Supercomputers—like those at the Ohio Supercomputer Center—are simply the fastest, most powerful computers at a particular point in time. But what is it really? A tool: A super-fast, sleek tool designed to help researchers turn questions into answers. Those answers are applied to many wonderful purposes, such as making us healthier and smarter. From fighting off viruses to comprehending how we learn and to making our children safer on a drive to school, researchers are harnessing the tools at OSC to break ground on a wide swath of incredible knowledge.

Epidemic Forecasting

## Rao seeks to proactively reduce the impacts of widespread disease

When life-threatening weather events loom, forecasters warn citizens days, even weeks, beforehand so they can take action.

It seems to work: We clear supermarket shelves, board up windows and even evacuate to higher ground ahead of the impending tempest to avoid danger. Blind to bias in its threat to human life is another force of nature—epidemics.

Unfortunately, we often do not know they are imminent until the disease has already infiltrated an area. Dhananjai Rao, Ph.D., an assistant professor in the department of computer science and software engineering at Miami University, is using machine learning and simulation through resources at the Ohio Supercomputer Center to apply the concepts of weather forecasting to epidemiology. He hopes this work will shift the paradigm for disease control from reactive to proactive.

"If we are able to forecast when, where and what an epidemic is going to be, then we can focus our energies and judiciously plan to mitigate the impact of those epidemics in those parts of the world so that the overall impact of the epidemic is significantly reduced," Rao said.

▶ D.J. Rao, Ph.D., has developed an award-winning disease forecasting model that could be used to take proactive steps to guard against infectious diseases.
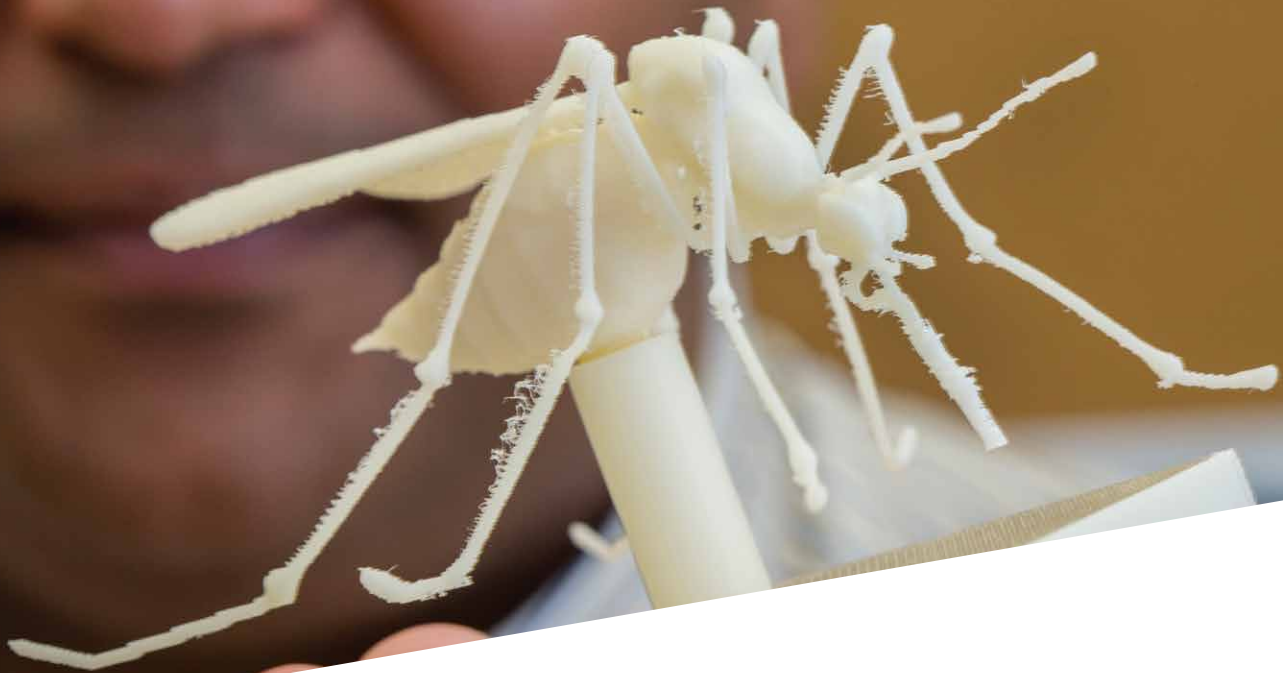
▶ Rao was recognized by DARPA in the CHIKV Challenge, a competition to accelerate the development of new infectious disease forecasting methods.

Rao has recently been studying mosquito-borne diseases, particularly chikungunya and the Zika virus. While there have been some efforts to forecast other known diseases, these existing methods rely on historical data trends to project the next stages of the disease. Unfortunately, for newly discovered or recently spreading diseases, such as the Zika virus, there is no historical data on which to base these statistical models.

"The gotcha with that is the statistical regression models will not be able to tell you what are the ecological processes that are happening. It can only tell you what is the final outcome of the system," Rao said. "Statistical (regression) models you can think of as a black box. You don't really understand what is going on inside."

Rao's approach of fundamentally modeling the disease ecology provides a better idea of what these new diseases are, where they are going and, from there, how to effectively contain them. To create a disease forecast, Rao layers data from multiple fields that collectively create the "perfect storm" of an epidemic. These include weather patterns, mosquito population and life cycle, human population density, air travel and socioeconomic data.

There are also critical parameters for which there are no known data, such as the probability that a mosquito will bite a human. To account for unknown parameters, Rao uses machine learning and parallel simulation to estimate what these values should be. By looking at a combination of continuous and discrete event simulations, Rao uses OSC's Oakley Cluster to study the overall ecology of a disease and how the disease propagates.

This disease model analysis required 3.5 million simulations. This would take approximately 1,000 hours of CPU time, or about 90 days' work, on a single computer.

"Three months. In three months, the epidemic would have long passed through the Americas, or the region that we're looking at," Rao said. "On Oakley, we were able to pull it off in about 12 hours, even on peak load."

Rao was one of 11 teams or individuals recognized recently in a competition organized by the U.S. Department of Defense's Defense Advanced Research Projects Agency (DARPA). The competition, known as the CHIKV Challenge, seeks to accelerate the development of new infectious disease forecasting methods. The challenge identified gaps in current disease forecasting and, with help from Rao's groundbreaking model, DARPA as well as other government agencies can begin to look at how to mitigate the spread and affect of infections diseases.

The next step of the process will involve advising health interventions and public policies to help contain epidemics. Rao said these policies will differ depending on the region and timing of the disease spread.

"There is not a day when we step out of the house without checking the weather, and that's where we want to go," Rao said. "For weather forecasting, it took about 100 years to come to where we are. With epidemic forecasting, we are hoping to be able to accelerate to where we want to be with supercomputing." •