Puzzling over blue-green algae

Sep. 30

burlingtonfreepress.com

Puzzling over blue-green algae

How do blooms begin? Why? What keeps them going? And what makes them toxic - or not? A new, cutting-edge collaboration at UVM aims to answer the riddles.

University of Vermont research technician Susan Fuller (left) and graduate student Rebecca Gorney carry samples of water from St. Albans Bay in early September. / GLENN RUSSELL/FREE PRESS

As a life form, blue-green algae — those occasionally toxic scums that appear and vanish without warning in Lake Champlain — are both simple and confounding.

These algae subsist on warm water, an elemental diet and sunlight; they make their own food via photosynthesis. That's the simple part.

But what gets a bloom started? What keeps it going? And just as important, why is it sometimes toxic — toxic enough to close beaches, toxic enough to kill dogs that drink the water it thrives on — and sometimes not?

These are questions that have preoccupied Mary Watzin for years. She's an aquatic scientist at the University of Vermont, but she's also a kind of detective, probing the mysteries of blue-green algae.

Starting about 10 years ago, she started getting grants to study cyanobacteria, as blue-green algae are otherwise known. Monitoring stations were set up around Lake Champlain, and data collection began.

Watzin knew soon enough that Burlington's and South Burlington's public water systems were not in jeopardy, because they drew from deep in the lake, far deeper than cyanobacteria generally were found. The potential public health problem lay, rather, along shorelines and in shallower intakes where the blooms appeared.

"My curiosity was always around what's causing the blooms, what's driving the blooms," Watzin recalled recently. "The common assumption was phosphorus. Clearly, there's more phosphorus in Missisquoi Bay than there is in most other parts of the lake, and that's where the blooms are the worst. Is that all of the story? Not when we start to think about all the other things."

She knew that blue-green algae also feed on nitrogen, which like phosphorus is contained in runoff washing into the lake, and she knew that the ratio of nitrogen to phosphorus is an important benchmark.

She was aware that some blue-green algae, when photosynthesyzing at the surface, can pull nitrogen directly out of the atmosphere — a capacity that gives them an advantage over nitrogen-dependent plant forms with which they compete. But Watzin also learned of something else blue-green algae can do: migrate downward 15 or 20 feet below the surface, which in Missisquoi Bay means the bottom. What was that about? Were blue-green algae bottom-scroungers?

More unanswered questions arose. Take the most problematic genus of cyanobacteria, Microcystis. Some samples of Microcystis contain the gene that's necessary for producing toxin, some don't. Of those that contain the gene, sometimes they produce the toxin, but sometimes they don't. How to explain that?

And overshadowing all these puzzles was the biggest one of all — the summer of 2007, known to Watzin and her colleagues as "the year without a bloom."

That year alone in Missisquoi Bay, Watzin said, "there never was a bloom of any significance. What is it that was different about 2007? That is the burning question, to figure out what is driving these blooms."

The data grew, and so did the hypotheses — but the puzzles remained.

Enter Donna Rizzo. She's an associate professor of engineering at UVM. She knows next to nothing about blue-green algae.

What Rizzo does know something about, however, is Artificial Neural Networks, which form the basis of computational techniques that can be used to characterize, and even predict, various complex phenomena.

An Artificial Neural Network (commonly referred to by its acronym, ANN) attempts to mimic the response of the human nervous system, or the human brain, to multiple stimuli. The stimuli take the form of data, which the researcher feeds into a computer programmed to do the processing. The processing formula, or algorithm, is designed to function as a network — a network comprising multitudes of parallel nodes (simulated neurons) that respond individually and collectively to the data by firing, or not. The result, or outcome, can take the form of clusterings, or patterns.

Rizzo, in fact, had used an ANN model to analyze the underground flow of contaminants from a landfill in upstate New York.

Could ANN modeling provide any insight into the blue-green algae blooms in Missisquoi Bay? Watzin wasn't sure, but she was willing to give it a try.

ANNs in action

Artificial Neural Networks are being used by researchers in a variety of fields these days.

• Josh Bongard, assistant professor of computer science at UVM, designs robots that have a capacity to learn or adapt.

"The robots that we create in my lab contain ANNs," he wrote in an email. "The ANNs make sense of the sensory information flowing into the robot and dictate how the robot should move in response. For example, if a two-legged robot starts to fall over, the robot's ANN detects the fall by changes in the robot's sensors and commands the robot to move its legs to catch the fall."

ANNs are coming into fashion among computer scientists, Bongard said, because simulating the biological brain requires enormous computing power that only the latest, advanced computers can handle.

"Also, theoretical advances have been made that help us decide what kinds of ANNs to use for what kinds of problems," he wrote. "ANNs can be used to model and predict, as can biological brains: our brain models the flight path of a ball and predicts where the ball will be when it reaches us. Our brain then commands our hand to move to that position to catch it."

• Lance Besaw worked with Rizzo to study groundwater flows using ANNs when he was a doctoral student at

UVM. He notes that recent advances in computers have facilitated ANN applications in the everyday world.

“ANNs have been used to understand content in pictures (like to find faces in digital camera displays) and recognize speech (such as in popular smart phone apps),” Besaw wrote in an email. “But their uses go beyond today’s must-have gadgets.”

Besaw works as a senior scientist for Applied Research Associates in Randolph, where he uses ANN modeling for a range of contractual assignments.

“I have used ANNs in applications that include humanitarian de-mining in Cambodia, anti-terrorism work for the Department of Homeland Security and oil slick cleanup for BP,” he wrote.

• Hugh Garavan, associate professor of psychology at UVM, has begun a collaboration with Rizzo that will use ANNs in quite a different way: to study adolescent behavior, such as hyperactivity.

Garavan and his colleague Robert Whelan have tons of data on adolescents, measures of activity from different brain regions, behavioral and cognitive measures — so much data, according to Garavan, that it’s “quite overwhelming” for a human observer trying to find which features are critical to an adolescent’s having ADHD (attention deficit hyperactivity disorder) or being a smoker.

The hope is that ANNs can sift through the mass of data to find which combinations accord, say, with impulsivity.

“The big picture here,” Garavan wrote in an email, “is that numerous sciences are now overwhelmed by having too much data (brain science, economics, weather prediction), so are very keen on importing machine learning techniques such as ANN to help discern what might be subtle and sophisticated patterns in their data.”

Recognizing patterns

The landfill that Rizzo, the UVM engineering professor, studied was in Schuyler Falls, N.Y. Contaminants such as lead and arsenic were known to be leaching out of the landfill underground. The main question was: Could the movement of the contaminants be accurately predicted? The subsurface migrations could depend on many things — soil and contaminant types, groundwater flow patterns — once again, a ton of data.

But Rizzo and her colleagues decided to focus on data from a subset of observations relatively easy to obtain: samples of microbes found in groundwater monitoring wells. Using those data, their ANN algorithm turned out a map that corresponded to where contaminants were known to have migrated — even though the computer had been given no information about the measured contamination. The result suggested that microbes could be used as proxies in assessing where the pollution was flowing.

Essentially, the computer was using the ANN to recognize patterns. Pattern-recognition is a form of intelligence. It’s something the brain does. It’s something that an advanced computer has the potential to do even better.

But the human brain’s superiority shows up in another form of intelligence, according to Rizzo.

“The thing that separates us from other species,” she said, “is that we can remember and interpret a long history of data and come up with new hypotheses. That’s not something the ANN model can do.”

In the Lake Champlain research, her hope was that the computer would recognize patterns in blue-green algae blooms. If it did, the task of of interpreting those patterns so as to form new hypotheses would fall to the brain of Mary Watzin.

The a-ha moment

After years of reading the literature and combing the data, Watzin still couldn't be sure what it was that drove the blue-green algae blooms. She knew a lot, but she was sort of stuck, so she was hoping for a little help from the computer.

"Give me another clue so I can figure this out," she recalls thinking when she started her ANN collaboration with Rizzo. "Tell me if some of the ideas that I think might be going on but that I don't have the data to prove might be true."

Watzin had not worked with ANNs before, but they seemed an appropriate way to approach the problem.

"When scientists don't really understand what's driving something, the first thing they do is classify," she said recently. "When you classify for a while, you see a pattern, and then you can try and figure out what drives the pattern. When you think about taxonomists, before we figured out evolution — they classified. They classified, they classified, they classified, and then Charles Darwin came along and said, whoa, there's a pattern here in these various classes.

"And so really, we're jump-starting the classifying to find the pattern when we use these Artificial Neural Networks," she continued. "They can take 15 things at once and tell you, here's a pattern. Now, you go back and decompose and figure out what it's telling you about what's important."

Here's what Watzin and Rizzo did: They took three years of Missisquoi Bay sampling data for about 20 environmental variables: phosphorus, nitrogen, temperature, turbidity, dissolved ions in the water, total nitrogen to total phosphorus ratio and more.

One piece of data they did not include, however, was whether blooms were present in the samples. They wanted the computer to be left in the dark on that.

The data had to be "scrubbed" — that is, put in the right format for the computer. Watzin and Rizzo focused on the data gathered when the water temperature was at least 19 degrees Celsius (66.2 degrees Fahrenheit). Below that temperature, blooms never occur.

Of the many ANN algorithms available, Rizzo chose one that asks the computer to process the data into clusters — in this case, two distinct clusters.

When the data were ready — a few thousand numbers — they were fed into a computer at UVM's advanced computing center. How long did the processing take?

"About a minute," Rizzo said.

The result was stunning.

The computer churned out two sets of clusters, represented graphically by irregularly shaped blobs of two colors. One color corresponded to the samples where blue-green algae blooms had occurred. The other corresponded to samples without blooms.

In a joint interview with Rizzo, Watzin was asked what she thought when she saw the outcome, in the form of one graphic display superimposed on another.

"We got bloom, no bloom, and it's like, Holy cow!" she said, chuckling. "It's magic!"

"Remember," Rizzo said, "we never told the computer 'bloom, no bloom.' It did that by itself."

"It tells you something's there," Watzin said. "It's telling us, there is some unique combination of environmental characteristics that are associated with a bloom, and you idiot, Watzin, just haven't figured them out yet!"

Factors emerge

When she looked at the data that went into the "bloom" clusters, several things jumped out. One was the presence of dissolved manganese at the interface between the water and the sediment at the bottom of the bay.

Dissolved manganese was coming out of the sediment, and the computer was associating that with blooms. Watzin hadn't made the connection before, but it started to make sense when she thought it through:

Dissolved manganese comes out of sediment when the oxygen in the sediment has been consumed by bacteria, and that typically happens when the sediment is rich in organic matter, as from farm and stormwater runoff. And when the oxygen is depleted, two other things come out of the sediment: dissolved phosphorus and dissolved nitrogen, in the form of ammonium. Perhaps those dissolved forms were fueling the algae.

"That's when we started to say, they're going down to the bottom, and it's giving them an advantage in Missisquoi Bay because it's so shallow" Watzin said. "We suspected the release of phosphorus and nitrogen from the sediment, but we weren't able to measure that because it's so transitory. That's what we think is happening: the algae are going down to the bottom at night, scarfing up that nitrogen and phosphorus, and going back up to the surface."

That led to experiments in Watzin's lab. Sure enough, blue-green algae thrived on ammonium. Watzin calls it their "rocket fuel."

Bottom scroungers, indeed.

The ANN work continued, a back-and-forth between the computer and Watzin. Withholding the data on which samples were toxic, Rizzo provided the computer with an algorithm that called for three clusters. The three that came back matched up with: no bloom, toxic blooms, non-toxic blooms. What factors did the toxic-bloom group have in common? Watzin suspected competition with other forms of algae, and that led to more experiments. She and her lab are still at it.

But what about 2007? Why weren't there any blooms that summer? Some combination of factors led to that, but the ANNs aren't revealing it. Not enough data, perhaps, or the right kinds.

"We have all the data for 2007 that we had for the rest of the years," Rizzo said. "So, we're missing something important. Unfortunately, you don't collect data when you don't think you need it."

They've started collecting data they didn't have before. Watzin wants to know more about what washes into the lake after the spring snowmelt.

The mystery of 2007 continues.

Contact Tim Johnson at 660-1808 or tjohnson@burlingtonfreepress.com. Follow him on Twitter at www.twitter.com/flyonthe.