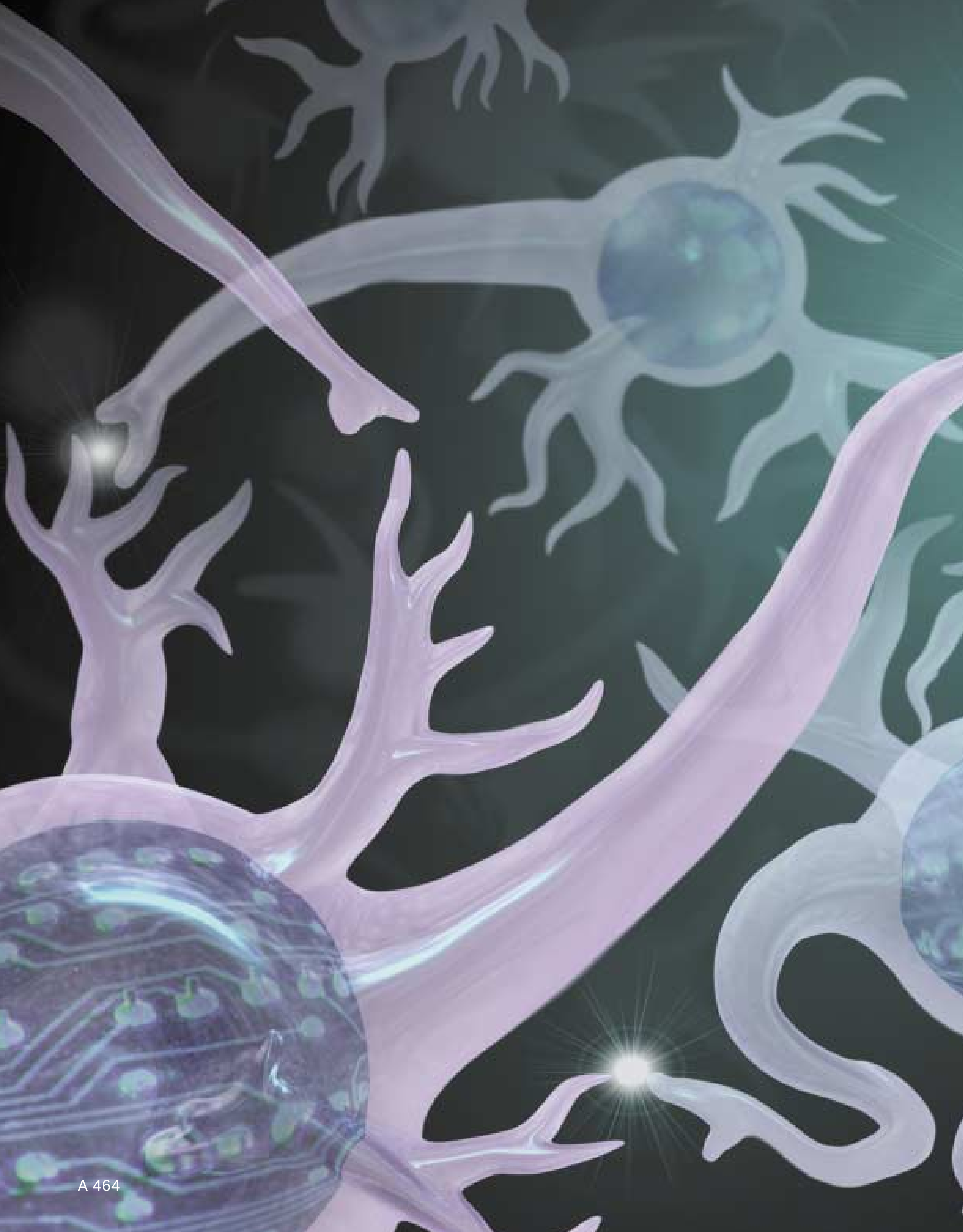

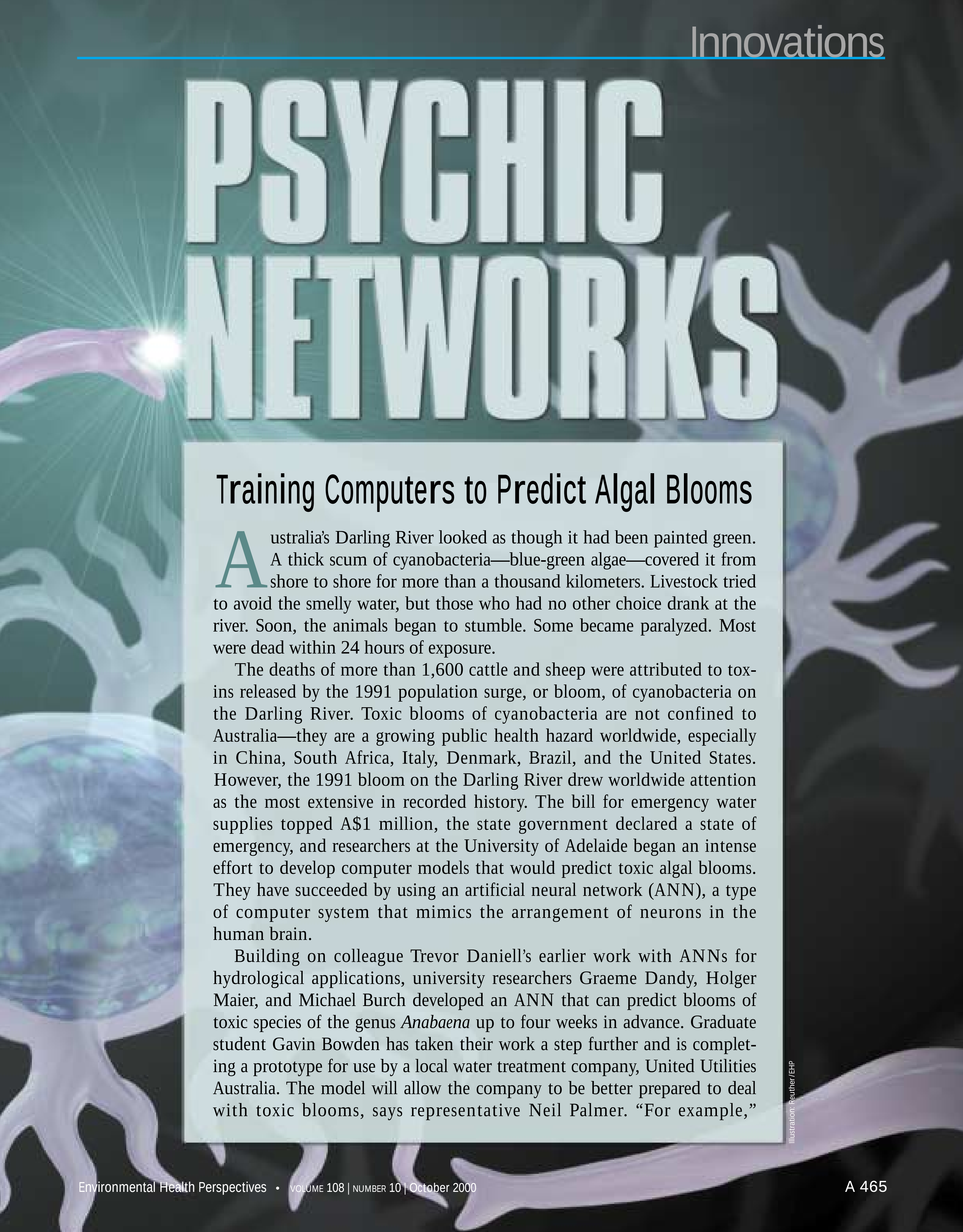

Illustration: Reuther/EHP

PSYCHIC NETWORKS

Training Computers to Predict Algal Blooms

Australia's Darling River looked as though it had been painted green. A thick scum of cyanobacteria—blue-green algae—covered it from shore to shore for more than a thousand kilometers. Livestock tried to avoid the smelly water, but those who had no other choice drank at the river. Soon, the animals began to stumble. Some became paralyzed. Most were dead within 24 hours of exposure.

The deaths of more than 1,600 cattle and sheep were attributed to toxins released by the 1991 population surge, or bloom, of cyanobacteria on the Darling River. Toxic blooms of cyanobacteria are not confined to Australia—they are a growing public health hazard worldwide, especially in China, South Africa, Italy, Denmark, Brazil, and the United States. However, the 1991 bloom on the Darling River drew worldwide attention as the most extensive in recorded history. The bill for emergency water supplies topped A$1 million, the state government declared a state of emergency, and researchers at the University of Adelaide began an intense effort to develop computer models that would predict toxic algal blooms. They have succeeded by using an artificial neural network (ANN), a type of computer system that mimics the arrangement of neurons in the human brain.

Building on colleague Trevor Daniell's earlier work with ANNs for hydrological applications, university researchers Graeme Dandy, Holger Maier, and Michael Burch developed an ANN that can predict blooms of toxic species of the genus *Anabaena* up to four weeks in advance. Graduate student Gavin Bowden has taken their work a step further and is completing a prototype for use by a local water treatment company, United Utilities Australia. The model will allow the company to be better prepared to deal with toxic blooms, says representative Neil Palmer. "For example,"

Ian Falconer, University of Adelaide

Ian Falconer, University of Adelaide

Blue-green blight. The shore waters of the Chaffey Reservoir, part of the Murray–Darling river system in New South Wales, Australia, are coated with toxic blooms of *Anabaena* (left). Reservoir visitors are warned of the health dangers of cyanobacteria (above). With the help of artificial neural networks, water utilities may be able to better prepare for imminent algal blooms.

he says, "we can make sure we have stocks of powdered activated carbon—which is quite expensive—to treat the water, and that all equipment is maintained and ready."

Bowden is also using ANNs to forecast salinity and other factors that affect water quality in the Murray–Darling river system, Australia's major source of surface water. "If you knew a few weeks in advance that a large volume of highly saline water was moving down the Murray, water could be pumped to the Adelaide reservoirs before or after that to avoid pumping the saline water," says Bowden.

Although the researchers' immediate charge is to help improve water quality in the Murray–Darling river system, their primary goal is to develop protocols for using ANNs that can be applied to problems in other geographic areas as well as with other types of water quality problems. "There are no established methodologies or guidelines for designing and implementing ANNs in water resources applications," says Bowden. "Much more research needs to be carried out in this area to explore the strengths and weaknesses of ANNs and identify how they can best be applied."

Cyanobacteria in Bloom

The organisms the Adelaide researchers are trying to track are some of the most ancient on earth. Although commonly referred to as algae, cyanobacteria are actually photosynthesizing bacteria of a type in existence for more than three billion years. These microorganisms can live in fresh- or saltwater—almost anywhere there is moisture—and often link together in long, gelatinous chains that coat stream banks and beaches with slick mats of growth. When conditions are right, they can also outcompete more advanced microorganisms. "When cyanobacteria do bloom, they can form almost 100% of the algal mass—not just one species, but one strain of one species," says Wayne Carmichael, a professor of aquatic biology and toxicology at Wright State University in Dayton, Ohio, and an internationally known expert on toxic cyanobacteria.

Cyanobacteria have several characteristics that allow them to outcompete more complex organisms. First, they have very simple environmental requirements. This enables them to exist in extreme environments where more advanced organisms cannot survive. Second, they can fix nitrogen from the atmosphere, like legumes do in terrestrial environments, so absence of this nutrient is not a limiting factor to their growth. When cyanobacteria are present even in low numbers, their nitrogen-fixing capabilities may enhance the fertility of marine and freshwater environments; in fact, cyanobacteria are sometimes used to fertilize rice paddies. Third, cyanobacteria contain gas vesicles that allow them to maintain a position at the water surface, near the light that fuels their photosynthesis and growth. Finally, some strains of cyanobacteria produce potent toxins that may discourage their being consumed by other marine life such as zooplankton.

As is the case with all algal blooms, explosive overgrowths of cyanobacteria are encouraged by excess nutrients in the water—primarily phosphorus and nitrogen—from sources such as agriculture, urban and rural runoff, and wastewater. Cyanobacteria also require warm (over 68°F), calm, stagnant water to bloom. Eventually, the organisms run out of something they need—a nutrient, perhaps, or adequate sunlight—which can cause them to die.

It's at this stage in their life cycle that cyanobacteria release their toxins. When cyanobacteria are alive, the toxins are contained within their cells, and animals would only be exposed if they ate the whole, living cell. But once cyanobacteria begin to decompose, their cells walls no longer have the strength or integrity to contain the toxins. The toxins released can include neurotoxins, which damage the nervous system, and hepatotoxins, which target liver cells.

A. circinalis, the species that caused the 1991 bloom on the Darling River, produces neurotoxins similar to saxitoxins, the paralytic shellfish poisons released by certain marine dinoflagellates during so-called red tides. One cyanobacterial neurotoxin, anatoxin-a(s), is similar in structure to the insecticides parathion and malathion. In high doses, neurotoxins can overload the nervous system, causing paralysis and respiratory failure.

The hepatotoxins produced by cyanobacteria include cyclic peptides called microcystins and nodularins. At high doses, these toxins destroy liver cells, causing fatal hemorrhaging. Low doses can disrupt the enzymes that control cell division, contributing to cancer. "The extraordinarily high rates of liver cancer in parts of China may be tied to repeated exposure to cyanobacterial microcystins in drinking water," says Carmichael. Children and people who already suffer from liver or kidney damage are especially vulnerable to cyanobacterial toxins.

Ian Falconer, an emeritus professor of biochemistry in the Department of Clinical and Experimental Pharmacology at the University of Adelaide and a project leader and research supervisor at the Cooperative Research Centre for Water Quality and Treatment, has documented several cases where cyanobacterial toxins passed through standard water treatment (including chlorination) to cause outbreaks of gastroenteritis, painful liver enlargement, and bloody diarrhea. According to Falconer, filtering can remove the toxins while they are still contained in live cyanobacterial cells. But once the cells die, the toxins pass into solution in the water column and are much more difficult to remove.

Municipalities could take effective action if they were prepared. For example, extra filtering through granulated activated carbon can be effective at removing toxins. Or, if water districts knew that conditions favorable to a cyanobacterial bloom were imminent, they could take measures to discourage a bloom, such as increasing aeration or water flows to disrupt the water column so cyanobacteria couldn't form mats at the surface. Recreational areas could be closed before a bloom hit to protect swimmers and boaters, who can suffer rashes and other allergic reactions from contact with cyanobacteria. This is where prediction models such as those developed by the Adelaide researchers come in.

Enter ANN

In their efforts to predict cyanobacterial blooms, the Adelaide researchers chose to experiment with ANNs, a relatively new type of statistical model that does not have to be programmed. "You let the ANN decide what kind of statistical model best fits the shape of your data," says Maier.

ANNs detect patterns through parallel processing among simple, interconnected processing units called nodes. The model for this structure is the human brain, in which processing is shared among many interconnected single cells, or neurons. The more a connection between neurons is used, the more it is reinforced. If a connection is not used, it atrophies. ANNs mimic this biological process by giving numerical values, or "weights," to connections between nodes, with more influential factors having higher weights than less influential factors. Each node receives inputs from many other nodes via these weighted connections, then performs calculations to transform these multiple inputs into one numerical output. This single output is then passed back through the weighted input connections of many other nodes.

ANNs set weights through "training," or exposure to data sets. At the start of training, the weights among nodes are small and random. The ANN is then presented with a set of data, such as water conditions over a given period, and develops an output that is compared to what actually happened—for example, whether a bloom occurred or not. The program calculates the error between its output and the data, and works backward through the weighted connections to adjust them. The process is repeated with more training examples until the ANN minimizes the error. A key factor in the researchers' decision to use an ANN was the fact that they had decades of data on algal blooms to work with. "If you have lots of variation in your data that covers all possible events, the ANN will find patterns. If you don't have a long data record, you can't expect the model to do well," says Maier.

A disadvantage of the ANN approach is that it can be difficult to determine what patterns the system has found and to figure out how it arrived at its conclusions. An ANN will deliver a prediction—for example, that a bloom will occur in two months if current conditions hold—but it won't single out the most important factor causing the bloom. There are ways to find out which factors are most important, primarily by changing the inputs into the ANN. If an ANN considered 12 factors in a prediction, the most influential factors could be determined by submitting the factors to the system individually or in small groups. For example, if the 2 most important factors were submitted, the system should yield a prediction close to the one obtained by using all 12 factors. If 2 relatively unimportant factors are presented to the system, it should yield a very different answer. Users can query the system by analyzing all the interactions among all the nodes, but this a complex and time-consuming process.

"It's not a straightforward process to go back and analyze the weights and connections," says Oscar Schofield, an assistant professor of biological oceanography at the Institute of Marine and Coastal Sciences at Rutgers University in New Brunswick, New Jersey, who also is experimenting with the use of ANNs to predict algal growth. Adds Maier, "If you use a really big network, the equations that you get are so complex that it's difficult to interpret them." However, right now Maier and his colleagues are less interested in describing algal blooms than they are in predicting them. Therefore, a model that is, to use Schofield's words, "set up to slurp in the data and spit out a prediction" best serves their immediate needs.

Like many statistical models, the University of Adelaide ANN is region-specific; its results can only be applied to the Murray–Darling basin, where the data used to train it were collected. However, the Adelaide researchers hope that their research will be of use in other regions. "We're trying to develop a robust methodology—which inputs to use for an ANN, how to structure it, how best to train it," says Maier.

Cyanobacterial blooms have not been a research priority in the United States, but that may be changing. "Almost 80% of the surface drinking water reservoirs we tested in a recent study of 50 water utilities had blue-green toxins," says Carmichael. "The toxins were found at low levels, and in almost all cases the water treatment plants in the areas were capable of removing them from drinking water. But as our systems age, blue-green algae could become more of a problem here, the way it is now in Europe. Many U.S. surface reservoirs are at the age, about 40–50 years, when they start to turn eutrophic"—that is, they collect enough excess nutrients to support cyanobacterial blooms.

"We need to understand a whole lot about the ecology and management of [cyanobacteria] in a hurry because we may be able to do preventative work to keep them from getting established in some areas," says Hans Paerl, Kenan professor of marine and environmental sciences at the Institute of Marine Sciences in Morehead City, North Carolina. "Once they do get established, you often have to take some very radical and expensive measures to remediate them."

Once U.S. institutions are ready to fund research into cyanobacterial blooms, they may go to Adelaide for tutoring in using ANNs. "All you need is the right data for a particular water supply and you can apply the same modeling process," claims Bowden. "Wherever water quality is of concern, [the model] can be used."

Kris Freeman

Suggested Reading

Carmichael WW. The toxins of cyanobacteria. Sci Am 270(1):78–86 (1994).

Falconer IR. An overview of problems caused by toxic blue-green algae (cyanobacteria) in drinking and recreational water. Environ Toxicol 14:5–12 (1999).

Maier HR, Dandy GC, Burch MD. Use of artificial neural networks for modelling cyanobacteria *Anabaena* spp. in the River Murray, South Australia. Ecol Model 105:257–272 (1998).

Recknagel F, French M, Harkonen P, Yabunaka K-I. Artificial neural network approach for modelling and prediction of algal blooms. Ecol Model 96:11–18 (1997).