Journal of Geovisualization and Spatial Analysis (2024) 8:36
https://doi.org/10.1007/s41651-024-00197-x

Visualizing Convergent Pressures on Arctic Development

Ana Rivera[1,2] · **Scott R. Stephenson**[2] · **Abbie Tingstad**[3,4]

Accepted: 5 September 2024

Abstract
A rapidly changing Arctic has impacted biophysical and human systems while creating new economic opportunities. Spatially identifying locations with development potential in this changing environment requires characterizing convergences in critical enabling/constraining factors occurring in a particular place. However, mapping techniques based on simple overlays of spatially heterogeneous data may result in visual clutter, compromising legibility, and increasing the likelihood of interpretation errors. To overcome these limitations, we introduce *Pythia*, a tool that combines geographic statistical analysis with a subtractive color model to enable bi- or tri-variate data analysis. Three case studies showcase this visualization tool. Case study 1 identifies locations where temperature and population are projected to increase by 2040. Case study 2 reveals locations with a significant presence of major roads and high NO_2 concentrations but few hospitals and clinics. In case study 3, a combination of transportation infrastructure, protected areas, and travel and tourism infrastructure signals challenges for the future Alaskan tourism industry. Comparing these results allows for further geographic characterization of locations, aiding policymakers in identifying areas lacking resources and infrastructure, exploring possible futures, and supporting long-term strategic planning.

Keywords Spatial statistics · Hotspot · Cluster · CYMK · Multivariate · Arctic

Introduction

Climate change significantly impacts biophysical and human systems, particularly in the Arctic. In the Bering Sea, for example, the frequency of marine heat waves has increased substantially during the last decade (2010–2019). This phenomenon has further decreased sea ice thickness and enhanced warming over Alaska, leading to snow melting and permafrost thawing (Carvalho et al. 2021). Consequently, permafrost thaw is damaging operational infrastructure (e.g., oil and gas pipelines), including roads and highways (Hjort et al. 2022), and exacerbating coastal erosion (Brady & Leichenko 2020). Some Indigenous Alaskan coastal villages are threatened by erosion and flooding, forcing some populations to migrate (Bronen 2010). Rural-to-urban migration has been cited as a primary cause of long-term demographic decline in small Alaskan communities (Martin 2009).

A warming Arctic also creates new economic opportunities for expanded natural resource extraction, particularly oil and natural gas, and increased shipping activities (VanderBerg 2018). The Arctic hosts large oil and natural gas deposits, and its exploration and production are expected to increase as sea ice declines, allowing additional investment in energy extraction activities (Nong et al. 2018). Despite its economic potential, natural resource extraction entails important environmental and health risks (McLoone et al. 2021). NO_2 (an indicator for the larger group of NO*x*) from oil production and gas flaring can aggravate respiratory diseases and increase susceptibility to respiratory infections (US EPA 2023). In Alaska, chronic obstructive pulmonary disease (COPD) is the fourth-highest cause of death and disability, with an increase in deaths per 100,000 of 12.3% from 2009 to 2019 (IHME 2020).

Tourism has played an important role in the economy of Alaska (Zegre et al. 2012) and has been proposed as a pathway for Indigenous sustainable development (Hillmer-Pegram 2016). The State of Alaska 2022–2027 strategic

✉ Ana Rivera
rivera59@msu.edu

[1] Department of Geography, Environment, and Spatial Sciences, Michigan State University, East Lansing, MI, USA

[2] RAND Corporation, Santa Monica, CA, USA

[3] RAND Corporation, Boston, MA, USA

[4] Center for Arctic Study and Policy, U.S. Coast Guard Academy, New London, CT, USA

Published online: 04 October 2024

plan considers the growth and development of the state as a world-class visitor destination "similar to Iceland" (Alaska Department of Commerce 2022). The main attractions in polar destinations are landscapes, ice, and wildlife (Bystrowska & Dawson 2017). Thus, tourism peaks coincide with natural phenomena (e.g., wildlife migration; Grenier 2007). However, efforts to open the Arctic National Wildlife Reserve for oil exploration and extraction (Sherval 2013) could damage natural landscapes and hinder the growth of the tourist industry.

These three themes—climate change, natural resource extraction, and tourism—constitute a small subset of a multitude of economic, political, and environmental pressures shaping Arctic development in complex, often competing ways. Accordingly, "development" itself is a contested concept in the region, implicitly evoking a range of processes spanning from colonialist capital accumulation to preserving traditional Indigenous lifeways. As numerous scholars have noted (e.g., Bennett 2016; Veland & Lynch 2017; Graybill and Petrov 2020), the nature of the intersections of these pressures on development is explicitly spatial. Development always occurs (or fails to occur) some*where*, driven (or inhibited) by a confluence of enabling (or constraining) factors ranging from global to local that converge on a particular location. Identifying places with potential for a specific mode of development thus becomes, partially, an exercise in locating and characterizing convergences in pre-existing factors that tell a story of past and present development through a modal lens.

Development centered on resource extraction in the Arctic, for example, would depend on the presence of raw materials (e.g., hard minerals, oil, and gas) and could be facilitated by infrastructure (e.g., pipelines), a robust transportation network (e.g., seaports, railways, roads), and other amenities (e.g., healthcare facilities, schools/universities) to enable the import/export of inputs/outputs as well as to attract and retain a skilled workforce (Myllylä et al. 2016). Even so, such factors are only incomplete indicators of development potential. Resource extraction may justify the need for novel accessways where few roads exist, and a limited range of amenities in remote communities may underrepresent fluctuations in human presence by a transient workforce.

A comprehensive examination of the development potential of a location typically involves fieldwork-informed case study analysis (e.g., Pashkevich et al. 2016). However, an initial screening of sites based on a small number of critical enabling/constraining factors is possible through mapping and spatial analysis. Despite the recent proliferation of publicly available geospatial data, thematic mapping of development indicators as an analytical technique has tended to focus on a single variable or metric, whether relating to human presence directly (e.g., population change, as in Heleniak 2021) or to second-order effects of human activity (e.g., urban night lights, as in Morshed et al. 2022).

This traditional approach to data visualization is sensible from the perspective of minimizing information that could potentially distract from the core message of the map. It is also insufficient for mapping complex concepts such as "vulnerability" or "development" that have multiple definitions and may be represented through various indicators/proxies, depending on context. The overlay or stack of multiple datasets on a single map may mitigate this limitation. This process, however, may also result in visual clutter, as the excess display of items can cause crowding, thus interfering with the ability of the map user to gather visual information (Rosenholtz et al. 2007). Map legibility can also be compromised, as symbols may overlap, resulting in the occlusion of relevant features (Touya et al. 2015). Choices in symbolization (e.g., size, shape, and color) can also distort the perception of distributional patterns (e.g., apparent low/high concentration of attributes; Dent et al. 2009).

Alternative mapping approaches and tools are thus needed to enable visualization and exploration of data in a systematic yet flexible way while minimizing arbitrary cartographic decisions and assigning exploratory and response variables *a priori* (e.g., Percival et al. 2022).[1] Such mapping approaches should also facilitate the analysis of overlapping "lenses" of development at multiple scales within a consistent visualization scheme to support the study of sustainable development. Datasets relevant to Arctic development, for instance, typically cover a wide range of thematic areas (e.g., climate change, demography, land use rights) and vary greatly in spatial scale, geometry, and format.

For this reason, methods that aim to constrain analysis along networked spaces to avoid detection of spurious clusters (e.g., Yamada and Thill 2010) are, therefore, inappropriate for analysis of these complex and variegated spatial phenomena. Furthermore, due to the sparse spatial distribution of infrastructure and amenities relevant to development in the Arctic, as well as the vast size of geographic administrative units in the region, mapping approaches that highlight where these overlapping factors are relatively highly concentrated are needed. These mapping approaches would allow users to focus attention on those locations where an intersection of complementary (or opposing) factors make development more (or less) likely (Eliasson et al. 2017).

[1] Existing spatial analysis approaches for exploring the overlap of thematic layers include techniques such as the reclassification of attributes to be used in weighted overlay analysis (Rahman et al. 2023). This technique, however, relies on the arbitrary selection of classification schemes, number of classes, and weighting of variables. Geographically Weighted Regression (GWR) and Multi-Scale Geographical GWR are other methods to explore multivariate interactions, but these require *a priori* definition of exploratory and response variables (Li 2022; Lotfata 2022).

This paper introduces a new multivariate visualization tool for geospatial data named *Pythia* that combines a hotspot analysis with a subtractive color model. Like Pythia, the high priestess in the Temple of Apollo at Delphi, for which the tool is named, the tool ingests input datasets and returns an image that experts would then interpret. Hotspot analysis has been used previously for univariate visualization (e.g., measuring rural development in Gupta et al. 2020; urban growth in Morshed et al. 2022; spread of COVID in Bleha & Ďurček 2023) and bivariate visualization (network-constrained phenomena in Tinghua et al. 2017). It has not, however, been used previously with a tri-variate color blending model that allows users to combine multiple datasets to represent a single variable.

The authors developed the tool to facilitate the identification, characterization, and visualization of current and potential places of human activity based on specific variables selected for analysis to illustrate the use of the tool. These variables represent a particular lens, and their visualization allows for the creation of maps indicating development potential (or lack thereof). The maps from case studies presented in this paper are intended to be interpreted primarily by subject matter experts in Arctic development, who may be best positioned to extract insights from an exploratory spatial data analysis while minimizing the risk of misinterpretation. Outputs from this tool may aid policymakers in identifying areas lacking particular resources and infrastructure (e.g., healthcare facilities and programs) that could allow communities to increase their resilience to current and future climate change threats.

Data and Methods

The State of Alaska, the largest state by area in the US, was selected as the study area to showcase the *Pythia* multivariate visualization tool. Publicly accessible geospatial data was acquired to explore three themes of development. The methodology combines geographic statistical analysis with a tri-variate cartographic technique to calculate an aggregate data value within square grid cells and assign a color and saturation scheme for each selected variable.

The *pygeoda* python library, based on GeoDa software, was used to run G^* local spatial statistics to detect local clusters of high values, or hotspots, for each variable analyzed. A subtractive color blending technique was then used to visualize where hotspots overlapped. The combination of these techniques allows for the detection of potential areas for Arctic development and the investigation of nature-society relationships.

Study Area

Alaska (Fig. 1) has experienced negative net migration since 2013, a decline in the working-age population since 2012, and a decrease in births since 2016 (Alaska Department of Labor 2023). According to the US Census Bureau (2022), the state has a population of about 733,000. Of the total population 25 years and older, 30.6% have a bachelor's or higher degree, smaller than the nationwide (35.7%) and North Dakota (31.8%) averages but greater than the average educational attainment of Wyoming (29.6%) (US Census Bureau 2022). In 2021, the women-to-men's earnings ratio was 79.1%, lower than the US (83.1%) and North Dakota (79.5%) averages but greater than the women-to-men's ratio of Wyoming (75.2%) (US Bureau of Labor Statistics 2023).

The US Bureau of Economic Analysis reports that the National Real Gross Domestic Product (GDP) increased by 1.9% in 2022. The real GDP of Alaska and North Dakota, however, decreased by 1.4% and 1.1%, respectively, while the real GDP of Wyoming grew by 1%. In 2022, the mining, quarrying, and oil and gas extraction industries accounted for 17.7, 17.9, and 19.1% of the GDP of Alaska, North Dakota, and Wyoming, respectively. Although GDP from mining, quarrying, and oil and gas extraction increased by 44.8% in the US and by 5% in Wyoming from 2021 to 2022, such industries declined in Alaska (− 12.8%) and North Dakota (− 13.1%). (US Bureau of Economic Analysis 2023).

Data Sources

Open-source geospatial data was acquired to represent indicators/proxies for a set of variables representing three case studies of development to be visualized in *Pythia*. The first case study considers the changing population in Alaska in the context of climate change. The second case study showcases the identification of sites with a significant presence of major roads and high NO_2 concentrations but few health facilities. The third case study explores possible locations for future tourism development.

Case Study 1: Future Climate Change and Changing Populations

The effects of a warming climate are particularly acute in the Arctic region. Rising temperatures have resulted in warmer oceans and amplified ice loss with a deleterious impact on local communities, including increased food insecurity (Loring & Gerlach 2009; Reza & Sabau 2022).

- Projected population: Data identifying locations with projected increasing populations by 2040 in a socioeconomic scenario of high challenges for climate mitigation and adaptation (SSP3) were acquired from the NASA

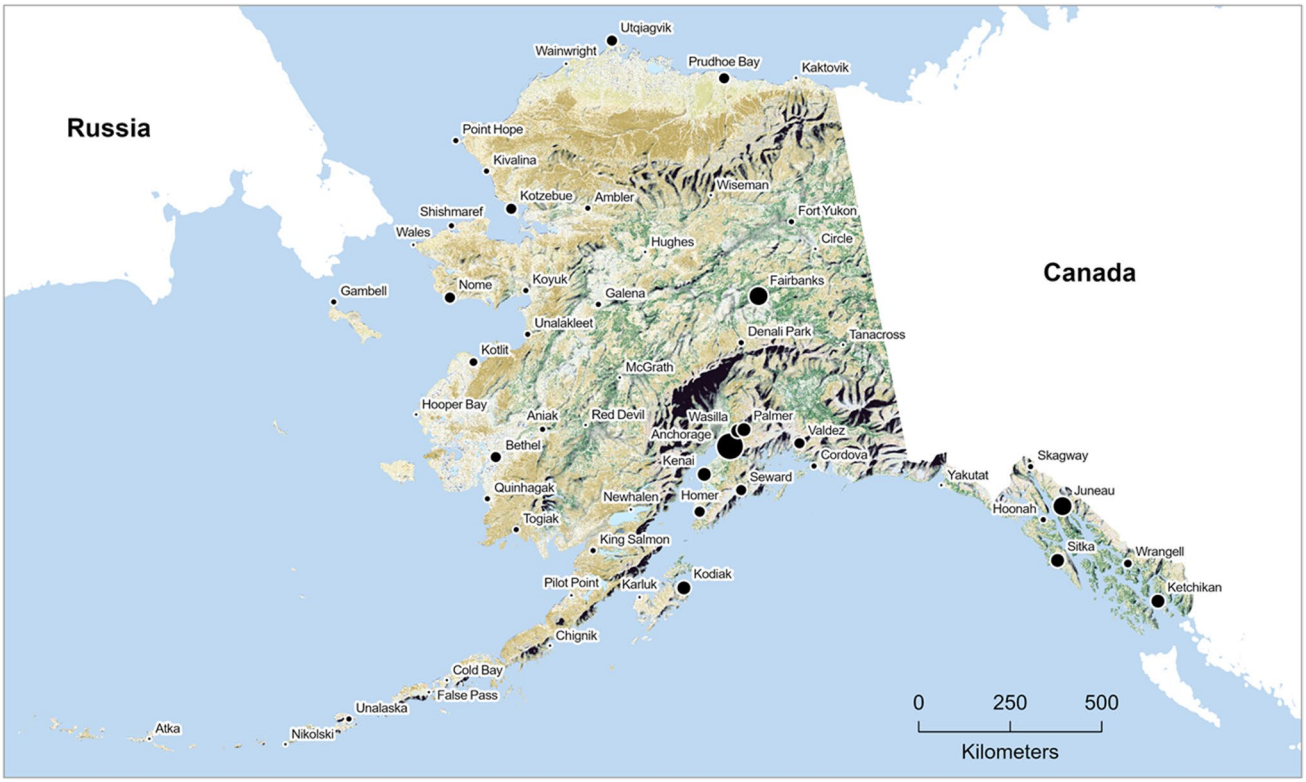

Fig. 1 The State of Alaska and its major populated locations. Data sources: Natural Earth (2022), Nelson (2010), and National Land Cover Dataset 2016 (Dewitz 2019)

Socioeconomic Data and Applications Center (sedac.ciesin.columbia.edu). The projection is relative to the base year 2000 and is available at one-eighth-degree spatial resolution (Jones and O'Neill 2020).

- Climate change: The projected annual mean temperature anomaly for 2040 under the SSP3 scenario was selected as an indicator of climate change. Gridded values, calculated relative to a 1995–2014 reference period, were acquired at 0.5° spatial resolution from the Climate Change Knowledge Portal (climateknowledgeportal.worldbank.org) developed by the World Bank (2021).

Case Study 2: Environmental Health

The accumulation of chemical stressors in the environment can affect communities that may be already overburdened by socioeconomic inequalities and a lack of resources, including access to healthcare. The identification of hotspots of NO_2 concentrations from activity such as road traffic, along with access to healthcare facilities to treat conditions such as asthma that may be exacerbated by air pollution, is therefore paramount for addressing health disparities, which are also projected to be exacerbated by climate change.

- Major roads: Locations of major roads were obtained from Natural Earth (2022) from the website naturalearthdata.com.
- Nitrogen dioxide (NO_2): Gridded annual mean NO_2 concentrations from Copernicus Sentinel-5P (2018) from January 01, 2021, to December 31, 2021, with 1-km spatial resolution, were acquired via Google Earth Engine.
- Hospitals and clinics: Locations of hospitals and clinics were obtained from OpenStreetMap (OSM) via the Geofabrik Download Portal (download.geofabrik.de). Hospitals and clinics were defined as features classified as "hospital" or "clinic" (OpenStreetMap & Geofabrik 2023).

Case Study 3: Tourism Potential

According to the Resource Development Council for Alaska (2023), in 2018, tourism in the state generated more than $126 million in state revenues and $88 million in municipal revenues. The same source reports that most visitors arrive by cruise ship (58%) and the rest by highway/ferry or air to spend money on tours, public land permits, campgrounds, and hotels.

- Transportation infrastructure: Major seaport, airport, road, and railroad locations were obtained from Natural Earth (2022) at naturalearthdata.com. Railroad data were last updated in 2017; all other datasets were updated in 2022. The airport dataset was combined with data acquired from the State of Alaska Geoportal (gis.data.alaska.gov) to include 237 facilities owned and operated by the State of Alaska Department of Transportation and Public Facilities (Alaska DOT&PF 2024).
- Hospitality and attractions: Hospitality facilities and tourist attractions were obtained from OpenStreetMap via the Geofabrik Download Portal (download.geofabrik.de). Hospitality facilities were defined as features classified as "guesthouse," "hostel," "hotel," or "motel." Tourist attractions were defined as features classified as "attraction," "campsite," "museum," "travel agent," or "tourist information." (OpenStreetMap & Geofabrik 2023).
- Protected areas: Terrestrial, coastal, and marine protected areas datasets were obtained from the World Database on Protected Areas (WDPA) portal (protectedplanet.net) for June 2022. The dataset is collected and made available by the UN Environment Program and the International Union for Conservation of Nature (IUCN & UNEP-WCMC 2022).

Data Analysis

Arctic development potential can be characterized partially by the presence (or absence) of enabling (or constraining) variables. *Pythia* allows one or more datasets to represent one variable characterizing an enabling/constraining factor for a specific path of development. Each variable, in turn, is represented by a circle in a Venn diagram with overlapping colors representing areas where two or more variables with statistically significant local clusters of high values overlap (hotspots). Optionally, these colors can be displayed with varying shades of saturation, representing the degree to which multiple datasets representing a single variable achieve statistical significance.

The identification and characterization of spatial hotspots for each variable requires a three-step process (Fig. 2): (a) the aggregation or synthesis of geographic data by a unit of analysis (e.g., cell), (b) the detection of local clusters, and (c) the visualization of detected clusters.

Data Aggregation

Geospatial data representing Arctic development factors come from multiple, unrelated sources and generally lack a standard format. A data aggregation method that could accommodate a variety of formats (vector vs. raster) and

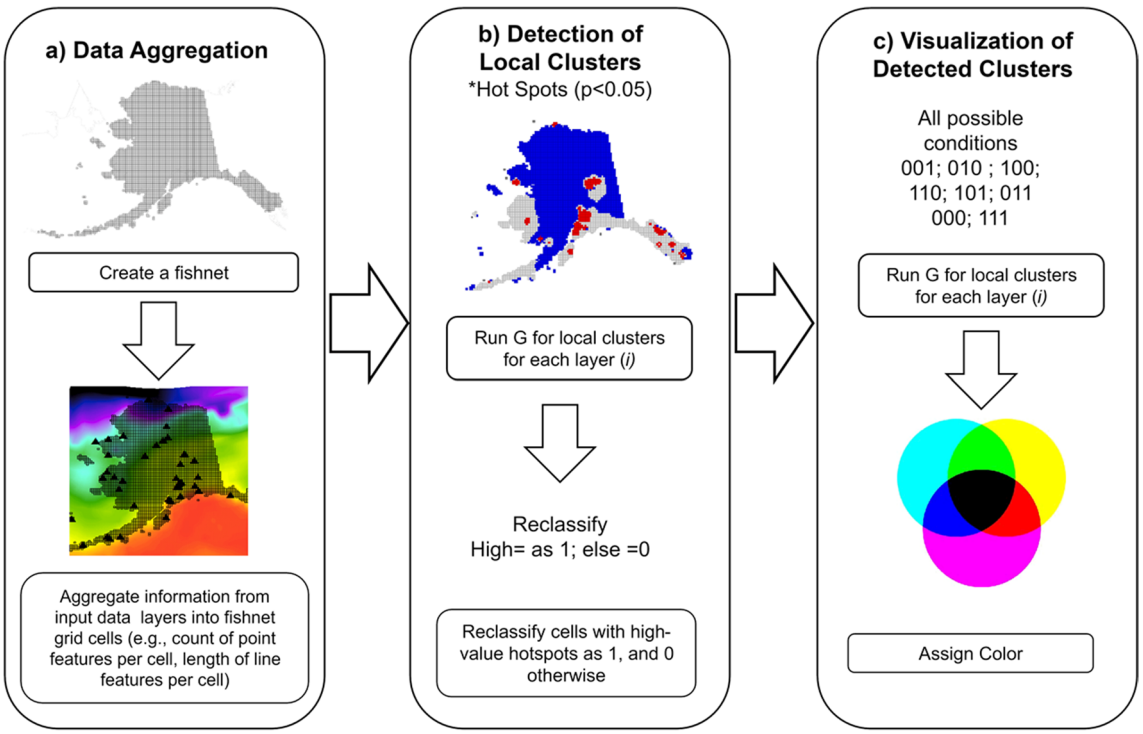

Fig. 2 An overview of the methodology to identify and characterize locations for a specific type of development using the *Pythia* multivariate visualization tool

geometries (point, line, polygon) and enable hotspot analysis on a standard geographic unit of measurement is, therefore, required. To this end, a "fishnet" grid, with rectangular cells of 20 × 20 km, was created to cover the extent of the study area. The cell size selection for this analysis was a compromise between resolution and over-aggregation time. While 20-km resolution limits the utility of the tool at local scales, it is sufficient to enable the exploration of broad-state and region-scale spatial patterns.

The grid was first generated at a spatial extent exceeding that of the study area to avoid the loss of data at the boundaries. Grid cells that at least partially intersected with the study area were selected via the spatial selection tool in ArcGIS Pro© (version 3.0.3). The selected cells were then retained as the complete fishnet grid for use in hotspot analysis. A manual inspection was performed to confirm that no portion of the study area was omitted. Geographic data were aggregated by cell according to the type (point, line, polygon, raster), as follows:

- Point data (e.g., seaports; hospitals): Features were aggregated with the use of the spatial join tool available in ArcGIS. Each cell was assigned a count value corresponding to the number of points falling inside it. Cells with no point features were assigned values of 0 for that variable.
- Line and polygon data (e.g., roads; marine protected areas): Features were split by portions overlapping the fishnet cells using an intersect tool in ArcGIS. The total length (for line data) or area (for polygon data) of the features intersecting each cell was computed and assigned as the value of the cell. Cells with no overlapping line segments or polygon fragments were assigned values of 0 for that variable.
- Raster data (e.g., temperature; nitrogen dioxide): A zonal statistics tool was used to summarize (e.g., mean, median, sum) the values of raster cells whose centroids fall in each fishnet grid cell. Summary statistics were then assigned to each fishnet cell using a table join. Where input raster data contained cells with "no data," no data value was assigned to the overlapping fishnet grid cells, and these areas were masked from hotspot analysis.

Detection of Local Clusters

The identification of hotspots or clusters was performed via the estimation of the G^* statistic. Other statistics allow for the detection of high values within the data. The G^* statistic, however, allows for the detection of statistically significant clusters only if a feature has other surrounding high values of a feature. A Kernel Density Estimation, for example, does not calculate the p-value. It is then unclear if the detected clusters are statistically significant or not (Kalinic & Krisp 2018). We, therefore, incorporate the G^* statistic into the *Pythia* tool.

Local clusters were detected using *pygeoda*, a Python library for spatial analysis based on GeoDa software. The tool allows for the application of local spatial autocorrelation statistics, specifically G or G^*, to identify statistically significant clustering of cases or "hotspots" based on calculated distances between each feature and the location of its nearest neighbor (Ord & Getis 1995). These distances are stored on a matrix and serve as spatial weights, integrating spatial relationships among cells in the dataset. These spatial relationships can be based on Rook contiguity, where only common sides of cells are considered to define neighboring relationships, or on Queen contiguity, where common vertices are additionally considered to define neighboring relationships (Anselin & Rey 2014).

The *Pythia* tool allows for the specification of polygon contiguity and order, where first-order contiguity analyzes the immediate neighbors, and second-order contiguity considers spatial relationships beyond immediate neighbors (e.g., neighbors of neighbors). For this exploratory analysis, a first-order Queen contiguity was selected. Based on the generated spatial weights, a local spatial autocorrelation statistic G is calculated for each column or dataset (i). The statistic measures the degree of association that results from the concentration of features, in this case, aggregated in an area and all other areas within a radius distance (neighborhood). The G statistic is then standardized, producing a Z_G score, with the formula defined as (Getis & Ord 1992)

$$z_G = \frac{(G - E[G])}{\sqrt{(V[G])}} \tag{1}$$

where G is the calculated G statistic describing the spatial dependency or autocorrelation of a feature, E is the expectation of G, and V is the variance of G. The Z_G score is evaluated under the null hypothesis (H_0) of complete spatial randomness (CSR) and the alternative hypothesis (H_A) of spatial dependency. Under CSR, the expected number of observations within any cell would be the same.

A p-value is also calculated and used to assess statistical significance at a 95% confidence level (p-value < 0.05). The 95% confidence level is a somewhat arbitrary standard for statistical significance; in this case, a p-value less than 0.05 indicates that the observed spatial pattern is unlikely to have occurred by random chance 95% of the time. While stricter standards (e.g., p-value < 0.01) may be applied to further narrow the set of locations identified as local clusters of geographic features (hotspots), we adopt the 95% standard owing to

its widespread use in statistical analysis and its general acceptance as a reasonable cut-off for statistical significance (Andrade 2019). Reducing the probability of Type I errors (rejecting the H_0 when it is true) commonly involves setting a lower significance level (e.g., 0.01). This significance level may, however, increase the risk of Type II errors (failing to reject H_0 when it is false). The opposite outcome may occur with a higher significance level (e.g., 0.1). Thus, a significance level of 0.05 may also balance both.

High positive z-values may indicate possible local clusters of high values (hotspots) of the variable being analyzed, whereas low negative z-values may indicate potential clusters of low values (cold spots). The resulting clusters are identified by an integer (Fig. 3), where 0 indicates non-significant clusters ($-1.96 < Z_G < 1.96$), 1 indicates significant hotspots (cluster of high values or $Z_G > 1.96$), and 2 indicates significant cold spots (cluster of low values or $Z_G < -1.96$). Statistically significant (p-value < 0.05) high-value clusters are extracted using a binary reclassification. A value of 0 is assigned to statistically significant low values (cold spots) to remove them, along with non-significant clusters from further analysis.

This G calculation and binary reclassification is performed for each aggregated column or dataset (i) within the fishnet. These binary-recoded values are then concatenated into a single column, resulting in a string comprised of 0 s and 1 s. The number of characters thus depended on the number of datasets analyzed. For the analysis of three datasets, one representing a constraining or enabling variable, a cell with no significant detected hotspots would have a concatenated value of "0,0,0." A cell with a string of "1,0,0" indicates significant hotspots detected for the first dataset, whereas "1,0,1" indicates significant hotspots were found for the first and third datasets, and "1,1,1" indicates significant hotspots were found for all three datasets.

Visualization of Detected Clusters

A multivariate CYMK color method is employed to visualize multivariate cluster detection. This process, therefore, translates qualitative and quantitative differences between datasets into distinct visual effects. Variations among datasets are represented by a changing hue and in the level of significance by saturation. This approach allows map users to easily distinguish among two or three datasets visualized simultaneously while also enabling visualization of varying levels of statistical significance in clusters aggregated from individual datasets.

The US Census Bureau used a color composite approach in the 1970s to analyze bivariate correlation in choropleth maps (Olson 1981; Strode et al. 2020). Variables visualized with such a method can be continuous (interval/ratio) or categorical (ordinal/binary; Carstensen 1984). The creation of tri-variate maps was discussed in the 1980s, which addressed the possibility of combining saturation and brightness using the Ostwald Triangle (Trumbo 1981). Color blending methods (e.g., RGB or CMYK) are now often used in spectral image analysis and geochemical mapping, highlighting the spatial relationship between variables (Agyeman et al. 2023).

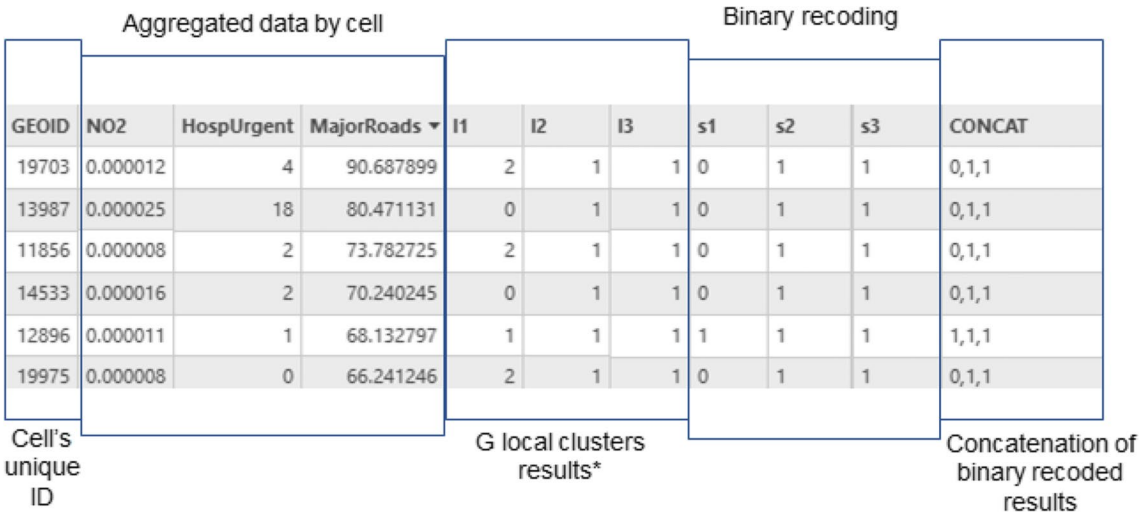

Fig. 3 Snapshot of the attribute table of a fishnet computed by *Pythia* for three datasets analyzed, each representing one variable. The table contains the unique ID of the cell (GEOID), aggregated data (i), cluster results* (I1–I3), binary reclassification (s1–s3), and concatenation (CONCAT) of binary results. *0: non-significant clusters; 1: statistically significant hotspots (cluster of high values); 2: statistically significant cold spots (cluster of low values)

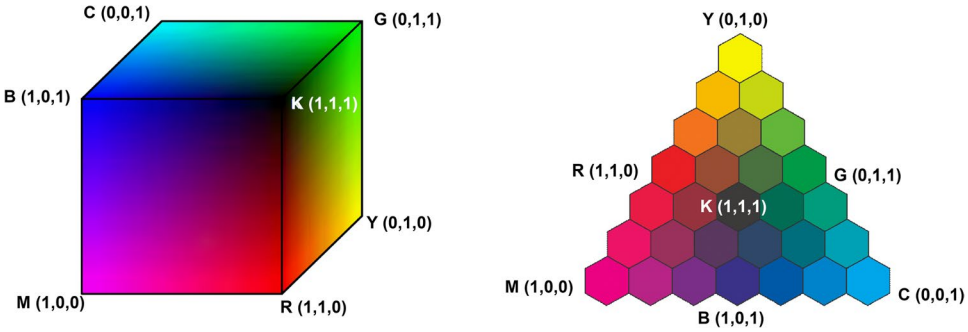

Fig. 4 Examples of CYMK color schemes represented in a cube or equilateral triangle

In a cyan, magenta, yellow, and black (CMYK) color scheme, each color represents an individual variable and, when combined, results in an array of colors (Albanese et al. 2011). The generated colors (Fig. 4) can be presented as a cube, where the origin vertex has a black color (K), with increasing saturation and changing hue along the axes (Albanese et al. 2011; Guagliardi et al., 2020). CYMK-generated colors can also be presented as an equilateral triangle, with each vertex representing a primary color and the center gradually turning black (K), indicating a full combination of colors (Kebonye et al. 2023).

The *Pythia* tool uses cyan, magenta, and yellow as the primary colors, each defining one variable for which statistically significant hotspots were detected. The variables are represented in a Venn diagram (Fig. 5), with each circle shaded with a primary CMY color. An RGB color scheme results if a hotspot is detected for two of the analyzed variables. A hotspot detected in all three variables results in the full combination of colors (K).

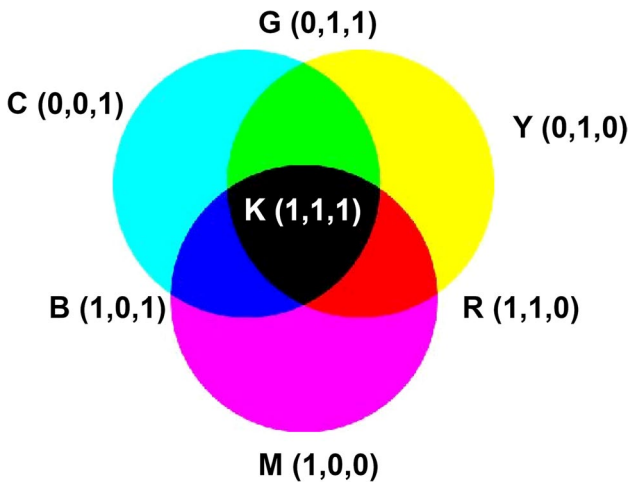

Fig. 5 CYMK color scheme used by *Pythia* to visualize places where hotspots overlap

Using More than One Dataset per Variable

Optionally, the *Pythia* tool allows for the incorporation of more than one dataset (i) per analyzed variable (x) and for hotspot analysis to be performed on each dataset individually, with the combined result representing the presence of a significant hotspot of both, either, or neither dataset. For such analyses, the *G* calculation and binary reclassification are performed for each column or dataset (i) within the fishnet, as described above (see "Detection of Local Clusters"). For two datasets (i) representing a single variable (x), results of the binary reclassification are added together before proceeding with the concatenation (Fig. 6), resulting in values ranging from 0 (no hotspots detected in either dataset) to 2 (hotspots detected in both datasets).

The resulting values are then concatenated. As before, the number of characters in the concatenated series depends on the number of variables (x) analyzed, but each character now represents the sum of the recoded results. For three variables and two datasets representing each, a cell with no significant detected hotspots in any dataset still results in a concatenated value of "0,0,0." A cell with a string of "1,0,0" indicates significant hotspots detected on one dataset of one variable only. A series of "2,0,1" shows significant hotspots detected for the two analyzed datasets in the first variable and only in one dataset of the third variable. A string of "2,2,2" indicates significant hotspots detected for both datasets in all three variables.

Each circle or primary color in the Venn diagram still represents a variable for which statistically significant hotspots are detected. The level of saturation of each color, however, depends on the number of datasets found to be significant (the addition of the binary-recoded results). For example, a cell with a string of "1,0,0" may be visualized in magenta with only 50% saturation (Fig. 7) because significance was found for only one of the two datasets. In contrast, a string of "2,0,0" presents as magenta with 100% saturation. A cell with a series of "2,0,1" or "1,0,2" may be visualized as blue with 75% saturation, given that significance was found in three of the four datasets, while a string of "1,0,1" results in blue with 50% saturation. A string of "2,2,2" indicates full saturation of all three colors, visualized as black.

Aggregated data by cell

Binary recoding

Concatenation of added recoded results

	x 1		x 2		x 3		x 1		x 2		x 3		x 1		x 2		x 3		x 1	x 2	x 3	
	i1	i2	i3	i4	i5	i6	i1	i2	i3	i4	i5	i6	i1	i2	i3	i4	i5	i6	s1+s2	s3+s4	s5+s6	
GEOID	Ports_Air	Rail_Road	Attraction	Lodging	WPDA_SQKM	WPDAM_SQKM	l1	l2	l3	l4	l5	l6	s1	s2	s3	s4	s5	s6	s_l1	s_l2	s_l3	CONCAT
13715	0	22.294363	7	0	0.13211	0.159215	1	1	1	1	1	0	1	1	1	1	1	0	2	2	1	2,2,1
13986	0	0	0	0	0.180738	0.427962	1	1	1	1	0	1	1	1	1	1	0	1	2	2	1	2,2,1
13716	0	54.970895	6	5	0.617865	0	1	1	1	1	1	0	1	1	1	1	1	0	2	2	1	2,2,1
13212	1	22.995602	43	4	0.000034	0	1	1	1	1	2	0	1	1	1	1	0	0	2	2	0	2,2,0
13211	0	0	0	0	0.005926	0	1	1	1	1	2	0	1	1	1	1	0	0	2	2	0	2,2,0

Cell's unique ID

G local clusters results

Add binary recoding by topic

Fig. 6 Snapshot of attribute table of a fishnet computed by *Pythia* for the analysis of three variables, each with two datasets. The resulting table contains the unique ID of the cell (GEOID), aggregated data by cell, cluster results (l1–l6), binary reclassification (s1–s6), the addition of binary recoded by variable (s1 + s2, s3 + s4, and s5 + s6), and concatenation results (CONCAT)

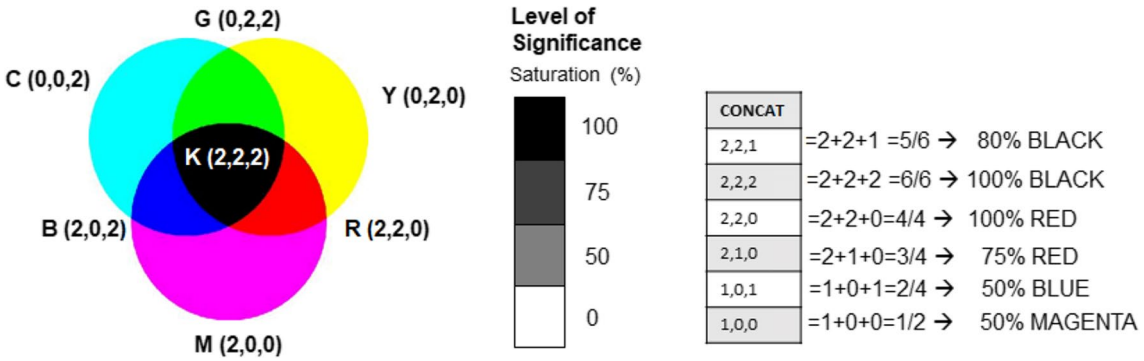

Fig. 7 The CYMK color and saturation scheme used by *Pythia* to visualize places where variables overlapped, considering two datasets to analyze each variable

Results

Here, we describe the results of *Pythia* as a supporting tool facilitating analysis of the three case studies: (1) future climate change and changing populations, (2) environmental health, and (3) tourism potential in Alaska. In each case, *Pythia* revealed statistically significant clusters or hotspots of key features and facilitated visual interpretation of convergent factors of development.

Case Study 1: Future Climate Change and Changing Populations

Mapping projected changes in population and increasing temperature data with traditional overlays illustrates the challenge of interpreting a coherent narrative from multiple overlapping raster datasets. The spatial distribution of areas with projected increasing populations by 2040 (Fig. 8) is shown with the use of a geometric interval scheme, given that the distribution of the data is heavily skewed to the right by a predominance of pixel values equal to 0. Areas with projected increases appeared primarily concentrated in Fairbanks, Anchorage, and to the east around Bethel and Nome.

Projected mean temperature anomalies appear to be more severe in the North Slope region compared to the rest of the state. The temperature values, however, are visualized with an equal interval classification scheme, resulting in some classes with more pixels than others. A different classification method (e.g., standard deviation) with a different number of classes would produce a different pattern, potentially obscuring hotspots of high-mean temperature anomalies or misidentifying some areas as hotspots. This traditional visualization method also raises potential questions about the meaning of "significant" increases in population and temperature anomalies, which could be left to the personal interpretation of the map user. An overlay of both datasets also complicates the detection of hotspots.

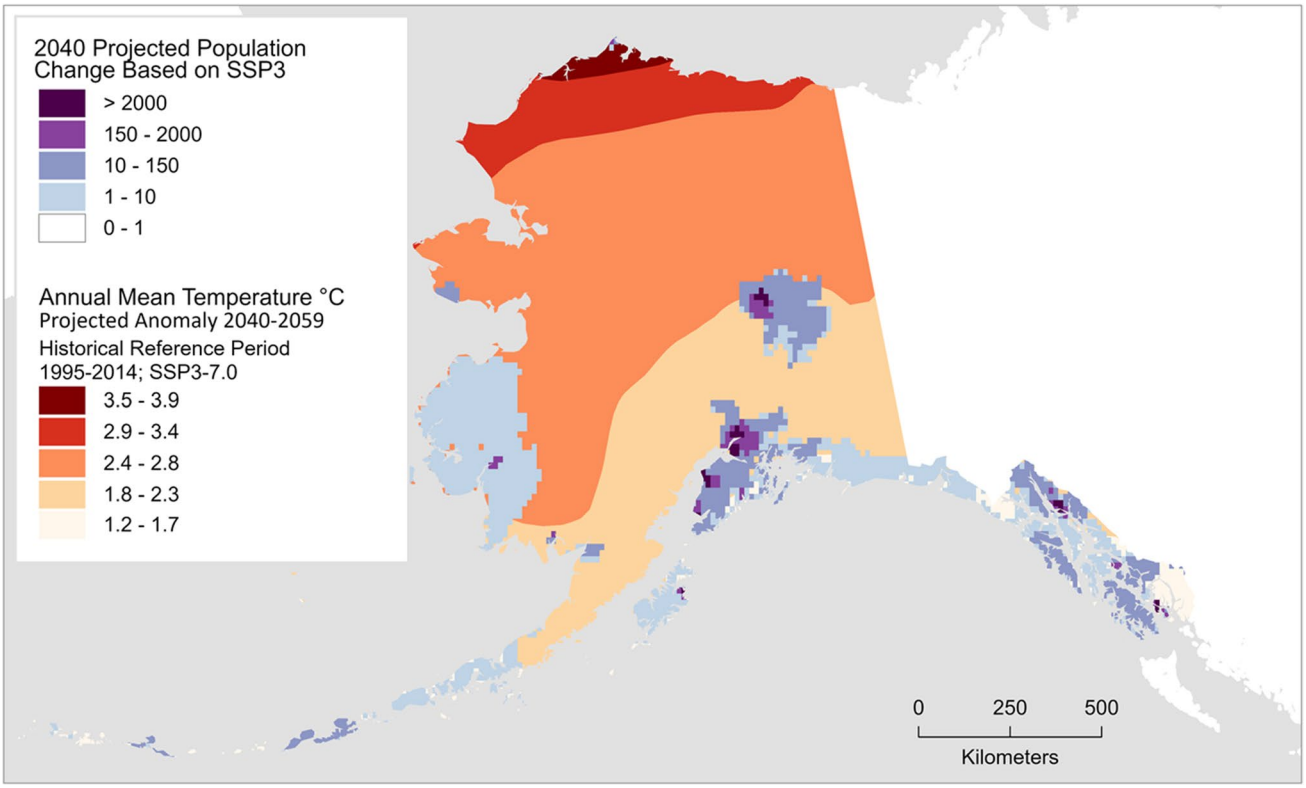

Fig. 8 Spatial distribution of projected increasing populations and annual mean temperature anomaly by 2040 under an SSP3 scenario. Data source: NASA Socioeconomic Data and Applications Center (Jones and O'Neill 2020) and World Bank (2021)

Visualization of these data with *Pythia* simplifies the interpretation of these phenomena. *Pythia* aggregates data into uniform polygonal units of analysis (cells), highlighting areas of projected increasing population. In addition, because the tool detects statistically significant hotspots based on neighboring values and distance between cells, statistical relationships determine the "threshold" for a high-value location, reducing potential human errors of interpretation.

Figure 9 clearly depicts locations with statistically significant high values of projected increasing mean temperature anomalies in northern Alaska (magenta). Areas with projected increasing population by 2040 were found along the eastern coast, in Fairbanks and southern Alaska (cyan). Places identified as projected to experience both increased temperature anomalies and increasing populations by 2040 were found only in Nome (blue).

Case Study 2: Environmental Health

The spatial distribution of major roads (Fig. 10) shows a high density in Fairbanks and southern Alaska, near Anchorage and Prudhoe Bay. Hospitals and clinics extracted from OSM are mainly concentrated in Anchorage and south towards Kenai. The apparent high concentration of point symbology in these areas can produce spatial bias. The resulting spatial pattern could lead a map user to focus an inordinate amount of attention on these areas of apparent high concentration of features.

Concentrations of NO_2 appeared higher in the North Slope and near Denali. The apparent location of places with the highest NO_2 concentrations, however, depends on the classification method used to distinguish the raster values (e.g., quantile, standard deviation) and the number of classes. The detection of clusters sensitive to the distribution of spatially heterogenous datasets thus requires a standard spatial unit (e.g., cell) and a spatial statistical test, which have been integrated into *Pythia*.

Using *Pythia*, hotspots of hospitals and clinics (magenta) were observed on the western coast, in the Kenai Peninsula towards Valdez and Cordova, and on the southeastern coast in Sitka. Additional hospital and clinic hotspots were detected near Utqiagvik, Kotzebue, and Koyukuk, each of which appeared as single facilities using simple overlays (Fig. 10). Concentrations of major roads (cyan) were mainly observed from Prudhoe Bay towards Fairbanks, near Nome, in southern Alaska near Valdez, and along the southeastern coast. Statistically significant locations of high NO_2 concentrations (yellow) were found in northern Alaska, in the interior surrounding Denali towards the Kenai Peninsula, along the Aleutian Islands, and on the southeastern coast.

Detection of statistically significant hotspots with *Pythia* allows for the identification of overlapping hotspots that

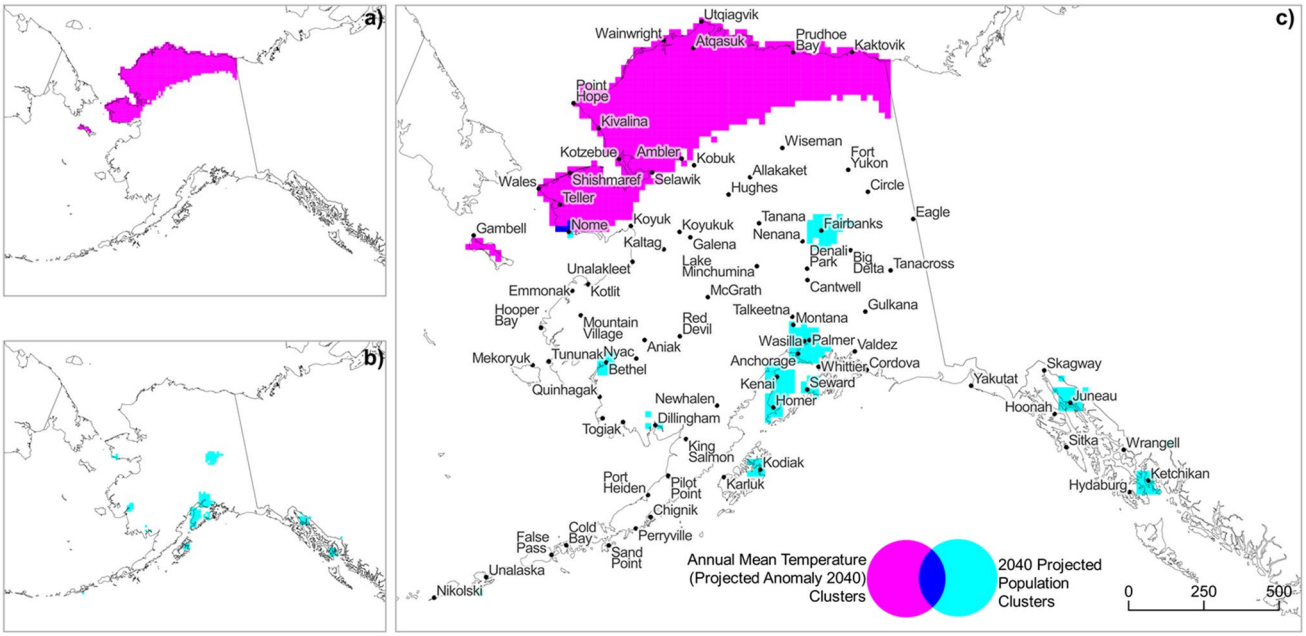

Fig. 9 Locations identified in *Pythia* as experiencing statistically significant **a** increased mean temperature anomalies (magenta), **b** increasing population (cyan), and **c** both (blue) by 2040 for an SSP3 scenario

otherwise may have gone undetected (Fig. 11). In this case, the high incidence of hospitals and clinics on the southeast coast may have been obscured by the overlapping point symbology. Likewise, the identification of statistically high NO_2 concentrations in the North Slope no longer depends on cartographic choices such as the classification method or

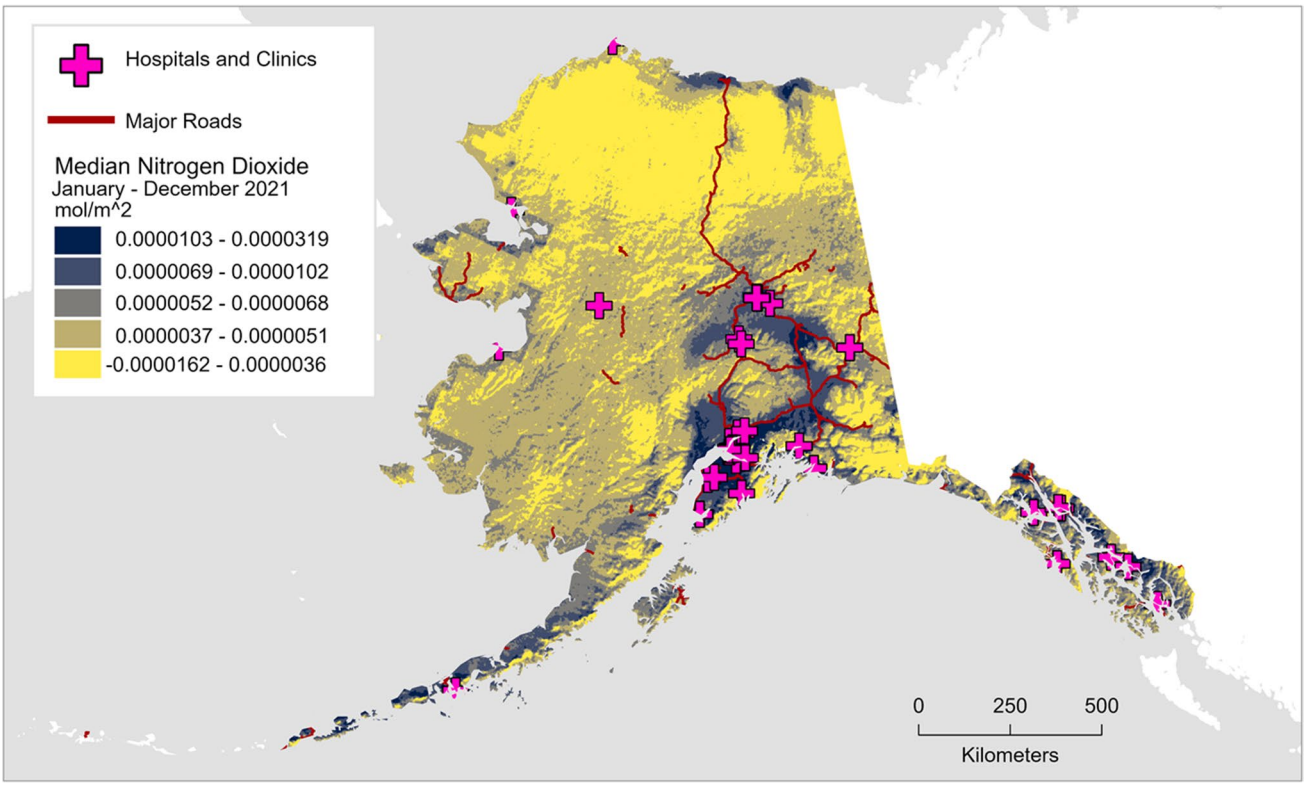

Fig. 10 Spatial distribution of hospitals and clinics, major roads, and median nitrogen dioxide concentrations for January to December 2021 in Alaska. Data source: OpenStreetMap & Geofabrik (2023); Natural Earth (2022); Copernicus Sentinel-5P (2018)

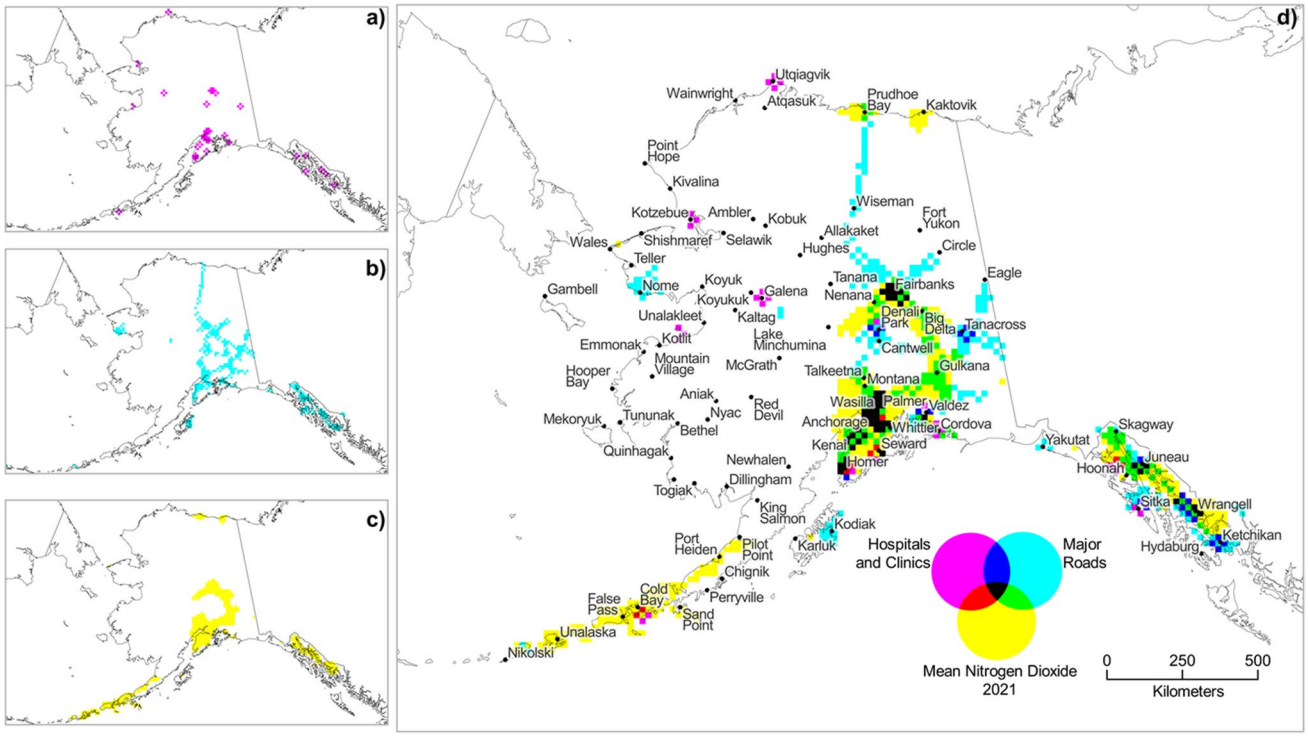

Fig. 11 Locations identified in *Pythia* as hotspots of **a** hospitals and clinics (magenta), **b** major roads (cyan), **c** mean concentration of nitrogen dioxide (yellow), and **d** overlapping combinations of the three variables

the number of classes. In the Anchorage area, the separation between overlapping clusters of high NO_2 values and hospitals and clinics (red) and overlapping clusters of high NO_2 values and major roads (green) is clearly visible, whereas these features may appear to fully intersect when visualized with simple overlays (Fig. 10). The overlap of major roads, hospitals, and clinics occurs in Tanacross, which stands out conspicuously as the only location visualized in blue.

Case Study 3: Tourism

Initial visualization of spatial data to analyze the tourism industry in Alaska resulted in considerable symbology overlap (Fig. 12). Major seaports are shown in Nome, Kodiak, and to the southeast from Skagway to Ketchikan. Major airports are located to the north in Utqiagvik and Prudhoe Bay, to the east in Nome and Bethel, in the interior near Fairbanks, and to the south near Kodiak. No specific spatial pattern, however, is observed when all public DOT&PF-controlled airports ($n = 237$) are displayed.

Major transportation lines (railroads and roads) run north–south from Prudhoe Bay to Anchorage, with hospitality facilities (e.g., hotels, motels, and guesthouses) located along the same route. A similar pattern appeared for tourist attractions (e.g., campsites, museums, travel agents, tourist information) but with other scattered locations throughout the state. Terrestrial protected areas appeared throughout Alaska, while coastal and marine protected areas are shown west in Bethel and southeast in Glacier Bay.

While these patterns are worthy of investigation, their presentation as simple overlays compromises the identification of hotspots of each variable. This phenomenon is particularly apparent where point datasets of hospitality facilities and tourist attractions are stacked over each other. These features, as well as railroads and roads represented as lines, cannot overlap in reality, as depicted on the map, because they cannot occupy the same physical space. Detection of hotspots of protected areas is also challenging due to the fragmentation of multiple polygon features of varying shapes and sizes that comprise these datasets. Hotspot detection, therefore, requires the construction of a uniform "neighborhood" unit of analysis, allowing for spatial aggregation and summarization of multiple features with different spatial dimensions.

The results from *Pythia* (Fig. 13) show a significant hotspot (high saturation) of terrestrial, coastal, and marine protected areas around the Bethel region and southeast in Glacier Bay (yellow). Most travel and tourism infrastructure hotspots (cyan) were found along the corridor from Anchorage, Valdez, and Cordova, with other hotspots detected to the North in Utqiagvik and along the southeastern part of Alaska. Transportation clusters (magenta) were found to the

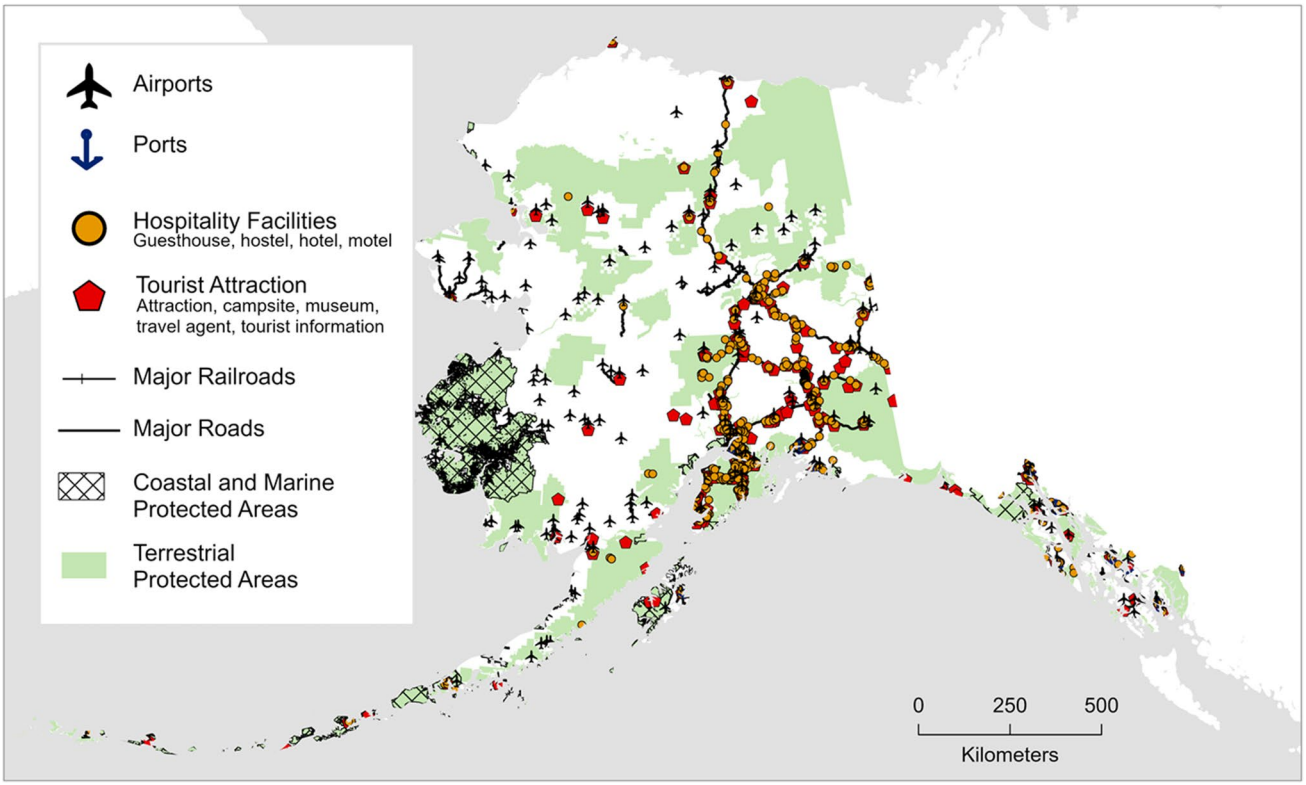

Fig. 12 Spatial distribution of major transportation infrastructure (airports, seaports, railroads, and roads) relevant for the tourism industry, hospitality and tourist attractions, and terrestrial, coastal, and marine protected areas. Data source: Alaska DOT&PF (2024); Natural Earth (2022); OpenStreetMap & Geofabrik (2023); IUCN & UNEP-WCMC (2022)

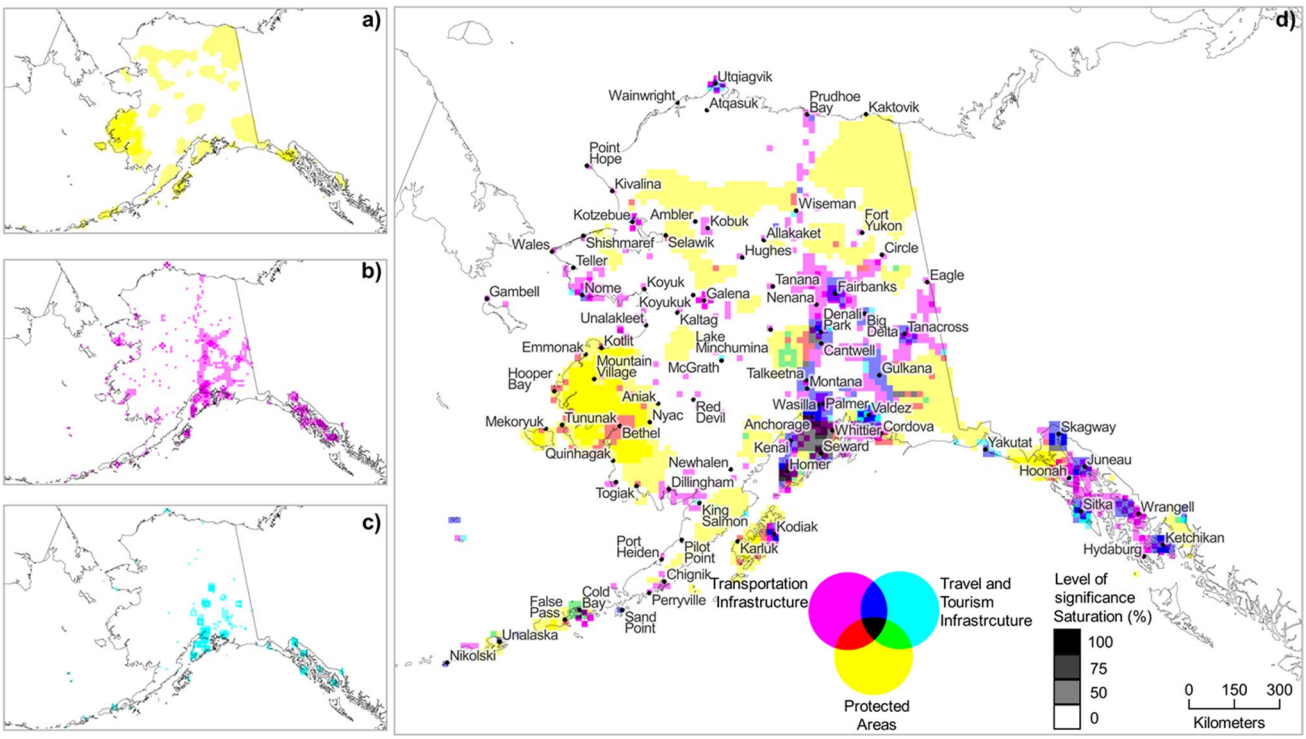

Fig. 13 Locations identified in *Pythia* as hotspots of **a** terrestrial, coastal, and marine protected areas (yellow); **b** transportation infrastructure (magenta); **c** travel and tourism infrastructure (cyan); and **d** overlapping combinations of the three variables

west in Nome, from Fairbanks north towards Prudhoe Bay, and along the eastern coast.

While Fig. 12 may appear to show a potential travel and infrastructure cluster near Kobuk Valley National Park, results from *Pythia* reveal that this potential hotspot is not statistically significant due to the consideration of distance between these features in the calculation of the G^* statistic. *Pythia* also allows the removal of statistically non-significant clusters, further aiding map interpretation. Such apparent clusters of terrestrial protected areas were removed along the Brooks Range and Cape Krusenstern National Monument. Several non-significant clusters of transportation infrastructure (e.g., Utqiagvik) were also removed.

As shown in this example, the aggregation of data into uniform spatial units and the removal of not statistically significant hotspots reduces visual clutter, aiding the identification of locations with tourism potential. Overlapping hotspots of transportation infrastructure and protected areas (red) were detected in Bethel and Cordova. Overlapping hotspots of protected areas and travel and tourism infrastructure (green) were detected in Cold Bay. Overlap of hotspots of transportation and travel and tourism infrastructure (blue) was primarily located in Fairbanks, Wasilla, and Palmer. Additional locations were found to the southeast in Skagway, Juneau, and Ketchikan. All variables overlapped (black) only in the Kenai Peninsula but with less than 100% color saturation.

Testing Sensitivity to Cell Size

Cell size determines the scale at which data are aggregated. The choice of cell size, therefore, can significantly affect the results of spatial clustering (Vieux & Needham 1993; Hengl 2006). A sensitivity analysis was performed to test the effect of cell size on results from *Pythia*. Data inputs from case study 2 (hospitals and clinics; major roads; NO_2 concentrations) were aggregated into 10-km, 30-km, and 40-km cells, representing a decrease and increase, respectively, in cell size over the baseline case (20-km). Case study 2 was selected because its three inputs represent three different types of spatial data (point, line, and raster), each of which might be affected by a change in cell size in a different way.

Results of the sensitivity analysis (Fig. 14) show similar spatial patterns among the clusters represented with finer to coarser (10-km, 30-km, 40-km) cell sizes compared to the 20-km baseline map (Fig. 14b). In general, clusters of the individual datasets (and their areas of overlap) that appear in the 20-km map are also present in the 10-km, 30-km, and 40-km maps but with varying levels of visibility. Individual clusters become easier to see as cell size increases. Thus, clusters may appear "larger" on the 40-km map than on the 10-km map, even when the total number of clusters is approximately equal. Conversely, a smaller cell size enables clusters to be depicted with greater spatial precision but may appear "smaller" than on the 40-km map.

An important exception to the overall similarity among these maps is that road clusters in the coarser grid are limited to urban areas and do not depict inter-urban connections, such as the highway from Fairbanks to Prudhoe Bay. This finding suggests that relatively small concentrations of spatial features are less likely to be classified as clusters as the level of data aggregation increases. In contrast, the finer grid represents clusters of inter-urban road connections with a higher level of detail than the baseline grid. A possible downside of this enhanced detail is that clusters of features, which may have high concentrations in fewer locations, may appear less prominent than other features with a more widespread distribution (e.g., road infrastructure).

Overall, results from this sensitivity analysis suggest that broad spatial patterns of statistically significant clusters may be detected using a range of cell sizes. Each cell size, however, has trade-offs relative to the others. In general, a coarser grid will enhance the visibility of clusters but may fail to detect hotspots of a relatively low concentration of features. In contrast, a finer grid is more likely to depict low-concentration spatial features as clusters but may appear to over-represent features with a broad geographic distribution relative to more localized features. These trade-offs should be carefully considered when selecting a cell size for a particular analytical context.

Testing User Interpretation

The interpretation of maps created with *Pythia* depends on whether the user can associate the combinations of the three subtractive colors (CYM) with combinations of development factors or variables. We, therefore, tested the reasonableness of this assumed association with a simple cognition experiment. Twenty graduate students were shown the three case study maps (Figs. 9, 11, and 13) with the legends removed. They were then asked to infer what each color on each map represented. The aim was to determine whether combinations of colors could convey the concept of spatial intersection of multiple datasets even without the benefit of a legend making this explicit.

Results showed strong indications of an intuitive association between colors and variables. Most respondents correctly inferred that red, green, blue, and black represented combinations of cyan, yellow, and magenta (Fig. 9: 72%; Fig. 11: 70%; Fig. 13: 75%). For Figs. 11 and 13, half of the respondents could associate each color with the correct combination of datasets. For Fig. 13, 65% of respondents intuited that darker shades indicated higher values of the respective datasets being intersected. When shown the figure with the legend and caption, 85% of respondents correctly identified that the level of shading corresponded to the level of statistical significance of

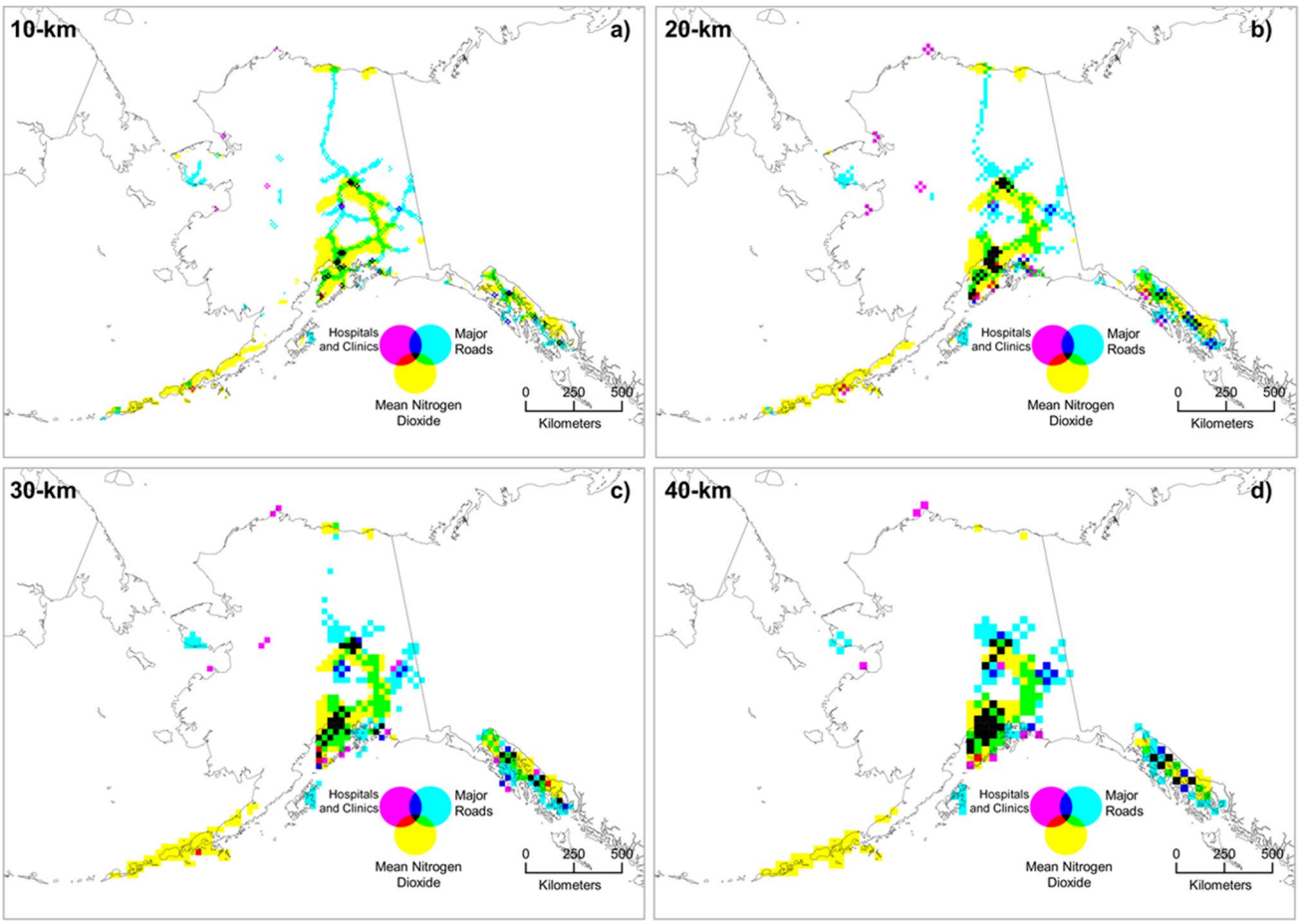

Fig. 14 Locations identified in *Pythia* as hotspots of hospitals and clinics (magenta), major roads (cyan), mean concentration of nitrogen dioxide (yellow), and overlapping combinations of the three variables, using a cell size of **a** 10×10 km, **b** 20×20 km, **c** 30×30 km, and **d** 40×40 km

the underlying data, even though they had not been previously informed that colored areas represented statistically significant hotspots. These results support the notion of an implicit association between CYMK colors and the spatial datasets they represent, lending confidence to the interpretability of maps created using this method. Therefore, we would expect an even larger number of respondents to correctly associate colors and variables if a legend had been shown.

Discussion and Conclusion

The initial identification of locations in the Arctic with characteristics that would enable/constrain a specific mode of development is possible through mapping and spatial analysis. However, mapping techniques based on simple overlays of spatially heterogeneous data may result in visual clutter, compromising the legibility of the map and increasing the likelihood of interpretation errors. The overrepresentation of particular features, like airports, may overwhelm the map and obscure other key features, which may exhibit statistically significant clustering but risk being overlooked due to having fewer features overall. Analyzing multiple raster datasets in a single map may also result in map illegibility, as cells of overlapping raster datasets may not be easily distinguished.

The *Pythia* tool was developed to address these challenges and allow for the overlay of multiple datasets to visualize and explore the spatial distribution of features relevant to a particular mode or lens of Arctic development. The tool facilitates the identification, characterization, and visualization of current and potential places of human activity within the Arctic based on specific variables selected for analysis. Based on spatial relationships, the tool uses the G or G^* statistic to identify significant locations or hotspots where particular features occur.

The G statistic and its corresponding z- and p-values are calculated for each dataset. Statistically significant (p-value < 0.05) high clusters are then extracted for each dataset, assigned a value of 1, and concatenated to a series. A CYMK is assigned to each variable, and consequently, an RGB color scheme results if a hotspot is detected for two of

the analyzed variables. A hotspot detected in all three datasets results in the full combination of colors (black). If more than one dataset is used per analyzed variable, the saturation of the colors will depend on the combined significance of the input datasets.

Three case studies were conducted to showcase this multivariate visualization tool. Case study 1 identifies Nome as a location projected to experience an amplified temperature anomaly and increased population by 2040 under an SSP3 climate change scenario. Case study 2 revealed statistically significant locations with major roads and high NO_2 concentrations but no hospitals and clinics in Prudhoe Bay, Big Delta, Gulkana, and near Anchorage. In case study 3, an overlap of transportation infrastructure and protected areas was detected in Bethel and Cordova, but with no significant travel and tourism infrastructure, which could signal challenges for the tourism industry.

Comparing the results of these case studies allows for further characterization of locations. The results indicate positive and negative lenses of development, such as the capacity to absorb increased tourism activities vs. environmental health consequences that may result from such activities. These results, however, are also complementary, highlighting combinations of critical enabling/constraining factors occurring in a specific location. Case study 1 shows projected increased mean temperature anomaly in Utqiagvik, while case studies 2 and 3 show this location as a hotspot for travel and tourism, as well as health infrastructure. Increasing temperatures, however, will reduce sea ice, consequently impacting coastlines (e.g., flooding and erosion) and local populations (e.g., displacement and drinking water contamination), which may constrain development.

In Prudhoe Bay, increased temperatures could result in reduced sea ice, permafrost thawing, and coastal erosion, furthering the potential for increased pollution due to oil spills. In Nome, the projected increase in temperatures may exacerbate already occurring coastal erosion, placing existing infrastructure at risk of severe damage. A hotspot of transportation and protected areas was found in Bethel, but no statistically significant travel and tourism infrastructure was located. While the lack of tourism infrastructure could be a possible constraining factor for the growth of tourism in the southwest, the region, particularly Cold Bay, did not show a projected increase in population (case 1) and was identified as a hotspot in case 2, illustrating some potential for growth in other sectors.

The flexibility to choose a cell size also enables users to define the scale of analysis from global to local. Because datasets are aggregated into a cell, results also depend on the dimension of the cell; larger cells will increase the level of aggregation of the features analyzed. The dimension of the cell, however, would also be dependent on the scale of analysis and the size of the study area. An optimal resolution would maximize cluster visibility and allow for efficient user interpretation without overaggregating the datasets since the purpose is to locate extremes or hotspots (Hengl 2006).

Each combination of variables represented a particular lens to showcase this visualization tool. Results will, therefore, depend on variables the user selects and the quality of the acquired datasets. Because *Pythia* assumes that input datasets accurately reflect "ground truth," using incomplete (or incorrect) input data may result in maps that miss "true" hotspots in some areas while detecting "false" hotspots in other areas. As a result, users are advised to exercise caution when combining data obtained from multiple sources across study areas, as data availability and quality standards vary significantly across the Arctic region (Schwoerer et al. 2021).

Pythia should only be used as an initial screening of sites based on a small number (1–3) of variables and datasets (1–6) of critical enabling/constraining factors. While this tool may aid policymakers in identifying areas lacking particular resources and infrastructure (e.g., travel and tourism infrastructure), it does not replace a comprehensive examination of the development potential of a location through fieldwork or deeper analysis beyond an exploratory phase.

The tool can also be used to explore possible futures with the use of projected data, which could lead to policy changes and strategic planning in Arctic contexts. Scenarios are often used as a tool in long-range planning and policy analysis (e.g., Lempert et al. 2003; Moniz 2006) and have frequently been utilized to understand potential trajectories of climate change and impacts of alternative policies on mitigation and adaptation measures (e.g., Ebi et al. 2014). Yet many available scenarios to support Arctic strategy, planning, and policy rely primarily on qualitative and fairly generalized narratives (e.g., Middleton et al. 2021) or focus mainly on quantitatively modeling physical changes with limited intersections with sociocultural or economic data (e.g., Hjort et al. 2018).

Due to a high degree of uncertainty in the region, researchers are advised to consider the various development trajectories from an interdisciplinary position when constructing Arctic scenarios (Zaikov et al. 2019). Furthermore, the Arctic is not a homogenous region (e.g., Young and Einarsson 2004); thus, different parts of the Arctic may experience different changes and responses to climate change. *Pythia* can facilitate the overlay of multiple spatial datasets, each exploring the spatial distribution of features relevant to a trajectory using a solid statistical approach, particularly in cases like the ones presented here, where multiple factors could become strong determinants of the future of a region. The results of this approach provide a more accessible means of distilling a bottom line for researchers, policymakers, Arctic residents, commercial entities, and other stakeholders.

In future research, we envision that *Pythia* can help baseline present-day characteristics and provide insights into future changes given specific assumptions. For example, *Pythia* can provide the basis for geographically depicting future scenarios of human migration due to climate change, development of health crises, and stimulation of tourism by assuming a continuation of the patterns seen in the three case studies, respectively, described in this paper. Policymakers and stakeholders could then use the results of such analysis, as the final visualization would represent a possible "future" map and "look back over the years to see how they got to where they are today" (Gordon 2021, p. 237).

Funding Open access funding provided by SCELC. This research was supported by the National Science Foundation through grant NNA 2022523.

Declarations

Ethics approval This research adheres to ethical standards. This research did not involve human subjects, animals, or sensitive data.

Consent to participate Informed consent was not applicable to this research.

Competing interests The authors report there are no competing interests to declare.

References

Agyeman PC, Kebonye NM, John K, Haghnazar H, Borůvka L, Vašát R (2023) Compositional mapping, uncertainty assessment, and source apportionment via pollution assessment-based receptor models in urban and peri-urban agricultural soils. J Soils Sediments 23(3):1451–1472. https://doi.org/10.1007/s11368-022-03417-3

Alaska Department of Commerce. (2022) Alaska statewide: comprehensive economic development strategy 2022–2027 (Issue October). https://www.commerce.alaska.gov/web/Portals/0/pub/CEDS/Final 2022–2027 Alaska Statewide CEDS.pdf

Alaska Department of Labor (2023) Alaska economic trends: the working-age population decline (Issue March). https://live.laborstats.alaska.gov/trends-magazine/2023/March/the-decline-in-working-age-alaskans

Alaska DOT&PF (2024) Airports. https://gis.data.alaska.gov/datasets/AKDOT::airports-1/about. Accessed Jan 2024

Albanese S, Cicchella D, De Vivo B, Lima A, Civitillo D, Cosenza A, Grezzi G (2011) Advancements in urban geochemical mapping of the Naples metropolitan area: colour composite maps and results from an urban brownfield site. In: Mapping the Chemical Environment of Urban Areas, pp. 410–423. https://doi.org/10.1002/9780470670071.ch24

Andrade C (2019) The P value and statistical significance: misunderstandings, explanations, challenges, and alternatives. Indian J Psychol Med 41(3):210–215

Anselin L, Rey SJ (2014) Modern spatial econometrics in practice: a guide to GeoDa, GeoDaSpace and PySAL. GeoDa Press LLC, Chicago, IL, USA

Bennett MM (2016) Discursive, material, vertical, and extensive dimensions of post-Cold War Arctic resource extraction. Polar Geogr 39(4):258–273. https://doi.org/10.1080/1088937X.2016.1234517

Bleha B, Ďurček P (2023) Unambiguous linkage between the vaccination coverage and the spread of COVID-19: geostatistical evidence from the Slovak LAU 1 Regions. J Geovis Spat Anal 7(1):1–11. https://doi.org/10.1007/s41651-023-00144-2

Brady MB, Leichenko R (2020) The impacts of coastal erosion on Alaska's North Slope communities: a co-production assessment of land use damages and risks. Polar Geogr 43(4):259–279. https://doi.org/10.1080/1088937X.2020.1755907

Bronen R (2010) Forced migration of Alaskan Indigenous communities due to climate change. In Afifi T & Jäger J (eds.), Environment, forced migration and social vulnerability. Springer, Berlin Heidelberg, pp. 87–98. https://doi.org/10.1007/978-3-642-12416-7_7

Bystrowska M, Dawson J (2017) Making places: the role of Arctic cruise operators in 'creating' tourism destinations. Polar Geogr 40(3):208–226. https://doi.org/10.1080/1088937X.2017.1328465

Carstensen LW (1984) Perceptions of variable similarity on bivariate choroplethic maps. Cartogr J 21(1):23–29. https://doi.org/10.1179/caj.1984.21.1.23

Carvalho KS, Smith TE, Wang S (2021) Bering Sea marine heatwaves: patterns, trends and connections with the Arctic. J Hydrol 600:126462. https://doi.org/10.1016/j.jhydrol.2021.126462

Copernicus Sentinel-5P (2018) TROPOMI level 2 nitrogen dioxide total column products. Version 01. European Space Agency. https://doi.org/10.5270/S5P-s4ljg54

Dent BD, Torguson J, Hodler TW (2009) Cartography: thematic map design. McGraw-Hill Higher Education, New York

Dewitz, J. (2019) National Land Cover Database (NLCD) 2016 products: US Geological Survey data release, https://doi.org/10.5066/P96HHBIE.

Ebi KL, Kram T, Van Vuuren DP, O'Neill BC, Kriegler E (2014) A new toolkit for developing scenarios for climate change research and policy analysis. Environment 56(2):6–16. https://doi.org/10.1080/00139157.2014.881692

Eliasson K, Ulfarsson GF, Valsson T, Gardarsson SM (2017) Identification of development areas in a warming Arctic with respect to natural resources, transportation, protected areas, and geography. Futures 85:14–29

Getis A, Ord JK (1992) The analysis of spatial association by use of distance statistics. Geogr Anal 24(3):189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

Gordon HSJ (2021) Ethnographic futures research as a method for working with Indigenous communities to develop sustainability indicators. Polar Geogr 44(4):233–254. https://doi.org/10.1080/1088937X.2021.1881647

Graybill JK, Petrov AN (eds) (2020) Arctic sustainability, key methodologies and knowledge domains: a synthesis of knowledge I. Routledge, New York

Grenier AA (2007) The diversity of polar tourism. Some challenges facing the industry in Rovaniemi, Finland. Polar Geography 30(1–2):55–72. https://doi.org/10.1080/10889370701666622

Guagliardi I, Zuzolo D, Albanese S, Lima A, Cerino P, Pizzolante A, Thiombane M, De Vivo B, Cicchella D (2020) Uranium, thorium and potassium insights on Campania region (Italy) soils: sources patterns based on compositional data analysis and fractal model. J Geochem Explor 212:106508. https://doi.org/10.1016/j.gexplo.2020.106508

Gupta S, Dharmaraj T, Reddy KM, Ravisankar T (2020) Spatial-temporal analysis and visualization of rural development works implemented under world's largest social safety programme in India—a case study. J Geovis Spat Anal 4(2). https://doi.org/10.1007/s41651-020-00062-7

Heleniak T (2021) The future of the Arctic populations. Polar Geogr 44(2):136–152. https://doi.org/10.1080/1088937X.2019.1707316

Hengl T (2006) Finding the right pixel size. Comput Geosci 32(9):1283–1298

Hillmer-Pegram K (2016) Integrating Indigenous values with capitalism through tourism: Alaskan experiences and outstanding issues. J Sustain Tour 24(8–9):1194–1210. https://doi.org/10.1080/09669582.2016.1182536

Hjort J, Streletskiy D, Doré G, Wu Q, Bjella K, Luoto M (2022) Impacts of permafrost degradation on infrastructure. Nature Reviews Earth and Environment 3(1):24–38. https://doi.org/10.1038/s43017-021-00247-8

Hjort J, Karjalainen O, Aalto J, Westermann S, Romanovsky VE, Nelson FE, Etzelmüller B, Luoto M (2018) Degrading permafrost puts Arctic infrastructure at risk by mid-century. Nat Commun 9(1). https://doi.org/10.1038/s41467-018-07557-4

IHME (2020) Health research by location. https://www.healthdata.org/research-analysis/health-by-location/profiles/united-states-alaska. Accessed Feb 2024

IUCN & UNEP-WCMC (2022) The World Database on Protected Areas (WDPA) https://www.protectedplanet.net/en/search-areas?filters%5Bdb_type%5D%5B%5D=wdpa&geo_type=region, Cambridge, UK: UNEP-WCMC. Available at www.protectedplanet.net. Accessed June 2022

Jones B, O'Neill BC (2020) Global one-eighth degree population base year and projection grids based on the shared socioeconomic pathways, revision 01. NASA Socioeconomic Data and Applications Center (SEDAC), New York. https://doi.org/10.7927/m30p-j498

Kalinic M, Krisp JM (2018) Kernel Density Estimation (KDE) vs. Hot-Spot Analysis—Detecting Criminal Hot Spots in the City of San Francisco. In: Proceedings of the 21st International Conference on Geographic Information Science (AGILE 2018), Lund, Sweden, pp 1–5. https://www.researchgate.net/profile/Maja-Kalinic-2/publication/325825793_Kernel_Density_Estimation_KDE_vs_Hot-Spot_Analysis_-_Detecting_Criminal_Hot_Spots_in_the_City_of_San_Francisco/links/5b27de230f7e9b332a31af55/Kernel-Density-Estimation-KDE-vs-Hot-Spot-Analysis-Detecting-Criminal-Hot-Spots-in-the-City-of-San-Francisco.pdf

Kebonye NM, Agyeman PC, Seletlo Z, Eze PN (2023) On exploring bivariate and trivariate maps as visualization tools for spatial associations in digital soil mapping: a focus on soil properties. Precision Agric 24(2):511–532. https://doi.org/10.1007/s11119-022-09955-7

Lempert R, Popper S, Bankes S (2003) Scenario building for policy. In Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis (p. 187). RAND Corporation. https://doi.org/10.7249/MR1626

Li T (2022) A spatiotemporal analysis of rock concerts associated with demographics and leisure and hospitality employment. J Geovis Spat Anal 6(1). https://doi.org/10.1007/s41651-022-00116-y

Loring PA, Gerlach SC (2009) Food, culture, and human health in Alaska: an integrative health approach to food security. Environ Sci Policy 12(4):466–478. https://doi.org/10.1016/j.envsci.2008.10.006

Lotfata A (2022) Using geographically weighted models to explore obesity prevalence association with air temperature, socioeconomic factors, and unhealthy behavior in the USA. J Geovis Spat Anal 6(1):1–12. https://doi.org/10.1007/s41651-022-00108-y

Martin S (2009) The effects of female out-migration on Alaska villages. Polar Geogr 32:61–67

McLoone P, Dyussupov O, Nurtlessov Z, Kenessariyev U, Kenessary D (2021) The effect of exposure to crude oil on the immune system. Health implications for people living near oil exploration activities. Int J Environ Health Res 31(7):762–787. https://doi.org/10.1080/09603123.2019.1689232

Middleton A, Lazariva A, Nilssen F, Kalinin A, Belostotskaya A (2021) Scenarios for Sustainable Development in the Arctic until 2050. Northern Research Forum. https://oulurepo.oulu.fi/handle/10024/45521

Moniz AB (2006) Scenario-building methods as a tool for policy analysis. In Innovative comparative methods for policy analysis. Kluwer Academic Publishers, pp. 185–209. https://doi.org/10.1007/0-387-28829-5_9

Morshed MM, Chakraborty T, Mazumder T (2022) Measuring Dhaka's urban transformation using nighttime light data. J Geovis Spat Anal 6(2). https://doi.org/10.1007/s41651-022-00120-2

Myllylä Y, Kaivo-Oja J, Juga J (2016) Strong prospective trends in the Arctic and future opportunities in logistics. Polar Geogr 39(3):145–164. https://doi.org/10.1080/1088937X.2016.1184723

Natural Earth (2022) 1:10m Cultural Vectors. https://www.naturalearthdata.com/

Nelson D (2010) Shaded Relief Archive: Alaska (Nelson). https://www.shadedreliefarchive.com/alaska-nelson.html

Nong D, Countryman AM, Warziniack T (2018) Potential impacts of expanded Arctic Alaska energy resource extraction on US energy sectors. Energy Policy 119:574–584. https://doi.org/10.1016/j.enpol.2018.05.003

Olson JM (1981) Spectrally encoded two-variable maps. Ann Assoc Am Geogr 71(2):259–276. https://doi.org/10.1111/j.1467-8306.1981.tb01352.x

OpenStreetMap & Geofabrik (2023) OpenStreetMap Data Extracts. http://download.geofabrik.de/. Accessed Feb 2024

Ord JK, Getis A (1995) Local spatial autocorrelation statistics: distributional issues and an application. Geogr Anal 27(4):286–306. https://doi.org/10.1111/j.1538-4632.1995.tb00912.x

Pashkevich A, Stjernström O, Lundmark L (2016) Nature-based tourism, conservation and institutional governance: a case study from the Russian Arctic. Polar Journal 6(1):112. https://doi.org/10.1080/2154896X.2016.1171000

Percival JEH, Tsutsumida N, Murakami D, Yoshida T, Nakaya T (2022) Exploratory spatial data analysis with gwpcorMapper: an interactive mapping tool for geographically weighted correlation and partial correlation. J Geovis Spat Anal 6(1):17

Rahman MM, Kamruzzaman M, Shahid S, Thorp KR, Rahaman H, Shahriyar MM, Islam AKMS, Huda MD (2023) A GIS framework to demarcate suitable lands for combine harvesters using satellite DEM and physical properties of soil. J Geovis Spat Anal 7(2):1–18. https://doi.org/10.1007/s41651-023-00156-y

Reza MS, Sabau G (2022) Impact of climate change on crop production and food security in Newfoundland and Labrador, Canada. J Agric Food Res 10. https://doi.org/10.1016/j.jafr.2022.100405

Rosenholtz R, Li Y, Nakano L (2007) Measuring visual clutter. J vis 7(2):1–22. https://doi.org/10.1167/7.2.17

Resource Development Council for Alaska (2023) Alaska's tourism industry. https://www.akrdc.org/tourism

Schwoerer T, Spellman KV, Davis TJ, Lee O, Martin A, Mulder CPH, Swenson NY, Taylor A, Winter G, Giguere N (2021) Harnessing the power of community science to address data gaps in Arctic observing. Arctic 74:1–14

Sherval M (2013) Arctic Alaska's role in future United States energy independence. Polar Geogr 36(4):305–322. https://doi.org/10.1080/1088937X.2013.827756

Strode G, Morgan JD, Thornton B, Mesev V, Rau E, Shortes S, Johnson N (2020) Operationalizing Trumbo's principles of bivariate choropleth map design. Cartogr Perspect. https://doi.org/10.14714/CP94.1538

Tinghua A, Mengjie Z, Xiaoming L (2017) Mining co-location pattern of network spatial phenomenon based on the law of additive color mixing. Acta Geodaetica et Cartographica Sinica 46(6):753–759. https://doi.org/10.11947/j.AGCS.2017.20160324

Touya G, Decherf B, Lalanne M, Dumont M (2015) Comparing image-based methods for assessing visual clutter in generalized maps. ISPRS Ann Photogramm Remote Sens Spat Inf Sci 2(3W5), 227–233. https://doi.org/10.5194/isprsannals-II-3-W5-227-2015

Trumbo BE (1981) A theory for coloring bivariate statistical maps published by : Taylor & Francis, Ltd. on behalf of the American Statistical Association Stable URL : http://www.jstor.org/stable/2683294 All use subject to http://about.jstor.org/terms A Theory for Colori. The American Statistician, 35(4), 220–226.

US Bureau of Economic Analysis (2023) Economic profile for Alaska. https://apps.bea.gov/regional/bearfacts/?f=02000&a=3

US Bureau of Labor Statistics (2023) Women's earning in Alaska – 2021. https://www.bls.gov/regions/west/news-release/womensearnings_alaska.htm

US Census Bureau (2022) 2021 American Community Survey, 1-Year Estimates, Table S1501; https://data.census.gov/profile/Alaska?g=040XX00US02

US EPA, 2023 US EPA (2023) Nitrogen dioxide pollution. https://www.epa.gov/no2-pollution

VanderBerg JD (2018) Optimal Arctic Port locations: a quantitative composite multiplier analysis of potential sites. Polar Geogr 41(1):55–74. https://doi.org/10.1080/1088937X.2017.1400604

Veland S, Lynch AH (2017) Arctic ice edge narratives: scale, discourse and ontological security. Area 49(1):9–17. https://doi.org/10.1111/area.12270

Vieux BE, Needham S (1993) Nonpoint-pollution model sensitivity to grid-cell size. J Water Resour Plan Manag 119(2):141–157

World Bank (2021) Climate Change Knowledge Portal: current climate, trends and variability. https://climateknowledgeportal.worldbank.org/country/united-states/trends-variability-historical. Accessed Feb 2024

Yamada I, Thill J-C (2010) Local indicators of network-constrained clusters in spatial patterns represented by a link attribute. Ann Am Assoc Geogr 100(2):269–285

Young OR, Einarsson S (2004) A human development agenda for the Arctic: major findings and emerging issues. In: Arctic Human Development Report. pp 229–242 https://oaarchive.arctic-council.org/items/f6c63158-401c-4a14-a8d2-5bc9f10710bf/full

Zaikov K, Kondratov N, Kudryashova E, Lipina S, Chistobaev A (2019) Scenarios for the development of the Arctic region (2020–2035). Arctic North 35(35):5–24. https://doi.org/10.17238/issn2221-2698.2019.35.5

Zegre SJ, Needham MD, Kruger LE, Rosenberger RS (2012) McDonaldization and commercial outdoor recreation and tourism in Alaska. Manag Leis 17(4):333–348. https://doi.org/10.1080/13606719.2012.711604