Heliyon 9 (2023) e18351

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Clustering energy support beliefs to reveal unique sub-populations using self-organizing maps

Heather Bedle [a,*], Christopher R.H. Garneau [b], Alexandro Vera-Arroyo [a]

[a] *School of Geosciences, University of Oklahoma, USA*
[b] *Department of Sociology, University of Oklahoma, USA*

A R T I C L E I N F O

Keywords:
Energy preferences
Self-organizing maps
Renewable energy
Fossil fuels

A B S T R A C T

Americans' support for energy sources is quite complex and dependent on a range of socio-demographic characteristics. In a novel approach to investigate energy opinions on renewables and fossil fuels, self-organizing maps are employed to cluster individuals solely on their energy beliefs using social survey data collected from Pew Research in the Spring of 2021. These energy preference clusters are then used in regression models to examine attitudes regarding energy policy in the United States. Results from the self-organizing map (SOM) analysis reveal four distinct clusters: energy traditionalists who oppose renewable sources due to partisan ideologies; energy renewers who strongly prefer investment in only renewable energy sources; energy universalists who universally support all forms of energy; and the aberrant cluster, individuals who prefer solar power greatly over wind energy but demonstrate no other energy preference patterns. Results from regression analyses reveal that SOM clusters are highly predictive of attitudes regarding energy policy. Taken together, these results reveal the unique capability of machine learning to categorize human attitudes – which should be of particular interest to energy policymakers when considering the opinions of the electorate.

1. Introduction

With the onslaught of information over the last decade creating concern for the byproducts and unintended consequences of fossil fuel usage [1], American opinions regarding energy sources – from hydrocarbon-based, to nuclear, solar, and wind - remain varied [2–6]. While acceptance of renewable energy sources is not as polarizing an issue as climate change, policy initiatives such as the Green New Deal have incited partisan opinions on energy sources [7]. While there is support amongst Americans for the transition to renewable energy sources in the future [2,8–11], opposition to moving away from fossil fuels exists in a substantial portion of the population, as renewable energy is linked to global climate issues that are more political and ideological in nature [12–17]. Investigating differences of opinion with traditional social science methods provides valuable insights into the views of Americans, but focusing only on support for energy as a general concept instead of investigating the complex interplay of beliefs among clusters of people is a limited method to capture overall attitudes toward energy preferences. Instead, to characterize Americans' opinions on energy strategies, the multidimensional belief systems of a highly heterogeneous population must be considered. We achieve this by utilizing self-organizing maps to cluster Americans based solely on their attitudes toward the expansion of six types of energy: wind farms, solar farms, hydraulic fracturing, coal mining, offshore drilling, and nuclear energy. From the resultant clusters that emerge

* Corresponding author.
E-mail address: hbedle@ou.edu (H. Bedle).

https://doi.org/10.1016/j.heliyon.2023.e18351
Received 12 December 2022; Received in revised form 3 July 2023; Accepted 13 July 2023
Available online 15 July 2023

based on individuals' energy preferences, the socio-demographics of each cluster are examined to reveal meaningful patterns.

Furthermore, we regress attitudes regarding energy policy on the SOM clusters to reveal how the multidimensional positioning of groups of Americans greatly impacts support for or against alternative energy sources (i.e. alternatives to fossil fuels) at a time when energy policy in the United States is transitioning to a more carbon neutral future. The energy transition requires policy and research implementation at the national level, and in a democratic society (such as the United States), public opinion is essential for enduring changes to energy policy - especially when change is not in the short-term economic interest of the public [3,5,18,19]. From a practical standpoint, this research is highly consequential for energy policy. The ability to identify groupings of Americans by energy-preference commonalities can assist policymakers and stakeholders in reaching portions of the general public who might be more likely to support renewable energy through targeted public relations campaigns.

1.1. General trends

Focusing on environmental concerns, Wave 89 of the Pew American Trends Panel Survey was collected from a representative sample of U.S. adults in April 2021. Many general trends emerge in the data and serve as an excellent starting point for our analysis [20]. Results for support for or against expanding a variety of energy sources and political affiliation and orientation present some of the more substantial trends in energy attitudes (Fig. 1), with Democrats of all political orientations largely supporting renewable energies. Conversely, renewable support diminishes greatly for Republicans – especially conservative Republicans, although a majority do favor renewables. Hydrocarbon-related energy practices, from hydraulic fracturing to coal and offshore drilling are robustly opposed as future energy sources by Democrats. While support for non-renewables is greater for Republicans compared to Democrats, it is clear that non-renewables are far more favored by conservative Republicans. Overall, these patterns show that Democrats of all orientations are united in support of renewable energy and opposition to non-renewable sources. Republican support is far more sensitive to orientation with conservative Republicans showing less support for renewable energy sources and more support for non-renewable sources relative to moderate/liberal Republicans.

Additional findings reveal generational trends with Millennials and Generation Z favoring renewable energy expansion, and Boomers and older generations more heavily supporting fossil fuels. Sex, ethnicity, and education also play minor roles in energy preferences as females, Hispanic Americans, and college educated Americans are substantially more in favor of expanding renewable energy sources. Many of these trends have been noted in previous studies (e.g. Refs. [4,5,21–24]). What remains a greater mystery, hidden in the multi-dimensional variables that attempt to define and encapsulate an individual, is how differing groups of people perceive and rank support for energy sources and then use those perceptions to develop ideas about policy. In this endeavor, we are not simply concerned with whether they favor one type of energy and oppose another, but we seek to gain access into a more topographically varied view of energy support – where some groups of individuals may support most energy types, but vehemently oppose others. Or rather, have no preference at all, and only worry about the monetary cost of their monthly energy bills, as climate change is largely unperceivable in day-to-day affairs. In sum, we assert that assessing policy positions as an outcome of support for energy types is best suited by machine learning methods that can disentangle unique clusters of individuals based on their support for energy types – a task that is ill-suited for a theory-based strategy prone to researcher bias. As such, we employ self-organizing maps to identify meaningful clusters of individuals assigned by their preferences of future energy sources to address our general hypothesis that organizing survey respondents into SOM clusters is a useful and novel strategy for analyzing survey data with traditional social science

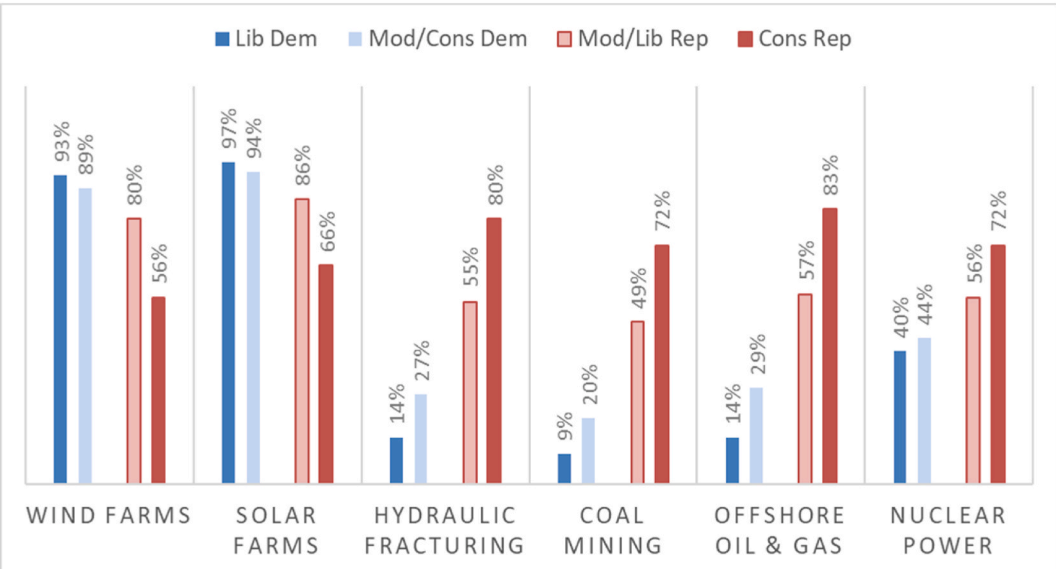

Fig. 1. Trends of energy source expansion by ideology and partisanship.

methods. To accomplish this, we use an analytical phase with multiple stages: 1) utilize self-organizing maps to identify clusters from energy preference variables, 2) use descriptive statistics to describe the energy preferences of the emergent clusters, 3) use descriptive statistics to reveal unique demographic characteristics of each cluster, and 4) employ binary logistic regression analyses to demonstrate that energy preference clustering significantly predicts energy policy positions.

2. Methodology

2.1. Data

Data for this study come from the Pew Research Center's American Trends Panel (ATP), which was first created in 2014. Data only from Wave 89 of the ATP were used in this study. For Wave 89, data were collected from April 20 to April 29, 2021. Details of the collection protocol, data quality checks, and weighting protocol are available from PEW (see Data Availability). After employing listwise deletion of non-answers, 6920 observations remained in the dataset.

Table 1
Variable Recodes from Pew survey identifiers.

	Pew data identifier	
	Variable name	Value
SOM Variables		
Wind farms	ENV2f, ENV2g	–
Solar panel farms	ENV2d	–
Hydraulic fracturing	ENV2e	–
Coal mining	ENV2c	–
Offshore oil & gas	ENV2a	–
Nuclear power plants	ENV2b	–
Dependent Variables		
Alternative energy	EN1	1,2
Clean energy only	EN2	1,2
Political Variables		
Liberal Democrat (ref)	PARTYSUMIDEO_FINAL	4
Mod/Con Democrat	PARTYSUMIDEO_FINAL	3
Lib/Mod Republican	PARTYSUMIDEO_FINAL	2
Conservative Republican	PARTYSUMIDEO_FINAL	1
Control Variables		
Income Cat	INC_SDT1	–
Female/Other	GENDER	2,3
White (ref)	RACE	1
Hispanic	RACE	3
African American	RACE	2
Other	RACE	4,5
Degree	EDUCCAT	1
Region South (ref)	CREGION	2
Region NE	CREGION	1
Region Midwest	CREGION	3
Region West	CREGION	4
Attend Church	ATTEND	–
Born Again	BORN	1
Silent Generation	GENERATIONS	2
Boomer Generation (ref)	GENERATIONS	3
Generation X	GENERATIONS	4
Millennials	GENERATIONS	5
Generation Z	GENERATIONS	6
c) Economic Variables		
Economy now	ECON1	–
Economy future	ECON1B	–
Grow jobs, economy	CCPROPd	–
Consumer costs low	CCPROPe	–
d) Energy/Climate Variables		
Wind cost	WINDa	–
Wind reliability	WINDb	–
Wind good for envir.	WINDc	–
Wind knowledge	WINDHRD	–
Solar cost	SOLARa	–
Solar reliability	SOLARb	–
Solar good for envir.	SOLARc	–
Solar knowledge	SOLARHRD	–
Humans cause climate change	EN7	–

*WIND and SOLAR questions were asked to half the survey respondents.

2.2. Data recoding

Several variables were recoded from the original dataset which we summarize in Table 1. When specific values were used in our recoding of dummy variables, those values are noted in the third column to show which categories in the original variable were recoded with a '1' as the affirmative category.

2.3. Self-organizing maps

As independent variables used in an analysis increase in number, visual interpretation becomes challenging because it is difficult to visualize data in more than three dimensions. Self-organizing maps (SOMs) are employed to help discriminate subsets of otherwise seemingly homogeneous populations in the survey data. SOMs are a common technique used in many fields to represent a multi-dimensional dataset in a two-dimensional visualization which can be understood and properly interpreted [25].

SOMs provide a technique for pattern identification in multidimensional space and are a handy tool for organizing large amounts of data that extend over multiple variables, such as in survey data. SOMs have been used in the social sciences to organize patterns in individual behaviors or beliefs, from tracking tourist behaviors [26], to understanding student behaviors related to learning [27–30], and to visualize the geography of climate beliefs in the United States [31]. This unsupervised machine learning method uses an artificial neural network to classify and cluster data in groups, introducing no bias or supervision from an interpreter in the process. SOMs are calculated in two steps – first, organizing multi-dimensional data in a two-dimensional grid of nodes that represent the data space [32], and second, classifying those nodes into similar clusters.

After performing listwise deletion on the dataset (n = 6920), and using the Kohonen package for R [33,34], SOM is run on six variables that assess the respondents' support (recoded as '1') or opposition (recoded as '0') to whether the United States should expand upon solar panel farms, wind turbine farms, nuclear power plants for electricity generation, coal mining, offshore oil and gas drilling in US waters, and hydraulic fracturing for oil and gas. After optimizing the nodal grid size so that no nodes remained empty based on the data distribution, heat maps of each normalized variable were generated (Fig. 2) showing the distribution of the six SOM variables in two-dimensional space. A higher value in the heat map for that node represents support for the expansion of that energy source. The wedge-shaped pieces in each node relay the relative importance of each of the six variables in that specific node.

After the SOM is generated, nodes are clustered using hierarchical clustering. By clustering, we can investigate the demographics of survey respondents in each cluster. While it is interesting to note how energy preferences cluster in Fig. 2b, the exciting revelations lie within understanding the characteristics of the people in each cluster. Elucidation of this information can be performed using traditional social statistical techniques to reveal details of the sub-populations regarding their energy preferences and policy insights.

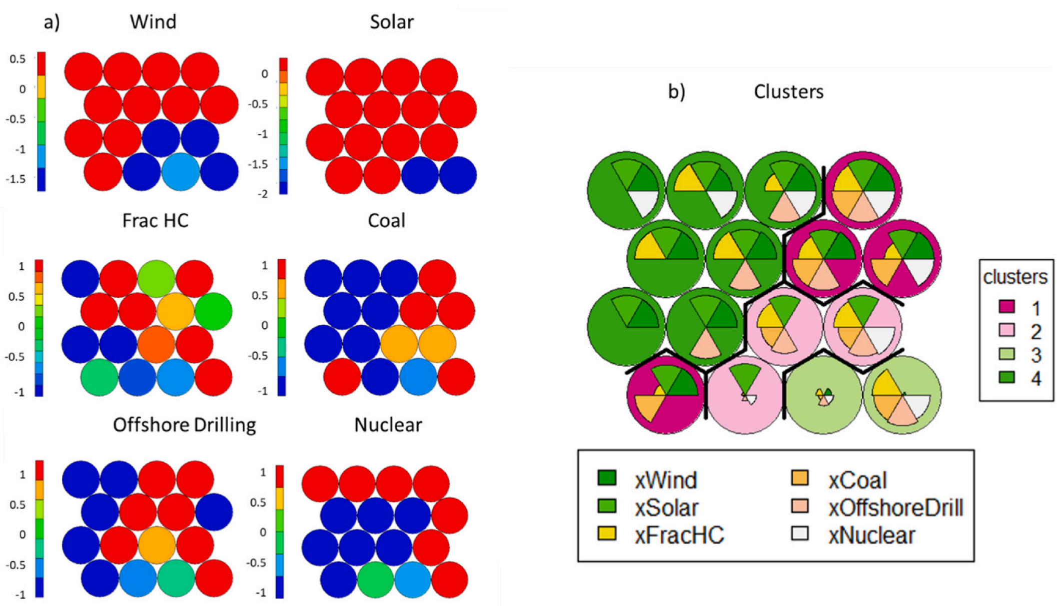

Fig. 2. a) Heat maps with the standardized value of each SOM input variable and its distribution in the two-dimensional nodal space; b) Clustered nodes, colored by assigned cluster.

3. Results

3.1. Clusters of energy support

The resultant four clusters and mean values of their sub-population's favorability toward the six energy types are summarized at the top of Table 2. Keeping in mind that only energy favorability was used in the SOM analysis, representing 27.7% of respondents, Cluster 1 (who we will refer to as universalists), consists of a population that is more favorable towards all energy types than the average population. Cluster 2 (aberrants), composing 11.2% of respondents, are those who fully support solar but oppose expanding wind energy, all while having slightly above average support for expanding hydrocarbon-related activities – making this cluster the least cohesive. At 16.8% of the sample, Cluster 3 (traditionalists), almost fully opposes expanding renewable energy, solar and wind, while supporting the expansion of hydrocarbon and nuclear energy capabilities. Finally, Cluster 4 (renewers), at 44.3% of the survey, strongly supports the expansion of wind and solar energy, while strongly opposing the expansion of hydrocarbon-related energy. Support for nuclear energy is below that of the other three clusters for Cluster 4.

To determine if the SOM clusters are statistically meaningful, binary logistic regressions are conducted using two dependent variables related to energy policy to evaluate cluster contribution (Table 3). Initially, we investigate the odds of supporting the position of developing alternative sources. The baseline model (Model 1) without SOM clusters shows that among a variety of control variables, politics is a primary driver of alternative energy exploration. Compared to liberal Democrats, the odds of supporting renewable exploration are far lower for all Republicans and also lower for moderate/conservative Democrats, to a lesser extent. With the addition of the SOM clusters into the regression model (Model 2), some control variables remain significant, such as Republican affiliation, having a college degree, and generational membership. With the introduction of SOM clusters, the difference between liberal and moderate/conservative Democrats is fully mediated – suggesting that the difference in variation has more to do with energy support clustering than political orientation among Democrats. The inclusion of the clusters (as shown in Model 2) also fully mediates the effect of income on support for developing alternative energy. With regard to the main effects of cluster variables, the odds of supporting alternative energy are 96.8% lower for traditionalists, 86.6% lower for universalists, and 89.4% lower for aberrants compared to renewers. While the clusters were created based on opinions for or against a variety of energy sources, the strong significance of these clusters in explaining the support of alternative energy policies demonstrates the usefulness of creating clusters of respondents with similar energy preferences from a larger set of energy types.

Regression results (shown in Model 3 and Model 4 of Table 3) report the odds ratios for a preference to "phase out the use of oil, coal, and natural gas completely, relying instead on renewable energy sources such as wind and solar power only" rather than the option to "use a mix of energy sources including oil, coal and natural gas along with renewable resources." As shown in Model 4, the clusters again have some of the strongest effects with all clusters reporting less support for policies to phase out non-renewables compared to the renewer cluster. While the self-organizing map algorithm is an unsupervised method of machine learning and is not designed to predict any particular dependent variable, but rather to organize the survey respondents into meaningful clusters with similar points of view, the significance of the cluster variables in these regressions clearly demonstrate that SOMs can be employed to extract information about subpopulations from survey data.

3.2. Demographics of energy support

Table 4 displays demographic information for each SOM cluster. Along with means for each demographic by cluster group, we include notation to demonstrate that a particular demographic variable significantly differs between cluster groups. Statistical significance is determined by conducting four (switching the reference between each of the four cluster groups) fully-controlled multinomial logistic regressions (full results not shown) where significance for the odds of being in one cluster group versus the reference is regressed on all focal and control variables. For example, results in Table 3 shows that 22% of universalists are moderate or conservative Democrats, which is a significantly lower percentage than is observed in the renewers cluster (39%) – as noted by the superscripted letter above the percentage. While not a focal component in our analyses, the multinomial logistic regressions demonstrate meaningful differences in demographic composition between SOM clusters.

Results on demographics show that renewers are more likely to identify as Democrat, regardless of ideology. While universalists, aberrants, and traditionalists are all more Republican, with traditionalists as the most conservative cluster. Renewers have high family

Table 2
The means by SOM clusters for each energy source.

	Full Sample	Cluster 1	Cluster 2	Cluster 3	Cluster 4
SOM Variables					
Wind farms	.74	**1.00**	**.00**	**.15**	**.99**
Solar panel farms	.82	**.97**	**1.00**	**.00**	**.99**
Hydraulic fracturing	.51	**.76**	.59	**.77**	**.25**
Coal mining	.46	**1.00**	.55	**.73**	**.00**
Offshore oil & gas	.54	**.81**	.59	**.78**	**.27**
Nuclear power	.55	**.66**	.55	**.71**	**.42**

Note: To ease interpretation, cluster group means substantially above the overall mean are colored green, while those substantially below the overall mean are colored red.

Table 3
Binary logistic regression demonstrates that SOM clusters based on energy source preference.

	Develop Alternative Energy		Phase out oil, coal, and gas	
	Model 1	Model 2	Model 3	Model 4
	O.R.	O.R.	O.R.	O.R.
Political Variables				
Lib. Democrat (ref)				
Mod/Con Democrat	.659**	.906	.518***	.576***
Lib/Mod Republican	.151***	.296***	.167***	.245***
Cons. Republican	.041***	.110***	.058***	.120***
Control Variables				
Income Cat	1.030**	1.003	1.006	.977
Female/Other	1.072	1.063	1.323***	1.453***
White (ref)				
Hispanic	1.101	1.234	1.456***	1.654***
African American	.928	1.267	.743**	.851
Other Race	1.262	1.094	1.289	1.283
College Degree	1.682***	1.436***	1.290***	1.081
Region South (ref)				
Region Northeast	1.090	1.148	.968	.931
Region Midwest	1.149	1.211*	.855	.834*
Region West	1.222*	1.173	1.092	1.011
Service Attendance	.959	.975	.964	.961
Born Again	.842*	.874	1.139	1.231**
Silent Generation	1.071	1.081	1.487**	1.424*
Boomer Gen. (ref)				
Generation X	1.197*	1.300**	1.515***	1.725***
Millennials	1.862***	2.153***	1.760***	1.965***
Generation Z	1.701***	1.753***	1.831***	1.991***
SOM Clusters				
Cluster 1 (Universalists)		.134***		.250***
Cluster 2 (Aberrants)		.106***		.264***
Cluster 3 (Traditionalists)		.032***		.164***
Cluster 4 (Renewers)				
Pseudo R^2	.400	.565	.279	.349
−2 Log Likelihood	6428.696	5187.219	6013.312	5620.394

Note: All data are weighted to adjust for non-response.
Significance markers: *** = >0.001; ** = >0.01; *= >0.05.

income and education attainment compared to other clusters. Females are over-represented in the aberrant cluster. Aberrants and renewers have higher proportions of Hispanics than universalists and traditionalists, while African Americans are overrepresented in the renewers, and underrepresented in the traditionalists. Individuals from other races are under-represented in aberrants and overrepresented in renewers. Overall, renewers have the highest percentage of non-whites. This clustering in renewers aligns with recent research that shows that whites have the lowest levels of reported concern for the environment compared to African Americans, Hispanics, and Asian-Americans [35,36]. The geographical Census regions are also significant in the SOM clusters. Those in the northeastern and western regions are slightly more represented in the renewers, especially when compared to universalists. Universalists are far less likely to be northeasterners when compared to traditionalists. Religion also plays a role in clustering, with traditionalists having higher rates of church attendance relative to universalists and renewers. Born-again Christians are also more likely to be clustered in traditionalists and universalists compared to renewers. Both religious measures are lowest in renewers.

There are clear generational effects on the clusters with traditionalists having the largest proportion of Boomers. The proportion of Generation X is significantly higher among the universalists as compared to every other cluster, and Millennials and Generation Z are least likely to be traditionalists. Overall, demographic differences between SOM clusters are clearly established in the survey data. These findings support research by Chapman et al. [37,38] who have highlighted the importance of factors such as cultural group, income, education, and energy technology awareness on the perception of individuals when considering energy systems and policies. Next, we investigate the role of income and economics in our clusters.

3.3. Economic concerns

To further demonstrate the usefulness of SOM clusters in survey data analysis - and to investigate some unexpected findings, we assess differences in economic views between cluster groups. Two questions regarding the respondents' perceptions of the US economy were asked in the survey, both currently (Spring 2021), and one year in the future. Higher values represent good or better perceptions of economic conditions, while lower values are indicative of poor and declining economic conditions. Additionally, two questions inquire about the respondents' priorities in consideration of U.S. proposals to reduce the effects of climate change. The first considers the importance of increasing job and economic growth while reducing climate change effects, and the second assesses the importance of keeping consumer costs low. In both cases, higher values signify higher importance for that consideration.

Table 4
Descriptive statistics by SOM Clusters for dependent variables.

	Full Sample		Cluster 1		Cluster 2		Cluster 3		Cluster 4	
			Universalists		Aberrant		Traditionalists		Renewers	
	M	SD	M	SD	M	SD	M	SD	M	SD
Political Variables										
Lib. Democrat (ref)	.17		.06		.09		.04		.30	
Mod/Con Democrat	.28		.22[d]		.25[d]		.11[d]		.39[abc]	
Lib/Mod Republican	.16		.20[d]		.21[d]		.11[d]		.14[abc]	
Cons. Republican	.38		.49[cd]		.45[cd]		.73[abd]		.17[abc]	
Control Variables										
Income Cat	4.89	3.06	4.57[d]	2.97	4.44[d]	3.03	4.87[d]	3.02	5.20[abc]	3.10
Female/Other	.52		.52		.56[cd]		.46[b]		.54[b]	
White (ref)	.68		.69		.68		.79		.63	
Hispanic	.13		.12[bd]		.17[acd]		.08[b]		.14[ab]	
Black	.14		.14[d]		.13[d]		.09[d]		.16[abc]	
Other Race	.05		.05[b]		.02[acd]		.04[b]		.07[b]	
Degree	.33		.24[bd]		.28[ad]		.27[d]		.42[abc]	
Region South (ref)	.42		.47		.42		.42		.39	
Region NE	.15		.13[cd]		.14		.15[a]		.17[a]	
Region Midwest	.23		.24		.24		.23		.22	
Region West	.20		.16[bd]		.20[a]		.20		.22[a]	
Attend Church	3.51	1.56	3.45[bc]	1.58	3.60[a]	1.57	3.85[ad]	1.59	3.37[c]	1.52
Born Again	.44		.50[d]		.47		.53[d]		.35[a]	
Silent Generation	.07		.06		.06		.11		.07[b]	
Boomer Gen. (ref)	.36		.31		.37		.41		.36	
Generation X	.27		.31[bcd]		.25[a]		.24[a]		.25[a]	
Millennials	.22		.23[cd]		.23		.17[a]		.23[a]	
Generation Z	.08		.08[c]		.09		.07[a]		.09	

Note: All data are weighted to adjust for non-response.
a = significantly different from Universalists.
b = significantly different from Aberrants.
c = significantly different from Traditionalists.
d = significantly different from Renewers.

Table 5 displays mean scores for economic, climate, and energy variables for each cluster. Statistical significance is again assessed with results from fully-controlled multinomial logistic regressions (results not shown). As we look at how economic perceptions are distributed between clusters, we note that those who score higher on belief that the economy is and will remain strong are more likely

Table 5
Descriptive statistics by SOM Clusters for economic, energy, and climate variables.

	Full Sample		Cluster 1		Cluster 2		Cluster 3		Cluster 4	
			Universalists		Aberrants		Traditionalists		Renewers	
	M	SD	M	SD	M	SD	M	SD	M	SD
Economic Variables										
Economy now	2.28	.71	2.24[c]	.73	2.20	.77	2.11[ad]	.74	2.38[c]	.66
Economy in 1 year	2.08	.86	1.91[bcd]	.83	1.82[acd]	.85	1.52[abd]	.74	2.45[abc]	.75
Grow jobs, economy	3.57	.66	3.62	.64	3.62[c]	.66	3.58[bd]	.76	3.54[c]	.63
Consumer costs low	3.50	.65	3.63[bd]	.51	3.56[acd]	.67	3.64[bd]	.62	3.36[abc]	.66
Climate Variable										
Humans cause CC	3.00	.94	2.75[cd]	.84	2.72[cd]	.91	2.13[abd]	.89	3.55[abc]	.63
Energy Variables										
Wind cost	1.99	.80	1.95[bcd]	.74	2.29[acd]	.77	2.56[abd]	.65	1.73[abc]	.77
Wind reliability	1.64	.73	1.62[cd]	.72	1.41[c]	.66	1.18[abd]	.44	1.87[ac]	.75
Wind good for envir.	2.48	.70	2.52[bcd]	.58	1.87[ad]	.71	1.79[ad]	.70	2.84[abc]	.43
Wind knowledge	2.05	.60	1.94[d]	.59	2.03[d]	.66	2.20[d]	.62	2.07[abc]	.56
Solar cost	2.02	.83	2.12[bcd]	.83	1.93[ac]	.84	2.47[abd]	.70	1.81[ac]	.81
Solar reliability	1.82	.78	1.78[c]	.78	1.86[c]	.79	1.29[abd]	.52	2.05[c]	.74
Solar good for envir.	2.60	.63	2.57[cd]	.62	2.62[cd]	.56	1.88[abd]	.69	2.89[abc]	.33
Solar knowledge	2.00	.60	1.94[d]	.61	1.90[d]	.59	2.04[d]	.66	2.06[abc]	.57

Note: All data are weighted to adjust for non-response.
a = significantly different from Universalists.
b = significantly different from Aberrants.
c = significantly different from Traditionalists.
d = significantly different from Renewers.

to be renewers (strong support for expanding wind and solar, and not hydrocarbon-based energies). Conversely, those with the most pessimistic views of the current and future economy are significantly more likely to be traditionalists, who strongly support expanding only hydrocarbon-related energies. Respondents with weaker than average views on the strength of the economy are significantly more likely to be aberrants or traditionalists. The second set of variables regarding economic priorities are more elucidating for discriminating between universalists, aberrants, and traditionalists. Results show that those who more strongly prioritize growing jobs and the economy are more likely to be universalists or aberrants. Additionally, those with a strong preference for keeping consumer costs low are more likely to be universalists. While a complete economic assessment of the clusters is not possible given the survey questions, these significant differences coupled with the significantly lower family income, suggest that the preference of universalists to expand all energy types, without discrimination for climate consequences, may be based on economic and financial considerations, despite the universalists' recognition that human activities related to the burning of fossil fuels are causing climate change (Table 5, variable 'Humans cause CC'). Economic standing has been researched as a critical role in the acceptance and favoritism for renewable energies [39], and the mitigation of energy poverty in the United States [40] through governmental policies could play a role in the universalists preferences of energy types.

While the economic attitudes may explain differences in the universalist group, the startling lack of support for wind energy (Table 2) in the aberrant cluster remains mysterious. To address this issue, variables regarding the respondents' opinions of wind and solar energy are examined. For both wind and solar, the survey asked: 1) if that type of energy was of a higher cost to consumers (higher value represents higher perceived cost); 2) if that alternate energy source is reliable (higher value is higher reliability); 3) if the effect is a net positive for the environment, where higher values represent that energy is perceived to be better for the environment, and 4) how much the individual had read or heard about that energy source (higher values represent more perceived knowledge).

Examining perceptions of wind energy, aberrants, believe that wind is more expensive to generate than other energies, and is less reliable while having a worse overall effect on the environment than other energy sources. Additionally, aberrants rank the lowest in having read or heard much about wind energy, ranking themselves on average as only hearing 'a little.' The aberrants' responses concerning solar power are much more favorable, believing that is it slightly less expensive than other energy sources, and while slightly less reliable than other energies, it is better for the environment than those sources. Noting that the aberrant cluster is more female, Hispanic, and poorer than other clusters, as well as having a higher western population than universalists, and the lack of knowledge about wind energy, these SOM clusters identify a target population that could become supportive of wind energy with more awareness. Such tailoring of messaging was proposed as an effective method for enhancing policy support [41]. The demographics of this group align with previous results [35], finding that wind energy policies have higher support among Millennials, non-whites, the college educated, and liberals. Although the aberrants are more female, studies in the past have found that gender was not a significant factor in wind energy support [42–44].

4. Conclusions

SOM is an unbiased and efficient method to organize and cluster survey respondents based on their opinions of an issue, in this case, support for energy sources in the United States. We find that SOM clustering on social survey data allows for the complex analysis of multiple and divergent factors simultaneously. When analyzing support for or opposition to various energy sources in the US, clustering revealed three sub-groups that were more accepting of expanding fossil fuel energy sources: one group that was strongly opposed to solar and wind energy likely motivated by partisan ideology, another group that favored all types of energy expansion due to economic concerns for inexpensive energy, and a third group opposed to wind farms, most likely due to their beliefs about the cost and reliability of wind energy. Furthermore, SOM clusters are highly adaptive and can work in tandem with traditional social science methodology – including multivariable regressions analyses. As machine learning has been underutilized in the social sciences [45–47], AI techniques like SOM clustering offer great benefits to social scientists who want to understand data problems from a multidimensional perspective. SOM clustering is another "tool in the toolkit" for social science research, allowing the clustering of respondents by attitudes, but then understanding their groupings based on demographic indicators.

As the United States set a goal in April 2021 to halve the US greenhouse emissions by 2030, with a net zero emission goal by 2050 [48], this technique can also be highly valuable in policy research for revealing sub-populations. For example, the aberrant cluster members demonstrate unique positions on wind-power. Policymakers can target this cluster with information campaigns and improved community engagement - methods noted to be influential for wind turbine acceptance [49–51]. Clustering demographics can also be used for public relations targeting. For instance, the aberrant group is primarily female, Hispanic, and lower income – essential demographics to assist policymakers when constructing targeted educational campaigns. Likewise, universalists have tied their energy preferences to economics, suggesting that policies for renewable energy may be accepted and preferred by this subgroup, should the personal cost to them be minimized. Furthermore, identifying clusters of individuals on energy preferences can engage stakeholders in mobilizing resources to reach groups who may be "moveable" on issues of renewable energy sources through public relations campaigns. These are just a few examples of how SOM clustering can not only reveal important dividing lines amongst the public but provide another tool to impact change and work towards social acceptable energy policies. For these reasons, we strongly advocate for broader use of machine learning in survey data-driven social science research.

Author contribution statement

Heather Bedle, Christopher R.H. Garneau: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Alexandro Vera-Arroyo: Contributed reagents, materials, analysis tools or data; Wrote the paper.

Data availability statement

Data included in article/supplementary material/referenced in article.

Additional information

No additional information is available for this paper.

Code availability

Kohonen SOM code is available at: https://cran.r-project.org/web/packages/kohonen/index.html.

Funding

This work was partially funded by the University of Oklahoma's Data Institute for Societal Challenges (DISC) and the AASPI Consortium at the University of Oklahoma for method development.

Declaration of competing interest

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

Acknowledgements

The authors would like to thank the Pew Research Foundation for making their survey data available.

References

[1] IPCC, in: P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, B. Zhou (Eds.), Climate Change 2021: the Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., Cambridge University Press, 2021.
[2] S. Batel, Research on the social acceptance of renewable energy technologies: past, present and future, Energy Res. Social Sci. 68 (2020), 101544.
[3] S. Ansolabehere, D.M. Konisky, Cheap and Clean: How Americans Think about Energy in the Age of Global Warming, The MIT Press, 2016.
[4] L.C. Hamilton, E. Bell, J. Hartter, J.D. Salerno, A Change in the wind? US public views on renewable energy and climate compared, Energy, Sustainability Soc. 8 (2018) 11.
[5] L.C. Stokes, H.L. Breetz, Politics in the US energy transition: case studies of solar, wind, biofuels and electric vehicles policy, Energy Pol. 113 (2018) 76–86.
[6] L.C. Stokes, C. Warshaw, Renewable energy policy design and framing influence public support in the United States, Nat. Energy 2 (2017) 1–6.
[7] S.A. Gustafson, M. Rosenthal, T. Ballew, M.H. Goldberg, P. Bergquist, J.E. Kotcher, E.W. Maibach, A. Leiserowitz, The development of Partisan polarization over the green new deal, Nat. Clim. Change 9 (11) (2019) 940–944.
[8] J. DeCicco, T. Yan, F. Keusch, D.H. Muñoz, L. Neidert, US consumer attitudes and expectations about energy, Energy Pol. 86 (2015) 749–758.
[9] S. Bigerna, C.A. Bollino, S. Micheli, P. Polinori, A new unified approach to evaluate economic acceptance towards main green technologies using the meta-analysis, J. Clean. Prod. 167 (2017) 1251–1262.
[10] A. Buylova, B.S. Steel, C.A. Simon, Public perceptions of energy scarcity and support for new energy technologies: a western US case study, Energies 13 (1) (2020) 238.
[11] B. Sütterlin, M. Siegrist, Public acceptance of renewable energy technologies from an abstract versus concrete perspective and the positive imagery of solar power, Energy Pol. 106 (2017) 356–366.
[12] C.A. Simon, M.C. Moltz, Going native? Examining nativity and public opinion of environment, alternative energy, and science policy expenditures in the United States, Energy Res. Social Sci. 46 (2018) 296–302.
[13] A. Leiserowitz, Climate change risk perception and policy preferences: the role of affect, imagery, and values, Climatic Change 77 (1) (2006) 45–72.
[14] C. Fraune, M. Knodt, Sustainable energy transformations in an age of populism, post-truth politics, and local resistance, Energy Res. Social Sci. 43 (2018) 1–7.
[15] A.M. McCright, R.E. Dunlap, C. Xiao, Perceived scientific agreement and support for government action on climate change in the USA, Climatic Change 119 (2) (2013) 511–518.
[16] M. Mildenberger, J.R. Marlon, P.D. Howe, A. Leiserowitz, The spatial distribution of Republican and Democratic climate opinions at state and local scales, Climatic Change 145 (3) (2017) 539–548.
[17] C.E. Clarke, P.S. Hart, J.P. Schuldt, D.T. Evensen, H.S. Boudet, J.B. Jacquet, R.C. Stedman, Public opinion on energy development: the interplay of issue framing, top-of-mind associations, and political ideology, Energy Pol. 81 (2015) 131–140.
[18] T.R. Peterson, J.C. Stephens, E.J. Wilson, Public perception of and engagement with emerging low-carbon energy technologies: a literature review, MRS Energy Sustain 2 (2015).
[19] J.C. Besley, The state of public opinion research on attitudes and understanding of science and technology, Bull. Sci. Technol. Soc. 33 (1–2) (2013) 12–20.
[20] Pew Research Center, Gen Z, Millennials Stand Out for Climate Change Activism, Social Media Engagement with Issue, May 2021. Washington, D.C.
[21] S.O. Hazboun, H.S. Boudet, Public preferences in a shifting energy future: Comparing public views of eight energy sources in North America's Pacific Northwest, Energies 13 (8) (2020) 1940.
[22] J.W. Stoutenborough, L. Shi, A. Vedlitz, Probing public perceptions on energy: support for a comparative, deep-probing survey design for complex issue domains, Energy 81 (2015) 406–415.
[23] D. Bell, T. Gray, C. Haggett, J. Swaffield, Re-visiting the 'social gap': public opinion and relations of power in the local politics of wind energy, Environ. Polit. 22 (1) (2013) 115–135.
[24] A. Leiserowitz, E. Maibach, C. Roser-Renouf, S. Rosenthal, M. Cutler, Politics and Global Warming, Spring 2016. Yale Program on Climate Change Communication, Yale University, New Haven, CT, USA, 2017.

[25] T. Kohonen, Self-organizing Maps, vol. 30, Springer Science & Business Media, 2012.
[26] K. Taczanowska, L.M. González, X. García-Massó, A. Zięba, C. Brandenburg, A. Muhar, J.L. Toca-Herrera, Nature-based tourism or mass tourism in nature? Segmentation of mountain protected area visitors using self-organizing maps (SOM), Sustainability 11 (5) (2019) 1314.
[27] C. Peiró-Velert, A. Valencia-Peris, L.M. González, X. García-Massó, P. Serra-Añó, J. Devís-Devís, Screen media usage, sleep time and academic performance in adolescents: clustering a self-organizing maps analysis, PLoS One 9 (6) (2014), e99478.
[28] M.À. Pans Sancho, L. Millán González, J. Úbeda Colomer, J. Devís Devís, Screen time among Spanish university students with disabilities: a self-organizing maps analysis, BMC Publ. Health 1 (2019) 19.
[29] N.B. Ahmad, U.F. Alias, N. Mohamad, N. Yusof, Principal component analysis and self-organizing map clustering for student browsing behaviour analysis, Procedia Computer Science 163 (2019) 550–559.
[30] J. Molina-García, X. García-Massó, I. Estevan, A. Queralt, Built environment, psychosocial factors and active commuting to school in adolescents: clustering a self-organizing map analysis, Int. J. Environ. Res. Publ. Health 16 (1) (2019) 83.
[31] J.W. Sugg, Exploratory geovisualization of the character and distribution of American climate change beliefs, Weather, Climate, and Society 13 (1) (2021) 67–82.
[32] T. Kohonen, Self-Organizing Maps, Springer, 1995.
[33] R. Wehrens, L.M. Buydens, Self-and super-organizing maps in R: the Kohonen package, J. Stat. Software 21 (2007) 1–19.
[34] R. Wehrens, J. Kruisselbrink, Flexible self-organizing maps in Kohonen 3.0, J. Stat. Software 87 (2018) 1–18.
[35] J. Crowe, Explaining popular support for wind energy in the United States, Journal of Rural Social Sciences 35 (2) (2020) 2.
[36] A.R. Pearson, J.P. Schuldt, R. Romero-Canyas, M.T. Ballew, D. Larson-Konar, Diverse segments of the US public underestimate the environmental concerns of minority and low-income Americans, Proc. Natl. Acad. Sci. USA 115 (49) (2018) 12429–12434.
[37] A. Chapman, Y. Shigetomi, S.C. Karmaker, B.B. Saha, K. Huff, C. Brooks, J. Stubbins, The cultural dynamics of energy: the impact of lived experience, preference and demographics on future energy policy in the United States, Energy Res. Social Sci. 80 (2021), 102231.
[38] A. Chapman, Y. Shigetomi, S.C. Karmaker, B. Saha, C. Brooks, Cultural and demographic energy system awareness and preference: implications for future energy system design in the United States, Energy Econ. 112 (2022), 106141.
[39] M. Shafiullah, M.D. Miah, M.S. Alam, M. Atif, Does economic policy uncertainty affect renewable energy consumption? Renew. Energy 179 (2021) 1500–1521.
[40] Q. Wang, M.P. Kwan, J. Fan, J. Lin, Racial disparities in energy poverty in the United States, Renew. Sustain. Energy Rev. 137 (2021), 110620.
[41] D.W. Hine, J.P. Reser, M. Morrison, W.J. Phillips, P. Nunn, R. Cooksey, Audience segmentation and climate change communication: Conceptual and methodological considerations, Wiley Interdisciplinary Reviews: Clim. Change 5 (4) (2014) 441–459.
[42] D. Bidwell, The role of values in public beliefs and attitudes towards commercial wind energy, Energy Pol. 58 (2013) 189–199.
[43] J. Firestone, A. Bates, L.A. Knapp, See me, Feel me, Touch me, Heal me: wind turbines, culture, landscapes, and sound impressions, Land Use Pol. 46 (2015) 241–249.
[44] J.B. Jacquet, R.C. Stedman, Perceived impacts from wind farm and natural gas development in northern Pennsylvania, Rural Sociol. 78 (4) (2013) 450–472.
[45] M. Molina, F. Garip, Machine learning for sociology, Annu. Rev. Sociol. 45 (2019) 27–45.
[46] J. Grimmer, M.E. Roberts, B.M. Stewart, Machine learning for social science: an agnostic approach, Annu. Rev. Polit. Sci. 24 (2021) 395–419.
[47] J. Zhang, W. Wang, F. Xia, Y.R. Lin, H. Tong, Data-driven computational social science: a survey, Big Data Research 21 (2020), 100145.
[48] Federal Administration of the USA, FACT SHEET: President Biden Sets 2030 Greenhouse Gas Pollution Reduction Target Aimed at Creating Good-Paying Union Jobs and Securing U.S. Leadership on Clean Energy Technologies, 2021. https://www.whitehouse.gov/briefing-room/statements-releases/2021/04/22/fact-sheet-president-biden-sets-2030-greenhouse-gas-pollution-reduction-target-aimed-at-creating-good-paying-union-jobs-and-securing-u-s-leadership-on-clean-energy-technologies/.
[49] M. Aitken, C. Haggett, D. Rudolph, Practices and rationales of community engagement with wind farms: awareness raising, consultation, empowerment, Plann. Theor. Pract. 17 (4) (2016) 557–576.
[50] J. Firestone, A.W. Bates, A. Prefer, Power transmission: where the offshore wind energy comes home, Environ. Innov. Soc. Transit. 29 (2018) 90–99, 2018.
[51] B. Hoen, J. Firestone, J. Rand, D. Elliot, G. Hübner, J. Pohl, K. Kaliski, Attitudes of US wind turbine neighbors: analysis of a nationwide survey, Energy Pol. 134 (2019), 110981.