 2012), which had the added benefit of being a familiar resource to the target audience.

In addition, conveying model error and uncertainty in future climate projections in a way that is meaningful to the DSS target audience was difficult and proved to be an unfinished task during the DSS development. This was due to two main issues: visualization and terminology. While these visualizations and their interpretations were not completely intuitive when shared among the PINEMAP team (see examples in Figs. 8 and 9), it was hoped that sufficiently good explanations could guide users as they displayed and interacted with model error on DSS tools. However, choosing the appropriate language for these explanations was even more difficult, as it required balancing scientifically or statistically correct

TABLE 2. Aggregated responses ($n = 15$) from beta testers to the question, “Based on how easy it was to determine individual values (e.g., the multimodel mean for the 2040–59 period under current/high emissions levels), rank the six time series plots from 1 (the worst) to 6 (the best).” One respondent did not answer this question.

Chart type	Number of responses per rating						Average Score
	1 (worst)	2	3	4	5	6 (best)	
Box and whiskers	1	2	5	5	2	0	3.33
Column and-whiskers	4	4	0	2	0	5	3.33
Column and thick bars	3	3	5	1	3	0	2.87
Scatter range: Mean plus and minus two std dev	1	5	2	1	2	4	3.67
Scatter range: Mean plus and minus one and two std dev	5	2	2	2	2	2	3.00
Faded bars	2	0	4	4	3	2	3.80

TABLE 3. Aggregated responses ($n = 15$) from beta testers to the question, “Based on the overall success of clearly and concisely presenting information, rank the six time series plots from 1 (the worst) to 6 (the best).” One respondent did not answer this question.

Chart type	Number of responses per rating						Average score
	1 (worst)	2	3	4	5	6 (best)	
Box and whiskers	1	0	2	2	7	3	4.53
Column and whiskers	3	1	1	3	3	4	3.93
Column and thick bars	5	4	3	0	2	1	2.53
Scatter range: Mean plus and minus two std dev	1	4	5	3	2	0	3.07
Scatter range: Mean plus and minus one and two std dev	5	5	2	1	0	2	2.47
Faded bars	0	1	2	6	1	5	4.47

terminology with language that was meaningful to the DSS target audience. It was decided not to use “error” to refer to these metrics since users may think they had made an error in using the tool or there was an error in displaying the data. Other options such as “reduced skill,” “high uncertainty,” and “low confidence” were also rejected due to possible interpretations that *all* models lacked skill or had low confidence in *all* locations. Instead, “model limitations” was preferred and implemented in a development version of the DSS, contextually used on the page in phrases such as, “For areas shaded in gray, values are not meaningful given model limitations.” Before these displays and this language could be formally evaluated, the PINEMAP project ended so no further work was done and the operational version of the PINEMAP DSS does not include model error metrics. However, due to the increasing prevalence and reliance upon future climate

projections, appropriately displaying and conveying model error or uncertainty remains an important and unsolved issue that warrants future research.

This communication challenge with the model error was complicated by the fact that the eye-tracking study on the DSS revealed that participants generally did not fully read the introductory content despite our best efforts to refine some of the technical language (as shown in Fig. 10). Using findings on the common “F shaped” reading pattern by Bergstrom and Schall (2014) and Pernice et al. (2018), we adjusted the order in which key information was presented. Based on recommendations from Misenheimer (2015), tool tips were added that defined unfamiliar terms and provided navigational hints, such as how the side maps could be enlarged. However, the eye-tracking study revealed that most users did not explore these tool tips either. Only three out of 30 study participants looked at any of

FIG. 8. The Minimum Temperature Thresholds tool with model error displayed. Areas shaded in gray denote locations where the magnitude of the model error exceeds the projections.

FIG. 9. A time series plot for the Minimum Temperature Thresholds tool with model error displayed. The orange dashed lines indicate the magnitude of the model error around zero (no change), and any future projections within those lines can thus be deemed not as meaningful for decision-making purposes.

the tool tips during their free exploration period or while solving tasks using the DSS. Previous studies suggest this lack of reading text is common (Morkes and Nielsen 1997; Spool et al. 1997) and is likely due to an inability to find the background information/tool tips or a desire to explore DSS data instead of reading this introductory content. Additional testing is needed to further explore the efficacy of the revised DSS, especially the design of tool tips, colors, and model error interpretation.

d. Limitations

The methods described here were not designed initially to answer specific research questions about

how users interpret language and graphics used to communicate climate projections. While the experience shared here is useful, an experimental design to address specific research questions would likely improve on these conclusions and recommendations. Similarly, while we think the experiences, conclusions, and recommendations are relevant for a wide audience of potential users of climate projections, the audiences we used to test the words, graphics, and layout design are not large and possibly represent a narrow group of natural resource professionals. Testing with larger audiences and other disciplinary users might produce more generalized conclusions and recommendations.

FIG. 10. A heat map showing where participants ($n = 17$) focused their attention while interacting with the DSS Introduction. Red denotes more attention given, and blue denotes less attention given. In this case, participants read the paragraph in an F-shaped pattern, meaning the majority of them read the entire first line, most of the second line, and decreasing amounts of the following lines of text. This finding is common (e.g., Bergstrom and Schall 2014; Pernice et al. 2018) and supported adjusting the text such that the most important take-away information is provided first.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. government.

4. Conclusions and recommendations

As outlined by the [National Research Council \(2010a\)](#) and [GAO \(2009, 2015\)](#), many challenges exist related to the access, use, and interpretation of historical and future climate datasets by non-climate-scientist audiences. Three key concepts presented here in developing this DSS can help improve communication of these complex datasets and assist with decision-making:

- 1) *a three-map layout* demonstrates the spread of future climate projections spatially,
- 2) *location-specific information* is accessible both spatially on a map and temporally on a time series plot, and
- 3) *model error metrics* may be useful for demonstrating the utility of these future datasets.

Through the use of usability testing and engagement techniques including eye tracking, we explored how a group of resource professionals (foresters) interpret and understand a series of graphical representations of future climate change that, at least through qualitative analysis, address the limitations identified in other tools. We find that a three-map layout effectively communicates the spread of future climate projections spatially and improves interpretation of uncertainty (following guidance by [Deitrick 2012](#); [Deitrick and Wentz 2015](#); [Kinkeldey et al. 2015](#)). We find that location-specific information is effectively communicated if depicted both spatially on a map and temporally on a time series plot. We find that model error metrics *may* be useful for partially communicating uncertainty and in demonstrating the utility of these future climate datasets, but more work is needed to confirm these findings.

While the audience of beta testers likely represents a more technical audience than those used in eye tracking, we consider them to be representative of natural resource managers and stakeholders who are interested in actually incorporating climate change projections into decision-making. However, these beta testers may not be (probably are not) representative of the wider audience of stakeholders interested in understanding climate change and its impacts. The beta testing group included individuals who we considered to be sophisticated web users and those who were very novice web users. Beta testers were also more heavily invested in specific applications, uses, and tools included in this DSS, and therefore may have been more likely to scrutinize each word, color, and layout choice to meet their needs.

The effective communication of scientific content remains a challenge, but the process used for developing these visuals ensured that it met the needs identified in [National Research Council \(2010a\)](#) and [GAO \(2009, 2015\)](#) for broader audiences. Future tools developed will also likely need to take a similar approach due to the increasing complexity of scientific data particularly with climate projections. With a clear vision and routine feedback from experts and stakeholders, the visuals developed for this DSS are able to fill a gap in the knowledge base of professional foresters to help them become better prepared and more resilient in the face of a changing climate.

Recommendations

From the vision, development, and testing of the DSS, we can offer several lessons for how to create more effective visual communications for scientific or technical information with a nonscientific audience:

- 1) Invest time learning about the range of potential users and uses before beginning development to have a clear vision for what the tool should accomplish. The final design of the DSS was actually the third overall iteration; earlier versions were for location-specific display with more emphasis on past climate information and less on future projections. However, the vision changed when we learned that several forestry industrial cooperatives already had proprietary tools with such information. After more a more rigorous assessment and understanding of forester needs and climate-related questions, the research and PINEMAP leadership team decided a tool more focused on future climate outcomes was determined to be more useful.
- 2) Survey the existing landscape of resources. We did not want to duplicate resources already available to our target audience. We also did not want to recreate a tool for climate projections that already existed. This required surveying existing tools, identifying their advantages and limitations, and determining the additional features needed. That resulted in the vision for a three-map layout and time series plot to show both regional and local projections, including the spread of model outcomes, which filled a niche that no other single tool had filled.
- 3) Plan for the time needed to iteratively solicit and incorporate feedback from actual users. In a relatively short time period of about nine months, the DSS went from design mock-ups to a usable interface showing actual data. However, the design process lasted longer due to the iterative feedback mechanisms we incorporated, and the final tool was much

stronger because of it. Whenever possible, we recommend including actual stakeholders in the development process to ensure that the tools meet their needs and answer their questions. Lacking that, surrogates such as extension agents or professionals can also provide meaningful guidance.

- 4) Consider a variety of usability testing—we found it to be extremely useful and provided unique insights. Working with communications experts during the initial stages of development can inform the sort of technical language that is appropriate. Testing with other scientists reassured us that even our simplified approach to presenting complex details, such as simplifying 20 climate models into three metrics, was scientifically and statistically sound. Beta testing at several stages with users during the development process helped ensure that details (e.g., layout, wording, and color choices) were effective. Finally, the eye-tracking studies showed us what no other testing could: which features users were *not* looking at, and therefore parts of the DSS that needed to be reworked.
- 5) While users often want something familiar, do not ignore the chance to test different ideas. The three-map layout and time series plot are likely new and unfamiliar to some first-time DSS users, which gave us periodic concerns about their accessibility by nontechnical users. However, testing showed that even novice users were generally able to complete tasks using the DSS. This provided support for our vision of trying something innovative.

These recommendations, and the experiences we developed in this effort that are shared here, build on the experiences by others cited here. We strongly encourage future endeavors to improve methods to effectively communicate climate science, and especially climate projections, in ways that foster correct interpretation of climate risk. Developers of such products and tools should be very cautious when using graphics—while powerful, they can potentially foster incorrect and misleading interpretations. Developers should make every effort to work with their intended users to test their graphics and designs to ensure viewers are able to correctly interpret the scientific data provided.

Acknowledgments. This research was partially supported under the Pine Integrated Network: Education, Mitigation, and Adaptation project (PINEMAP), a Coordinated Agricultural Project funded by the USDA National Institute of Food and Agriculture, Award 2011-68002-30185. The authors want to thank Ms. Carrie Misenheimer for guidance as a Technical Communications

graduate student at North Carolina State University, Laura Melenric with Southern Region Extension Forestry at University of Georgia for providing feedback on design elements, and the PINEMAP team and beta testers who provided many hours of dedicated testing and input to ensure the tools would be relevant to the targeted user audience.

REFERENCES

- Abatzoglou, J. T., 2013: Development of gridded surface meteorological data for ecological applications and modelling. *Int. J. Climatol.*, **33**, 121–131, <https://doi.org/10.1002/joc.3413>.
- , and T. J. Brown, 2012: A comparison of statistical downscaling methods suited for wildfire applications. *Int. J. Climatol.*, **32**, 772–780, <https://doi.org/10.1002/joc.2312>.
- Alder, J. R., and S. W. Hostetler, 2013: National Climate Change Viewer. U.S. Geological Survey, accessed 22 March 2019, <https://www2.usgs.gov/landresources/lcs/nccv/viewer.asp>.
- Allen, G. L., C. R. Miller Cowan, and H. Power, 2006: Acquiring information from simple weather maps: Influences of domain specific knowledge and general visual-spatial abilities. *Learn. Individ. Differ.*, **16**, 337–349, <https://doi.org/10.1016/j.lindif.2007.01.003>.
- Barfield, W., 1986: Expert-novice differences for software: Implications for problem-solving and knowledge acquisition. *Behav. Inf. Technol.*, **5**, 15–29, <https://doi.org/10.1080/01449298608914495>.
- Bergstrom, J. R., and A. Schall, Eds., 2014: *Eye Tracking in User Experience Design*. Elsevier, 400 pp.
- Blodgett, D., N. Booth, T. Kunicki, J. Walker, and R. Viger, 2011: Description and testing of the geo data portal: A data integration framework and web processing services for environmental science collaboration. USGS Open File Rep. 2011-1157, 9 pp., <https://pubs.usgs.gov/of/2011/1157/>.
- Boby, L., W. Hubbard, M. Megalos, and H. Morris, 2016: Southern foresters' perceptions of climate change: Implications for educational program development. *J. Ext.*, **54**, 6RIB3, https://www.joe.org/joe/2016december/pdf/JOE_v54_6rb3.pdf.
- Bojko, A., 2006: Using eye tracking to compare web page designs: A case study. *J. Usability Stud.*, **1**, 112–120.
- Breuer, N., C. Fraisse, and P. Hildebrand, 2009: Molding the pipeline into a loop: The participatory process of developing agroclimate, a decision support system for climate risk reduction in agriculture. *J. Serv. Climatol.*, **3**, 1–12, <https://doi.org/10.46275/JoASC.2009.10.001>.
- Brown, A., 1992: Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *J. Learn. Sci.*, **2**, 141–178, https://doi.org/10.1207/s15327809jls0202_2.
- Budescu, D. V., H.-H. Por, S. B. Broomell, and M. Smithson, 2014: The interpretation of IPCC probabilistic statements around the world. *Nat. Climate Change*, **4**, 508–512, <https://doi.org/10.1038/nclimate2194>.
- Canham, M., and M. Hegarty, 2010: Effects of knowledge and display design on comprehension of complex graphics. *Learn. Instr.*, **20**, 155–166, <https://doi.org/10.1016/j.learninstruc.2009.02.014>.
- Chen, S. Y., J.-P. Fan, and R. D. Macredie, 2006: Navigation in hypermedia learning systems: Experts vs. novices. *Comput. Hum. Behav.*, **22**, 251–266, <https://doi.org/10.1016/j.chb.2004.06.004>.
- Collins, A., D. Joseph, and K. Bielaczyc, 2004: Design research: Theoretical and methodological issues. *J. Learn. Sci.*, **13**, 15–42, https://doi.org/10.1207/s15327809jls1301_2.

- Deitrick, S., 2012: Evaluating implicit visualization of uncertainty for public policy decision support. *Proc. AutoCarto 2012*, Columbus, OH, Cartography and Geographic Information Society, https://www.cartogis.org/docs/proceedings/2012/Deitrick_AutoCarto2012.pdf.
- , and E. A. Wentz, 2015: Developing implicit uncertainty visualization methods motivated by theories in decision science. *Ann. Assoc. Amer. Geogr.*, **105**, 531–551, <https://doi.org/10.1080/00045608.2015.1012635>.
- Fiedler, S., and A. Glockner, 2012: The dynamics of decision making in risky choice: An eye tracking analysis. *Front. Psychol.*, **3**, 1–18, <https://doi.org/10.3389/fpsyg.2012.00335>.
- Fischhoff, B., 1996: Public values in risk research. *Ann. Amer. Acad. Pol. Soc. Sci.*, **545**, 75–84, <https://doi.org/10.1177/0002716296545001008>.
- GAO, 2009: Climate change adaptation: Strategic federal planning could help government officials make more informed decisions. U.S. Government Accountability Office, GAO-10-113, 86 pp., <https://www.gao.gov/assets/300/296526.pdf>.
- , 2015: Climate information: A national system could help federal, state, local, and private sector decision makers use climate information. U.S. Government Accountability Office, GAO-16-37, 53 pp., <https://www.gao.gov/assets/680/673823.pdf>.
- Gegenfurtner, A., E. Lehtinen, and R. Saljo, 2011: Expertise differences in the comprehension of visualizations: A meta-analysis of eye tracking research in professional domains. *Educ. Psychol. Rev.*, **23**, 523–552, <https://doi.org/10.1007/s10648-011-9174-7>.
- Hegarty, M., M. S. Canham, and S. I. Fabrikant, 2010: Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *J. Exp. Psychol. Learn. Mem. Cognit.*, **36**, 37–53, <https://doi.org/10.1037/a0017683>.
- Hmelo-Silver, C. E., S. Marathe, and L. Liu, 2007: Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. *J. Learn. Sci.*, **16**, 307–331, <https://doi.org/10.1080/1058400701413401>.
- Holmqvist, K., M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer, 2011: *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford, 560 pp.
- Jarodzka, H., K. Scheiter, P. Gerjets, and T. van Gog, 2010: In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learn. Instr.*, **20**, 146–154, <https://doi.org/10.1016/j.learninstruc.2009.02.019>.
- Kastens, K. A., T. F. Shipley, A. P. Boone, and F. Straccia, 2016: What geoscience experts and novices look at, and what they see, when viewing data visualizations. *J. Astron. Earth Sci. Educ.*, **3**, 27–58, <https://doi.org/10.19030/jaese.v3i1.9689>.
- Kinkeldey, C., A. M. MacEachren, and J. Schiewe, 2014: How to assess visual communication of uncertainty? A systematic review of geospatial uncertainty visualisation user studies. *Cartogr. J.*, **51**, 372–386, <https://doi.org/10.1179/1743277414Y.0000000099>.
- , —, M. Riveiro, and J. Schiewe, 2015: Evaluating the effect of visually represented geodata uncertainty on decision-making: Systematic review, lessons learned, and recommendations. *Cartogr. Geogr. Inf. Sci.*, **44**, 1–21, <https://doi.org/10.1080/15230406.2015.1089792>.
- Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl, 2010: Challenges in combining projections from multiple models. *J. Climate*, **23**, 2739–2758, <https://doi.org/10.1175/2009JCLI3361.1>.
- LaFrance, M., 1989: The quality of expertise: Implications of expert-novice differences for knowledge acquisition. *SIGART Bull.*, **108**, 6–14, <https://doi.org/10.1145/63266.63267>.
- Maudlin, L. C., K. S. McNeal, H. Dinon-Aldridge, C. Davis, R. Boyles, and R. M. Atkins, 2020: Website usability differences between males and females: An eye-tracking evaluation of a climate decision support system. *Wea. Climate Soc.*, **12**, 183–192, <https://doi.org/10.1175/WCAS-D-18-0127.1>.
- McNie, E., 2013: Delivering climate services: Organizational strategies and approaches for producing useful climate-science information. *Wea. Climate Soc.*, **5**, 14–26, <https://doi.org/10.1175/WCAS-D-11-00034.1>.
- Melillo, J. M., T. C. Richmond, and G. W. Yohe, Eds., 2014: *Climate Change Impacts in the United States: The Third National Climate Assessment*. U.S. Global Change Research Program, 841 pp., <https://doi.org/10.7930/J0Z31WJ2>.
- Misenheimer, C., 2015: Developing PINEMAP: An environmental risk assessment and decision-making tool for southeastern forest managers. North Carolina State University Tech. Rep., 19 pp.
- Morkes, J., and J. Nielsen, 1997: Concise, SCANNABLE, and objective: How to write for the web. Nielsen Norman Group, <https://www.nngroup.com/articles/concise-scannable-and-objective-how-to-write-for-the-web/>.
- Morris, H., M. Megalos, W. Hubbard, and L. Bobby, 2014: 2013 Climate change attitudes of southeast forestry professionals: Implications for outreach. PINEMAP Research Summary: Extension, 2 pp., http://www.pinemap.org/publications/research-summaries/extension/2013_Climate_Change_Attitudes_SE_Forestry_Professionals.pdf.
- Moser, S. C., 2010: Communicating climate change: History, challenges, process and future directions. *Wiley Interdiscip. Rev.: Climate Change*, **1**, 31–53, <https://doi.org/10.1002/wcc.11>.
- Moss, R., and Coauthors, 2010: The next generation of scenarios for climate change research and assessment. *Nature*, **463**, 747–756, <https://doi.org/10.1038/nature08823>.
- National Research Council, 1996: *Understanding Risk: Informing Decisions in a Democratic Society*. The National Academies Press, 264 pp., <https://doi.org/10.17226/5138>.
- , 2010a: *Informing an Effective Response to Climate Change*. The National Academies Press, 346 pp., <https://doi.org/10.17226/12784>.
- , 2010b: *Adapting to the Impacts of Climate Change*. The National Academies Press, 292 pp., <https://doi.org/10.17226/12783>.
- Pachauri, R. K., and Coauthors, 2014: *Climate Change 2014: Synthesis Report*. Cambridge University Press, 151 pp., https://www.ipcc.ch/site/assets/uploads/2018/02/SYR_AR5_FINAL_full.pdf.
- Patt, A., and S. Dessai, 2005: Communicating uncertainty: Lessons learned and suggestions for climate change assessment. *C. R. Geosci.*, **337**, 425–441, <https://doi.org/10.1016/j.crte.2004.10.004>.
- Pernice, K., K. Whitenon, and J. Nielsen, 2018: *How People Read Online: The Eyetracking Evidence*. Nielsen Norman Group, 412 pp., <https://www.nngroup.com/reports/how-people-read-web-eyetracking-evidence/>.
- Pierce, D. W., T. P. Barnett, B. D. Santer, and P. J. Gleckler, 2009: Selecting global climate models for regional climate change studies. *Proc. Natl. Acad. Sci. USA*, **106**, 8441–8446, <https://doi.org/10.1073/pnas.0900094106>.
- Sabatia, C. O., and H. E. Burkhart, 2014: Predicting site index of plantation loblolly pine from biophysical variables. *For. Ecol. Manage.*, **326**, 142–156, <https://doi.org/10.1016/j.foreco.2014.04.019>.

- Schmidtling, R. C., 2001: *Southern Pine Seed Sources*. U.S. Department of Agriculture, Forest Service, Southern Research Station General Tech. Rep. SRS-44, 25 pp.
- Simmons, P. E., and V. N. Lunetta, 1993: Problem-solving behaviors during a genetics computer simulation: Beyond the expert/novice dichotomy. *J. Res. Sci. Teach.*, **30**, 153–173, <https://doi.org/10.1002/tea.3660300204>.
- Somerville, R., and S. Hassol, 2011: Communicating the science of climate change. *Phys. Today*, **64**, 48–53, <https://doi.org/10.1063/PT.3.1296>.
- Spool, J. M., T. Scanlon, W. Schroeder, C. Snyder, and T. DeAngelo, 1997: *Web Site Usability: A Designer's Guide*. User Interface Engineering, 46 pp.
- The Nature Conservancy, 2009: Climate wizard. The Nature Conservancy, accessed 5 December 2018, <http://www.climatewizard.org>.
- USDA, 2012: Plant hardiness zone map. U.S. Department of Agriculture Agricultural Research Service, <http://planthardiness.ars.usda.gov>.
- U.S. Global Change Research Program, 2017: *Climate Science Special Report: Fourth National Climate Assessment*. Vol. I, U.S. Global Change Research Program, 470 pp., <http://doi.org/10.7930/J0J964J6>.
- WCRP, 2011: World Climate Research Program, coupled model intercomparison project—Phase 5 (CMIP5). *CLIVAR Exchanges Newsletter*, No. 56, International CLIVAR Project Office, Southampton, United Kingdom, 52 pp., <http://www.clivar.org/sites/default/files/documents/Exchanges56.pdf>.