Collaborative Exploration of Scientific Datasets using Immersive and Statistical Visualization

Preprint

Nicholas Brunhart-Lupo,[1] Brian Bush,[1] Kenny Gruchalla,[1]
Kristen Potter,[1] and Steve Smith[2]

1 National Renewable Energy Laboratory
2 Los Alamos Visualization Associates

Prepared for the SEA's Improving Scientific Software Conference 2020

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-2C00-76783
September 2020

Collaborative Exploration of Scientific Datasets using Immersive and Statistical Visualization

Preprint

Nicholas Brunhart-Lupo,[1] Brian Bush,[1] Kenny Gruchalla,[1] Kristen Potter,[1] and Steve Smith[2]

1 National Renewable Energy Laboratory
2 Los Alamos Visualization Associates

Suggested Citation
Brunhart-Lupo, Nicholas, Brian Bush, Kenny Gruchalla Kristen Potter, and Steve Smith. 2020. *Collaborative Exploration of Scientific Datasets using Immersive and Statistical Visualization: Preprint*. Golden, CO: National Renewable Energy Laboratory. NREL/CP-2C00-76783. https://www.nrel.gov/docs/fy20osti/76783.pdf

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-2C00-76783
September 2020

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

Collaborative Exploration of Scientific Datasets using Immersive and Statistical Visualization

NICHOLAS BRUNHART-LUPO, BRIAN BUSH, KENNY GRUCHALLA, and KRISTIN POTTER,

National Renewable Energy Laboratory

STEVE SMITH, Los Alamos Visualization Associates

We discuss the value of collaborative, immersive visualization for the exploration of scientific datasets and review techniques and tools that have been developed and deployed at the National Renewable Energy Laboratory (NREL). We believe that collaborative visualizations linking statistical interfaces and graphics on laptops and high-performance computing (HPC) with 3D visualizations on immersive displays (head-mounted displays and large-scale immersive environments) enable scientific workflows that further rapid exploration of large, high-dimensional datasets by teams of analysts. We present a framework, PlottyVR, that blends statistical tools, general-purpose programming environments, and simulation with 3D visualizations. To contextualize this framework, we propose a categorization and loose taxonomy of collaborative visualization and analysis techniques. Finally, we describe how scientists and engineers have adopted this framework to investigate large, complex datasets.

Additional Key Words and Phrases: datasets, collaborative visualization, statistical graphics, scientific workflow

1 INTRODUCTION

Immersive visualization is poised to advance analysis for certain classes of complex scientific and engineering data, through rapid advances in virtual reality (VR) and augmented reality (AR). In our daily usage of the large-scale immersive virtual environment at the National Renewable Energy Laboratory (NREL), we have observed how immersive visualizations enhance scientific workflows [27]. Many scientists and engineers across NREL are beginning to adopt commodity AR/VR head-mounted displays (HMDs). While these immersive displays may provide unique qualitative insights, they are relatively limited for quantitative examinations. To deepen our analyses, we have been blending statistical interfaces and statistical graphics on traditional displays with 3D visualizations in immersive displays, providing a collaborative visualization framework with the objective of gaining the best of both worlds.

Collaboration has long been named one of the grand challenges for visual analytics [63]. There has been significant research into distributed and co-located visualization and sense-making [5, 30, 41]. However, this collaborative visualization research has focused on a *shared* visual representation, either synchronously or asynchronously, among analysts. We are specifically interested in supporting workflows where analysts collaborate through heterogeneous imagery, as shown in Figure 1. The views into the data may be different, tailored to each analyst, but the views are directly linked together. For example, we describe use cases where some analysts have immersive views of a dataset, while other analysts have more quantitative statistical views of that same dataset. When one analyst introduces derived data or flags regions of interest, those actions are visible and available in all the views.

2 BACKGROUND

Increases in computational power has lead to growth in the size and complexity of scientific datasets. There has been a corresponding growth in both need and opportunity for teams, diverse in terms of location and expertise, to extract

Authors' addresses: Nicholas Brunhart-Lupo, nicholas.brunhart-lupo@nrel.gov; Brian Bush, brian.bush@nrel.gov; Kenny Gruchalla, kenny.gruchalla@nrel.gov; Kristin Potter, kristi.potter@nrel.gov, National Renewable Energy Laboratory, 15013 Denver West Pky, Golden, CO, 80401; Steve Smith, sas@lava3d.com, Los Alamos Visualization Associates, 3 Bundy Rd, Otowi, NM, 87506.

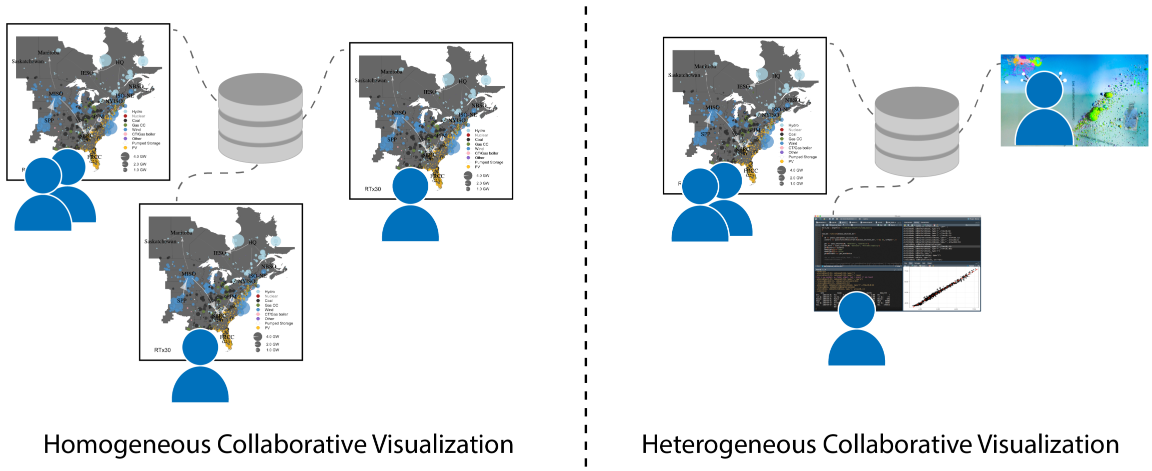

Fig. 1. Left: In traditional collaborative visualization users share common imagery. Right: We propose linking different types of imagery of the same data, facilitating analysis by providing to stakeholders visualizations customized to their tasks and expertise, while directly linking and coordinating the data and actions of the team.

usable knowledge from these sets. The technical progress in collaborative visualization has been likewise advancing; however, the challenges in creating compelling, useful, and accessible visualization environments hold ubiquitous collaborative visualization still at arm's length. These obstacles include technical challenges relating to factors like a limitation of network bandwidth [29], trade-offs between centralized or disparate data management [61], and the targeting of different types of display hardware. Several surveys have addressed the evolution of these challenges throughout the years [10, 19, 21, 26, 30, 57]; however, many of the software tools and environments have become outdated. Similarly, there has been research on human issues relating to how people operate in immersive spaces [69] and how they use shared visualization [68].

Experimentation and practice at NREL have revealed myriad considerations for designing workflows to explore datasets using collaborative, immersive visualization. A primary concern is whether collaborators are physically present in the same room and, if so, whether they jointly use the same immersive display or individually use separate displays. If they jointly use a display, then the collaborators may share the point of view with a single, privileged collaborator or have their own points of view for the shared display. (Multiple points of view for the same display pose technical challenges for walk-in immersive spaces, but not for individual AR or VR headsets.) Furthermore, with the joint use of a display, either a single collaborator or multiple collaborators might have control, being able to manipulate the view with user-input devices such as gamepads, joysticks, gloves, or wands. For display hardware that supports multiple points of view, each collaborator can be presented with a customized rendering of the model. For instance, some collaborators might view a simpler rendering of the dataset whereas others might see richer renderings that include uncertainty information or high-density details, but the basic geometric skeleton of the renderings likely would be common to all of the collaborators.

In situations where collaborators do not jointly use the same display, each collaborator might see and manipulate radically different renderings and subsets of the underlying raw dataset. For example, some renderings might be scatterplots, others might be maps, and still others might be abstract statistical graphics. Selections or highlighting of data records by one collaborator might automatically propagate to the views of the other collaborators; this linkage of data records across views might be total, targeted, or absent. Collaborators might work on subsets of the data, each having their own scene graph. The combination of disparate views with linked selections or highlights enables simultaneous, coordinated exploration among collaborators probing different aspects of the same dataset, essentially combining the insightfulness of the team and opening dialogs scrutinizing patterns, correlations, and hypotheses regarding the dataset.

Immersive collaboration can extend beyond the confines of a single room and a particular time. Moreover, recording, annotation, and bookmarking of visualization sessions allows for their revisiting and replay at later times, perhaps by different subsets of the collaborative team. These techniques further extend the exploratory dialog beyond single sessions and allow supplemental "offline", asynchronous exploration between the immersive sessions. Moreover, collaborators might be geographically distributed, so all of the aforementioned interactions can occur at distant locations.

Real-time statistical and data-analytic tools can supplement and drive immersive explorations. When collaborators find interesting paradigmatic features or form hypotheses regarding some aspect of a dataset, an analyst at a workstation or laptop might locate additional interesting features based on the paradigm or perform tests of the hypothesis, respectively, "pushing" the statistical results into the immersive display for scrutiny by the other collaborators. One might sometimes have several collaborators working in immersive spaces and several others working in statistical applications such as R, Python, or Julia, but with all of the collaborators' work linked in a common information space where results propagate among renderings. This approach combines the power of statistics and machine learning with the intuitive interactivity of immersive visualization.

2.1 Taxonomy

While a full taxonomy of work in collaborative visualization is out of scope for this publication, we present a loose classification of literature in the area to help establish the focus of previous works in this field and highlight directions for future work. We have broken our sampling of previous work into eight categories: location, imagery, viewpoint, data sources, number of display devices, type of display devices, interaction, and epoch. These categories are meant to summarize aspects of collaborative visualization and may not reflect details of those aspects.

The first category in our taxonomy is *location*, which refers to where participants are joining the collaboration. This category is loosely divided into remote and colocated participation. However, there has been a collection of research that focuses on mixed-presence [33], that is, support for both remote and in-person participation. Many of the challenges for remote and colocated participation are quite disparate: remote participation requires high-bandwidth, low-latency networking in models where data and imagery are shared from a centralized server, or for users to have high-powered computers to do that processing at each remote location. With the advances in new web-based technologies, some of these concerns are being alleviated; however, for large-scale scientific computing, these issues are still paramount. For colocated users, more prevalent issues include how to facilitate multiple user inputs and interactions, how to physically locate users in a space, and how to target imagery for different user tasks in a single location. There has also been work on representing remote participants as virtual avatars [59]. However, we did not capture that work in this taxonomy as the challenges to using avatars include technical issues related to motion capture of remote users and representational challenges of including avatars in a visualization environment, as well as issues relating to human interactions with avatars and other human factors. While this work is very relevant to collaborative visualization, the broad scope of such work is outside the focus of this paper.

The second category of our taxonomy is *imagery*: what the users are seeing while using a system. We classify the imagery as either *constant*, that is all users see the same representation of data, or *targeted* where imagery is designed for particular types of users or tasks. For example, the SAGE and SAGE2 frameworks [52, 61] aim at providing a workspace for multiple users to present windows from their individual machines to a powerwall. Thus, users can look at the same data but design disparate visualizations that investigate different aspects of that data and share those visualizations on a single powerwall. Similarly, systems could target volume rendering type displays for users whose main interest is the 3D aspects of the data, while also providing statistical charts and graphs for analysts looking at general 1 and 2D summaries of a data set.

Next, we look at *viewpoint*, which defines where in the scene a user is looking. The two classes in this category are *privileged* and *independent*. Privileged viewpoints constrain all users to see what the privileged user sees (WISIWYS—"what I see is what you see") with no support for multiple view points. Independent viewpoints allow users to individually explore the data, and oftentimes share their viewpoint with collaborators. Thus, one can imagine privileged viewing of a a volume rendering that is similar to a guided tour, versus allowing a user to actually navigate around the rendering. The challenges in disparate viewpoints include the computational expense of those different rendering viewpoints as well as how to indicate disparate viewpoints across all collaborators.

Interaction is related to viewpoint in that support for independent viewpoints often provides support for *asynchronous* interactions. However, asynchronous interactions may also refer to actions taken on a data set such as filtering, etc., and may be reflected back to all users in a privileged viewpoint environment, possibly through a queue of interactions or through sharing of a viewpoint. *Synchronous* interactions can be thought of as passing the baton such that only one user is performing an interaction and all other users are essentially an audience.

The *data sources* category considers whether the virtual environment provides methods for looking at different data sources simultaneously. Homogeneous environments support a single data source, while heterogeneous environments can handle disparate data sets. For example, a materials scientist may want to look at a volume rendering of a proposed molecule, whereas a techno-economist may want to look at cost sheets and graphs for the new material. The visualizations that support multiple data sets may actually use constant imagery across all users or can target visualizations for data set or user type. NREL is actively looking at how to support these sorts of heterogeneous data sets in collaborative spaces. The main challenges include how to visually connect disparate data sets and how to manage these data sets at scale.

The *number of displays* category simply defines how many display devices are targeted by a system. For example, is the collaboration environment a single cave-like immersive environment or powerwall, or is there a network of display devices, either remotely or co-located? Much of the early work on immersive displays focused on hardware and software issues in the design of single displays and not on the collaborative aspects of such spaces, and thus much of that work is out of scope of this taxonomy. However, the utility of collaboration quickly shifted the need to connect multiple users and thus multiple display devices. This category is not solely captured by the location category, specifically because multiple displays may be used in both remote and in-person collaborations.

The *type of devices* category is separated as its own category to capture targeting homogeneous or heterogeneous display types. For example, connecting two immersive cave-like displays is homogeneous, whereas connecting HMDs and laptops is a heterogeneous environment. This type of situation is becoming more prevalent as HMDs, table-top displays, and other state-of-the-art displays are becoming more main stream.

Finally, the last category, *epoch* relates to when a user is collaborating within a system. Most often, *realtime* collaboration is what is thought of in terms of immersive spaces, where people gather at the same time and supplement technological collaboration with personal interactions such as speech. However, there is also a group of work looking at *playback* collaborations where users can use the collaborative system independently, and share their work with others.

Year	Author	Title
1994	Anupam et al.	Distributed and collaborative visualization [1]
1995	Pang et al.	CSpray: A collaborative scientific visualization application [47]
	Wood et al.	CSCV-computer supported collaborative visualization [66]
1996	Rantzau et al.	Collaborative and interactive visualization in a distributed high performance software environment [50]
1997	Pang et al.	Collaborative 3D visualization with CSpray [48]
2000	Brewer et al.	Collaborative geographic visualization: Enabling shared understanding of environmental processes [9]
	Childers et al.	Access grid: Immersive group-to-group collaborative visualization [17]
2004	Koyamada et al.	VizGrid: collaborative visualization grid environment for natural interaction between remote researchers [35]
2005	Casera et al.	A collaborative extension of a visualization system [13]
2006	Song et al.	Paraview-based collaborative visualization for the grid [59]
2007	Ryu et al.	Collaborative object-oriented visualization environment [54]
2010	Dupont et al.	Collaborative scientific visualization: The COLLAVIZ framework [22]
	Filho et al.	An immersive and collaborative visualization system for digital manufacturing [21]
2015	Li et al.	High performance heterogeneous computing for collaborative visual analysis [36]

Table 1. Collaborative visualization taxonomy of systems with default configurations. Specifically, systems in this table support remote locations, constant imagery, privileged viewpoints, simultaneous interactions, homogeneous data sources and display types, multiple display devices, and a realtime epoch

The categorization of our taxonomy is fluid in that some categories are clearly disparate while others are closely related, but broken out based on findings in the literature. Some of this is due to the maturation of the technologies behind collaborative environments. Hardware challenges of the late 1990s, including networking and compute bandwidth, have, for the most part, been solved and replaced with challenges of new and novel display technologies, as well as expectations regarding speed of interactions and targeted support of certain devices, such as mobile. We expect this taxonomy to evolve and perhaps combine or divide categories along with the advances enabling this technology.

Tables 1 and 2 lay out a selection of related publications that target collaborative visualization systems. These works, for the most part, look at software systems to support collaboration, and papers discussing the use of these software for specific applications have not been included. Tables 1 and 2 are divided by the most common and novel categorizations. The most common configurations, shown in Table 1, consists of systems that support collaborations with remote locations, constant imagery, privileged viewpoints, simultaneous interactions, homogeneous data sources and display types, multiple display devices, and a realtime epoch. Many of these systems are older and reflect the state of the art for mid-1990s to mid-2000s systems. Table 2 shows a taxonomy for novel configurations that vary across all categories. Many of these systems reflect innovations in computational power and bandwidth, visualization hardware and software, and ways of using collaborative environments. Future work for this taxonomy includes researching applications in which collaboration systems have been used and understanding the success rate of those efforts. In addition, it will be useful to understand which systems support newer rendering clients (see Figure 2) such as virtual and augmented reality goggles, touch-panels, and web browsers.

Year	Author	Title
1997	Wood et al.	Collaborative visualization [65]
1998	Raje et al.	CEV: collaborative environment for visualization using Java RMI [49]
2000	Sawant et al.	The tele-immersive data explorer: A distributed architecture for collaborative interactive visualization of large data-sets [57]
2004	Benko et al.	Collaborative mixed reality visualization of an archaeological excavation [4]
	Borgeat et al.	Collaborative visualization and interaction for detailed environment models [7]
	Ellis et al.	A collaborative annotation system for data visualization [23]
	Pan et al.	A peer-to-peer collaborative 3D virtual environment for visualization [46]
2005	Chastine et al.	A collaborative multi-view virtual environment for molecular visualization and modeling [15]
	Mavris et al.	Advanced design of complex systems using the collaborative visualization environment (CoVE) [41]
2006	Taesombut et al.	Collaborative data visualization for earth sciences with the OptIPuter [61]
2007	Marchese et al.	Fostering asynchronous collaborative visualization [38]
2008	Bourke	Evaluating Second Life as a tool for collaborative scientific visualization [8]
	Klasky et al.	Collaborative visualization spaces for petascale simulations [34]
2009	Nakasone et al.	Astrosim: Collaborative visualization of an astrophysics simulation in second life [43]
	Renambot et al.	Enabling high resolution collaborative visualization in display rich virtual organizations [51]
2010	Kim et al.	Hugin: A framework for awareness and coordination in mixed-presence collaborative information visualization [32]
2011	Chen et al.	Supporting effective common ground construction in Asynchronous Collaborative Visual Analytics [16]
2012	Galambos et al.	VirCA NET: A case study for collaboration in shared virtual space [24]
	Marion et al.	Real-time collaborative scientific WebGL visualization with WebSocket [39]
2014	Badam et al.	PolyChrome: A Cross-Device Framework for Collaborative Web Visualization [2]
	Guo et al.	A collaborative large spatio-temporal data visual analytics architecture for emergence response [27]
2016	Klapperstuck et al.	ContextuWall: Peer Collaboration Using (Large) Displays [33]
	Su et al.	High-Resolution Interactive and Collaborative Data Visualization Framework for Large-Scale Data Analysis [60]
2019	Badam et al.	Vistrates: A Component Model for Ubiquitous Analytics [3]
	Zimmer	Guided Interaction and Collaborative Exploration in Heterogeneous Network Visualizations [69]

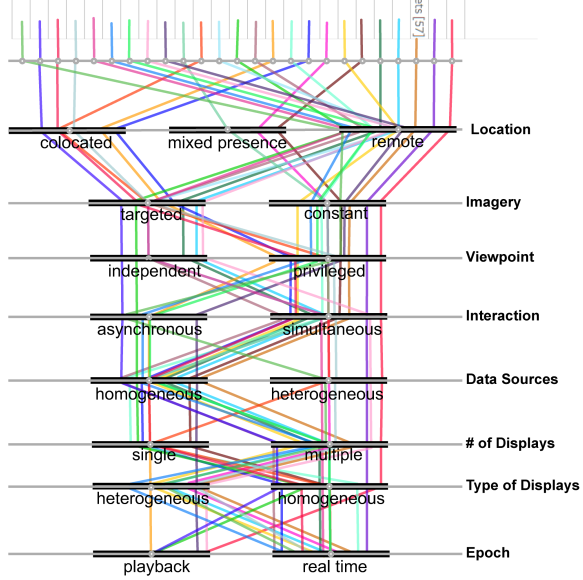

Table 2. Parallel Collaborative visualization taxonomy of novel system configurations.

Fig. 2. Example 3D rendering clients: a) web browser with WebVR support; b) virtual reality (VR) headset; c) NREL's immersive visualization environment; d) augmented reality (AR) headset. *(c) Photo by Dennis Schroeder, NREL 97821, (d) Photo by Dennis Schroeder, 34387*

3 PLOTTYVR

PlottyVR represents a collection of tools to facilitate collaborative visualization. In terms of our taxonomy, this framework provides real-time, synchronous colocated or remote collaboration, with distinct imagery, distinct viewpoints, homogeneous data sources, on multiple heterogeneous displays. The primary objective is to combine the qualitative reasoning and intuition of immersive displays with the quantitative power of programming languages like R and Python. The toolset provides a bidirectional link between the Python or R programming language environments and an immersive 3D visualization, supporting both VR HMDs and large-scale, walk-in immersive environments. Using this software package, an analyst at a laptop can push data into an immersive visualization environment. The representation of that data can take the form of primitives, like points or line segments, or of more advanced objects like images and text. Developing immersive visualizations mirrors plot development in base R graphics [43] but with a third dimension. For example, R users can instantiate a 3D scatter plot in a connected immersive display by issuing a plot command with three arrays *x,y,z* for each data point, optionally providing point colors and diameters (scaling each point in cardinal directions to create ellipsoids). The immersive users can then interact with the visualization by selecting or querying data points. In our scatter plot example, textual annotations can also be linked to points so that the immersive users can, using this immersive system's interaction device, "click" on a point to quickly query relevant information, like a record identifier or notes about that data point. The immersive users can select regions of interest, which immediately become available to the R users for further quantitative analysis of the selected cluster.

One of our primary guiding design principle was to reduce the friction of incorporating an immersive space into statistical analyses. We have noticed that when the cost (even just the perception of cost) of using VR or other technologies is too high (whether that be financial, intellectual, or technical), researchers will not make use of it. This principle itself is expressed in several forms: (1) Critically, we cannot demand that analysts discard their existing workflows. Because of this, we chose to avoid a stand-alone, foundational framework. A common complaint for existing visualization solutions, such as ParaView or other large software packages, is that the researcher must then answer the question "how do I get my data and decision flow *into* there?" and re-engineer their entire pipeline. The user must now figure out the acceptable formats for the package, possibly write scripts to translate data, and then figure out how to extract features in the data and translate that back into discoveries. This is in addition to the challenges of learning to navigate an immersive space. In light of this, we chose to integrate with the existing analysis pipelines, i.e., to augment, not replace. (2) This augmentation must also be friction-free. Even if the package does not require replacing a researcher's workflow, it can still be excessively intrusive; researchers may quickly become disaffected when required to spend tens of minutes setting up a run-time environment on a server, verifying connections, and copying data for five minutes of exploration. Turnaround and setup times must be kept to a minimum.

To meet the aforementioned requirements, we provide PlottyVR to users as a library for R, Python, and Julia[1]. All that a researcher must do is to download and install these libraries for the environment of their choice, using the language's standard package manager[2], and issue a plot command. If, for example, the user is at a cave-like installation (such as the one at NREL), the server will already be running and the time from setup to use is on the order of seconds. If they are on their own private system, there is the additional step of launching the server first (automatic launching is currently being explored).

These libraries connect via WebSocket to the immersive platform. To reduce friction further, we also provide the most common plot types in a form that mirror the base graphics in R (see Fig. 3). 3D scatter plots and 3D line plots are direct

[1]These are the most common environments in use at NREL. There is, however, no technical limitation to supporting other languages or systems.
[2]These packages are currently internal-only. They will be published in the future; see Section 5.

```r
# Load and connect to PlottyVR
plotty_init()

# Plot the data
plot_tech("Coal Capacity [GW]", "Wind Capacity [GW]", "Utility Solar Capacity [GW]", ids=ids[1:70])

#--------------------------------------------------------------------------
#' plot_tech
#'
#' @description An immersive plot to compare real vs predicted of three
#'            technologies. Real trajectories are plotted as lines,
#'            and the error between real and pred is represented as
#'            an ellipsoid, scale by the error in the three cardinal
#'            directions.
#'
#' @param techx Metric string for x-axis
#' @param techy Metric string for y-axis
#' @param techz Metric string for z-axis
#' @param ids vector of run ids to plot
#'
tech_pts = function(techx, techy, techz, ids=unqiue(test$ID))
{
  xlim = range(c(data$Real[data$Metric==techx], data$Pred[data$Metric==techx]))
  ylim = range(c(data$Real[data$Metric==techy], data$Pred[data$Metric==techy]))
  zlim = range(c(data$Real[data$Metric==techz], data$Pred[data$Metric==techz]))

  # initialize the plot axes
  iplot(c(0),c(0),c(0),xlim=xlim,ylim =ylim,zlim=zlim,xlab=techx,ylab=techy,zlab=techz)

  plot_traj = function(ID)
  {
    col =randomcoloR::randomColor()
    xr = data$Real[data$ID==ID & data$Metric==techx]
    yr = data$Real[data$ID==ID & data$Metric==techy]
    zr = data$Real[data$ID==ID & data$Metric==techz]
    xp = abs(xr-data$Pred[data$ID==ID & data$Metric==techx])/(xlim[2]-xlim[1])
    yp = abs(yr-data$Pred[data$ID==ID & data$Metric==techy])/(ylim[2]-ylim[1])
    zp = abs(zr-data$Pred[data$ID==ID & data$Metric==techz])/(zlim[2]-zlim[1])

    id = ipoints(xr,yr,zr,col=col) # plot the points
    iscale(xp/2,yp/2,zp/2, id) # scale the points into ellipsoids
    ilines(xr,yr,zr,w_size=0.002,col=col) # draw the trajectories
  }

  invisible(lapply(ids, plot_traj))
}
```

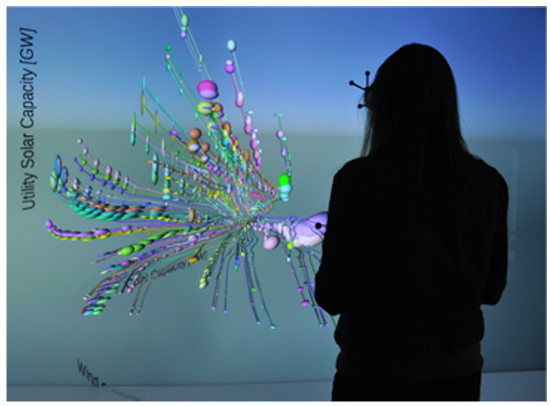

Fig. 3. R-code listing that generates the inset immersive visualization.

and straightforward. From just the point (an ellipsoid in 3D) and the line (a tube in 3D) primitives, we can construct a wide variety of complex plot types, such as plot trajectories and 3D parallel coordinate plots, such as parallel planes [12]. We support images and text, providing context like geospatial plots and custom annotations. These API function calls are transmitted to a server application, running an interface to the graphics engine that is on the immersive platform. For the immersive environment at NREL, this engine is *Isopach*, a custom immersive scene graph library. For HMDs, we have developed a server application in the Unity engine. The server library is tasked with transforming the WebSocket messages to the engine's graphics representation, and relaying selection and other manipulation back to the client-side.

By integrating into environments in this way, not only are researchers given easy access to immersive plotting, but they are also able to make use of the space as a primitive for larger compositions. An example of this is that researchers can link together the immersive space and Shiny R [54] webpages, either for using a tablet as a companion in large immersive environments to furnish supplemental 2D plots and provide more accessible control over what is being seen, or for providing views for other remote analysts. In a more operational context, streaming data can be pulled in and joined with the immersive space for real-time analysis.

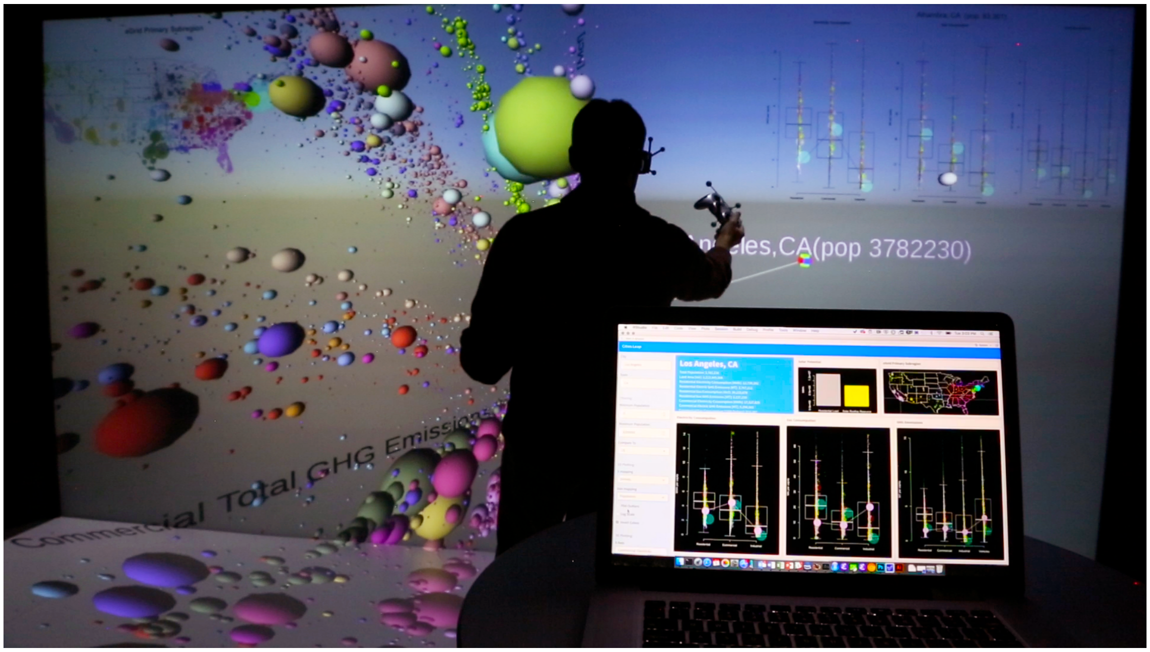

Fig. 4. Cities-LEAP analysis exploring a high-dimensional data set of city energy profiles. Foreground: real-time analysis in R-based Shiny web application. Background: immersive visualization.

4 USE CASES

Using this toolset, we have achieved workflows such as synchronous, collaborative hypothesis testing where a statistician pushes data into the VR space, a scientist constructs a hypothesis by manipulating plots in that space, and the statistician applies statistical tests to the hypothesis, all in real-time. This workflow combines the power of statistical analysis with the insightfulness of rich, multidimensional data visualizations, and we have used this workflow to support both static datasets and on-demand simulations. PlottyVR has been used at NREL to meaningfully explore multidimensional time-series in as many as twenty dimensions by teams of as many as six persons, allowing collaborative identification of features, anomalies, and patterns in datasets that would be difficult and tedious to explore in 2D displays that limit collaborative interaction. Workflows combining real-time statistical analysis in R-based Shiny web applications [54] with immersive visualization by PlottyVR have been the most popular, allowing rapid, interactive statistical exploration of the output of complex dynamic simulation models. We briefly describe three examples, but PlottyVR continues to be employed regularly for varied applications at NREL.

The Cities Leading through Energy Analysis (Cities-LEAP) [64] project has developed city energy profiles for over 23,400 U.S. Cities. We combined an R-based Shiny dashboard that allowed cities to be compared on a range of metrics with immersive 3D scatter plots that revealed correlations, trends, stratification, and outliers in the data (see Figure 4). An analyst seated at the desktop would choose the metrics of interest, specifying the three axes, color, and size for the scatter plot. A second analyst in the immersive environment would identify and select cities or clusters of interest. Those selections would automatically update on the R desktop. The two analysts would iterate, hypothesizing on relationships, generating new metrics, and applying those metrics to the scatter plot to test the hypotheses.

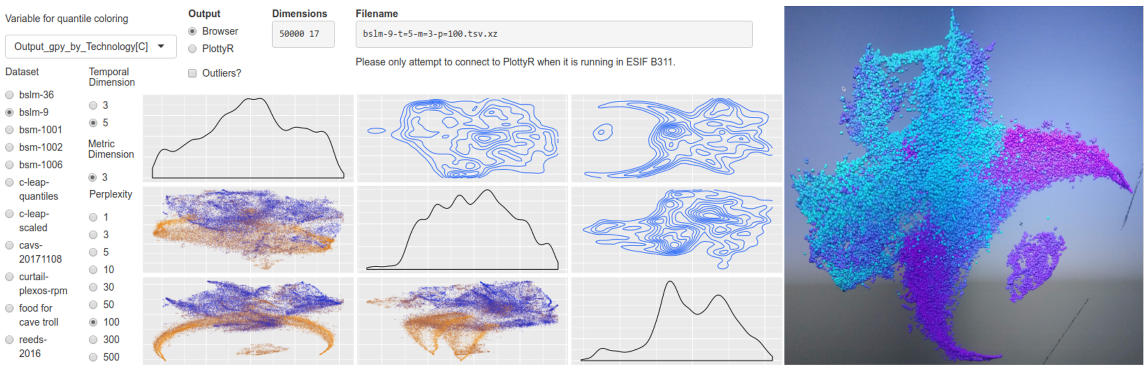

Fig. 5. R-based Shiny web application (left) and immersive visualization environment (right) for controlling and exploring self-organized maps of ensembles of simulation output [13].

In another application, the R program running on an analyst's laptop pushes ensembles of simulation output to the immersive environment displayed as 3D scatter plots. Working in the immersive space, analysts use a handheld tool to select regions of interest on the scatterplot and send that selection back to R on the laptop. The R program then performs Monte-Carlo filtering [56], which is a sensitivity analysis technique that infers which input variables most significantly influence whether data records fall inside versus outside of the selected region. The web application then displays (sometimes via a standard 2D projector on a wall next to the immersive environment) a ranked list of input parameters sorted in order of their influence on the output in the selected region. This enables rapid, interactive sensitivity analysis on ensembles of simulations: users can formulate, test, and discard hypotheses sequentially, gradually refining their understanding of the correlations and influences of input parameters on output results over ensembles of simulations. Interspersed with the hypothesis testing, analysts use the system to verify simulation behavior by scrutinizing input-output relationships and trends that might be present in the model.

We have deployed similar applications that combine Shiny web applications with immersive visualizations of self-organized maps (SOMs) of ensembles of high dimensional simulation output [13]. The web interface allows analysts to manipulate the parameters of the dimension-reduction algorithm and to view 2D projections of the SOMs on their laptops, but simultaneously to push the SOMs into 3D renderings in the immersive environment (see Figure 5). The user with the laptop can be in a location distant from the immersive environment. For this application, we have developed a parallel coordinates plotting WebVR client, specifically supporting multiple, simultaneous remote collaborators to view and jointly manipulate the same 3D scene of SOMs of simulation output.

5 FUTURE WORK

The most immediate improvement for the PlottyVR system is the publishing of server and client libraries to public-facing package repositories.

While PlottyVR has regularly proven useful, there are some limitations to the current implementation and protocol. First, the protocol originally was designed for a single client and a single server; while useful for a single researcher or a small team working in the same room, this has proven limiting in common situations at NREL, such as distributed analysis where many participants are not in the same locale (i.e. remote locations in the nomenclature of the taxonomy), and so is sub-optimal for multi-client or multi-server collaborative configurations. In order to share data outside of the single client and server model, the researcher must build their own methods of distributing data, updating that data, or coordinating clients or servers. Further, it was initially envisioned that only the server would be graphically based, and that there would be a library client.

In response to these limitations, NOODLES (NREL Object Oriented Data Layout and Exploration System) is a work-in-progress replacement for the protocol and data representation presently used in PlottyVR, updating the model to a single server with multiple clients. It functions closer to the scene-graph level, where a synchronized 3D scene can be shared across many clients with support for customizations for different form factors, as well as providing database-like access to the protocol to support clients[3] of any form factor (a more heterogeneous environment). Note that this would not change the interface for the user; they would still be supplied with a library to provide a low-friction path to simple plotting. Other clients can now join the session and register simple callbacks or software hooks to automatically watch for data changes from collaborators or add their own data and plots to the mix. Additional generalized primitives permit more applications to be developed beyond the domain of simple 3D plotting and statistics, all while still fulfilling the PlottyVR goals. An initial demonstration of the system can be found in [11].

ACKNOWLEDGMENTS

This work was authored in part by the National Renewable Energy Laboratory (NREL), operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. This work was supported by the Laboratory Directed Research and Development (LDRD) Program at NREL. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

A portion of This research was performed using computational resources sponsored by the Department of Energy's Office of Energy Efficiency and Renewable Energy and located at the National Renewable Energy Laboratory.

REFERENCES

[1] V. Anupam, C. Bajaj, D. Schikore, and M. Schikore. 1994-07. Distributed and collaborative visualization. *Computer* 27, 7 (1994-07), 37–43.

[2] Sriram Karthik Badam and Niklas Elmqvist. 2014-11-16. PolyChrome: A Cross-Device Framework for Collaborative Web Visualization. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces* (Dresden, Germany) *(ITS '14)*. Association for Computing Machinery, 109–118.

[3] Sriram Karthik Badam, Andreas Mathisen, Roman Rädle, Clemens N. Klokmose, and Niklas Elmqvist. 2019. Vistrates: A Component Model for Ubiquitous Analytics. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (2019), 586–596.

[4] Hrvoje Benko, Edward W. Ishak, and Steven Feiner. 2004. Collaborative mixed reality visualization of an archaeological excavation. In *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*. IEEE, 132–140.

[5] Mark Billinghurst, Maxime Cordeil, Anastasia Bezerianos, and Todd Margolis. 2018. Collaborative Immersive Analytics. In *Immersive Analytics*, Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H. Thomas (Eds.). Springer International Publishing, Cham, 221–257.

[6] Mark Billinghurst, Maxime Cordeil, Anastasia Bezerianos, and Todd Margolis. 2018. Collaborative Immersive Analytics. In *Immersive Analytics*. Springer International Publishing, 221–257.

[7] Louis Borgeat, Guy Godin, Jean-François Lapointe, and Philippe Massicotte. 2004. Collaborative visualization and interaction for detailed environment models. In *10th International Conference on Virtual Systems and Multimedia*.

[8] Paul Bourke. 2008. Evaluating Second Life as a tool for collaborative scientific visualization. *Computer Games and Allied Technology* 29 (2008), 2008.

[9] Isaac Brewer, Alan M. MacEachren, Hadi Abdo, Jack Gundrum, and George Otto. 2000. Collaborative geographic visualization: Enabling shared understanding of environmental processes. In *IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings*. IEEE, 137–141.

[3]It is up to the developer of the client library to determine how best to present this data for that platform. They can select 3D geometry, tabular representations, plots, or any mixture of these.

[10] K. W. Brodlie, D. A. Duce, J. R. Gallop, J. P. R. B. Walton, and J. D. Wood. 2004. Distributed and Collaborative Visualization. 23, 2 (2004), 223–251.

[11] Nicholas Brunhart-Lupo. 2020. NOODLES Demo. https://www.youtube.com/watch?v=qddOHC_WHr0. Accessed: 6-24-2020.

[12] N. Brunhart-Lupo, B. W. Bush, K. Gruchalla, and S. Smith. 2016. Simulation exploration through immersive parallel planes. In *2016 Workshop on Immersive Analytics (IA)*. 19–24.

[13] Bruce Bugbee, Brian W. Bush, Kenny Gruchalla, Kristin Potter, Nicholas Brunhart-Lupo, and Venkat Krishnan. 2019. Enabling immersive engagement in energy system models with deep learning. *Statistical Analysis and Data Mining: The ASA Data Science Journal* 12, 4 (2019), 325–337.

[14] Steve Casera, H.-H. Nageli, and Peter Kropf. 2005. A collaborative extension of a visualization system. In *First International Conference on Distributed Frameworks for Multimedia Applications*. IEEE, 176–182.

[15] Tom Chandler, Maxime Cordeil, Tobias Czauderna, Tim Dwyer, Jaroslaw Glowacki, Cagatay Goncu, Matthias Klapperstueck, Karsten Klein, Kim Marriott, Falk Schreiber, and Elliot Wilson. 2015. Immersive Analytics. In *2015 Big Data Visual Analytics (BDVA)*. 1–8.

[16] J.W. Chastine, Ying Zhu, J.C. Brooks, G.S. Owen, R.W. Harrison, and I.T. Weber. 2005-07. A collaborative multi-view virtual environment for molecular visualization and modeling. In *Coordinated and Multiple Views in Exploratory Visualization (CMV'05)*. 77–84.

[17] Yang Chen, Jamal Alsakran, Scott Barlowe, Jing Yang, and Ye Zhao. 2011. Supporting effective common ground construction in Asynchronous Collaborative Visual Analytics. In *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*. 101–110.

[18] Lisa Childers, Terry Disz, Robert Olson, Michael E. Papka, Rick Stevens, and Tushar Udeshi. 2000. *Access grid: Immersive group-to-group collaborative visualization*. Technical Report. Argonne National Lab., IL (US).

[19] Kristian Sons Christophe Mouton and Ian Grimstead. [n.d.]. Collaborative visualization: current systems and future trends. In *Proceedings of the 16th International Conference on 3D Web Technology* (2011). 101–110.

[20] Maxime Cordeil, Tim Dwyer, Karsten Klein, Bireswar Laha, Kim Marriott, and Bruce H. Thomas. 2017. Immersive Collaborative Analysis of Network Connectivity: CAVE-style or Head-Mounted Display? *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 441–450.

[21] Ciro Donalek, S. George Djorgovski, Alex Cioc, Anwell Wang, Jerry Zhang, Elizabeth Lawler, Stacy Yeh, Ashish Mahabal, Matthew Graham, and Andrew Drake. 2014. Immersive and collaborative data visualization using virtual reality platforms. In *2014 IEEE International Conference on Big Data (Big Data)*. IEEE, 609–614.

[22] Nelson Duarte Filho, Silvia Costa Botelho, Jonata Tyska Carvalho, Pedro de Botelho Marcos, Renan de Queiroz Maffei, Rodrigo Remor Oliveira, Rodrigo Ruas Oliveira, and Vinicius Alves Hax. 2010. An immersive and collaborative visualization system for digital manufacturing. *The International Journal of Advanced Manufacturing Technology* 50, 9 (2010), 1253–1261.

[23] Florent Dupont, Thierry Duval, Cedric Fleury, Julien Forest, Valérie Gouranton, Pierre Lando, Thibaut Laurent, Guillaume Lavoue, and Alban Schmutz. 2010. Collaborative scientific visualization: The COLLAVIZ framework.

[24] Sean E. Ellis and Dennis P. Groth. 2004. A collaborative annotation system for data visualization. In *Proceedings of the working conference on Advanced visual interfaces*. 411–414.

[25] Peter Galambos, Christian Weidig, Peter Baranyi, Jan C. Aurich, Bernd Hamann, and Oliver Kreylos. 2012. VirCA NET: A case study for collaboration in shared virtual space. In *2012 IEEE 3rd International Conference on Cognitive Infocommunications (CogInfoCom)*. 273–277.

[26] I.J. Grimstead, D.W. Walker, and N.J. Avis. 2005. Collaborative visualization: a review and taxonomy. In *Ninth IEEE International Symposium on Distributed Simulation and Real-Time Applications*. 61–69.

[27] Kenny Gruchalla and Nicholas Brunhart-Lupo. 2019. The Utility of Virtual Reality for Science and Engineering. In *VR Developer Gems*, William R. Sherman (Ed.). Taylor Francis, Chapter 21, 383–402.

[28] D. Guo, J. Li, H. Cao, and Y. Zhou. 2014. A collaborative large spatio-temporal data visual analytics architecture for emergence response. *IOP Conference Series: Earth and Environmental Science* 18 (2014), 012129.

[29] Andrei Hutanu, Gabrielle Allen, Stephen D. Beck, Petr Holub, Hartmut Kaiser, Archit Kulshrestha, Milos Liska, Jon MacLaren, Ludek Matyska, and Ravi Paruchuri. 2006. Distributed and collaborative visualization of large data sets using high-speed networks. *Future Generation Computer Systems* 22, 8 (2006), 1004–1010.

[30] Petra Isenberg, Niklas Elmqvist, Jean Scholtz, Daniel Cernea, Kwan-Liu Ma, and Hans Hagen. 2011. Collaborative visualization: Definition, challenges, and research agenda. *Information Visualization* 10, 4 (2011), 310–326.

[31] Timothy Jacobs and Sean Butler. 2001. Collaborative visualization for military planning. In *Java/Jini Technologies*, Vol. 4521. International Society for Optics and Photonics, 42–51.

[32] Nils Jensen, Stefan Seipel, Wolfgang Nejdl, and Stephan Olbrich. 2003. COVASE: Collaborative visualization for constructivist learning. In *Designing for Change in Networked Learning Environments*. Springer, 249–253.

[33] KyungTae Kim, Waqas Javed, Cary Williams, Niklas Elmqvist, and Pourang Irani. 2010. Hugin: A framework for awareness and coordination in mixed-presence collaborative information visualization. In *ACM International Conference on Interactive Tabletops and Surfaces*. 231–240.

[34] Matthias Klapperstuck, Tobias Czauderna, Cagatay Goncu, Jaroslaw Glowacki, Tim Dwyer, Falk Schreiber, and Kim Marriott. 2016. ContextuWall: Peer Collaboration Using (Large) Displays. In *2016 Big Data Visual Analytics (BDVA)*. 1–8.

[35] Scott Klasky, Roselyne Barreto, Ayla Kahn, Manish Parashar, Norbert Podhorszki, Steve Parker, Deborah Silver, and Mladen A. Vouk. 2008. Collaborative visualization spaces for petascale simulations. In *2008 International Symposium on Collaborative Technologies and Systems*. IEEE, 203–211.

[36] V. Ryuichi Matsukura V. Koji Koyamada, V. Yasuo Tan, and V. Yukihiro Karube V. Mitsuhiro Moriya. 2004. VizGrid: collaborative visualization grid environment for natural interaction between remote researchers. *FUJITSU Sci. Tech. J* 40, 2 (2004), 205–216.

[37] Jianping Li, Jia-Kai Chou, and Kwan-Liu Ma. 2015. High performance heterogeneous computing for collaborative visual analysis. In *SIGGRAPH Asia 2015 Visualization in High Performance Computing (SA '15)*. Association for Computing Machinery, 1–4.

[38] Thomas Ludwig, Tino Hilbert, and Volkmar Pipek. 2015. Collaborative visualization for supporting the analysis of mobile device data. In *ECSCW 2015: Proceedings of the 14th European Conference on Computer Supported Cooperative Work, 19-23 September 2015, Oslo, Norway*. Springer, 305–316.

[39] Francis T. Marchese and Natasha Brajkovska. 2007. Fostering asynchronous collaborative visualization. In *2007 11th International Conference Information Visualization (IV'07)*. IEEE, 185–190.

[40] Charles Marion and Julien Jomier. 2012-08-04. Real-time collaborative scientific WebGL visualization with WebSocket. In *Proceedings of the 17th International Conference on 3D Web Technology (Web3D '12)*. Association for Computing Machinery, 47–50.

[41] Roberto Martinez-Maldonado, Judy Kay, Simon Buckingham Shum, and Kalina Yacef. 2019. Collocated Collaboration Analytics: Principles and Dilemmas for Mining Multimodal Interaction Data. *Human-Computer Interaction* 34, 1 (Jan. 2019), 1–50.

[42] Dimitri Mavris, Patrick Biltgen, and Neil Weston. 2005. Advanced design of complex systems using the collaborative visualization environment (CoVE). In *43rd AIAA Aerospace Sciences Meeting and Exhibit*. 126.

[43] Paul Murrell. 2011. *R Graphics* (2nd ed.). CRC Press, Inc., USA.

[44] Arturo Nakasone, Helmut Prendinger, Simon Holland, Piet Hut, Jun Makino, and Ken Miura. 2009. Astrosim: Collaborative visualization of an astrophysics simulation in second life. *IEEE Computer Graphics and Applications* 29, 5 (2009), 69–81.

[45] Dorit Nevo, Saggi Nevo, Nanda Kumar, Jonas Braasch, and Kusum Mathews. 2015. Enhancing the Visualization of Big Data to Support Collaborative Decision-Making. In *2015 48th Hawaii International Conference on System Sciences*. 121–130. ISSN: 1530-1605.

[46] Jasminko Novak and Michael Wurst. 2005. Collaborative knowledge visualization for cross-community learning. In *Knowledge and Information Visualization*. Springer, 95–116.

[47] Yi Pan and Francis T. Marchese. 2004. A peer-to-peer collaborative 3D virtual environment for visualization. In *Visualization and Data Analysis 2004*, Vol. 5295. International Society for Optics and Photonics, 180–188.

[48] Alex Pang and Craig Wittenbrink. 1997. Collaborative 3D visualization with CSpray. *IEEE Computer Graphics and Applications* 2 (1997), 32–41. Publisher: IEEE.

[49] Alex Pang, Craig M. Wittenbrink, and Tom Goodman. 1995. CSpray: A collaborative scientific visualization application. In *Multimedia Computing and Networking 1995*, Vol. 2417. International Society for Optics and Photonics, 317–326.

[50] Rajeev R. Raje, Michael Boyles, and Shiaofen Fang. 1998. CEV: collaborative environment for visualization using Java RMI. *Concurrency: Practice and Experience* 10, 11 (1998), 1079–1085.

[51] D. Rantzau, U. Lang, R. Lang, H. Nebel, A. Wierse, and R. Ruehle. 1996. Collaborative and interactive visualization in a distributed high performance software environment. In *High Performance Computing for Computer Graphics and Visualisation*. Springer, 207–216.

[52] Luc Renambot, Byungil Jeong, Hyejung Hur, Andrew Johnson, and Jason Leigh. 2009. Enabling high resolution collaborative visualization in display rich virtual organizations. *Future Generation Computer Systems* 25, 2 (2009), 161–168.

[53] Nathalie Henry Riche, Kori Inkpen, John Stasko, Tom Gross, and Mary Czerwinski. 2012. Supporting asynchronous collaboration in visual analytics systems. In *Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI '12)*. Association for Computing Machinery, 809–811.

[54] RStudio, Inc. 2013. *Easy web applications in R.* URL: http://www.rstudio.com/shiny/.

[55] So-Hyun Ryu, Hyung-Jun Kim, Jin-Sung Park, Yong-won Kwon, and Chang-Sung Jeong. 2007. Collaborative object-oriented visualization environment. *Multimedia Tools and Applications* 32, 2 (2007), 209–234.

[56] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. 2008. *Global sensitivity analysis: the primer.* John Wiley & Sons.

[57] Ali Sarvghad, Narges Mahyar, and Melanie Tory. 2009. History tools for collaborative visualization. *Collaborative Visualization on Interactive Surfaces-CoVIS'09* (2009), 21.

[58] Nikita Sawant, Chris Scharver, Jason Leigh, Andrew Johnson, Georg Reinhart, Emory Creel, Suma Batchu, Stuart Bailey, and Robert Grossman. 2000. The tele-immersive data explorer: A distributed architecture for collaborative interactive visualization of large data-sets. In *Proceedings of the Fourth International Immersive Projection Technology Workshop*. 1–16.

[59] Ralph Schroeder and Ann-Sofie Axelsson. 2006. *Avatars at work and play: Collaboration and interaction in shared virtual environments*. Vol. 34. Springer Science & Business Media.

[60] Guanghua Song, Yao Zheng, and Hao Shen. 2006. Paraview-based collaborative visualization for the grid. In *Asia-Pacific Web Conference*. Springer, 819–826.

[61] Simon Su, Vincent Perry, Nicholas Cantner, Dylan Kobayashi, and Jason Leigh. 2016-10. High-Resolution Interactive and Collaborative Data Visualization Framework for Large-Scale Data Analysis. In *2016 International Conference on Collaboration Technologies and Systems (CTS)*. 275–280.

[62] Nut Taesombut, Xinran Ryan Wu, Andrew A. Chien, Atul Nayak, Bridget Smith, Debi Kilb, Thomas Im, Dane Samilo, Graham Kent, and John Orcutt. 2006. Collaborative data visualization for earth sciences with the OptIPuter. *Future Generation Computer Systems* 22, 8 (2006), 955–963.

[63] James J. Thomas and Kristin A. Cook (Eds.). 2005. *Illuminating the Path: The Research and Development Agenda for Visual Analytics.* National Visualization and Analytics Ctr, Los Alamitos, Calif.

[64] U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy 2018. *Cities-LEAP City Energy Profiles.* U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy. https://apps1.eere.energy.gov/sled.

[65] Andreas Wierse. 1995. Collaborative visualization based on distributed data objects. In *Workshop on Database Issues for Data Visualization*. Springer, 208–219.
[66] Jason Wood, Helen Wright, and Ken Brodie. 1997. Collaborative visualization. In *Proceedings. Visualization'97)*. IEEE, 253–259.
[67] Jason Wood, Helen Wright, and Ken Brodlie. 1995. CSCV-computer supported collaborative visualization. In *Proceedings of BCS Displays Group International Conference on Visualization and Modelling*. Citeseer, 13–25.
[68] Noráin Mohd Yusoff and Siti Salwah Salim. 2015. A systematic review of shared visualisation to achieve common ground. *Journal of Visual Languages & Computing* 28 (2015).
[69] Shanyang Zhao. 2003. Toward a Taxonomy of Copresence. *Presence: Teleoperators & Virtual Environments* 12 (2003), 445–455.
[70] Björn Zimmer. 2019. *Guided Interaction and Collaborative Exploration in Heterogeneous Network Visualizations*. Ph.D. Dissertation. Linnaeus University.