# Modeling Complex Systems: A Case-Based Approach

Brian Castellani, Ph.D.
Professor of Sociology
Kent State University, USA

- Over the past several years we have developed a case-based, mixed-methods, density approach to modeling the temporal and spatial complexities of big data.

- The platform for this approach is called the SACS Toolkit. In terms of simplifying assumptions, the Toolkit employs three novel solutions:
  - (1) it conceptualizes the complex causal organization of a system as a set of microscopic cases (k-dimensional vectors spaces);
  - (2) it clusters/groups cases to identify major and minor profiles and (discrete or continuous) trajectories
  - (3) it translates their high-dynamic microscopic trajectories into the movement of macroscopic, low-dynamic densities.

- The strengths of this approach are several.  It allows researchers to:
  - Model complex systems as sets of cases.
  - Explore these systems at multiple levels.
  - Examine the interactions between system and environment.
  - Explore the relationships amongst the cases (networks).
  - Address and combine both structure (organizational pattern) and agency.
  - Study complex causal structure.
  - Use small to big data.
  - Model these systems as static or longitudinal.
    - In terms of longitudinal, we can model as discrete or continuous
  - In terms of continuous modeling, we can:
    - map the complex, nonlinear evolution of ensembles (or densities) of cases;
    - classify major and minor clusters and time-trends;
    - visually identify dynamical states, such as saddles and attractor points;
    - plot the speed of cases along different states;
    - detect the non-equilibrium clustering of case trajectories during key transient times;
    - construct multiple models to fit novel data;
    - predict future time-trends and dynamical states; and, finally, in terms of impact,
    - generate results that are visually and conceptually intuitive to private/public sector users and policy makers.

# Cases Are Complex Systems

- Researchers in the social sciences currently employ a variety of mathematical/ computational models for studying complex systems.
- Despite the diversity of these models, the majority can be grouped into one of four types:
  - equation-based modeling,
  - stochastic (statistical) modeling,
  - computational modeling
  - network modeling.
- However, ***David Byrne*** and colleagues have added a fifth type: ***case-based modeling***

# Cases Are Complex Systems

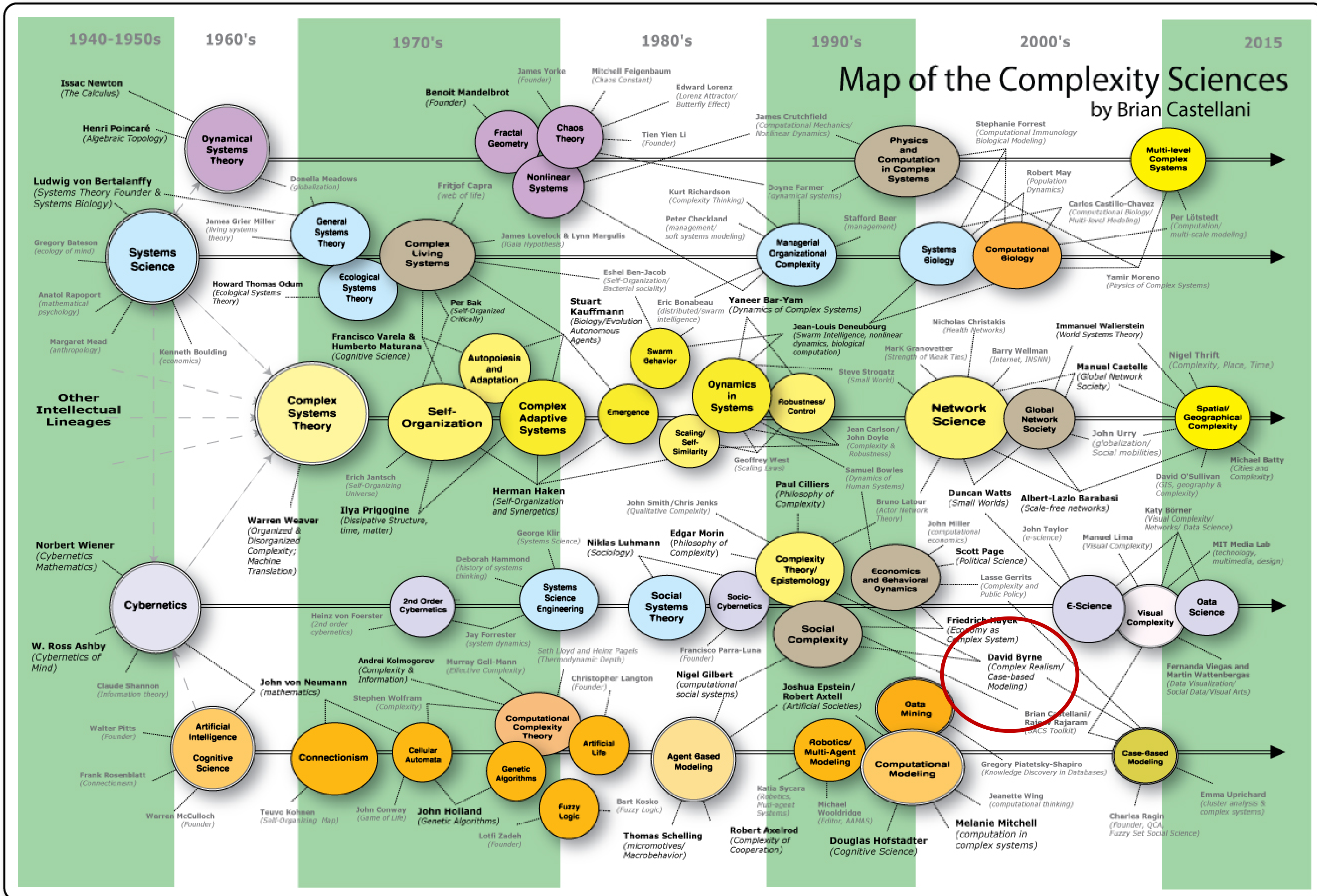
- ***Byrne*** is recognized, internationally, as a leading figure in what most scholars see as two highly promising but distinct fields of study:
  - (1) case-based method and
  - (2) the sociological study of complex systems.
  - An example of the former is Byrne's ***Sage Handbook of Case-Based Methods*** – which he co-edited with Charles Ragin, the creator of Qualitative Comparative Analysis.
  - An example of the latter is his widely read ***Complexity Theory and the Social Sciences*** – which Callaghan and he just significantly updated in 2013.

# Cases Are Complex Systems

- What scholars (including the current authors) are only beginning to grasp, however, is that Byrne sees these areas as conditional upon one another – that is, they are two sides of the same theoretical/methodological coin:

- His premise, while simple enough, is ground-breaking:
  - ***Cases are the methodological equivalent of complex systems; or, alternatively, complex systems are, theoretically speaking, cases and therefore should be studied as such.***

- With this premise – Byrne introduces an entirely new approach for modeling social complexity and the temporal and spatial dynamics of complex systems.



1940-1950s
1960's
1970's
1980's
1990's
2000's
2015
Map of the Complexity Sciences
by Brian Castellani
Issac Newton
(The Calculus)
Henri Poincaré
(Algebraic Topology)
Dynamical Systems Theory
Ludwig von Bertalanffy
(Systems Theory Founder & Systems Biology)
Gregory Bateson
(ecology of mind)
Systems Science
Anatol Rapoport
(mathematical psychology)
Margaret Mead
(anthropology)
Kenneth Boulding
(economics)
Other Intellectual Lineages
Norbert Wiener
(Cybernetics Mathematics)
Cybernetics
W. Ross Ashby
(Cybernetics of Mind)
Claude Shannon
(Information theory)
Walter Pitts
(Founder)
Artificial Intelligence
Cognitive Science
Frank Rosenblatt
(Connectionism)
Warren McCulloch
(Founder)
Donella Meadows
(globalization)
James Grier Miller
(living systems theory)
General Systems Theory
Ecological Systems Theory
Howard Thomas Odum
(Ecological Systems Theory)
Complex Systems Theory
Warren Weaver
(Organized & Disorganized Complexity; Machine Translation)
John von Neumann
(mathematics)
Connectionism
Teuvo Kohnen
(Self-Organizing Map)
James Yorke
(Founder)
Mitchell Feigenbaum
(Chaos Constant)
Edward Lorenz
(Lorenz Attractor/ Butterfly Effect)
Benoit Mandelbrot
(Founder)
Fractal Geometry
Chaos Theory
Tien Yien Li
(Founder)
Nonlinear Systems
Fritjof Capra
(web of life)
Complex Living Systems
James Lovelock & Lynn Margulis
(Gaia Hypothesis)
Eshel Ben-Jacob
(Self-Organization/ Bacterial sociality)
Per Bak
(Self-Organized Criticality)
Francisco Varela & Humberto Maturana
(Cognitive Science)
Autopoiesis and Adaptation
Self-Organization
Complex Adaptive Systems
Erich Jantsch
(Self-Organizing Universe)
Ilya Prigogine
(Dissipative Structure, time, matter)
Herman Haken
(Self-Organization and Synergetics)
Deborah Hammond
(history of systems thinking)
George Klir
(Systems Science)
2nd Order Cybernetics
Heinz von Foerster
(2nd order cybernetics)
Systems Science Engineering
Jay Forrester
(system dynamics)
Andrei Kolmogorov
(Complexity & Information)
Murray Gell-Mann
(Effective Complexity)
Seth Lloyd and Heinz Pagels
(Thermodynamic Depth)
Stephen Wolfram
(Complexity)
Computational Complexity Theory
Cellular Automata
Genetic Algorithms
Artificial Life
Fuzzy Logic
John Conway
(Game of Life)
John Holland
(Genetic Algorithms)
Lotfi Zadeh
(Founder)
Christopher Langton
(Founder)
Bart Kosko
(Fuzzy Logic)
Stuart Kauffmann
(Biology/Evolution Autonomous Agents)
Swarm Behavior
Emergence
Eric Bonabeau
(distributed/swarm intelligence)
Yaneer Bar-Yam
(Dynamics of Complex Systems)
Dynamics in Systems
Robustness/ Control
Scaling/ Self-Similarity
Geoffrey West
(Scaling Laws)
Jean-Louis Deneubourg
(Swarm Intelligence, nonlinear dynamics, biological computation)
Steve Strogatz
(Small World)
Jean Carlson/ John Doyle
(Complexity & Robustness)
Samuel Bowles
(Dynamics of Human Systems)
John Smith/Chris Jenks
(Qualitative Complexity)
Niklas Luhmann
(Sociology)
Edgar Morin
(Philosophy of Complexity)
Social Systems Theory
Socio-Cybernetics
Francisco Parra-Luna
(Founder)
Nigel Gilbert
(computational social systems)
Agent Based Modeling
Thomas Schelling
(micromotives/ Macrobehavior)
Robert Axelrod
(Complexity of Cooperation)
James Crutchfield
(Computational Mechanics/ Nonlinear Dynamics)
Physics and Computation in Complex Systems
Doyne Farmer
(dynamical systems)
Kurt Richardson
(Complexity Thinking)
Peter Checkland
(management/ soft systems modeling)
Managerial Organizational Complexity
Stafford Beer
(management)
Systems Biology
Paul Cilliers
(Philosophy of Complexity)
Complexity Theory/ Epistemology
Economics and Behavioral Dynamics
Bruno Latour
(Actor Network Theory)
Social Complexity
Joshua Epstein/ Robert Axtell
(Artificial Societies)
Data Mining
Robotics/ Multi-Agent Modeling
Computational Modeling
Katia Sycara
(Robotics, Multi-agent Systems)
Michael Wooldridge
(Editor, AAMAS)
Douglas Hofstadter
(Cognitive Science)
Melanie Mitchell
(computation in complex systems)
Jeanette Wing
(computational thinking)
Gregory Piatetsky-Shapiro
(Knowledge Discovery in Databases)
Stephanie Forrest
(Computational Immunology Biological Modeling)
Robert May
(Population Dynamics)
Carlos Castillo-Chavez
(Computational Biology/ Multi-level Modeling)
Computational Biology
Multi-level Complex Systems
Per Lötstedt
(Computation/ multi-scale modeling)
Yamir Moreno
(Physics of Complex Systems)
Nicholas Christakis
(Health Networks)
Mark Granovetter
(Strength of Weak Ties)
Barry Wellman
(Internet, INSNA)
Immanuel Wallerstein
(World Systems Theory)
Manuel Castells
(Global Network Society)
Network Science
Global Network Society
Nigel Thrift
(Complexity, Place, Time)
Spatial/ Geographical Complexity
John Urry
(globalization/ Social mobilities)
Michael Batty
(Cities and Complexity)
David O'Sullivan
(GIS, geography & Complexity)
Duncan Watts
(Small Worlds)
Albert-Lazlo Barabasi
(Scale-free networks)
Katy Börner
(Visual Complexity/ Networks/ Data Science)
John Miller
(computational economics)
John Taylor
(e-science)
Manuel Lima
(Visual Complexity)
MIT Media Lab
(technology, multimedia, design)
Scott Page
(Political Science)
Lasse Gerrits
(Complexity and Public Policy)
E-Science
Visual Complexity
Data Science
Friedrich Hayek
(Economy as Complex System)
David Byrne
(Complex Realism/ Case-based Modeling)
Brian Castellani/ Rajeev Rajaram
(SACS Toolkit)
Case-Based Modeling
Fernanda Viegas and Martin Wattenbergas
(Data Visualization/ Social Data/Visual Arts)
Emma Uprichard
(cluster analysis & complex systems)
Charles Ragin
(Founder, QCA, Fuzzy Set Social Science)

# Cases Are Complex Systems

- There are several strengths to this approach, **three** of which are crucial to the work Dr. Rajaram and I are doing:
  1. It embraces an interdisciplinary framework –with great thought given to the transport of theories, concepts, and methods between scientific and disciplinary boundaries, for the purposes of modeling social complexity and complex social systems.
  2. It employs a mixed-methods toolkit, including case-comparative analysis and many of the latest advances in computational and complexity science method.
  3. It provides an epistemological platform (grounded in complex realism) for constructing a cohesive ‘complex systems’ methodology, based on its concept of the case.

# Cases Are Complex Systems

- Pace Byrne, we seek to develop a mathematically-rigorous, computationally-based, mixed-methods platform for modeling social complexity and complex social systems.
- The purpose of this presentation (in combination with that of Dr. Rajaram) is to explore what we have so far accomplished – albeit tentatively.

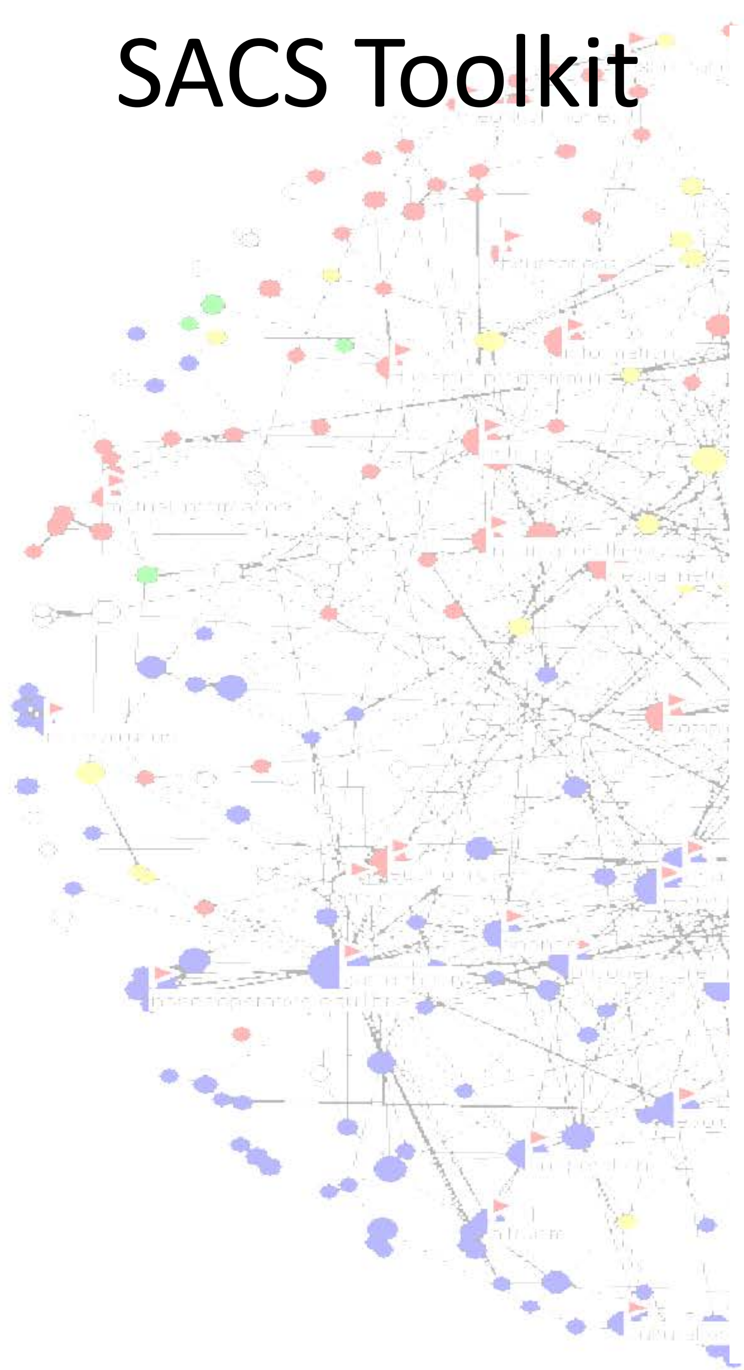
# Cases Are Complex Systems

- To begin, we have introduced two new terms:
  - ***case-based complexity science*** is the attempt to actively integrate case-based method with the latest developments in the complexity and social sciences for the purpose of modeling complex social systems as sets of cases.
    - It also revolves around a particular set of epistemological assumptions:
    - Complexity theory is not so much a substantive theory, as much as it is an epistemologically explicit attempt to model social life in complex systems terms.
    - It also revolves around complex realism
  - In turn, ***case-based modeling*** is the mixed-methods set of techniques scholars use to engage in case-based complexity science, particularly the latest developments in the computational and complexity sciences.
    - The key to this approach is that the methods serve the purpose of case-comparative analysis, from small to big data!

# Cases Are Complex Systems

- We also introduce a new methodological framework: the ***Sociology and Complexity Science (SACS) Toolkit***.
- ***The SACS Toolkit*** is a the case-based, mixed-methods, computationally-grounded platform for modeling socio-biological complexity and, more specifically, complex socio-biological systems.

# SACS Toolkit



Comput Math Organ Theory
DOI 10.1007/s10588-012-9114-1

## Case-based modeling and the SACS Toolkit: a mathematical outline

**Brian Castellani · Rajeev Rajaram**

© Springer Science+Business Media, LLC 2012

**Abstract** Researchers in the social sciences currently employ a variety of mathematical/computational models for studying complex systems. Despite the diversity of these models, the majority can be grouped into one of three types: agent (rule-based) modeling, dynamical (equation-based) modeling and statistical (aggregate-based) modeling. The purpose of the current paper is to offer a fourth type: case-based modeling. To do so, we review the SACS Toolkit: a new method for quantitatively modeling complex social systems, based on a case-based, computational approach to data analysis. The SACS Toolkit is comprised of three main components: a theoretical blueprint of the major components of a complex system (*social complexity theory*); a set of case-based instructions for modeling complex systems from the ground up (*assemblage*); and a recommended list of case-friendly computational modeling techniques (*case-based toolset*). Developed as a variation on Byrne (in Sage Handbook of Case-Based Methods, pp. 260–268, 2009), the SACS Toolkit models a complex system as a set of $k$-dimensional vectors (cases), which it compares and contrasts, and then condenses and clusters to create a low-dimensional model (map) of a complex system's structure and dynamics over time/space. The assembled nature of the SACS Toolkit is its primary strength. While grounded in a defined mathematical framework, the SACS Toolkit is methodologically open-ended and therefore adaptable and amenable, allowing researchers to employ and bring together a wide variety of modeling techniques. Researchers can even develop and modify the SACS Toolkit for their own purposes. The other strength of the SACS Toolkit, which makes it a very effective technique for modeling large databases, is its ability to compress data matrices while preserving the most important aspects of a complex system's structure and

B. Castellani (✉)
Dept. of Sociology, Kent State University, Ashtabula, OH 44004, USA
e-mail: bcastel3@kent.edu

R. Rajaram
Dept. of Mathematical Sciences, Kent State University, Ashtabula, OH 44004, USA

# SACS Toolkit

1. First, it is comprised of a theoretical blueprint for studying complex systems called it social complexity theory. Social complexity theory is not a substantive theory; instead, it is a theoretical framework comprised of a series of key concepts necessary for modeling complex systems. These concepts include field of relations, network of attracting clusters, environmental forces, negotiated ordering, social practices, and so forth. Together, these concepts provide the vocabulary necessary for modeling a complex system.

2. Second, it is comprised of a set of case-based instructions for modeling complex systems from the ground up called it assemblage. Regardless of the methods or techniques used, assemblage guides researchers through a seven-step process of model buildingwhich we review belowstarting with how to frame ones topic in complex systems terms, moving on to building the initial model, then on to assembling the working model and its various maps to finally ending with the completed model.

3. Third, it is comprised of a recommend list of case-friendly modeling techniques called the *case-based toolset*. The case-based toolset capitalizes on the strengths of a wide list of techniques, using them in service of modeling complex systems as a set of cases. Our own repertoire of techniques include k-means cluster analysis, the self-organizing map neural net, Ragins QCA, network analysis, agent-based modeling, hierarchical regression, factor analysis, grounded theory method, and historical analysis.

# SACS Toolkit

We begin our review of the SACS Toolkit with five opening points:

(1) For the SACS Toolkit, case-based modeling is the study of a complex system $S$ as a set of cases $c_i$ such that:

$$S = \{c_i : c_i \text{ is a case relevant to the system under study}\}. \tag{1}$$

(2) At minimum, $S$ is comprised of one case $c_i$.

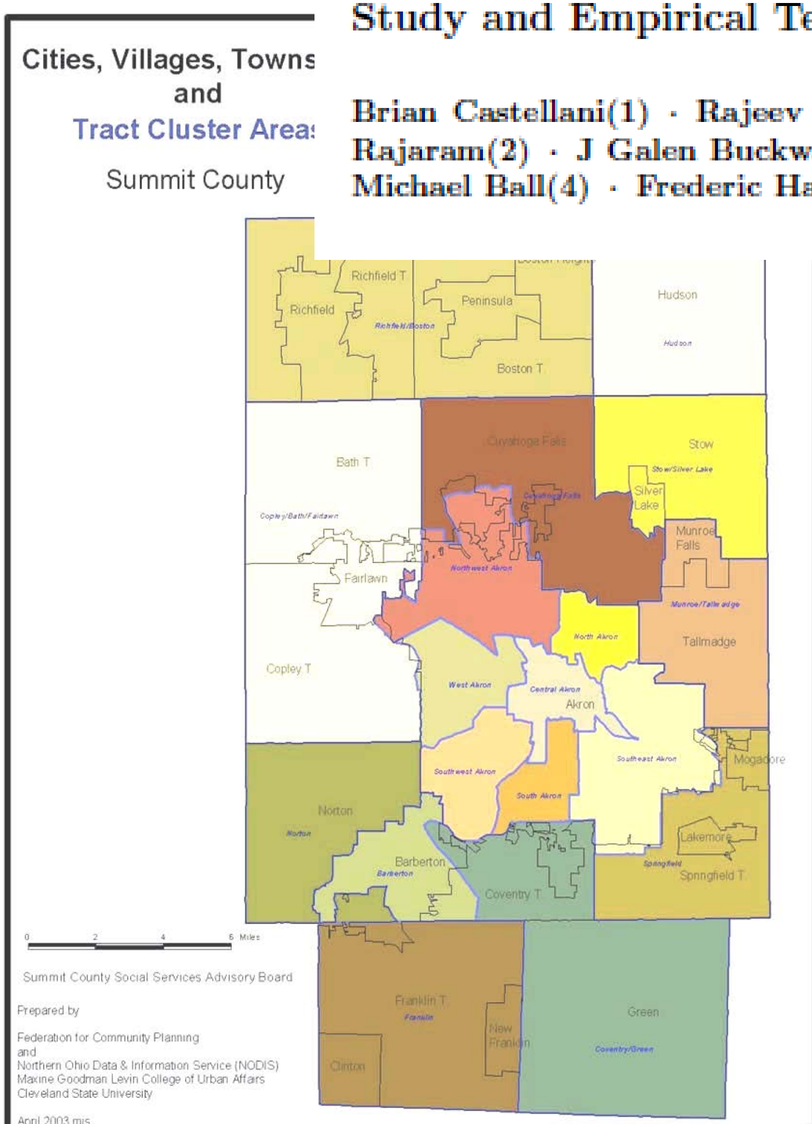
(3) While there is no predefined limit to the maximum number of cases that can be included in the study of a complex system, practically speaking the upper limit will be bounded, based on the particular set of cases identified for study—which is always an empirical issue.

(4) We denote the number of cases being studied by $n$.

(5) Each case $c_i$ in $S$ is a $k$ dimensional row vector $c_i = [x_{i1}, \ldots, x_{ik}]$, where each $x_{ij}$ represents a measurement on one of the variables being used to model a complex system.

# Place and Health as Complex Systems: A Case Study and Empirical Test

Brian Castellani(1) · Rajeev Rajaram(2) · J Galen Buckwalter(3) · Michael Ball(4) · Frederic Hafferty(5)



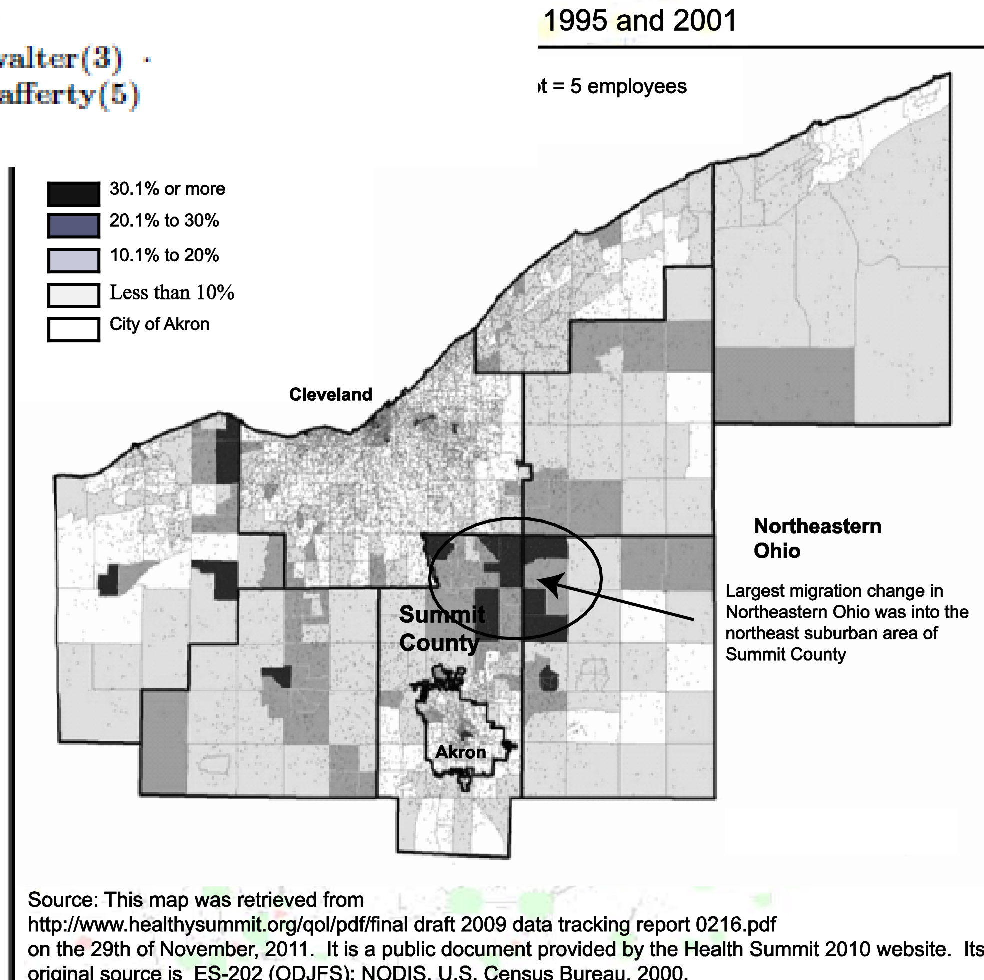
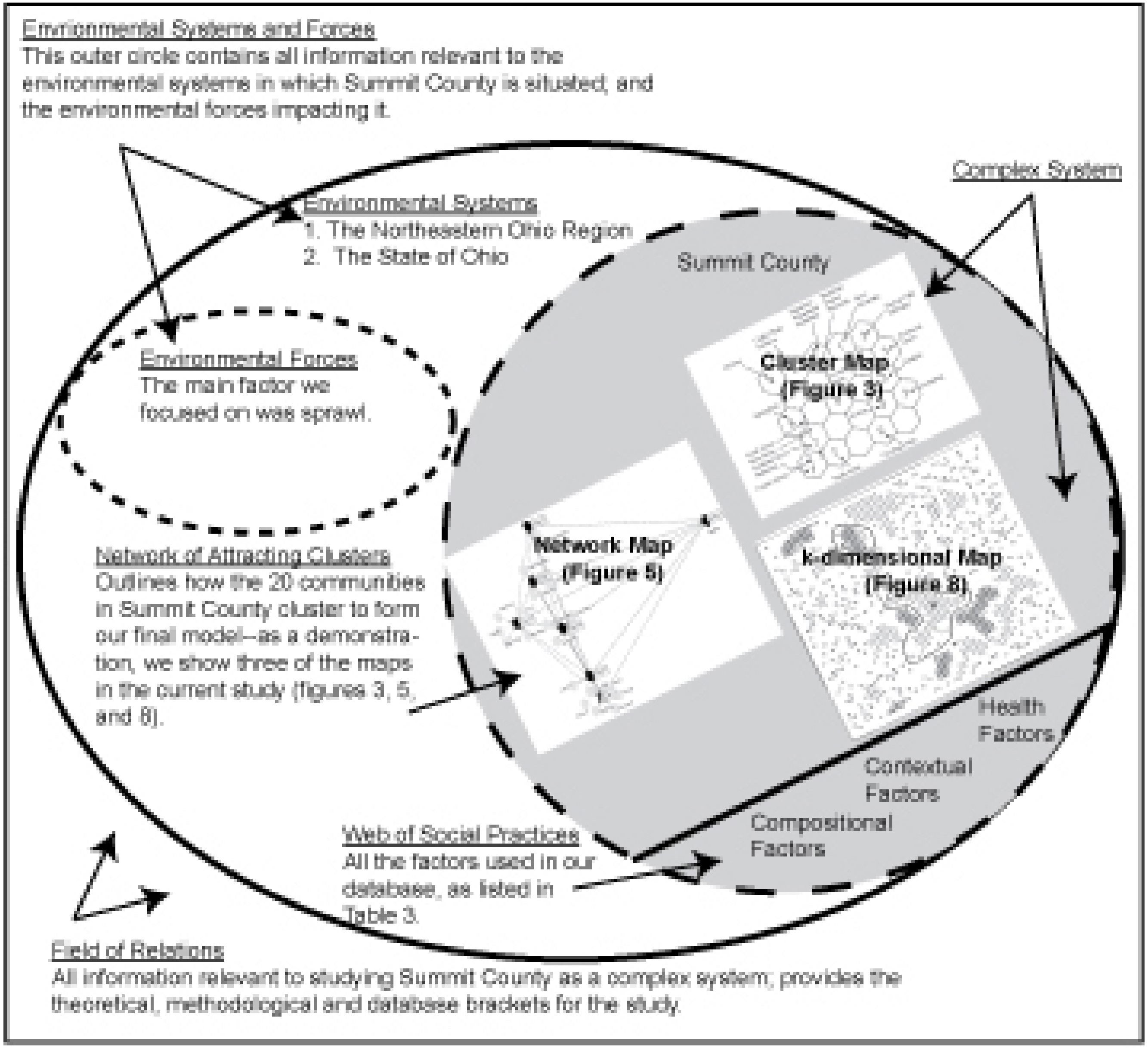



Figure 1
Example of the Final Map Created by the SACS Toolkit for Current Case Study



SACS Toolkit

TABLE 3
Variables Analyzed for the 20 Communities in the Summit County Database

| | |
|---|---|
| Compositional Factors | • Population 65 years of age of older[1]<br>• % White Population[1] (Defined as number of persons identifying themselves as "White" in response to the 1990 US Census or "White Alone" in response to the 2000 US Census)<br>• % African-American Population[1] (Defined as the number of persons identifying themselves as "Black or African-American" in response to the 1990 US Census or "Black or African-American Alone" in response to the 2000 US Census)<br>• Median Household Income[1] |
| Contextual Factors | • Overall Poverty[1] (Defined as the number of persons living "below the poverty level" as defined by the U.S. Census)<br>• Public Assistance[1] (Defined as the number of households receive public assistance as defined by the U.S. Census)<br>• Persons 25+ Years with High School Diploma[1]<br>• Net Job Growth[3] (Defined as the number of jobs in 2000 minus the number of jobs in 1990.<br>• Unemployment Rate[1] (Defined as unemployed civilian labor force)<br>• Housing affordability[1] (Defined as the percentage of households where mortgage/rent is greater than 30% of the household income)<br>• No Health Care Coverage[4] (An estimate of the number of individuals with no health care coverage based upon a statewide survey (Behavior Risk Factor Surveillance Survey – Centers for Disease Control and Prevention) |
| Health Outcomes | • No First Trimester Prenatal Care[4] (Defined as the number of births occurring to mothers from 1995 to and including 1998 for which no prenatal care was received during the first three months of the pregnancy)<br>• Teen Birth Rate[4] (Defined as the number of births occurring between 1995-1998 to mothers 15 to and including 17 years of age)<br>• Childhood Immunization Rate[5] (Defined as the percentage of children with a complete immunization series 4:3:1 by their second birthday based on the kindergarten retrospective study)<br>• Child Abuse/Neglect[6] (Defined as the number of referrals resulting in assessment per 1,000 childre under 18 years of age)<br>• Elder Abuse/Neglect[7] (Defined as the number of referrals received by the Department of Jobs and Family Services for abuse, exploitation, or neglect)<br>• Years of Potential Life Lost per Death[5] (Defined as the sum of the differences between the age at death and the life expectancy at age of death for each death occurring between 1990-1998 due to all causes divided by the number of deaths due to all causes within the census tract cluster borders where those borders are defined by United States Census Bureau census tracts) |

Data Sources: (1) United States Census Bureau 1990 and 2000 Decennial Censuses; (2) Ohio Department of Education; (3) NODIS; (4) Akron City Health Department, Office of Epidemiology; (5) Ohio Department of Health; (6) Children's Services Board; (7) Summit County Department of Jobs and Family Service.

# SACS Toolkit

## FIGURE 1

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | **Income per person** | **1950** | **1951** | **1952** | **1953** | **1954** | **1955** |
| 2 | Abkhazia | | | | | | |
| 3 | Afghanistan | 757.3188 | 766.7522 | 779.4 | | | |
| 4 | Akrotiri and Dhekelia | | | | | | |
| 5 | Albania | 1532.354 | 1598.493 | 1601. | | | |
| 6 | Algeria | 2429.214 | 2397.531 | 2449. | | | |
| 7 | American Sa | | | | | | |
| 8 | Andorra | | | | | | |
| 9 | Angola | | | | | | |
| 10 | Anguilla | | | | | | |
| 11 | Antigua and | | | | | | |
| 12 | Argentina | | | | | | |
| 13 | Armenia | | | | | | |
| 14 | Aruba | | | | | | |
| 15 | Australia | | | | | | |
| 16 | Austria | | | | | | |
| 17 | Azerbaijan | | | | | | |
| 18 | Bahamas | | | | | | |
| 19 | Bahrain | | | | | | |
| 20 | Bangladesh | 673.3711 | 675.3403 | 684.2 | | | |
| 21 | Barbados | 3245.073 | | | | | |
| 22 | Belarus | 2340.52 | 2309.686 | 2415. | | | |
| 23 | Belgium | 7990.466 | 8393.416 | 8343. | | | |

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | **life expectancy at birth** | **1950** | **1951** | **1952** | **1953** | **1954** | **1955** |
| 2 | Abkhazia | | | | | | |
| 3 | Afghanistan | 26.674 | 26.932 | 27.448 | 27.964 | 28.48 | 28.996 |
| 4 | Akrotiri and Dhekelia | | | | | | |
| | | | | 6 | 55.471 | 56.184 | 57.012 |
| | | | | 8 | 43.081 | 43.493 | 43.914 |
| | | | | | | | |
| | | | | | | | |
| | | | | 4 | 30.201 | 30.599 | 30.997 |
| | | | | | | | |
| | | | | 4 | 58.779 | 59.271 | 59.759 |
| | | | | 2 | 62.855 | 63.331 | 63.749 |
| | | | | 2 | 63.024 | 63.446 | 63.866 |
| | | | | 1 | 60.98 | 61.873 | 62.687 |
| | | | | 2 | 69.7 | 69.85 | 70.17 |
| | | | | 8 | 67.29 | 67.32 | 67.6 |
| | | | | 4 | 58.166 | 58.576 | 58.985 |
| 18 | Bahamas | 59.179 | 59.396 | 59.824 | 60.242 | 60.649 | 61.047 |
| 19 | Bahrain | 41.164 | 41.583 | 42.459 | 43.406 | 44.425 | 45.515 |
| 20 | Bangladesh | 42.676 | 43.038 | 43.376 | 43.739 | 44.127 | 44.541 |
| 21 | Barbados | 56.124 | 56.4 | 56.95 | 57.491 | 58.023 | 58.547 |
| 22 | Belarus | 65.022 | 65.247 | 65.692 | 66.125 | 66.548 | 66.958 |
| 23 | Belgium | 66.35 | 66.8 | 68 | 68.37 | 68.63 | 68.58 |
| 24 | Belize | 54.806 | 55.036 | 55.644 | 56.197 | 56.745 | 57.289 |
| 25 | Benin | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |

**The Utility of Nonequilibrium Statistical Mechanics, Specifically Transport Theory, for Modeling Cohort Data**

RAJEEV RAJARAM AND BRIAN CASTELLANI

*Departments of Mathematical Sciences and Sociology, Kent State University, Ashtabula, Ohio 44004*

*Shown here is a sample of the data used for the study, which consisted of two variables (K=2) taken from the Gapminder Website Database; namely, per capita GDP (x1(t)) and Life Expectancy (x2(t)) for 156 countries over 63 years (t).

# SACS Toolkit

Because $S$ consists of $n$ cases $\{c_i\}_{i=1}^{n}$, and each case $c_i$ has a vector configuration of $k$ dimensions, it is natural to represent $S$, at least initially and at its most basic, in the form of a data matrix $D$ as follows:

$$D = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} x_{11} & \dots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{nk} \end{bmatrix}. \qquad (6)$$

In the notation above, the $n$ rows in $D$ represent the set of cases $\{c_i\}$ in $S$, and the $k$ columns represent the measurements on some finite partition $\bigcup_{i=1}^{p} O_i$ of $W_s$ and $E_s$ as defined in Eq. (5) that couple to form the vector configuration for each $c_i$.

Simplifying Assumptions

Clustering and grouping to search for major and minor configurations/ profiles and trajectories (discrete or continuous)

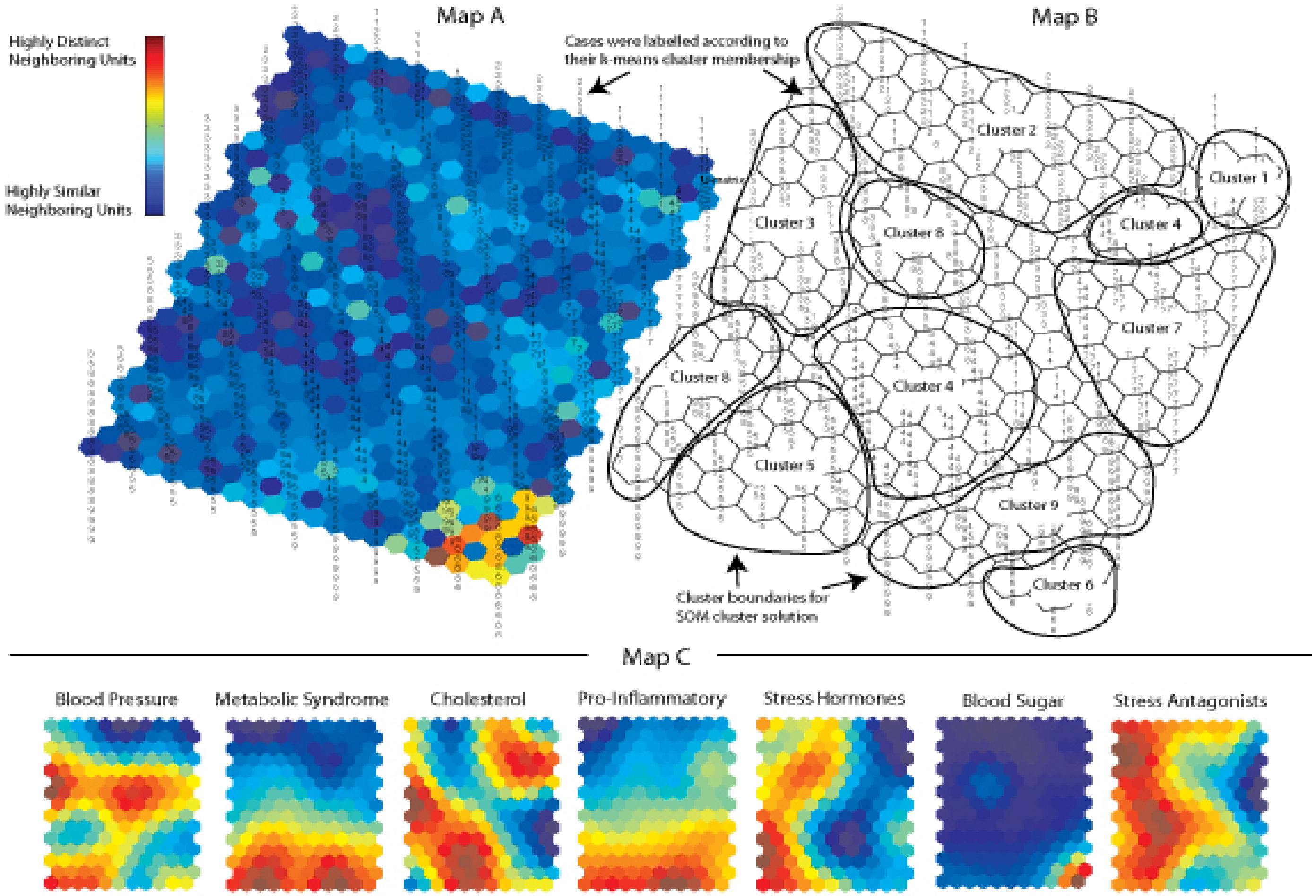



**Figure 1.**
*Map A* and *Map B* are graphic representations of the cluster solution arrived at by the Self-Organizing Map (SOM) Neural Net, referred to as the *U-Matrix*. In terms of the information they provide, *Map A* is a three-dimensional (topographical) u-matrix: for it, the SOM adds hexagons to the original 15X11 map to allow for visual inspection of the degree of similarity amongst neighboring map units; the dark blue areas indicate neighborhoods of cases that are highly similar; in turn, bright yellow and red areas, as in the lower right corner of the map, indicate highly defined cluster boundaries. *Map B* is a two-dimensional version of *Map A* that allows for visual inspection of how the SOM clustered the individual cases. Cases on this version of the u-matrix (as well as *Map A*) were labelled according to their k-means cluster membership (The 9 cluster solution showin Table 2) to see if the SOM would arrive at a similar solution. *Map C* is a graphic representation of the relative influence that the seven factors (shown in Table 1) had on the SOM cluster solution. The SOM generates a mini-map for the seven factors, each of which can be overlaid across maps A and B. Each of these mini-maps can then be inspected visually to examine what its rates are across the different neighborhoods (clusters of cases). Dark blue areas indicate the lowest rates for a factor; and the bright red areas indicate the highest rates for a factor. For example, looking at the mini-map for Factor 6 (*Blood Sugar*), its rates are extremely low across most of the map, except for the lower right corner, which is where (looking at *Map A* and *Map B*) the SOM placed Cluster 6.

TABLE 3
Final K-means Cluster Solution for 20 Communities in Summit County

| Variables (Unless otherwise noted, all data is from 1990—See Table 2) | Cluster | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| % Non-Hispanic Caucasian | 97.3* | 68.6 | 93.5 | 97.6 | 93.8 | 98.4 | 77.5 |
| % African-American | 1.7 | 28.0 | 5.6 | 1.0 | 4.7 | 1.0 | 21.2 |
| % Overall Poverty | 3.60 | 44.30 | 6.04 | 1.00 | 2.60 | 6.77 | 19.30 |
| 1990 household Income | 41464 | 11404 | 36021 | 68083 | 49144 | 30002 | 21688 |
| Job Growth (1993 to 2000) | 31.87 | 20.80 | 17.36 | 27.70 | 43.10 | 15.83 | .33 |
| % Civilian Labor Force (16+ old) | 96.17 | 85.90 | 95.22 | 96.60 | 95.70 | 94.73 | 90.82 |
| % Receiving Public Assistance | 2.8 | 25.8 | 4.3 | 1.4 | 2.6 | 5.6 | 13.8 |
| % No High School Degree (25yrs+) | 15.3 | 41.5 | 16.8 | 2.7 | 11.1 | 22.1 | 29.4 |
| % of households mortgage/rent is <30% of income | 16.0 | 43.4 | 17.6 | 15.8 | 19.0 | 18.1 | 27.4 |
| % Unemployed | 3.8 | 14.1 | 4.8 | 3.4 | 4.3 | 5.3 | 9.2 |
| % No 1st Trimester Care 1995-98 | 5.63 | 24.60 | 7.54 | 1.20 | 4.80 | 8.90 | 14.78 |
| Teen Pregnancies per 1000 births (1995-1998) | 5.80 | 66.00 | 12.54 | 1.30 | 3.50 | 12.33 | 47.72 |
| % children immunized by 2yrs of age | 74.1 | 40.0 | 76.5 | 86.1 | 72.9 | 78.1 | 60.7 |
| % No Health Care Coverage | 4.20 | 25.30 | 6.34 | 1.20 | 3.70 | 8.40 | 14.52 |
| Child Abuse/Neglect Rate per 1000 | 10.8 | 98.3 | 19.3 | 4.0 | 6.8 | 16.2 | 60.5 |
| Elder Abuse/Neglect Rate per 1000 | 4.1 | 53.8 | 4.9 | 2.1 | 4.8 | 9.1 | 9.3 |
| Years Lost per Death 1998 | 13.83 | 16.40 | 13.96 | 10.50 | 10.60 | 14.40 | 15.18 |

1. (*) The values listed in the columns for all 7 clusters represent the average value/measurement that the communities in that cluster scored for each variable listed in Column 1. In cluster analysis, these averages are called the cluster's centroids. 2. Community Membership for each of the 7 Clusters is as follows: Cluster 1: Stow/ Silverlake, Northfield/Macedonia/Sagamore, and Richfield/Peninsula; Cluster 2: Central Akron; Cluster 3: Twinsburg, Northwest Akron, Munroe Falls/Tallmadge, Norton and Franklin; Cluster 4: Hudson; Cluster 5: Copley/ Bath/Fairlawn; Cluster 6: Springfield, Coventry/Green and Cuyahoga Falls; Cluster 7: North, West, Southwest, South and Southeast Akron and Barberton City.

FIGURE 3: Final Cluster Centers Solution of 20 Communities in Summit County Using SOM

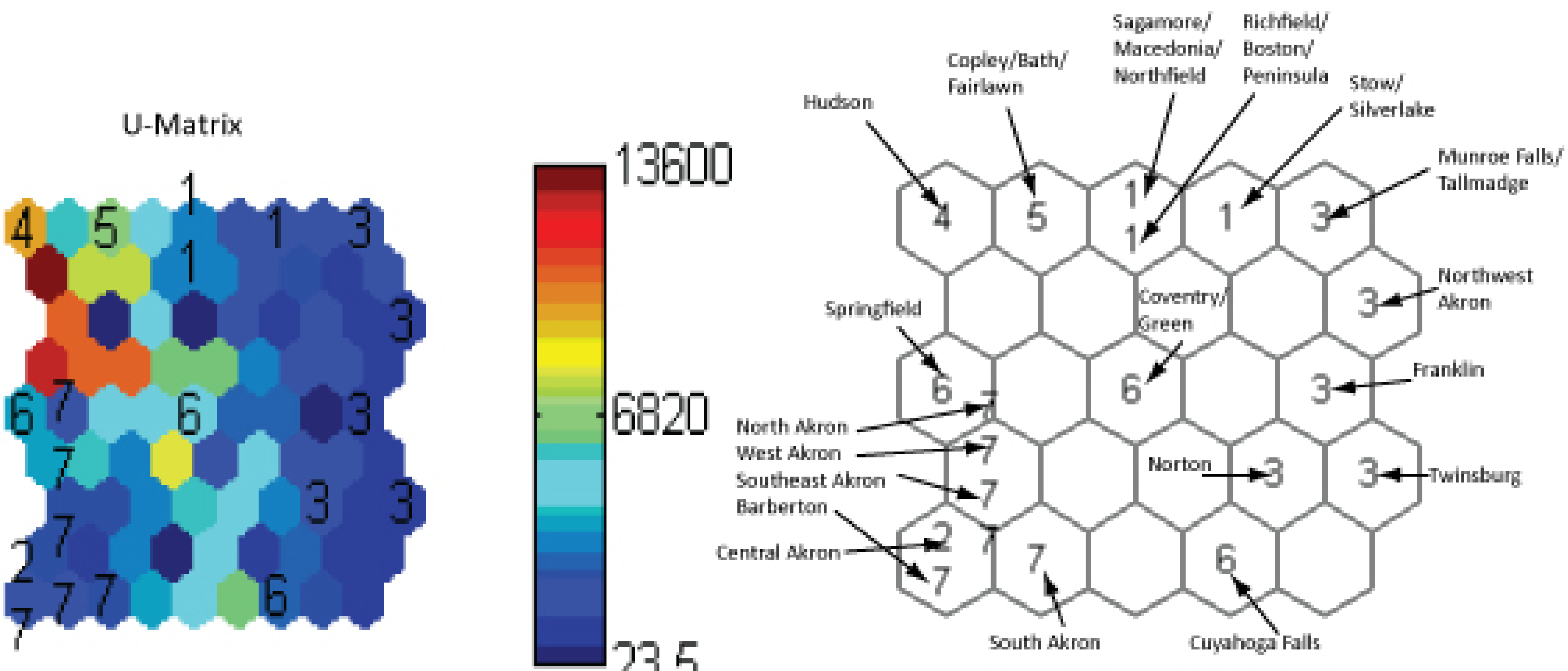



**Figure 3a** **Figure 3b**

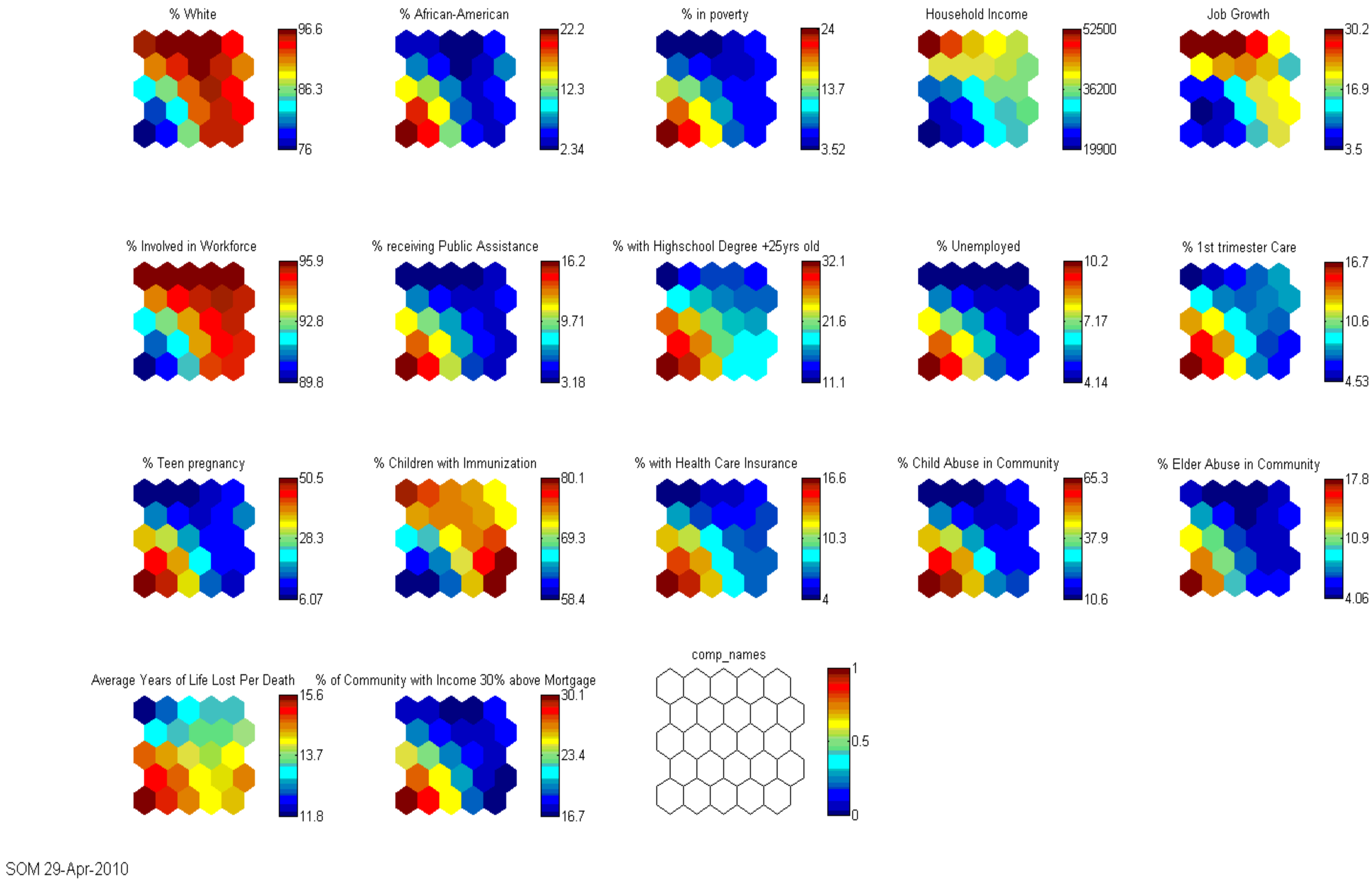
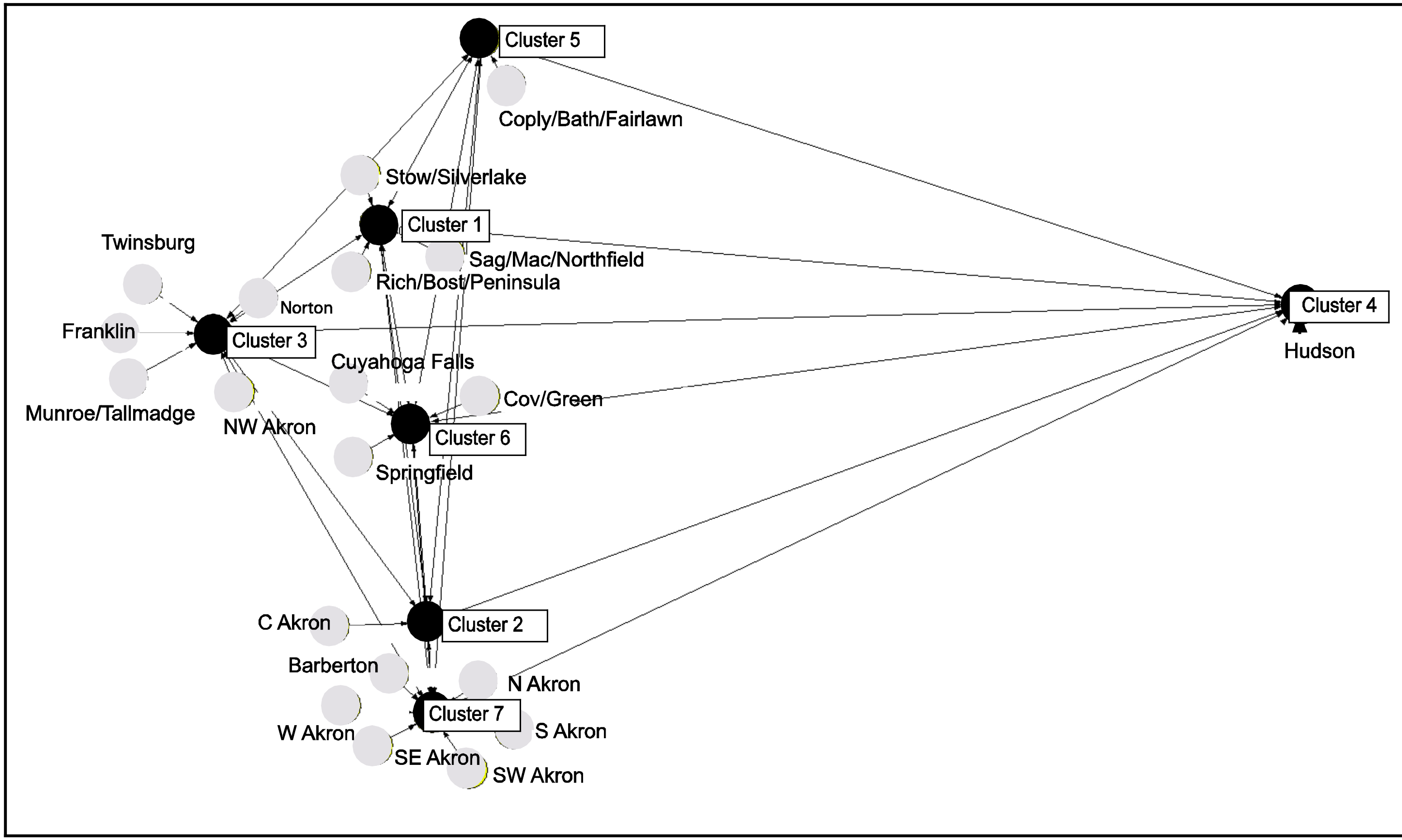

% White
96.6
86.3
76
% African-American
22.2
12.3
2.34
% in poverty
24
13.7
3.52
Household Income
52500
36200
19900
Job Growth
30.2
16.9
3.5
% Involved in Workforce
95.9
92.8
89.8
% receiving Public Assistance
16.2
9.71
3.18
% with Highschool Degree +25yrs old
32.1
21.6
11.1
% Unemployed
10.2
7.17
4.14
% 1st trimester Care
16.7
10.6
4.53
% Teen pregnancy
50.5
28.3
6.07
% Children with Immunization
80.1
69.3
58.4
% with Health Care Insurance
16.6
10.3
4
% Child Abuse in Community
65.3
37.9
10.6
% Elder Abuse in Community
17.8
10.9
4.06
Average Years of Life Lost Per Death
15.6
13.7
11.8
% of Community with Income 30% above Mortgage
30.1
23.4
16.7
comp_names
1
0.5
0
SOM 29-Apr-2010

Figure 4:
Network Map of the Seven Clusters in Summit County and their Respective Communities



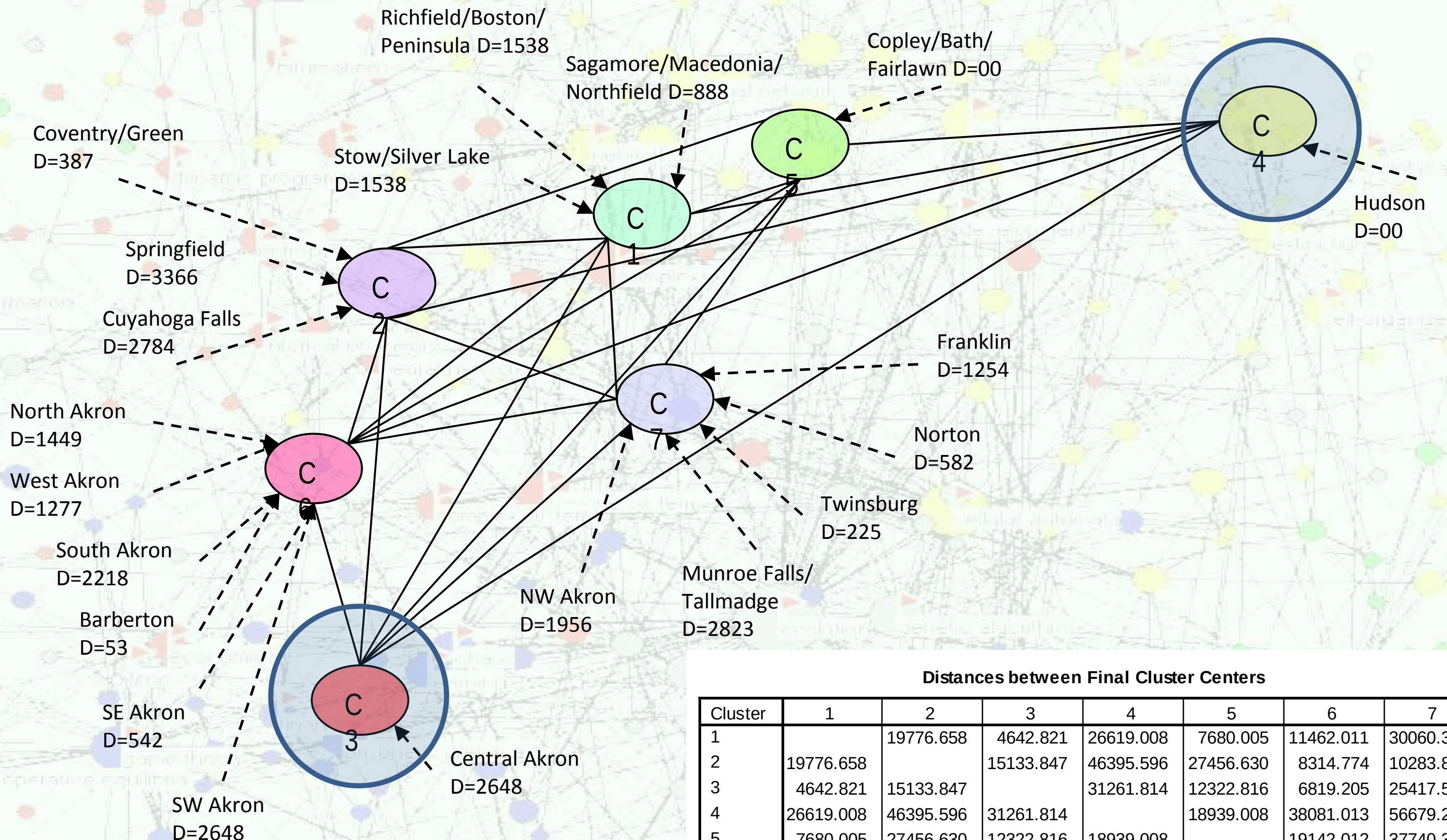

NOTE: Distances between clusters are based on Euclidian distances arrived at through k-means analysis. Distances within clusters for each community are based on within-cluster measures. All measures are non-standardized.

TABLE 6
Change in Final Cluster Solutions for 20 Communities in Summit County, 1990 to 2000

| COMMUNITY | YEAR<br>1990 Cluster Membership | <br>2000 Cluster Membership |
|---|---|---|
| **(Affluent Cluster) Hudson** | 4 | 4 |
| **(Affluent Cluster) Copley/Bath/Fairlawn** | 5 | 5 |
| **(Middle Class Cluster) Stow/Silverlake** | 1 | 1 |
| **Northfield/Macedonia/Sagamore** | 1 | 1 |
| **Richfield/Peninsula** | 1 | **5*** |
| **Twinsburg** | 3 | **1*** |
| **Northwest Akron** | 3 | 3 |
| **Munroe Falls/Tallmadge** | 3 | 3 |
| **Norton** | 3 | 6 |
| **Franklin** | 3 | 3 |
| **Springfield** | 6 | 6 |
| **Coventry/Green** | 6 | **3*** |
| **Cuyahoga Falls** | 6 | 6 |
| **(Poor Cluster) North Akron** | 7 | 7 |
| **West Akron** | 7 | 7 |
| **South Akron** | 7 | 7 |
| **Southwest Akron** | 7 | **2*** |
| **Southeast Akron** | 7 | 7 |
| **Barberton City** | 7 | 7 |
| **(Poorest Cluster) Central Akron** | 2 | 2 |

**1. (*)** The values listed in the columns for all 7 clusters represent the average value/measurement

# How did things change between 1990 and 2000?

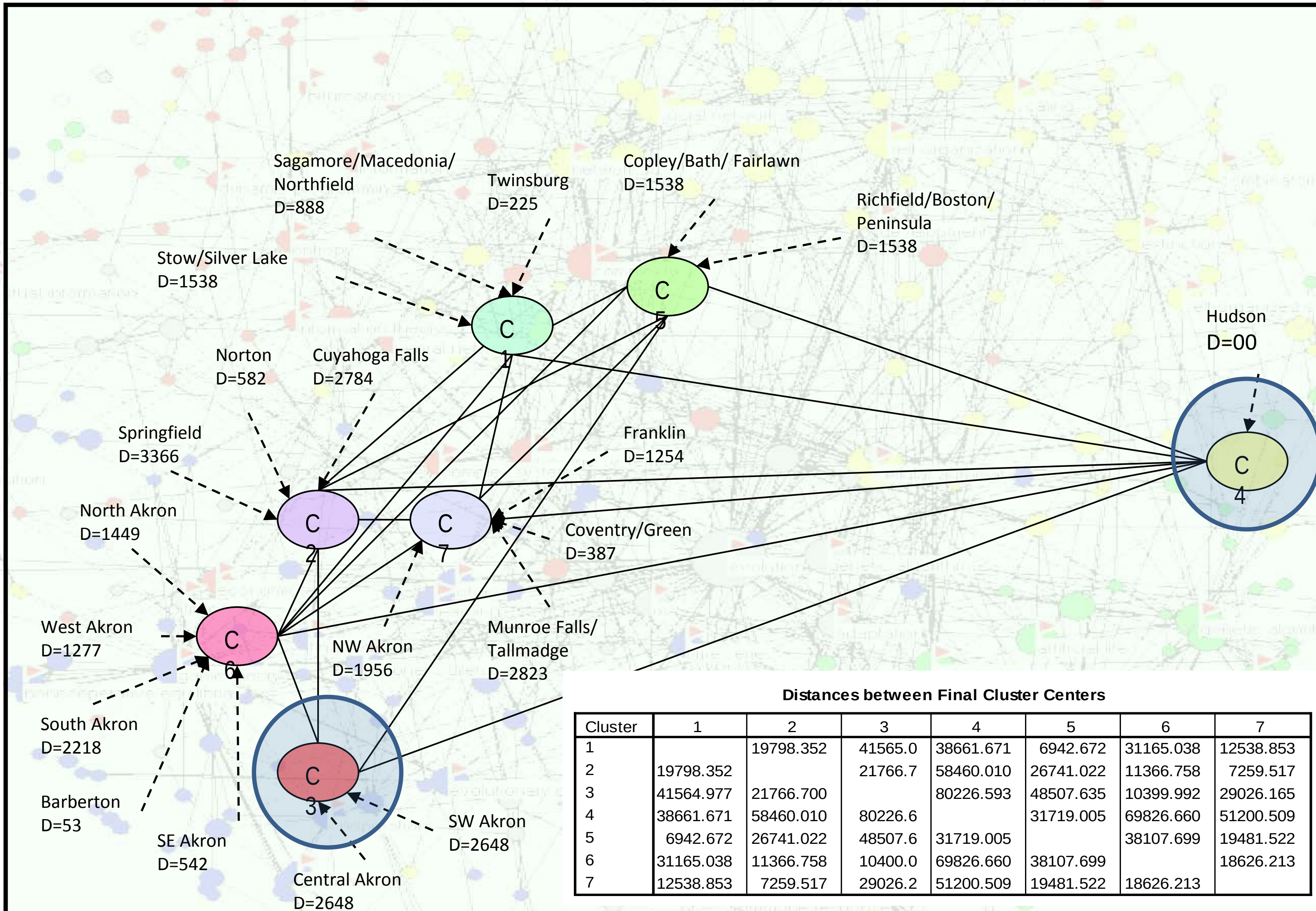



**Distances between Final Cluster Centers**

| Cluster | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | | 19776.658 | 4642.821 | 26619.008 | 7680.005 | 11462.011 | 30060.373 |
| 2 | 19776.658 | | 15133.847 | 46395.596 | 27456.630 | 8314.774 | 10283.802 |
| 3 | 4642.821 | 15133.847 | | 31261.814 | 12322.816 | 6819.205 | 25417.577 |
| 4 | 26619.008 | 46395.596 | 31261.814 | | 18939.008 | 38081.013 | 56679.240 |
| 5 | 7680.005 | 27456.630 | 12322.816 | 18939.008 | | 19142.012 | 37740.317 |
| 6 | 11462.011 | 8314.774 | 6819.205 | 38081.013 | 19142.012 | | 18598.509 |
| 7 | 30060.373 | 10283.802 | 25417.577 | 56679.240 | 37740.317 | 18598.509 | |

**Network of Attracting Clusters Yr = 1990** (Within and Between Euclidian Distance Measures)

# How did things change between 1990 and 2000?




**Distances between Final Cluster Centers**

| Cluster | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | | 19798.352 | 41565.0 | 38661.671 | 6942.672 | 31165.038 | 12538.853 |
| 2 | 19798.352 | | 21766.7 | 58460.010 | 26741.022 | 11366.758 | 7259.517 |
| 3 | 41564.977 | 21766.700 | | 80226.593 | 48507.635 | 10399.992 | 29026.165 |
| 4 | 38661.671 | 58460.010 | 80226.6 | | 31719.005 | 69826.660 | 51200.509 |
| 5 | 6942.672 | 26741.022 | 48507.6 | 31719.005 | | 38107.699 | 19481.522 |
| 6 | 31165.038 | 11366.758 | 10400.0 | 69826.660 | 38107.699 | | 18626.213 |
| 7 | 12538.853 | 7259.517 | 29026.2 | 51200.509 | 19481.522 | 18626.213 | |

**Network of Attracting Clusters Yr = 2000** (Within and Between Euclidian Distance Measures)

Figure 7: Summit-SIm Dashboard in Netlogo

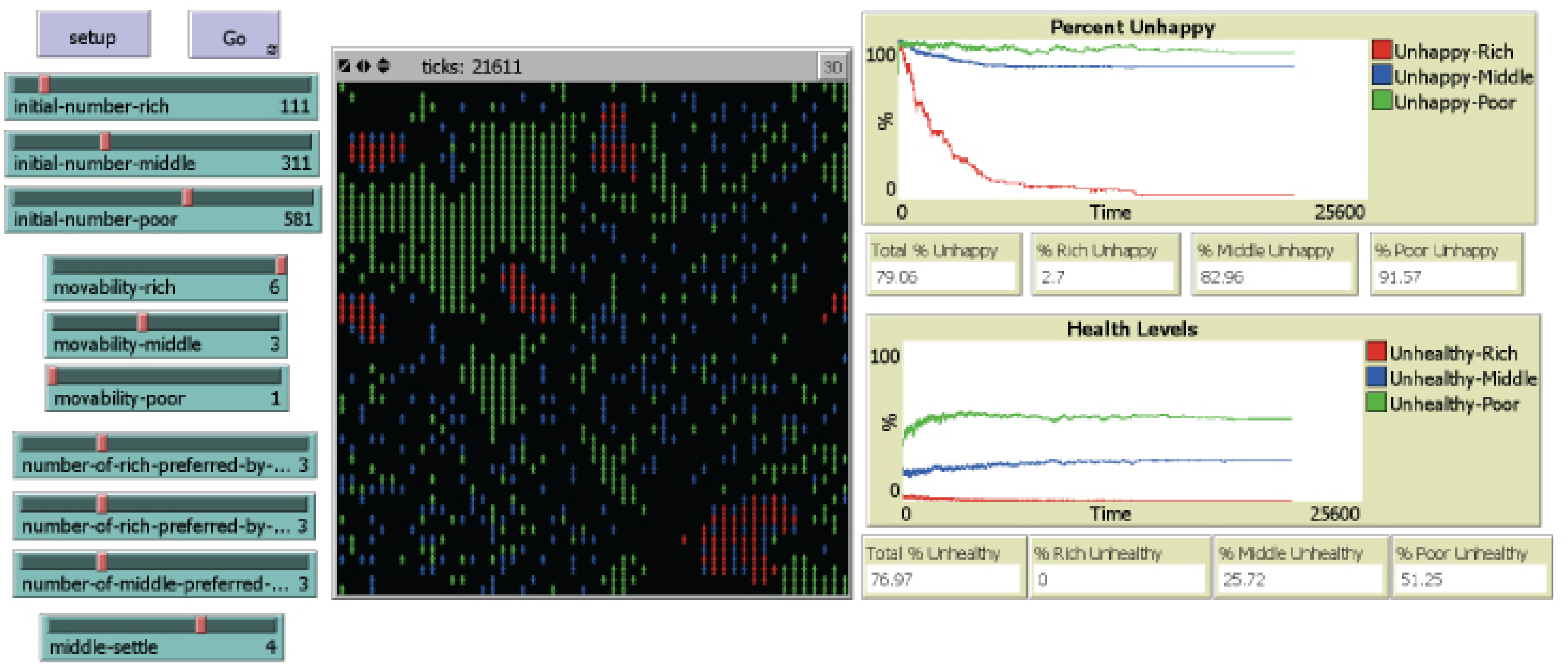


FIGURE 7:
Snapshot of SummitSim with a Preference Rating of 3 for all Agents


**NOTE:** Rich Agents = Squares; Middle Class Agents = Stars; and Poor Agents = Triangles. **Cluster A** identifies one of the dense clusters of rich agents. **Cluster B** identifies one of the dense clusters of poor agents; which complexity scientists would call a poverty trap.