The Models of Infectious Disease Agent Study (MIDAS) Research Network

Diglio A. Simoni, Philip C. Cooley, Diane K. Wagener and Eric Solano
RTI International
Research Triangle Park, North Carolina, 27709, USA
{dsimoni, pcc, dwagener, solano}@rti.org

Abstract

The Models of Infectious Disease Agent Study (MIDAS) is a collaborative network of scientists. The major aim of their research is to develop mathematical models that describe infectious disease outbreaks. The overall objective of this effort is to provide feasible responses to "what if" outbreak scenarios that are based on the best available science and data sources that support the science. The nature of providing a computational and informational resource for developing national models has demanded significant effort in many areas. These include cluster design, networking, security, software maintenance, system administration, and new methods for facilitating management and streamlining access to scientific data. These efforts are indicative of the level of commitment that is necessary to support contemporary modeling research in epidemiology.

Keywords: Infectious Diseases, Mathematical Models, Epidemiology, High Performance Computing, High Level Programming, Clusters, Networking.

Resumen

El grupo de análisis de modelos matemáticos para el estudio de mecanismos de dispersión de enfermedades infecciosas (MIDAS, por sus siglas en inglés) es una red de colaboración de científicos. El producto más importante de su investigación es el desarrollo de modelos matemáticos que describen la emergencia de enfermedades contagiosas. El objetivo principal de este esfuerzo es generar respuestas factibles a escenarios potenciales de emergencia de enfermedades. Estas respuestas están basadas en la mejor ciencia disponible y en bases de datos que fundamentan la ciencia. La acción de proveer recursos de computación e informática para el desarrollo de modelos matemáticos en los EEUU ha demandado un gran esfuerzo en muchas áreas. Estos esfuerzos incluyen el diseño de "clusters" o aglomerados de computadores, diseño de redes, seguridad, mantenimiento de "software", administración de sistemas, y métodos nuevos para la facilitación de gestión y acceso directo y simplificado a datos científicos. Estos esfuerzos son indicadores del nivel de compromiso que es necesario para dar ayuda y soporte a la investigación de modelos matemáticos en epidemiología.

Palabras Clave: Enfermedades Infecciosas, Modelos Matemáticos, Epidemiología, Computación de Alto Rendimiento, Programación de Alto Nivel, Conglomerados de Computadores, Redes.

1 Introduction

The Models of Infectious Disease Agent Study (MIDAS) is a research partnership between the United States National Institutes of Health (NIH) and the scientific community whose purpose is to develop computational models for policymakers, public health workers, and other researchers to assist them in making better-informed decisions about emerging infectious diseases, both anthropogenic and naturally-occurring. MIDAS researchers are working to develop models that may assist the public health community understand how best to respond during outbreaks and epidemics. MIDAS consists of seven Research Groups and one centralized Informatics Group.

2 MIDAS Research Groups

The Research Groups focus on models that address research questions that span many aspects of spatiotemporal/biological issues. These include host-pathogen relationships, disease epidemiology, disease surveillance methods and pandemic response strategies. The Research Groups focus on information-driven research rather than hypothesis-driven investigations. MIDAS model developers use real or simulated data that are widely available through the MIDAS web site (www.epimodels.org). The Informatics Group developed and maintains the web site, and manages the information and data resources that modelers may access.

3 MIDAS Informatics Group

The Informatics Group manages computational resources (e.g., Linux clusters and associated data storage systems) and information resources (e.g., information on personal transportation characteristics, demographics, epidemiological information, molecular biology data, and geospatial databases). The Informatics Group also conducts research on validation, organization, analysis and distribution of tools and models, and implements systems to support MIDAS research.

4 MIDAS Missions

4.1 MIDAS Research Mission

MIDAS Research Mission is to study the dynamics of emergence and spread of pathogens and their products, to provide identification and surveillance of infectious diseases, to analyze the effectiveness and consequences of intervention strategies, to study the host/pathogen interactions, and to assess the ecological, climatic, and evolutionary dimensions of infectious disease outbreaks. The MIDAS Research groups actively publish their results in prestigious peer reviewed journals (see [1], [3], [5] and [6]).

4.2 MIDAS Informatics Mission

MIDAS Informatics Mission is to develop large-scale computational resources to support complex models, to create information and knowledge management tools, to formulate analytical and statistical approaches for validating and comparing across model results, to create an archival repository of models, model outputs, and documentation, to acquire a variety of data relevant to modeling, and to test and validate models.

4.3 MIDAS Collaborative Mission

MIDAS collaborative mission is to catalyze discussions among modelers, policymakers, and the public health community that involve setting priorities and designing studies, to take leadership to ensure that MIDAS software is translated into useful tools for the public health community, to share results and resources with the MIDAS network, policymakers, public health officials, and the scientific community, and to take advantage of the intellectual capital within MIDAS to undertake projects that would be impossible for any single group.

5 MIDAS Infrastructure

5.1 MIDAS Portal, Cluster and TeraGrid

The MIDAS portal provides the entry point to both the database systems that store the simulation codes from the Research Groups, as well as to both the MIDAS cluster and TeraGrid-enabled computational and visualization capabilities. TeraGrid resources are available to the MIDAS investigators using a peering mechanism that allows them to checkout the codes and run them on TeraGrid-enabled nodes in a transparent manner.

5.2 MIDAS Repository

The MIDAS repository (MREP) stores and manages computerized models, model results, and relevant information. This resource, in combination with access to TeraGrid facilities is fundamental for integrating the results from all these models into a single modeling framework. MREP stores the simulation codes being developed by the MIDAS research groups into a professional, organized and controlled environment. MREP was designed to "fit" into and support an emergency response process. MREP is implemented using relational database management technology and a web-based Portal for submission to TeraGrid resources.

6 MIDAS Informatics Capabilities

Through a combination of storage and computational resources, the MIDAS Informatics Group offers many capabilities to the Research Groups. It provides a process for responding to an emergency event. Characteristics of MIDAS models across all modeling groups can be quickly identified and linked to relevant documentation. It includes a quality assurance mechanism for tracking, cataloging and locating the different versions of the models involved in different experiments that comprise MIDAS studies. It involves productivity features that allow modelers to efficiently locate previous model versions that use code to be reused in new models. It maintains an inventory of the work developed by the Research Groups in a locatable form.

The MIDAS Informatics Group uses high level programming tools to perform computational intensive tasks faster than any traditional language such as C, C++ and FORTRAN. Current activities include the use of MATLAB® to model the effect of influenza vaccination on target populations, and the MATLAB Distributed Computing Toolbox (DCT) to run computationally intensive models in the MIDAS cluster, by allocating multiple independent computations on separate cluster nodes.

7 MIDAS Scientific Activity

7.1 Models and Data

The Research Groups develop space-time models to detect emerging outbreaks before they spread to most parts of the area under surveillance. The models capture the natural temporal and geographical variations in different types of electronic health system infectious disease data. Most research teams develop and evaluate methods for estimating the instantaneous reproductive number of the infection, the effectiveness of control measures, and the time course of infectiousness from the onset of symptoms. Mathematical models are designed to optimize the response to an emerging disease before the onset of sustained human-to-human transmission, as well as during the early stages of human-to-human spread of the disease. They focus on how to target limited supplies of medication or vaccines for maximum control or containment.

7.2 Airborne and Vector-borne Diseases

The Research Groups have done extensive work in infectious diseases transmitted primarily by the respiratory route (influenza, smallpox, SARS, etc). Additional research has been performed for vector-borne diseases (malaria, dengue, West-Nile). One of the research teams develops mathematical modeling

tools that capture the key features of realistic disease transmission networks in models of intermediate complexity and spatially explicit, village-week level prediction models for annual epidemic malaria to optimize the timing of epidemic alerts for prevention.

One of MIDAS most relevant modeling tool includes a Disease module that represents the disease as a Markov chain. To date, the MIDAS teams have used the module to simulate smallpox, bubonic plague, pneumonic plague, influenza A, B, and H5N1, and anthrax contaminations. For example, one could use the module to model any vector-borne disease for which a Markov chain representation of the disease could be constructed.

Future MIDAS Research may involve the combination of Agent/Individual Based Models or Equation Based Models with ecological niche modeling to predict spatial dynamics of vector insects and human cases of vector borne diseases such as dengue, malaria, yellow fever or leishmaniasis. Some non-MIDAS research has used ecological niche modeling to study vector-borne diseases ([7] and [8]) while other efforts concentrate on the application of individual based models to analyze vector borne diseases ([2] and [4]).

References

[1] Bootsma MC, Ferguson NM (2007). The effect of public health measures on the 1918 influenza pandemic in U.S. cities. *Proc Natl Acad Sci U S A.* 2007 Apr 6.

[2] Depinay JM, Mbogo CM, Killeen G, Knols B, Beier J, Carlson J, Dushoff J, Billingsley P, Mwambi H, Githure J, Toure AM and McKenzie FE (2004). A simulation model of African Anopheles ecology and population dynamics for the analysis of malaria transmission. *Malaria Journal* 2004, 3:29

[3] Epstein JM, Goedecke DM, Yu F, Morris RJ, Wagener DK, et al. (2007). Controlling Pandemic Flu: The Value of International Air Travel Restrictions. *PLoS ONE* 2(5): e401.

[4] Gu W, Killeen GF, Mbogo CM, Regens JL, Githure JI, Beier JC (2003). An individual-based model of Plasmodium falciparum malaria transmission on the coast of Kenya. *Trans R Soc Trop Med Hyg.* 2003 Jan-Feb;97(1):43-50

[5] Hatchett RJ, Mecher CE, Lipsitch M (2007). Public health interventions and epidemic intensity during the 1918 influenza pandemic. *Proc Natl Acad Sci USA.* 2007 Apr 6.

[6] Lessler J, Cummings DA, Fishman S, Vora A, Burke DS. (2007). Transmissibility of swine flu at Fort Dix, 1976. *J R Soc Interface*. 2007 Apr 3.

[7] Peterson AT, Martínez-Campos C, Nakazawa Y, Martínez-Meyer E. (2005). Time-specific ecological niche modeling predicts spatial dynamics of vector insects and human dengue cases. *Trans R Soc Trop Med Hyg.* 2005 Sep; 99(9):647-55.

[8] Peterson, A. T., and J. J. Shaw. (2003). *Lutzomyia* vectors for cutaneous leishmaniasis in southern Brazil: Ecological niche models, predicted geographic distributions, and climate change effects. *International Journal of Parasitology* 33:919-931.