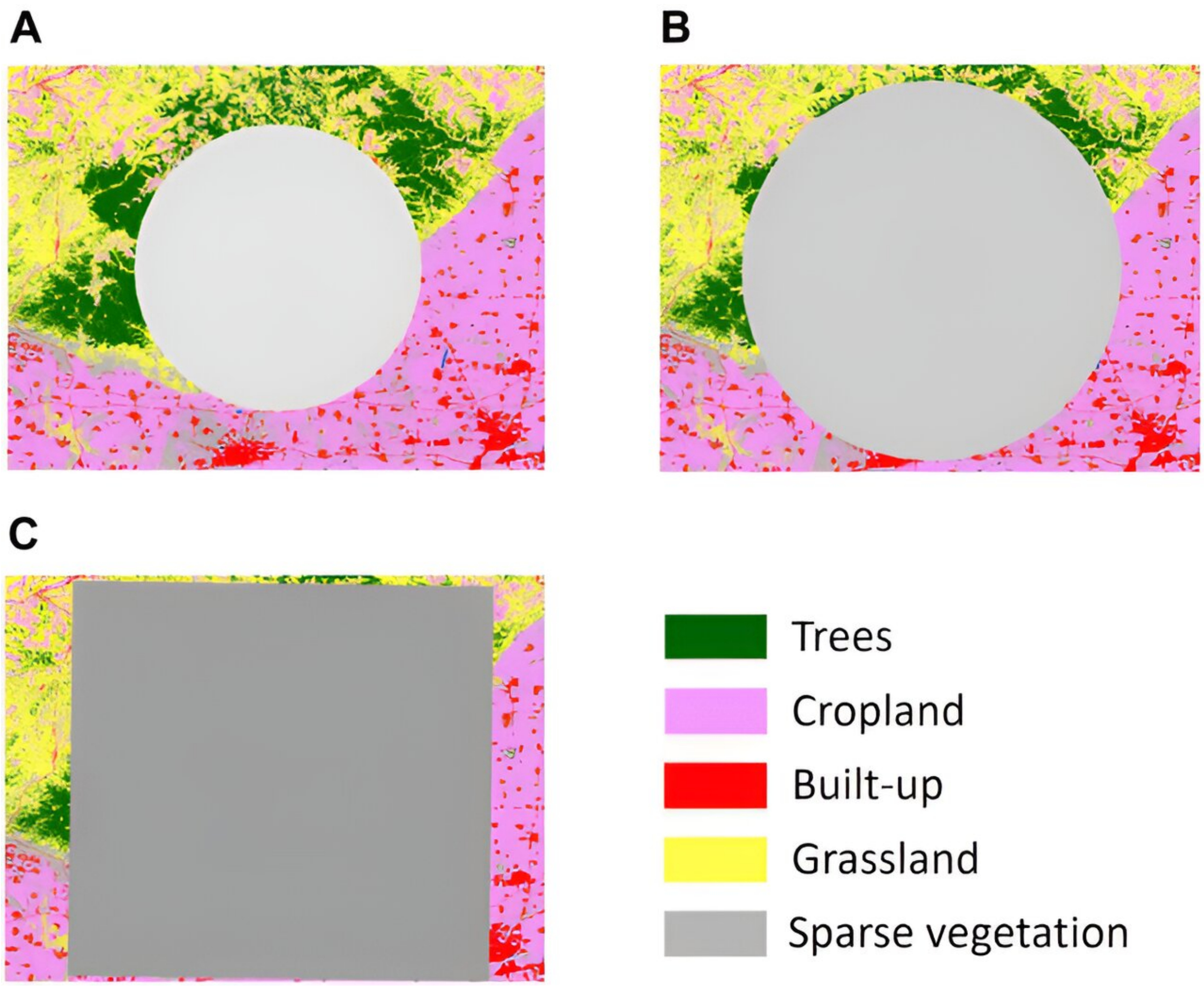
Machine learning and better radar solve the 'cloud cover' problem

August 6 2024



Three scales of LST deficiency for simulated cloud contamination. (A) Low-scale absence (missing 30%). (B) Medium-scale absence (missing 51%). (C) High-scale absence (missing 75%). Credit: *Journal of Remote Sensing* (2024). DOI: 10.34133/remotesensing.0071

Clouds have for decades been a bugbear for remote sensing of land surface temperature—one of the most important earth system metrics, used in everything from tracking climate change to predicting wildfires. A new approach incorporating machine learning appears to have solved this challenge

Land surface temperature tracking via remote sensing is often bedeviled by cloud cover. Traditional techniques that make educated guesses at temperatures beneath the clouds are helpful but suffer from significant errors. However, a new approach that incorporates a novel radar technique, better elevation models and machine learning looks set to be a game-changer.

A paper describing the new approach was published in the *Journal of Remote Sensing* on June 26.

Measurement of the surface temperature of the Earth by satellite or aerial sensors ("remote sensing") is crucial for monitoring a wide range of environmental conditions, from mapping vegetation health to assessing water stress, monitoring climate change and predicting wildfires. It is utterly essential within agriculture, where it informs irrigation practices, and is used by policy makers and planners to observe weather patterns, study the urban heat island effect, and predict natural disasters.

Land surface temperature is simply one of the most important metrics that society needs to track.

Unfortunately, cloud cover presents significant challenges in remote sensing by obstructing the view of the Earth's surface, leading to substantial gaps in data collection. This interference can result in inaccurate or incomplete measurements of land surface temperature. Particularly in regions with persistent cloudiness, researchers often struggle to obtain reliable data.

In order to overcome the cloud-cover problem, until now, researchers and others who monitor land surface temperature have used interpolation methods. Wherever there is a gap in the data collected as a result of cloud cover, the observations either side of the clouds are used to predict what the temperature must be under the clouds. This can be done temporally as well as spatially, where land surface temperature measurements from adjacent time periods are used to fill gaps during cloudy conditions.

"So long as the spatial or temporal gap is small, any errors are pretty small," said Jingbo Li, lead author of the study and a Ph.D. student with the Key Laboratory of Quantitative Remote Sensing at the Beijing Academy of Agriculture and Forestry Sciences. "But when the gap is big, we start seeing significant errors."

This is particularly a problem for regions with high variation in topography, as land surface temperature is profoundly affected by altitude, and wherever more complex cloud formations are found.

As a result, as much as half of all land surface temperature readings are contaminated by cloud cover problems. The situation is even worse in mid and low latitudes.

"This is not a minor issue only of interest to scientists. Poor land surface temperature remote sensing undermines decision making by political leaders, public health officials and security agencies," added Yang Guijun, a co-author of the study and professor in the same laboratory.

However, in recent years, advanced algorithms and models have been developed to reconstruct land surface temperature, incorporating machine learning and deep learning techniques to enhance prediction accuracy. In addition, synthetic aperture radar (SAR) has gained attention as a complementary tool, as it can penetrate cloud cover and provide valuable information for reconstructing optical data.

SAR operates by emitting microwave signals towards the ground and measuring the backscattered signals that bounce back from the surface. But unlike traditional radar systems that rely on a physical antenna size to determine resolution, SAR simulates a larger antenna by moving the radar platform (such as an aircraft or satellite) along a flight path. This movement allows the system to collect multiple radar echoes from the same target area, effectively creating a "synthetic" aperture that enhances image resolution.

This same ability to collect multiple radar echoes from the same target area means that it is more capable of penetrating clouds and fog, but also rain. Moreover, two or more SAR images can be compared to detect changes in the Earth's surface, such as land subsidence, deformation, or changes in vegetation.

Separately, what are called Digital Elevation Models (DEMs)—digital representations of the Earth's surface topography or terrain—can also assist. These consist of a grid of elevation values, where each cell in the grid corresponds to a specific geographic location and contains information about the height and slope of the terrain above a reference level, typically sea level. And elevation and slope, as well as where and when a given slope faces the sun, significantly influence local temperatures.

For example, higher elevations generally experience cooler temperatures, while slopes facing the sun (south-facing in the northern hemisphere) may receive more solar radiation, leading to higher temperatures. The topographic features captured in DEMs allow for a more nuanced understanding of how terrain affects temperature, improved interpolation, and lead to better predictions of land surface temperature in the very complex landscapes that have caused the biggest headaches.

And so the research team integrated DEMs with SAR in order to enhance the accuracy of their land surface temperature models, and also incorporated machine learning to aid in recognition of complex relationships and patterns that may not be observable by humans alone. In addition, machine learning models tend to be more resilient to noise and outliers in the data than traditional methods.

They call their new model, the Synthetic Aperture Radar and Digital Elevation Model-integrated Land Surface Temperature (SDX-LST) reconstruction model.

The authors of the paper developed a novel model called the Synthetic Aperture Radar and Digital Elevation Model-integrated Land Surface Temperature (SDX-LST) reconstruction model. It generates clear-sky land surface temperature (LST) data even in regions affected by extensive cloud cover, at high-resolution (down to 30 meters). They then validated the effectiveness against high-quality surface temperature data gathered not remotely from satellites or planes, but ground-based stations, and found it was consistent.

Finally, they tested their model across the Loess Plateau southeast of the Gobi Desert, the Qinghai-Tibet Plateau, China's northeast and northern plains, the Nanling Mountains, and finally Desert Rock, in Nevada. The aim was to trial it against as wide a range of longitude, latitude, topography, landform, and vegetation cover types as possible, and at multiple dates and times throughout the year.

Across areas, their SDX-LST model enabled precise predictions of land surface temperature regardless of level of cloud cover, vegetation type and complexity of terrain.

Taking their model to the next level, the researchers are now aiming to improve its responsiveness to temporal factors.

More information: Jingbo Li et al, Generating Spatiotemporal Seamless Data of Clear-Sky Land Surface Temperature Using Synthetic Aperture Radar, Digital Elevation Mode, and Machine Learning over Vegetation Areas, *Journal of Remote Sensing* (2024). DOI: 10.34133/remotesensing.0071

Provided by Journal of Remote Sensing