THE FORECAST IS CLOUDY

BY MICHAEL BEHAR

HURRICANES LIKE MATTHEW HAVE LAID BARE THE DIRTY SECRET OF THE NATIONAL WEATHER SERVICE: ITS TECHNOLOGIES AND METHODS ARE WOEFULLY BEHIND THE TIMES.

A t 11 o'clock on the night of Sept. 29, the National Hurricane Center in Miami posted an updated prediction for Hurricane Matthew. Using the latest data from a reconnaissance aircraft, the center's computerized models led meteorologists there to conclude, in a post on the center's website, that "only a slight strengthening is forecast during the next 12 to 24 hours." Their prediction proved to be astonishingly amiss: The following day, Matthew exploded from a Category 1 into a Category 5 hurricane, with winds gusting to 160 miles per hour, strong enough to flatten even the sturdiest homes.

This was hardly the first time that United States government forecasters significantly underestimated a storm's potential. Last year, 24 hours before Patricia reached Mexico's Pacific Coast, it unexpectedly mushroomed from a tropical storm to a Category 5 hurricane, its winds topping 215 miles per hour. Luckily, Patricia — officially the strongest hurricane on record in the Western Hemisphere — made landfall over a sparsely populated region. Matthew behaved similarly, its intensification also unforeseen and sudden, occurring just two days before it overwhelmed Haiti. Residents there had little time to flee, and the death toll exceeded 1,000. (More than 30 died in the United States.) The failure to make timely, accurate predictions about these storms would have had far deadlier consequences had they made landfall near a major metropolitan area. In South Florida, for example — where the initial forecasts for a storm of modest size would not have prompted hurricane-weary residents to evacuate — Matthew's rapid increase in power could have pinned down more than six million people in the region.

It's a situation that deeply troubles Cliff Mass, a meteorologist and professor of atmospheric sciences at the University of Washington. As he does after every major weather event, Mass deconstructed the bungled predictions for Matthew and Patricia on his popular website, "Cliff Mass Weather Blog," which he started in 2008. He called Patricia a "poster child, perhaps the worst case in a while, of a major problem for meteorologists," and in response to Matthew he posted a graph that showed how the National Hurricane Center's computer-forecasting model at one point was off by more than 325 nautical miles in predicting the storm's westward course.

Mass, who is 64, has become the most widely recognized critic of weather forecasting in the United States — and specifically the National Oceanic and Atmospheric Administration, which manages the National Weather Service and its underling agencies, including the National Centers for Environmental Prediction, where the nation's weather models are run. Mass argues that these models are significantly flawed in comparison with commercial and European alternatives. American forecasting also does poorly at data assimilation, the process of integrating information about atmospheric conditions into modeling programs; in the meantime, a lack of available computing power precludes the use of more advanced systems already operating at places like the European Center for Medium-Range Weather Forecasts, based in Reading, England. And there are persistent management challenges, perhaps best represented by the legions of NOAA scientists whose innovations remain stranded in research labs and out of the hands of the National Weather Service operational forecasters who make the day-to-day predictions in 122 regional offices around the country.

As Mass points out, accuracy is everything, often the difference between life and death, given that extreme weather — tornadoes, flash floods, heat waves — kills more than 500 Americans each year. "An incremental improvement

MAP BY FERNANDA B. VIÉGAS AND MARTIN WATTENBERG

would make a huge difference," he says. Industries like shipping, energy, agriculture and utilities lose money when predictions fail. Even slightly more precise wind-speed projections would help airlines greatly reduce fuel costs.

Mass participates regularly on government committees tasked with improving forecasts, where his message is consistently the same — in his words, "We could be doing much better, and it's outrageous that we're not." Last year, he notes, the Air Force began paying Britain's Met Office $100,000 a year to license its weather-modeling software. "That a U.S. government agency has decided that our capability is not good enough is pretty amazing," he says. On his blog, he's not shy about criticizing federal agencies in post-mortems after storms, often singling out the National Centers for Environmental Prediction.

William Lapenta, who heads the centers, welcomes the criticism: "His job through his blog is quite honestly to provoke people to respond and hopefully take action," he says. Indeed, Lapenta told me that the National Centers for Environmental Prediction, which directs the National Hurricane Center, might never have obtained additional funding from Congress to buy new supercomputers had Mass not drawn public attention to the center's inadequacies in the aftermath of Hurricane Sandy. "People didn't talk about this before," Al Roker, the "Today" show weather anchor, told me. "It wasn't up for discussion. People didn't care. Cliff has made this part of the conversation. He's like the weather's answer to Neil deGrasse Tyson."

You don't have to go far to get firsthand confirmation of Mass's concerns. Last summer, I visited the regional National Weather Service office in Boulder, Colo., where I live. The office serves three million people in the greater Denver area. Its director, Nezette Rydell, told me she knows many NOAA scientists who have developed better prediction technologies. But few are at her disposal. "There is so much stuff on the shelf that isn't being used," she said.

Rydell introduced me to a meteorologist on her staff who was tracking a thunderstorm inching across Colorado. I watched as he issued a "severe thunderstorm alert" to the public. Like every such warning, it was dispatched in all capital letters. When I asked why, he explained that the Weather Service interface was so primitive — the protocol was originally designed for the telegraph — it could only accommodate uppercase type.

T o make a forecast, meteorologists use a weather model, essentially a big software program that runs on a supercomputer. Multiple times a day, new atmospheric data — obtained mostly from weather stations and satellites — is automatically fed into a model, which generates a prediction by applying equations from fluid dynamics, physics and chemistry to data like temperature, wind direction and humidity. No model will be right every time, but faulty predictions can often be traced back to correctable factors: flawed or insufficient data, shoddy physics, inferior methods of data assimilation used to integrate that data into the model or a lack of the microprocessing power needed to run it effectively.

These are the sorts of deficiencies that can prompt forecasters to underestimate the severity of a hurricane, for instance — or overhype an end-of-the-world blizzard, like the one Gary Szatkowski anticipated on the last Sunday in January 2015. Szatkowski, now retired but then the head meteorologist for the National Weather Service office in Mount Holly, N.J., predicted that a cold front creeping westward toward the East Coast would evolve into an unprecedented nor'easter. The Weather Service issued an official warning: A "crippling and potentially historic blizzard" with "life-threatening conditions" and up to three feet of snow was coming.

Ten inches of snow fell. By late Monday night, Szatkowski, whose office forecasts the weather for more than 11 million people, accepted that he had made a huge blunder and said in two tweets: "My deepest apologies to many key decision makers and so many members of the general public. You made a lot of tough decisions expecting us to get it right, and we didn't."

Mass spent hours in advance of the January 2015 storm scrutinizing models on the National Weather Service website. After the debacle, he posted a critique on his blog and told me later that the forecaster "broke all the rules I teach my students."

While Mass is the most outspoken on the subject, many experts insist that if the Weather Service wants to meaningfully improve its predictions, it must employ a technique called ensemble forecasting. The basic premise is either to tweak the physics equations or to make repeated changes to a model's variables: You might bump up the temperature slightly, for example, and then run the model again. After a half-dozen or so reruns, you get a set, or "ensemble," of forecasts that can be compared with one another. When all the forecasts in an ensemble agree, it's a reasonably sure bet that the predictions will pan out.

Nobody I've spoken to doubts the superiority of ensembles. Yet they haven't been widely adopted in the United States at the resolution required to forecast localized, or "mesoscale," events — specifically, thunderstorms, flash floods and tornadoes — because high-resolution ensembles require more computing power than the National Weather Service can currently provide. Higher-resolution ensembles translate to greater accuracy in the same way that HDTVs are clearer than analog sets. I met with a scientist at the National Center for Atmospheric Research in Boulder who showed me a prototype mesoscale ensemble for the United States. But at the moment, he can't exploit its full potential because the supercomputing cluster at the Weather Service simply couldn't handle the load.

Mass also contends that the Weather Service should be spending far more to exploit Tropospheric Airborne Meteorological Data Reporting, or Tamdar, developed in the late 1990s. Regional airlines, like SkyWest, have Tamdar sensors on their aircraft, capturing data for Panasonic Weather Solutions' new global model, which often outperforms Weather Service predictions. In 2008, the Weather Service began buying this proprietary data, but budget constraints in 2013 put an end to that. The Federal Aviation Administration funded NOAA to study the value of Tamdar observations. The results were staggering: Without it, forecast accuracy plummeted up to 50 percent. Last year, the Weather Service found a budgetary workaround that let it purchase a small amount of limited, low-resolution Tamdar data, but it's nowhere near enough to make a difference in the accuracy of its models.

ONE SUNNY SPRING MORNING last year, I sat in on an undergraduate course at the University of Washington called Weather Prediction and Advanced Synoptic Analysis. Mass, doe-eyed and gangly, with finger-thick eyebrows and a pronounced aquiline nose, arrived for the hourlong lecture perspiring heavily. "I just squeezed in a run," he told the class, apologetically. On an overhead projector, he showed some past National Weather Service predictions; about one-third of them were wrong.

"Oh, and the fun doesn't end there," he declared. Next he displayed screen grabs from the agency's webpages and mocked the archaic design. They received more online ridicule from him earlier this month, after the National Hurricane Center's website crashed just as Hurricane Matthew approached the Florida coast.

After class, I joined Mass in his narrow sixth-floor office. (I took a couple of courses from him when I was a Washington student in the 1980s.) As we talked, he pulled up a digital photo on his computer that he'd scanned from 1974, taken on Mass's graduation day at Cornell, where he earned a bachelor's degree in physics. Mass, then 22, stood smiling shoulder to shoulder with Carl Sagan. "There is a little Carl Sagan in me — billions and billions," Mass said, mimicking the renowned astronomer.

"I loved weather," Mass said. "I also loved astronomy. I wasn't sure which way I wanted to go." He'd been experimenting with building computerized weather models while taking Sagan's class on planetary atmospheres. "I decided to try modeling Mars." Mass pitched the concept to Sagan. "He was all excited about it so he said, 'Let's write a paper together.'"

The Journal of Atmospheric Sciences published the study; Mass and Sagan shared a byline. To be paired with Sagan was quite an accomplishment, and the two remained friends until Sagan's death in 1996. "Carl was an activist in a lot of ways," Mass said. "He basically taught me how scientists have to go directly to the people."

A few months before Sagan died, Mass began volunteering weekly at Seattle's KUOW public radio, hosting a show about the intricacies of weather forecasting. With his sonorous baritone, Mass brought gravitas to even the most mundane meteorological concepts. Two years later, in 1998, he created a webpage to push for bringing Doppler radar to Washington's coast. Doppler can reveal a lot about an approaching storm — the intensity of its rainfall and wind, for instance. It provides a three-dimensional picture of a storm's internal structure at a given moment, which satellites generally cannot do because their orbits result in time delays; they're lousy, too, at seeing the atmosphere vertically, supplying something more like a bird's-eye view than a cross section of a layer cake.

The National Weather Service installed most of the country's Doppler radars — more than 150 of them — during the mid-1990s. But gaps in coverage included a swath along Washington's coastline. "We couldn't see what was coming in from the ocean," Mass said. "It was all blank out there."

Mass on the roof of the Atmospheric Sciences-Geophysics Building at the University of Washington in Seattle.
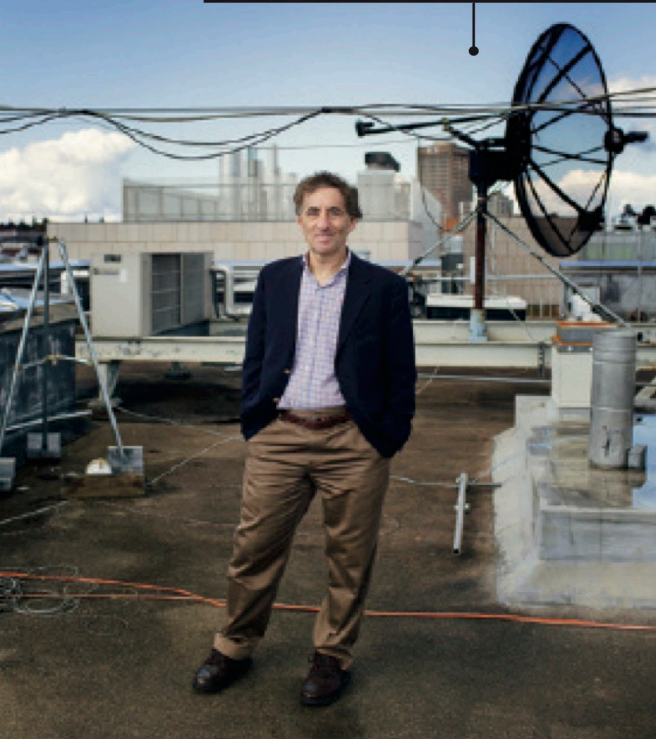
When Mass promoted his radar initiative on KUOW, viewers supported him, he said — including Maria Cantwell, the state's Democratic senator. Mass implored her to demand an explanation from Jack Hayes, then the director of the National Weather Service. Within a few months, Hayes had tracked down a surplus $4 million Doppler unit from the Air Force. "Mass is a very powerful voice," Hayes told me. "You might quibble with how he pounds his fist on the table, [and] his arguments are a little more emotional than they need to be. But at the end of the day, I always appreciate Cliff."

In 1995, Mass founded the Northwest Regional Modeling Consortium in Seattle to create better local forecasts. At the time, the National Weather Service used models with resolutions of 80 kilometers, or about 50 square miles. In Washington State, a sample that large might encompass topography ranging from 10,000-foot-high snowcapped volcanoes all the way down to the Pacific Ocean; the resulting forecast, forced to produce sweeping overgeneralizations, renders itself essentially useless. "My first test was at 27-kilometer resolution, and then I started doing it higher and I found something absolutely magical," Mass says. "We could actually forecast the local weather."

Jeff Renner, the chief meteorologist for Seattle's NBC television affiliate KING 5, who retired in April, told me, "It really allows us to see some beautiful detail." KING pays more than $20,000 a year to buy model data from the consortium, which it incorporates into its on-air forecasts. "[Mass] was doing a much better job than the Weather Service, where you were basically getting vanilla ice cream," Renner says. With the consortium, "it was like walking into Ben & Jerry's."

HURRICANES ARE PERHAPS the deadliest phenomena that forecasters are charged with modeling, but the high stakes mean that they can be useful in catalyzing change. In October 2012, the European Center's supercomputing cluster — the most powerful forecasting system in the world — correctly plotted Hurricane Sandy's path into the Mid-Atlantic United States eight days in advance, while the National Hurricane Center predicted the storm would veer harmlessly offshore. Al Roker told me, "In a sad way, it took something like Sandy for people to say, Wait a minute, this is crazy."

Because hurricanes form over oceans, where there are very few weather stations, predicting their routes and strength requires satellite data. What frustrates Roker and other meteorologists is that many satellites carry outdated technology and are nearing the ends of their life spans. "There is stuff circling the earth that's been up there for 20 to 30 years," Roker says. Replacements were scheduled to be sent into orbit last year, but their launches were postponed in part because the inspector general at the Department of Commerce, which oversees NOAA, found engineering and manufacturing defects in their components.

Mass argues that technological shortcomings are not the sole problem, however. He also blames "poor organization and poor leadership at NOAA — their efforts are divided into uncoordinated groups, each trying to protect their turf." According to Conrad Lautenbacher, who led NOAA from 2001 to 2008, "There is no orderly process to take some really great idea somebody has in research and turn it into something that the weather service can use." Dysfunctional, compartmentalized bureaucracy gets in the way.

Five months after Hurricane Sandy, Roker reported a story for "NBC Nightly News" in which he interviewed Mass about the inadequacies of the Weather Service. Lapenta, at the National Centers for Environmental Prediction, was on a treadmill at his gym a few days after NBC aired the report. "The woman next to me was working out with her trainer and talking about the difference between the G.F.S." — Global Forecast System, the National Weather Service's go-to model — "and the European model," Lapenta says. It was the first time in his 28-year career as an atmospheric scientist that Lapenta overheard the merits of competing weather models discussed in casual chitchat. He credits the change largely to Mass.

Even though Lapenta got his new supercomputers in January, most of the system remains idle. "It's extraordinary," Mass says. "They are only using a small portion of it." He also notes that the upgrades still aren't enough to run high-resolution ensembles effectively. Lapenta, however, remains optimistic; he's involved with NOAA's Next Generation Global Prediction Initiative, and he formed an advisory committee to evaluate the National Weather Service's numerous models and come up with a plan to build better ones. He invited Mass to join the 14-member team, which met for the first time in the summer of 2015.

"I came away very sobered," Mass says. "It's a real mess. They are running way too many models." At last count, the centers managed at least a dozen models, some in development, others already operating. None work very well. (They recently spent eight years and more than $100 million trying to fix their main hurricane-forecasting model. But it still performs so poorly that meteorologists inside the agency want to scrap it altogether.) Mass's advice was to focus all efforts on designing a single high-performance weather model that can be adapted to

Weather
(Continued from Page 29)

predict weather at a regional scale, similar to what the European Center and Britain's Met Office are already doing. But there is heavy resistance from the National Weather Service union to any associated downsizing. In a blog post Mass wrote, "When I talk to middle-level managers in the N.W.S., they complain about [its] powerful unions that slow down innovation and new ways of doing business."

In the summer of 2015, Mass visited Boulder to give a presentation at a conference on weather modeling. His 12-minute talk focused on the poor performance of regional models, which typically bring higher resolution to smaller geographical areas, and should, as a result, be better at predicting localized events, like flash floods and hurricanes. But they're not — a fact Mass demonstrated with PowerPoint slides of statistics. He concluded his presentation with a photo of a man doing a face palm, above which he had typed, "Em-bar-rass-ment: the shame you feel when your inadequacy or guilt is made public." The audience of 200 groaned. But when I chatted with scientists during a lunch break, they told me the chiding was expected — after all, this was Cliff Mass at the lectern. As a research meteorologist with NOAA put it, "Cliff is very excellent at being provocative."

The next evening I had dinner with Mass at the Chautauqua Dining Hall, which has a wraparound patio looking onto the Flatirons, a succession of towering rock slabs. We chose an outdoor venue because a forecast earlier in the day predicted a dry evening. "I'm hoping to get in a hike later," he said. Midway through our meal, lightning crackled overhead, followed by pelting rain. This was, Mass said, a classic example of a model's failing at the most basic level — it couldn't even forecast the weather just a few hours in advance.

As we ate, Mass ranted about the storms the Weather Service flubbed over the past winter. But Mass got most excited when I asked him to discuss high-resolution ensembles. The violent thunderstorms and frequent tornadoes (there were 1,259 in 2015) that routinely thrash the Great Plains and Midwest — killing more than 100 people each year over the last decade — could be predicted far more skillfully with high-resolution ensemble forecasts. So, too, could hurricanes like Matthew. "But we don't even have enough computer power to do it," Mass said. He opened an app on his iPhone called RadarScope, made by a private company, that projects where a storm is most likely headed.

"It looks like it's moving away," he said. "I can still get that hike in." Mass excused himself and set off on a nearby trail. He walked warily, stopping often to scan the clouds for clues to an impending downpour, all while imagining a forecast that someday would be able to simply tell him when and where it will rain. ◆