Highlight Insert Colormaps: Luminance for Focused Data Analysis

Francesca Samsel[1], Trinity Overmyer[2], Paul A. Navrátil[1]

[1]The University of Texas at Austin [2]Purdue University

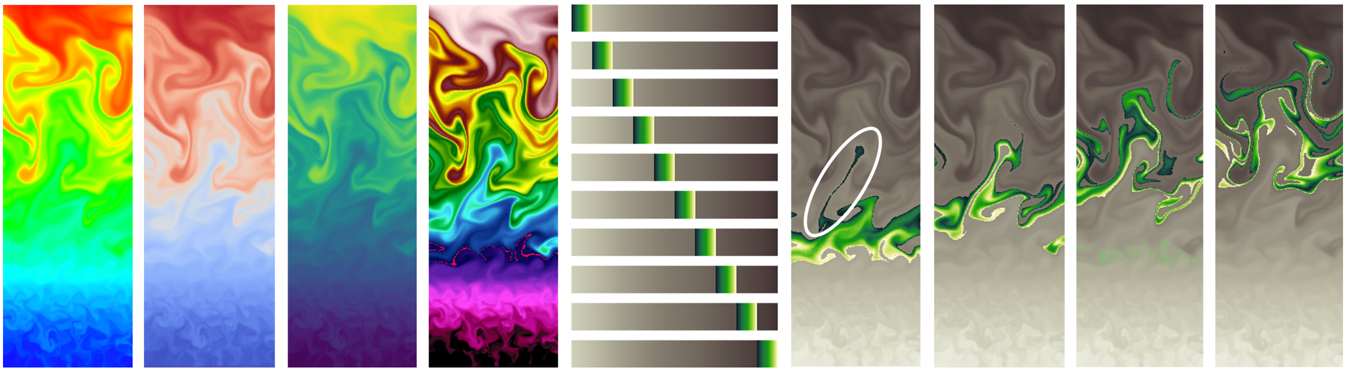

Figure 1: *Comparison of highlight insert colormaps to common colormaps: the colormaps* `rainbow`*,* `cool-warm`*,* `viridis` *and* `thermal` *applied to a turbulence region from the LANL MPAS-Ocean model [WR17b]. Highlight insert colormaps (HICs), such as the set of ten colormaps that highlight successive 10% subsets of the data range* (center)*, emphasize a fine band of features while still providing muted context in the remaining regions. Applying a sequence of HICs reveal structures not easily seen in the full-range maps* (right)*, such as the filament structure circled in panel 6. HICs maintain clear expression of the turbulence in the gray regions at top and bottom.*

Abstract

Color provides the primary conduit through which we extract insight from data visualizations. As the dynamic range of data grows, extracting salient features from surrounding context becomes increasingly challenging. Default colormaps provided by visualization software are poorly suited to perform such reductions of visual data. Here we present sets of highlight insert colormaps *(HICs) that provide scientists with the means to quickly and easily render a detailed overview of their data, create detailed scans of their data, and examine the outer ranges of data in detail. This method builds on the long understood discriminatory power of luminance and in the highlight region provides* $3\times$ *to* $10\times$ *the discriminative power of common colormaps.*

1. Introduction

As the types and dimensions of scientific data continue to grow, new methods of locating and examining data detail within areas of interest are needed. Colormaps are the primary means for visually translating discrete and scalar data into a cognitively digestible form. Extensive research has been conducted into understanding the perceptual properties of colormaps for general data representation. Traditionally, colormaps are designed to render all areas of the data in equal focus. As data grows in size, the cost of equal allocation of contrast increases. Large-scale data analysis presents challenges that strain the capabilities of general colormaps: wide dynamic range (blurring features over a general colormap range), small feature size relative to global data extent (losing features in a map with insufficient discriminatory power), and dense feature arrangement (obscuring features due to color simultaneity). Such data can benefit from colormaps that *visually reduce* data to focus analysis and to highlight data sections for further processing (e.g. highlight data for extraction or decimation).

Our research examines the specific challenges facing scientists with extremely large data who need to comprehensively examine their data. Similar to how an adaptive mesh refinement (AMR) simulation increases data resolution in regions of greatest interest, we examine colormaps with focused discriminatory power allocated in a narrow band overlaid on a linear grayscale map that provides complimentary discrimination outside the highlight region. Our HIC maps enable a scientist to focus the colormap on a precise data region (highlight insert) while preserving context for global analysis (grayscale). Just as AMR allocates computational resources to simulation regions most likely to generate insight, our insert colormap system concentrates the highest discriminatory power to re-

DOI: 10.2312/evs.20191170

gions of most interest. Our HIC maps provide a means and workflow for exploiting the power of luminance, the strongest, most intuitive channel for conveying scalar data [WWP13, ZH16]. These colormaps act as a "scanning microscope" enabling scientists to quickly identify critical features for further analysis. Using both a statistical analysis of colormap perceptual power and domain scientist interviews, we demonstrate how our HICs provide significantly more discriminative power than commonly-used colormaps available in visualization software. Note that we are *not* evaluating colormaps either by task or perceptual properties, nor are we advocating our highlight insert colormaps as general defaults. The work presented here is designed for scientists with high dynamic range data who need to *visually reduce* data in order to see feature detail. Figure 1 illustrates this principle, including the "scanning microscope" capability enabled by a set of our maps: a scientist can easily iterate through the colormap set to explore their data and identify regions of interest. The contributions of this paper are (1) to demonstrate the power of condensing luminance within colormaps for large data scientific visualization and (2) to provide scientists means to incorporate such methods into their existing workflows to facilitate exploration and knowledge extraction.

2. Prior Work

Much effort has been invested in creating effective colormapping solutions for scientific visualization. Areas of exploration include evaluating discriminatory power, uniformity and speed [RT96, BTS*18, WTS*17, WTB*18, Mor09, Rhe00]. Solutions such as isoluminant colormaps (colormaps that advance at regular hue intervals keeping luminance constant) such as `viridis` and `plasma` are popular [LH18, ZH16] but there has been little previous work into how to the power of luminance itself might be exploited as in HICs [ZH16]. Segmentation colormaps, used to classify structures in biomedical imaging, use isoluminant color blocks to distinguish structure (e.g. components of a cell), and to our knowledge do not consider luminance in color selection. In contrast, our HICs are constructed with a focused high-contrast color region with rapid luminance ramp *combined* with a linear grayscale to provide muted but distinguishable data context to the focus region (see Figure 2).

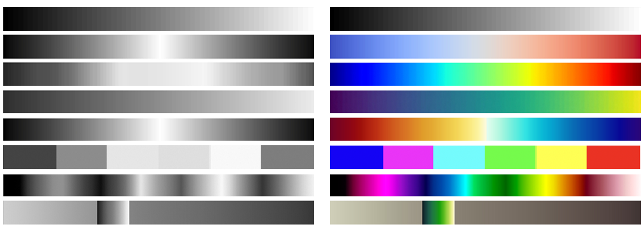

Figure 2: *Luminance distributions of commonly used colormaps, top to bottom:* `grayscale`*,* `cool-warm`*,* `rainbow`*,* `viridis`*,* `blue-orange divergent`*,* `blot` *from ParaView,* `thermal`*; and an example of our ten-percent HIC. The HIC provides a high contrast at the insert region while still providing contextual distribution equivalent to* `grayscale` *outside of the highlight region.*

Harnessing Luminance — While the `rainbow` colormap carries metric information well (quantitative tasks), Ware [War88] demonstrates how it fails at conveying form (qualitative tasks). Testing by Ware and Turton confirmed the shortcomings not only of `rainbow` but also of `cool-warm` and `viridis` colormaps, specifically in middle ranges. Samsel's blue-orange divergent, shown in Figure 2, slowed narrowly in the mid-range but provided statistically significant higher speed in all areas [TWSR17, WTB*18]. Lui and Heer recently found that `viridis` and `linear blue` colormaps are superior in speed and accuracy over `jet` and `blue-orange divergent` colormaps [LH18], however their study focused on quantitative comparisons more typically found in information visualization (e.g. comparing the color of two nodes in a shaded connected graph) rather than qualitative feature determination common in scientific visualization.

Rogowitz, Ware and others [WTS*17, RK01, ZH16] have established that luminance is the most powerful color channel able to differentiate detailed information, and data that is small in size [War12]. Recently, Ware has confirmed the primary role of luminance in feature detection, specifically in large datasets with high dynamic range [WTB*18]. All scientists we interviewed confirmed that the availability of extra luminance distribution would significantly benefit their science, which motivates this work.

Workflows — Colormap generation tools [HB03, NAS, MJSK15, RO86, SKR18] all require significant breaks from data analysis to tune the colormap. The scientist must: (1) interrupt their workflow within visualization software; (2) move to the colormap generating software; (3) design the map; (4) export it in an acceptable format; and then (5) import back into the visualization software, hoping that it will meet their needs. Since most scientists lack formal training in color theory, this process will likely be repeated multiple times and will ultimately yield suboptimal results. While some colormapping tools enable scientists to test the new colormaps directly on their data, the interruption in the workflow makes it impractical. In contrast, our highlight insert colormaps can be directly incorporated into workflows and provide immediate value without modification (e.g., data scanning, data narrowing, visual emphasis) within existing visualization platforms.

3. Highlight Insert Colormaps

Using the principles of artistic and perceptual color theory, particularly the discriminative power of luminance, we have designed sets of *highlight insert* colormaps (HIC) that contain precisely what the name implies: within a neutral linear color scale, we embed another color scale within a narrow data range (e.g. 10% or 20%) that both provides a hue-spanning saturated color scale to highlight data and contrasts with the muted color scale of the underlying map to extend discriminative power. The center of Figure 1 shows a set of ten HICs that span the data range in 10% increments. Such a set can be used to quickly scan data for interesting features (Figure 1, right).

The left column of Figure 3 shows luminance distributions of common colormaps and an example of our highlight insert maps with this contrasting luminance clearly visible. None of the common colormaps pack as much luminance variance over as narrow a band. Note that our colormaps are for scientists who need to see detail and nuance within narrow ranges of their data. They are not suitable for use as default colormaps where analysis priority is speed and uniformity [BTS*18, WTS*17, WTB*18, War12]. The comparison shown in Figure 1 illustrates the application of HIC to reveal greater detail within specific data regions.

HICs serve to narrow the expressed range of a data variable. Once loaded into a visualization framework, scientists can explore detailed regions of their data by sequentially moving through an HIC set to highlight particular ranges of interest. The righ side of Figure 1, shows the application of a focus colormap set in 10% of the data. We have generated HIC sets where the focal points cover 10%, 20%, and 33.3% of the data, which enables a scientist to select a set with the level of detail appropriate to their exploration needs. These maps are avaialble at `SciVisColor.org`.

4. Evaluation

To evaluate our colormaps, we employed a mathematical analysis [WTB*18]to quantify the information conveyance power of HICs compared to commonly used colormaps; and we conducted a set of informal scientist interviews to qualify our colormap performance over the status quo.

Color Speed Formula — We evaluate the discriminatory power of our structured colormaps against a sampling of common colormaps using Ware et al.'s weighted CIELAB formula for measured contrast sensitivity [WTB*18], which has been shown to provide a good approximation for human perception. This formula was derived from perceptual studies Ware performed to measure colormap discriminatory power [WTS*17]. To apply the formula, first translate each colormap from RGB space to CIELAB space and then sample 30 color values uniformly distributed across the map (the same sapling rate used by Ware). Then calculate the color difference ΔE for each sample interval Δs by computing the Euclidean distance between sample points in weighted CIELAB space:

$$\Delta E = \sqrt{(\Delta L^*)^2 + (w_a \Delta a^*)^2 + (w_b \Delta b^*)^2} \qquad (1)$$

where ΔL^* expresses luminance change between samples, and Δa^* and Δb^* express change in the red-green and blue-yellow CIELAB channels, respectively. We set $w_a = w_b = 0.1$, which provides the best fit to observed user data [WTB*18]. Use the color difference to compute Ware's contrast sensitivity over each interval:

$$c = 3.4(\Delta E/\Delta s)^{0.879} \qquad (2)$$

then plot the resulting discriminative for each colormap.

Results from Ware's original study [WTB*18] are shown in the left-hand graph of Figure 3. The colormaps represented are: `blue orange divergent` (BOD), `cool-warm` (CW), `rainbow` (R) `thermal` (TH), `viridis` (VI) and `grayscale` (Gray). We separate VI and Gray for clarity and to emphasize comparison to HICs. This graph clearly shows the significantly lower resolving power of the most commonly used colormaps (`cool-warm` and `rainbow`) especially in the middle ranges. Our highlight insert colormaps have 3× to 10× the resolving power in the highlight regions, and they maintain consistent resolving power across the neutral linear underlay region without the central drop of `cool-warm` and `rainbow`. While `thermal` has high discriminatory power, it creates objectionable color artifacts in high-frequency data (see Figure 1). We note that `viridis` and `grayscale` exhibit relatively uniform dicriminatory power across their range, comparable to the non-highlight region of HICs.

Observed and Simulated vs. Synthetic Data — Most colormap validation experiments use synthetic data as this provides an equal distribution able to be analytically assessed. However, this approach masks the complexities present in simulated and observed data. Our discussions with scientists has brought to light several needs which are not covered by the current types of colormaps assessment, mainly how effective is a colormap on large data with narrow areas of interest. Figure 1 illustrates the issue by demonstrating the application of common colormaps to real data. Even the thermal colormap, which demonstrates high discriminatory power on synthetic data [WTB*18], can obscure features in high-frequency data; such features are revealed by our system due to the focused highlight region.

Scientist Interviews — In order to understand the benefits and drawbacks of structured colormaps in scientific visualization, we interviewed domain scientists and visualization experts from Los Alamos National Laboratory (LANL). For each scientist, we rendered their data using ParaView [AGL05] in a range of static HIC colormaps downloaded from `SciVisColor.org`. We report on feedback from these conversations, which centered around colormaps as facets of their workflows and the ways in which these new luminance structures are generative tools for data insight in a range of domains and tasks. The information distilled here comes from senior LANL personnel, including: Dr. Phillip Wolfram, Research Scientist in the Climate Ocean-Sea Ice Modeling team [WR17a, WRM*15, WR17b]; Dr. Nicole Jeffrey, Research Scientists in the Climate Ocean-Sea Ice Modeling team [EJH*17, JEHL16]; Dr. Joseph Schmidt, Senior Scientist at LANL; Li-Ta Lo, Software Engineer in the Computer, Computational and Statistical Sciences group specializing in data visualization; Dr. Boonthanome Nouanesengsy, Staff Scientist in the Data Science at Scale team; and Dr. Erica Fogerty, Computational Physicist in the Center for Theoretical Astrophysics. Overall, conversations revealed that the luminance structures were elegant solutions to complex problems scientists had failed to solve using more time-consuming ways. Once scientists were able to see the structured colormaps overlaid on data, the solution seemed to be an obvious one that had previously been overlooked. According to these interviews, the structured colormaps provided three primary benefits: they offer an alternative to visualizing data with a logarithmic scale, they provide a highly detailed view of boundaries and outliers, and they were useful for both early exploratory phases of research and later phases when data structures have already been identified.

Benefits — *An alternative to logarithmic scale:* While most scientists often turn to a log scale to look at highly concentrated areas, Jeffrey found the colormap with contrast concentrated on one end useful for seeing both the difference and the minimum and maximum edges of her data. "The benefits of this color structure are that you could look at the data in a regular scale, and not have to scale it logarithmically, and it will give you more resolution in the regions of interest. This structure helps me scale things out and it offers the resolution needed to explore more. I don't currently have any tools like that." The luminance structures, being concentrated and aligned with dense data ranges, helps scientists see more detail without visualizing it on a logarithmic scale. Schmidt, whose team studies pressure distributions and density movement during shock charges, was able to see structures using these luminance distributions that he had not been able to differentiate previously.

Enhanced visual detail: The structures offer a level of visual detail that scientists have not been able to see with default color maps. For Dr. Phillip Wolfram, the maps illuminated critical structures, not seen in the four years of exploring this specific data (see Figure 4). "These are actually interesting because they are showing

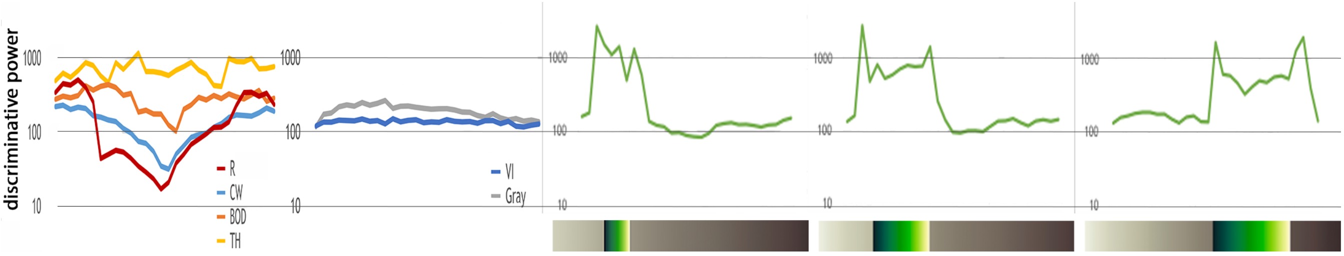

Figure 3: *Ware's weighted CIELAB formula for measured contrast sensitivity [WTB*18] (log scale). The graphs at left plot the discriminatory power of: R - `rainbow`; CW - `cool-warm`; BOD - `blue orange divergent`; TH - `thermal`; VI - `viridis`; and `grayscale`. While `thermal` has high discrimination power, it creates objectionable color artifacts in high-frequency data (see Figure 1); `cool-warm` and `rainbow` suffer a distinct power drop through their middle ranges. Our linear green inset colormaps with insets covering 10, 20 and 30 percent of the colormaps contain $3\times$ to $10\times$ peak resolving power than the standards while maintaining discriminatory power similar to `viridis` and `grayscale` outside the highlight regions. See Figure 2 for the comparison colormaps and their respective luminance.*

filament structures of the mixing at different times. I have never been able to see this filament structure before. I didn't even know the question concerning the filaments was here until I saw this visualization. But now that I see it, I know the colormap is useful." Some of the challenges of visualizing his data have been in finding appropriate colors where the eddies mix. In previous visualizations, the team used white lines on a blue colormap to denote mixing patters, but solid lines denote more structured and definite boundaries than actually exist. However, the colormap set shown in Figure 4 allowed him to zoom in on these areas and look at the various boundaries and mixing levels, rather than a single, solid line. According to his work with research groups at the lab, Dr. Nouanesengsy reported that, like Dr. Wolfram, the majority of users want to look at boundaries between parameters or the differences that occur when various values change, and yet they primarily begin with the ParaView default colormap (`cool-warm`, Figure 2 map 2) which "muddies" the intermediate data ranges, so the boundaries can be difficult to differentiate. While linear structures are useful for smaller datasets, they do not have the resolving power needed for large data, particularly in the central ranges of the colormap. Often when users find that `cool-warm` does not work, Nouanesengsy noted, users get frustrated and return to the `rainbow` map (e.g., Figure 2 map 3), even when they are aware of its biases. Our map sets provide an easily-used alternative to both the `cool-warm` and `rainbow` maps.

Useful at multiple research phases: Researchers identified two situations when the luminance distributions would be particularly fruitful. First, Jeffrey mentioned that in the early phases of research, the structures would help her identify areas of importance, because they allow her to navigate through targeted areas in more detail. It is often in these early phases when scientists need tools that make exploration more efficient and generative. Alternatively, Wolfram was most interested in how the luminance structures perform in the later phases of research, after he is already familiar with the data structures: "If you know there's some sort of physical threshold in your data, you can use luminance structures to show the differences in behavior. That's where this becomes really useful—when you are trying to see something in detail that you already know is there. The system shown in Figure 4 lets you continuously scan the data, to see features in detail, in every time step and isolate it purposefully. When you are trying to pick something specific out, then the structure is great." He also noted their value to illustrate data narratives, as the luminance structures emphasize specific parameters and interactions in the data for outside audiences.

5. Conclusion and Future Work

Our highlight insert colormaps provide scientists with a simple, easily applied, and powerful tool to visually reduce and emphasize data for initial exploration, for comparative analysis, and for communication of findings to broader audiences. Our maps provide both exceptional discriminative power ($3\times$ to $10\times$ that of commonly used maps) in the highlight region while preserving global context and avoiding regions of significant power loss (present in the middle region of common maps). Our maps are available for download today and can be easily loaded into common visualization tools. Many of the scientists interviewed commented that they would like to be able to slide the points at which the luminance scale changes to align with changes in their data. This feature is the foundation of colormap construction tools [MJSK15, SKR18], but it is only achievable in outside of visualization software, thus interrupting workflow and requiring effort far beyond practical usage.

HICs are not designed to be application defaults. They are designed for scientists with large, highly-dynamic data sets needing detailed views within their data. Other situations where we recommend caution using HICs: when the colormap legend cannot be displayed clearly; if an equally weighted representation of the data across the entire dataset is required; if the changes in luminance scales fall in locations that cause distortion or misdirection of attention. There are other issues to be addressed which are beyond the scope of this paper, such as: the role and placement of saturation within the structured colormaps; the impact of attention hierarchies; the effects of feature size on structured map selection; the perceptual interactions between hues; and automatic selection and tuning of an HIC to match a certain feature range. We will address these aspects in future work.

The authors thank: Dr. Colin Ware for his guidance and usage of his colormap evaluation techniques; Dr. Phillip Wolfram, LANL, for the DOE BER sponsored MPAS-Ocean data and visualization. This work is funded in part by the US DOE Office of Science ASCR program (Dr. Laura Biven, Program Manager) and by an Intel Corporation IPCC award.

References

[AGL05] AHRENS J., GEVECI B., LAW C.: *ParaView: An End-User Tool for Large Data Visualization*. Visualization Handbook, Elsevier, 2005. 3

[BTS*18] BUJACK R., TURTON T. L., SAMSEL F., WARE C., ROGERS D. H., AHRENS J.: The good, the bad, and the ugly: A theoretical framework for the assessment of continuous colormaps. *IEEE transactions on visualization and computer graphics 24*, 1 (2018), 923–933. 2

[EJH*17] ELLIOTT S., JEFFERY N., HUNKE E. C., DEAL C., JIN M., WANG S., SMITH E. E., OESTREICHER S.: Strategies for the simulation of sea ice organic chemistry: Arctic tests and development. *Geosciences 7*, 52 (2017). 3

[HB03] HARROWER M., BREWER C.: Colorbrewer.org: An online tool for selecting colour schemes for maps. *The Cartographic Journal 40*, 1 (2003), 27–37. 2

[JEHL16] JEFFERY N., ELLIOTT E. C., HUNKE W. H., LIPSCOMB A. T.: Biogeochemistry of cice: The los alamos sea ice model documentation and software user's manual, 2016. 3

[LH18] LIU Y., HEER J.: Somewhere over the rainbow: An empirical assessment of quantitative colormaps. In *ACM Human Factors in Computing Systems (CHI)* (2018). 2

[MJSK15] MITTELSTÄDT S., JÄCKLE D., STOFFEL F., KEIM D. A.: Colorcat: Guided design of colormaps for combined analysis tasks. In *Proc. of the Eurographics Conference on Visualization (EuroVis 2015: Short Papers)* (2015), vol. 2. 2, 4

[Mor09] MORELAND K.: Diverging color maps for scientific visualization. In *Proceedings of the 5th International Symposium on Advances in Visual Computing, Part II* (2009), ISVC '09, pp. 92–103. 2

[NAS] NASA: Nasa color tool website. https://colorusage.arc.nasa.gov/ColorTool.php. 2

[Rhe00] RHEINGANS P.: Task-based color scale design. In *28th AIPR Workshop: 3D Visualization for Data Exploration and Decision Making* (2000), International Society for Optics and Photonics, pp. 35–43. 2

[RK01] ROGOWITZ B., KALVIN A.: The "which blair project": A quick visual method for evaluating perceptual color maps. In *Proceedings Visualization 2001* (2001), VIS'01, p. 183ff. 2

[RO86] ROBERTSON P. K., O'CALLAGHAN J. F.: The generation of color sequences for univariate and bivariate mapping. *IEEE Computer Graphics and Applications 6*, 2 (1986), 24–32. 2

[RT96] ROGOWITZ B., TREINISH L.: How not to lie with visualization. *Computers in Physics 10*, 3 (1996), 268–273. 2

[SKR18] SAMSEL F., KLAASSEN S., ROGERS D. H.: Colormoves: Real-time interactive colormap construction for scientific visualization. *IEEE Computer Graphics and Applications 38*, 1 (Jan./Feb. 2018), 20–29. 2, 4

[TWSR17] TURTON T., WARE C., SAMSEL F., ROGERS D.: A crowdsourced approach to colormap assessment. In *EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization (EuroRV3)* (2017), Lawonn K., Smit N., Cunningham D., (Eds.), The Eurographics Association. 2

[War88] WARE C.: Color sequences for univariate maps: theory, experiments and principles. *Computer Graphics and Applications, IEEE 8*, 5 (1988), 41–49. 2

[War12] WARE C.: *Information Visualization: Perception for Design*, 3rd ed. Morgan Kaufman, San Francisco, CA, 2012. 2

[WR17a] WOLFRAM P. J., RINGLER T. D.: Computing eddy-driven effective diffusivity using lagrangian particles. *Ocean Modelling 118* (2017), 94–106. 3

[WR17b] WOLFRAM P. J., RINGLER T. D.: Quantifying residual, eddy, and mean flow effects on mixing in an idealized circumpolar current. *Journal of Physical Oceanography 47*, 8 (2017), 1897–1920. 1, 3

[WRM*15] WOLFRAM P. J., RINGLER T. D., MALTRUD M. E., JACOBSEN D. W., PETERSEN M. R.: Diagnosing isopycnal diffusivity in an eddying, idealized midlatitude ocean basin via lagrangian, in situ, global, high-performance particle tracking (light). *Journal of Physical Oceanography 45*, 8 (2015), 2114–2133. 3

[WTB*18] WARE C., TURTON T. L., BUJACK R., SAMSEL F., SHRIVASTAVA P., ROGERS D. H.: Measuring and modeling the feature detection threshold functions of colormaps. *IEEE Transactions on Visualization and Computer Graphics (*to appear*)* (2018). 2, 3, 4

[WTS*17] WARE C., TURTON T. L., SAMSEL F., BUJACK R., ROGERS D. H.: Evaluating the perceptual uniformity of color sequences for feature discrimination, June 2017. Accepted into EuroRV3: International EuroGraphics Workshop on Reproducibility, Verification, and Validation in Visualization. 2, 3

[WWP13] WARE C., WRIGHT W., PIOCH N. J.: Visual thinking design patterns. *Distributed Multimedia Systems Proceedings* (2013), 150–155. 2

[ZH16] ZHOU L., HANSEN C.: A survey of colormaps in visualization. *IEEE Transactions on Visualization and Computer Graphics 22*, 8 (2016), 2051–2069. 2