Proceedings of the 2010 Winter Simulation Conference
B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Yücesan, eds.

EFFECTIVE REAL-TIME ALLOCATION OF PANDEMIC INTERVENTIONS

Catherine Dibble

Aiki Labs, Ltd.
5505 Connecticut Ave, NW, #375
Washington, DC 20015, USA

ABSTRACT

We address the integration of computational laboratories, spatial agent-based simulation, and real time situation updates to provide pandemic risk assessments and optimal intervention and prevention strategies. Our goal is to support decisions that save lives by helping to integrate real-time feedback and coordinate effective responses. Computational laboratories using super computing resources allow us to explore and optimize deployments of scarce resources and disruptive interventions for controlling pandemic influenza. We have developed an agent based model for simulating the diffusion of pandemic influenza via carefully calibrated inter-city airline travel. This and related simulation models at community scales can be used to learn vital lessons based on CPU-intensive virtual experience from millions of simulated pandemics. Real-time situation updates can greatly enhance the strategic usefulness of simulation models by providing accurate interim conditions for adapting effective deployments of interventions as a pandemic unfolds.

1 INTRODUCTION

> "Floods of data are pounding down all around us in torrents. How will we cope?" (Gelernter 1992, page 107)
>
> "There's a great deal at stake. … At a deeper and more general level, *power requires control*. The power and complexity of our technical infrastructure is exploding, and our control systems have not kept pace. When you double an engine's horsepower, you'd better improve your brakes, tires, suspension and steering as well. Interpreting this data instead of ignoring it is our main chance of beefing up the brakes and the steering. Without adequate control systems, we face real danger." (Gelernter 1992, page 112)

Tightly coupled global systems of twentieth-century transportation and communication technologies ensure that cascading effects of economic, public health, and environmental disasters propagate rapidly with potentially dire consequences for citizens of many countries. Winter Simulation Conference and related venues provide opportunities to explore and evaluate state-of-the-art computational systems for sensing, simulating, and guiding such system-of-systems cascading dynamics toward beneficial or at least less harmful directions.

2 SIMULATION MODELING FOR RISK ANALYSIS AND PREPARATION

While it does not yet spread easily among humans, mutation among flu strains is likely to generate a human-adapted variant capable of causing a worldwide influenza pandemic, worse than the pandemic of 1918-1919. Recent projects have developed fine-grain computational models with each individual mod-

eled separately, interacting with peers according to local contact processes (NIH 2007). Germann et al. (2006) and Ferguson et al. (2006), both presented models of the entire United States calibrated to community level demographics and behavior, providing a previously impossible level of detail. However, their computational requirements are immense and their complexity makes it difficult to study the sensitivity of the model to its assumptions and input parameters.

The Institute of Medicine of The National Academies of Science evaluated the usefulness of detailed community-level simulation models for effective community containment of pandemic influenza (Institute of Medicine 2006). The report emphasized the importance of complementary simulations with simpler models rich enough to incorporate surveillance information, yet fast enough to provide policy-relevant updates and risk evaluations during a pandemic (Institute of Medicine 2006, page 13). Integration of computational laboratories, spatial agent-based simulation, and real time situation is needed to provide pandemic risk assessments and optimal intervention and prevention strategies (Figure 1).

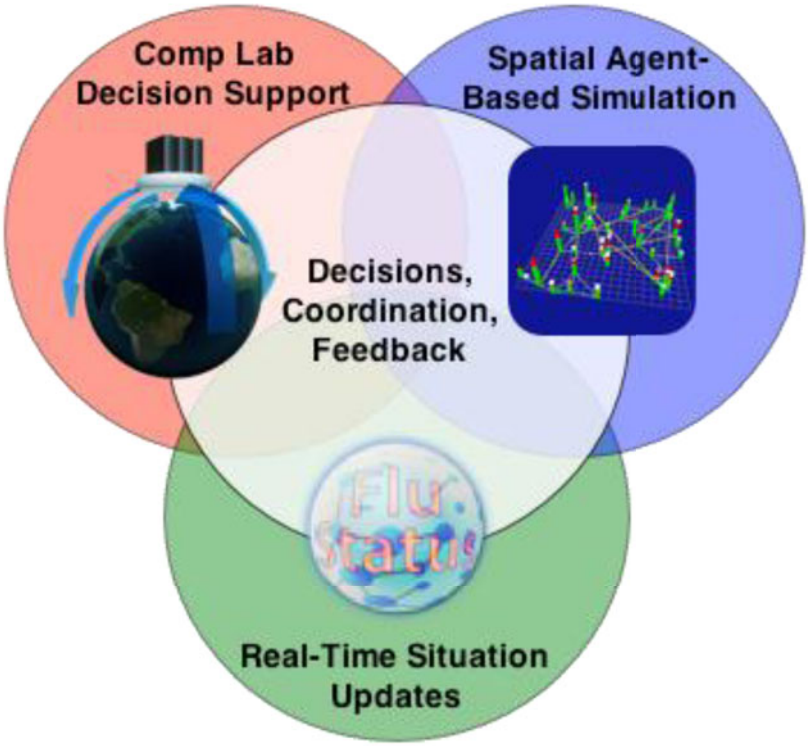

Figure 1: Integration of Spatial Agent-Based Simulation Models, Computational Laboratory Spatial Evolutionary Optimization, and Real-Time Situation Updates for effective allocation of disaster interventions.

Here we present a model that complements the high-resolution models used to evaluate community level containment measures. We use calibrated airline networks and travel patterns to simulate the spread of pandemic influenza among major cities in the continental US, focusing on the rapid, hierarchical diffusion of pathogens. The simplicity of our model supports many thousands of replications for full sensitivity analyses, clearly distinguishing relative pandemic risks among US cities. These patterns are remarkably stable with respect to the geographic location of initial pandemic cases, and understanding these patterns will help direct the allocation of scarce resources, thus facilitating appropriate community preparation.

2.1 Inter-City Model of Pandemic Influenza

We developed a compact, efficient agent-based model of the geographic spread of influenza within the United States and placed it in a computational laboratory to evaluate the impact of the uncertainty resulting from epidemiological characteristics, stochastic travel and transmission behavior.

There exists a rich history of work on the spread of disease, including considerable recent work with agent-based simulation models. The widely used SEIR model provides differential equations which approximate disease spread (Anderson and May 1991). The SEIR model predicts the mean quantity of four categories of individuals--Susceptible, Exposed, Immune, and Recovered--who interact over time within a single perfectly mixed population. Keeling and Grenfell (2000) examined the role of discretization in SEIR models, finding that discretization more accurately approximates disease spread. Lloyd and May (1996) examined the impact of spatial heterogeneity and stochasticity on the spread of the disease, similarly verifying the importance of these extensions.

The epidemiological component consists of a discretized, spatialized, agent-based SEIR simulation model in which agents track carefully calibrated travel behavior, each having a set of internal clocks which govern the progression of disease phases from susceptible through exposed, infectious, and recovered or removed. These agents are proportionally allocated to the fifty most populous metropolitan areas of the United States. Agent travel among these areas is calibrated according to a representative sample of US air travel behavior. Within each metropolitan area, the agents interact randomly with one another. The model thus builds on previous work on stochastic, discrete, spatial SEIR models, but adds the capacity for agent heterogeneity and customized travel behavior.

The fifty metropolitan areas included in the model represent roughly half of the US population, or 150 million agents. By testing the sensitivity of the model's results to the number of agents, we found that a population as small as 100,000 agents provided unbiased results.

For each simulation we introduce 1 flu case per city, then model flu diffusion within and between cities via airline travel of infected individuals not yet too sick to travel. Diffusion within cities is agent-based and calibrated to other MIDAS community level papers and mathematical SEIR models.

For each of the fifty cities, 5,000 pandemic flu simulations (with 100 random number seeds to simulate quirks of human behavior) is an initial starting point for the pandemic in the continental US. Slightly less than half (2,335) of the simulations resulted in epidemics with an incidence greater than 1% (> 1,000 cases).

2.2 Relative Pandemic Risks of US Cities

We analyzed pandemic risks for simulations, resulting in pandemic influenza affecting more than 1,000 agents and spreading to at least two other cities. We defined pandemic risk for each city according to the average number of infected individuals arriving at that city during each pandemic, averaged across all simulated pandemics, producing 51 rankings of the 50 cities. One ranking is for overall pandemic risk, averaged across all the epidemics, regardless of where the initial case was introduced. This establishes overall pandemic risk when the location of initial case(s) is yet unknown. Each of the other 50 rankings shows the relative risks for each city depending on the location of the initial case, ranking the relative risks for each of the US cities where the first outbreak occurs.

Significantly, *regardless where the pandemic begins, the relative risks across cities remain stable*. This is clear in Figure 2, where minor changes exist in the rank order, but where the membership of the top 15 high-risk cities is consistent, with only one city in the bottom of the origin-specific rankings. The Spearman Rank Correlation Coefficient (SRCC) across city rankings was ~98% (see Figures 3 and 4).

We also performed sensitivity analyses on the density of links in the network (for minimum 10K, 1K, and 10 passengers per year for each airline route included in the network), the numbers of random number (Monte Carlo) repetitions for each scenario, travel probabilities, incubation period, and the standard reproductive number R_0 (i.e. the expected number of secondary cases from the first case in a fully-mixed susceptible population).

Initial City Unknown	Initial City Los Angeles	Initial City New York	Initial City Washington	Initial City Atlanta
Los Angeles, CA	Los Angeles, CA	Los Angeles, CA	Los Angeles, CA	Los Angeles, CA
New York, NY	New York, NY	New York, NY	New York, NY	New York, NY
San Francisco, CA	San Francisco, CA	Chicago, IL	San Francisco, CA	Washington, DC
Chicago, IL	Washington, DC	San Francisco, CA	Chicago, IL	Chicago, IL
Washington, DC	Chicago, IL	Washington, DC	Washington, DC	San Francisco, CA
Miami, FL	Miami, FL	Miami, FL	Miami, FL	Miami, FL
Dallas, TX	Dallas, TX	Dallas, TX	Atlanta, GA	Dallas, TX
Atlanta, GA	Atlanta, GA	Atlanta, GA	Dallas, TX	Seattle, WA
Seattle, WA	Boston, MA	Boston, MA	Boston, MA	Atlanta, GA
Boston, MA	Seattle, WA	Seattle, WA	Seattle, WA	Boston, MA
Philadelphia, PA	Denver, CO	Houston, TX	Houston, TX	Philadelphia, PA
Houston, TX	Houston, TX	Philadelphia, PA	Philadelphia, PA	Houston, TX
Denver, CO	Philadelphia, PA	Detroit, MI	Minneapolis, MN	Denver, CO
Detroit, MI	Phoenix, AZ	Phoenix, AZ	Phoenix, AZ	Phoenix, AZ
Phoenix, AZ	Detroit, MI	Denver, CO	Denver, CO	Detroit, MI
Minneapolis, MN	Minneapolis, MN	Minneapolis, MN	Detroit, MI	Minneapolis, MN
San Diego, CA	San Diego, CA	San Diego, CA	San Diego, CA	San Diego, CA
Tampa, FL	Orlando, FL	Portland, OR	Orlando, FL	Tampa, FL
Orlando, FL	Portland, OR	Orlando, FL	Tampa, FL	Orlando, FL
Portland, OR	Tampa, FL	Tampa, FL	St. Louis, MO	Portland, OR

Figure 2: Relative pandemic risks of US Cities, according to the city where the initial case appears. City risks are defined as the number of active pandemic cases arriving via airline travel during the course of the pandemic. City risks are remarkably insensitive to the geographic location of the index case, with important and beneficial implications for benefits of *a priori* mitigation efforts for high-risk cities.

Generalizing Results to Other Landscapes

Benchmark Regression on Node Risk (as Log # Infected Arrivals)

	Estimate	Std. Err	T-value	Beta	Pr(>|t|)	
(Intercept)	-7.454	0.558	-13.36	(NA)	< 2e-16	***
mCPL_D	-0.007	0.001	-6.22	-0.212	1.4e-07	***
Degree	0.013	0.002	5.48	0.274	1.7e-06	***
log(empiricalPop)	0.614	0.041	15.04	0.738	< 2e-16	***

Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05

R-Squared: 0.952 Adjusted R-squared: 0.949

We can capture 95% of the variation in risk between nodes with a parsimonious model, where our network measures provide additional and significant explanatory power over node population alone.

Figure 3: Spearman Rank Correlation Coefficient (SRCC) for Node Risk Rankings for simulations with population ranging from 500 to 50,000,000. Sensitivity analyses determined that with 100,000 agents, the simulation produces virtually the same pattern of node risks as with 50 million agents (SRCC = 0.984).

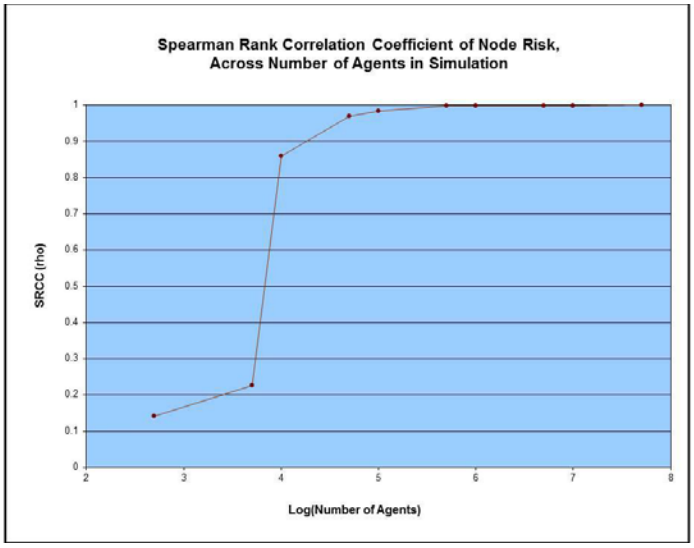

Figure 4: Spearman Rank Correlation Coefficient (SRCC) for Node Risk Rankings for simulations with population ranging from 500 to 50,000,000. Sensitivity analyses determined that with 100,000 agents, the simulation produces virtually the same pattern of node risks as with 50 million agents (SRCC = 0.984).

2.3 Discussion

This work complements the community-level pandemic models of MIDAS and other *Nature*, *Science*, and *PNAS* papers referenced in this paper by providing a lighter, faster model through which regional and global patterns of pandemic diffusion and risks can be evaluated.

The spatial structure of inter-city transportation networks is emphasized at the expense of spatial structure within cities and complements the highly detailed community level models devised by Longini et al. (2005), Ferguson et al. (2006), and Lee et al. (2008). This is also a proof-of-concept for the usefulness of inter-city analyses based on rapid diffusion of pandemic influenza via airline travel. This approach does not yet include analysis of diffusion via slower, localized land travel such as interstate highways and rail systems, and presumes travel behavior would continue as normal during a pandemic. However, because behavior during an actual pandemic would likely be limited voluntarily (even if not mandated by the CDC or other restrictions), thereby reducing the risk to most cities, it would be important to explore the possibility of increased local diffusion due to panic travel or other local/ground travel to nearby cities.

3 EFFECTIVE ALLOCATION OF INTERVENTIONS

We gauge the effectiveness of the treatments relative to the no-intervention control case in terms of the "indirect protection" of the intervention, i.e., the decrease in morbidity among members of the population who were not directly targeted by the intervention. The subsequent sections provide more detail on GA optimization, the intervention strategies, and the post-optimization evaluation of interventions

3.1 Effective Allocation of Limited Resources or of Highly Disruptive Interventions

Genetic Algorithms use evolutionary methods to solve otherwise intractable optimization problems. An optimization problem is represented as a population of strings of parameters; initializing the first generation of strings with valid but otherwise random values; evaluating the fitness of each string according to an objective function; creating child strings by selecting parent strings according to their relative fitness, applying crossover and (rare) mutations to parent strings to create child string(s); and repeating the process for subsequent generations (see Figure 5). Fitness is determined by running sufficient numbers of

simulations to distinguish statistically significant differences. Here, we evolve and analyze the most effective geographic deployment of intervention resources to critical cities, to compare against inaction (no intervention) or pro rata allocation. Specifically, strings specify interventions and geographic deployments for pandemic control. The fitness function maximizes indirect protection; intuitively, the number of people protected indirectly by each direct vaccination.

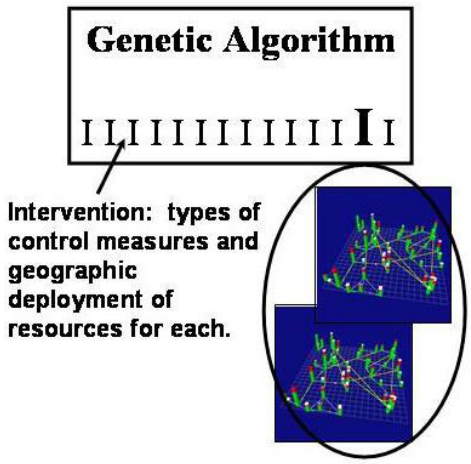

Figure 5: A supervisory Genetic Algorithm evolves populations of strings of simulation parameters by evaluating combinations of simulation parameters according to the results of one or more simulation runs based on those parameters.

The GA employed in this study includes a number of features specific to this task. The GA first processes the simulation model's input parameters, and executes the no-intervention control case using those specific parameters for each of thirty stochastic seeds Roughly 25% of the seeds used in the control case stochastically fail to spark a pandemic, and are discarded (see Dibble, Wendel, and Carle (2007)). The GA then generates a random population of fifty candidate interventions. For each candidate intervention string, the GA executes fifteen *sets* of simulations with the disease originating in each of the (fifteen) highest risk cities from Dibble, Wendel, and Carle (2007). For each intervention candidate and origin city, the GA executes the simulation across the remaining stochastic seeds (~23). Since the intervention process changes the stream of random numbers used in the simulation, roughly 25% of the seeds used for intervention simulations fail to spark a pandemic because of stochastic variation and are discarded. The fitness value of the string is the average indirect protection across the twice-filtered seeds and origin cities. The GA then generates each new generation of interventions using a tournament selection process; the GA completes when the fitness of the best intervention has been constant for fifty generations.

3.2 Intervention Strategies

Our first set of interventions deploys vaccines to a targeted subset of metropolitan areas selected by the Genetic Algorithm. We test subsets of 6 cities. For each level of deployment, we consider two forms of distribution: uniform distribution across the selected metropolitan areas and distribution proportional to the population of the selected metropolitan areas.

Interventions occur at the *start* of the simulation, and provide *permanent* immunity to the disease as though through allocation of effective vaccinations. As in the risk model presented in Section 1, the agents are perfectly mixed within each metropolitan area; vaccines are allocated randomly among agents in a given metropolitan area. If the number of vaccines in a given metropolitan area exceeds the population there, the vaccines are considered wasted and do not contribute to halting the pandemic.

3.3 Evolving Effective Allocations of Pandemic Interventions

The GA evaluates optimal strategies in terms of the indirect protection generated by the intervention, i.e., in terms of its fitness function. This process raises three issues that are important in the ex-post evaluation of optimal strategies. First, since the GA is optimizing over a static set of stochastic seeds, it is feasible that it optimizes interventions that are effective for those particular seeds and not others. With a sufficiently large set of stochastic seeds, and an underlying process which converges at the limit, this possibility is vanishingly small. We investigated various numbers of stochastic seeds, and determined that 30 seeds, as used in the analysis here, would be sufficient to avoid such artifacts. However, "sufficiently large" is not well defined for this simulation model, and thus we check ex-post for the validity of the GA's optimal strategies under new stochastic seeds. All results presented below in Figures 6 and 7 are based on these "new" (non-optimized) seeds.

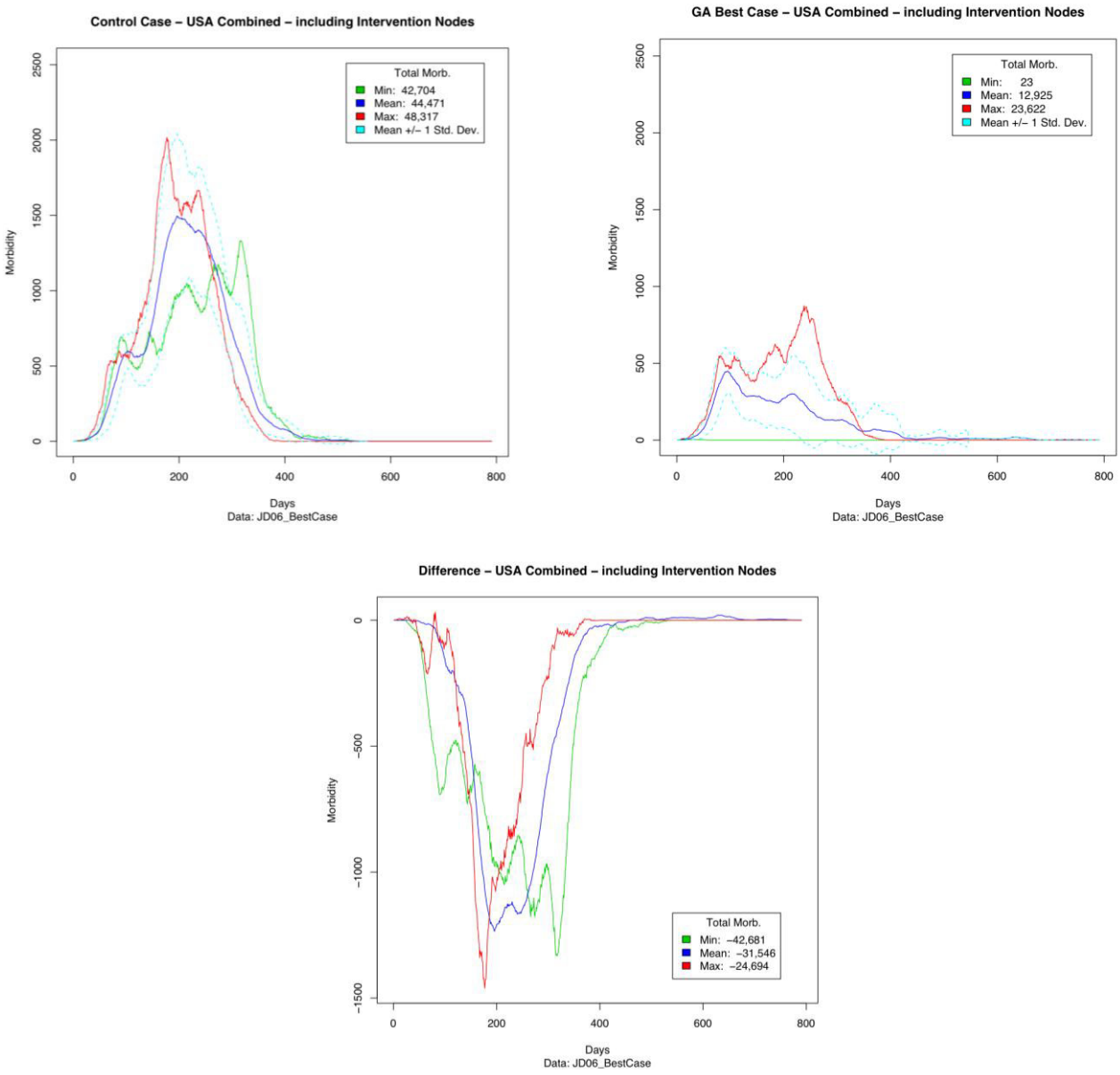

Figure 6: Collective effects of interventions on all US Cities, including 6 cities receiving pandemic interventions and 44 cities that did not receive any interventions.

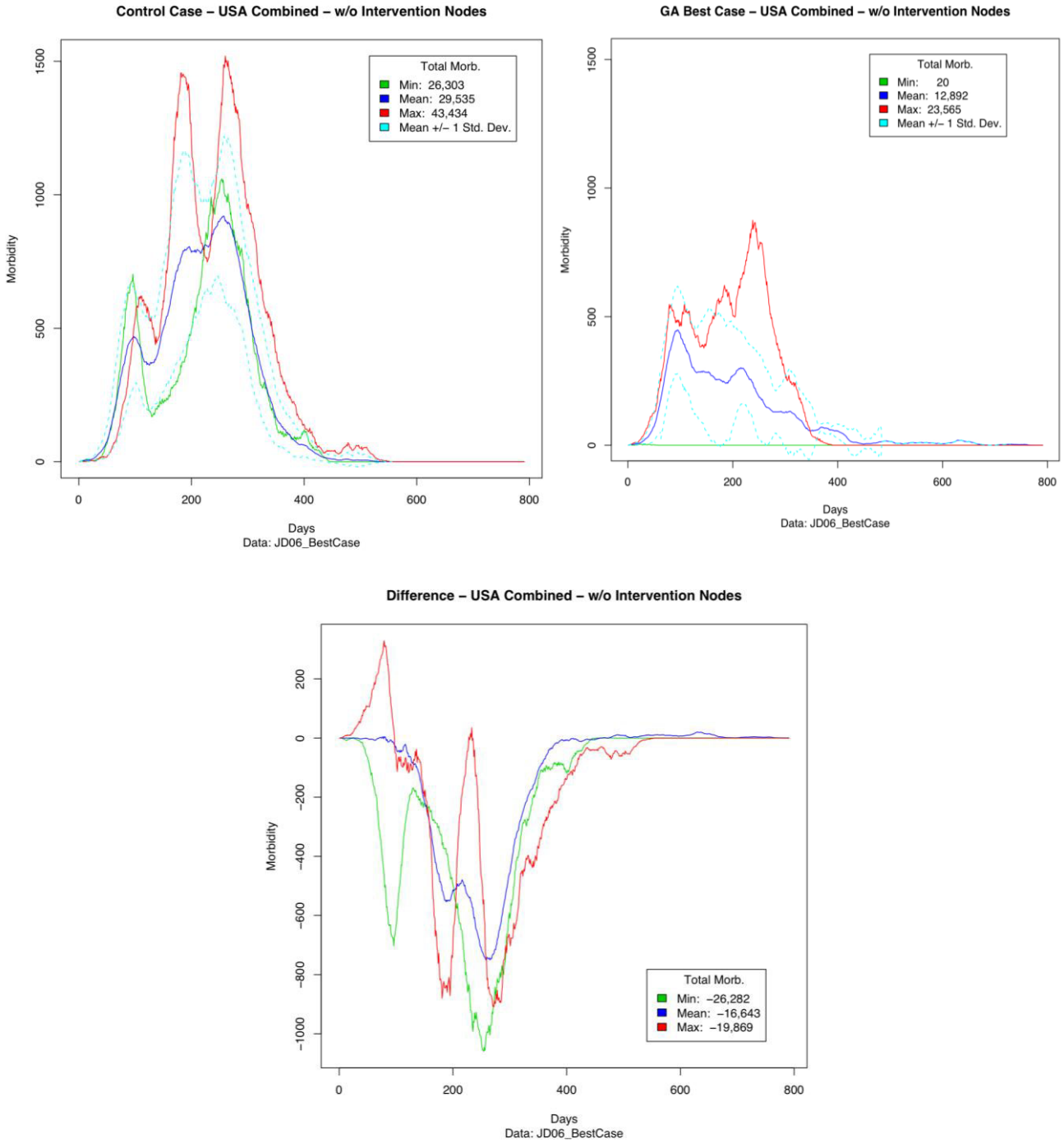

Figure 7: Protective effects of interventions on the 44 US Cities that did *not* receive any interventions.

ACKNOWLEDGMENTS

I would like to thank the Department of Health and Human Services National Vaccine Program Office, the Office of Naval Research, Grant N000140310062, the National Institutes of Health, National Institute of General Medical Sciences, Models of Infectious Disease Agent Study (MIDAS), Grant 1 U01

GM070698, for supporting our work on controlling pandemics. I would also like to thank the TeraGrid Development Allocation Committee (DAC), the University of Maryland Office of Information Technology (OIT), and the University of Maryland Lattice Project for providing computational resources. Kristofor Carle and Stephen Wendel provided expert research assistance for the simulation modeling and computational laboratory evaluations.

REFERENCES

Anderson, R.M., and R.M. May. 1991. *Infectious diseases of humans: dynamics and control*. Oxford University Press.

Dibble, C., S. Wendel, and K. Carle. 2007. Simulating pandemic influenza risks of US cities. In *Proceedings of the 2007 Winter Simulation Conference*, eds. S. G. Henderson, B. Biller, M.-H Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 1548-1550. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Ferguson, N.M., D.A. Cummings, C. Fraser, J.C. Cajka, P.C. Cooley, and D.S. Burke. 2006. Strategies for mitigating an influenza pandemic. *Nature* 442(7101): 448-52.

Gelernter, D. 1992. *Mirror Worlds*. New York: Oxford University Press.

Germann, T.C., K. Kadau, I.M. Longini, and C.A. Macken. 2006. Mitigation strategies for pandemic influenza in the United States. In *Proceedings of the National Academy of Sciences* 103(15): 5935-40.

Institute of Medicine, 2006. *Modeling Community Containment for Pandemic Influenza: A Letter Report*. Committee on Modeling Community Containment for Pandemic Influenza. Washington D.C.: The National Academies Press.

Keeling, M.J., and B.T. Grenfell. 2000. Individual-based perspectives on R_0. *Journal of Theoretical Biology* 203:51-61.

Lee, B.Y., V.L. Bedford, M.S. Roberts, and K.M. Carley. 2008. Virtual epidemic in a virtual city: simulating the spread of influenza in a U.S. metropolitan area. *Translational Research* 151(6):275-287.

Lloyd, A.L., and R.M. May. 1996. Spatial heterogeneity in epidemic models. *Journal of Theoretical Biology* 179:1-11.

Longini, I.M., A. Nizam, S. Xu, K. Ungchusak, W. Hanshaoworakul, D.A.T. Cummings, and M. E. Halloran. 2005. Containing Pandemic Influenza at the Source. Science 309(5737):1083-1087.

NIH 2007. Models of Infectious Disease Agents Study (MIDAS). National Institutes of Health, National Institute of General Medical Sciences. Available via <http://www.epimodels.org> [accessed October 13, 2010].

The TeraGrid Project 2007. TeraGrid Home Page Available via <http://www.teragrid.org/> [accessed July 2007].

AUTHOR BIOGRAPHY

CATHERINE DIBBLE, PhD, is Senior Research Scientist at Aiki Labs in Washington, DC. She has been working professionally with spatial evolutionary algorithms, location-allocation optimization, and adaptive spatio-temporal relevance filters since the early 1990s, and with policy-relevant spatial agent-based computational laboratories since the late 1990s. She has served on the International Steering Committee for the GeoComputation Conference Series since its inception in 1996, on the US National Academy of Sciences National Research Council Committee on Organizational Modeling from Individuals to Societies for 2005-2007, on an NSF Review Panel for Cyber-Infrastructure, and was a co-Investigator with the NIH-funded MIDAS project for controlling epidemics. Dr. Dibble's accomplishments include a 1981 National Bureau of Economic Research (NBER) simulation model to evaluate the baby boom's collective effects on the US Social Security System's balance of payments out to 2050; 1990 shared patents for clinical-complexity-adjusted evaluation of health care providers (Peer-A-Med®); a 1993 genetic algorithm (GA) to optimize the geographic location of facilities such as

emergency centers, clinics, or vaccine stockpiles; the 1994 invention of a genetics-based machine learning relevance filter representation for expert detection of relevant events in terabytes of spatio-temporal surveillance data; the 1999 invention of a new measure to quantify the socio-economic influence of pandemic vulnerability of cities or the leadership effectiveness of individuals in an organization or social network; the 2000 implementation of a supervisory genetic algorithm (meta-GA) for multi-objective optimization of network structure; the design and funded development of four versions of the GeoGraph 3D spatial agent-based computational laboratory; a 2007 genetic algorithm for Disaster-Resilient Location-Allocation optimization; and the 2006-present quantitative evaluation of inter-city pandemic risks and spatio-temporal optimization of pandemic interventions. Her email address is <cath@aikilabs.com>.