

Groundwater studies can be tainted by 'survivor bias'

September 5 2019



Credit: CC0 Public Domain

Bad wells tend to get excluded from studies on groundwater levels, a problem that could skew results everywhere monitoring is used to decide government policies and spending.

Researchers at the University of Waterloo uncovered the problem while examining a discrepancy between [scientific data](#) and anecdotal evidence in southern India.

Reports on thousands of wells and [satellite images](#) taken between 1996 and 2016 suggested groundwater levels were rising, good news in an area where it is vitally important for agriculture.

At the same time, however, fieldworkers were hearing more stories from farmers about wells running dry, suggesting levels were actually declining.

Researchers solved the apparent paradox by first obtaining [census data](#) that backed up the anecdotal evidence. It showed, for example, that more farmers were digging expensive deep wells in the hard-rock aquifer.

"If indeed groundwater levels are going up, why would farmers choose to pay more and dig deeper wells?" asked Nandita Basu, a civil and environmental engineering professor. "It didn't make sense."

Researchers then examined the well data and found that those with missing water level data were often excluded from analysis because they were considered unreliable.

When the excluded wells were added back into the mix, the results confirmed the evidence from farmers that [groundwater](#) levels were decreasing, not increasing.

"They were systematically picking the wells with a lot of data and potentially ignoring the wells that were going dry because they had incomplete data," said Tejasvi Hora, an engineering Ph.D. student who led the research.

The culprit was identified as something called 'survivor bias,' a statistical phenomenon that results in the exclusion of negative data.

When wells ran dry, there were no water levels to report. That created gaps in reports for those wells, and their incomplete data was then discarded as inferior to the complete data from good wells that hadn't run dry.

Basu, also a professor of earth and environmental sciences and a member of the Water Institute at Waterloo, said the lesson from southern India is applicable anywhere in the world that [groundwater levels](#) are monitored and analyzed.

"Our main point is that bad data is good data," she said. "When you have wells with a lot of missing data points, that is telling you something important. Take notice of it."

"Whenever you're focusing only on complete data, you should take a step back and ask if there is a reason for the incomplete data, a systematic bias in your data source," Hora said.

Basu and Hora collaborated with Veena Srinivasan, a researcher at an environmental think tank in India.

A paper on their work, The Groundwater Recovery Paradox in South India, appears in the journal *Geophysical Research Letters*.

More information: Tejasvi Hora et al, The Groundwater Recovery Paradox in South India, *Geophysical Research Letters* (2019). [DOI: 10.1029/2019GL083525](https://doi.org/10.1029/2019GL083525)

Provided by University of Waterloo

Citation: Groundwater studies can be tainted by 'survivor bias' (2019, September 5) retrieved 27 April 2024 from <https://phys.org/news/2019-09-groundwater-tainted-survivor-bias.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.