Analysis of Remotely Sensed Datasets to Detect Changes in Waste Sites

December 2017

J. Tagestad, P. Royer, K. Larson
DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTEN
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01839

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 52, Oak Ridge, TN 37831 0062:
ph: (865) 576 8401
fax: (865) 576 5728
e-mail: reports@adonis.ornl.gov

Available to the public from the National Technical Information Service
500 Shannon Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
e-mail: nits@ntis.gov <http://www.ntis.gov/about/form.aspx>
Online ordering: http://www.ntis.gov

This document was printed on recycled paper.
(8/2013)
Analysis of Remotely Sensed Datasets to Detect Changes in Waste Sites

J. Tagestad, P. Royer, K. Larson

December 2017

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99352
1.0 Executive Summary

There is an ongoing concern about possible compromise at burial sites containing radioactive waste. The need to provide timely alerts related to land surface change has emerged as a management priority following the Hanford Site PUREX partial tunnel collapse event in May, 2017. This work evaluated various remote sensing techniques for detecting surficial changes and land deformation at the U.S. Department of Energy’s Hanford Site.

LiDAR-derived land surface data was evaluated for a portion of the Site from 2008 and 2010. LiDAR was very effective in detecting changes in elevation related to waste activities. The high accuracy and easy interpretation of LiDAR make it a powerful tool for mapping vertical structure across relatively large areas (airborne) or in small focused missions (unmanned aircraft). However, the high cost of mobilizing airborne LiDAR missions, and security and airspace approval considerations for unmanned missions, require further investigation in order to propose a path forward using LiDAR.

Change detection analysis was performed using satellite multispectral data from two sources. Broad scale analysis using moderate-resolution (10 meter) satellite data was used to detect changes in surficial condition across the entire Hanford Site between July and August 2017; including burned areas, construction activities and landfill modification. A higher resolution analysis using commercial high-resolution (3 meter) multispectral data for the 200 areas was used to detect changes in surface appearance at a site coincident with the PUREX tunnel collapse. The low cost, mature algorithms and broad selection of spatial, temporal, and spectral resolution available from multispectral satellite platforms makes it an attractive tool for monitoring.

Potential land surface deformation was also detected at focal locations on the Hanford Site using publically available satellite Synthetic Aperture Radar (SAR) imagery, although the PUREX tunnel collapse was not detected using this method, likely because the spatial resolution of SAR imagery was too coarse. The localities of potential deformation observed in SAR analysis were primarily in undeveloped portions of the Hanford landscape in areas known to experience wind-driven soil erosion. It is important to note that we performed this analysis without ground-based observations or direct validation, and as such, there is a level of uncertainty about detected areas of change, because we lack a priori data about activities that may have occurred during the time periods examined. Ground-based cross validation using historical, ongoing, or planned management activities would add rigor and substantial support for this work. Various options for verification and validation are described in sections 7 and 8 of this document.

Findings of this study suggest that LiDAR, multispectral imagery and space-borne SAR can be used and are complementary to detect surficial changes that may be of potential concern in cleanup and waste management operations. These methods are capable of detecting changes that may go unnoticed for potentially long periods of time using traditional ground observations. These methods can feasibly be implemented to achieve a continuous and potentially automated monitoring program, which may save resources and time as well as reduce risk associated with ground-based monitoring.
2.0 Background

The May 2017 partial collapse of the PUREX tunnel, which holds eight rail cars loaded with highly radioactive contaminated equipment, has underscored the need for continuous burial site monitoring across the Hanford Site. Current ground-based monitoring is time consuming, often performed ad hoc, and expensive to perform across large areas. Consequently, investigation of indirect monitoring methods that are less cost and time prohibitive has emerged as a research priority. Remote sensing technologies such as aerial or space-borne multispectral and hyperspectral imagery, Synthetic Aperture Radar (SAR), and Light Detection and Ranging (LiDAR) can be used to provide early warning for many types of waste site compromise. The ability to automate analysis of data streams from these technologies also makes it possible to develop continuous monitoring solutions.

Monitoring for waste site compromise is challenging because target phenomena that may indicate compromise are not known ahead of time and may manifest in multiple forms. For example, it may manifest as rapid subsidence, loss of vegetation, increased soil moisture or many other indicators. Multi-modal remote sensing is well suited to the challenge of detecting these types of phenomena because the different technologies are sensitive to many different events. The aforementioned types of remote sensing data are also georeferenced, lending their use to be easily combined with other spatial data in a Geographic Information System (GIS) to provide additional context to the results.

Both LiDAR and SAR can be used to detect ground elevation changes, while multispectral data can be used to detect other indicators of change (increased soil moisture, vegetation die-off, erosion, etc.). It is well documented that LiDAR can provide very precise (up to ±0.15 meters in this study) measurements of infrastructure and surrounding land, but as an airborne sensor LiDAR missions have a comparatively high cost. Satellite SAR data can be acquired for free at availability or by fee on-demand and provide complete coverage of the Hanford Site per acquisition, although it requires more oversight during processing to yield the best results. There are many sources of multispectral data, both free and paid, with a wide variety of spatial and temporal resolutions. An optimal early-detection system would likely rely on several modalities to create the most robust approach to inform targeted ground surveys. These remote sensing approaches also offer potential for a significant advancement and cost-savings compared to current ground-based methods. A major benefit of the remote sensing approach is that is inherently scalable and can be used to monitor entire landscapes.

3.0 Impact

Remote sensing approaches offer potential for a significant advancement and cost-savings compared to ground-based methods. Algorithmic approaches are necessary to fully exploit the volumes of data available from the increasing number of satellite sensors. Such approaches can vary in complexity, utilizing one or many platforms to create a more robust solution. A major benefit of the synoptic remote sensing approach is that while it operates at the pixel-scale, the methods are scalable to hundreds, thousands or millions of acres.

The chance discovery of the tunnel collapse at the PUREX facility highlighted the importance of early detection and continued monitoring of surficial and structural elevation. If the tunnel site had been further away from human activity areas, the collapse may have gone unnoticed for many more days, weeks or even months. By developing remote techniques capable of high-temporal wide area surveillance we can improve response to unexpected, potentially dangerous events.
The remainder of this document describes the datasets and workflow used to: 1) calculate the difference in LiDAR measured elevations 2) analyze the visual changes in multispectral satellite data and 3) apply interferometric processing to SAR data to detect subsidence.

4.0 LiDAR

This section describes the use of LiDAR to assess potential ground surface elevation changes on the Hanford Site. Relevant details on the source data and processing methods are provided, as well as a summary of LiDAR analysis results.

4.1 Data
LiDAR datasets were acquired from 2 sources; April 2008 LiDAR survey by Aerometric over Central Hanford, and 2010 Washington DNR LiDAR for Rattlesnake Area flown by Watershed Sciences. We selected the digital terrain model (DTM) for each of these surveys. The DTM is a raster representation of the earth surface with vegetation and man-made structures (buildings) removed. The 2008 Aerometric survey data had a horizontal resolution of 0.5 m with a stated vertical accuracy of 7 cm. The Rattlesnake data source resolution was ~1 m (pixel size of 3 feet) and were processed at 1 meter with a stated vertical accuracy of 7 cm. Elevations in each dataset were in North American Vertical Datum 1988 (NAVD88).

4.2 Processing
ArcGIS Spatial Analyst was used for all processing. The 2010 DNR LiDAR elevation dataset was converted to meters to standardize it to the same vertical units as the 2008 elevation data. The 2008 elevation data were subtracted from that from 2010 to determine the delta between years. Thus, positive values in the resultant map represent an increase in elevation and negative values represent a decrease. A correction factor was determined by calculating the mean delta value of the resultant map (Mean = 0.084 meters, sample size > 100 million pixels) and subtracting it from all delta values in the map. This correction factor was assumed to represent the average error between the different datasets and was applied to establish a more reliable estimate of elevation change in areas that actually experienced change. Finally, the delta raster was convolved with a 3x3 m mean filter to remove potential noise due to horizontal registration error.

4.3 Results
The results of the LiDAR change analysis contain both positive and negative pixel values. These values were mapped to categories representing 0.1, 0.2, 0.3, 0.5, 0.7, 1 and > 1 meter positive and negative change. Figure 1 shows these categories over 2015 aerial photo.

The results show that there are considerable changes between the elevations represented in the two datasets. Given that the vertical accuracies of the 2008 and 2010 LiDAR data were ±0.07 and ±0.15 meters respectively, we would expect a detection limit of ±0.22 meters. From the patterns of change apparent in figure 1, it would appear that the sensitivity of detection is in that range (0.1-0.2 meters). It should be noted that the LiDAR validation done by the providers was performed using hard flat surfaces (roadways) for survey validation points, and the expected accuracy over uneven natural surfaces is less certain. In light of this fact, the apparent sensitivity of detection is better than expected.
5.0 Multispectral Analysis of Satellite Images

5.1 Data
Multispectral satellite imaging systems are capable of high spatial resolution imaging of large areas with
daily or even higher frequency. This new paradigm of high temporal imaging allows for rapid detection of
very subtle localized events or phenomena, but approaches for doing so are still emerging as analytic
capabilities adapt to this new paradigm. We have an in-depth understanding of the reflectance of surfaces
due to changes in moisture, vegetative cover, phenophase, physiognomy and condition; all of which are
potential indicators of a compromised waste site. A physically-based change detection approach can
reasonably be expected to highlight these phenomena and utilization of machine learning neural networks
may improve results significantly.

To perform a site-wide change detection we acquired Sentinel-2 multispectral 10 m satellite data from
USGS. The scene dates were July 29 and Aug 18, 2017. These very large footprint scenes covered the
entire Hanford Site (in addition to a portion of the surrounding area). Using image differencing we
performed a broad area change detection and flagged the large changes inside the waste area polygons
(figure 2).

To look specifically for indication of the PUREX tunnel collapse we needed high-resolution scenes as
close to the collapse date as possible. A search for archived data around the time of the collapse, turned
up several scenes collected by Planet Labs. We acquired sample multispectral dataset from Planet Labs,
Inc. The scenes consisted of 4 bands (blue, green, red and near infrared) from May 4, 2017 and May 8,
2017 were used to visualize changes at the surface (figure 3).

5.2 Processing
We used ERDAS Imagine and ArcGIS Spatial Analyst for processing all multispectral images. Simple
subtraction was performed on the images, subtracting data values of the May 4th image from the May 8th
image. The overall changes were not evident at a broad scale. The changes were classified by their
magnitude and spectral direction (determining whether intensity increased or decreased) and represent
different types of changes on the ground. To rapidly characterize the wide variety of changes evident in
the difference, we performed a k-means clustering on the subtraction image. By classifying the
differences we can increase the understanding of the phenomena that may be causing the changes.

5.3 Results
Subtle changes over a large area were automatically extracted via change detection analysis. Though the
images were collected only 4 days apart; Thursday May 4th and Monday May 8th, 2017, there are a
considerable number of changes detected via the algorithm. By assessing the location, shape, and spectral
type of the changes we can intuit certain phenomena including digging activities, road treatment for dust
abatement, and vehicle positions in parking lots (figure 4). While these changes are not the signal we are
looking for, they can inform a model for identification of unexplained changes; those areas that may need
to be considered for field reconnaissance.

The collapse at the Purex tunnel (roughly 6 m by 6 m) would be difficult to see at this scale even if the
changes were obvious. Zooming into the area around the Purex tunnel illustrates the challenge of
highlighting the subtle, rare events and ignoring natural and manmade changes that are not of interest (figure 5).

To increase the ability to discern meaningful changes from noise, we stratified the analysis to areas where waste is located or suspected. Limiting our detection to areas inside waste site polygons, we reduce the variation in the dataset allowing for an effective increase in signal-to-noise. By zooming in to the area of the collapse we can see a visual indication of changes that have happened over the 4 days which are likely a result of the collapse (figure 6). The simple difference image, even constrained by the tunnel boundary shows very subtle difference in the area where the collapse is known to have occurred (figure 6). More sophisticated analyses, using multispectral principal component analysis (PCA), are able to highlight these changes (figure 7). Figure 7 illustrates the promise of statistical analysis and warrants some description. This figure is a result of constrained analysis using the polygon surrounding the Purex tunnel to select pixels from both before and after images. The selected pixels were used in a PCA; a method of condensing inter-correlated bands into new bands. By investigating each PCA band looking for evidence of the collapse it was apparent that component 6 showed evidence of increased signal at the known collapse location. Figure 7 is a color composite image of PCA band 6 with the highest values shown in a pink hue. It should be noted that principal components operations transform the original brightness bands into a complex signal that is statistically unique, but difficult to interpret physically. This method shows promise and should be investigated further before being used in a rapid alert system.

This example analysis illustrated the challenge of highlighting non-obvious signals over a large diverse area. Stratification and higher dimensional analyses can be employed to assist in the reduction of false positives and identification of relevant changes. It should be noted however that this example is overly simplified by only using images from two time periods. In an operational system, tens of images per year could be ingested into the analysis matrix. Such an approach provides an opportunity (as well as additional challenges). However by relying on high frequency acquisitions and machine learning approaches we could iteratively reduce false positives and highlight the rare phenomena for further on-the-ground or aerial investigation.

6.0 Synthetic Aperture Radar Imaging

While multispectral imaging can provide visible indication of surficial changes to land or structures, it cannot provide a direct measurement of topographic change. The efficacy of multispectral imaging is also dependent on the clarity of the scene as captured from above by the sensor, which can be obscured by clouds. Satellite-based Synthetic Aperture Radar (SAR) imaging is a powerful capability that is complimentary to multispectral imaging because it is an active sensor technology (i.e., actively transmits and receives radar signals) that can penetrate cloud cover and can provide direct measurements of elevation and topographic motion (subsidence or uplift) using a technique known as interferometry. Interferometric SAR (InSAR) techniques use two or more SAR images to either generate maps of elevation or surface deformation based on differences in the phase of the radar waves returning to the sensor. For this paper we focus on the latter due to its obvious relevance to monitoring waste sites.

InSAR techniques can potentially measure centimeter-scale changes in elevation over spans of days to years across spatial scales ranging from buildings to entire landscapes. These techniques are commonly used for geophysical monitoring of natural hazards such as earthquakes, volcanoes, and landslides, but also have been successfully used to monitor the stability of built structures and many other uses. To our knowledge, this capability has not been used to assess potential topographic motion of waste sites and subsequent risks to human or environmental health. PNNL has expertise in remote sensing applications of
SAR as well as fundamental research in radar physics and development of radar technologies and signal processing techniques.

There are many subtle variations and uses of InSAR techniques, each generally being tailored to a specific application or even location. InSAR algorithms for measuring deformation are generally very similar but can be differentiated into two application categories based on the magnitude and duration of deformation that is expected; i) detecting rapid deformation over short periods, limited to the minimal repeat cycle (6 days for Sentinel 1) ii) subtle deformation over long periods. The primary difference between these two application categories is the amount of SAR imagery required to reliably detect and accurately quantify deformation. Methods for detecting rapid deformation over short time periods focus on making interferometric comparisons between two to three SAR images taken in close time proximity to each other, whereas methods aimed at detecting subtle deformation over long time periods require a longer time series of SAR images spanning several months to multiple years. The short-term deformation monitoring method can be applied on a continual basis by repeatedly generating interferometric comparisons of new before-and-after image pairs.

The sensitivity and accuracy of InSAR for measuring deformation is dependent on a variety of factors, the most important of which are those that affect the ability to accurately co-register images used for generating the interferogram. The co-registration process is used to ensure that the same ground targets are contributing to the same pixel among the images being used. Poor co-registration is typically due to differences in the viewing angles [of a given location] of the sensor that result from slight variations in the orbital path of the satellite between acquisitions; thus, it is important to utilize imagery captured at similar orbital view angles. This consideration can sometimes present another challenge of finding suitable imagery, however, this is becoming less difficult with the addition of new satellite platforms with stable, well-defined orbits such as the European Space Agency’s (ESA) Sentinel-1A and Sentinel-1B that together provide global coverage on a 6-day repeat cycle (figure 8).

The following sections illustrate the use of InSAR derived from the Sentinel-1B platform for measuring rapid deformation for two short time periods (~2-3 week span each) over a large portion of the Hanford Site and surrounding area. One of the selected time periods overlaps the timeframe when the Purex tunnel collapsed, although we did not anticipate that this event would be detected in our analysis because the spatial resolution of InSAR products derived from Sentinel-1B (~14-m) is slightly larger than the approximate size of the collapse. The primary purpose of these examples is to illustrate InSAR deformation products for an area familiar to the reader and explain their interpretation and potential use.

6.1 Processing

The processing chain for InSAR products varies according to the software used and intended application, but will usually include the following steps. We used ESA’s free open source toolboxes available through their SNAP software architecture (http://step.esa.int/main/) for all the following steps. SAR image pairs were first co-registered to account for orbital differences among the images. An initial interferogram was then formed. Subsequent steps were aimed at removing non-target contributions to the initial phase difference, which include contributions due to the curvature of the Earth, topography, atmosphere, and noise. The purpose of removing these phase contributions is to isolate the contribution of interest, which in this case was potential phase difference due to deformation. A useful bi-product of these processing steps is an estimate of coherence – measure of similarity between the before and after image – for each pixel in the scene. The coherence product subsequently can be used to mask out areas where accuracy is likely to be poor due to low coherence.
The resulting interferogram at this point represents a two-dimensional relative phase signal that must be “unwrapped” to be useful for visual interpretation. There are many proposed techniques for performing phase unwrapping. For this study we performed phase unwrapping using the SNAPHU software package (v. 1.4.2, http://nova.stanford.edu/sar_group/snaphu/) that employs a Statistical-Cost, Network-Flow Algorithm.

The last two steps in the processing chain are to convert the units of the unwrapped phase product to actual displacement (e.g., meters, centimeters, millimeters) and apply a geometric correction to create a “birds-eye” view of the image that can be used in a GIS for interpretation and mapping. In the unit conversion step displacement is measured as a line-of-sight change in the distance between the ground target and sensor.

6.2 Results

An important step in interpretation of InSAR products such as the deformation maps demonstrated here, is to evaluate coherence of pixels within the scene to determine what areas likely to yield accurate or inaccurate deformation estimates. A common rule-of-thumb for an acceptable coherence threshold is greater than 0.3, where coherence values range from 0 (very poor) to 1 (very good). We used this threshold to mask out pixels in the final deformation maps that were coincident areas of low coherence. In our two date pair interferograms, we were able to achieve acceptable coherence levels for much of the Hanford Site and surrounding area (Figure 9). Coherence also generally improved across much of the Hanford Site in the later image pair (19-May to 31-May). Vegetated areas exhibit lower coherence whereas developed or barren areas exhibit higher coherence. Vegetation and its associated changes over time due to growth can strongly affect coherence; thus, it is suspected that coherence improved in the later image pair due to the natural reduction in growth and senescence of vegetation in the landscape.

As shown in Figure 10, the estimated deformation of nearly all pixels (with acceptable coherence values) in the Hanford Site boundary for both date pair interferograms was -.01 and 0.01 meters. To better isolate potential deformations we removed zero values from the deformation maps and inspected remaining pixel clusters at finer spatial resolutions. In general, there were relatively few locations of potential deformation detected within the portion of the Hanford Site for which our image pairs covered. Potential deformation was not detected at the Purex tunnel collapse location in the 25-April to 19-May image pair. The range of estimated displacement at the identified locations is small, ranging up to -0.072 meters in the 25-April to 19-May image pair and -0.039 meters in the 19-May to 31-May image pair. Most potential deformation locations appear in vegetated areas south of the 200 Areas (Figures 11 and 12). It is uncertain at this time whether these displacements are actual or artifact, because we lack a priori data for this demonstration. However, they are not unrealistic given the presence and mobility of sandy soils at many of these locations. Note that more significant levels of deformation appeared to be more systematic in open, non-vegetated areas, where it is known, based on conventional wisdom, that wind driven soil displacement occurs regularly in April to early May time frame (Figure 12A). Natural changes in the vegetation structure may be another contributing factor to these potential deformations. Closer inspection of the 200 Areas revealed several locations of potential deformation in a vegetated area in the northeast quadrant of the 200-West Area (Figure 12). Again, there is a high level uncertainty about areas of potential deformation identified in this demonstration because we lack a priori data about topographic changes that may have occurred during the time periods examined.
7.0 Proposed Future and Ongoing Tasks

Although the implications for SAR-based analysis at the area coincident PUREX tunnel were constrained by the resolution of Sentinel data, our results support the hypothesis that detection of physical, biophysical, and structural properties on land surface is feasible using airborne and space borne remote sensing analysis. This finding in itself is not novel within the remote sensing research community, but the applicability to the Hanford Site highlights the extent to which these methods can be used to detect subtle changes, and can be correlated with episodic type changes in land surface profile. This is a new finding, and is highly relevant to monitoring priorities at the Hanford site.

The strongest secondary line of evidence would be to corroborate these results with ground-based verification, or run a similar analysis with an equally weighted field-based component. Field based validation could be performed with either historical, or present/ongoing information. An approach using historical data for SAR-based analysis would be limited to the initial time of launch for the Sentinel 1 satellite (April 2014). This approach might consider, for example;

- Drill pads for newer wells. An area surrounding the vicinity of wells, nominally larger than 100 m² is cleared and leveled before wells are installed (Figure 13). For wells installed after April 2014 both multispectral and SAR imagery before and after well installation could be evaluated.

- Past excavation activities, after April 2014. The timing of this study was specific to the partial collapse of the PUREX tunnel, but did not necessarily consider other concurrent events involving removal of surface material, which is a typical management activity at the Hanford site. A more systematic review of past management activities in concert with collection and timing of satellite imagery could be used to ground truth “after the fact”.

Validation with ground-based data might also consider ongoing and/or planned activities, for example, continued ground based activity at ERDF. This is a site that accepts low-level radioactive, hazardous, and mixed waste. The last and largest expansion of the ERDF was completed in January 2011.

Surficial change, which can be indicative of geologic subsidence, can be monitored to a spatial resolution of 10m and 3m with the imagery used in this study, but can be measured at resolutions < 1m with other imagery sources. Geologic subsidence, as measured here in terms of deformation, or change in elevation, was observed using LiDAR and SAR imagery, and outcomes were consistent with expectations based on land features and characteristics. Whereas LiDAR can cost prohibitive, SAR data can be acquired for no-cost, although it requires significant processing to convert into maps of displacement.

Future ongoing work could be undertaken either continuously, or upon request, similar to the manner in which this work was conducted. A continuous monitoring program would be ongoing, based on an established frequency, for example, every 2 weeks. A request-based (on demand, or as required) approach would be initiated every time the analysis is requested. It is highly recommended, for several reasons, that a continuous monitoring approach be taken;

- Initial costs per analysis would drop significantly as the analytic workflow would benefit from repeat analysis. Algorithms would be streamlined and automated concurrent with technological maturation.
• Long term data acquisition and analysis would lead to “deep learning”, or “machine learning”, by which the system itself would become more adept and capable of discerning episodic type events in surface change as opposed to expected variation in surface elevation, for example mobility of sandy soils.

• Early warning system could be developed. This would include, for example automated email alerts and or text, based on established criteria.

8.0 Addendum on Machine Learning

Machine learning (ML) refers to the approach of deriving actionable insights from vast data by automatically developing algorithms. Recent ML studies have made astounding advancements to traditionally difficult problems (e.g. facial recognition, language processing, and image captioning), meeting and exceeding human performance in these tasks. Some of the earliest successful applications of ML were for searching for, finding and flagging anomalies in financial transactions, and the fraud detection case provides a very effective model for development of an early detection system of waste site anomalies. Why hasn’t ML been harnessed for detecting geographic anomalies? One key reason is data complexity. Financial transaction data may contain complex patterns (amount, vendor, location, etc.) but the data themselves are relatively simple; structured data representing time, amount of purchase, etc. Data related to geographic observables certainly encapsulates complex patterns, but also is complex in form; pixels representing amount of vegetation, lines representing roads, text describing construction schedules, etc. To integrate all of the potentially meaningful data into a rule-based model is virtually impossible because the interaction (and correlation) between the datasets is not known. Therein lies the strength of ML, the computer can learn the associations and those can be exploited in a rapid and sensitive detection model.

Why use machine learning for detection of waste site anomalies?

1. Speed – In rule-based systems, experts develop rules and thresholds to determine what is abnormal; a time-consuming, labor intensive process. Machine learning is continuously analyzing and processing new data to discover complex correlations. Furthermore, ML neural networks can autonomously update the model to ignore the natural and man-made activities causing changes across the landscape.

2. Scale – Machine learning algorithms and models become more effective with increasing data volume and richness; not so with rule-based models. Scaling rule-based models, to another site say, requires adjustment of the rules. More human time, and a game of algorithmic whack-a-mole. For a ML approach, one just simply needs to confirm any additional known events. This feedback can then be used to automatically improve subsequent detections.

3. Efficiency – It simply not possible to scan for changes in every pixel over every image. Machine algorithms employed for data analysis would only escalate decisions to humans when their input adds insights. Moreover, they can be more effective than humans at detecting subtle or non-intuitive patterns that may go unseen by even the keenest human observer.

Though experimental, there is some precedent in the usefulness of ML for geographic feature recognition. PNNL is well suited to advance the integration of ML into geographic analysis with geospatial expertise across many domains including high performance computing, data science and deep learning. A relatively modest project could be undertaken with only a few requirements; 1) geospatial data, 2) ancillary data (e.g. seismic data, demolition schedules, etc.) and 3) known historic anomalies. Such a project could be
scaled into a persistent monitoring program, where the data collected from an ongoing monitoring program could eventually be used as training data, developing a self-improving feedback loop.

A truly robust solution must allow information to come from any and all sensors, and should exploit the rapidly increasing information content from all digital data. Development, visualization, and utilization of ML neural networks has been subject of numerous projects at PNNL, and integration with geospatial imagery represents a novel application of the technology. By combining our physical understanding with PNNL’s capability in ML and high-performance computing we can develop sophisticated and automated change detection methods to improve outcomes across the waste disposal community.
Figure 1. Map of LiDAR DTM surface delta between 2008 and 2010 in the northern analysis region T31 and T34 (left panel) and southern analysis region U10 and U11 (right panel).
Figure 2. Detected change between Jul 29 and Aug 18, 2017 across the site (yellow) and in the 200 Areas (red). Insets shows a close up of changes at locations across the Hanford Site.
Figure 3. False-color images of a subset of Hanford from May 4, 2017 (left) and May 8, 2017 (right). False-color vegetation appears red and developed areas blueish. The Purex tunnel is in the far right of each image as a black polygon.

Figure 4. Classified changes over the May 8th, 2017 image. Changes in yellow represent increase in brightness, those in blue represent decrease in brightness.
Figure 5. Left panel: difference image from May 4 to May 8th, 2017 with outline of the Purex tunnel in black. Notice the 3 light blue areas indicating change. Right panel: same area with air photo background. Red circles indicate areas of most change.

Figure 6. Left panel: May 4th false color image of the northern portion of the Purex tunnel. Right panel: May 8th false color image of the northern portion of the Purex tunnel. Small star is the known location of the collapse.
Figure 7. Principal components image composite showing the collapse area
Figure 8. Illustration of the scene boundary for the SAR swath used in our analyses.
Figure 9. Green pixels indicate acceptable InSAR coherence (>0.3) for deformation analysis. Note the improved areal coverage of pixels with acceptable coherence in the later image pair.
Figure 10. Hanford Site-scale view of potential deformation represented as line-of-sight displacement measured in meters. Note that areas of potential deformation are poorly to not visible at this scale as displacement was 0 meters for most of the pixels in the scene.
Figure 11. Locations of potential deformation identified on the Hanford Site Central Plateau.
Figure 12. Locations of potential deformation identified within and around the Hanford Site 200 Areas, and other areas that show significant change in elevation.
Figure 13. Wells drilled after April 2014 (above) in the Hanford site, and drill pads in well area (below)