Complexity and its Implications to the Systems Engineering Process

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

Abstract

- Systems Engineers are moving ahead with various initiatives to enable Model Based Systems Engineering. But these initiatives may not be capturing the idea that these systems are Complex. INCOSE and ISO-15288 process areas capture best practices to address complicated systems, but that there are additional methods require to address the emergent behavior found in Complex Systems. This presentation will discuss some of the deficiencies in the current ISO-15288 processes, then provide an overview of the SACS Toolbox approach to viewing a complex system.

Message

- A popular understanding of Systems Engineering is if an organization is mature in the INCOSE Process Areas, then, the products and services it produces will be optimal (low cost, high reliability, optimal stakeholder satisfaction.)
- This is an optimization of a static, complicated system, but not necessarily an optimization of a complex, dynamic system, with emergent behavior – particularly in terms of *sociotechnical systems!*

Background on INCOSE / Systems Engineering Practiced by This Professional Group…

Origins of the Process Areas

- In the beginning, there were **Mistakes** made when developing complex products and services.
- People developed **Checklists** to try to make sure these didn't happen again.
- Sometimes the Checklists were moved up front in development and called **Requirements or Standards**.
- Taxonomists and Ontologists (well, groups of engineers) got together and started to group these Checklist activities into focus areas, and called them **Process Areas**.
- Soon, **Communities of Practice** and event departments started to spring up named after these Process Areas…
- **Systems Engineering** as a business function was born.
- As Computers improved to the present day, Systems Engineers are advancing the use of the software tools and methodology to assist in the practice of these process areas: **MBSE – Model Based Systems Engineering** was born.

Note: Notice the Cartesian approach: Break things into its component parts, sum the parts into a whole…

INCOSE Process Areas

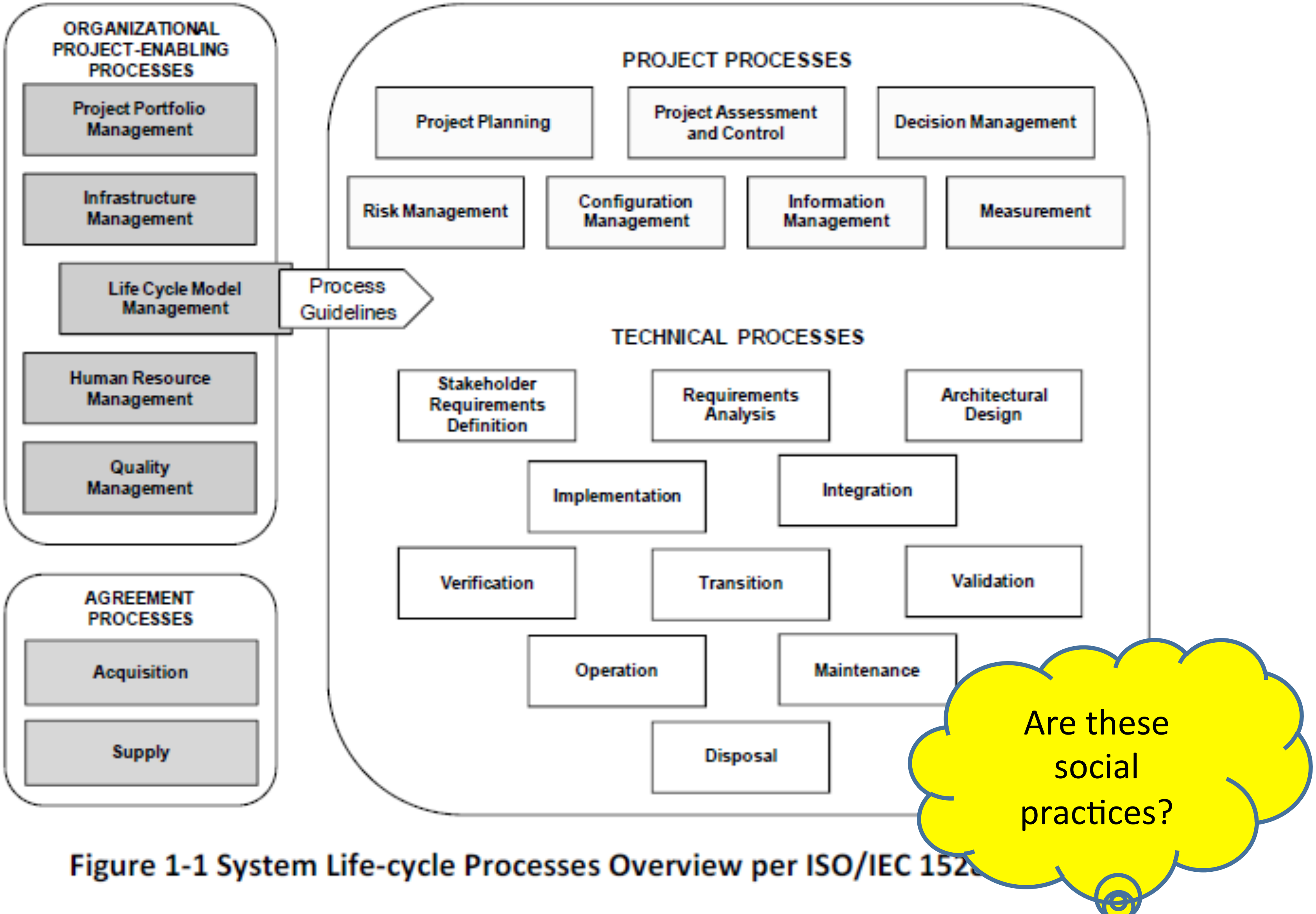


Figure 1-1 System Life-cycle Processes Overview per ISO/IEC 152

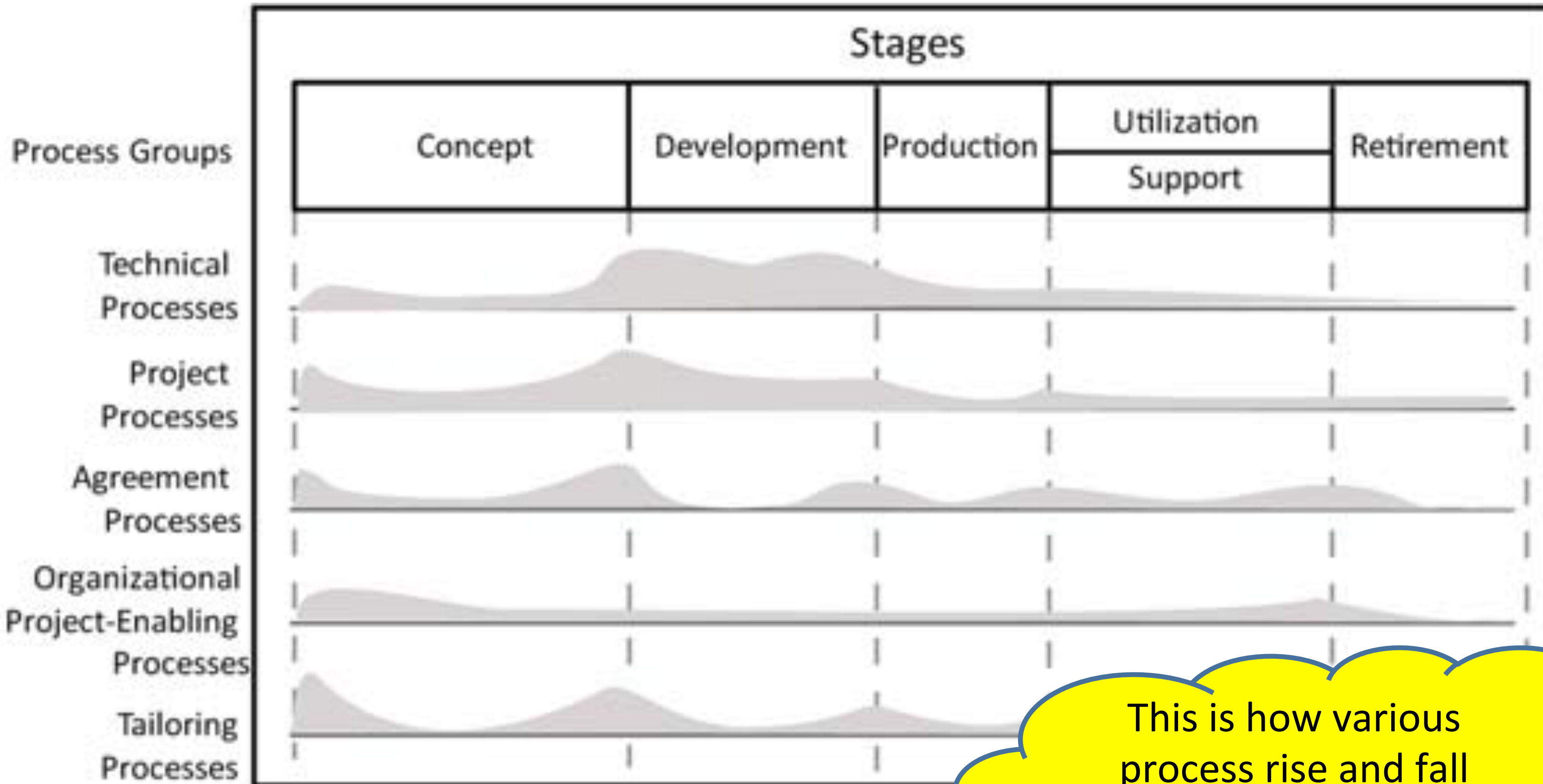
INCOSE Processes – Communities of Practice

Process Area	Binary Code (Heuristic)	Programs	Artifacts of the Process	Typical Process Participants
CONFIGURATION MANAGEMENT	Controlled	Change Control	**Configurations** with Key Grid Assets	**CM Staff,** CM Managers, IT Specialists, Maintenance Workers
INFORMATION MANAGEMENT	Information	Ontology/Workflow	**Documents**, Records, Wikis, Journals containing Grid Information	**IT Specialists**, Document or Records Control Staff, Functional Managers, HR Knowledge Mgrs, Trainers
HR MANAGEMENT	Productive Trained	Performance Evaluations Skills	**Employees** critical to Grid Reliability	**HR Staff**, HR Managers, Managers
IMPLEMENTATION	Ready	Build Instructions	**Systems** Recently Implemented or Under Implementation impacting Grid Reliability	**Installers,** Lead Engineers, End Users of System
INTEGRATION	Connected	State Machines	**Systems** Recently Integrated or Under Integration impacting Grid Reliability	**Systems Engineers**, Project Managers, Test Technicians, Integration Technicians
OPERATIONS	Operating	Operations Plans, MRP Runs	**Control Centers**, Control Rooms, Different Shifts, Local/Backup	**Operations** Managers, Operations Staff
MAINTENANCE	Repaired	Maintenance Schedule and Procedures	**Maintenance Crews**, Maintenance Project Teams, Different Shifts	**Maintenance** Managers, Maintenance Staff
PROJECT CONTROL	Healthy	Schedule, Resources Allocation	**Standard Work and Project Work** (including scheduling/monitoring) impacting Grid Reliability	**Project Managers**, Line Managers
EXTERNAL DEPENDENCIES	Dependent	Relationship Maps	**List of External Resources** Critical to Grid Reliability and their Interdependencies	**Security Managers**, Business Managers, Operators, Maintenance Workers, Executives, etc
MEASUREMENT	Accurate	Analytics	**List of Measurement Objectives** QM	**Analysts**, IT Staff, Managers, Kaizen Leaders, Executives
VERIFICATION	True	Testing	**Systems** recently verified, or under verification that impact grid reliability	**Test Technicians**, Project Managers, Systems Engineers, Line Managers
VALIDATION	Desired	Surveys, Tracability	**Systems** recently validated or under validation that impact grid reliabilty	**System Customer Liasons**, System Customers, System Operators, Stakeholders with Approval Authority

Process Area	Binary Code (Heuristic)	Programs	Artifacts of the Process	Typical Process Participants
QUALITY MANAGEMENT	Quality	Auditing	**Policy, Procedures, and Objectives** for achieving Grid Reliability	**Compliance** Managers, Quality Managers, Compliance Staff, Quality Staff
RISK MANAGEMENT	Risky	Fault Trees, FMEA, Root Cause Analysis	List of **Grid Reliability Risks/Threats**	**Risk Analysts**, Systems Engineers, Security Managers, Project Managers
DECISION MANAGEMENT	Selected	MCDA, Decision Trees	List of **Key Decisions**, Decision Management Policy / regarding Grid Reliability	**Decision Makers**, Managers of Decision Makers
PROJECT PLANNING	Planned	Resource List, Scheduling Program	**List of Projects** currently in planning, or recently planned that impact Grid Reliability	**Project Managers**, Line Managers, Project Coordinators
SUPPLY	Good Deal	Supply Chain Analysis, Contracts	**List of Parts or Process** currently in sourcing	**Supply Chain Manager,** Commodity Managers
ACQUISITION	ROI	Investment Study, Balance Sheet	**Sales List** of parts/contracts	**Sales Force,** Business Managers, Salesforce
AKEHOLDER REQUIREMENTS DEFINITI	Desired	Requirements Tool	**Requirements List** for the program	**Systems Engineers,** Customers, Stakeholders
REQUIREMENTS ANALYSIS	Required	Trace Tool	**Tracability Matrix** of requirements	**Systems Analysts,** technicians
ARCHITECTURAL DESIGN	Functional	Various Design Programs	**Prints, Software, Concept** designs for systems	**Design Engineers,** Analysts
TRANSITION	Sufficient	Checkout Procedure, Readiness	**Checklist** of transition items	**Test Technicians,** Build crew
DISPOSAL	Scrapped	Material Analysis, Hazard Analysis	**Inventory List** of disposed items	**Facilities Manager**, Safety, Maintenance Staff
LIFE CYCLE MODEL MANAGEMENT	Value Added	Process Improvements	**Process Library** of policies, procedures, SOP	**Process Manager**, Process Improvement Staff
INFRASTRUCTURE MANAGEMENT	Capitalized	Capitalization Tool, Forecasting Tool	**Asset List** of buildings, equipment, resources	**CFO,** Leadership Team, Finance
PROJECT PORTFOLIO MANAGEMENT	ROI	Portfolio Tool	**Portfolio** of projects in processes and planned	**Portfolio Managers,** Leadership Team, Finance Team
TAILORING	Applicable	Process Library Trace	**Tailoring Guidelines, Tailoring Process**	**Process Manager**, Process Improvement Staff

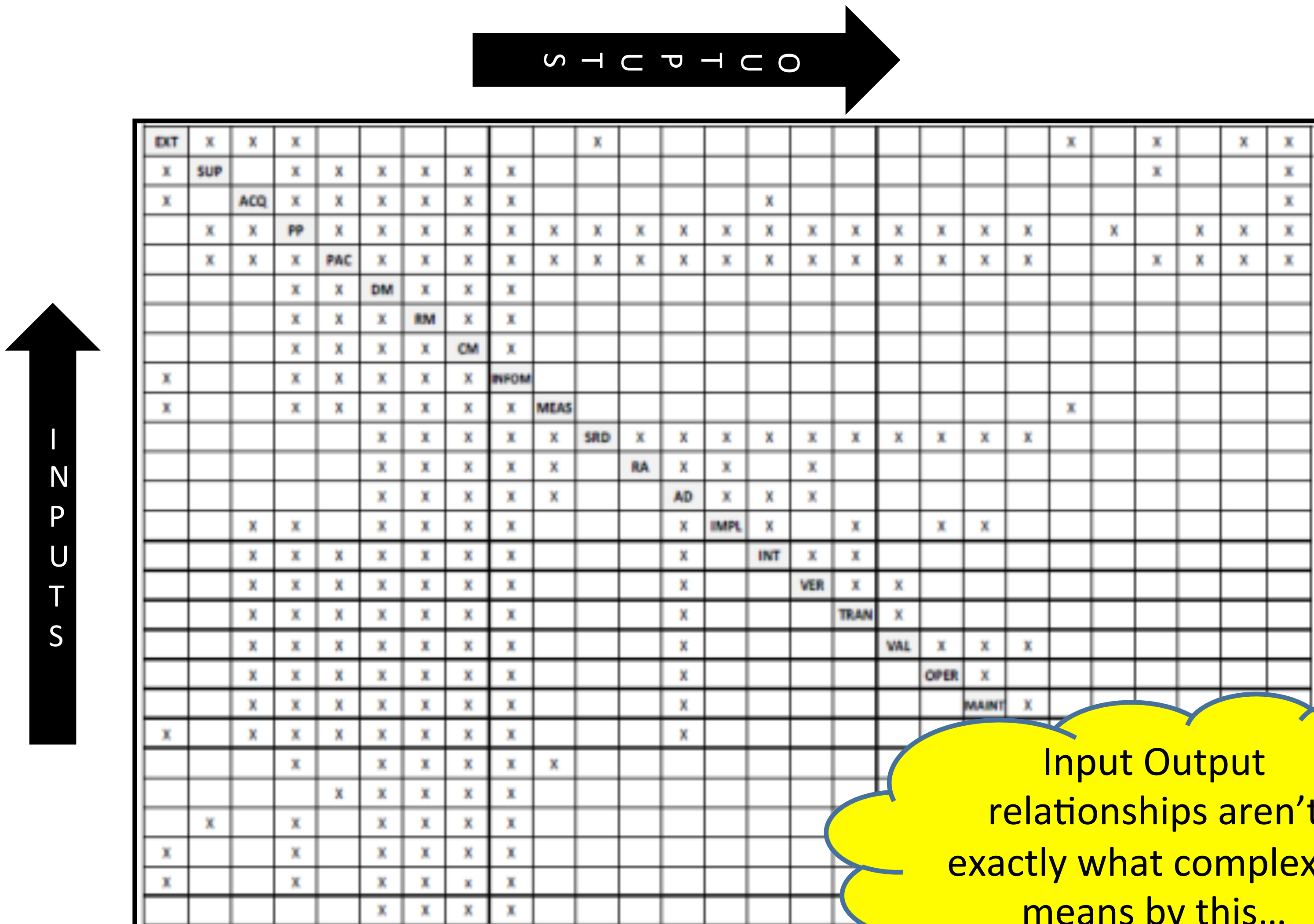
More details on who practices these areas an how they think

INCOSE Concept of Time



This is how various process rise and fall throughout a project, but not Dynamics…

INCOSE Concept of Interrelatedness



Input Output relationships aren't exactly what complexity means by this…

CMMI's Maturity Suggested Progression Through the Systems Engineering Process Areas

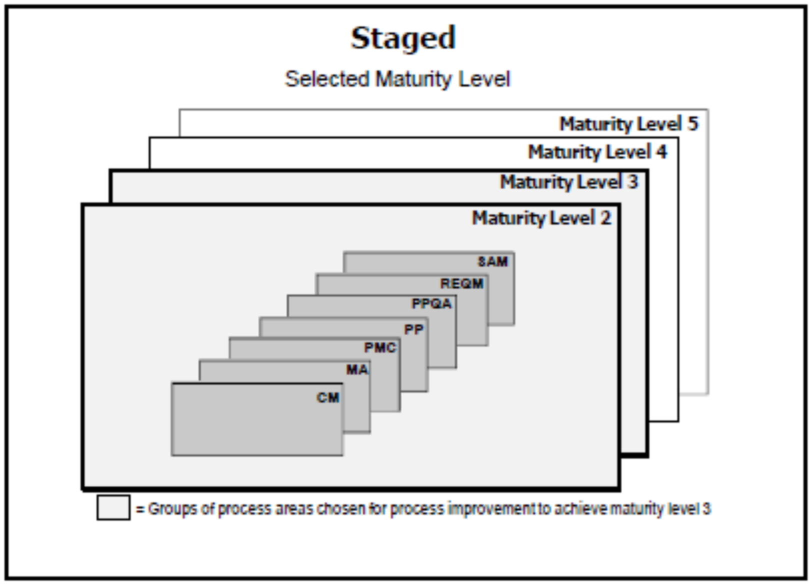


Figure 3.2: Process Areas in the Continuous and Staged Representations

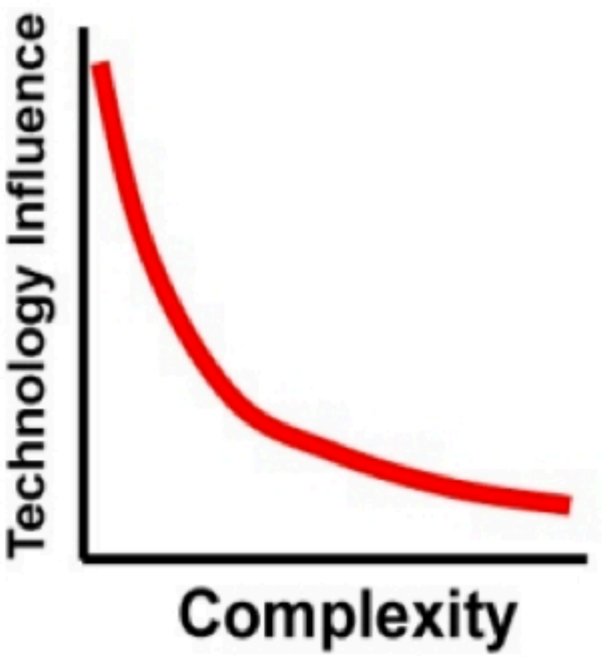
Table 3.2 Process Areas, Categories, and Maturity Levels

Process Area	*Category*	*Maturity Level*
Causal Analysis and Resolution (CAR)	Support	5
Configuration Management (CM)	Support	2
Decision Analysis and Resolution (DAR)	Support	3
Integrated Project Management (IPM)	Project Management	3
Measurement and Analysis (MA)	Support	2
Organizational Process Definition (OPD)	Process Management	3
Organizational Process Focus (OPF)	Process Management	3
Organizational Performance Management (OPM)	Process Management	5
Organizational Process Performance (OPP)	Process Management	4
Organizational Training (OT)	Process Management	3
Product Integration (PI)	Engineering	3
Project Monitoring and Control (PMC)	Project Management	2
Project Planning (PP)	Project Management	2
Process and Product Quality Assurance (PPQA)	Support	2
Quantitative Project Management (QPM)	Project Management	4
Requirements Development (RD)	Engineering	3
Requirements Management (REQM)	Project Management	2
Risk Management (RSKM)	Project Management	3
Supplier Agreement Management (SAM)	Project Management	2
Technical Solution (TS)	Engineering	3
Validation (VAL)	Engineering	3
Verification (VER)	Engineering	3

Just improving pieces doesn't mean the emergent whole will be better…

Complex Systems Engineering?

We have found that *as complexity has increased, the ability of technology to dominate the solution space is diminished*. We lack capabilities to effectively address multidisciplinary problems emerging in the 21st century.



In order to engage high complexity systems problems the hard technological perspective must be expanded to include a soft perspective that accounts for human, political, organizational, managerial and policy elements associated with the complex systems problem.

IEEE 9th International System of Systems Engineering Conference (SoSE 2014) – Outline Program

Time	Mon 9 Jun – Parallel Session 1	Mon 9 Jun – Parallel Session 2	Tue 10 Jun – SoSE Modelling Stream	Tue 10 Jun – Defence Stream	Tue 10 Jun – Academic Stream	Wed 11 Jun – Single Session
7:30	Registration		Registration			Registration
8:00	Tutorial 1a (Half Day) **Complex Systems: How to Recognize Them and Engineer Them: Part A** Dr Brian E White *CAU<--SES*, USA		Welcome Opening			Keynote: **Why Intelligent Systems are Complex Systems** Dr Chee-Peng Lim, *Deakin University*, Australia
8:30			Keynote: **Systems of Systems: Perspectives, Pain Points and Prospects** Dr Judith Dahmann, *Mitre Corporation*, USA			
			Keynote: **Complex Enterprise Reform in Defence and Health** Dr Richard Hodge, *Brooke Institute*, Australia			Panel Session: **An Australian Industry perspective of how Systems of Systems issues are treated within Major Projects**
10:15	Morning Break		Morning Break			
10:45	Tutorial 1a (Half Day) Continued		Paper Session	Special Session	Paper Session	Keynote: **Application of Complex Systems Science to the Management of Systems of Systems** Prof Vernon Ireland, *University of Adelaide*, and Prof Stephen Cook, *University of South Australia*, Australia
						Panel Session: **Engineering System of Systems for Future Smart Cities**
12:30	Lunch	Registration	Lunch			
13:30	Tutorial 1b (Half Day) **Complex Systems: How to Recognize Them and Engineer Them: Part B** Dr Brian E White *CAU<--SES*, USA	Tutorial 2 (Half Day) **Influencing the Organization System of Systems Through Strategy** Mr Mark A Wilson *Strategy Bridge International*, USA	Keynote: **Complex Systems Technology and What's Next** Dr Terry Stevenson, *Raytheon Australia*, Australia			Local Tour
			Paper Session	Special Session	Tutorial 3 **Using the Incremental Commitment Model to Evolve SoS Capabilities**	
15:15	Afternoon Break		Afternoon Break			
15:45	Tutorial 1b (Half Day) Continued	Tutorial 2 (Half Day) Continued	Panel Session: **To set up complex undertakings for success, how much up front design is enough?**		Dr Jo Ann Lane *U of Southern California* & Dr Rich Turner *Stevens*, USA	Wine Tasting
18:00						Conference Dinner with **After Dinner Speaker:** Dr Charles Keating
18:30	Welcome Reception					
22:30						

Time	Thu 12 Jun – Civil Industry Stream	Thu 12 Jun – Defence Stream	Thu 12 Jun – Enterprise & Applications Stream	Fri 13 Jun – Parallel Session 1	Fri 13 Jun – Parallel Session 2
7:30					
8:00	Registration			Registration	
8:30	Keynote: **Research Challenges in SoSE** Prof Michael Henshaw, *Loughborough University*, UK			Tutorial 4a (Half Day) **Complex Adaptive Methods for SOS: Part A** Dr John Findlay *Maverick & Boutique, USA*	Site Visit with Keynote: **The Air Warfare Destroyer** Mr Peter Croser, *Department of Defence*, Australia
	Special Session	Paper Session	Paper Session		
10:15					
10:45	Special Session	Special Session	Paper Session	Tutorial 4a (Half Day) Continued	
12:30	Lunch				
13:30	Keynote: **The Supply Chain as a Complex System in German Manufacturing** Prof Frank Schultmann, *Karlsruhe Institute of Technology*, Germany			Tutorial 4b (Half Day) **Complex Adaptive Methods for SOS: Part B** Dr John Findlay, *Maverick & Boutique, USA*	
	Paper Session	Special Session	Paper Session		
15:15					
15:45	Paper Session	Special Session	Paper Session	Tutorial 4b (Half Day) Continued	
16:45					
17:30					
22:00					

Dartmouth Engineer Magazine

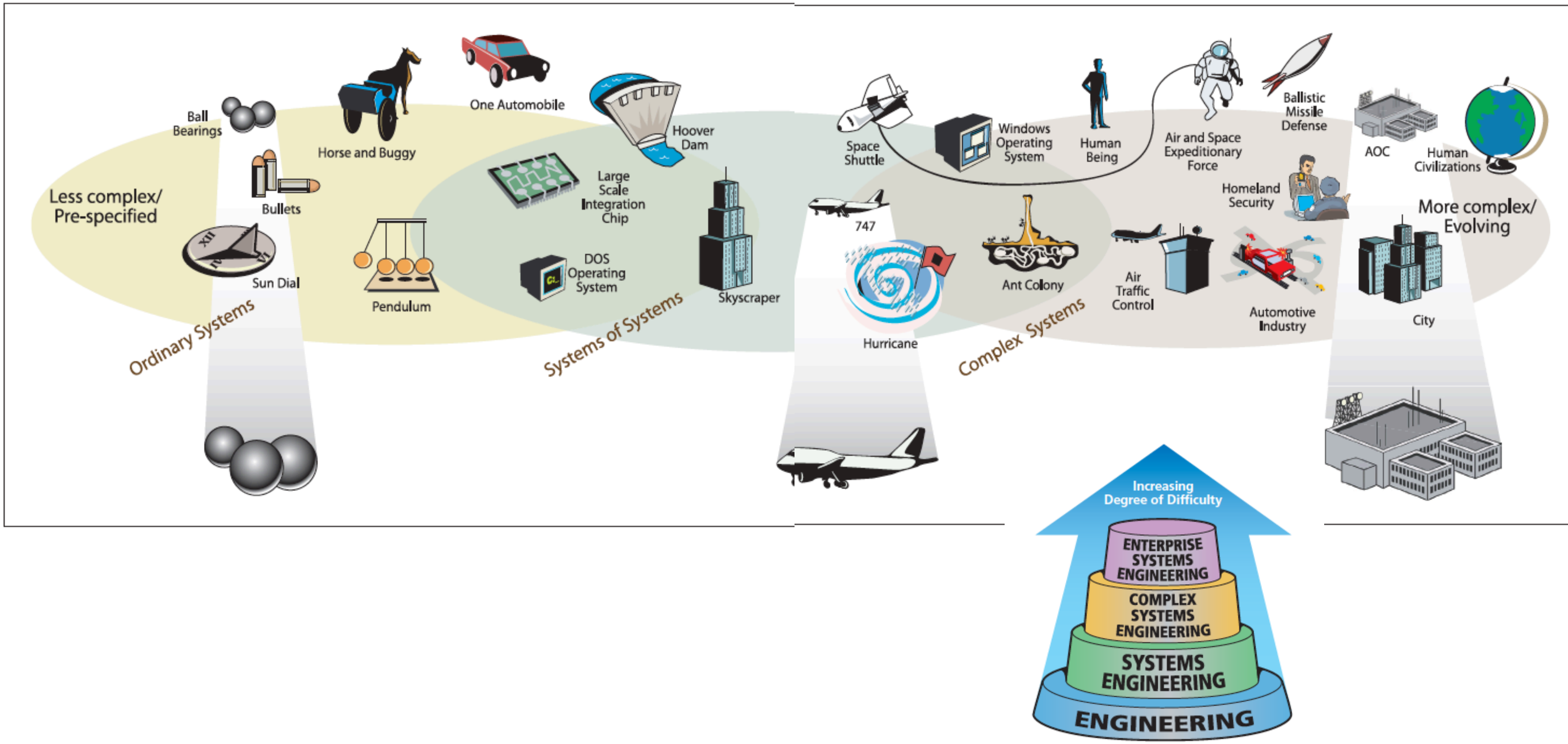
Complex Systems

Tackling surprises in multi-component systems, from human behavior to robotic smarts

BY LEE MICHAELIDES AND KAREN ENDICOTT
COVER ART BY MICHAEL AUSTIN

Don't worry if you're not sure what a complex system is. Even the people who study multi-component systems, such as the internet, communication networks, industrial processes, and interacting teams of robots, define complex systems in various ways.

Some see complex systems as having so many components that they are difficult or impossible to model. Others emphasize that interacting components produce unexpected emergent properties that make the overall system tough to model. Still others see complex systems as intricate interfaces between humans, nature, and technologies. One of Thayer's three research focus areas (the other two are engineering in medicine and energy), complex systems provides room for creative new ways of thinking about the world around us. Here we look at some of the complex system challenges that Thayer professors are trying to understand and solve.



Engineering Discipline Sets

Cybernetics – Which Systems can INCOSE's Processes Address Well?

- 1st Order – Systems as objects
- 2nd Order – How to measure being a part of the system (dynamics)
- 3rd Order – Complex, open-minded, dynamic, self organizing, emergent – system reproduces itself while adapting to change
- 4th Order – System redefines itself entirely

INCOSE Complexity Working Group's Advice

Assertion: Complex-systems can only be engineered by intervention, not by specification and then development. - Brian White:

Complex Adaptive Systems Engineering (CASE) - Complex Systems Engineering Principles:

- **1. Bring Humility**
- **2. Follow Holism**
- **3. Achieve Balance**
- **4. Utilize Trans-Disciplines**
- **5. Embrace Political, Operational, Economic, and Technical Factors**
- **6. Nurture Discussions**
- **7. Pursue Opportunities**
- **8. Formulate Heuristics**
- **9. Foster Trust**
- **10. Create Interactive Environment**
- **11. Stimulate Self-Organization**
- **12. Seek Simple Elements**
- **13. Enforce Layered Architecture**

Some INCOSE members may be familiar with the Complexity Working Group's advice in this area

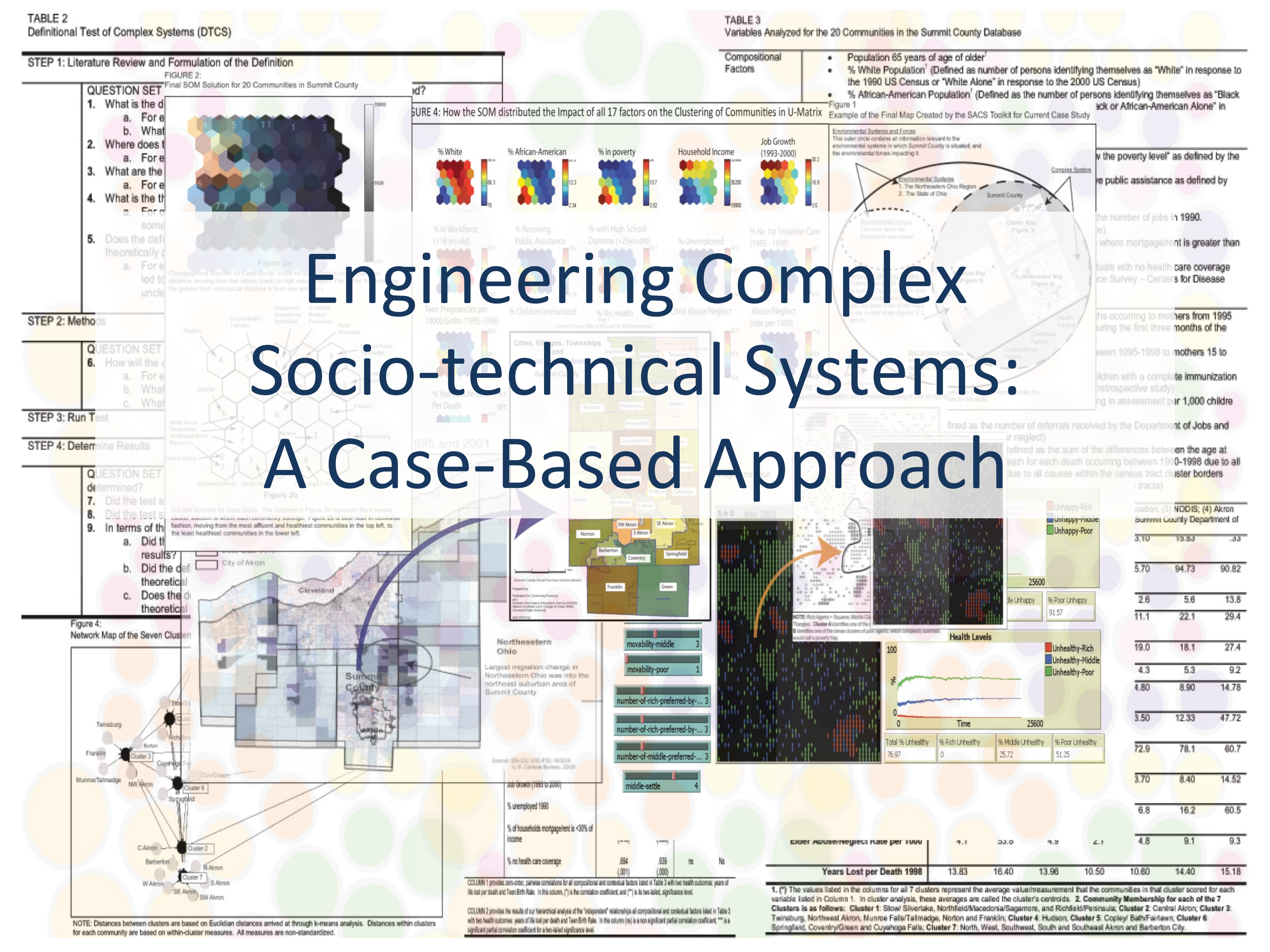
Systems Engineering – Applying Complex Systems Thinking

- System Engineers cannot control complex systems development. They can only influence projects by targeted communications.
- System Engineers need dynamic models of the Social Behavior of their teams in order to steer them through targeted communications
- System Boundaries should be drawn around the stakeholders and environment of the system, not the development team
- System Engineers need to bridge the gap between the Natural Sciences, the Social Sciences, and the Humanities

Suggested Competencies to Consider

- Applied Complexity Science
- Coursework in Non-linear Dynamics and Chaos
- Coursework in Communications
- Coursework in performing Surveys and analyzing their results correctly
- Coursework in Causal Mapping
- Coursework in Data Mining and Data Analytics
- Experience with leading Self-Organizing Teams (e.g. Open Space, World Café, Kaizens – not just Waterfall or Agile Project Management Processes.)
- Practice applying tools to Social Systems, not just Technical Systems

Engineering Complex Socio-technical Systems: A Case-Based Approach



- 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

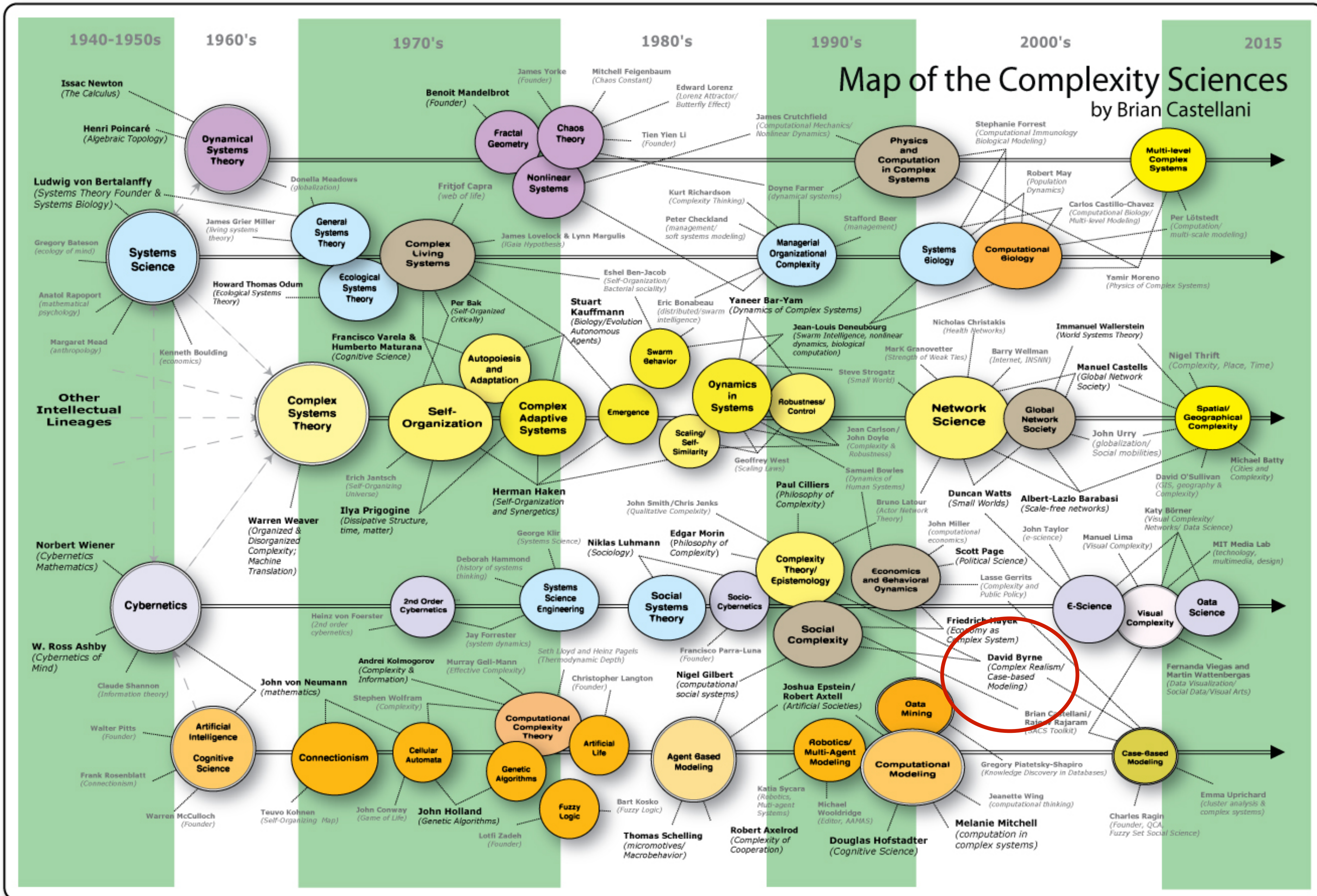
- ***Byrne*** is recognized, internationally, as a leading figure in what most scholars see as two highly promising but distinct fields of study:
  - (1) case-based method and
  - (2) the sociological study of complex systems.
  - An example of the former is Byrne's ***Sage Handbook of Case-Based Methods*** – 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)
Systems Science
Gregory Bateson (ecology of mind)
Anatol Rapoport (mathematical psychology)
Margaret Mead (anthropology)
Kenneth Boulding (economics)
James Grier Miller (living systems theory)
Donella Meadows (globalization)
General Systems Theory
Ecological Systems Theory
Howard Thomas Odum (Ecological Systems Theory)
Fritjof Capra (web of life)
Complex Living Systems
James Lovelock & Lynn Margulis (Gaia Hypothesis)
Benoit Mandelbrot (Founder)
Fractal Geometry
James Yorke (Founder)
Chaos Theory
Mitchell Feigenbaum (Chaos Constant)
Edward Lorenz (Lorenz Attractor/ Butterfly Effect)
Tien Yien Li (Founder)
Nonlinear Systems
James Crutchfield (Computational Mechanics/ Nonlinear Dynamics)
Physics and Computation in Complex Systems
Doyne Farmer (dynamical systems)
Stephanie Forrest (Computational Immunology Biological Modeling)
Multi-level Complex Systems
Robert May (Population Dynamics)
Carlos Castillo-Chavez (Computational Biology/ Multi-level Modeling)
Per Lötstedt (Computation/ multi-scale modeling)
Kurt Richardson (Complexity Thinking)
Peter Checkland (management/ soft systems modeling)
Stafford Beer (management)
Managerial Organizational Complexity
Systems Biology
Computational Biology
Yamir Moreno (Physics of Complex Systems)
Eshel Ben-Jacob (Self-Organization/ Bacterial sociality)
Per Bak (Self-Organized Criticality)
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)
Francisco Varela & Humberto Maturana (Cognitive Science)
Autopoiesis and Adaptation
Swarm Behavior
Dynamics in Systems
Robustness/ Control
Other Intellectual Lineages
Complex Systems Theory
Self-Organization
Complex Adaptive Systems
Emergence
Scaling/ Self-Similarity
Network Science
Global Network Society
Spatial/ Geographical Complexity
Jean Carlson/ John Doyle (Complexity & Robustness)
John Urry (globalization/ social mobilities)
Michael Batty (Cities and Complexity)
David O'Sullivan (GIS, geography & Complexity)
Geoffrey West (Scaling Laws)
Samuel Bowles (Dynamics of Human Systems)
Erich Jantsch (Self-Organizing Universe)
Herman Haken (Self-Organization and Synergetics)
Ilya Prigogine (Dissipative Structure, time, matter)
Warren Weaver (Organized & Disorganized Complexity; Machine Translation)
Paul Cilliers (Philosophy of Complexity)
Duncan Watts (Small Worlds)
Albert-Lazlo Barabasi (Scale-free networks)
Katy Börner (Visual Complexity/ Networks/ Data Science)
John Smith/Chris Jenks (Qualitative Complexity)
Bruno Latour (Actor Network Theory)
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)
Deborah Hammond (history of systems thinking)
Scott Page (Political Science)
Lasse Gerrits (Complexity and Public Policy)
Cybernetics
2nd Order Cybernetics
Systems Science Engineering
Social Systems Theory
Socio-Cybernetics
Complexity Theory/ Epistemology
Economics and Behavioral Dynamics
E-Science
Visual Complexity
Data Science
Heinz von Foerster (2nd order cybernetics)
Jay Forrester (system dynamics)
Social Complexity
Friedrich Hayek (Economy as Complex System)
W. Ross Ashby (Cybernetics of Mind)
Andrei Kolmogorov (Complexity & Information)
Murray Gell-Mann (Effective Complexity)
Seth Lloyd and Heinz Pagels (Thermodynamic Depth)
Francisco Parra-Luna (Founder)
David Byrne (Complex Realism/ Case-based Modeling)
Fernanda Viegas and Martin Wattenbergas (Data Visualization/ Social Data/Visual Arts)
Claude Shannon (Information theory)
John von Neumann (mathematics)
Stephen Wolfram (Complexity)
Christopher Langton (Founder)
Nigel Gilbert (computational social systems)
Joshua Epstein/ Robert Axtell (Artificial Societies)
Data Mining
Brian Castellani/ Rajeev Rajaram (SACS Toolkit)
Walter Pitts (Founder)
Artificial Intelligence
Cognitive Science
Frank Rosenblatt (Connectionism)
Connectionism
Cellular Automata
Computational Complexity Theory
Genetic Algorithms
Artificial Life
Fuzzy Logic
Agent Based Modeling
Robotics/ Multi-Agent Modeling
Computational Modeling
Case-Based Modeling
Gregory Piatetsky-Shapiro (Knowledge Discovery in Databases)
Warren McCulloch (Founder)
Teuvo Kohnen (Self-Organizing Map)
John Conway (Game of Life)
John Holland (Genetic Algorithms)
Lotfi Zadeh (Founder)
Bart Kosko (Fuzzy Logic)
Katia Sycara (Robotics, Multi-agent Systems)
Michael Wooldridge (Editor, AAMAS)
Jeanette Wing (computational thinking)
Emma Uprichard (cluster analysis & complex systems)
Charles Ragin (Founder, QCA, Fuzzy Set Social Science)
Thomas Schelling (micromotives/ Macrobehavior)
Robert Axelrod (Complexity of Cooperation)
Douglas Hofstadter (Cognitive Science)
Melanie Mitchell (computation in complex systems)

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

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

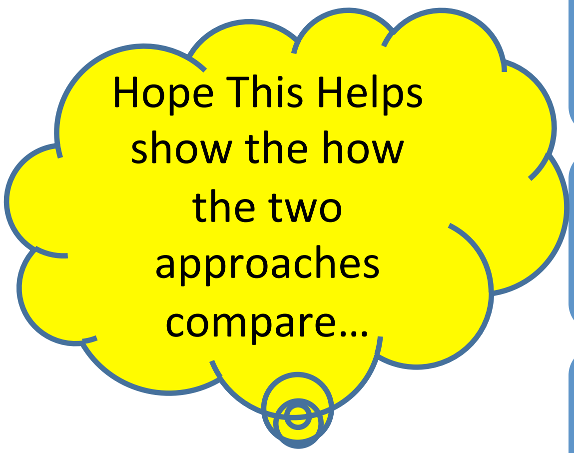
B. Castellani (✉)
Dept. of Sociology, Kent State University, Ashtabula, OH 44004, USA
e-mail: bcastel3@kent.edu

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

Comparing SACS vs. MBSE

MBSE Tools:

- Requirements Management
- Swimlanes - Activity Diagrams
- Business Processes
- A Project
- Non-linear Technical Only Models (e.g. Simulink / FEA) - Typically not showing emergence



SACS Toolkit:

- Active Data Management
- Structure/Agency Dualism
- Social Practices
- A Case
- Non-linear Socio-Technical Models sensitive to Initial Conditions and capable of showing emergent behavior

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.

This is our first simplifying assumptions

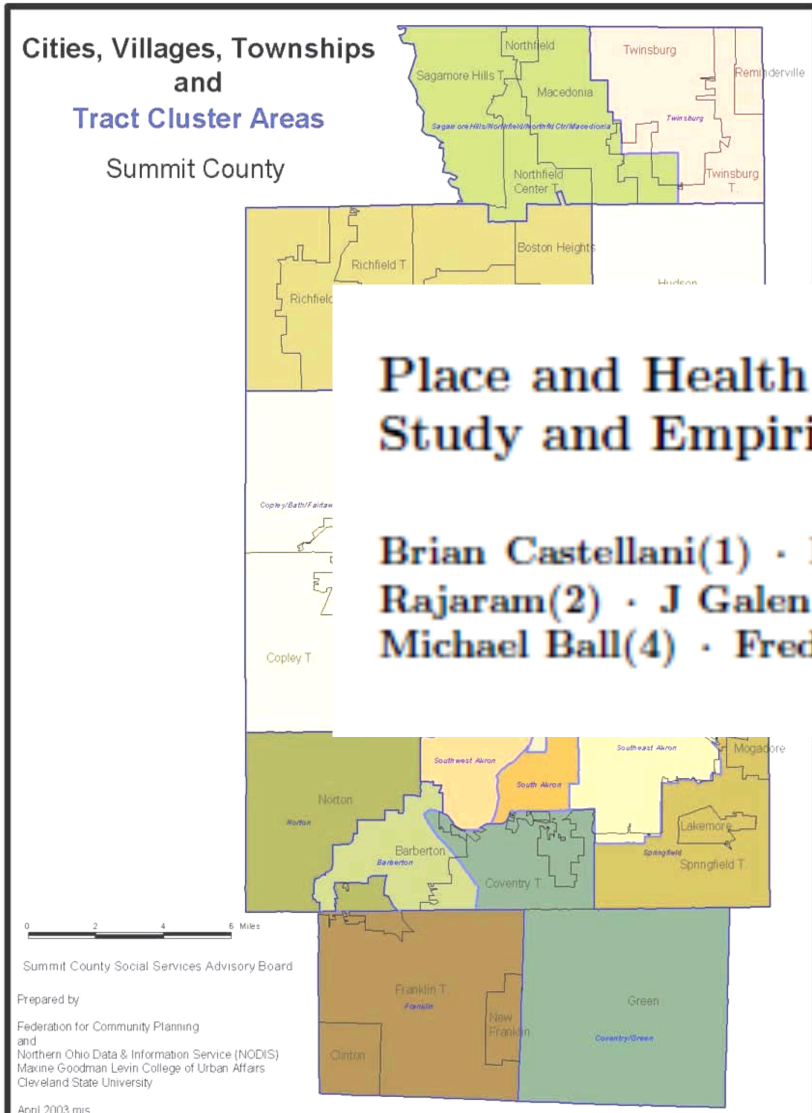
aximum number of cases that can be ... , practically speaking the upper limit will be bounded, base on the particular set of cases identified for study—which is always an empirica issue.

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

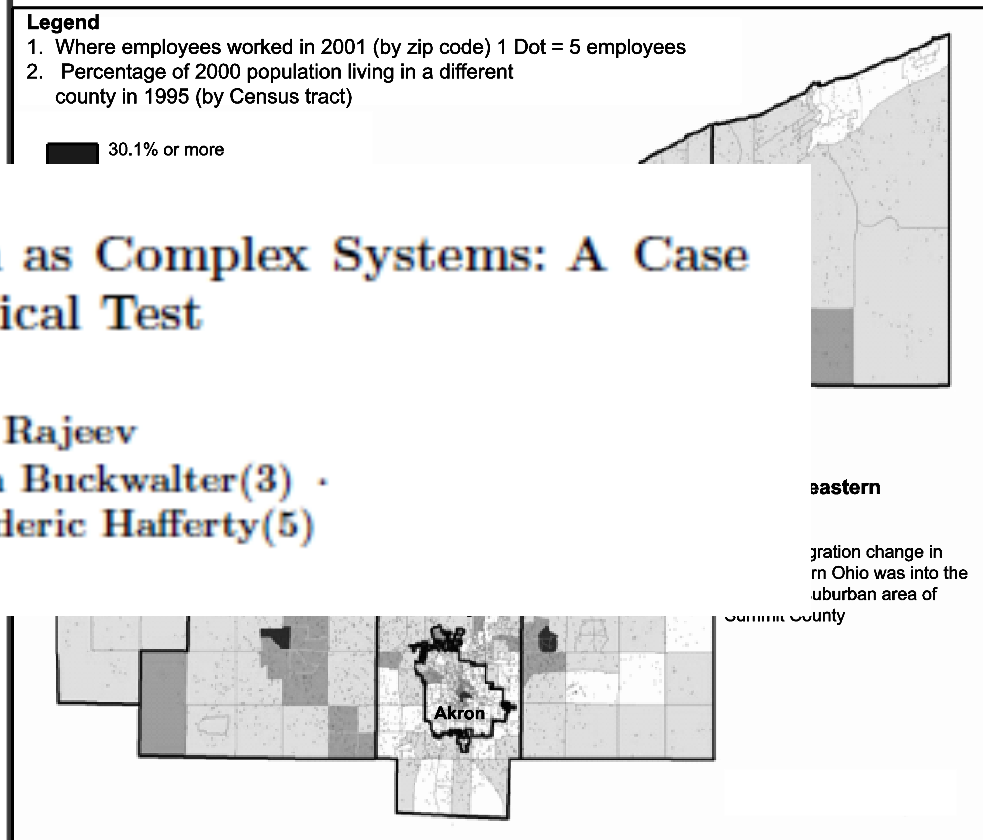
(5) Each case c_i in S is a k dimensional row vector $c_i = [x_{i1}, \ldots, x_{ik}]$, where each x_{ij} represents a measurement on one of the variables being used to model a complex system.

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)

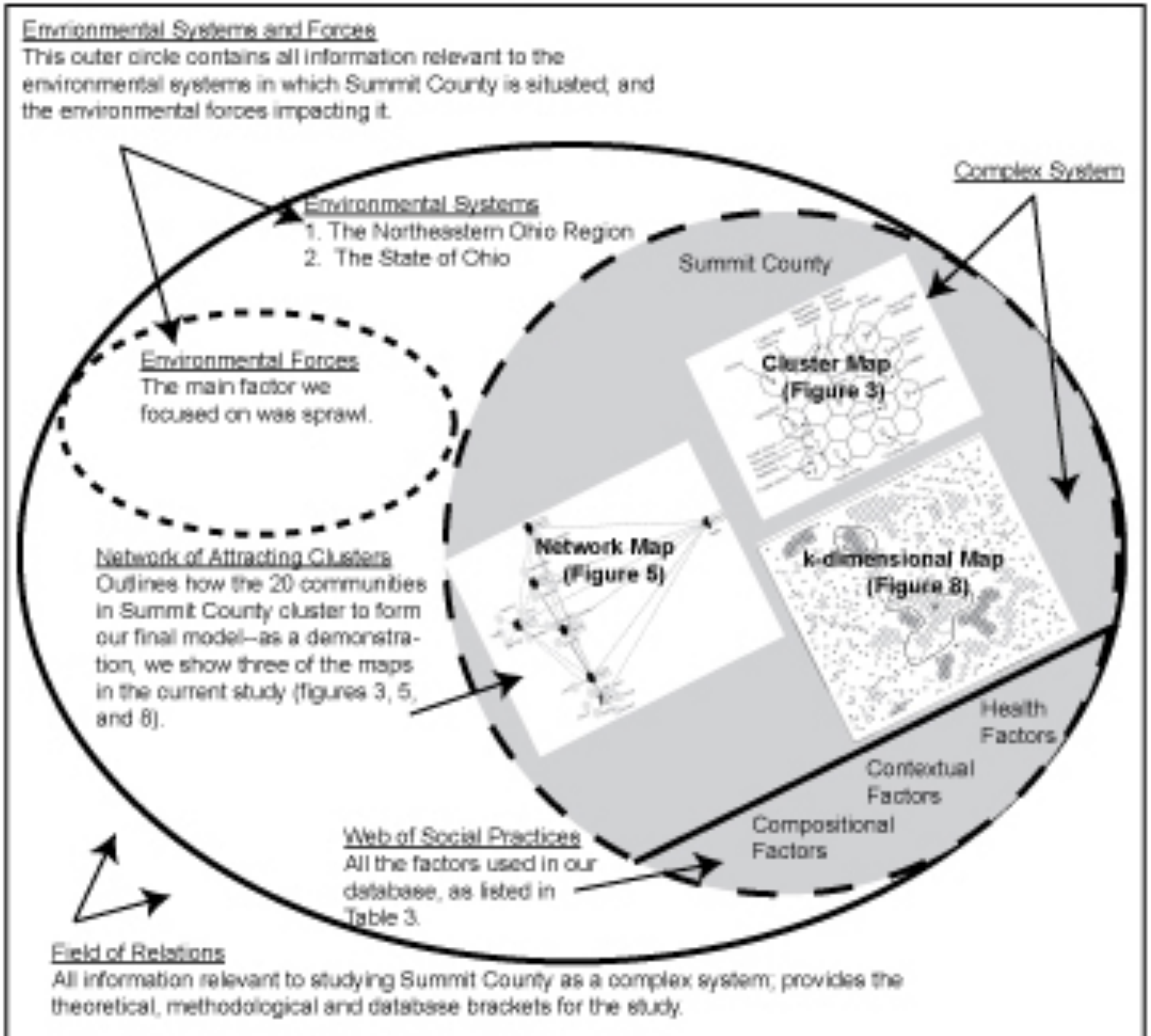


MAP 2
Change in Residential Mobility Between 1995 and 2001



Source: This map was retrieved from http://www.healthysummit.org/qol/pdf/final draft 2009 data tracking report 0216.pdf on the 29th of November, 2011. It is a public document provided by the Health Summit 2010 website. Its original source is ES-202 (ODJFS); NODIS, U.S. Census Bureau, 2000.

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] • % 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 the U.S. Census) • Persons 25+ Years with High School Diploma[1] • Net Job Growth[3] (Defined as the number of jobs in 2000 minus the number of jobs in 1990. • Unemployment Rate[1] (Defined as unemployed civilian labor force) • Housing affordability[1] (Defined as the percentage of households where mortgage/rent is greater than 30% of the household income) • 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) • Teen Birth Rate[4] (Defined as the number of births occurring between 1995-1998 to mothers 15 to and including 17 years of age) • 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

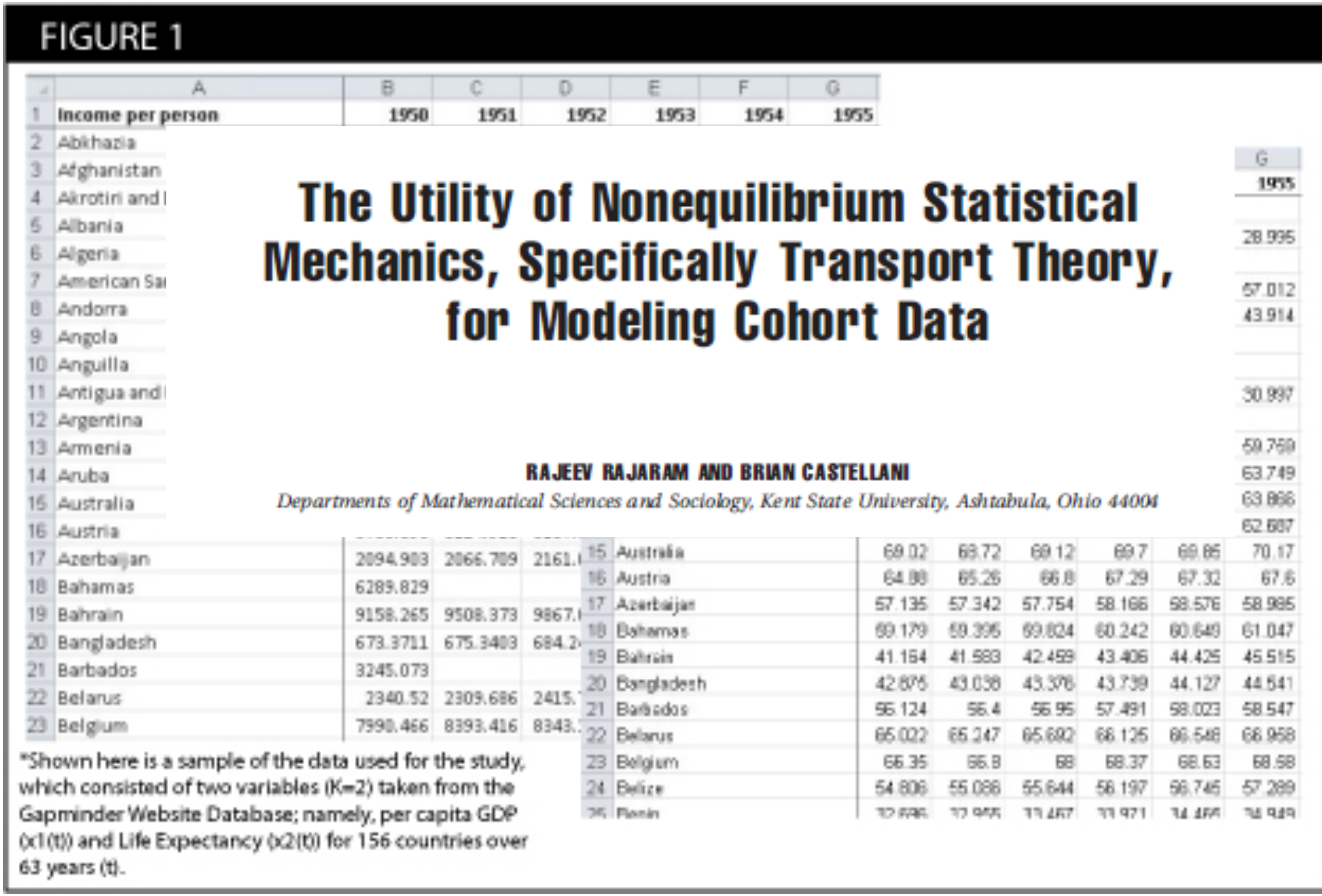


FIGURE 1

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

	A	B	C	D	E	F	G
1	Income per person	1950	1951	1952	1953	1954	1955
2	Abkhazia						
3	Afghanistan						
4	Akrotiri and						
5	Albania						
6	Algeria						
7	American Sa						
8	Andorra						
9	Angola						
10	Anguilla						
11	Antigua and						
12	Argentina						
13	Armenia						
14	Aruba						
15	Australia						
16	Austria						
17	Azerbaijan	2094.903	2066.709	2161.			
18	Bahamas	6289.829					
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.			

G
1955
28.995
57.012
43.914
30.997
59.759
63.749
63.866
62.687

15	Australia	69.02	68.72	69.12	69.7	69.86	70.17
16	Austria	64.88	65.26	66.8	67.29	67.32	67.6
17	Azerbaijan	57.135	57.342	57.754	58.166	58.576	58.985
18	Bahamas	59.179	59.395	59.824	60.242	60.649	61.047
19	Bahrain	41.164	41.583	42.459	43.406	44.425	45.515
20	Bangladesh	42.675	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

*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

FIGURE 1

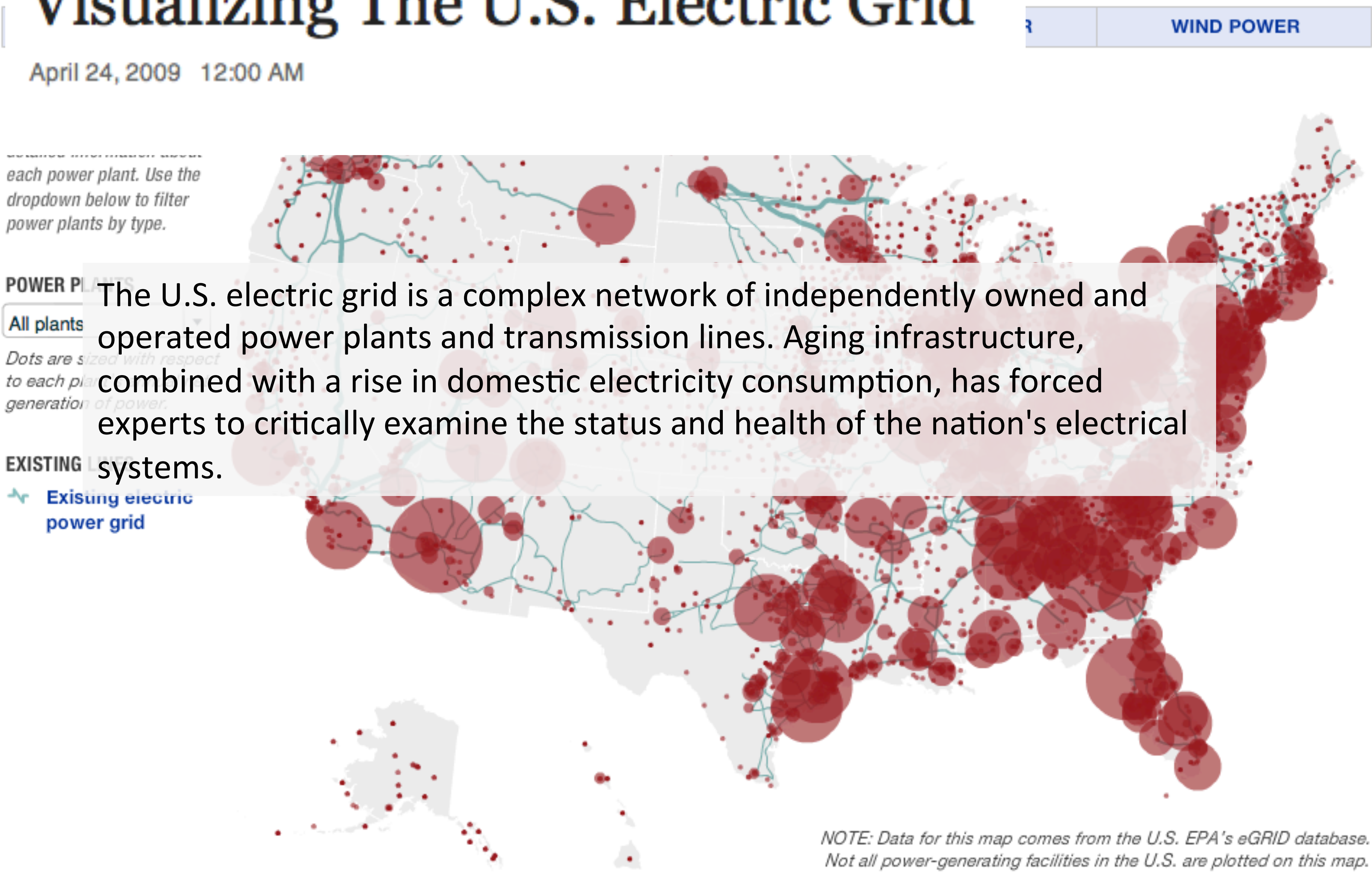
	A	B	C	D	E	F	G
1	**Income per person**	**1950**	**1951**	**1952**	**1953**	**1954**	**1955**
2	Abkhazia						
3	Afghanistan	757.3188	766.7522	779.4			
4	Akrotiri and Dhekelia						
5	Albania	1532.354	1598.493	1601.			
6	Algeria	2429.214	2397.531	2449.			
7	American Samoa	4465.145					
8	Andorra						
9	Angola	3363.022	3440.901	3520			
10	Anguilla						
11	Antigua and Barbuda						
12	Argentina	6252.859	6362.126	5911.			
13	Armenia	1366.372	1346.227	1405.			
14	Aruba						
15	Australia	10031.12	10160.74	1003			
16	Austria	5733.098	6124.928	6137.			
17	Azerbaijan	2094.903	2066.709	2161.			
18	Bahamas	6289.829					
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**
2	Abkhazia						
3	Afghanistan	26.674	26.932	27.448	27.964	28.48	28.995
4	Akrotiri and Dhekelia						
5	Albania	54.191	54.399	54.876	55.471	56.184	57.012
6	Algeria	42.089	42.283	42.678	43.081	43.493	43.914
7	American Samoa						
8	Andorra						
9	Angola	29.209	29.407	29.804	30.201	30.599	30.997
10	Anguilla						
11	Antigua and Barbuda	57.536	57.786	58.284	58.779	59.271	59.759
12	Argentina	61.418	61.728	62.32	62.855	63.331	63.749
13	Armenia	61.965	62.178	62.602	63.024	63.446	63.866
14	Aruba	58.419	58.962	60.01	60.98	61.873	62.687
15	Australia	69.02	68.72	69.12	69.7	69.85	70.17
16	Austria	64.88	65.26	66.8	67.29	67.32	67.6
17	Azerbaijan	57.135	57.342	57.754	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

*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).

Visualizing The U.S. Electric Grid

April 24, 2009 12:00 AM



The U.S. electric grid is a complex network of independently owned and operated power plants and transmission lines. Aging infrastructure, combined with a rise in domestic electricity consumption, has forced experts to critically examine the status and health of the nation's electrical systems.

1. The Power Grid is a Complex System of Multiple Networks
As a major study on the American energy grid recently explained (See *Nature Physics*, Aug 2013), a complex system such as the power grid is a dense web of interconnected and interdependent networks—each a different type:

There are, for example, the human social networks among key players in the power industry (CEOs, engineers, technicians, etc).

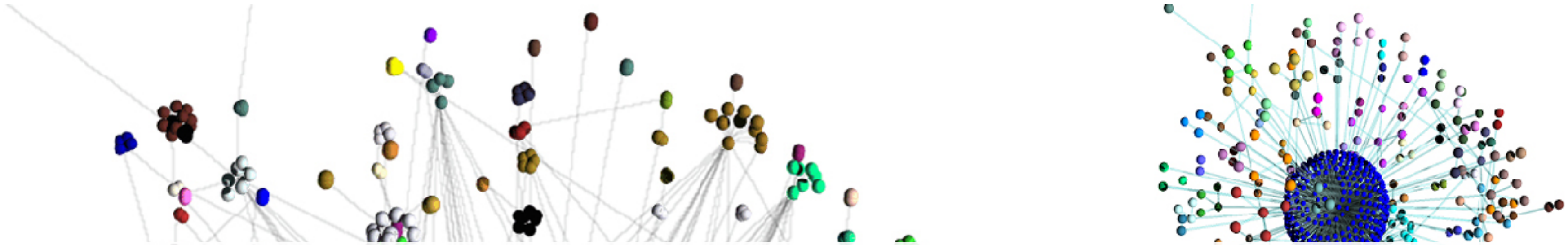
There are also the organizational networks among the institutions that run the power grid, including distributors, transmission companies, generating companies, etc.

There are inter-organizational networks within companies, which exist among, between and across different people and departments.

Then there is the actual power grid equipment, comprised of the cables, generators, power stations, etc that generate and provide power to people.

Finally, there are the environmental networks in which the power grid is situated, including local ecosystems and weather.

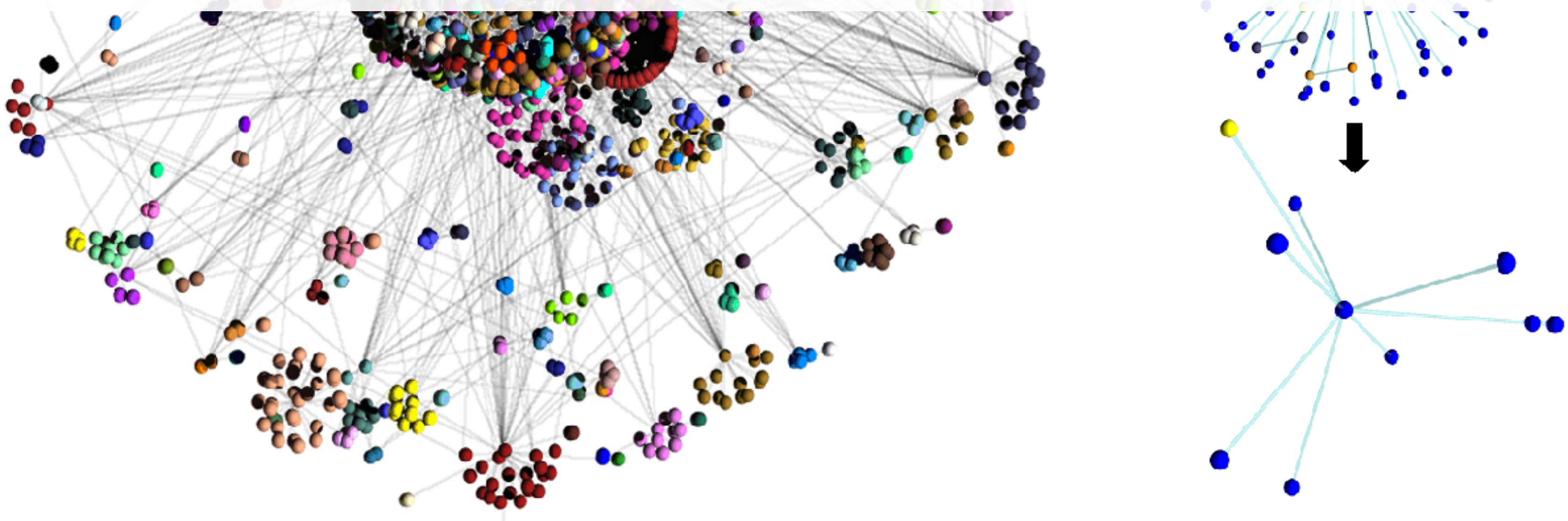
Beyond this web of networks are also the political and economic networks in which the power industry is situated, from local towns and communities to the larger business community and economy to cities, states and the federal government.



2. A Complex System of Embedded Networks is Highly Sensitive to Power Failure

As the aforementioned major study in *Nature Physics* also found, such a complex system is highly sensitive to violations, misoperations, errors and risks—which can lead to system-wide power outages and collapse.

This sensitivity is due, in large measure, to the complexity of this system and the embedded nature of its multiple networks (listed above), each overlapping and influencing the others.



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).

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

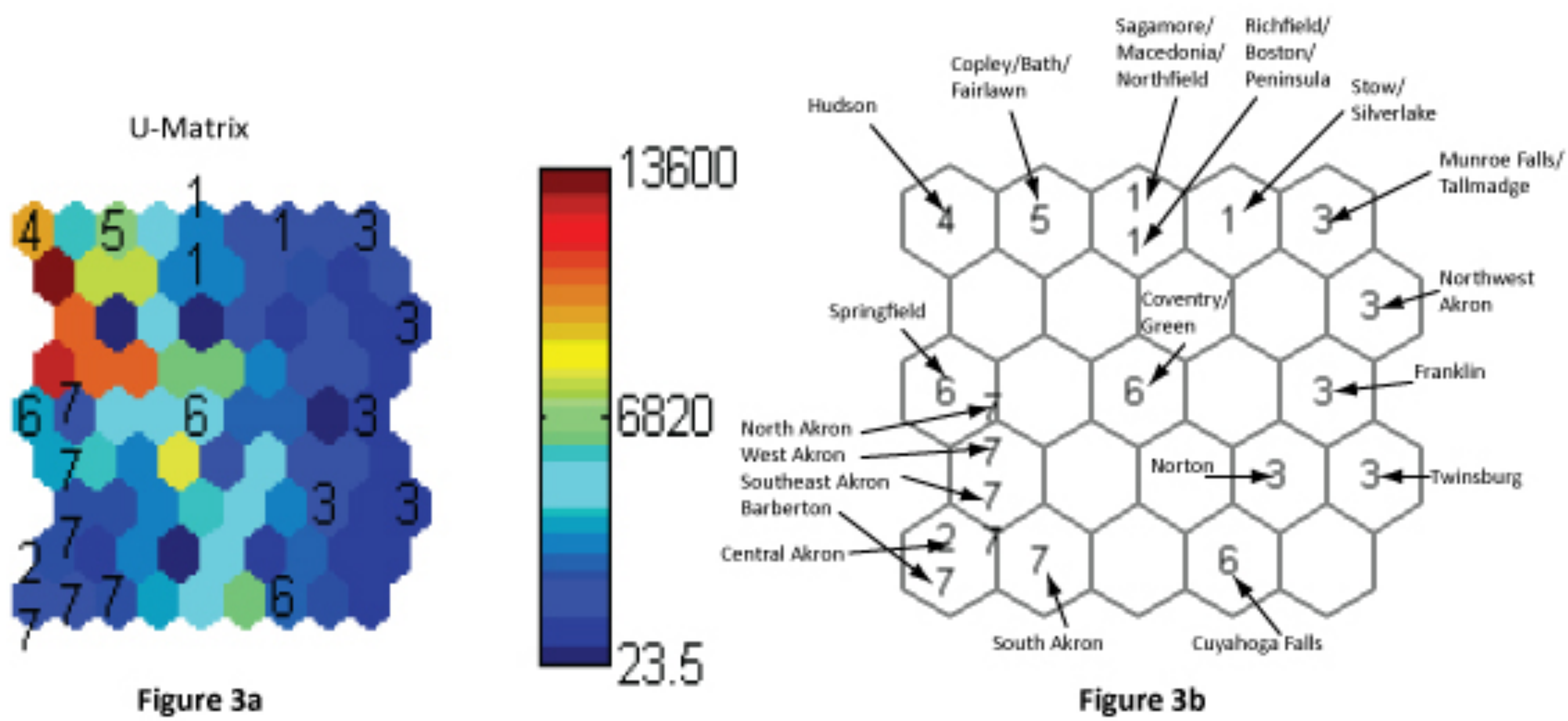
Clustering and grouping to search for major and minor configurations/ profiles and trajectories (discrete or continuous)

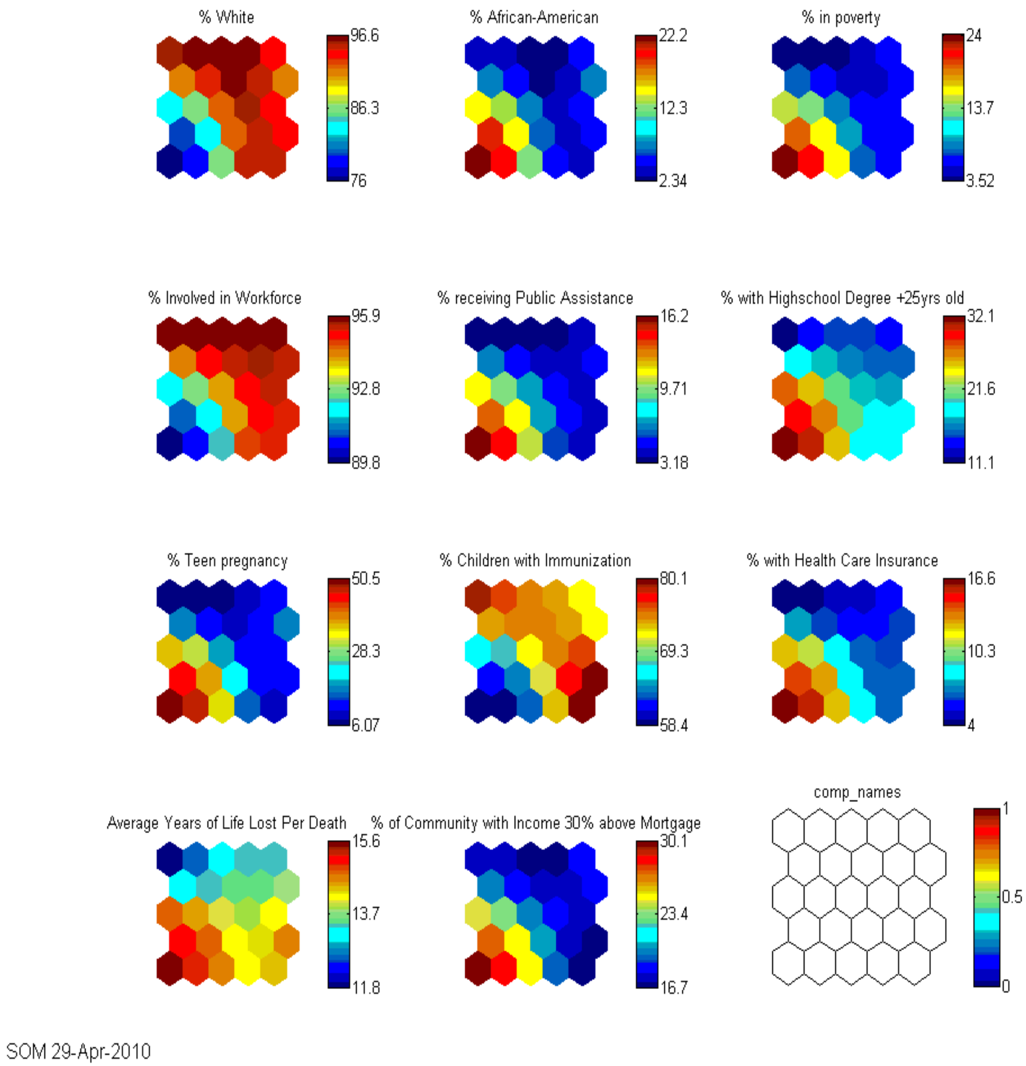
TABLE 3
Final K-means Cluster Solution for 20 Communities in Summit County

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.

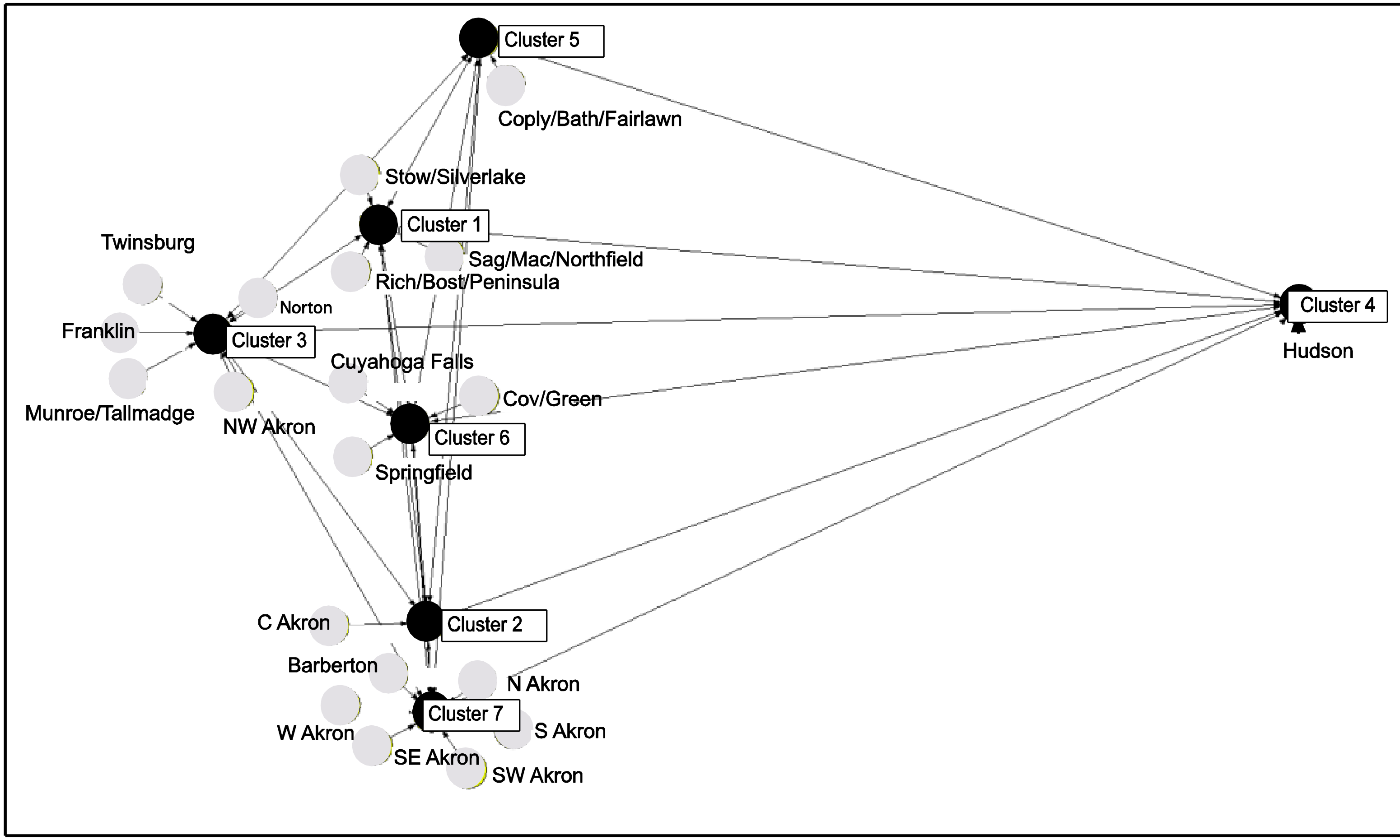
FIGURE 3: Final Cluster Centers Solution of 20 Communities in Summit County Using SOM





These figures show how the 17 factors in this study impact the topic – blue is low impact, and red is high impact

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



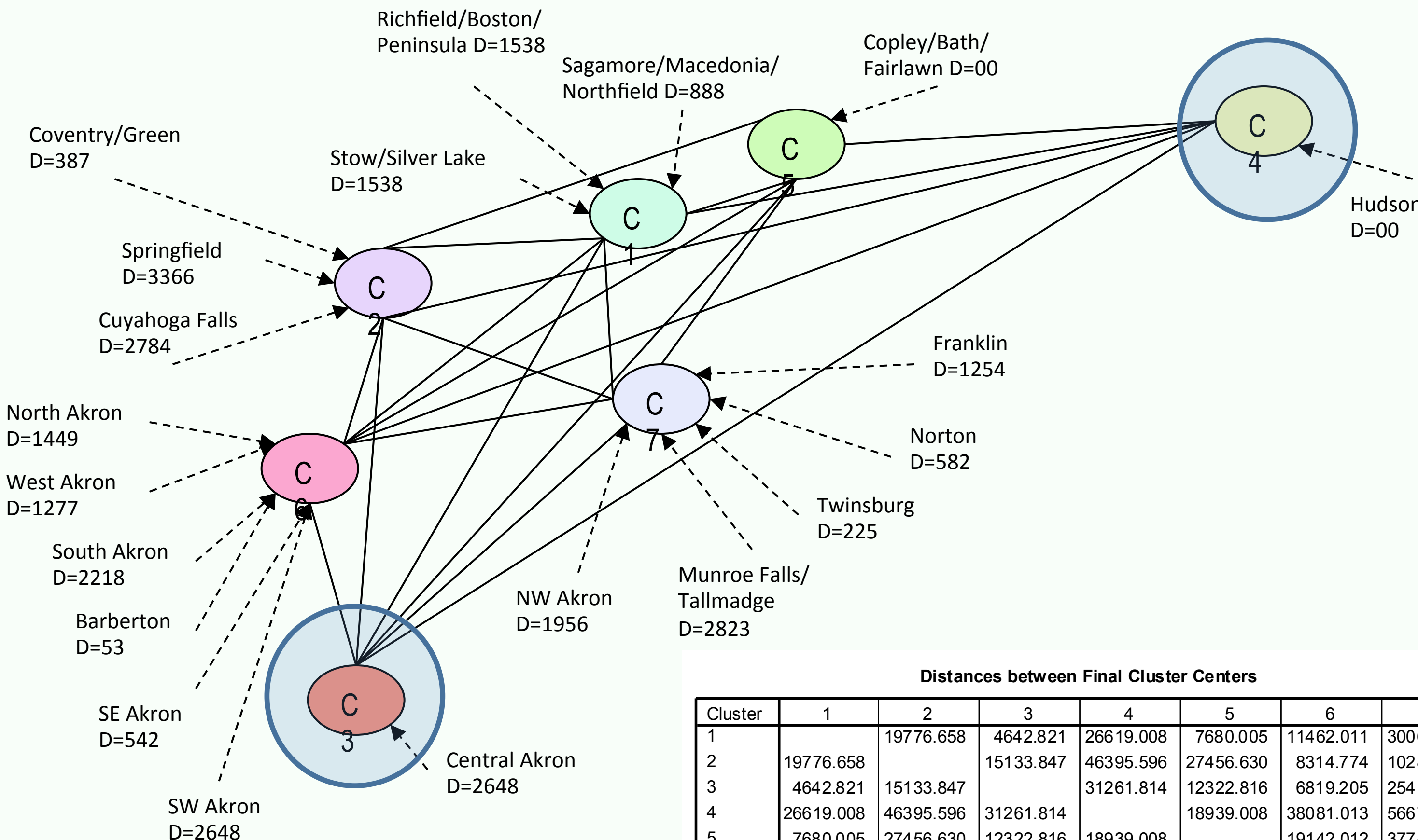
NOTE: Distances between clusters are based on Euclidian distances arrived at through k-means analysis. Distances within clusters for each community are based on within-cluster measures. All measures are non-standardized.

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

COMMUNITY	YEAR 1990 Cluster Membership	YEAR 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?

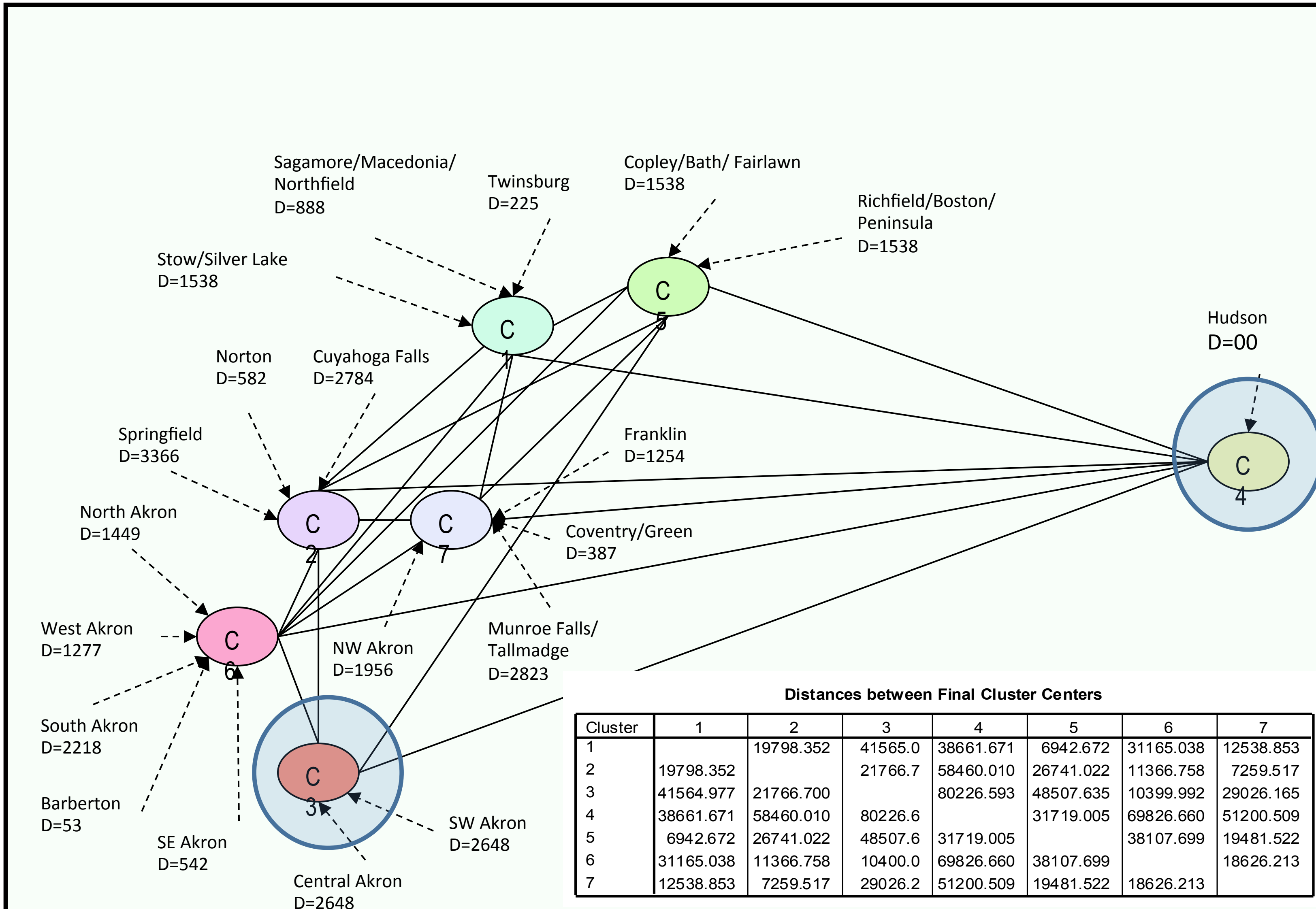


Distances between Final Cluster Centers

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)

Network of Attracting Clusters for 1990 with Summit County

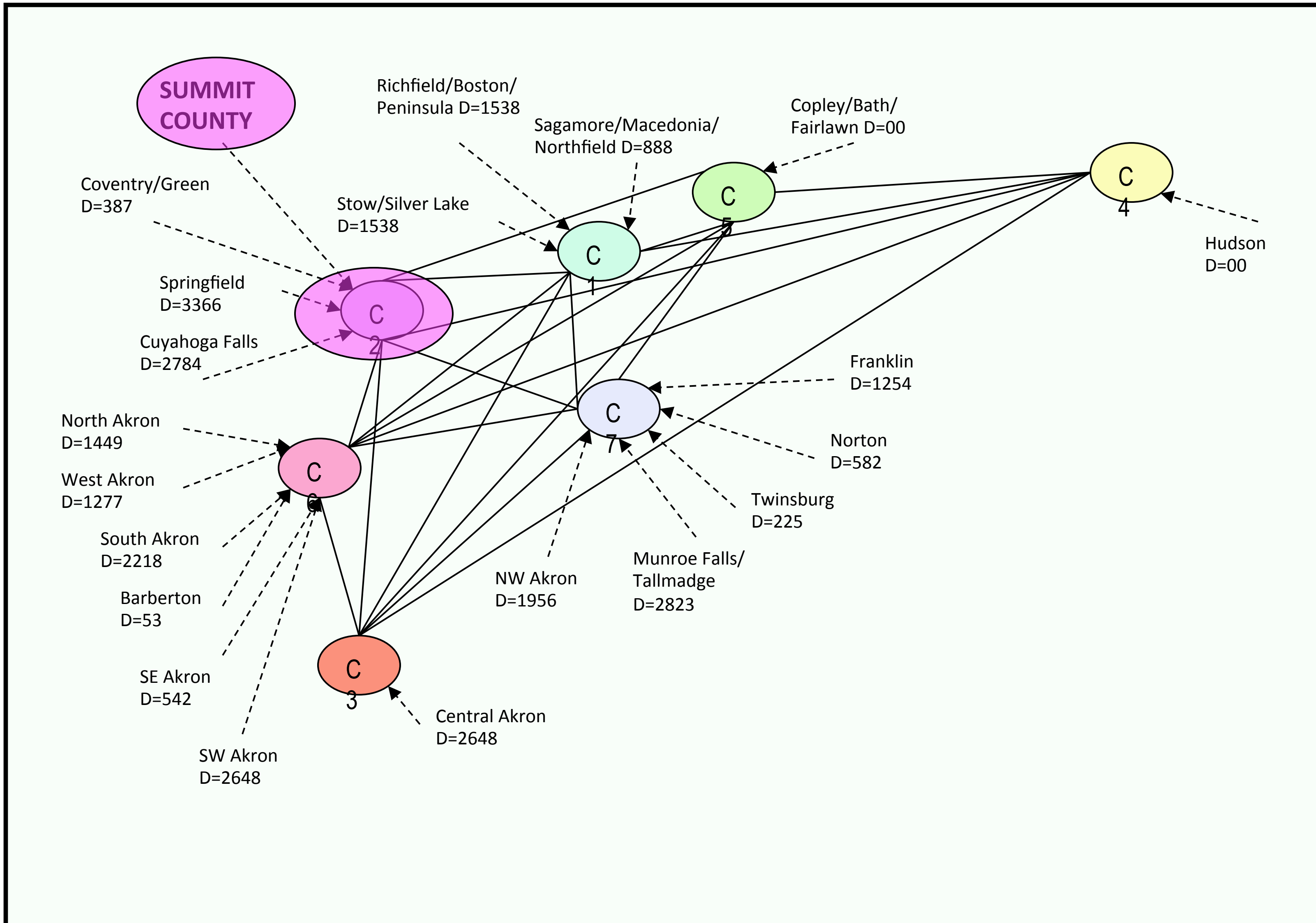


Figure 7: Summit-SIm Dashboard in Netlogo

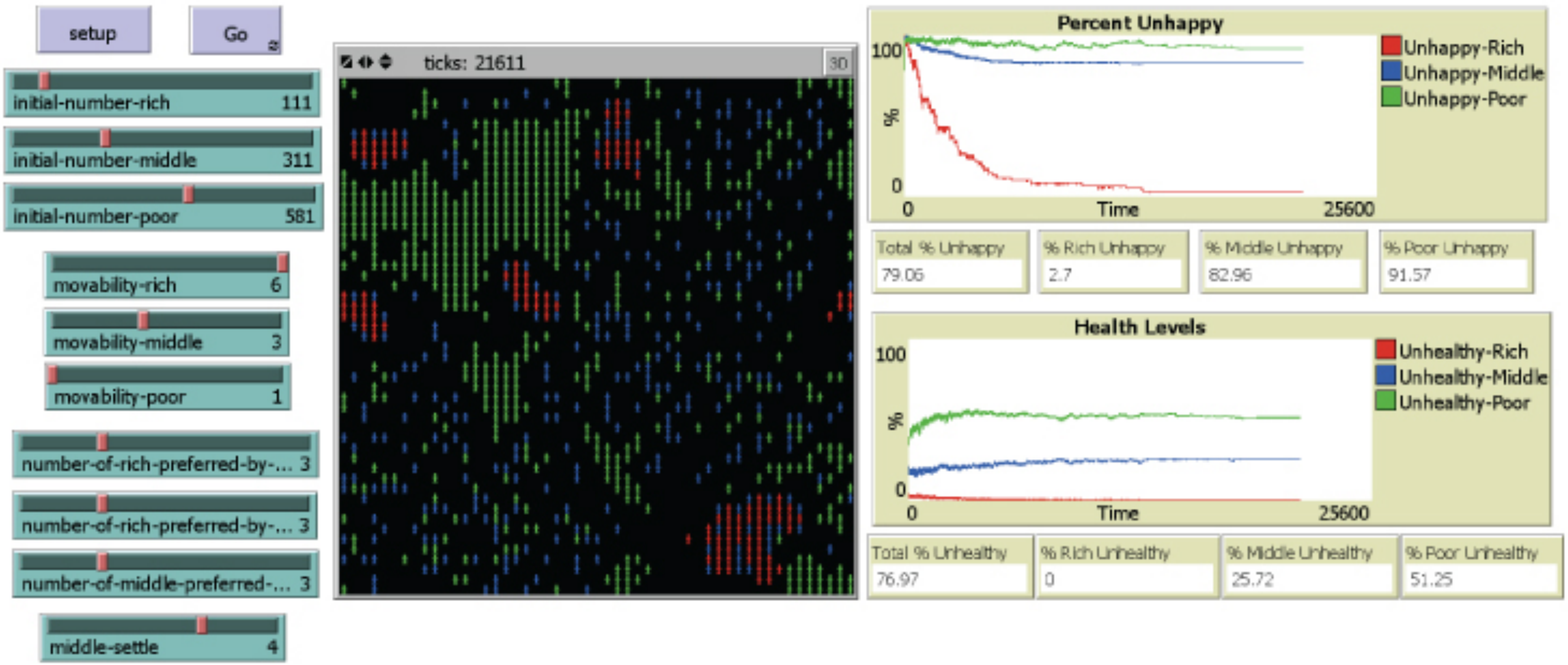
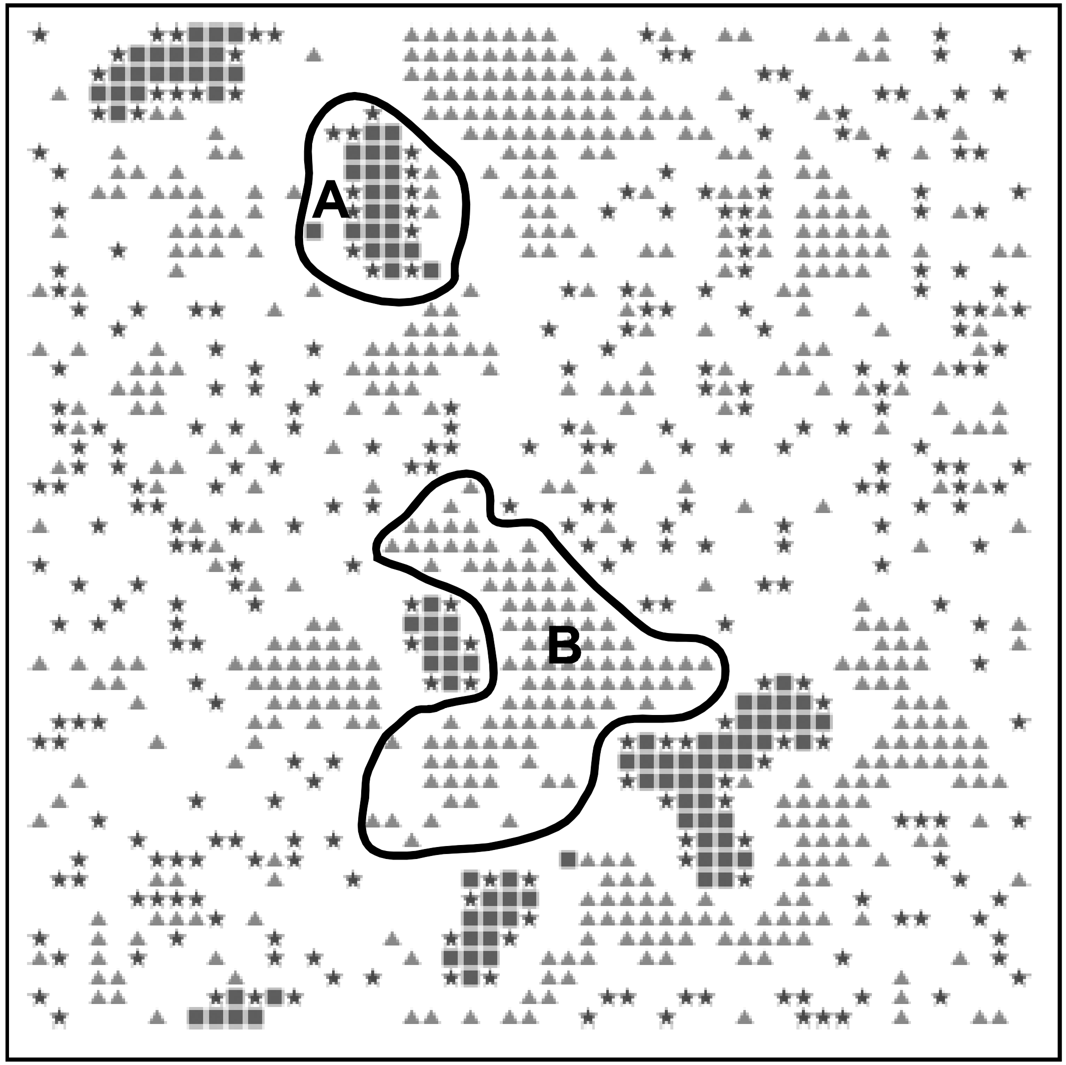
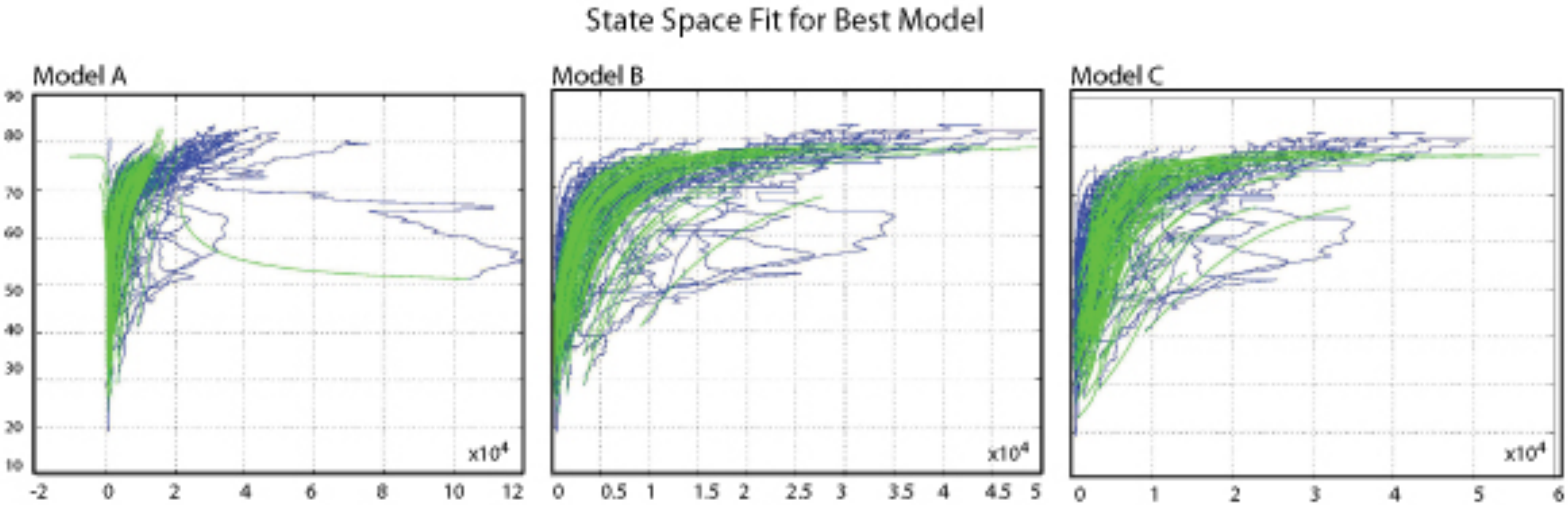


FIGURE 7:
Snapshot of SummitSim with a Preference Rating of 3 for all Agents



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.

FIGURE 4



State space fit for the best model.

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

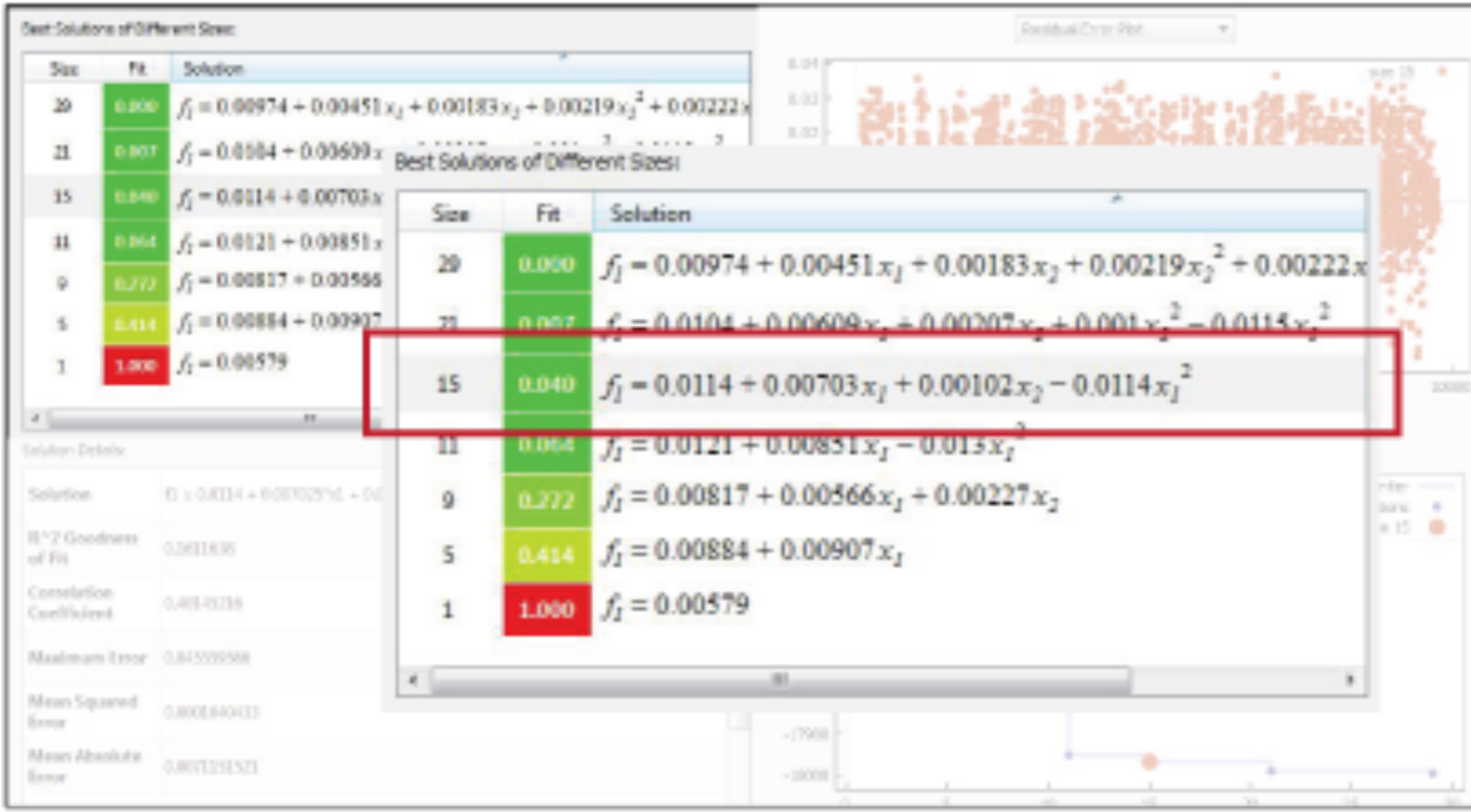
State space fit for the best model without Kuwait or Luxembourg, but with *time* as an independent variable.

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

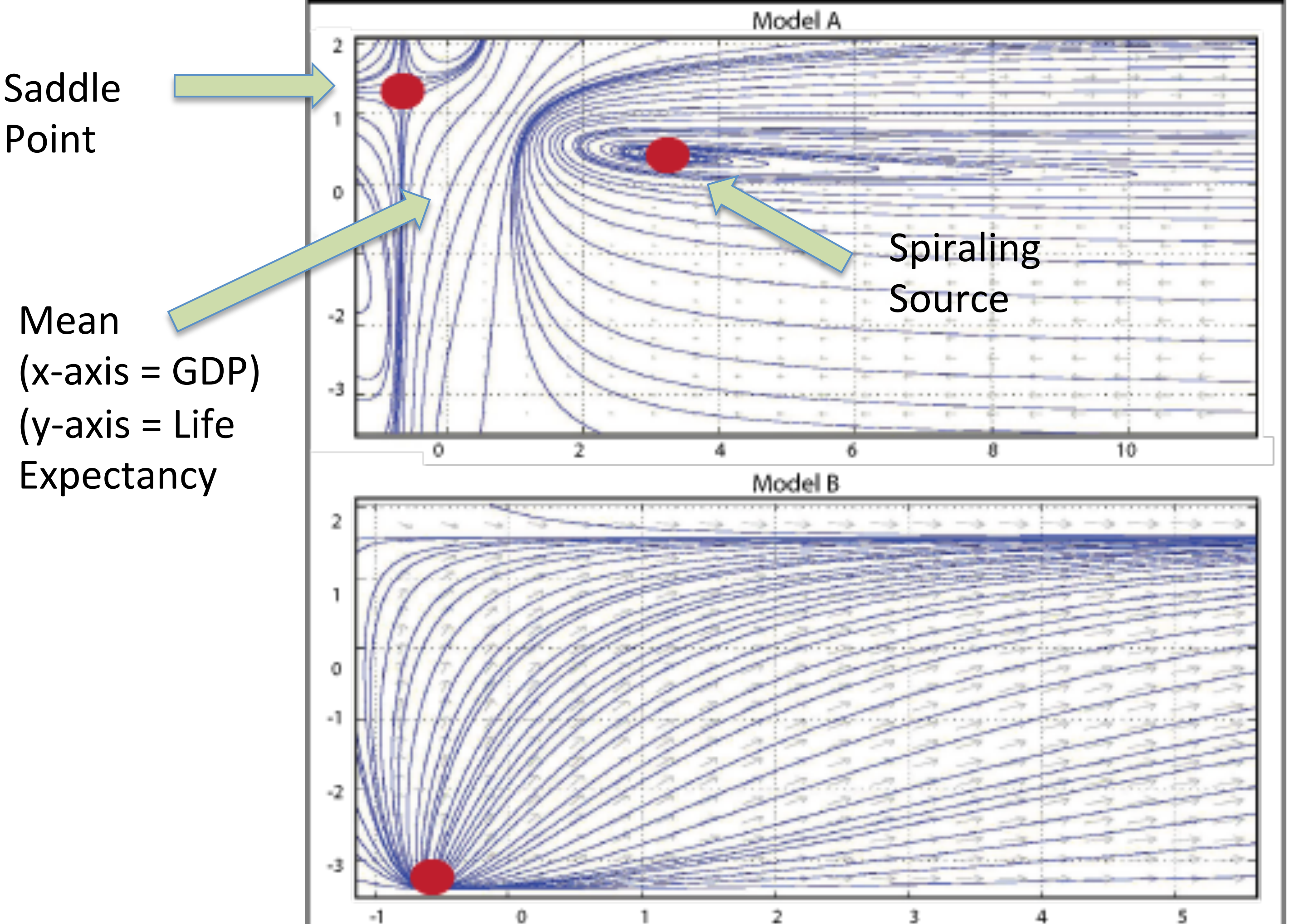
Third, we perform this translation by fitting the time trajectories of these cases using an autonomous (and, in some instances nonautonomous) ordinary differential equation (ODE) (1). In most cohort studies, be they network studies or otherwise, the laws governing their macroscopic dynamics are not known [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

- 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

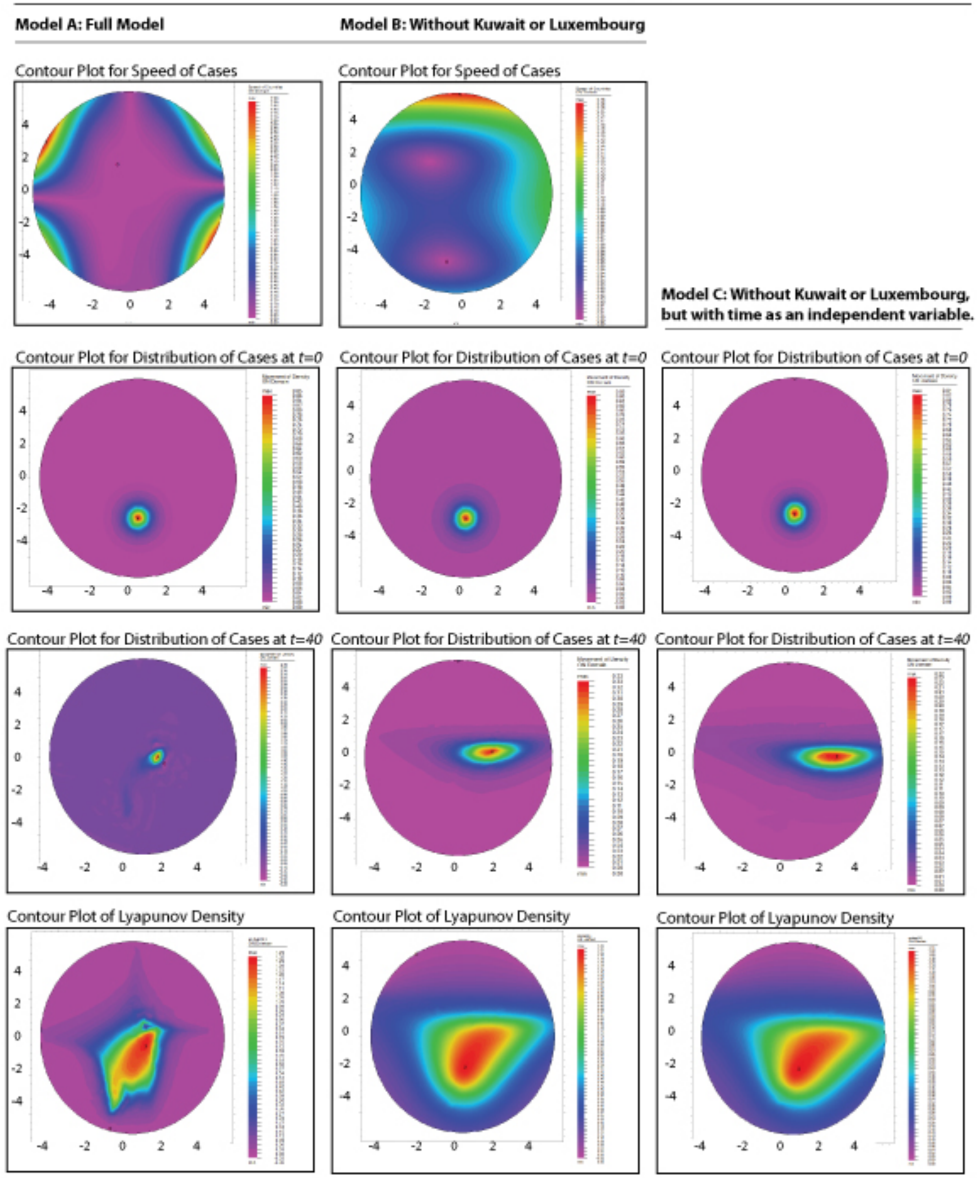
$$\rho_t + \nabla \cdot (f\rho) = 0; \quad \rho|_{\Gamma_1} = 0; \rho(x,y,0) = \rho_0(x,y).$$

- Tr f
  ca
- R field
  of on of
  th
  (c
- U pill),
  tr

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

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 converteted 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)

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

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

Strengths

- Prediction of 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!

For more information, visit me at:

http://www.personal.kent.edu/~bcastel3/brian%20castellani.html