Where Should I Sample Next? Adaptive Sampling Algorithm for Rapid and Efficient Autonomous Screening of Soil Contaminant Distributions at Large Sites

Gregory V. Lowry (glowry@cmu.edu), Shane Deng, Sean Wang, and Aaron Johnson
(Carnegie Mellon University, Pittsburgh, PA)
Sung-woo Lee, Kavish Ganesh, and James Culp (Geosyntec Consultants, Pittsburgh, PA)
Thomas Hoelen and Natasha Sihota (Chevron Technology Company)

An accurate understanding of contamination distribution in the soil is critical for the selection of remediation methods at impacted sites. This is typically done by trained personnel manually sampling soils at the site, followed by either field analysis with a portable sensor or shipping selected samples to an off-site laboratory for analysis. The data are then used to create a contaminant distribution, potentially leading to one or more additional rounds of sampling to delineate the site. This process can take a long time and can be costly. Such a manual process is also very expensive, and time-consuming and often relies on intuition-based sampling location selection by operators rather than scientific methods. Recent advancements in terrestrial robot mobility, sensors, and machine learning algorithms provide more possibilities for robotic systems to autonomously explore the unknown environment with obstacle avoidance, intelligently select the next sampling location, and manipulating the soil to perform accurate analysis. Here, we developed an adaptive sampling algorithm that can be used to rapidly delineate the contaminant distribution at the site based on in situ field measurements by a portable sensor attached to a robotic system or measured by hand. The algorithm uses a Gaussian process to identify the next sampling location that will provide the most value of information and lower the uncertainty in the contaminant distribution. Algorithms can be optimized to delineate “hot spots” or can be optimized to provide maximum coverage of the overall investigation area. By building in constraints like minimizing the number of samples, or minimizing the total travel distance of the robot, the algorithm can efficiently delineate the site contamination.