A visual analytics framework for conservation planning optimization

Rui Zhang [a], Yafeng Lu [a], Katherine Adams [b], Jorge A. Sefair [a,*], Haley Mellin [c], Miguel A. Acevedo [d], Ross Maciejewski [a]

[a] *School of Computing and Augmented Intelligence, Arizona State University, USA*
[b] *Department of Industrial and Systems Engineering, University of Wisconsin-Madison, USA*
[c] *Conserve.org, Art to Acres, USA*
[d] *Department of Wildlife Ecology and Conservation, University of Florida, USA*

ARTICLE INFO

ABSTRACT

Approximately 15% of global land is currently in some state of protection. Recent conservation research suggests the need for a drastic increase of protected lands by 2050. In order to reach this target, an additional 35% of lands need to be conserved or restored in a cost-effective and time-efficient manner in order to support the resiliency of our planet and its climate. While many individuals and foundations continue raising much-needed funding for the environment, the development of conservation portfolios is a complex multi-dimensional task. Agencies have limited resources for investing in new conservation areas and have differing priorities for conservation in terms of species, land cover, human activities, etc. We present an interactive conservation portfolio development system that combines visualization, multicriteria analysis, optimization, and decision making that enables conservation planners and scientists to efficiently construct, compare, and modify conservation portfolios under multiple constraints.

1. Introduction

Biodiversity is declining at rapid rates due to human-driven habitat loss and landscape deterioration (Stokstad, 2010). Human activities have resulted in species extinctions at 10–100 times normal 'background' extinction levels (Sala et al., 2000; Pimm et al., 1995). This rapid biodiversity decline threatens the provision of key ecosystem services such as food, clean water, and crop pollination, resulting in negative consequences for economies and human health (Mace et al., 2012). Therefore, protecting remaining natural areas is fundamental to preserve biodiversity and to mitigate the negative consequences of ongoing environmental change (Johnson et al., 2017). While 15% of the Earth is in some kind of protection (Belle et al., 2018), this is still insufficient due to substantial gaps in land coverage and increasing threats (Rodrigues et al., 2004). Recent studies suggest the need of a drastic increase in protected lands by 2050 to maintain current rates of resource extraction (Watson and Venter, 2017; Dinerstein et al., 2019). This ambitious goal contrasts with the limited resources available to local institutions to design and implement networks of protected areas (Bicknell et al., 2017).

Conservation biologists apply systematic conservation planning approaches to design protected area networks that are cost-effective while meeting conservation goals. This systematic process is composed of six steps that include: (1) biodiversity data collection and analysis, (2) identification of conservation goals, (3) analysis of current conservation areas, (4) identification of a set of additional areas, (5) implementation of proposed conservation actions, and (6) preservation of required conservation values (Margules and Pressey, 2000). Out of these, we focus on Step (4), the identification of additional conservation areas—one of the most challenging steps in the conservation planning process. These conservation areas are selected with multiple conservation goals in mind, such as maximizing biodiversity representation while attenuating future threats, and remaining within a limited budget (Luck et al., 2012). This is a complex selection process that, if performed using inadequate quantitative tools, may result in landscape or seascape portfolios that are not optimal in terms of their budget and priorities. Therefore, the success of systematic conservation planning rests, in part, in the development of appropriate data-driven methodologies for designing protected area networks at the regional level (Williams et al., 2005).

---

* Corresponding author.
E-mail addresses: rzhan100@asu.edu (R. Zhang), yafeng@asu.edu (Y. Lu), kbadams@wisc.edu (K. Adams), jorge.sefair@asu.edu (J.A. Sefair), haley.mellin@gmail.com (H. Mellin), maacevedo@ufl.edu (M.A. Acevedo), rmacieje@asu.edu (R. Maciejewski).

Web-based geographic information systems (WB-GIS) provide an ideal setting to translate the result of complex spatial mathematical models used in systematic conservation planning into simple qualitative visual scenarios (Dragicevic, 2004). These WB-GIS can summarize multiple layers of information, allowing planners to analyze various future hypothetical scenarios (Rao et al., 2007). While multiple mathematical models are available to prioritize areas for conservation (Sarkar et al., 2006; Moilanen et al., 2009), designing a network of protected areas requires the quantification, visualization, and adjustment of multiple hypothetical scenarios almost simultaneously (Tress and Tress, 2003; Pettit et al., 2011). In this context, typical questions faced by conservation analysts include, what areas should be selected as part of a network of protected areas to have all species of conservation concern under protection while minimizing the acquiring costs? If we decrease the budget by 10%, which areas should be protected? What happens if, instead, we increase the budget by 5%? Therefore, there is a need for WB-GIS applications for conservation planning that combine cost optimization with efficient visualization tools that can provide alternative future scenarios in real time (Portman, 2014).

In this paper, we present an interactive conservation portfolio development system that combines visualization, multicriteria analysis, optimization, and decision making that enables conservation planners and scientists to explore different land purchasing portfolios under a variety of constraints in real time. Our system incorporates a multi-layer map view, a parallel coordinates attribute view, a control area for optimization modeling, and a multiple portfolio visualization for solution comparison. To support automatic portfolio optimization, we implemented a median ranking algorithm to allow parcel filtering by an aggregated indicator of all the attributes and an integer programming model to generate land purchase recommendations given user-defined constraints and objective function. The visual analytics system is designed to support the efficient selection of conservation areas by enabling portfolio generation and interactive modification. Multiple land portfolios can be generated and saved for comparison. Our system complements the existing body of tools by providing new visual, analytical, and mathematical features, while also allowing loading of a (shape compatible) conservation plan obtained with any other tool for further visual analysis.

From the software systems perspective, we propose a novel combination of visualization components, where our design has focused on featuring credibility, saliency, and legitimacy (White et al., 2010). The software tool itself serves as a boundary object to enable decision making. Our design ranges from providing detail-on-demand for the data source to enable analysts to determine credibility of data layers, interactive selection of optimization criteria, and provenance analysis. Specifically, for supporting provenance analysis and comparison, we propose new visualization designs to capture different portfolios and provide comparison between them. Along with novel integration of techniques and the proposal of a visualization design, we have also designed a pre-processing scheme to match data across different levels of granularity. Our down-sampling technique allows data comparisons at a high-resolution level, and supports land purchases which can only occur at the parcel level. The human-machine combination is also innovative, where our framework is designed to present an optimal solution within the problem formulation; however, the problem formulation needed to be computationally efficient for real-time exploration. By providing an optimal conservation portfolio as a first pass, we allow users to refine their choices in a human-machine teaming process. Our system enables conservation planners to develop consecutive portfolios in real-time and adjust the outputs of the multiple criteria optimization selections. From an optimization perspective, our proposed approach uses the analyst's preferences to drastically reduce the problem size. By supporting interaction with the optimization results, planners can utilize species specific knowledge to enforce different landscape features (i.e., connectivity, compactness, corridor width) which can be challenging for automatic optimization either due to the difficulty in representing the

corresponding constraints or the computational complexity of the resulting model. Although we propose a simple (and fast) optimization model to support land acquisition decisions, the proposed system can accommodate the results from other spatial models. As a result, our system can be seen as a visualization framework that supports user interaction with an optimal solution. To our knowledge, currently there is no tool available with the proposed features to support conservation decisions.

2. Related work

Our system is designed to support decision making through multi-criteria analysis and solution comparison. In this section, we summarize previous work in visual analytics for decision making and multicriteria optimization for conservation planning.

Recent visualization work has focused on how to best display multiple attributes for analysis, a key component of multicriteria analysis. Turkay et al. (2014) explored the geographic variation of multivariate data and developed attribute signatures consisting of dynamically generated graphs to summarize the change of statistics over a sequence of geospatial data selection. Pajer et al. (2017) proposed WeightLifter, a technique that allows the exploration of weight space with up to ten criteria and helps to explore the sensitivity of candidate solutions to the change of weights. Weng et al. (2018b) designed a visual analytics system, ReACH, which helps analysts identify their ideal home given multiple purchasing constraints. Common amongst many of these systems are the use of parallel coordinates plots, and a wide variety of extensions to parallel coordinate plots. Systems extending the parallel coordinated plots (e.g., Lind et al. (2009); Johansson et al. (2005); Rosenbaum et al. (2012); Xie et al. (2017)) provided a means to easily explore multivariate data. Other works have focused more on supporting analysis and decision making through the integration of interactive models. Afzal et al. (2011) developed a decision support environment to evaluate disease control strategies by predicting the course of an outbreak and analyzing the response measures. The severity of the epidemic is visualized by different color intensities on the map, and a custom split timeline is used to show the solution path. Konev et al. (2014) proposed an automatic simulation-based approach for flood management. The decision trees are automatically generated and visualized by clustered timelines. Rinner (2007) developed a geographic visualization system to support multi-criteria decision making. An index rank for each tract is calculated, and users can explore attributes through a linked parallel coordinate plot. Similar to the work of Rinner (2007), Cassol et al. (2017) proposed a framework to explore the optimal evacuation plan for crowd egress based on multiple factors, which were taken as input by the proposed metric to calculate the optimal plan. In both systems, interactive optimization methods are not fully considered. These systems support multi-criteria analysis through interactions with a parallel coordinate plot and quality indices (similar to our use of median ranking). However, portfolio comparisons and interaction with the optimization results are limited.

Other major issues underlying such decision support systems are the mechanisms used to compare across candidate solutions. The work by Gleicher (2018) summarized the basic designs of comparison into three categories, juxtaposition—i.e., which places the compared items are in different screen spaces, superposition—i.e, which places the compared items fit into the same screen space, and explicit encoding—i.e., visualization of the relationship between the compared items. Kehrer et al. (2013) and Munzner et al. (2003) utilized juxtaposition design for their comparisons of bar charts, lists, and trees. Dasgupta et al. (2015) combined juxtaposition and superposition for climate model comparison. Law et al. (2018) developed Duet, a visual analytics system for pairwise comparison integrating all three categories. Duet uses visualizations and textual descriptions to explain the recommended object groups which are similar to, or different from, the user-specified object with a focus on the similarity and difference. Weng et al. (2018a) proposed a spatial

ranking visualization technique to explore and analyze ranking datasets and annotate the cause of the ranking with spatial context, which involves the three design categories of comparison. From the optimization perspective, multi-criteria analysis and modeling have been integrated in a number of visual analytics systems for domains ranging from epidemiology to emergency response. However, little work in the visual analytics community has focused on conservation planning. Conservation planning requires the integration of optimization algorithms for conservation portfolios given the myriad of possible parcel configurations available. These conservation portfolios must allocate resources efficiently while considering current and future threats and their influence on the biodiversity assets. The problem of designing natural reserves has received considerable attention since the 1980s (Kirkpatrick, 1983; Cocks and Baird, 1989), mostly through the use of exact optimization models (Ando et al., 1998; Church et al., 1996; Polasky et al., 2001; Sefair et al., 2017; Acevedo et al., 2015) and heuristic approaches (Pressey et al., 1997; Arthur et al., 1997; Margules et al., 1988). The use of Operations Research techniques in this area have become more prevalent in recent years, including deterministic and stochastic approaches (see Moilanen et al. (2009) for a comprehensive review). Moreover, these methodological efforts have evolved into free software designed to support conservation planning decision-making processes (e.g., Zonation and MARXAN) (Lehtomäki and Moilanen, 2013; Ball et al., 2009). Although available tools cover several pressing issues in conservation problems, some contemporary challenges are still unsolved. Some of the existing approaches focus on cost-minimization subject to ecological outcomes, ignoring the more realistic *dual* problems of maximizing such outcomes subject to a given budget. Approaches that optimize the ecological performance of the conservation portfolio approximate the quality of candidate patches by species representation (i.e., whether a species is present in a patch) and other single static patch attributes (Toregas and Revelle, 1973; Underhill, 1994; Williams and ReVelle, 1996; Camm et al., 2002; Moilanen et al., 2009), ignoring the multiobjective nature of the conservation decisions.

Other works focus on desirable geographical properties of protected areas such as landscape connectivity (Önal and Briers, 2006; Dilkina and Gomes, 2010; Dissanayake et al., 2012; Jafari and Hearne, 2013) and compactness (Önal and Briers, 2002; Nalle et al., 2002; Dissanayake et al., 2012; Jafari and Hearne, 2013) but ignore the subjacent ecological processes. The majority of the works studying connected and compact reserves are mixed-integer programming models that are difficult to solve for realistic-size instances and that provide a single solution (i.e., a single connected and compact set of parcels to purchase). Without the visual support, analysts cannot easily modify an existing solution to incorporate expert knowledge and other attributes not included in the optimization model. Although optimization models in conservation planning are difficult to solve (Margules and Pressey, 2000), they are a fundamental part in the conservation decision process. However, ignoring other equally important components such as the interaction with experts for the inclusion of non-quantifiable or other aspects that are hard to express as constraints or objectives may reduce their applicability in real life conservation decisions.

3. Visual analytics framework

This section describes the design process and components of the proposed framework. The design of the proposed system is the result of a collaborative effort with a variety of stakeholders including donors, ecologists, and conservation planners. Through discussions and planning with domain experts, we identified key data needs, tasks, and design requirements. The proposed framework avoids the manual processing of the attributes of each candidate parcel to determine its relative convenience with respect to other parcels in the area of interest. It also consolidates the data processing, visualization, and optimization processes into a single intuitive tool. The interaction with potential users resulted in the following functionalities of our framework.

- Data Storage and Downscaling: Stores map data for conservation planning, including biological, physical or socio-economic attributes. Currently includes 12 attributes suggested by conservation planners, and is scalable to further attributes. The data set is categorized into land use, physical-geospatial, and biodiversity layers. Input data is downscaled to the parcel scale to facilitate the calculation of the quality of land portfolios.
- Multi-layer Map View: Allows the investigation of parcel attribute values and the visualization of one of more attributes over a common area of analysis.
- Attribute Selection View: Filters parcels whose attributes fall within a certain range of interest. Provides the distributions of attribute values in any selected search area and allows the user to turn on/off each attribute layer on the map, define the ranking order of each attribute (e.g., higher values are preferred), and filter parcels based on a ranking aggregation metric calculated using selected attributes.
- Conservation Portfolio Optimization: Allows the specification of requirements for the land purchase portfolio, such as area of interest on the map, desired criteria for candidate parcels, objective to optimize, constraints, and maximum budget. Embeds a multicriteria optimization functionality to automatically provide land purchase recommendations and allow the user to visually interact with a solution to induce other desirable performance metrics (e.g., landscape connectivity and compactness).
- Porfolio Comparison View: Provides comparison tools to help portfolio managers explore their criteria of interest, compare land purchase portfolios, and work together to realize their final solution space.

We build upon previous works on multicriteria analysis and visualization, integrating geographic visualization and optimization to recommend land portfolios. Our system is designed to support the comparison of candidate land portfolios generated between the optimization recommendation and the analyst adjustments. Similar to previous work, our system uses a color code to visually inform the analyst on the quality of patches and land portfolios, linking attribute analysis and filtering to a parallel coordinates plot. Instead of displaying a sequence of portfolios to illustrate the impact of parameter changes, our system provides a unique visualization method to help comparing the attributes and their differences between various candidate portfolios. Our target users are conservation planning decision-makers in the broad sense. This could be an analyst assessing the ecological benefits of land patches, an ecologist surveying alternatives to expand current reserves, or NGOs and government agencies deciding which patches of land to restore or purchase. Fig. 1 shows a snapshot of our system and its features, which is freely available at Zhang et al. (2021a). We have deployed this system to conservation planners, and our use cases demonstrate the effectiveness of optimizing their decision process given limited resources. A step-by-step demonstration video is available at Zhang et al. (2021b). The final product functionalities are explained in detail in Sections 3.1–4.4.

3.1. Data Storage and downscaling

Typical data for conservation planning is characterized by biological, physical or socio-economic attributes. Our framework includes 12 common attributes and is scalable to additional data. We use the state of Montana as an example to describe the properties of a typical dataset for this system and the data downscaling steps. Table 1 describes the used datasets and their attributes that were chosen by conservation experts.

Our system supports a wide variety of shapefiles, geotiffs, open street map layers, among other types, including conservation portfolios built in other tools (e.g., MARXAN) as long as they are compatible with the shapefiles in our system. We note that our system is flexible to any geographical data, where users only need to select a base layer for analysis. Typically, the parcel layer would be used for this purpose, as
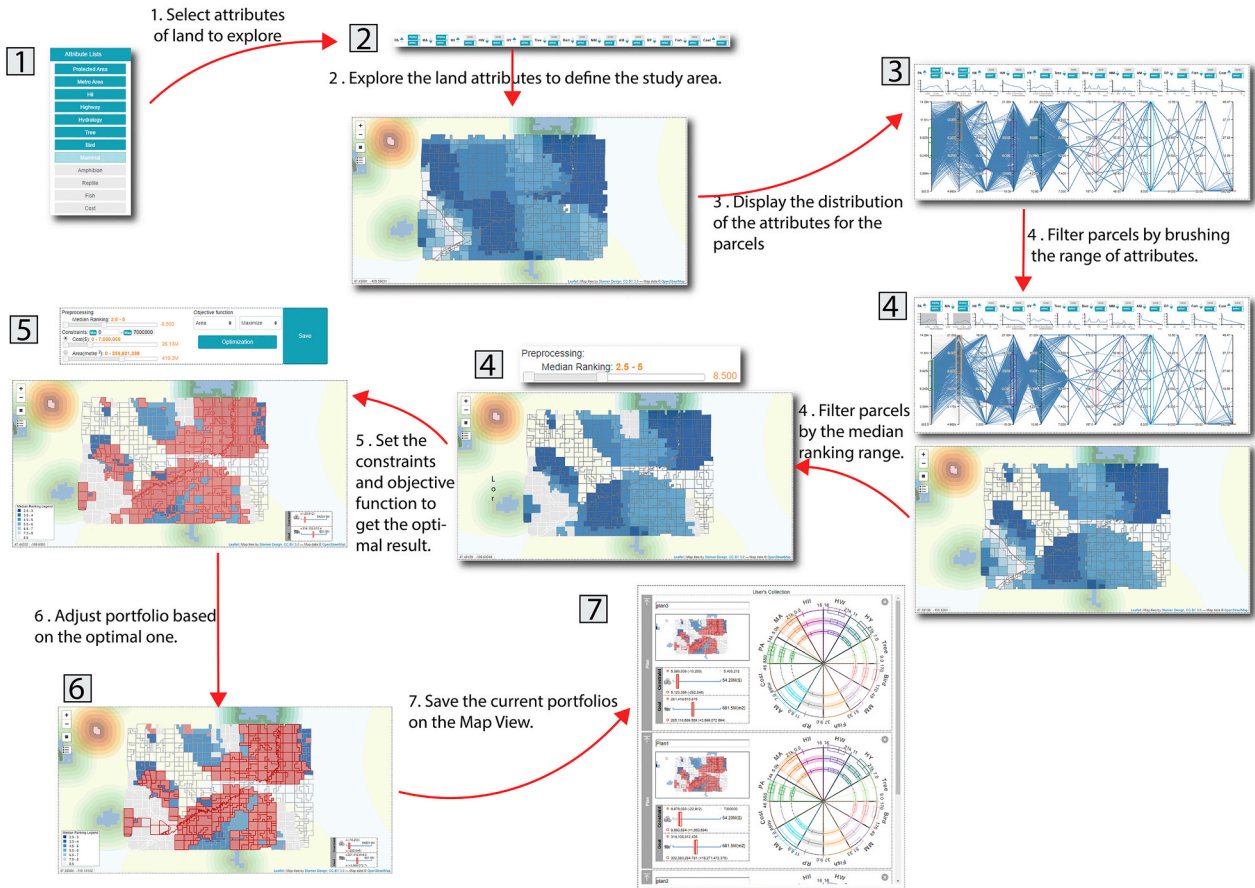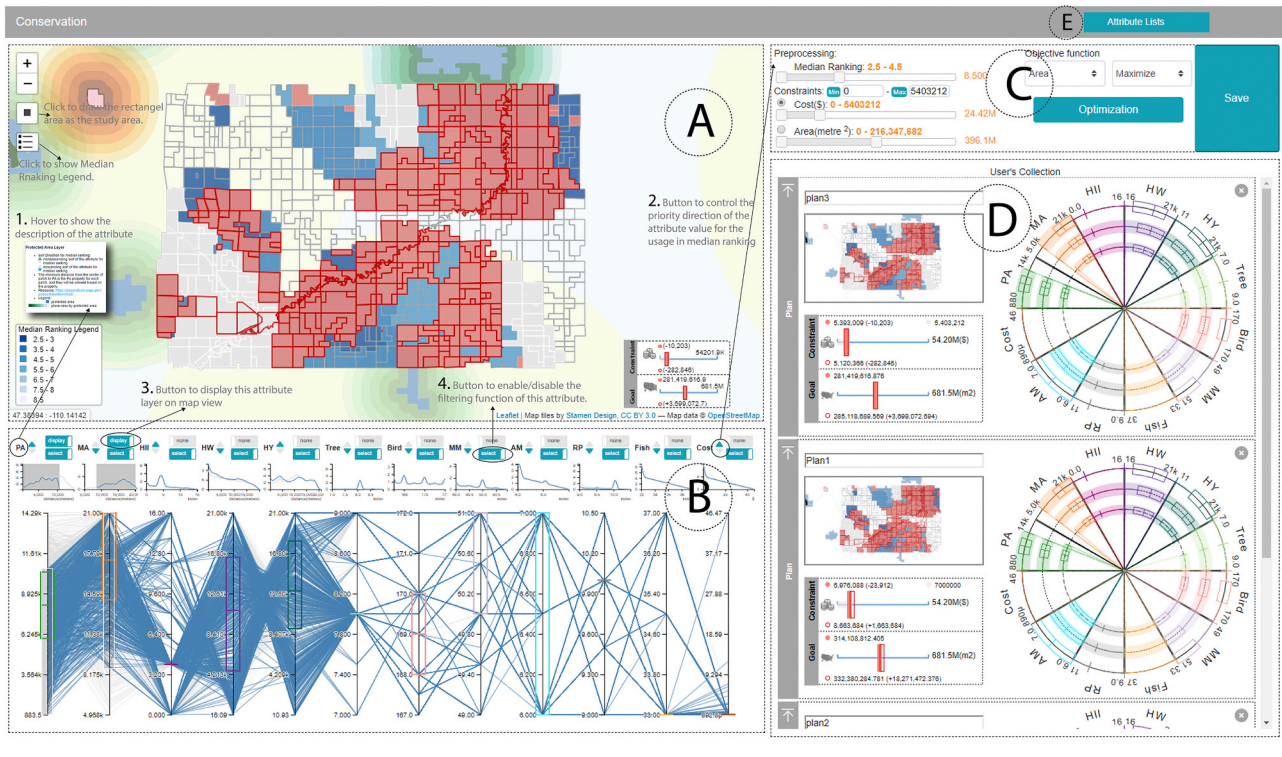
(caption on next page)

Fig. 1. A visual analytics framework for conservation planning. (a) Map View: Visualizes parcels falling in the analyst-defined search area. Colors in the map reflect the median ranking, attributes, or optimization results. Users can create a customized portfolio with the draw, pan, and zoom controls in the upper left corner. The legend and the differences of the constraints and goals reached by the portfolios are shown separately in the left and right bottom corner. (b) Attribute Analysis View: Attribute distributions are visualized by line charts and parallel coordinates integrated with a box plot, which can be brushed to filter parcels. (c) Optimization Configuration: The median ranking slider filters out low-quality parcels and reduce the optimization algorithm run-time. The radio and text box are used to input the optimization constraints and objective function. Users click the 'Optimization' button to run the algorithm and 'Save' the current portfolios on the map to the 'User's Collection' for comparisons. (d) Portfolio Comparisons: The generated portfolios are saved, and a screenshot of the selected portfolios and attributes are visualized. (e) Attribute Lists: Drop-down menu for selection. Event Sequence: 1> select the land attributes to explore. 2> Explore the land attributes to define the study area. 3> Display the distribution of the attribute for the parcels. 4> Filter parcels by brushing the range of attributes and the median ranking range. 5> Set the constraints and objective function to get the optimal result. 6> Adjust portfolio based on the optimal solution. 7> Save the current portfolio on the map view for future comparisons.
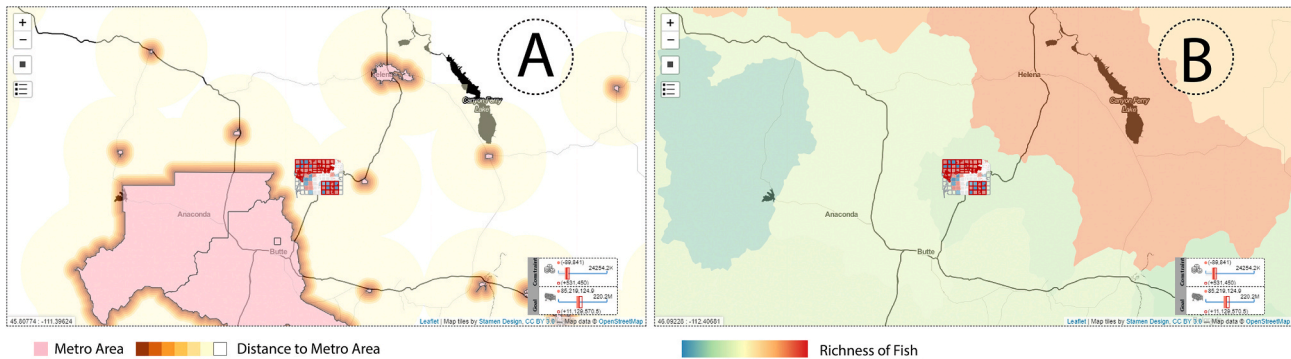

Fig. 2. Examples of multi-layer map views. (a) Visualization of distance to metro areas. The pink area is the metro area and the peripheral region is colored from brown to yellow based on the distance value of each patch. The red/blue highlights correspond to the user selected region of interest. (b) Visualization of fish species richness. The region with bluer color has lower species richness for fish.

this is the level at which land can be purchased. Once the base spatial unit is chosen, attributes are aggregated or dis-aggregated through a downscaling step to conform to the level of spatial granularity under analysis. For each data category, we use different processing rules to derive the corresponding attribute(s). Other than COST, which is directly provided in the parcel shapefile dataset, we downscale the remaining datasets to calculate the parcel-level attributes. We calculate the distances to the existing protected areas, metro area, highway and hydrology areas, and aggregate the HII and other biodiversity attributes. Some conservation attributes measure the distance from a parcel to a feature of interest. In our dataset, examples include PA, MA, HW, and HY, which require us to calculate the distance from the parcel to the areas described in the attribute datasets. We first discretize each dataset into $30 \times 30 \ m^2$ *patches* (a request from our conservation planning partners), which will be later used to calculate the attributes of the larger-sized *parcels*. Parcels can be different in shape and size, and there are a variety of geographic aggregation methods that can be employed to calculate their attributes out of the patch attributes (Unwin, 1996). Then, we calculate the distance from the center of a patch to its nearest feature of interest. From there, we can aggregate all patches that fall within a parcel using min, max, average, or other aggregation functions. In our system, we use the average value of all patches within a parcel. Other attributes focus on measurements and estimates from sensors, reports, and other sources. Examples of these attributes include TREE, BIRD, FISH, and other attributes in the biodiversity layer. We overlay the parcels onto the datasets and perform an aggregation operation to estimate the parcel attributes.

3.2. Multi-layer map view

In order to support the multicriteria analysis during the decision-making process, we incorporated a multi-layer map view to visualize each attribute and their combinations over space. For distance-based attributes (PA, MA, HW, and HY), a sequential color scheme is used. The darker color means a patch is closer to the feature of interest. As an illustration, Fig. 2A shows the visualization of the distance to the metro area (MA) attribute. The pink region is the metro area, and the

peripheral region around the metro area is colored based on the distance. The red and blue highlighted regions in Fig. 2A are the parcels in the user selected region of interest.

For region-based attributes (HII, COST, TREE, FISH, BIRD, AM, MM, and RP), the original datasets are overlaid on the map and colored based on their attribute values using diverging color schemes. The color scheme is designed to match the NASA analysis (SEDAC, 2018) and BiodiversityMapping.org (Jenkins, 2017). Fig. 2B shows the visualization of FISH. The region with redder color has higher species richness, while the region with bluer color has lower species richness for this variable. In the map view, the user can define their conservation area by drawing a rectangle on the map. Once the area is selected, the optimization algorithm suggests which parcels within this area to buy (the red/blue area seen in Fig. 2). The parcels are colored based on an aggregation of the parcel attributes through a ranking function (see Section 3.4), filtering updates, the optimization algorithm's solution, and other user modifications. Results in the selection are influenced by the Attribute Selection View.

3.3. Attribute selection view

The attribute selection view integrates parallel coordinates, line charts, and attribute controllers (see Fig. 1B). The user can explore value distributions of attributes in the search area, turn on/off each attribute layer on the map, define the ranking order of each attribute in the median ranking, and filter attributes for median ranking and based on the attribute value.

On each attribute controller, the user can click the top switch button (see Fig. 1B3) to turn on/off the corresponding attribute layer, and mouse over the attribute name (see Fig. 1B1) to see the explanation of the attribute and the color legend or the layer. The bottom switch button (see Fig. 1B4) is used to enable/disable the filtering function of this attribute. The user can still explore the value distribution of an attribute when its filtering function is disabled, but interactions on disabled attributes won't impact the map view or the optimization model. The triangles pointing up and down (see Fig. 1B2) are used to decide the priority direction of the attribute value when used in the median ranking

Table 1
Variables and data sources used.

Category	Attribute Name (Abbr.)	Explanation	Data Source
Land Use Layers	Distance to protected area (PA)	The average distance of a parcel to its nearest protected area(s)	The shapefile of protected area by state from the USGS (USGS, 2018b)
	Distance to a metro area (MA)	The average distance of a parcel to its nearest metro area(s)	The shapefile of the 129 incorporated cities and towns in Montanan from Montana.gov (State of Montana, 2018)
	Distance to highway (HW)	The average distance of a parcel to its nearest highway (s)	The shapefile of primary and secondary roads by state from Census.gov (USCB, 2017)
	Human influence index (HII)	HII values range from 0 to 64 and measure the direct human influence on terrestrial ecosystems (Sanderson et al., 2003). The average of the HII values within a parcel.	The shapefile of HII by North America from the socioeconomix data and applications center of NASA (SEDAC, 2018)
	Cost per square meter (COST)	total cost per square meter	The shapefile of parcels from Montana (Loveland Tech, 2018).
Physical-Geospatial Layer	Distance to hydrology area (HY)	The minimum distance from the center of the parcel to the nearest hydrology area	The shapefile of hydrology area by state from the USGS (USGS, 2018a).
Biodiversity Layers	Richness of trees (TREE)	The total species richness of trees in the parcel.	The TIF file of richness of trees by state from BiodiversityMapping.org (Jenkins, 2017).
	Richness of birds (BIRD)	The total species richness of birds in a parcel.	The TIF file of richness of birds by state from BiodiversityMapping.org (Jenkins, 2017).
	Richness of fishes (FISH)	The total species richness of fishes in a parcel.	The shapefile of richness of fishes by state from BiodiversityMapping.org (Jenkins, 2017).
	Richness of amphibians (AM)	The total species richness of amphibians in a parcel.	The shapefile of richness of amphibians by state from BiodiversityMapping.org (Jenkins, 2017).
	Richness of mammals (MM)	The total species richness of mammals in a parcel.	The TIF file of richness of mammals by state from BiodiversityMapping.org (Jenkins, 2017).
	Richness of reptiles (RP)	The total species richness of reptiles in a parcel.	The TIF file of richness of reptiles by state from BiodiversityMapping.org (Jenkins, 2017).

aggregation (e.g., whether near or far proximity is desirable). For example, if the user wants to buy parcels near a protected area, then the priority direction is non-decreasing. That is, the user prioritizes low PA values by turning on the up triangle for the PA attribute. By turning on their down triangles, the user prioritizes high values of TREE, BIRD, FISH and AM attributes.

The line charts and parallel coordinates display the value distribution of each attribute and support parcel filtering by attribute value. Such filtering is only active when the attribute's filtering function is enabled. To explore attribute correlations and observe patterns of the

data, the user can drag the axes of the parallel coordinates to change the order of the attributes. On each axis of the parallel coordinates, we add a box plot to help reveal the statistical distribution of the data. We use a categorical color scheme for the box plots to represent different attributes, and the attribute uses the same color in the portfolio comparison view, which we describe in detail in Section 3.5.

When the number of parcels increases, it is difficult to observe the distribution on the parallel coordinates due to visual clutters. Therefore, each attribute is also associated with a line chart where the x-axis represents the attribute value and the y-axis represents the frequency of the attribute value. The line chart is adjacent to each axis in the parallel coordinate plot and is used to show the value distribution of both the original data and the filtered data. To filter parcels, brush interaction is supported on the axes of the parallel coordinates as well as on the x-axis of the line chart. Parcels removed from the filtering will be grayed out on the parallel coordinates, while brushed parcels are highlighted in blue. On the top line chart, the black line shows the value distribution of all parcels in the search area, and, once filtered, a blue line is used to display the value distribution of the filtered parcels, and the original line will become gray.

In our system, all the interactions are coordinated with the map view. Once attributes are selected, the parcels in the user selected area will be colored based on their median ranking order. The legend for the median ranking results is in the left bottom of the map. The result of the median ranking depends on which parcels are selected and which filters have been applied to the data. The attributes of the parcels in the selected area are then used to generate a potential conservation portfolio.

3.4. Conservation Portfolio Optimization

Once the region and attributes are defined, our system employs a mathematical programming model to identify an optimal portfolio of patches for conservation. We define P as the set of candidate parcels eligible for purchase and A as the set of attributes of interest. We assume that all attributes are (or can be converted to) numerical values, and that all attributes are available for each parcel. We denote the value of attribute $j \in A$ for parcel $i \in P$ by a_{ij}. Depending on the discretization of the area of analysis (chosen by the user), the number of candidate parcels may be very large. To reduce the computational effort in our system, we implement two pre-processing techniques. Both aim to reduce the set of candidate parcels by ignoring some that are not of interest for the decision-maker. The first technique is based on user-defined attribute filters. In this case, the user explicitly sets thresholds for a subset of the attributes, and the system discards those parcels with attributes violating the thresholds. Mathematically, we denote the set of attributes with threshold values as $\bar{A} \subseteq A$, and the corresponding lower and upper threshold values by $\bar{a}_j$ and $\underline{a}_j$ for attribute $j \in \bar{A}$, respectively. Using these values, the set of eligible parcels can be calculated as $P = \{i \in P : \underline{a}_j \leq a_{ij} \leq \bar{a}_j, \forall j \in \bar{A}\}$. The $\bar{a}$- and $\underline{a}$-parameters are calculated via user interactions with the map and the attributes' value distribution.

Depending on the magnitude and meaning of an attribute, it may not be intuitive for the user to specify the $\bar{a}$- and $\underline{a}$-parameters. We also determine the set of eligible patches using a ranking-based procedure that describes the *relative* performance of a parcel with respect to other parcels. Parcels are sorted in non-decreasing order based on each attribute and then ranked such that $r_{ij} < r_{kj}$ if $a_{ij} < a_{kj}$, where $r_{ij} \in \{1, ..., |P|\}$ is the rank of candidate parcel $i \in P$ on attribute $j \in A$. In other words, the smaller the value of an attribute the higher the ranking of the parcel on that attribute (i.e., closer to 1). In the case where attributes with larger values are preferred (e.g., distance to human settlements), then the attribute values are sorted in non-increasing order and ranked such that $r_{ij} < r_{kj}$ if $a_{ij} > a_{kj}$. If two parcels have the same value on a particular attribute, then their ranking on that attribute is the same, i.e., $r_{ij} = r_{kj}$ if

$a_{ij} = a_{kj}$. The ranking describes the parcel's relative performance on each attribute. We aggregate such rankings into a single number $\tilde{r}_i$ using the median value of the rankings across attributes. In other words, $\tilde{r}_i =$ median$(r_{i,1}, \ldots, r_{i,|A|})$, $\forall i \in P$. We add the aggregated rank to the set of attributes for each parcel, allowing the user to specify more intuitive filters on the $\tilde{r}$-values. For instance, the user can choose to discard parcels that are not among the top k parcels–according to the median ranking–by setting the corresponding a-parameter to k. We use median ranking aggregation because, among other properties, it eliminates the effect of extreme r-values and it can be computed efficiently (Sculley, 2007). Our system is flexible to accommodate any other ranking aggregation model. For a review on ranking aggregation, readers are referred to Sculley (2007), Lin (2010), and Ailon et al. (2008) and the references therein. To find an optimal set of parcels for conservation, we use an integer programming model with variables x_i, where $x_i = 1$ if parcel i is recommended for purchase, and $x_i = 0$ otherwise, $\forall i \in P$. The model constraints represent conditions that a *portfolio* of parcels must satisfy, as opposed to the individual parcel conditions described in the pre-processing analysis. These include land purchase budget, minimum population area to protect, among others. We use linear constraints reflecting that the *aggregated* value of an attribute for the selected parcels must be less than (or greater than) or equal to a threshold value. We denote by $A^{\leq}$ and $A^{\geq}$, the set of attributes with a less than or equal to and greater than or equal to constraints, respectively. We use b_j as the threshold value for attribute $j \in A^{\leq} \cup A^{\geq}$. Note that not all attributes need to be included in such constraints, which means that $A^{\leq} \subseteq A$ and $A^{\geq} \subseteq A$. We pay special attention to the cost and area of each parcel, which we denote by c_i and α_i, $\forall i \in P$, respectively. Our optimization problem maximizes the total purchase area (1a), subject to attribute constraints (1b)–(1c), and variable-type constraints (1d). The optimal purchased area will be a subset of the available area given that the purchasing cost will be part of $A^{\leq}$, with a corresponding b-parameter equal to the budget available for land purchases.

$$\max \quad \sum_{i \in P} \alpha_i x_i \tag{1a}$$

$$\text{s.t.} \quad \sum_{i \in P} a_{ij} x_i \leq b_j, \quad \forall j \in A^{\leq} \tag{1b}$$

$$\sum_{i \in P} a_{ij} x_i \geq b_j, \quad \forall j \in A^{\geq} \tag{1c}$$

$$x_i \in \{0, 1\}, \quad \forall i \in P \tag{1d}$$

An alternative model minimizes the total purchase cost, subject to constraints (1b)–(1d). In this case, the area will be part of $A^{\geq}$, with a corresponding b-parameter equal to a minimum required area to conserve. Mathematically, this problem can be written as $\min \sum_{i \in P} c_i x_i$, subject to (1b)–(1d). Although some of the constraints in our models may indirectly induce some landscape attributes (e.g., landscape connectivity by selecting the distance to existing protected areas as an attribute), the conservation portfolio produced by our models may not satisfy some those landscape requirements. This is because of the complexity and computing demand of enforcing such constraints for any arbitrary sized area selected by the user. Instead, our system allows the user to interactively modify an existing solution (through clicks on the map) to induce these landscape features. This allows the exploration of solutions that are infeasible for the optimization model, but that provide a good compromise between the ecological values gained and the extra cost required. The user is allowed to add or remove attribute constraints, as well as select the objective function to optimize (maximize the protected area or minimize the purchasing cost). Our models produce an optimal purchasing plan that satisfies all the selected attribute constraints at the same time. Using the optimal values of the decision variables, denoted by x_i^*, we define an optimal conservation portfolio as $P^* = \{i \in P : x_i^* = 1\}$. These optimal portfolios are displayed for further user analysis. The analyst interacts with the optimization model through the configuration view (see Fig. 1C). The analyst can filter parcels based on their median ranking and sets the constraints and objective function of the optimization model. The analyst can also "save" the current portfolio to the comparison view for further exploration and comparison. The median ranking slider shows the rankings of all selected parcels, and the analyst can drag the two ends of the slider to remove low-ranked or high-ranked parcels. The filtering tool changes the parcels used in the automatic optimization algorithm. The sliders under "constraints" are used to set the constraints for the optimization model. Currently, our system is able to answer the questions: *What is the largest total area that can be protected given a fixed budget and other ecological constraints?* and *What is the least-expensive set of parcels to protect with an area of at least b km^2 while satisfying other ecological constraints?* Therefore, the analyst needs to select one variable between "cost" and "area" to be the constraints, and leave the other to be the objective function. Our mathematical algorithm and system can support multiple constraints. For both "cost" and "area", the maximum value of the slider updates to represent the sum of the cost and area of user-selected parcels. Dragging the ends of the slider can change the value range we set for the constraint. To set the objective function, the analyst can choose either to maximize or minimize the variable. When the configuration is done, the analyst can click on the "Optimization" button to run the algorithm. For an easy comparison of multiple optimal portfolios under different right-hand-sides of the constraints, the constraint value from the previous run of the optimization algorithm is recorded in the slider.

3.5. Portfolio comparison view

Multiple land purchasing portfolios may satisfy the planners' requirements under different attribute priorities. The analyst can make different modifications on top of the same suggested portfolio or change the selected parcels. Fig. 3 shows our portfolio comparison view which uses a multiple portfolio visualization to display all saved portfolios. Each portfolio visualization has three visual components, the map screenshot, the optimization setting, and the attribute pie. The map screenshot represents the exact status of the map view when the portfolio is saved, and it records the details of the parcel selection in the portfolio. The optimization setting uses the same design as the lower right legend on the map to present the constraint and objective function for the optimization algorithm. The attribute pie is a glyph designed to visualize the attribute distribution of selected parcels under the setting of each portfolio and allow the analyst to compare their customized portfolio to that suggested by the optimization model. The pie shows all attributes with evenly split sectors and each attribute is assigned one color. This is because even if not all the attributes are used to filter parcels, their value distribution may need to be considered in the final decision-making process.

To compare the influence of an attribute value in the portfolio, three circles with different radius are used. The outermost circle represents all the parcels in the search area, the middle circle represents the parcels suggested by the optimization model, and the inner circle represents the analyst selected parcels, which are those finalized in the portfolio. The three circles are arranged into the same value scale, which ranges from the minimum to the maximum of all parcels in the search area (the range of the outermost circle). For the middle and the innermost circle, the value range is also emphasized with a brushed color arc. Within this color arc, a box plot visualizes the attribute value's statistical distribution. We use a brushed color arc on the outermost circle to indicate the attribute value range that was used by the analyst to filter the attribute. If there is no such colored arc, it means the analyst did not filter on this attribute. The box plot on the outermost circle shows the quartiles of the
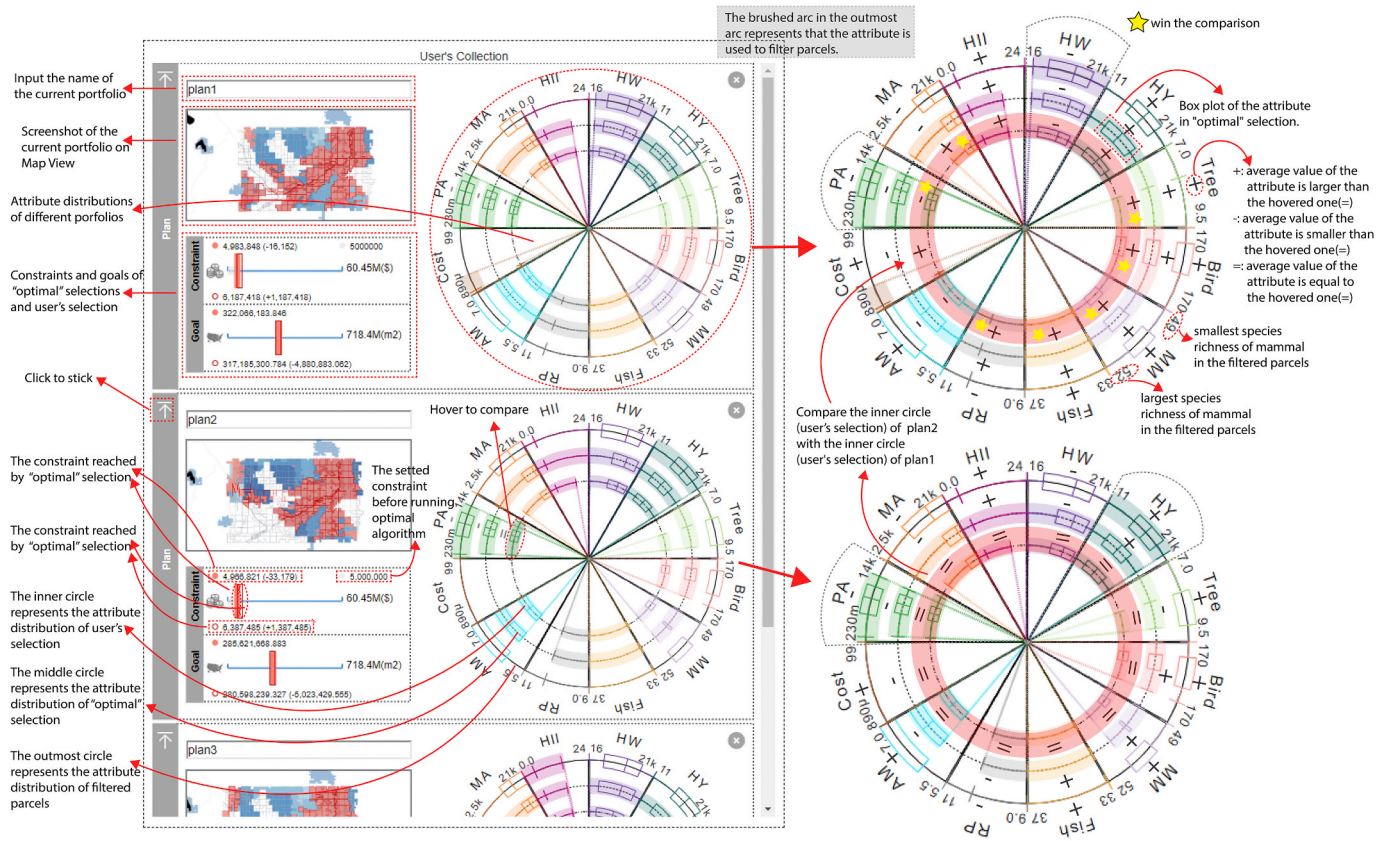
Fig. 3. Comparison of the average value for each attribute. The selected attribute with a gray arc is the PA attribute of the analyst's portfolio in the second row. All the comparison result of other portfolios with this selected attribute are given with three relations, "less than", "larger than", and "equal to". The corresponding signals in the comparison view are a minus sign, a plus sign, and an equal sign, respectively. For example, the average values of PA for an analyst's selected set of parcels in the first row are smaller than that of the analyst's selected set of parcels in the second row. Comparing analyst's portfolios of plan1 and plan2, plan1 can buy 314 million square meters area better than the 280 million square meters in plan2 with a similar cost. In addition, the 7 attributes out of 12 attributes in plan1 perform better than those in plan2. In general, plan1 is much better than plan2.

attribute value with these filtered parcels. By using these glyphs, analysts can compare the attribute distribution of filtered parcels, suggested parcels and the user selected parcels to explore how the choice of parcels affects the attribute distribution. Analysts can also directly compare different portfolios, allowing multiple analysts to provide input and serving as a mechanism for both provenance and analysis. The map screenshots of the portfolios provide an overview of the differences between search areas and parcel selection. A black vertical line across all saved portfolio appears when the analyst mouses over the optimal setting view so that the analyst can easily compare the value of the constraint and objective function (cost and area) for these portfolios. The analyst can also compare the attribute value distribution of different portfolios by mousing over one arc of an attribute to turn on the comparison signs of this attribute for all portfolios. In this case, the average attribute values are compared both between the parcels represented of the three circles within one portfolio and also between the parcel selections represented by the circles of other portfolios. The reference circle arc is colored gray. If the average value equals the reference value, it shows an = sign. When the average value of the parcels represented by the circle is larger than the reference value, a + sign will appear, and when the value is smaller, a − sign will appear.

4. Case studies

In this section, we illustrate the use of our tool for the selection of conservation areas in Montana, USA. The state of Montana has a long wildlife conservation tradition dating back to 1895 when the Game and Fish Commission was established (Brownell, 1987). The evolution of

wildlife legislation in this state reflects a serious commitment to the protection of wildlife; yet, less than 3.7% of its total area is designated as a wilderness protection area. Furthermore, most of the currently designated protected areas are composed of isolated mountain ranges clustered in a limited number of counties. Therefore, there is a need to complement existing protected areas by establishing new protection zones in counties that have limited designated wilderness areas and establishing corridors that facilitate movement and gene flow among wildlife populations living in isolated conservation areas (Hodgson et al., 2009).

4.1. Multi-species conservation scenarios for the Judith Gap in Montana

In this case study, a conservation planner (the "analyst" hereafter) selects a set of areas to acquire (or restore) near the Judith Gap in Wheatland County. This gap represents a region of unprotected land between protected areas in the Little Belt Mountains in the west and the Big Snowy Mountains in the east. The analyst's overall goal is to identify the largest possible total area to purchase subject to a budget constraint, while at the same time maximizing the number of terrestrial vertebrate species under protection within the corridor. There is evidence showing that in many instances the negative effects of human populations on protected areas decreases with distance to population centers (Mcdonald et al., 2009). In this case, our system allows the analyst to visually explore a variety of attributes related to human use. Using the distance to metro area (MA) layer, as shown in Fig. 4b, it is possible to assess the spatial relationship between existing protected areas and urbanized centers. The figure shows how protected areas are generally distant from

major urban centers. The highway layer (HW), as shown in Fig. 4c, illustrates how major roads may influence accessibility to protected areas. The figure shows how highway 91 is located between the two major conservation areas that the analyst seeks to connect. Alternatively, the analyst can visualize a human influence index (HII), as shown in Fig. 4d, which summarizes in a scale from 0 to 64 the overall influence of humans on terrestrial ecosystems. This view shows that areas near the metro and highway areas usually have high human influence index.

Meeting cost constrains is a central goal of conservation planning because resources for conservation are always limited (Naidoo et al., 2006). Fig. 4e shows the spatial distribution of costs and its relationship with existing protected areas or other attributes. The cost layer shows that the average cost to purchase land near the Big Snowy Mountains is higher than that near the Little Belt Mountains. After the exploratory spatial analysis of existing protected areas, human influence, and cost, the analyst can define a candidate region between the Big Snowy Mountains and Little Belt Mountains conservation areas in Fig. 4f using the drawing tool.

Protecting sites that are closer to existing conservation areas (both east and west) will encourage connectivity. Therefore, the analyst selected PA as an attribute for the ranking calculation and the optimization model, as well as terrestrial vertebrate species richness which includes mammals, birds, reptiles, and amphibians because the overall goal is to promote movement and gene flow of wildlife species among existing conservation areas. Mammal conservation is a regional

conservation priority, thus, using the brushed axis the analyst imposed a constraint to include sites that have a total richness of mammals index of at least 54 species. The model's goal is to maximize the area under protection while adding a budget as a constraint as it is a common practice in conservation planning (Cabeza and Moilanen, 2001; Williams et al., 2005). Acquiring the whole candidate region would cost $47M, which is higher than conservation budgets in many instances. Therefore, the analyst sets a target total cost of $0 — $10,000,000 to test if this budget range allows to meet the conservation goal of acquiring land to connect the conservation areas. The prescribed solution is shown in Fig. 5a. The figure shows that the current budget allows purchasing a limited number of isolated patches that will contribute little to the overall goal of promoting connectivity. The budget is increased to $15,000,000, obtaining the area in Fig. 5b, which better promotes connectivity between the two existing protected areas. This budget level also allows connecting the southern portion of the Little Belt Mountains.

As is common in conservation planning, the prescribed optimal set requires manual refinement by the analyst to incorporate expert opinion on attributes that are not necessarily accounted for in the optimization model. For example, a land parcel may be already zoned for other uses, it may be prone to fire or flood disturbance, or it may be spatially isolated and therefore not desirable as a conservation unit. This manual refinement is a key component of the conservation planning process that is lacking in many computational applications and is intuitively incorporated in this tool given its spatial nature. In these examples, the analyst
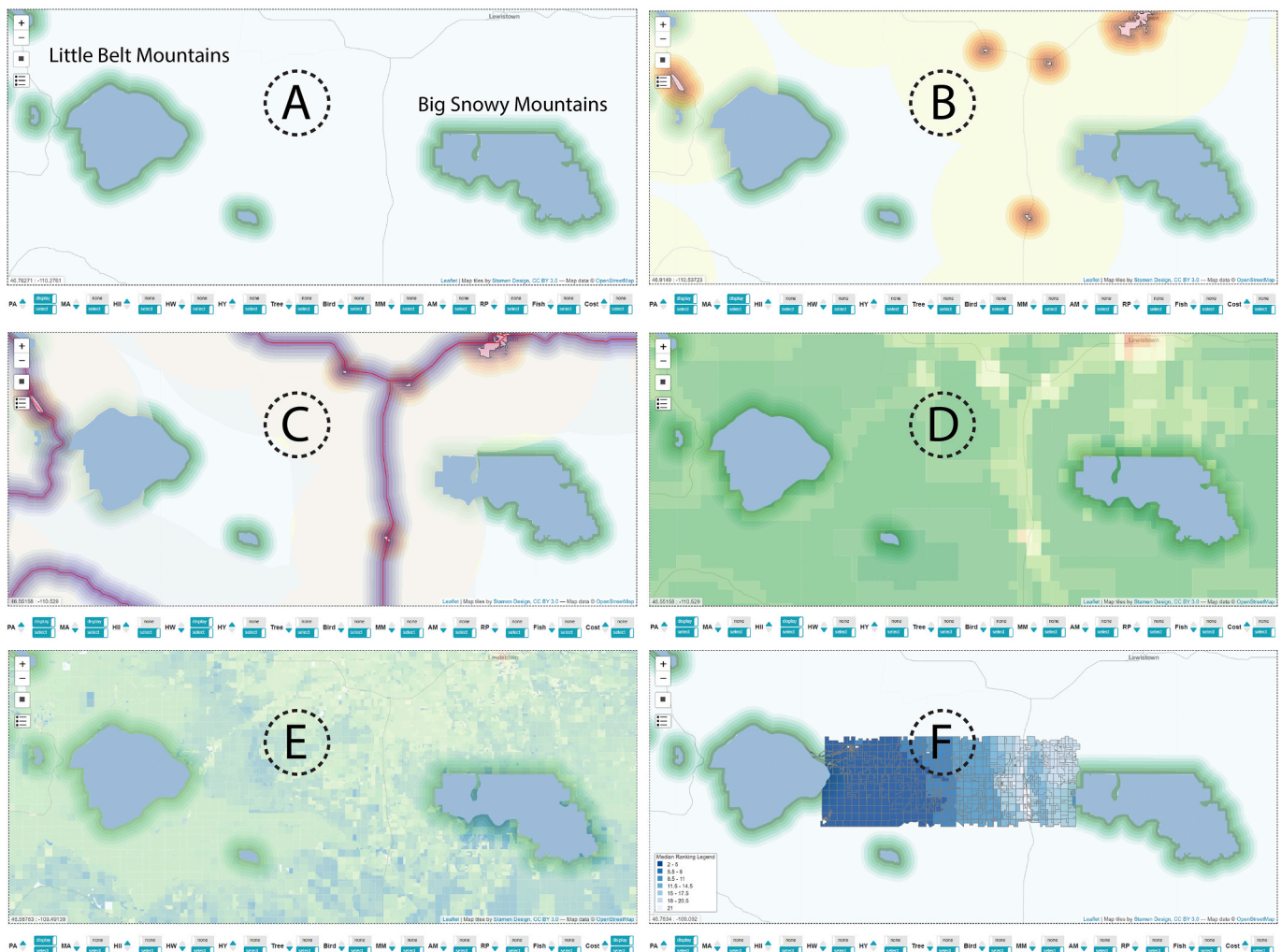

Fig. 4. Process to define the candidate region: (a) PA layer. (b) MA layer covering PA layer. (c) HW layer covering MA and PA layers. (d) HII layer covering PA layer. (e) Cost layer covering PA layer. (f) Selected candidate region with median ranking covering PA layer.

9

replaced isolated regions with little contribution wildlife movement with areas in the west that ensure connection to the Big Snowy Mountains. Because the map reports the total selected area and cost after any analyst action (e.g., selection or removal of a parcel), the analyst was able to select an area within the given budget while using the manual refinement tool. The customized portfolio in Fig. 5c results in a set of areas to protect of $\sim$ 564 km^2 and a cost of $16,142,555. Although this solution is not optimal (i.e., the optimal solution recommends the purchase of $\sim$ 589 km^2 within the same budget), it reflects the complementary insight of the mathematical model and expert judgment based on attributes such as landscape connectivity that are not included in the mathematical model.

We compare the three portfolios in the left part of Fig. 5 using the spatial and non-spatial information. The first two examples have the same search area and different budgets. The customized portfolio in the second example consists of a larger area (within its budget) than that in the first example, which exceeds its budget as shown in Fig. 5ab. The screenshot of the map shows how the parcels of each portfolio distribute. Besides the difference of constraints and goals reached by the portfolios, the change on the distribution of each attribute is visualized on the arcs in Fig. 5ab. We hover the inner arc to get the comparison result of the average attribute value among different portfolios. To show the hovered result of each attribute, we list five attributes we concern in the right part of Fig. 5ab. We observe that the customized portfolio in the second example has a higher average richness of amphibian, bird, and tree species. Moreover, it consists of a larger area within its budget to connect the two protected areas. The analyst decides that the second portfolio is better than the first one. With the same budget, we generate the third portfolio based on a larger search area. Fig. 5bc shows the comparison result of the second and the third portfolios. The third portfolio consists of a larger area within the same budget to connect three protected areas. In addition, it has a higher average richness of reptile species, and lower cost. Based on our analysis from Fig. 5, the analyst selects the third portfolio as the final choice. The comparison result of the customized and suggested portfolios in the third example, which is represented near the inner and middle arc, gives more evidence to support the analyst's decision. In Fig. 5bc, the average richness of mammal, tree and bird species is higher in the customized portfolio.

4.2. Creating a protected area in Montana's Park and Sweet Grass counties

In this section, we illustrate the creation of a protected area at the boundary of Montana's Park and Sweet Grass counties, between highways 89, 90, 191, and 371. The region of interest consists of federal land and other unprotected areas and is within 100 mi from urban areas such as Bozeman, Livingston, Big Timber, and White Sulphur Springs, as well as other unincorporated communities. While in the first case study we were interested in designing a conservation area distant from areas of human influence, in this case study we have an opposite goal. Recent studies argue for a positive role of nature parks and protected areas close to human population centers (More et al., 1988). Proximity to natural areas has been associated with improved mental health (Sturm and Cohen, 2014) and positive attitudes towards nature (Lin et al., 2014). Therefore, in this case the analyst is interested in creating conservation areas that promote the protection of biodiversity, while at the same time being accessible by the community. In addition to the mammals, reptiles, amphibians, and bird richness layers, the analyst includes the distance to metropolitan areas and the distance to highway as attributes in non-decreasing order using the Attribute Analysis View. In this way, areas closer to highways and metro areas are given a higher preference.

Using the selected attributes, Fig. 6a shows the initial ranking of areas within the region of interest. This ranking combines both biological and geographical features. Because the cost of purchasing the whole region of interest is prohibitively high ($\sim$ $32M), the analyst decided to exclude from the analysis such areas whose median ranking is larger than 5 for the selected attributes. In other words, discards those areas that are not ranked in the top five in at least half of the selected attributes. This was done using the pre-processing slider in the Optimization Configuration panel, which reduced the area from $\sim$863 km^2 to $\sim$420 km^2, with an updated total cost of $\sim$ $11.5M (see Fig. 6b). The optimization model's goal is to minimize the total purchasing cost subject to a minimum protected area of 300 km^2. The results of this baseline scenario are shown in Fig. 6c. The size of the optimal area is $\sim$ 299 km^2 with a total cost of $\sim$ $5.27M. This area is neither connected nor compact, having some isolated parcels and gaps inside the main cluster of selected areas. To improve the geographical properties of the selected area, the analyst manually induced these properties using the point-and-
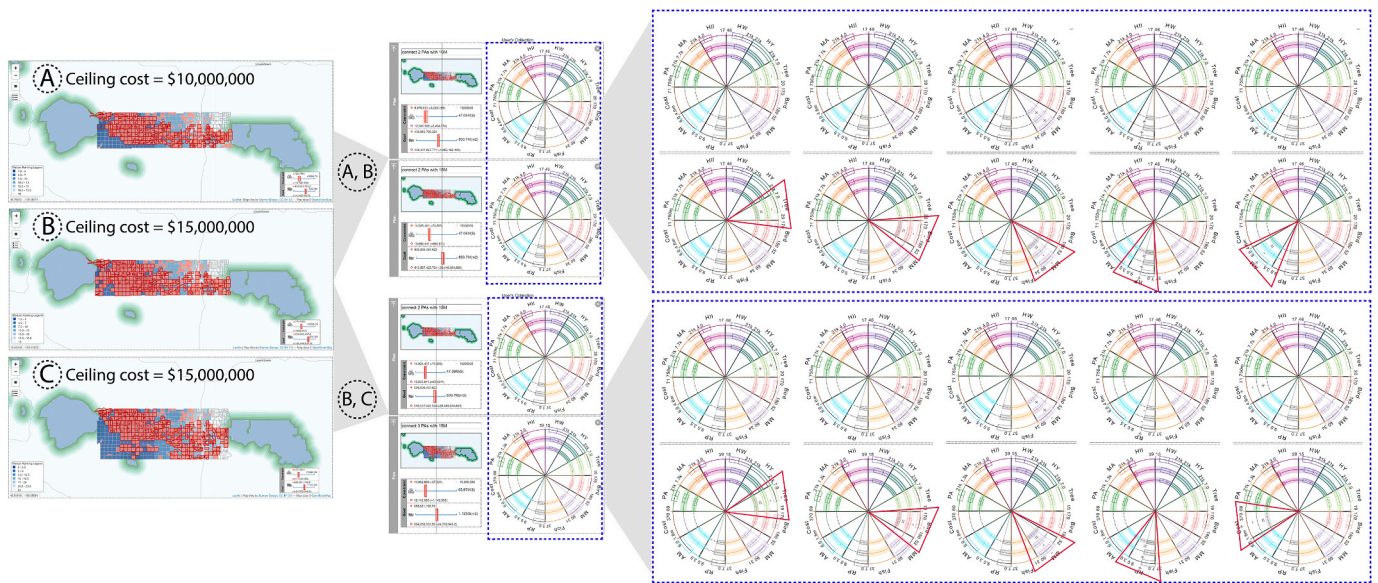

Fig. 5. Generated portfolios on the map. (A) Portfolios generated with $10M budget connecting two protected areas. (B) Portfolios generated with $15M budget connecting two protected areas. (C) Portfolios generated with $15M budget connecting three protected areas. All the parcels are prepossessed by filtering those with less than 54 mammal species richness. Portfolio comparisons: (A,B) The portfolios generated based on the small search area with the different budgets are compared. (B,C) With the same budget, the portfolios generated based on the small search area and larger search area are compared.
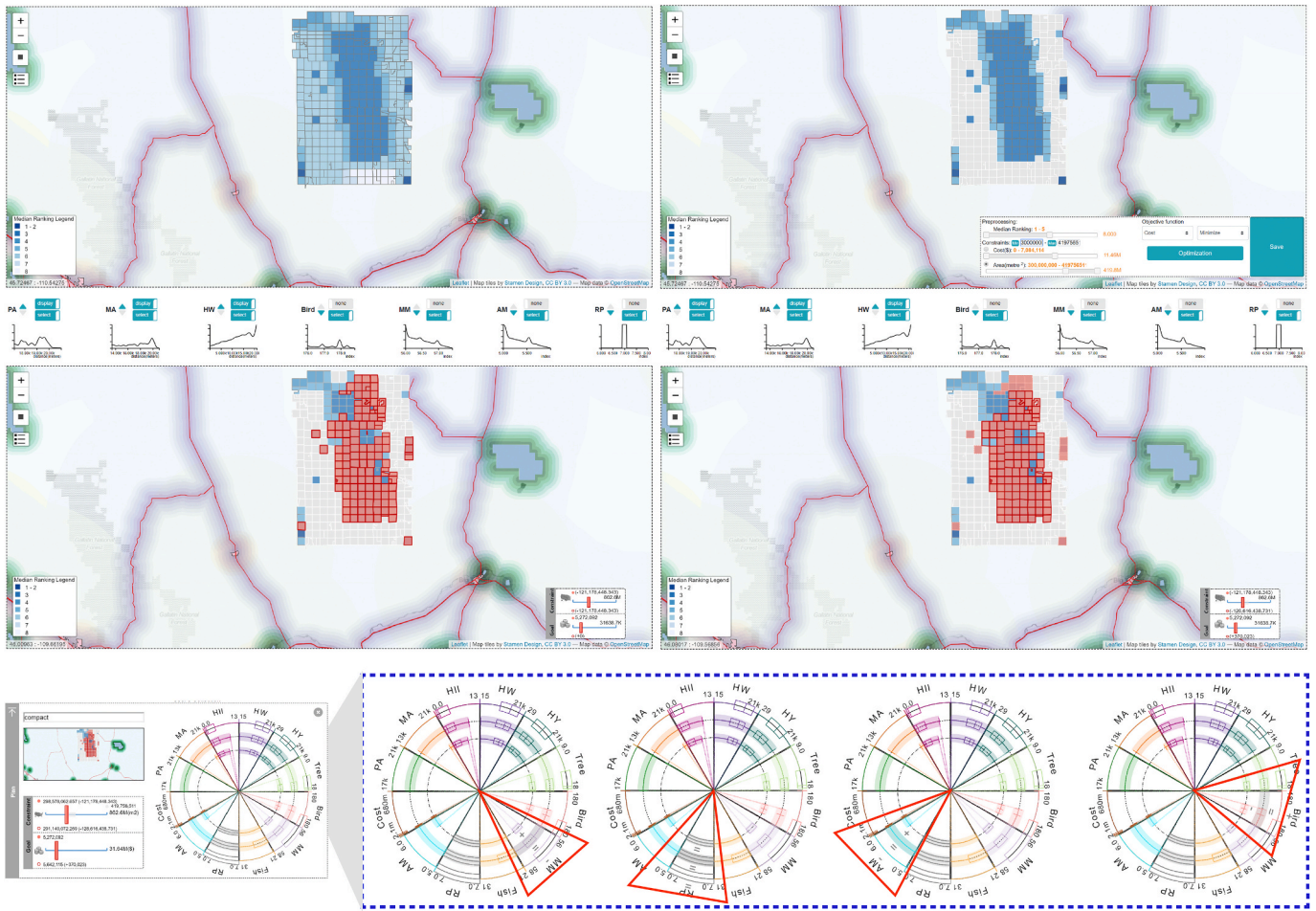
Fig. 6. Case Study 2: (A) Initial ranking given selected attributes in Montana's Park and Sweet Grass counties. (B) Pre-processed areas using median ranking. (C) Optimal results when total purchase cost is minimized subject to a minimum area of 300 km^2; (D) Results after manually inducing compactness in the protected area. (E) Comparison between the optimal portfolio and the connected and compact analyst-selected area.

click feature of our system, ultimately producing the area shown in Fig. 6d. In this case, the size of the analyst-selected area is ∼ 291 km^2 with a total cost of ∼ \$5.6M. Regarding the ecological features, the attribute comparison in Fig. 6e shows that the analyst-selected landscape has a higher average richness for mammals, reptiles, and amphibians, but not birds. In this case, the increase in some species coverage as well as the connectivity and compactness properties of the resulting landscape are achieved at the expense of a higher land purchase cost with respect to the baseline scenario (∼ \$370K).

5. Conclusions and future work

We propose a visual analytics framework to help conservation planners and scientists to explore, compare, and modify conservation portfolios under a variety of constraints. To explore the candidate parcels, our system proposes the multi-layer map view and the parallel coordinates-based attribute analysis view. The suggested portfolios and the user-defined portfolios are generated based on an optimization model and users' domain knowledge. The comparison between these portfolios is supported by the portfolio comparison view. Using our system, analysts can incorporate their decision preferences and add selection attributes that are not easily incorporated as constraints or objectives, or that delay the construction of a portfolio given the resulting model complexity. Currently, our optimization model is fast for moderately sized landscapes and allows the construction of what-if scenarios almost in real time.

Our framework has been validated by conservation experts through

two case studies, which demonstrate how our framework can help analysts to generate conservation portfolios for different goals under a variety of constraints. Moreover, our system has been received design feedback from multiple conservation experts including two co-authors and four external partners. Although the feedback received was generally positive, some limitations have been identified for future work. Specifically, analysts appreciated the option to compare portfolios; however, more automation for supporting detailed comparison could improve the analysis process. The analysts also noted that while the framework is flexible to the underlying optimization approach, an API that would allow users to directly integrate their own optimization routines could greatly enhance their workflow. A possible avenue is to explore alternative multi-objective approaches to explore the trade-off between objectives in the portfolio optimization (see, e.g., Miettinen (2012) and Sawaragi et al. (1985) for alternatives). Further work will focus on the automatic comparison of candidate portfolios and add customized algorithms to induce other spatial properties to the framework in case the user decides to use them (e.g., connectivity and compactness). An interesting conjecture is whether adding human interaction with the optimization helps with the run-time issues when spatial properties are enforced. Further studies exploring the tradeoffs between human input and ability to explore reasonable solutions is an interesting future direction. As of now, the analyst can manually load a candidate conservation portfolio for further analysis using a shapefile or a file specifying whether a parcel is selected. We will add modifications in this aspect to facilitate the upload and compatibility check of a candidate portfolio, as this will allow our system to complement the

analysis of other existing tools like MARXAN and Zonation. Although our framework focuses on conservation planning decisions, it can be extended to other spatial problems, including electoral districting, location of urban parks, and land-use planning. Such applications will require the proper data inputs and specification of the related optimization problems.

Acknowledgment

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Acevedo, M.A., Sefair, J.A., Smith, J.C., Reichert, B., Fletcher Jr., R.J., 2015. Conservation under uncertainty: optimal network protection strategies for worst-case disturbance events. J. Appl. Ecol. 52, 1588–1597.

Afzal, S., Maciejewski, R., Ebert, D.S., 2011. Visual analytics decision support environment for epidemic modeling and response evaluation. In: Visual Analytics Science and Technology. IEEE, pp. 191–200.

Ailon, N., Charikar, M., Newman, A., 2008. Aggregating inconsistent information: ranking and clustering. J. ACM 55, 23.

Ando, A., Camm, J., Polasky, S., Solow, A., 1998. Species distributions, land values, and efficient conservation. Science 279, 2126–2128.

Arthur, J.L., Hachey, M., Sahr, K., Huso, M., Kiester, A., 1997. Finding all optimal solutions to the reserve site selection problem: formulation and computational analysis. Environ. Ecol. Stat. 4, 153–165.

Ball, I.R., Possingham, H.P., Watts, M., 2009. Marxan and relatives: software for spatial conservation prioritisation. Spatial conservation prioritisation: Quantitative methods and computational tools 185–195.

Belle, E., Kingston, N., Burgess, N., Sandwith, T., Ali, N., MacKinnon, K., 2018. Protected Planet Report 2018. UNEP-WCMC, IUCN and NGS, Cambridge UK, Gland Switzerland and Washington, DC, USA. URL: https://livereport.protectedplanet.net/pdf/Protected_Planet_Report_2018.pdf.

Bicknell, J.E., Collins, M.B., Pickles, R.S., McCann, N.P., Bernard, C.R., Fernandes, D.J., Miller, M.G., James, S.M., Williams, A.U., Struebig, M.J., et al., 2017. Designing protected area networks that translate international conservation commitments into national action. Biol. Conserv. 214, 168–175.

Brownell, J.L., 1987. The Genesis of Wildlife Conservation in Montana. Ph.D. thesis. Montana State University-Bozeman, College of Letters & Science.

Cabeza, M., Moilanen, A., 2001. Design of reserve networks and the persistence of biodiversity. Trends Ecol. Evol. 16, 242–248.

Camm, J.D., Norman, S.K., Polasky, S., Solow, A.R., 2002. Nature reserve site selection to maximize expected species covered. Oper. Res. 50, 946–955.

Cassol, V.J., Testa, E.S., Jung, C.R., Usman, M., Faloutsos, P., Berseth, G., Kapadia, M., Badler, N.I., Musse, S.R., 2017. Evaluating and optimizing evacuation plans for crowd egress. IEEE Computer Graphics & Applications 60–71.

Church, R.L., Stoms, D.M., Davis, F.W., 1996. Reserve selection as a maximal covering location. Biol. Conserv. 76, 103–112.

Cocks, K., Baird, I.A., 1989. Using mathematical programming to address the multiple reserve selection problem: an example from the eyre peninsula, south Australia. Biol. Conserv. 49, 113–130.

Dasgupta, A., Poco, J., Wei, Y., Cook, R., Bertini, E., Silva, C.T., 2015. Bridging theory with practice: an exploratory study of visualization use and design for climate model comparison. IEEE Trans. Visual. Comput. Graph. 996–1014.

Dilkina, B., Gomes, C.P., 2010. Solving connected subgraph problems in wildlife conservation. In: International Conference on Integration of Artificial Intelligence (AI) and Operations Research (OR) Techniques in Constraint Programming. Springer, pp. 102–116.

Dinerstein, E., Vynne, C., Sala, E., Joshi, A., Fernando, S., Lovejoy, T., Mayorga, J., Olson, D., Asner, G., Baillie, J., Burgess, N., Burkart, K., Noss, R., Zhang, Y., Baccini, A., Birch, T., Hahn, N., Joppa, L., Wikramanayake, E., 2019. A global deal for nature: guiding principles, milestones, and targets. Science Advances 5, eaaw2869.

Dissanayake, S.T., Önal, H., Westervelt, J.D., Balbach, H.E., 2012. Incorporating species relocation in reserve design models: an example from ft. benning ga. Ecol. Model. 224, 65–75.

Dragicevic, S., 2004. The potential of web-based gis. J. Geogr. Syst. 6, 79–81.

Gleicher, M., 2018. Considerations for visualizing comparison. IEEE Trans. Visual. Comput. Graph. 24, 413–423.

Hodgson, J.A., Thomas, C.D., Wintle, B.A., Moilanen, A., 2009. Climate change, connectivity and conservation decision making: back to basics. J. Appl. Ecol. 46, 964–969.

Jafari, N., Hearne, J., 2013. A new method to solve the fully connected reserve network design problem. Eur. J. Oper. Res. 231, 202–209.

Jenkins, C.N., 2017. GIS layers of biodiversity data. URL: https://biodiversitymapping.org/wordpress/index.php/download/.InstitutodePesquisasEcológicas. (Accessed 10 August 2018).

Johansson, J., Cooper, M., Jern, M., 2005. 3-dimensional display for clustered multi-relational parallel coordinates. In: IEEE. Information Visualisation, pp. 188–193.

Johnson, C.N., Balmford, A., Brook, B.W., Buettel, J.C., Galetti, M., Guangchun, L., Wilmshurst, J.M., 2017. Biodiversity losses and conservation responses in the anthropocene. Science 356, 270–275. https://doi.org/10.1126/science.aam9317. URL: https://science.sciencemag.org/content/356/6335/270 https://science.sciencemag.org/content/356/6335/270.full.pdf. arXiv:

Kehrer, J., Piringer, H., Berger, W., Gröller, M.E., 2013. A model for structure-based comparison of many categories in small-multiple displays. IEEE Trans. Visual. Comput. Graph. 19, 2287–2296.

Kirkpatrick, J., 1983. An iterative method for establishing priorities for the selection of nature reserves: an example from tasmania. Biol. Conserv. 25, 127–134.

Konev, A., Waser, J., Sadransky, B., Cornel, D., Perdigao, R., Horváth, Z., Groller, M., 2014. Run watchers: automatic simulation-based decision support in flood management. IEEE Trans. Visual. Comput. Graph. 1–1.

Law, P.M., Basole, R.C., Wu, Y., 2018. Duet: helping data analysis novices conduct pairwise comparisons by minimal specification. IEEE Trans. Visual. Comput. Graph.

Lehtomäki, J., Moilanen, A., 2013. Methods and workflow for spatial conservation prioritization using zonation. Environ. Model. Software 47, 128–137.

Lin, B.B., Fuller, R.A., Bush, R., Gaston, K.J., Shanahan, D.F., 2014. Opportunity or orientation? who uses urban parks and why. PloS One 9, e87422.

Lin, S., 2010. Rank aggregation methods. Wiley Interdisciplinary Reviews: Comput. Stat. 2, 555–570.

Lind, M., Johansson, J., Cooper, M., 2009. Many-to-many relational parallel coordinates displays. In: Visualisation, Information (Ed.), IEEE, pp. 25–31.

Loveland Tech, 2018. LOVELAND: mapping every parcel on the planet. URL: https://landgrid.com/reports/parcels. (Accessed 10 August 2018).

Luck, G.W., Chan, K.M., Klien, C.J., 2012. Identifying spatial priorities for protecting ecosystem services. F1000Research 1.

Mace, G.M., Norris, K., Fitter, A.H., 2012. Biodiversity and ecosystem services: a multilayered relationship. Trends Ecol. Evol. 27, 19–26.

Margules, C.R., Nicholls, A., Pressey, R., 1988. Selecting networks of reserves to maximise biological diversity. Biol. Conserv. 43, 63–76.

Margules, C.R., Pressey, R.L., 2000. Systematic conservation planning. Nature 405, 243.

Mcdonald, R.I., Forman, R.T., Kareiva, P., Neugarten, R., Salzer, D., Fisher, J., 2009. Urban effects, distance, and protected areas in an urbanizing world. Landsc. Urban Plann. 93, 63–75.

Miettinen, K., 2012. Nonlinear Multiobjective Optimization, ume 12. Springer Science & Business Media.

Moilanen, A., Wilson, K.A., Possingham, H., 2009. Spatial Conservation Prioritization: Quantitative Methods and Computational Tools. Oxford University Press.

More, T.A., Stevens, T., Allen, P.G., 1988. Valuation of urban parks. Landsc. Urban Plann. 15, 139–152.

Munzner, T., Guimbretiére, F., Tasiran, S., Zhang, L., Zhou, Y., 2003. Treejuxtaposer: scalable tree comparison using focus+ context with guaranteed visibility. In: ACM Transactions on Graphics (TOG), ACM, pp. 453–462.

Naidoo, R., Balmford, A., Ferraro, P.J., Polasky, S., Ricketts, T.H., Rouget, M., 2006. Integrating economic costs into conservation planning. Trends Ecol. Evol. 21, 681–687.

Nalle, D.J., Arthur, J.L., Sessions, J., 2002. Designing compact and contiguous reserve networks with a hybrid heuristic algorithm. For. Sci. 48, 59–68.

Önal, H., Briers, R.A., 2002. Incorporating spatial criteria in optimum reserve network selection. Proc. R. Soc. Lond. B Biol. Sci. 269, 2437–2441.

Önal, H., Briers, R.A., 2006. Optimal selection of a connected reserve network. Oper. Res. 54, 379–388.

Pajer, S., Streit, M., Torsney-Weir, T., Spechtenhauser, F., Möller, T., Piringer, H., 2017. Weightlifter: visual weight space exploration for multi-criteria decision making. IEEE Trans. Visual. Comput. Graph. 23, 611–620.

Pettit, C.J., Raymond, C.M., Bryan, B.A., Lewis, H., 2011. Identifying strengths and weaknesses of landscape visualisation for effective communication of future alternatives. Landsc. Urban Plann. 100, 231–241.

Pimm, S.L., Russell, G.J., Gittleman, J.L., Brooks, T.M., 1995. The future of biodiversity. Science 269, 347–350.

Polasky, S., Camm, J.D., Garber-Yonts, B., 2001. Selecting biological reserves cost-effectively: an application to terrestrial vertebrate conservation in Oregon. Land Econ. 77, 68–78.

Portman, M.E., 2014. Visualization for planning and management of oceans and coasts. Ocean Coast Manag. 98, 176–185.

Pressey, R., Possingham, H., Day, J., 1997. Effectiveness of alternative heuristic algorithms for identifying indicative minimum requirements for conservation reserves. Biol. Conserv. 80, 207–219.

Rao, M., Fan, G., Thomas, J., Cherian, G., Chudiwale, V., Awawdeh, M., 2007. A web-based gis decision support system for managing and planning usda's conservation reserve program (crp). Environ. Model. Software 22, 1270–1280.

Rinner, C., 2007. A geographic visualization approach to multi-criteria evaluation of urban quality of life. Int. J. Geogr. Inf. Sci. 21, 907–919.

Rodrigues, A.S., Akcakaya, H.R., Andelman, S.J., Bakarr, M.I., Boitani, L., Brooks, T.M., Chanson, J.S., Fishpool, L.D., Da Fonseca, G.A., Gaston, K.J., et al., 2004. Global gap analysis: priority regions for expanding the global protected-area network. Bioscience 54, 1092–1100.

Rosenbaum, R., Zhi, J., Hamann, B., 2012. Progressive parallel coordinates. In: Pacific Visualization Symposium. IEEE, pp. 25–32.

Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., et al., 2000. Global biodiversity scenarios for the year 2100. Science 287, 1770–1774.

Sanderson, E., Jaiteh, M., Levy, M., Redford, K., Wannebo, A., Woolmer, G., 2003. The human footprint and the last of the wild. Bioscience 52, 891–904.

Sarkar, S., Pressey, R.L., Faith, D.P., Margules, C.R., Fuller, T., Stoms, D.M., Moffett, A., Wilson, K.A., Williams, K.J., Williams, P.H., et al., 2006. Biodiversity conservation planning tools: present status and challenges for the future. Annu. Rev. Environ. Resour. 31.

Sawaragi, Y., Nakayama, H., Tanino, T., 1985. Theory of Multiobjective Optimization. Elsevier.

Sculley, D., 2007. Rank aggregation for similar items. In: Proceedings of the 2007 SIAM International Conference on Data Mining. SIAM, pp. 587–592.

SEDAC, 2018. Last of the wild (geographic), v2 (1995–2004). URL: http://sedac.ciesin.columbia.edu/data/set/wildareas-v2-last-of-the-wild-geographic/data-download. Socioeconomic Data and Applications Center. (Accessed 10 August 2018).

Sefair, J.A., Smith, J.C., Acevedo, M.A., Fletcher Jr., R.J., 2017. A defender-attacker model and algorithm for maximizing weighted expected hitting time with application to conservation planning. IISE Transactions 49, 1112–1128.

State of Montana, 2018. Geographic information clearinghouse: administrative boundaries. URL: http://geoinfo.msl.mt.gov/Home/msdi/administrative_boundaries. (Accessed 10 August 2018).

Stokstad, E., 2010. Despite progress, biodiversity declines. Science 329, 1272–1273. https://doi.org/10.1126/science.329.5997.1272. URL: https://science.sciencemag.org/content/329/5997/1272 https://science.sciencemag.org/content/329/5997/1272.full.pdf. arXiv:

Sturm, R., Cohen, D., 2014. Proximity to urban parks and mental health. J. Ment. Health Pol. Econ. 17, 19.

Toregas, C., Revelle, C., 1973. Binary logic solutions to a class of location problem. Geogr. Anal. 5, 145–155.

Tress, B., Tress, G., 2003. Scenario visualisation for participatory landscape planning—a study from Denmark. Landsc. Urban Plann. 64, 161–178.

Turkay, C., Slingsby, A., Hauser, H., Wood, J., Dykes, J., 2014. Attribute signatures: dynamic visual summaries for analyzing multivariate geographical data. IEEE Trans. Visual. Comput. Graph. 2033–2042.

Underhill, L., 1994. Optimal and suboptimal reserve selection algorithms. Biol. Conserv. 70, 85–87.

Unwin, D.J., 1996. Gis, spatial analysis and spatial statistics. Prog. Hum. Geogr. 20, 540–551.

USCB, 2017. TIGER/Line® shapefiles: Roads. URL: https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2017&layergroup=Roads. (Accessed 24 May 2020).

USGS, 2018a. The national map. URL: http://prd-tnm.s3-website-us-west-2.amazonaws.com/?prefix=StagedProducts/Hydrography/NHD/State/HighResolution/Shape/. United States Geological Service. (Accessed 10 August 2018).

USGS, 2018b. Protected areas database of the United States (PAD-US) data download. URL: https://gapanalysis.usgs.gov/padus/data/download/. United States Geological Service. (Accessed 10 August 2018).

Watson, J.E., Venter, O., 2017. Ecology: a global plan for nature conservation. Nature 550, 48–49.

Weng, D., Chen, R., Deng, Z., Wu, F., Chen, J., Wu, Y., 2018a. Srvis: towards better spatial integration in ranking visualization. IEEE Trans. Visual. Comput. Graph.

Weng, D., Zhu, H., Bao, J., Zheng, Y., Wu, Y., 2018b. Homefinder revisited: finding ideal homes with reachability-centric multi-criteria decision making. In: Proceedings of the CHI Conference on Human Factors in Computing Systems. ACM, p. 247.

White, D., Wutich, A., Larson, K., Gober, P., Lant, T., Senneville, C., 2010. Credibility, salience, and legitimacy of boundary objects: water managers' assessment of a simulation model in an immersive decision theater. Sci. Publ. Pol. 37, 219–232.

Williams, J., ReVelle, C., 1996. A 0–1 programming approach to delineating protected reserves. Environ. Plann. Plann. Des. 23, 607–624.

Williams, J.C., ReVelle, C.S., Levin, S.A., 2005. Spatial attributes and reserve design models: a review. Environ. Model. Assess. 10, 163–181.

Xie, C., Zhong, W., Mueller, K., 2017. A visual analytics approach for categorical joint distribution reconstruction from marginal projections. IEEE Trans. Visual. Comput. Graph. 23, 51–60.

Zhang, R., Lu, Y., Adams, K., Sefair, J.A., Mellin, H., Acevedod, M.A., Maciejewski, R., 2021a. A Visual Analytics Framework for Conservation Planning. https://github.com/VADERASU/A-Visual-Analytics-Framework-for-Conservation-Planning.

Zhang, R., Lu, Y., Adams, K., Sefair, J.A., Mellin, H., Acevedod, M.A., Maciejewski, R., 2021b. A Visual Analytics Framework for Conservation Planning. https://www.youtube.com/watch?v=NU5QoRY3ZhE.

Rui Zhang is a Ph.D. student in the School of Computing and Augmented Intelligence at Arizona State University. She received her B.S. degree from Zheng Zhou University, Henan, China, M.S. degree from Beijing University of Posts and Telecommunications, Beijing, China. Her research interests include Data Visualization and geovisualization.

Yafeng Lu is a Postdoctoral Research Assistant in the School of Computing and Augmented Intelligence at Arizona State University. Her current research interests are in social media and predictive visual analytics.

Katherine Adams completed a bachelor's and master's degree in Industrial Engineering at Arizona State University. She is currently pursuing a doctoral degree in Industrial Engineering at the University of Wisconsin-Madison. Her research interests include multi-objective optimization and decision-making applied to sustainability, healthcare, and social good.

Jorge A. Sefair is an Assistant Professor of Industrial Engineering at Arizona State University. His research interests include network optimization, multistage optimization, and combinatorial optimization, with applications in environment, public policy, and transportation.

Haley Mellin is an artist and conservationist. She received her Ph.D. from the Steinhardt School at NYU with a digital-visual specialization. Her research interests include large-scale conservation with an emphasis on core reproduction zones and climate-balancing forests.

Miguel A. Acevedo is an Assistant Professor of Quantitative Ecology in the Department of Wildlife Ecology and Conservation at the University of Florida. His research interests include conservation planning, spatial ecology and disease ecology all under a mathematical and statistical framework.

Ross Maciejewski is an Associate Professor of Computer Science at Arizona State University. His primary research interests are in the areas of geographical visualization, social media mining, and predictive analytics.