Abstract:

The article focuses on paleontology and discusses how luck has played a big role in many of the world's great fossil discoveries. It also discusses a new technique that improves the odds of finding ancient bones. It reports on the use of modern computer models that look for hidden patterns in satellite images and generate maps of fossils sites, helping fossil hunters narrow their search.

Keywords: Paleontology | Computer simulation | Remote-sensing images | Fossils | Paleoarchaeology

Article:

PALEONTOLOGY

Luck has played a big part in many of the world's great fossil discoveries New models predict where the bones are and put serendipity in the backseat

ON A BROILING DAY IN JULY 2009, A CARAVAN OF FOUR-WHEEL-DRIVE VEHICLES traveled a faint, two-track dirt road in southwestern Wyoming's Great Divide Basin. The expedition was headed for an area known as Salt Sage Draw in search of buried treasure: fossils dating to between 55 million and 50 million years ago, at the start of the Eocene epoch, when the ancestors of many modern orders of mammals were beginning to replace the more archaic mammals that had existed during the earlier Paleocene epoch. One of us (Anemone) had been leading field crews of anthropologists, paleontologists and geologists to the basin since 1994, and Salt Sage Draw had proved a fruitful hunting ground over the years, yielding fossils at several localities. Yet this time I was having trouble finding the site. It dawned on me that the road we were on was not the one we had used in previous years. My error would turn out to be very fortunate indeed.

As the tracks began to disappear in the sagebrush and tall grass, I stopped the caravan and walked a ways to see if I could spot the road ahead. Rounding a small hill, I noticed an extensive bed of sandstone in the near distance and the elusive road right alongside it. Because sandstone in the Great Divide Basin and many other sedimentary basins in the American West often harbors fossils, I decided to spend some time searching these deposits before we resumed our trip to Salt Sage Draw. After about an hour of systematically scanning the rock on hands and knees, my then graduate students Tim Held and Justin Gish shouted that they had found a couple of nice mammal jaws. I eagerly joined them. Fossil jaws with teeth are prized because they contain enough information to identify the kind of animal they came from, even in the absence of other parts of the skeleton, and because they reveal what the animal ate.

What came next can only be described as every paleontologist's dream. My students had located a fossil "hotspot." But this was no ordinary hotspot with a handful of jaws or a few dozen teeth and bones eroding out of the sandstone. Rather they had found an extraordinary trove from which we have now collected nearly 500 well-preserved jaws and several thousand teeth and bones from more than 20 different fossil mammal species that lived here approximately 50 million years ago. We call the spot "Tim's Confession," and today it remains not only our best site in the Great Divide Basin but also one of the richest caches of early Eocene mammals in the entire American West.

Mine is hardly the first team to make a major fossil discovery more or less by accident. The history of paleontology is littered with such tales of serendipity. In fact, the ways that vertebrate paleontologists attempt to locate productive fossil sites have not changed much since the earliest days of our science. Like the 19th-century pioneers of our field, we use geologic and topographic evidence to determine where we might have the best chance of finding fossils eroding out of ancient sediments. But beyond that, whether we hit pay dirt is still largely a matter of luck, and more often than not the hard work of looking for fossils goes unrewarded.

Our experience at Tim's Confession got me thinking about whether there might be a better way to determine where my field crew should spend its efforts searching for new fossil sites. We knew that the fossils we were interested in occur in sandstone dating to between 55 million and 50 million years ago, and we knew where in the basin some of these sedimentary layers were exposed and thus suitable for exploration. But although that information helped to narrow our search somewhat, it still left thousands of square kilometers of ground to cover and plenty of opportunities to come up empty-handed.

Then one night in camp, an idea began to germinate. Out in the field, kilometers away from the nearest source of light pollution, we often noticed satellites passing overhead. I wondered whether we could somehow combine our expert knowledge of the local geology, topography and paleontology of the Great Divide Basin with a satellite's view of the entire 10,000-square-kilometer area to, in essence, map its probable fossil hotspots. Perhaps satellites could "see"

features of the land invisible to the naked eye that could help us find more sandstone outcrops and distinguish those that contain accessible fossils from those that do not.

EYES IN THE SKY

OTHER PALEONTOLOGISTS, of course, have speculated about whether satellite imagery might improve our ability to find fossils in the field. As a specialist in the fossil record of primate and human evolution, I knew that in the 1990s, Berhane Asfaw of the Rift Valley Research Service and his colleagues had used such images to identify rock exposures in Ethiopia that might yield fossils of human ancestors. At around the same time, Richard Stucky of the Denver Museum of Nature & Science demonstrated that different rock units in the fossil-rich Wind River Basin in central Wyoming could be distinguished and mapped based on analysis of satellite imagery of the region. Both these projects involved collaborations between paleontologists and remote-sensing specialists from NASA and proved the value of such cross-disciplinary efforts. But I wondered if there was a way to tease more information out of the satellite images and thus better focus our search.

I turned to a geographer, the other author of this article (Emerson), and the two of us soon sketched out a plan. We would obtain freely available images of the basin from the Landsat 7 satellite and its so-called Enhanced Thematic Mapper Plus sensor, which detects radiation reflected or emitted from the earth's surface in wavelengths spanning the electromagnetic spectrum -- from the blue to the infrared -- and represents it in eight discrete spectral bands. The bands can be used to distinguish soil from vegetation, for example, or to map mineral deposits. Then we would develop a method that would allow us to characterize the radiation profiles of known productive fossil localities in the Great Divide Basin based on satellite imagery and see if they shared a telltale spectral signature. If so, we could search the entire Great Divide Basin from our computers to locate new sites that share this spectral signature and thus have a high probability of bearing fossils. We could then visit those places (as well as places with different spectral signatures) in person and exhaustively search them for fossils to test the model.

Determining whether our known fossil sites shared a distinctive spectral signature was no small task, because for each site we had to assess the combination of values in six bands of the electromagnetic spectrum provided by the Landsat data. Our problem was essentially one of pattern recognition in multiple dimensions, something that humans do not do particularly well but that computers excel at. So we enlisted a so-called artificial neural network -- a computational model capable of learning complex patterns.

Our artificial neural network revealed that the basin's known fossil sites do indeed share a spectral signature, and it was able to easily tell these sandstone localities apart from other types of ground cover, such as wetlands and sand dunes. But the model had its limitations. Neural networks, by their very nature, are analytical "black boxes," meaning they can distinguish patterns, but they do not reveal the actual factors that allow different patterns to be distinguished.

So whereas our neural network could easily and accurately distinguish fossil localities from wetlands or sand dunes, it could not tell us how the spectral signatures of different land covers actually differed in the six bands of the Landsat data -- information that could conceivably help us conduct a more targeted search. Another limitation of the neural network approach is that it is based entirely on the analysis of individual pixels. The problem is that the area of an individual Landsat pixel, which measures 225 square meters, does not necessarily correspond to the size of a fossil locality: some localities are larger than an individual pixel; some are smaller. Thus, the neural network's predictions about the location and extent of potential fossil sites (or a certain type of ground cover, for that matter) do not always match up with reality.

To overcome these constraints, we needed to be able to analyze multiple adjacent and spectrally similar pixels and to statistically describe the distinctive spectral signature of the entire area, whether it was a fossil site or a forest. We turned to a technique known as geographical object-based image analysis and to commercially available, high-resolution satellite imagery in which individual pixels were less than one meter in diameter. Unlike an artificial neural network, this approach allows satellite images to be segmented into image objects -- that is, groups of spectrally homogeneous pixels -- that can then be characterized by statistical parameters such as mean or median brightness or texture. These image objects more closely match points of interest on the ground, such as fossil sites or stands of forest. Using this image-analysis technique, we were able to develop an independent set of predictions about where to find fossils.

MOMENT OF TRUTH

BOTH OUR PREDICTIVE MODELS yielded maps of the Great Divide Basin that pinpointed unexplored areas whose spectral signatures most closely resembled those of the known localities. Although the models exhibited a good degree of overlap in their predictions, they also diverged in some cases. We chose to focus on those places that both models identified as high-priority potential sites. Maps in hand, we headed out to Wyoming during the summers of 2012 and 2013 to see if our models would lead us to new fossil caches in the Great Divide Basin. Gratifyingly, they did exactly that.

The artificial neural network model turns out to be extremely efficient at identifying sandstone deposits, which are almost always worth exploring because so many of the ones in this basin contain fossil vertebrates. One of the first sandstones it led us to in July 2012 yielded a dozen fossils of characteristic Eocene mammals, including the five-toed horse Hyracotherium, the early primate Cantius and several other creatures belonging to an extinct group of hoofed mammals known as the Condylarthra. The neural network also guided us to several spots that yielded aquatic fossil vertebrates, including fish, crocodiles and turtles.

Our geographical object-based image analysis model took us to new sites, too. After a slow start in which the first three or four places the model pointed us to gave up no fossils, we moved to the northern part of the Great Divide Basin, near a place called Freighter Gap, for a week of

intensive "ground truthing" of our new technique. Graduate student Bryan Bommersbach, who a week before had led us on a long hike to a place that was entirely barren of fossils (we dubbed it "Bryan's Folly"), took the lead in choosing which areas to survey based on the model's predictions. Almost immediately, we began to find bones at many of these locations. We searched for remains at 31 separate places on the landscape that our model indicated were spectrally similar to known localities and found vertebrate fossils at 25 of these places, which is a much higher success rate than is typical when surveying without the help of a predictive map. Mammal fossils emerged from 10 of these localities, one of which dates to the latest part of the Paleocene -- an extremely rare find.

We have every reason to believe that predictive models akin to the ones we developed will work in regions other than the Great Divide Basin. In fact, they should work virtually anywhere in the world. In theory, as long as one has satellite images of the region in question and a handful of known fossil localities with which to train the model, one can generate a custom map showing those spots in the region that are likely to contain fossils of interest.

In a conservative test of this approach, we used the neural network we developed for the Great Divide Basin to predict the locations of fossil-bearing sedimentary deposits in the nearby Bison Basin, which is known to harbor Paleocene mammal fossils. (We did not train the model with fossil sites specifically from Bison Basin, because it contains the same kinds of fossil deposits as the Great Divide Basin.) Encouragingly, our neural network predicted the three most productive fossil localities known in the Bison Basin. Thus, a field crew exploring this vast area for the first time using our predictive model would have had a far better chance of discovering these sites than a crew using traditional survey methods.

Our trial runs in 2012 and 2013 in Wyoming showed that the use of satellite imagery in combination with geospatial predictive models greatly increased the effectiveness of our fieldwork, helping us to find more fossils in less time. But we still have more to do. We are now focused on refining our models to better characterize and differentiate the spectral signature of productive localities. And we are working on ways to apply more constraints to our predictive models to limit the number of false positive results in the maps we generate and thus improve our ability to determine the highest-priority areas to survey.

We are convinced that with these tools we can put the future of paleontological exploration on a more secure and scientific footing and reduce the role of serendipity in finding important fossils. Achieving that goal will be well worth the effort required. Piecing together the origin and evolution of life on earth is far too interesting and important an endeavor to leave to chance. And we can't afford to wait another 15 years to find the next Tim's Confession.

MORE TO EXPLORE

GIS and Paleoanthropology: Incorporating New Approaches from the Geospatial Sciences in the Analysis of Primate and Human Evolution. R. L Anemone, G. C. Conroy and C. W.

Emerson in American Journal of Physical Anthropology, Vol. 54, Supplement No. 53, pages 19-46; 2011.

Finding Fossils in New Ways: An Artificial Neural Network Approach to Predicting the Location of Productive Fossil Localities. Robert Anemone, Charles Emerson and Glenn Conroy in Evolutionary Anthropology, Vol. 20, No. 5, pages 169-180; September/October 2011.

An Artificial Neural Network-Based Approach to Identifying Mammalian Fossil Localities in the Great Divide Basin, Wyoming. Charles W. Emerson and Robert L. Anemone in Remote Sensing Letters, Vol. 3, No. 5, pages 453-460; 2012.

For more on this research, including its funding source, visit ScientificAmerican.com/may2014/anemone

FROM OUR ARCHIVES

First of Our Kind. Kate Wong; April 2012.

IN BRIEF

For more than a century paleontologists have used geologic and topographic information to inform their search for fossils. Yet the discovery of fossils is still largely a matter of luck.

New computer models that look for hidden patterns in satellite images can generate maps of where fossils are likely to be located, thus helping fossil hunters narrow their search.

Ground truthing of such predictive maps in the American West has shown that they do indeed improve the odds of finding fossil sites. In theory, this approach could be used anywhere in the world.

FINDINGS

Treasure Map

Computer models can analyze satellite images of an area's known fossil sites to identify their shared radiation profile. The models can then assess the broader region to find other spots that share that profile and thus may harbor fossils of interest. This technique enabled the two of us to generate a predictive map of fossil localities (red) in the Great Divide Basin that helped to guide our search for fossils there (yellow). Restricting our surveying to focus on some of these areas (blue) greatly increased our success rate in finding fossil sites (green), compared with that of previous expeditions conducted without such a map.

PICTURE IS OMITTED FROM THIS FORMATTED DOCUMENT

MAP: Treasure Map

PICTURE IS OMITTED FROM THIS FORMATTED DOCUMENT

PHOTO (COLOR): FOSSIL HUNTING in an area as vast as Wyoming's Great Divide Basin (pictured here) has long been akin to searching for the proverbial needle in the haystack. But a new technique improves the odds of finding ancient bones.

PICTURE IS OMITTED FROM THIS FORMATTED DOCUMENT

PHOTO (COLOR): PALEONTOLOGISTS ACCIDENTALLY found a trove of 50-million-year-old fossils at a site dubbed Tim's Confession in Wyoming's Great Divide Basin in 2009.

PICTURE IS OMITTED FROM THIS FORMATTED DOCUMENT

PHOTO (COLOR): Among the finds were hundreds of well-preserved mammal jaws.

PICTURE IS OMITTED FROM THIS FORMATTED DOCUMENT

PHOTO (COLOR): Computer models have since enabled the team to focus its efforts in those areas most likely to yield fossils, including this spot south of an extinct volcano known as the Boar's Tusk (below).

PICTURE IS OMITTED FROM THIS FORMATTED DOCUMENT

PHOTO (COLOR)

PICTURE IS OMITTED FROM THIS FORMATTED DOCUMENT

PHOTO (COLOR)