Bee Population Visualization

Georgia Argyrou (102807651)

December 12, 2024

Abstract

The current report is a description behind the story that we like to say with a visualization about bees population. It is an endeavor to represent and analyze the decline in bees' population by examining the factors that lead to it as well as the importance of this decline for the US. The current visualization was created as a project for Aalto's course "CS-E4450 - Explorative Information Visualization" for the year 2024-2025, under the supervision of Prof. Tomi Kauppinen.

1 Motivation

Bees play a crucial role in maintaining biodiversity and ensuring the health of our ecosystems. As primary pollinators, they significantly impact the production of a wide range of crops and the stability of natural habitats. However, recent reports have highlighted alarming declines in bee populations globally [ZA21], driven by factors such as pesticide use, habitat loss, climate change, and disease. Understanding these trends and their underlying causes is essential for implementing effective conservation strategies and ensuring the sustainability of both agricultural and natural ecosystems. The proposed "Bee Population Visualization" aims to provide a comprehensive visualization of bee population trends, enabling stakeholders to identify critical areas of concern and track the effectiveness of conservation efforts. By integrating data from NASS(National Agricultural Statistics Service), the dashboard will help raise awareness about the importance of bees, support informed decision-making, and promote actions to protect these vital pollinators.

2 Approach

For the visualizations in this dashboard, we aimed to create an intuitive, engaging, and responsive way to explore and understand bee population trends and colony health over time across different states of US. Here is an overview of our approach to design, interaction, and technical implementation:

2.1 Data Analysis and Preparation

We began by examining the dataset, which contains information about bee populations, colony losses, and potential causes across different states and quarters from 2015 to 2022. Key metrics like average colony numbers, quarterly losses, and impact factors (e.g., pests, diseases, pesticides) were identified as primary variables to visualize. To facilitate trend analysis, we calculated year-over-year changes and organized data by state and season, allowing viewers to filter and compare data efficiently.

2.2 Design of the Components

Our approach prioritized accessibility by breaking down complex data into focused and interpretable visuals:

- Choropleth Map: To show geographical trends in bee populations, we implemented a choropleth map of the United States. This map highlights each state's average colony count per year, allowing for quick geographical comparisons. To improve visibility, we adjusted color gradients to accentuate differences across regions.

- Line Chart for Colony Loss Over Time: For temporal trends, we used a line chart that tracks colony loss year-over-year, enabling users to observe patterns over time and across states. This helps in identifying state-specific trends or national shifts in colony losses.
- Bar Chart for Seasonal Colony Condition: We employed a bar chart to illustrate the number of colonies by season within each selected year.
- Bar Chart for Causes of Colony Loss: We created a time dynamic bar chart to illustrate the primar causes of bees' colonies loss. The main reasons examined are "varroa mites", "other pests and parasites", "diseases", "pesticides", "other reasons" and "unknown reasons".

2.3 Interactive Elements and Filters

To make the dashboard user-centric, we incorporated several interactive components:

- Year Slider: The slider allows users to select a year and instantly update all visualizations. This interactivity encourages comparisons across different years and highlights annual changes.
- State Filter Dropdown: Users can select specific states to focus on regional data or analyze trends across multiple states simultaneously. This filtering ensures that insights can be tailored to specific regions of interest.

2.4 Expectations

The visualization is expected to reveal several key insights:

- Population Trends: Clear trends in the growth or decline of bee colonies over time, potentially identifying periods of significant change.
- Regional Variations: Geographic variations in bee population health, highlighting regions where conservation efforts may be needed most.
- Influencing Factors: Relationships between bee population changes and external factors such as pesticide use and habitat loss, aiding in the identification of key threats.

3 Themes and Topics

Some of the most permanent themes that are interesting to include in the visualization are going to be discussed below. These topics constitute the goal without being restricted by the availability of data.

- Bees' population: This theme focuses on the overall trends in bee populations, showing fluctuations in the number of bees over time and across regions. It highlights critical periods of decline or recovery, helping to understand the extent of bee population challenges globally.
- Bees' species: Visualizing the health and status of different bee species (honeybees, bumblebees, solitary bees) allows for a better understanding of species-specific vulnerabilities.
- Countries: Comparing bee population trends across countries highlights the geographical variations in bee health. This theme can reveal which regions are facing the greatest changes.
- Pesticides: Examining the use of pesticides, particularly harmful chemicals like neonicotinoids, helps to understand their impact on bee populations. Visualizing pesticide use alongside population trends can reveal potential links between chemical exposure and bee declines
- Climate factors: Climate data, including temperature changes, rainfall, and extreme weather events, are critical to understanding how environmental shifts impact bee populations. This theme shows how changes in local climates affect bee activity, food availability, and colony survival.

- Diseases: Visualizing the spread and impact of diseases like Colony Collapse Disorder (CCD), Varroa mites, and other pathogens allows for a deeper understanding of health challenges faced by bee colonies. This theme can reveal disease hotspots and show correlations with declines in bee populations.
- Habitat changes/ loss: This theme highlights the effects of habitat destruction, such as deforestation, urbanization, and agricultural expansion, on bee populations. It illustrates how shrinking natural habitats reduce bees' access to food and nesting sites, contributing to population declines.

4 Tools

In this project, we used several tools and libraries to create an interactive and visually appealing dashboard for visualizing bee population data. More specifically, we used the following tools:

1. Python: Python was our primary programming language for data processing, visualization, and building the interactive dashboard.
2. Pandas: We used Pandas to load, clean, and process the data, enabling us to filter by year, state, and specific metrics (like colony loss).
3. Plotly: We used Plotly to create a variety of visualizations, including choropleth maps, bar charts, line charts and scatter plots. These visualizations are interactive, allowing users to hover over data points for more information, zoom into specific areas, and see patterns and trends.
4. Dash: Dash enabled us to build an interactive dashboard where users can explore data by selecting specific years and states.
5. CSS and HTML: We used HTML and CSS to structure and style our dashboard respectively.

5 Information Usability

We believe that environmentalists, researchers, and policymakers will gain a comprehensive understanding of bee population trends when using our Bee Population Visualization tool. Through interactive maps and timeline graphs, they will learn how bee populations fluctuate across different regions and how these fluctuations correlate with environmental factors like pesticide use or habitat loss.

When doing specific tasks, such as exploring regional trends by zooming in on a particular state, comparing population changes over time using the timeline slider, or correlating bee declines with external factors like pesticide use via overlays, users will be able to pinpoint critical periods or regions where conservation efforts may be needed. These tasks help users identify trends, threats, and potential solutions to the bee population decline.

In addition to the use of the visualization of experts and people relative to the subject, we would like for it to be used by everyday people that find an interest in the topic or want to learn more and be motivated to participate in the protection of the species.

We will know that our visualization is effective when we see data showing users frequently engaging with these key features, such as selecting regions on the map, adjusting timelines, and turning on/off correlation overlays. Additionally, feedback from users regarding improved decision-making or policy proposals based on their findings will confirm that visualization successfully aids in addressing bee population challenges.

6 Human-Computer Interaction

Users interact with the Bee Population Visualization through several intuitive tools, as seen in Figure 1.The core features include:

- Interactive Map: Users can hover over or click on different states to see the bee colony distribution, with color-coded regions representing colony density. A tooltip provides specific colony counts for each state. This map is paired with a zoom feature that allows users to dive deeper into specific geographic areas or compare adjacent states.

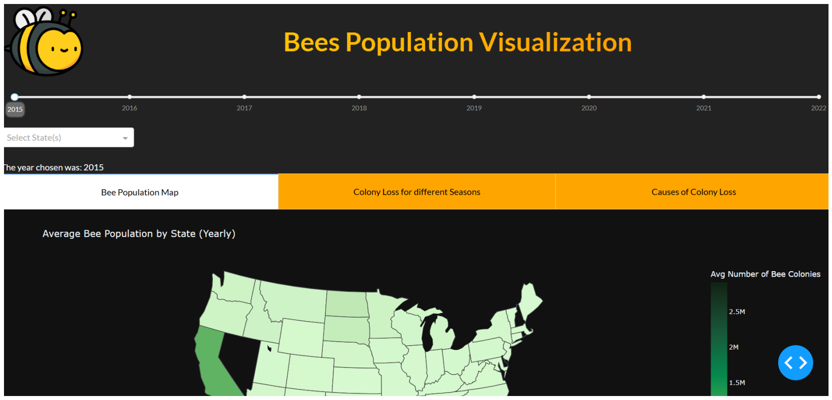


Figure 1: Dashboard's Main Page

- Timeline Slider: A timeline slider at the bottom of the dashboard enables users to view how bee populations have changed. By dragging the slider, users can move through the years and see real-time updates on both the map and any connected charts.
- Tabs: Users can change tabs to check each time the infromation they are interested in.
- States Selection: Users are able to choose more than one states to compare them.

In terms of time needed, initial exploration, where users might interact with the map and familiarize themselves with basic bee population trends, typically takes around 5-10 minutes. More detailed tasks, such as selecting a specific state, analyzing how bee population changes relate to pesticide use, or comparing trends over two decades, could take between 15-30 minutes, depending on how many variables the user decides to explore.

7 Datasets

The data driving the Bee Population Visualization comes primarily from the United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS), covering the years 2015 to 2022. The dataset includes comprehensive statistics on bee colonies, hive counts, and honey production across various U.S. states.It is important that the dataset is reliable, with observations provided directly by beekeepers and agricultural bodies. Issues like missing data are handled by employing statistical interpolation techniques where possible.

8 Time

Time in my visualization signifies the change in bee populations, hives number but also the passage of time when no change happens. We aim to identify patterns in these variables over time, highlighting the yearly impacts on bee populations and hives due to factors such as pesticide use, diseases, pests, and parasites. Beyond the traditional timeline, we also want to show these changes in relation to the growth of other simultaneous factors, providing a richer context for understanding the dynamics affecting bee health and hive sustainability in time.

The story we're telling focuses on the cumulative impact of these environmental and biological pressures on bee health. By layering these changes against other concurrent variables, such as agricultural growth, climate variations, or shifts in pesticide usage, we'll visually explore how these elements interact with and amplify each other. This approach lets us go beyond a simple timeline, giving a nuanced view of how each year's events contribute to long-term trends in bee population sustainability.

Following the three dimensions of our design space for expressive storytelling with timelines: representation, scale, and layout, presented in [BLB+17], we can analyze the time components of our visualization. A linear representation 2 might work well, as these will allow viewers to track changes over the years in an easily understandable format. Regarding the scale it will be chronological between different events. We could also implement a sequential scale that demonstrates the relationship

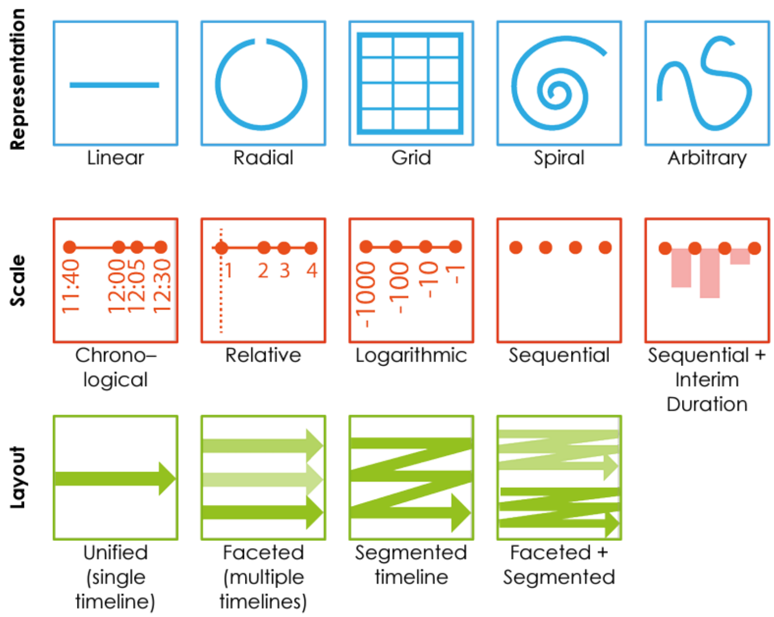


Figure 2: The three dimensions of design space for expressive storytelling with timelines: representation, scale, and layout.[BLB+17]

between certain events such as extreme pesticide use and decrement in population. We will also consider a faceted layout to separate different variables. This could allow viewers to see how each factor changes over time and make comparisons across categories. Alternatively, a faceted and segmented layout could highlight distinct periods or trends

9 Space

In creating a dashboard to visualize the trends in bee population health, we considered spatial design elements to ensure that users can quickly interpret patterns and spot key geographic insights. In this section, an overview of our approach to space in the visualization is presented, focusing on our selection of map elements, symbols, color coding, and scale.

First and foremost, a simple map constitutes the most elegant way to represent our data, as presented in 3. The choropleth map at the heart of the dashboard provides a geographic breakdown of bee colony populations by state, enabling users to understand regional variations at a glance. We selected a U.S. map with state-level detail, keeping the boundaries of each state visible. This layout allows us to present granular data for individual states while maintaining the continuity of the broader national context. The map is centrally positioned in the dashboard to draw immediate focus. We adjusted zoom limitations to prevent over-zooming, which helps users maintain a consistent view of the data's scope without losing context.

States are colored based on the average number of bee colonies, which dynamically updates based on the selected year. The use of a color gradient enhances contrast across different states, quickly highlighting states with higher or lower bee populations.

10 Use Cases

The Bee Population Visualization dashboard provides users with interactive tools to explore trends in bee populations, analyze colony health, and understand the factors affecting bee sustainability. Below are two main use cases illustrating how users can interact with the dashboard:

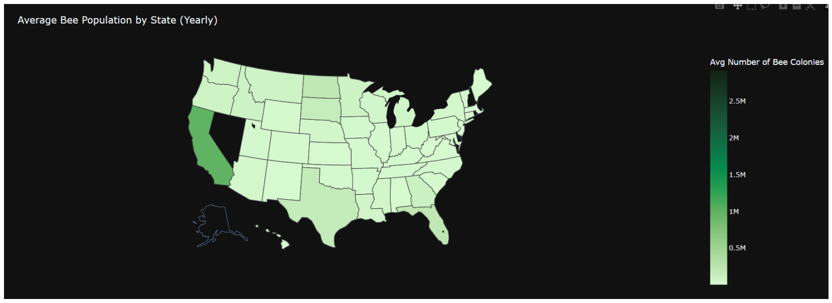


Figure 3: Map for Bees Population

10.1 No Selected States

In this scenario, the user explores the dashboard without selecting specific states, focusing instead on national trends.

- **Analyzing National Trends:** A user, such as an environmental researcher, selects a year using the timeline slider to observe nationwide bee population trends. The choropleth map displays the average number of colonies per state as seen in Figure 4a, while the line chart reveals overall colony losses across the United States as observed in Figure 4b. The user can identify patterns such as years with significant colony losses and compare them with causes of loss (e.g., pesticide use or diseases) visualized in the bar chart.

- **Examining Seasonal Impact:** By observing the bar chart for seasonal colony losses in Figure 4d, the user identifies which seasons experience the highest losses. This insight helps uncover the influence of environmental or agricultural factors on bee populations during different times of the year.

- **Evaluating Causes of Colony Loss:** Using the bar chart for colony loss causes presented in Figure 4c, the user explores the relative impact of factors like varroa mites, pesticides, and diseases on national bee populations in the selected year. The user can spot which factor has the highest contribution to colony losses and track changes across years.

10.2 Selected States

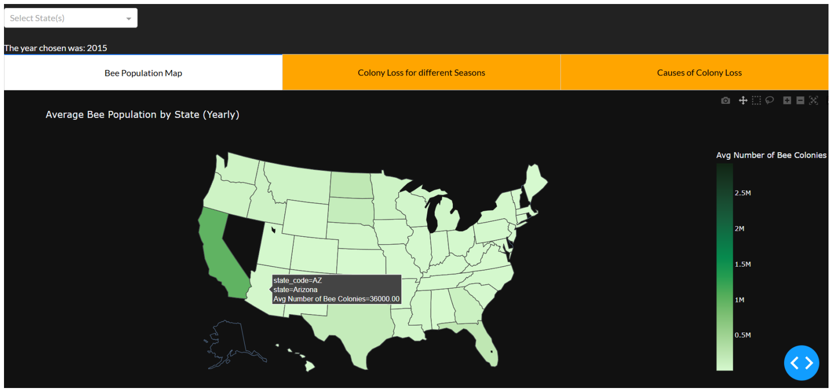
Here, the user focuses on specific states, comparing their data with national trends or other states.

- **Comparing Regional Trends:** A policymaker selects multiple states using the dropdown menu to compare regional differences in bee colony numbers. The choropleth map updates to highlight the selected states, and the line chart provides a state-specific view of colony losses over time as seen in Figure 5a. This helps identify states with higher or lower rates of decline, supporting targeted conservation efforts.

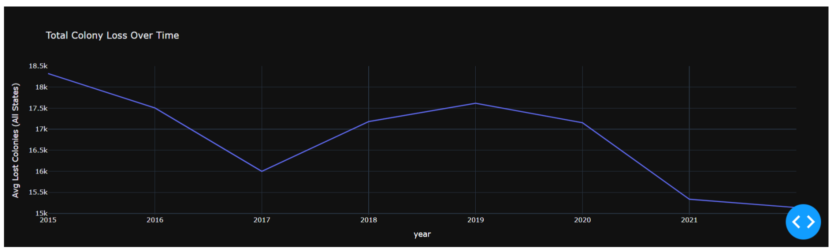
- **Exploring Seasonal Variations:** By focusing on the bar chart for seasonal colony losses provided in Figure 5d, the user compares how seasons impact bee populations across the selected states. For instance, a user may observe that certain states experience higher losses during winter due to harsher climates.

- **Investigating Local Causes of Loss:** The bar chart for causes of colony loss updates dynamically to reflect the selected states, as seen in Figure 5c. The user evaluates the primary factors contributing to bee population declines in specific regions, such as pesticide use in agricultural states or disease prevalence in others. This information supports regional initiatives to address these challenges.

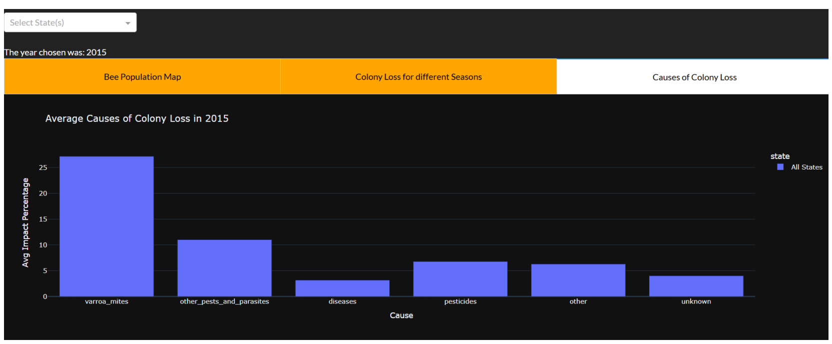
These use cases demonstrate the versatility of the Bee Population Visualization dashboard, catering to diverse audiences such as researchers, policymakers, environmental advocates, and the general public. By enabling both broad and granular analyses, the tool supports informed decision-making and awareness campaigns to protect bee populations and their vital role in ecosystems.



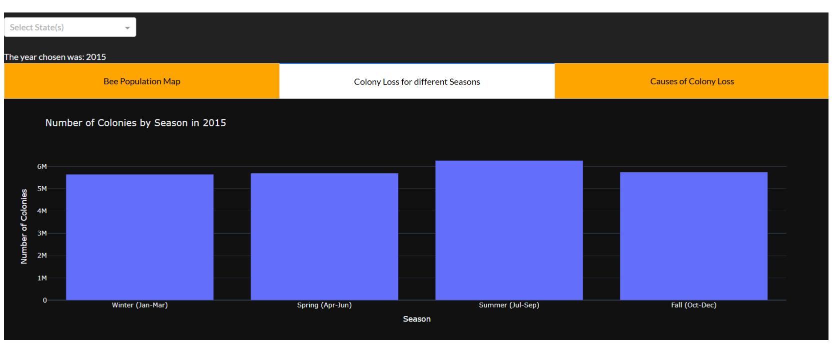
(a) Map of Average Colony Numbers Across All States



(b) Line Chart of Total Colony Loss Across All States

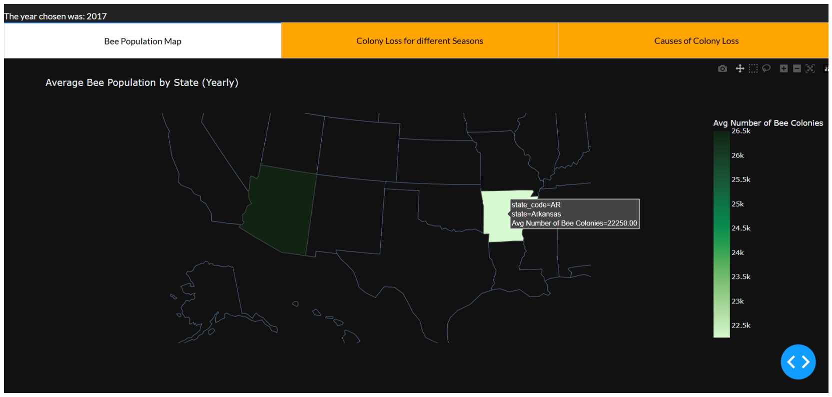


(c) Impact of Different Factors on Colony Loss (All States)

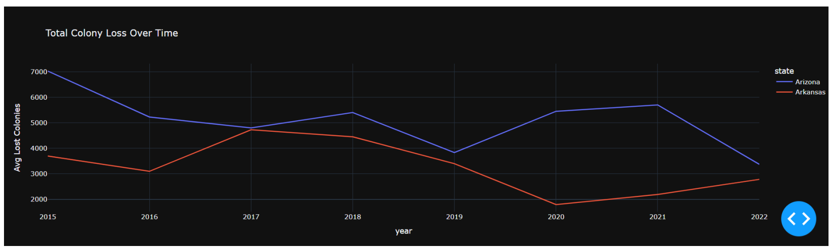


(d) Colony Numbers by Season Across All States

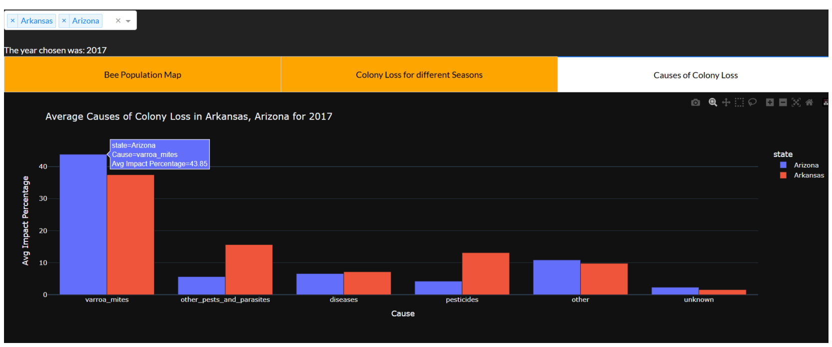
Figure 4: Visualizations for the Use Case Where No State Is Selected



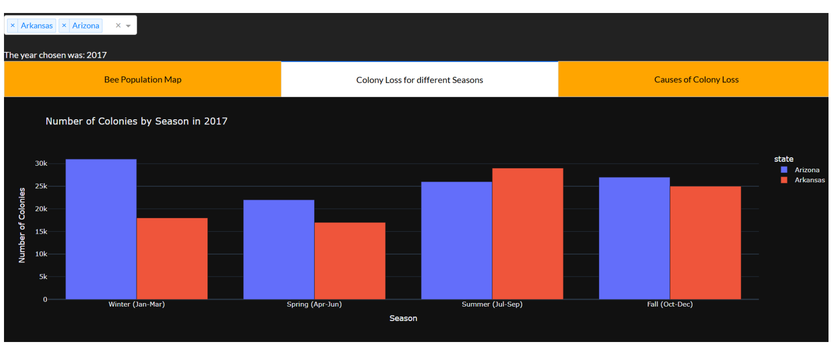
(a) Map of Average Colony Numbers Across Selected States



(b) Line Chart of Total Colony Loss Across Selected States



(c) Impact of Different Factors on Colony Loss (Selected States)



(d) Colony Numbers by Season Across Selected States

Figure 5: Visualizations for the Use Case Where Some States Are Selected

References

[BLB+17] Matthew Brehmer, Bongshin Lee, Benjamin Bach, Nathalie Henry Riche, and Tamara Munzner. Timelines revisited: A design space and considerations for expressive storytelling. *IEEE Transactions on Visualization and Computer Graphics*, 23(9):2151–2164, 2017.

[ZA21] Eduardo E. Zattara and Marcelo A. Aizen. Worldwide occurrence records suggest a global decline in bee species richness. *One Earth*, 4(1):114–123, 2021.