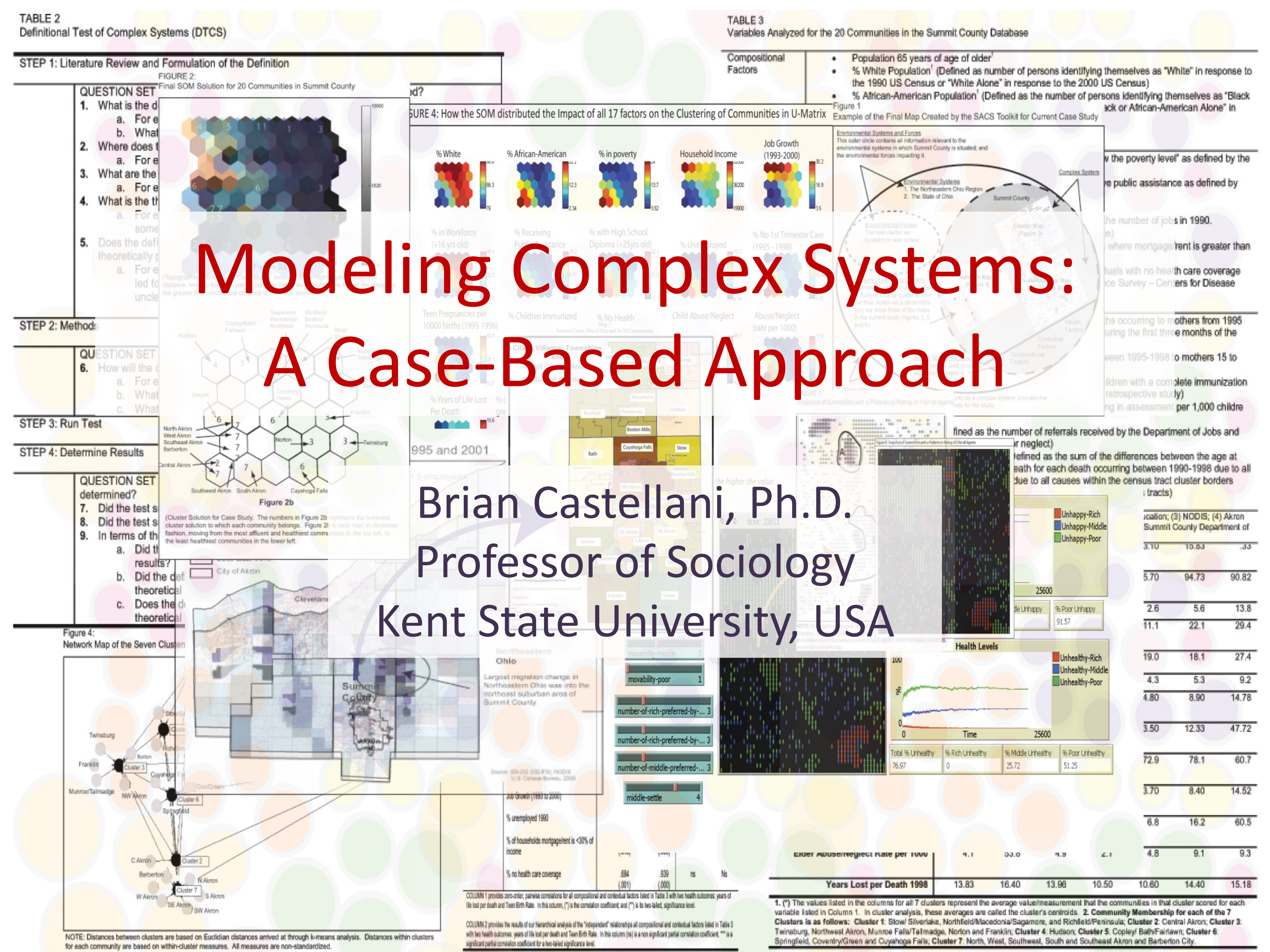


Modeling Complex Systems: A Case-Based Approach

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 linear 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

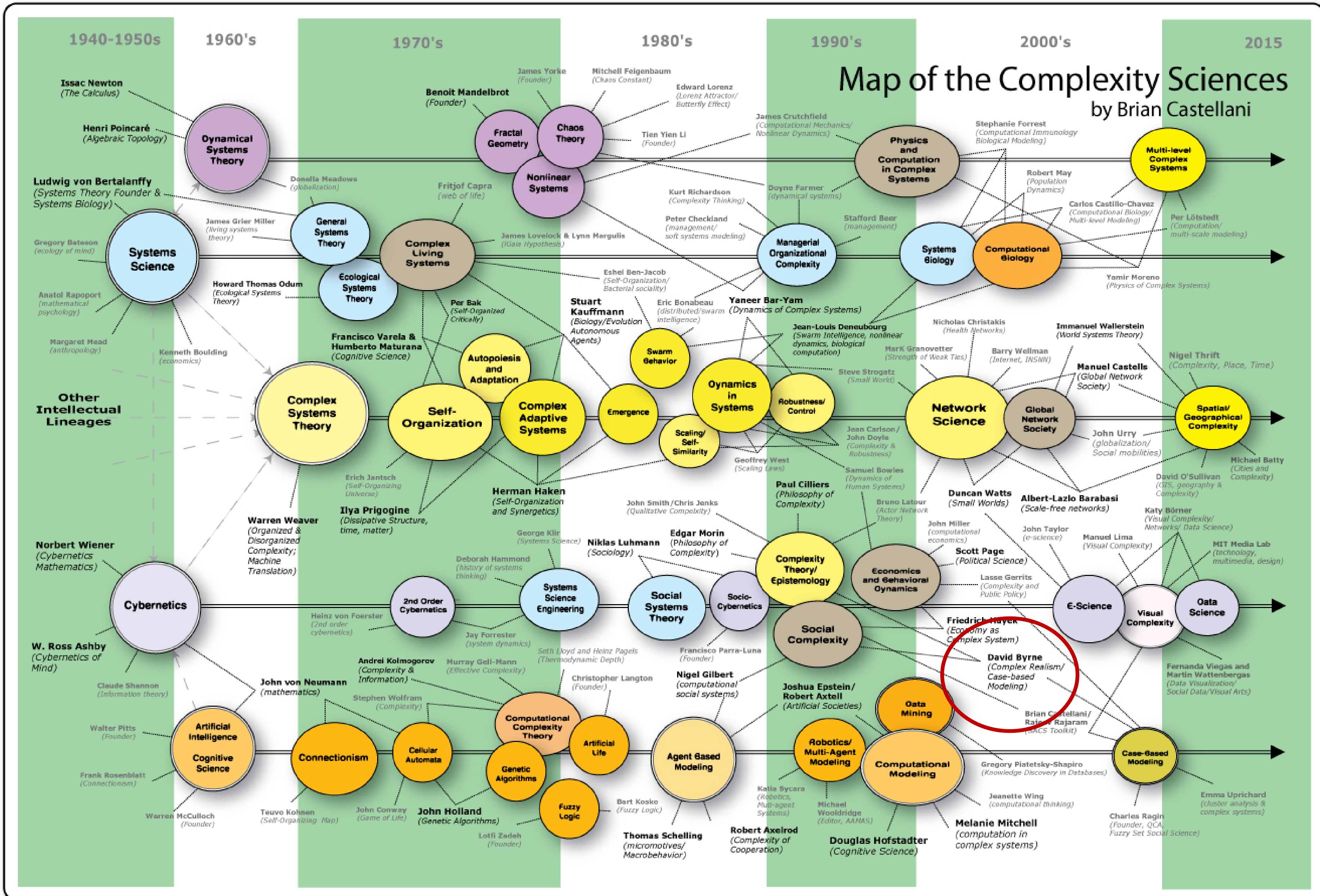
- ***Byrne*** is recognized, internationally, as a leading figure in what most scholars see as two highly promising but distinct fields of study:
  - (1) case-based method and
  - (2) the sociological study of complex systems.
  - An example of the former is Byrne's ***Sage Handbook of Case-Based Methods*** – 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.



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

Cases Are Complex Systems

- There are several strengths to this approach, **three** of which are crucial to the work my colleagues 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 (my colleagues and I) seek to develop a mathematically-rigorous, computationally-based, mixed-methods platform for modeling social complexity and complex social systems.
- The purpose of this presentation is to explore what we have so far accomplished – albeit tentatively.

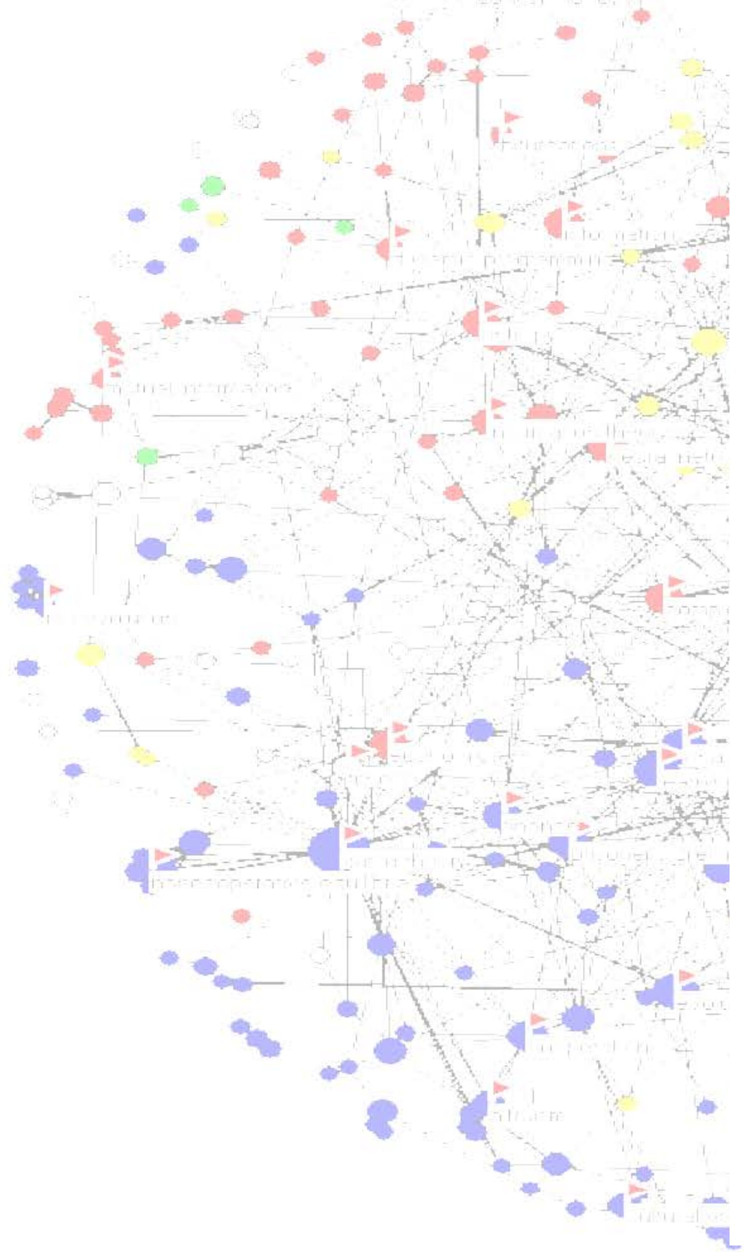
Cases Are Complex Systems

- 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

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:

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

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

(3) … maximum number of cases that can be … practically speaking the upper limit will be bounded, base… on the particular set of cases identified for study—which is always an empirica… issue.

This is our first simplifying assumptions

(4) We denote the numbe… 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.

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] • % 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) • % 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) • 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) • Public Assistance[1] (Defined as the number of households receive public assistance as defined by

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)

s the number of jobs in 1990.
rce)
ls where mortgage/rent is greater than

iduals with no health care coverage
ance Survey – Centers for Disease

irths occurring to mothers from 1995
during the first three months of the

tween 1995-1998 to mothers 15 to

	• 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) • Child Abuse/Neglect[6] (Defined as the number of referrals resulting in assessment per 1,000 childre under 18 years of age) • Elder Abuse/Neglect[7] (Defined as the number of referrals received by the Department of Jobs and Family Services for abuse, exploitation, or neglect) • 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						
19	Bahrain	9158.265	9508.373	9867.			
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**
				.448	27.964	28.48	28.995
				.876	55.471	56.184	57.012
				.678	43.081	43.493	43.914
				.804	30.201	30.599	30.997
				.284	58.779	59.271	59.759
				2.32	62.855	63.331	63.749
				.602	63.024	63.446	63.866
				0.01	60.98	61.873	62.687
				9.12	69.7	69.85	70.17
				66.8	67.29	67.32	67.6
				.754	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	42.876	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	55.086	55.644	56.197	56.745	57.289
25	Benin	32.696	32.955	33.467	33.971	34.466	34.949

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:

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

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)

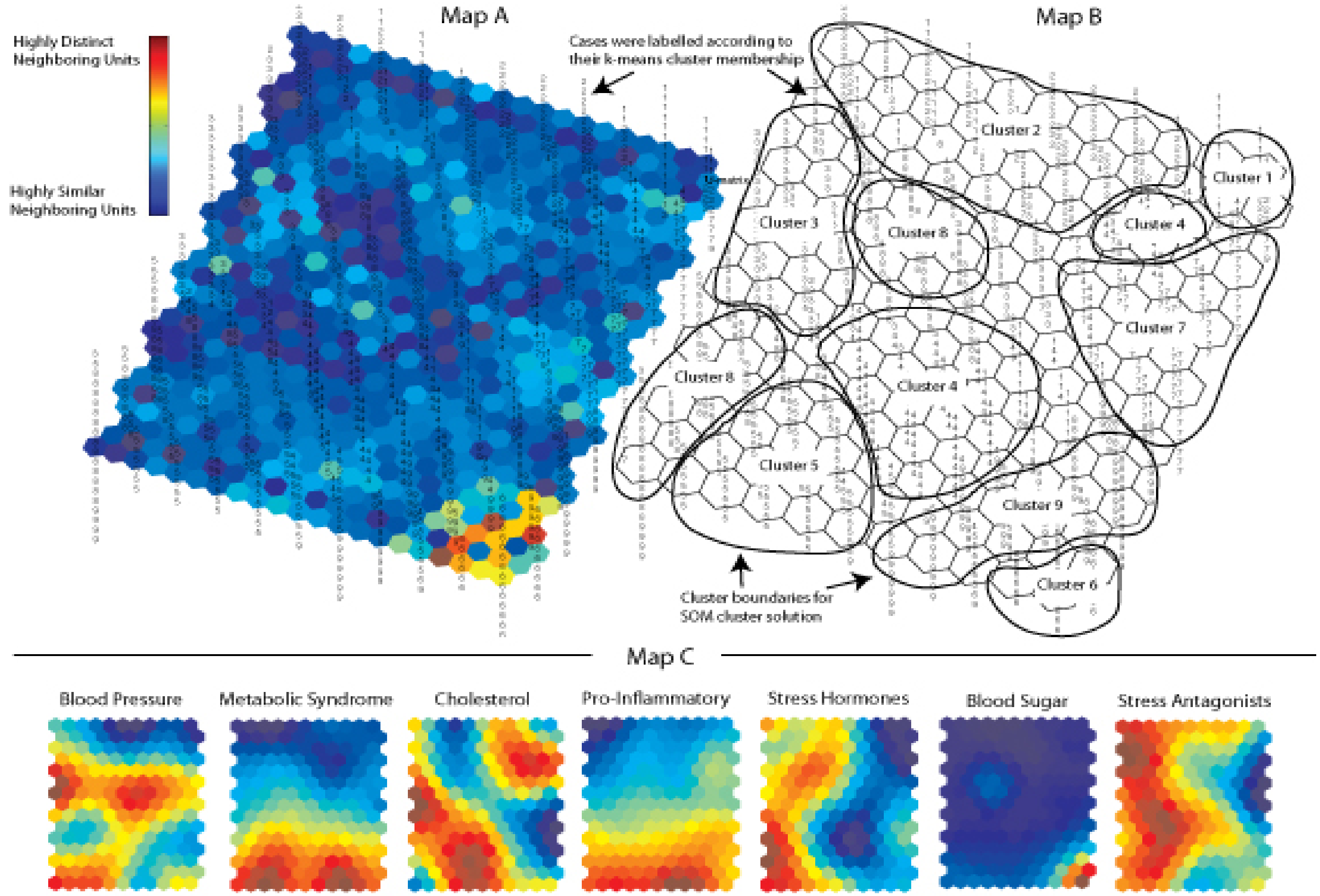


Figure 1.
Map A and *Map B* are graphic representations of the cluster solution arrived at by the Self-Organizing Map (SOM) Neural Net, referred to as the *U-Matrix*. In terms of the information they provide, *Map A* is a three-dimensional (topographical) u-matrix: for it, the SOM adds hexagons to the 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.

Simplifying Assumptions

Our approach (which combines what is known in physics and applied mathematics as the inverse and direct problem) is novel in four important ways: first, we take a unique, data-driven view of the cases in a cohort, which we define as K dimensional vectors, where the velocity vector for each case is computed according to its particular measurements on some set of empirically defined social, psychological, or biological variables.

Second, we translate the data-driven, nonlinear trajectories of these microscopic cohort constituents (cases) into the linear movement of macroscopic trajectories, which take the form of densities.

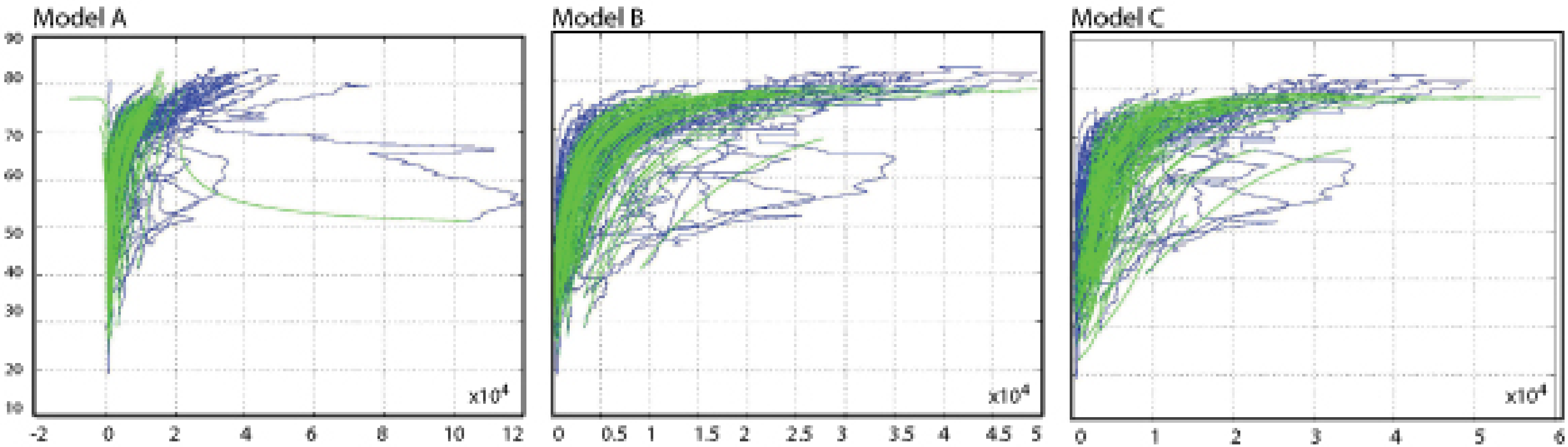
Here, we are drawing on Haken's synergetics and the idea that self-organizing macroscopic trajectories are less dynamic, generally speaking, than microscopic trajectories, which are high dynamic, out of which the former emerge.

Simplifying Assumptions

For our empirical case, we drew our data from the Gapminder website. The Gapminder website (created by Ola Rosling, Anna Rosling Rnnlund, and Hans Rosling) provides researchers, teachers, students, and the general public a wealth of time-series data (often starting back in the early 1900s) on the economic, political, cultural, social, biomedical, and health development of countries throughout the world, which it converts into a series of two-dimensional (2D) animations and interactive graphics (see http://www.gapminder.org/). For the sake of demonstration, therefore, we consider a database with two variables ($K = 2$) from Gapminder; namely, per capita GDP ($x_1(t)$) and life expectancy ($x_2(t)$) for 156 countries over 63 years (t).

FIGURE 4

State Space Fit for Best Model



State space fit for the best model.

State space fit for the best model without Kuwait or Luxembourg.

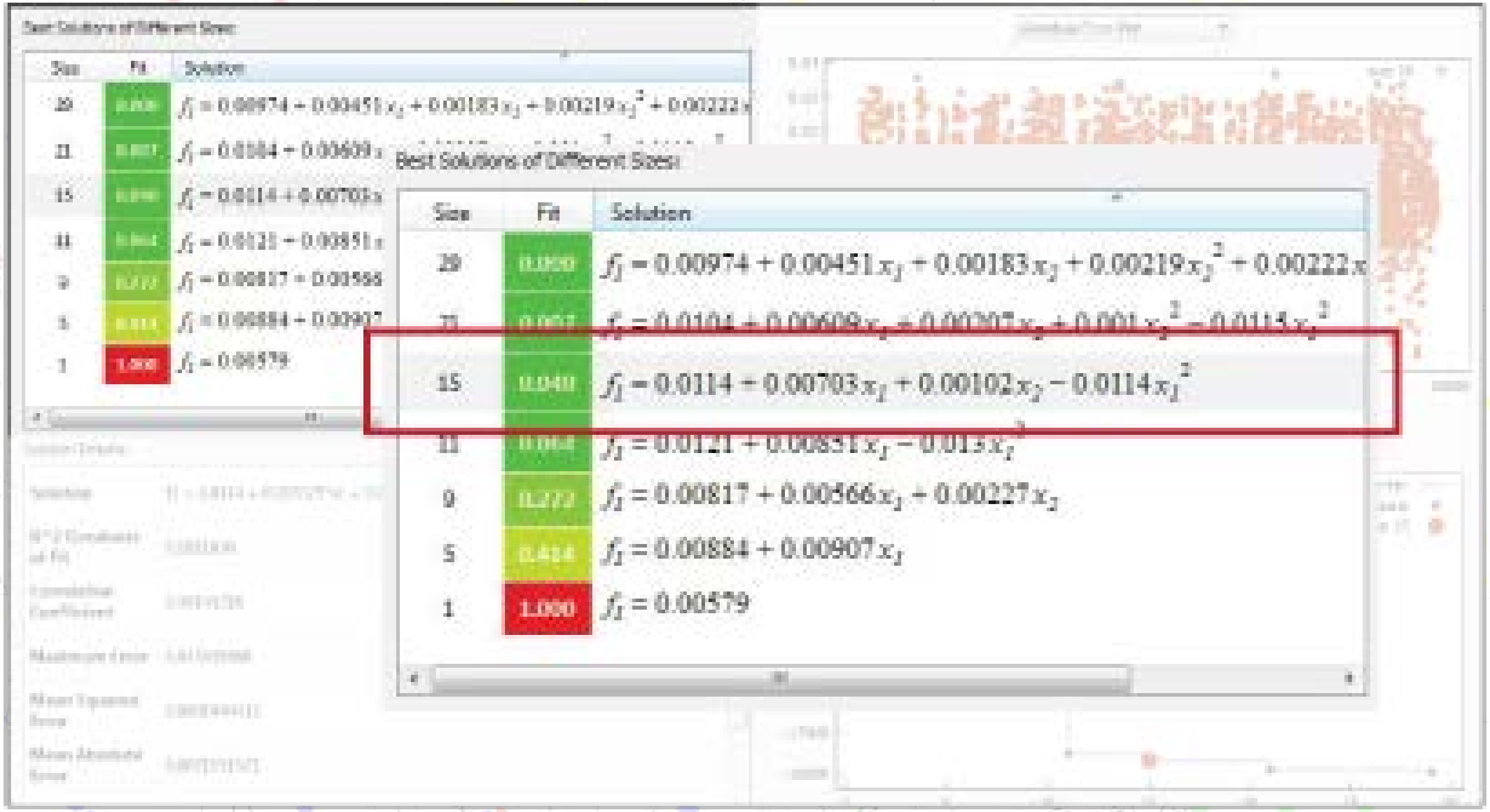
State space fit for the best model without Kuwait or Luxembourg, but with *time* as an independent variable.

Shown here are several computed Matlab models for the first component of velocity vector f1. Models were created using the ordinary differential equation solution from Eureqa. In all three models, the X-axis represents ***GDP***; and the Y-axis represents ***Life Expectancy***. In the models, the blue trajectories are from the data; green trajectories are the fitted model

Simplifying Assumptions

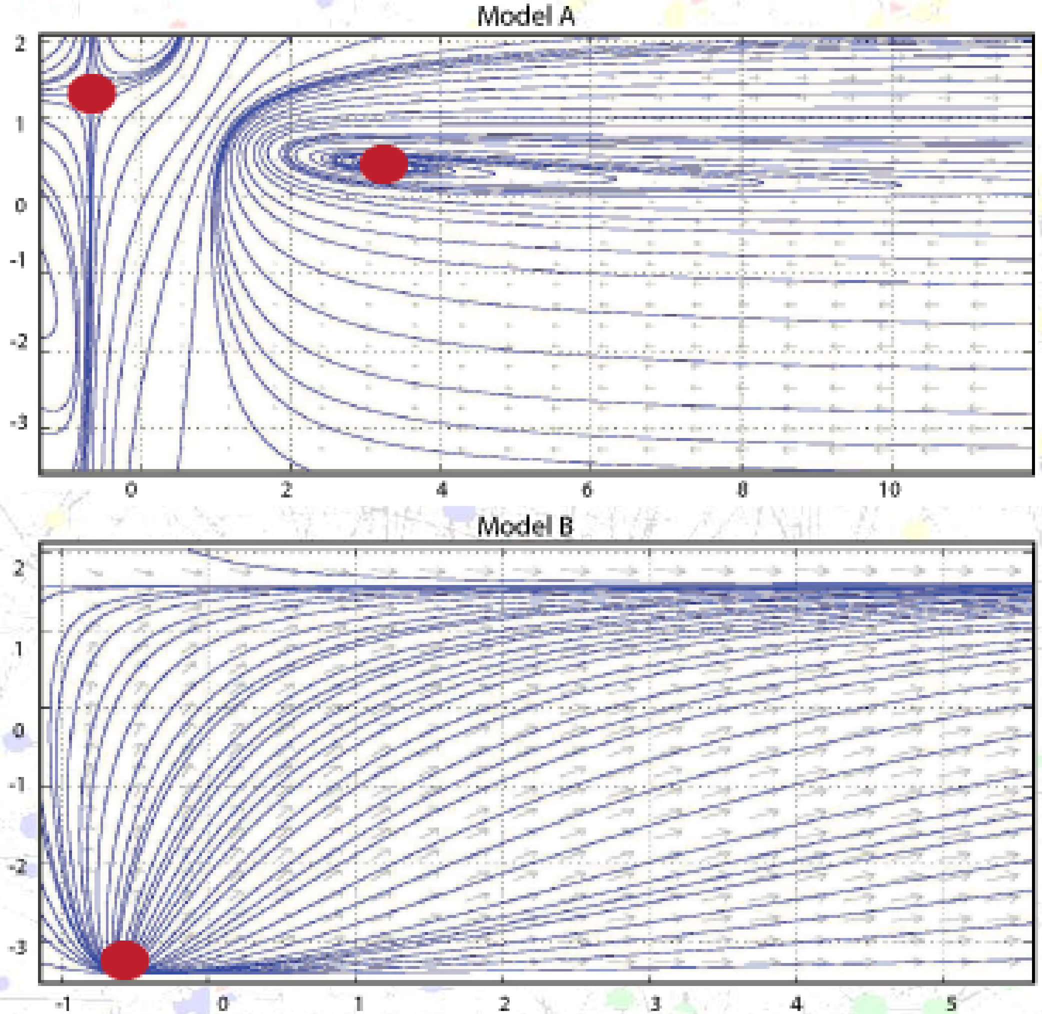
Third, we perform this translation by fitting the time trajectories of these cases using an autonomous (and, in some instances nonautonomous) ordinary differential equation (ODE) (1). In most cohort studies, be they network studies or otherwise, the laws governing their macroscopic dynamics are not known [17,18]. Fitting functions with an autonomous ODE must, therefore, be entirely data driven and based on a "goodness of fit" model. Our unique solution to this data-driven problem is to employ a genetic algorithm, as it does not require any a priori knowledge of the laws governing the data [19]. Instead, it uses the data to evolve to an optimal solution. It can do so because a genetic algorithm is a "brute force" search, but in an efficient way. Also, it finds global minima (as opposed to local minima), hence, it is a global optimization routine. As such, a genetic algorithm allows researchers to find the novel, mechanical laws of motion for social science and biomedical cohort data—with the knowledge that, often, each new dataset presents a new search for new laws, hence the study of complexity [20].

FIGURE 2



*Eureqa gives multiple models for the vector field of velocities. Figure 2 shows several computed models for the first component of velocity vector f1. The best fit model (#15 in our case, shown above) is usually the one that has a mid-level complexity in terms of number of polynomial symbols and the error values in the mid range amongst all models.

FIGURE 3



Shown here is are the state space trajectories for two of the models we settled on using Eureqa. In both models (A and B), the X-axis represents ***GDP*** (converted to z-scores); and the Y-axis represents ***Life Expectancy*** (converted to z-scores). In both models, the arrows show the direction of the trajectories; the larger the arrow the higher the vector's velocity. ***Model A:*** In this model, all countries are included; the red dot located in the top-left sectionof Figure 3 shows a saddle point; and the red dot located in the top-middle sectionof Figure 3 shows a spiraling source. ***Model B:*** In this model, the minority trajectories of Luxembourg and Kuwait were removed; the red dot here is a source.

Simplifying Assumptions

Fourth, using the vector field thus obtained, we use the advection PDE to simulate the evolution of a distribution of cases (as densities) across time (2). The advection PDE has been used extensively in fluid mechanics and electromagnetism to model the transport of physical quantities such as mass and charge, respectively [21,22].

Advection equation – transport of density of cases

- Transforms the motion of individual cases to the motion of a density of cases.

$$P(t) = \iint_{\Omega} \rho(x,y,t)\,dxdy = \iint_{\Omega} \rho_0(x,y)\,dxdy = P_0. \quad (7)$$

- Requires the initial distribution of case profiles, and the velocity vector field of cases (same as the one used in the ODE), and can compute the motion of the initial density assuming that the total number of cases is a constant (called mass conservation property).

$$P_t\rho = \rho(\phi_{-t}(x)) \left| \frac{d\phi_{-t}(x)}{dx} \right|, \quad (8)$$

- Used in modeling of transport phenomena such as fluid dynamics (oil spill), traffic on streets.

$$\rho_t + \nabla \cdot (f\rho) = 0; \quad \rho|_{\Gamma_1} = 0; \rho(x,y,0) = \rho_0(x,y).$$


AP

Advection equation – transport of density of cases

- Notion of transport is applicable to a variety of topics in sociology such as residential mobility and health trajectories.

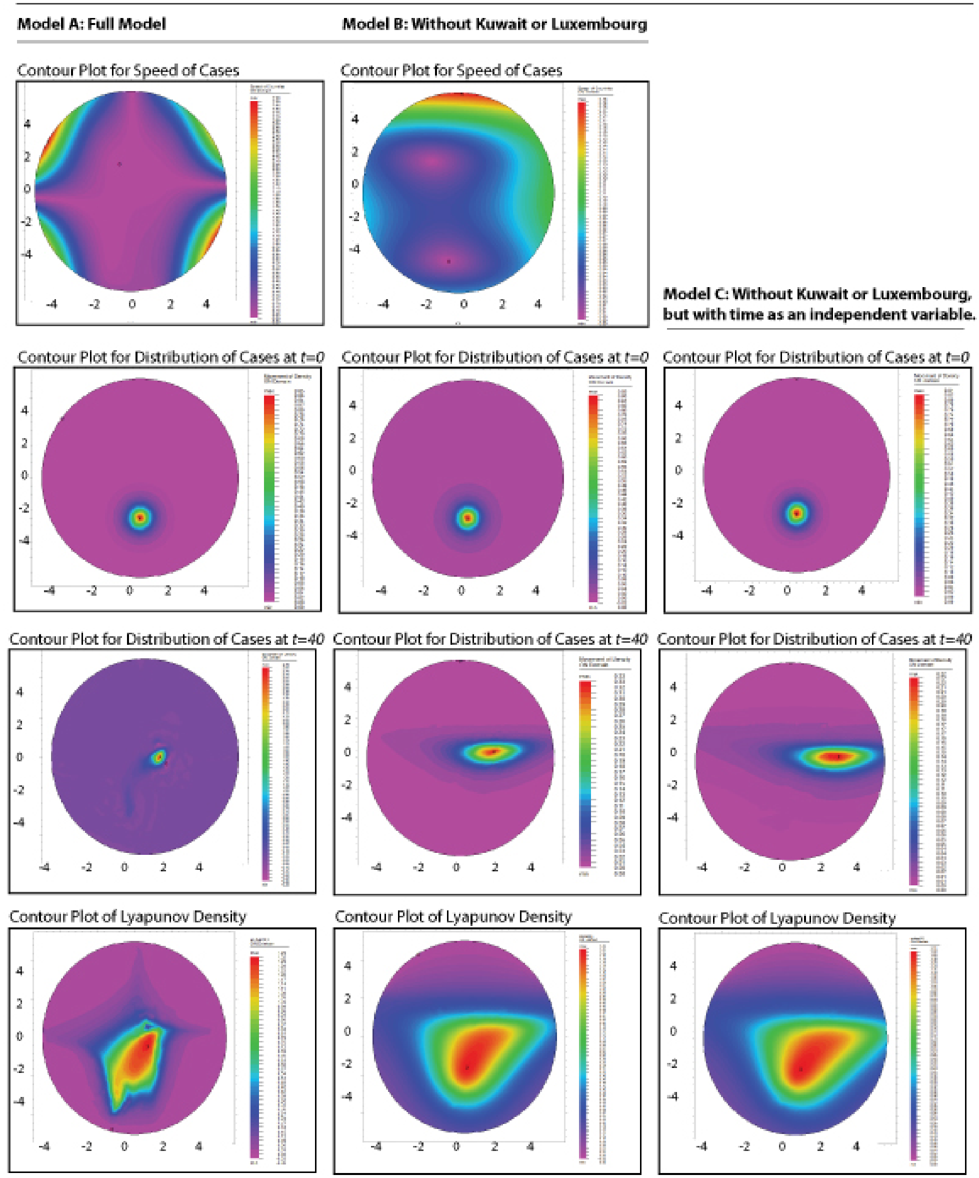
- Residential mobility – variables are actual geographical ones. Trajectories are in physical coordinate space.

- Health trajectories- Variables are biological, sociological markers – state space is more abstract

$$\rho_t + \nabla \cdot (f \rho) = 0; \quad \rho|_{\Gamma_1} = 0; \rho(x, y, 0) = \rho_0(x, y).$$

FIGURE 7

FlexPDE Contour Plots for GDP and Longevity



Shown here are several computed models for FlexPDE. Models were created using the advection equation. In all three models, the X-axis represents ***GDP*** and the Y-axis represents ***Life Expectancy***; both scores on the axes were convereted to z-scores for noramlization and comparison. NOTE: In the Lyapunov Density, higher values mean more cases go through that region)

Uniqueness of our approach

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0; \rho(x,0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

- Continuous time modeling
- Deterministic modeling
- Differential equations (both ODE and PDE)
- Gradation of state space based on velocity of motion
- Non-equilibrium clustering using the Lyapunov density plot

$$\dot{x} = f(x); x(0) = x_0$$

$$\rho_t + \nabla \cdot (f\rho) = 0, \rho(x, 0) = \rho_0(x); \rho|_{\Gamma_i} = 0$$

Strengths

- Prediction of longitudinal evolution of cases with multiple variables across time
- Studying complexity in dynamical motion of cases in the form of saddles, sources, sinks, or periodic orbits
- Gradation of the state space into regions where cases move faster (or slower) from the velocity contour plot
- Non-equilibrium clustering of trajectories from the Lyapunov density plot (higher values mean more trajectories have squeezed through)

Strengths

- Prediction of majority trends in trajectories for novel choices of initial profiles or densities
- Multiple models to describe the same phenomena allowing for a choice of better ones
- Ease of incorporation of new data into the modeling process to fit the database as it grows

Thanks!