Science Societies

More accurate predictions of daily soil temperature

By Kaine Korzekwa

July 11, 2022



Photo courtesy of Adobe Stock/shintartanya.

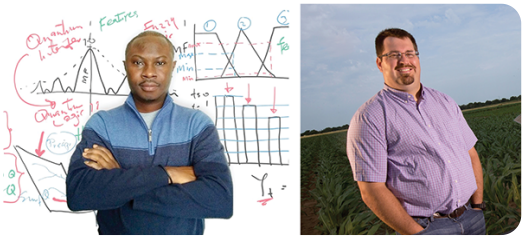
Accurately predicting daily soil temperature is essential for understanding everything from crop production to carbon sequestration. However, predictions become especially difficult as the soil depth increases.

The ability to predict soil temperature at deeper depths across multiple environments can help with many areas of agronomy, soil science, climate science, and geology.

By implementing machine-learning approaches and unique data transformations, new research published in *Vadose Zone Journal* describes a way to maintain prediction accuracy of daily soil temperature with increasing soil depth.

Accurately predicting daily soil temperature is essential for understanding everything from crop production to carbon sequestration. These predictions become especially difficult as the soil depth increases. By implementing machine-learning approaches and unique data transformations, new **research published in Vadose Zone Journal** describes a way to maintain prediction accuracy of daily soil temperature with increasing soil depth.

The crux of their work lies in combining what are called knowledge-based and machine learning approaches with novel uses of data that take into account on-the-ground knowledge and relatively simple theories of heat transfer in soil.



Olufemi Abimbola (left) was lead author and Trenton Franz (right) lead faculty member of the Vadose Zone Journal article. Photos courtesy of Olufemi Abimbola (left) and the University of Nebraska–Lincoln (right).

"The aim of the research was to find a way to improve the prediction of daily soil temperature, especially in the root zone of the typical crop," explains lead author Olufemi Abimbola, who is now a senior data scientist with Syngenta. "Most models out there don't do a good job of predicting the daily soil temperature in the root zone, and I learned that is because of their approach. The few models that had high prediction accuracies used monthly weather data for predicting monthly soil temperature."

He adds that the equations typically used in soil temperature prediction models use parameters that are hard to measure and also vary widely in space, such as across just a few feet. There was a need to capture all of these variations without having to know exactly what those parameters were, he says.

Looking at other models, the researchers found that the vast majority had the problem of increasing prediction errors at increasing depth—it simply gets more difficult to use something like air temperature above the soil or soil surface temperature to predict soil temperature at deeper levels.

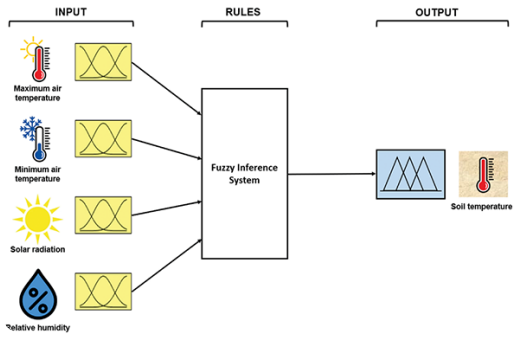
"I discovered the problem had nothing to do with the kind of model they used, and it had little to do with how complex the algorithms and models were," Abimbola explains. "More complexity and more data and more algorithms did not help increase the

accuracy at deeper depths. It had more to do with the need to bring in a knowledge-based approach that accounted for the way heat transfers through the soil."

Harnessing the Power of Machine-Learning Approaches

Trenton Franz, an associate professor and hydrogeophysicist at the University of Nebraska–Lincoln, served as the lead faculty member on this work. He notes the power of machine-learning approaches and their use for expanding soil temperature observations.

"We are always limited by how many observations we can feed to weather models and improve their accuracy," he says. "Machine-learning approaches appear to be a huge advance in statistical techniques compared with classical regression techniques. However, they need to be constrained and vetted with discipline knowledge, so the results can pass the basic 'smell test.' Use of machine learning without vetting by discipline knowledge is not advisable and can lead to poor predications as they are often black box approaches."



The adaptive neuro-fuzzy inference system in this study uses inputs like maximum temperature, minimum temperature, solar radiation, and relative humidity in conjunction with a fuzzy inference model to generate an output. Illustration courtesy of

Olufemi Abimbola.

The most accurate model the researchers came up with does not use special inputs that are different from what has been used in existing models. The only difference is the way Abimbola transformed the input variables. When looking over high quality weather station data, he noticed a lag between, for example, how a hot day impacts soil on the surface versus at deeper depths. It wasn't enough to know how the daily temperature was going up and down; it was also important to know how that heat made its way into the soil.

"I could see these lags from the topsoil to the next layer and the next layer," he says. "So I could see, say, 1 m down what the specific lags were from the peak temperature at the soil surface. A spike in air temperature that strikes the soil surface will take time to travel to 1 m down. The heat is traveling down and spreading out and weakening. This means that if I wanted to predict the soil temperature 1 m down on a particular day, I wouldn't use that day's air temperature. I would need to go back in time using the lag data to see what day's air temperature may have filtered down to 1 m by now. In my analysis, I found out how far to go back in time in the model for each depth. The lags had been used before but not quite in this way with other transformations."

The researchers' data also showed that as the depth increased, the variability and fluctuations are smaller than they are on the surface. The insulating effects of the soil mean that amplitude is very high at the topsoil as the temperature responds more fully to the air temperature, but the effect evens out the deeper the soil becomes.

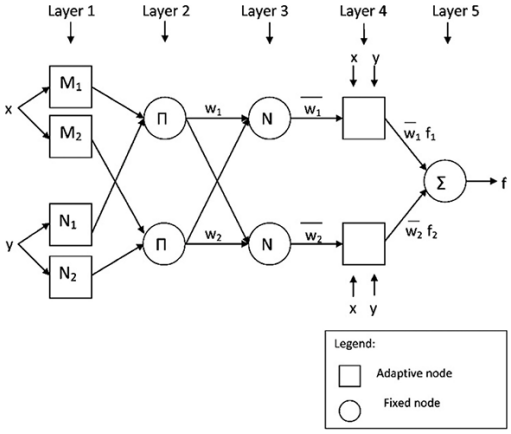
In order to achieve more generalizable results, the researchers also incorporated moving averages of meteorological features—such as maximum temperature, minimum temperature, mean temperature, solar radiation, and relative humidity—that were best to use at each soil depth. This smoothed the data and filtered out noise from random

daily fluctuations. For example, this uniquely accounted for how a shorter moving average may work best near the soil surface, but a longer one may be more predictive at deeper depths.

Adaptive Neuro-Fuzzy Inference System

The exact machine-learning algorithm the scientists found worked the best is called an adaptive neuro-fuzzy inference system. This combines the benefits of both an artificial neural network and fuzzy logic. Artificial neural networks have been used to predict soil temperature in other studies and function by using individual nodes, modeled after neurons, that function independently and signal each other to generate an output. Fuzzy logic is a method of generating models that is able to incorporate expert knowledge and does not require extremely clear-cut data. Combining the two into an adaptive neuro-fuzzy inference system was the best of both worlds, Abimbola says.

"Rather than a more crisp approach where something is one or the other, a 1 or a 0 in the model, fuzzy logic allows for more variation—a fuzzy middle where it can be partially 0 and partially 1 with different degrees of membership," he explains. "The sky may be blue today, and my shirt may be a different blue. But how blue is each blue? What version of blue will we consider blue? Rather than being a 1 or a 0, there can be a scale between 1 and 0 of how blue are they. The same can be applied to temperature and other variables."



A schematic of an adaptive neuro-fuzzy inference system. It illustrates the different nodes and levels of the model. Photo courtesy of Olufemi Abimbola.

The researchers were also able to have their model generalize to the whole United States. Abimbola found that he could use the latitude of any site to know the solar inclination, which is the intensity of the sunlight hitting the soil in that area. It is most intense at the Equator, for example. By using the maximum temperature and the latitude, he learned the lag times, and hence, the soil temperature can be predicted for any location in the United States. The model is generally accurate regardless of climate or soil type, he adds. They hope that interested researchers will apply this model in their work.

“Overall, soil properties and precipitation do not disrupt the physics enough to matter in general,” he says. “However, the model does not account for if soil is in an urban environment or something like a forested area. There may be a need to dive deeper into one location for more extreme accuracy, but we found our model to be generally accurate enough for many uses using our inputs and transformations.”

Benefits to Multiple Disciplines

The ability to predict soil temperature at deeper depths across multiple environments can help with many areas of agronomy, soil science, climate science, and geology, Franz says. While the biggest application may be a better soil temperature dataset and method, submodels that focus on soil carbon, yield forecasting, and hydrological work will also benefit from a more robust dataset.

Soil temperature is critical for seed germination, for example, and later in the growth cycle the root zone deeper in the soil is performing a myriad of chemical reactions, such as those in the nitrogen cycle, that are tied to soil temperature. Soil processes driven by soil temperature also play a role in the assessment of soil carbon, which has implications for climate modelling and initiatives like carbon credits, which are growing in popularity.

“Everything is related to soil, even something like carbon credits,” Abimbola says. “It will become increasingly important to have a way to predict what is going on deeper in the soil that impacts carbon storage and movement.”

Abimbola and Franz believe that they and others are just scratching the surface of how machine learning can be applied to multiple fields.

“I think the idea of taking machine-learning approaches and using them to investigate problems in other disciplines is wide open,” Franz says. “It is thrilling to see what new insights can be found.”

Dig deeper

Read the original article, “Knowledge-Guided Machine Learning for Improving Daily Soil Temperature Prediction Across the United States,” in *Vadose Zone Journal* at **https://doi.org/10.1002/vzj2.20151**.

More science

Back to issue

Back to home

Text © . The authors. CC BY-NC-ND 4.0. Except where otherwise noted, images are subject to copyright. Any reuse without express permission from the copyright owner is prohibited.