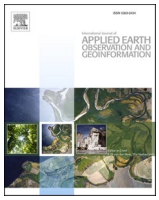
International Journal of Applied Earth Observations and Geoinformation 103 (2021) 102463

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Enhancing environmental enforcement with near real-time monitoring: Likelihood-based detection of structural expansion of intensive livestock farms☆

Ben Chugg [1], Brandon Anderson [1], Seiji Eicher, Sandy Lee, Daniel E. Ho [*]

Stanford University, United States

ARTICLE INFO

Keywords:
Structural expansion
Time series
Maximum likelihood
Animal feeding operations

ABSTRACT

Much environmental enforcement in the United States has historically relied on either self-reported data or physical, resource-intensive, infrequent inspections. Advances in remote sensing and computer vision, however, have the potential to augment compliance monitoring by detecting early warning signs of noncompliance. We demonstrate a process for rapid identification of significant structural expansion using Planet's 3 m/pixel satellite imagery products and focusing on Concentrated Animal Feeding Operations (CAFOs) in the US as a test case. Unpermitted building expansion has been a particular challenge with CAFOs, which pose significant health and environmental risks. Using new hand-labeled dataset of 145,053 images of 1,513 CAFOs, we combine state-of-the-art building segmentation with a likelihood-based change-point detection model to provide a robust signal of building expansion (AUC = 0.86). A major advantage of this approach is that it can work with higher cadence (daily to weekly), but lower resolution (3 m/pixel), satellite imagery than previously used in similar environmental settings. It is also highly generalizable and thus provides a near real-time monitoring tool to prioritize enforcement resources in other settings where unpermitted construction poses environmental risk, e.g. zoning, habitat modification, or wetland protection.

1. Introduction

The protection of land, air, and water depends critically on the enforcement of environmental laws (Gray and Shimshack, 2011). There is, however, mounting evidence of serious challenges facing environmental regulators (Evans and Malcom, 2019; GAO, 2008; Purdy, 2010). Conventionally, the mainstay of enforcement has consisted of low-frequency visits to permitted facilities (e.g., once every 2–5 years under the U.S. Clean Water Act, the federal law governing water pollution) and self-reported information, placing a heavy burden on state and federal agencies in light of growing environmental challenges and budgetary threats. This has lead to increasing interest in leveraging machine learning to augment regulatory capacity (Glicksman et al., 2017; Hino et al., 2018).

Academic interest has focused on satellite imagery given rapid progress in the field of computer vision (Gauthier et al., 2007; Weinstein, 2018) and the dramatic increase in the availability of satellite imagery (Handan-Nader et al., 2020). In contrast to aerial imagery, which can be hard to obtain and limited to a specific region, optical satellite imagery has become available for many locations at a moderate to low cost. The application of deep learning, particularly of convolutional neural networks (CNNs), to such imagery has led to recent breakthroughs in a wide range of regulatory areas including animal tracking (Laradji et al., 2020; Xue et al., 2017), industrial farm detection (Handan-Nader et al., 2020; Handan-Nader and Ho, 2019), monitoring habitat change (Evans and Malcom, 2019), tracking oil spills (Krestenitis et al., 2019; Nieto-Hidalgo et al., 2018; Bianchi et al., 2020), and deforestation mapping (Maretto et al., 2020; de Bem et al., 2020).

☆ This research was supported by Schmidt Futures, Stanford Impact Labs, the Chicago Community Trust, and Stanford's Woods Institute for the Environment. We gratefully acknowledge the Stanford Data Science for Social Good program, in particular contributions from Narasimhan Balasubramanian, Qijia Jiang, and David Kang, and exceptionally helpful comments from Elinor Benami.

* Corresponding author.
E-mail addresses: benchugg@law.stanford.edu (B. Chugg), banderson@law.stanford.edu (B. Anderson), seicher@stanford.edu (S. Eicher), sl269@stanford.edu (S. Lee), dho@law.stanford.edu (D.E. Ho).

[1] Joint Contribution.

https://doi.org/10.1016/j.jag.2021.102463
Received 27 May 2021; Received in revised form 30 June 2021; Accepted 21 July 2021
0303-2434/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Fig. 1. CAFO in Unionville, Missouri housing large volumes of livestock. Each of the four buildings in the foreground are CAFO sheds. Source: Socially Responsible Agriculture Project.

Despite these successes and the availability of imagery, a practical barrier in its broader use for regulation stems from the trade-off between temporal and spatial resolution. To date, imagery that is both available to policymakers and at high resolution rarely revisits the same location in the time required for regulatory monitoring. The United States, for instance, has regularly acquired 1 m/pixel resolution in the National Agriculture Imagery Program, but this data is only refreshed every few years. Methods that rely on such high spatial resolution imagery are thus ill-suited for augmenting regulatory capacity. Coarser resolution imagery, however, is increasingly available at weekly, if not daily, revisit times (see, e.g., Sentinel-2 Imagery). Such data could aid, for instance, in detecting zoning violations, building construction, and wetlands destruction, but requires a method for detecting changes in time series that align with the spatial and temporal resolution of what is more broadly available today.

In this work, we present a technique intended to address this important gap by statistically leveraging repeated observations of medium-resolution (3 m/pixel), but high-cadence (weekly), imagery to augment environmental enforcement.[2] The method builds on any given segmentation approach and requires only simple (and generalizable) assumptions about object permanency.

To demonstrate the technique, we focus on the task of monitoring the capacity of Concentrated Animal Feeding Operations (CAFOs), which can generate serious environmental and health consequences (see Section 2). We first train an image segmentation model to differentiate facilities from background, and then test for change points within a time series of satellite images.

Our contributions are:

1. A new time series dataset of 145,053 hand-validated segmented satellite images (GeoTiffs) for 1,513 CAFO facilities from January 1 to December 31, 2019[3];
2. A maximum likelihood model that takes as inputs a time series of segmented images and provides a test statistic that captures the likelihood that a facility expanded;
3. A demonstration that this model can detect the expansion of CAFO facilities in Indiana with high-cadence satellite imagery;
4. A comparison to show that this model provides substantial improvement relative to changepoint detection baselines.

Replication code and data are available at: https://github.com/reglab/building_expansion.

The remainder of the paper is structured as follows. Section 2 provides background on CAFOs, the substantial environmental impact of intensive livestock farming, and the significant gap in environmental monitoring. Section 3 discusses related work in detecting structural change in satellite imagery. Sections 4 and 5 provide background on the data and the details of our proposed algorithm. Section 6 presents the performance we use to analyze the results, and Section 7 provides results. Section 8 discusses implications, interprets the differences in results, and offers future directions for research.

2. Background on CAFOs

Large-scale industrial farming is increasingly responsible for livestock production in the United States (MacDonald and McBride, 2009; MacDonald et al., 2018; Hribar, 2010). CAFOs, responsible for raising large numbers of animals at high densities, represent a prominent feature of this transformation. They have been estimated to produce anywhere from 40% to 99% of all livestock in the United States (Copeland, 2010; Gurian-Sherman, 2008; Anthis, 2019), with uncertainty stemming from lack of regulatory monitoring. Fig. 1 provides an example of a large CAFO, with distinctive large sheds that confine animals. Under federal law, a large CAFO, for instance, contains 125,000 or more heads of poultry or 2,500 or more heads of hog or 1000 beef cows.[4]

CAFOs pose a variety of social and environmental risks with limited and infrequent oversight. For example, a large CAFO can produce some 1.6 M tons of waste a year – which is equivalent to a large (1 M resident) city – but is not subject to the same regulation as human wastewater (GAO, 2008). Such large levels of animal waste are thought to be serious contributors to water contamination, due to potentially poorly-constructed manure storage pits or run-off from the application of waste to fields (Burkholder et al., 2007; Gurian-Sherman, 2008), as well as air pollution (Burkholder et al., 2007; Hribar, 2010). Consequences of this contamination include the degradation of aquatic systems, harmful algal blooms, and nitrate contamination of drinking water (Copeland, 2010; Hribar, 2010). These health and environmental risks can are frequently borne by poorer and minority communities, where such facilities are disproportionately sited (Nicole, 2013; Son et al., 2021).

The regulation of CAFOs has proven challenging. As noted by the Government Accountability Office (GAO), the "EPA does not have comprehensive, accurate information on the number of permitted CAFOs nationwide. As a result, EPA does not have the information it needs to effectively regulate these CAFOs" (GAO, 2008). There is substantial evidence that many CAFOs often undergo unpermitted construction and expansion, in the process evading environmental review (Howard, 2019; Marshall, 2015; Guay, 2012; Merced County, 2012). While we draw upon building expansion in CAFOs as a focused case where evasion is acute, this type of regulatory avoidance also exists in other areas of environmental law, where limited monitoring systems and the reliance on self-reporting can impede regulatory efforts (OIG, 2005;

[2] We acknowledge that describing 3 m/pixel data as "medium" resolution is relative. We describe it as such because existing work in similar environmental settings (e.g., detecting CAFOs) has relied on imagery with at least 1 m/pixel resolution.

[3] Our underlying imagery comes from Planet Labs's PlanetScope Daily Imagery.

[4] See EPA, Regulatory Definitions of Large CAFOs, Medium CAFO, and Small CAFOs, https://www.epa.gov/sites/production/files/2015-08/documents/sector_table.pdf.

Purdy, 2010). Given the potential social and environmental impact, it is critical to develop methods for closer to real-time monitoring.

3. Related work

For general change detection techniques on lower-cadence imagery, we refer the reader to the surveys of Zhu (2017) and Namoano et al. (2019). Previous work focusing on longer time series is typically concerned with larger scale changes than the addition of buildings, e.g., urban expansion (Wan et al., 2019), changing land cover (Zhu and Woodcock, 2014), forest disturbance (Kennedy et al., 2010; Huang et al., 2010), or vegetation (Verbesselt et al., 2010; Browning et al., 2017). A common approach for longer sequences of imagery is to reduce each image to a metric — a vegetation or drought index for instance — and apply more traditional changepoint detection techniques to the resulting time series. The Normalized Difference Vegetation index (NDVI) is one popular metric, used for example by both Wan et al. (2019) to detect urban change in Landsat imagery and by the Breaks for Additive Season Trend (BFAST) algorithm (Verbesselt et al., 2010).[5] Setiawan et al. (2016) proposes a median moving window and linear interpolation to reduce noise in the NDVI time series. Ye et al. (2021) propose a state space model with a Kalman filter and take a time series approach to detect forest disturbance.

Another strand of research uses regression methods for anomaly detection. The influential approach by Zhu and Woodcock (2014) develops a Continuous Change Detection and Classification algorithm, using ordinary least squares time series to classify changes at the pixel level. Koltunov et al. (2009) describes the Dynamic Detection Model (DDM), which combines a series of basis images into a prediction image. This prediction image is then compared with the true test image for anomaly detection. The DDM and related anomaly detection approaches have inspired a wide range of applications (e.g., Tang et al., 2020; Koltunov et al., 2020). Another approach is by Fytsilis et al. (2016), which expands anomalous pixels into homogeneous regions and calculates a difference metric between two images based on the maximum spatial correlation of the shape with each image as a whole.

When sufficient data are available, deep learning based detection methods have also proven useful. Varghese et al. (2018) develop ChangeNet, a CNN to detect changes between pairs of images. Peng et al. (2019) develop an end-to-end (as opposed to image-by-image or pixel-by-pixel) change detection method using UNet++, and Sefrin et al. (2021) conjoin two deep learning architectures to detect land cover change. While promising, such deep learning approaches cannot easily be applied in our setting due to the large number of free parameters and scarce positive examples.

The most closely related work is that of Robinson et al. (2021), who proposed a semi-supervised algorithm for retroactively determining construction dates using high-resolution satellite imagery. While the approach is complementary, it presupposes a set of final structure footprints, within which it looks backwards in time for localized spectral shifts that represent construction events. Having these footprints, however, would imply that the expanded or new structures were already detected at a reasonable significance. Instead, the premise of our work here is that this is a challenging task at lower resolutions. Instead of focusing on the detection of individual buildings, our method tests the broader hypothesis that the distribution of pixels classified as a structure undergoes a discrete change over time. This allows us to draw more reliable inferences using a statistical framework for building expansion in high-cadence, lower-resolution imagery.

In sum, while much of the above work has demonstrated that detecting changes in local building footprints is possible, we are unaware of work that has attempted to provide near real-time (weekly), actionable insights for regulatory enforcement. Below, we compare existing approaches, showing substantial gains to our likelihood approach tuned to this setting.

[5] NDVI gives a pixel-wise value (between −1 and 1) corresponding to the amount of greenery in the pixel, and can be reduced to a single value for the image by averaging.

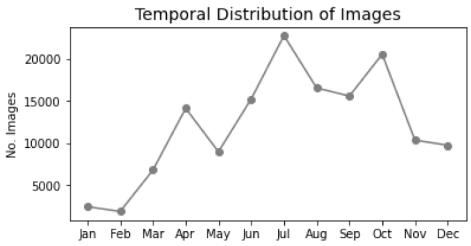


Fig. 2. Total number of images in each month across all locations.

4. Data

4.1. Imagery

Our dataset consists of 1,516 ground truth Indiana CAFO coordinates validated by the Environmental Working Group (EWG). At each location, we queried the Planet Labs API for 4-band (red, green, blue, and near-infrared) imagery between January 1, 2019 and December 31, 2019, requiring that each be at least 95% clear (e.g., of cloud cover and cloud shadows). This yielded roughly 80 to 130 images per location, with 175,736 images total. Due to the cloud cover filter, images were less frequent in January, February, and December, and most frequent in the summer (see Fig. 2). We clipped each resulting scene to a centered 600 m × 600 m (200 × 200 pixel) area, and discarded those with over 15% missing pixel values (missing pixels occur if the image was on the edge of the area being considered). Overall, post-processing removed just over 30,000 images, leaving 145,053 corresponding to 1,513 final locations.

Fig. 3 illustrates both the resolution of the imagery and gives two examples of CAFO expansion. We note that these images are of lower spatial resolution (3 m/pixel) than previously used for CAFOs and other similar settings, which poses the main inferential challenge for real-time monitoring. To the human eye, the construction of barns is visible (i.e., a third rectangular roof to the south of facilities in the left panel and a second rectangular roof to the east of the single barn in the right panel). The images, however, also illustrate some of the complications for computer-based object segmentation, such as the cloud cover in panel (2) and the snow in panel (3).

4.2. Building footprints

Many applications may be able to take advantage of pre-existing segmentation models. Conventional building segmentation models, however, have focused in large part on urban areas, which may result in performance degradation in rural areas. To provide building footprints that are tailored to CAFOs — and not other types of structures — we train our own segmentation model (see Section 5.1). This approach required ground truth building footprints corresponding to known CAFO locations. We procured such information from (1) the Microsoft Building Footprint dataset Microsoft (2018), and (2) a United States Geological Survey of poultry barns in the DelMarVa (Delaware - Maryland - Virginia) peninsula (Soroka and Duren, 2020). Because the Microsoft dataset includes all building types, we removed those footprints with area less than 450 m^2 that are below the threshold of a reasonable size for large CAFO operations. When combined with the proximity to a known CAFO location, this filter left us with a fairly pure training set. The most significant remaining source of noise stemmed from date mismatch between the Microsoft and EWG labels. For instance, a CAFO may have been constructed and tagged *after* the building footprints were

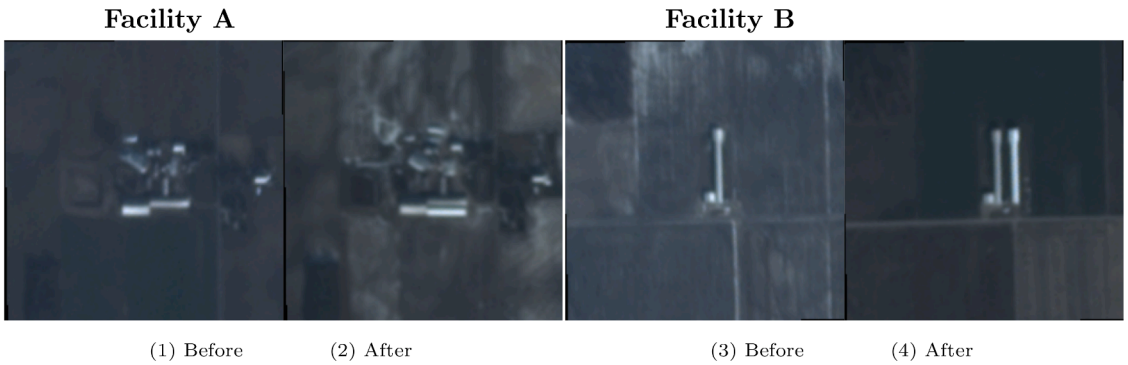


Fig. 3. Two examples of expanded CAFO facilities. Both examples exhibit environmental factors which can complicate the segmentation—cloud cover in panel (2) and snow in the panel (3).

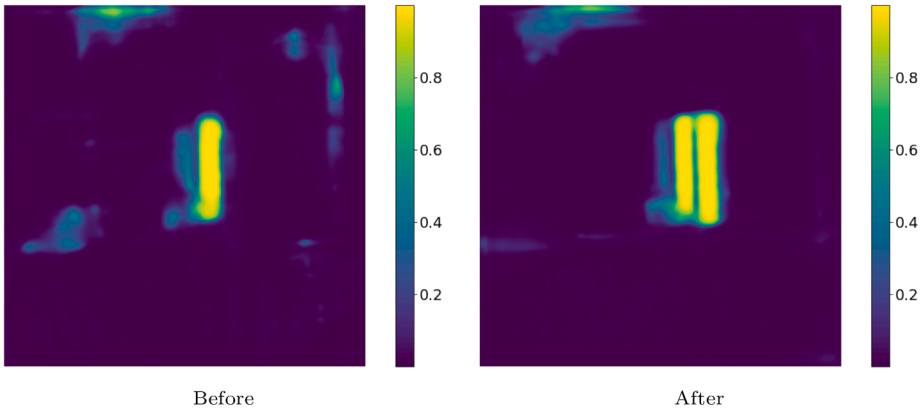


Fig. 4. Pixel-level class probabilities determined by the segmentation model of the example in the right of Fig. 3. Higher values indicate a higher likelihood of being part of a CAFO.

generated. Similarly, the DelMarVa data was collected in 2016–2017 – these barns may have been demolished or moved between the time of data collection and the acquisition dates of our Planet imagery datasets (2019). As far as we are aware, such changes are rare, but they would reduce the accuracy of our image segmentation and hence attenuate results of subsequent changepoint detection.

4.3. Expansion labels

Each image time series corresponding to our CAFO locations was hand-labeled by our research team to indicate whether a facility expanded during the observation window (1/2019–12/2019). This was done by comparing the first and last available images in the time series, with the review of intermediate images if any structural changes were observed. In total, we found 22 sites that had constructed at least one new CAFO building, 1,414 with no evidence of new construction, and 77 with either a no-longer-existent CAFO or indeterminate expansion. It is worth noting that despite the strong class imbalance, the sheer number of farm operations in the US — just shy of one million according to the National Agricultural Census (USDA, 2017), excluding aquaculture — implies that there are potentially on the order of 10,000 such expansion events every year.

5. Methods

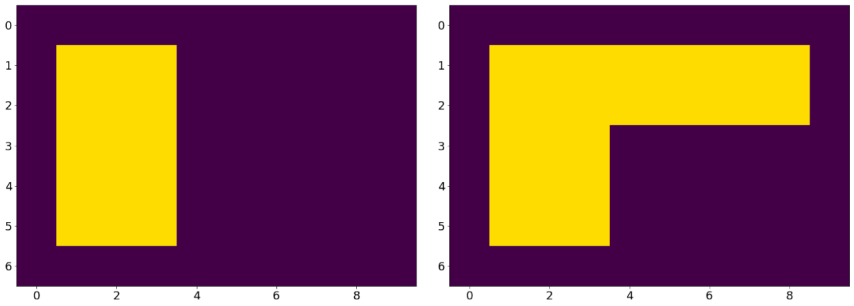
Our procedure for detecting structural expansion can be broken down into three steps. In the first, we apply a segmentation model (Ronneberger et al., 2015a) to each image in order to obtain pixel-level probabilities of class membership, with higher values indicating that the pixel is more likely to belong to a CAFO shed. We emphasize that our approach is independent from the choice of segmentation model as long as it yields such probabilities.[6] In the second step, we fit a time-dependent building footprint model using maximum likelihood estimation (MLE) to best match the class probabilities. Finally, we compare this time-dependent model to a restricted model fit under the hypothesis of no building expansion. A test statistic is then computed comparing the two fits. Intuitively, the more dissimilar the two fits, the more likely there was expansion.

5.1. Segmentation

In order to take advantage of both the spectral and spatial signatures of a CAFO structure, we employ a U-Net architecture (Ronneberger et al., 2015b). We trained the model on about 2,611 labeled images, composed of 1,176 EWG location images with post-processed Microsoft building footprint labels and 1,435 images of DelMarVa peninsula poultry houses with labels from their shapefile. DelMarVa data, in particular, appeared to boost the performance of the segmentation model. We trained on one image per location, with an even split throughout each month of the year to regularize the model against seasonality.

The model was trained with a batch size of 8, using the Adam optimizer with learning rate 5e-4, and weight decay 1e-7. The training data was split into 70% train, 15% validation, and 15% test sets. We

[6] A notable alternative for an image segmentation approach would be DeepLabv3+ (Chen et al., 2018), which has shown small performance gains of roughly 1% in accuracy relative to U-Net (Ahmed et al., 2020; Cai et al., 2021; Jiwani et al., 2021). In A, we conduct an ablation study showing that performance of MLE is comparable across reasonable ranges of segmentation accuracy.



$$\mathbf{F}^0 = \begin{bmatrix} 0&0&0&0&0&0&0&0&0&0\\ 0&1&1&1&0&0&0&0&0&0\\ 0&1&1&1&0&0&0&0&0&0\\ 0&1&1&1&0&0&0&0&0&0\\ 0&1&1&1&0&0&0&0&0&0\\ 0&1&1&1&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0 \end{bmatrix} \quad \mathbf{F}^+ = \begin{bmatrix} 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&1&1&1&1&1&0\\ 0&0&0&0&1&1&1&1&1&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0 \end{bmatrix}$$

Fig. 5. Idealized CAFO footprints before and after expansion (top left and right, respectively). Yellow (light) squares represent the CAFO shed. The entries with a 1 in $\mathbf{F}^0$ (rows 1–5, columns 1–3 assuming a zero-index) represent yellow pixels in the footprint prior to expansion. The entries with 1 in $\mathbf{F}^+$ represent the pixels of the added building (rows 1–2, columns 4–8).

performed random flips and rotations on the training set for data augmentation. The loss function converged after 20 epochs. Our loss function was binary cross-entropy with logits, and we weighted CAFO to background loss 30:1 to compensate for the class imbalance. The final results on the validation set had a recall of 0.864, a precision of 0.482, and an Intersection over Union (IoU) of 0.448.

Our change detection algorithm (explained below) does not take as input the final segmentation, but rather pixel-level class probabilities for each image (CAFO vs. no-CAFO). Instead of passing along the probabilities directly from the U-Net, however,[7] a final post-processing step was applied: A $k \times k$ smoothing kernel ($k = 3$) was applied to each image, whereby the probability of each pixel was average over the k^2 nearest pixels. Smoothing was applied in order to remove small artifacts generated by the segmentation model. For each location, the corresponding series of probability maps was then converted into a tensor in order to apply MLE as described next. Fig. 4 illustrates the probabilities generated by the segmentation model.

5.2. Footprint modeling and change detection

Once the imagery has been segmented, we can test each CAFO location for significant changes in building footprint over a series of time-sequenced snapshots. For clarity of demonstration in our test case we make the following simplifying assumptions:

1. CAFO sheds are constructed and not removed.
2. The time required to construct a CAFO shed is short compared to the length of the period considered.
3. Only one expansion event can take place during the period considered.

These assumptions may be relaxed, affecting the parameterization of the likelihood approach. Formally, let $\{\mathbf{X}^1, \ldots, \mathbf{X}^T\}$ with $\mathbf{X}^t \in \mathbb{Z}_{\geqslant 0}^{w \times h \times 3}$ be the series of (3-band) images for some fixed location at each timestep $t \in \{1, \ldots, T\}$. Here, $w = 200$ is the width of the image and $h = 200$ its height in pixels. For each image $\mathbf{X}^t$ the U-Net produces a matrix of class-membership probabilities $\mathbf{P}^t \in [0,1]^{w \times h}$, where $p_{ij}^t = \mathbf{P}_{ij}^t$ is the probability that pixel (i,j) belongs to a CAFO shed.

Let $\mathbf{F}^0 \in \{0,1\}^{w \times h}$ be the *footprint* of the CAFO shed prior to any expansion. That is, $\mathbf{F}_{i,j}^0 = 1$ if pixel (i,j) belongs to a CAFO shed, and 0 otherwise. Let $\mathbf{F}^+ \in \{0,1\}^{w \times h}$ define the pixels which belong to a CAFO shed *after* expansion, but did not before. Thus, $\mathbf{F}^0 + \mathbf{F}^+$ defines which pixels belong to the expanded shed. If there was no expansion then $\mathbf{F}^+$ is the all-zeroes matrix. Fig. 5 illustrates how $\mathbf{F}^0$ and $\mathbf{F}^+$ correspond to an expanded shed.

Using $\{\mathbf{P}^1, \ldots, \mathbf{P}^T\}$, the goal of the expansion model is to determine $\mathbf{F}^0, \mathbf{F}^+$ and at what timestep (if any) the transition occurs. To this end, we define a function which, given the building footprints and transition time t^*, captures the transition:

$$Z_{\mathbf{F}^0,\mathbf{F}^+,t^*,\alpha}(t) = \mathbf{F}^0\left(1 - S_\alpha\left(t - t^*\right)\right) + \mathbf{F}^+ S_\alpha\left(t - t^*\right), \tag{1}$$

where $S_\alpha(x) = \alpha/(1 + e^{-x})$ is the sigmoid function and t^* is the timestep at which the CAFO shed expands. Here, α controls the transition speed: bigger values of α imply a longer building period. For notational convenience, $Z_{\mathbf{F}^0,\mathbf{F}^+,t^*,\alpha}(t)$ will be written as $Z(t)$ or simply Z^t. Note that for $t < t^*, Z(t) \approx \mathbf{F}^0$ (the original CAFO shed), and for $t > t^*, Z(t) \approx \mathbf{F}^0 + \mathbf{F}^+$ (the expanded shed).

In what follows, we use maximum likelihood estimation (MLE) to estimate the parameters of Z which are best described by the class probabilities $\mathbf{P}^1, \ldots, \mathbf{P}^T$. This is performed twice: once with no restrictions on the parameters, and once with the restriction $\mathbf{F}^0 = \mathbf{0}$ (enforcing no expansion).

We judge the fitness of a given set of parameters by forming a likelihood comparing Z to the pixel-level class probabilities yielded by the segmentation model. In other words, for each image frame the model will predict a class membership vector for each pixel. We take the corresponding class probability vector from the segmentation model p_{ij}^t and form a likelihood with their product. Functionally, this takes the form:

$$\mathscr{L}_t = \prod_{i,j} p_{ij}^t \cdot Z_{ij}^t. \tag{2}$$

Taking the product over all frames in the time sequence to get the likelihood $L = \prod_t L_t$, the log-likelihood becomes:

$$\log \mathscr{L} = \sum_{i,j,t} \log\left(p_{ij}^t \cdot Z_{ij}^t\right). \tag{3}$$

Observing that Z^t is differentiable due to the sigmoid function, we adjust its parameters to maximize Eq. (3). For the unrestricted model, we solve:

$$\widetilde{\mathbf{F}^0}, \widetilde{\mathbf{F}^+}, \widetilde{t^*} = \arg\max_{\mathbf{F}^0,\mathbf{F}^+,t^*} \log \mathscr{L}\left(\mathbf{F}^0, \mathbf{F}^+, t^* \middle| \mathbf{P}^1, \ldots, \mathbf{P}^T\right).$$

[7] Our U-Net implementation did not provide class probabilities directly, but rather class *confidences* in the range [-7,7]. We applied the sigmoid function to these values in order to obtain class probabilities.

For the static model, we enforce no expansion and solve:

$$\overline{\mathbf{F}^0}, \bar{t}^* = \arg\max_{\mathbf{F}^0, t^*} \log \mathscr{L}\left(\mathbf{F}^0, \mathbf{0}, t^* \middle| \mathbf{P}^1, \ldots, \mathbf{P}^T\right).$$

From here, we form a test statistic—the delta log-likelihood—comparing the two fits:

$$TS = \log \mathscr{L}\left(\widetilde{\mathbf{F}^0}, \widetilde{\mathbf{F}^+}, \tilde{t}^* \middle| \mathbf{P}^1, \ldots, \mathbf{P}^T\right) - \log \mathscr{L}_0\left(\overline{\mathbf{F}^0}, \mathbf{F}^+ = \mathbf{0}, \bar{t}^* \middle| \mathbf{P}^1, \ldots, \mathbf{P}^T\right). \quad (4)$$

This statistic enables us to assess whether there is a difference in distributions between the images with and without expansion. Intuitively, we expect that the locations exhibiting expansion will larger test statistic (in absolute value) than those that do not, because those without expansion should be well modeled with $\mathbf{F}^+ = \mathbf{0}$. In practice, plotting the values of the test statistic for facilities known not to have expanded gives us a null distribution against we can compare the values of other facilities. The further a facility deviates from this null distribution, the more likely it is (according to the model) to have expanded.

5.3. Baselines

We compare our method against several baseline methods. Two use Bayesian changepoint detection (BCP), a recommended method when analyzing univariate timeseries (van den Burg and Williams, 2020). We apply BCP to both the sequence of NDVI values obtained from the images over time (see Section 3), and to the number of pixels identified as belonging to CAFOs by the segmentation model in each image. We refer to these as BCP-NDVI and BCP-PC, respectively. Another baseline is the Breaks For Additive Seasonal and Trend (BFAST) algorithm by Verbesselt et al. (2010). As BFAST was originally designed to examine vegetation response, we apply it to the timeseries of NDVI values obtained from the images.

Our final baseline is based on the Dynamic Detection Model (DDM). DDM is a localized anomaly detection method, contrasting each pixel/band value at time t with a predicted value generated by an optimized combination of m prior observations. Specifically, let Γ^t_{ijk} be a pixel value within a spectral band k of image $\mathbf{X}^t$. A linear DDM fits coefficients $\gamma(t) \equiv \{\gamma_0(t), \ldots, \gamma_m(t)\}$ to form an estimator

$$\widehat{\Gamma}^t_{ijk} = \gamma_0\left(t\right) + \sum_{\tau=1}^{m} \gamma_\tau\left(t\right)\Gamma^{t-\tau}_{ijk} \quad (5)$$

that minimizes the RMSE $\sigma_k(t)$ relative to the true value Γ^t_{ijk}. Note the explicit dependence on t, which indicates that each set of γ is tied to an absolute time and can therefore be used to describe dynamics such as seasonality.[8] The localized anomaly score Z^t_{ijk} is then generated by

$$Z^t_{ijk} = \frac{\Gamma^t_{ijk} - \widehat{\Gamma}^t_{ijk}}{\sigma_k(t)}. \quad (6)$$

To provide an image-level anomaly score Z^t comparable to our other methods, we simply sum Z^t_{ijk} over pixels and the three visible bands $k \in \{R, G, B\}$, leveraging the fact that true expansion events should involve multiple simultaneous pixel changes. From here, we employ two techniques to determine the probability of an expansion event taking place during a given period at this location. The first is to check if any score Z^t is above a predefined threshold. We call this DDM-MAX VALUE, or DDM-MV. The second is to apply BCP to the time-series $\{Z^t\}$, and similarly classify the series as exhibiting an expansion event whenever BCP returns a probability over a given threshold. We call this DDM-BCP. When reporting the accuracies for both methods, we scan over the set of all thresholds and report the maximum value.

It is worth noting that none of the baseline methods were tailored specifically for the purpose of detecting discrete image-level changes in an ongoing series. Limitations of these approaches reflect the lack of research in this setting, not a general drawback to the methods. BFAST, for instance, was designed for timeseries over recurring seasons and large scale changes. And while sensitive to the initial appearance, DDM does not inherently make use of subsequent observations that reinforce the hypothesis of a new structure.

6. Performance metrics

We evaluate our proposed method in addition to three baseline methods. The evaluation uses a combination of metrics, several traditional in machine learning, and several designed to capture resource efficiency. All four methods assign a "confidence" of expansion to each location: a posterior probability for BCP, a magnitude for BFAST, and the test statistic (Eq. 4) for MLE. This is important from a resource allocation perspective, as we would like to know not only which facilities the model think expanded, but which are most likely to have done so. It allows enables us to examine the performance of the various methods as a function of their confidence.

Balanced Accuracy & F1-score. Accuracy and F1-score are traditional performance metrics in machine learning. Due to the severe class imbalance, we report balanced accuracy as opposed to overall accuracy. Balanced accuracy is the mean of sensitivity and specificity, and the F1-score is the harmonic mean of precision and sensitivity. Accuracy and F1-score can be calculated using any given confidence as a threshold: Locations with an assigned confidence above the threshold are predicted as expansions. For each algorithm, we report the maximum balanced accuracy and F1-score achieved over all confidence values.

ROC & AUC. Receiver Operating Characteristic (ROC) Curves and the corresponding area underneath the curve (AUC) are another common metric in machine learning. The ROC Curve plots the false positive rate against the true positive rate as the threshold (in this case, the confidence) of the classifier is varied. The larger the area under the curve the better the overall performance, with an area of one being optimal. In addition to the area underneath the ROC curve, we also report the area under the precision-recall curve (PR-AUC).

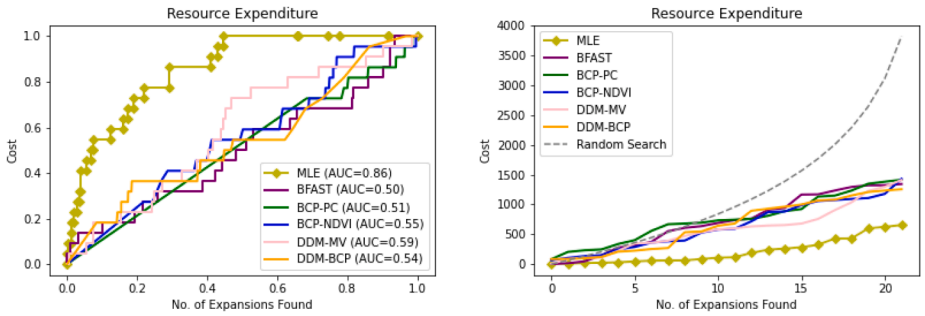
Confidence-Size Correlation. Ideally, higher model confidence in an expansion is correlated with a larger expansion. If so, sorting by confidence would allow enforcement agencies to focus on the largest expansions first. Moreover, a high correlation suggests that the model is picking up on the features indicative of an expansion. We thus examine the Pearson correlation between the size of the expansion and each model's confidence.

Resource Expenditure. Finally, we examine to what extent the four methods save costs compared to a random search — a reasonable baseline given extensive reliance on random field visits — for building expansions. Suppose that examining a facility which did not expand has a cost of 1, while correctly identifying an expanded facility costs 0. A random search is inefficient due to the severe class imbalance. Let T_n be the number of trials until n expanded facilities are found. Then $T_n =$

Table 1
Balanced Accuracy, macro F1-scores, and Pearson correlation between confidence and size of expansion. Proposed method is in bold. In the correlation column, a star (*) indicates a p-value of less than 0.001.

Model	*Accuracy*	*F1-score*	*Size Correlation*
BCP-NDVI	56.9%	0.45	0.21
BCP-PC	52.1%	0.25	−0.20
BFAST	55.2%	0.55	0.00
DDM-MV	62.9%	0.53	−0.02
DDM-BCP	58.9%	0.50	0.17
MLE	78.6%	0.60	0.67*

[8] To account for finite cadence and gaps in observation, we fit γ on aggregated imagery at a 1-week resolution.



(a) (b)

Fig. 6. (a) The ROC Curve using confidence values as the prediction threshold. Locations above a given value of the test statistic are labelled as expansion. (b) Reduction in cost achieved by sorting images based on their test statistic compared to random search. The *y* axis is the cost of finding *n* expansion events. Examining an image not exhibiting expansion is a cost of 1.

$\sum_{i=1}^{n} T_{i-1,i}$ where $T_{i-1,i}$ is the time taken to find the i-th expansion after the $(i-1)$-st expansion has been found. For random search, $T_{i-1,i}$ is a geometric random variable with success probability $\frac{M-i}{N-i}$ where M is total number of expansions, and N the total number of images. Hence,

$$\mathbb{E}\left[T_n\right] = \sum_{i=1}^{n} \mathbb{E}\left[T_{i-1,i}\right] = \sum_{i=1}^{n} \frac{N-i}{M-i}.$$

For us, $N = 1,414$ and $M = 22$. We examine how much the four models improve over this baseline cost.

7. Results

We first report on the distribution of confidence values. For MLE, expansion events have a substantially higher average test statistic than non-expansions — 23,205 versus 5,427, a difference of over 300% (p-value < 0.001 using a t-test). Meanwhile, none of the other three methods had statistically significant differences between the likelihoods of expansions versus no-expansions.

Next we move onto the performance metrics. Table 1 gives the maximum balanced accuracy, F1-score, and likelihood-size correlation of each method. MLE achieves a balanced accuracy of 78.6%. BCP on NDVI is the runner-up at 56.9%. Due to the strong class imbalance, macro F1-scores are relatively low for all models, but at 0.60, MLE again represents a significant improvement over the other methods. MLE achieves its maximum accuracy using a threshold of $t = 5{,}308$, at which point it correctly classifies 86.4% of the expansions, and 70.1% of the non-expansions. However, many thresholds give significantly worse ratios, resulting in a low area under the PR Curve (0.1). There is hence still significant room for improvement. We emphasize though that the goal of this approach is to help *prioritize* and augment enforcement resources, which is why the cost metric is particularly salient. With respect to the size correlation, at 0.67 (p-value < 0.001) MLE is the only method which exhibits significant correlation between the likelihood and the size of the expansion. All other methods have a correlation below 0.22, none of which are statistically significant.

Fig. 6a illustrates the ROC Curves for all methods. MLE has the highest AUC at 0.86, followed by DDM-MV at 0.59, BCP-NDVI at 0.55, DDM-BCP at 0.54, BCP-PC at 0.51 and BFAST at 0.50. Regarding cost reduction, Fig. 6b demonstrates that all methods improve over random search. For each method, we rank the locations by confidence of expansion and examine the images in that order. We assume that examining a facility which did not expand has a cost of 1, and examining an expansion has a cost of 0. A random search involves examining 3,836 false positives (facilities which did not expand) in expectation in order to find all 22 expansions. MLE examines 654 false positives, a decrease of over 80%. The three baselines perform approximately equally, decreasing the cost by between 62% and 64%. Overall, sorting by test statistic is a substantial improvement over random search. While MLE outperforms the other three methods, the savings exhibited by any of the four algorithmic approaches over an ad-hoc scan demonstrates the potential of such automated methods.

8. Discussion

In this work, we have demonstrated the possibility of enhancing traditional enforcement techniques with dynamic monitoring of satellite imagery. Most promising is that the methods developed herein demonstrate the possibility of near real-time environmental monitoring, using 3 m/pixel resolution imagery—available at high cadence and increasingly low-cost.

Our MLE approach appears to outperform the main available baselines (BCP, BFAST, and DDM) by a substantial margin. One way to think about the difference between our approach and these baselines is that our algorithm is specifically tailored to building expansion. First, DDM (and similar algorithms, such as CCDM) can detect anomalies with a single image. Our MLE approach, in contrast, leverages the full time series before and after a posited changepoint. Second, our algorithm works particularly effectively in conjunction with available building segmentation techniques. Other approaches using NDVI and DDM use distinct methods (e.g., denoising, image basis) to reduce dimensionality, but, as a result, do not perform changepoint detection in a way that is as tailored to building changes.

Because it is a lightweight approach that can be used in conjunction with segmented images, another strength is that it is not demanding computationally. On NVIDIA's Tesla v100 GPU, our method required approximately 30 s per location (after segmentation). Running sequentially thus yields a time of approximately twelve hours for 1,500 CAFOs. This could, of course, be sped up by running multiple GPUs in parallel.

These results are important substantively, as the gains of employing this approach would be a major improvement over the status quo of random, physical inspections. As illustrated by Fig. 6b, ordering locations by their test statistic and proceeding down the list would constitute a large gain over random inspections. The U.S. Department of Agriculture estimates that there are approximately 1 million farm operations and if our base rate of expanded facilities holds, then approximately 10,000 of these facilities may be expanding each year.[9] Augmenting conventional inspections with remote sensing can help address limitations in the resourcing of environmental protection agencies (Gray and Shimshack, 2011).

[9] This is based on a simple estimate multiplying the rate of expansion observed in our dataset ($\frac{22}{1436}$) with the number of facilities. We recognize that this simple estimate does not account for many other factors that are associated with expansion.

To illustrate, we manually searched in local records (e.g., county assessor's office, permits offices) for a sample of expanded facilities. While these records are highly decentralized and hence not easily searchable for many localities, we quickly found several facilities that were classified as "vacant land," both before and after significant expansion events. The comparison is stark: comparable permitted facilities can pay over 50 times the property tax. With access to permit records, our method could hence easily scale the approach pioneered by Massachusetts, which manually compared satellite imagery against permit records to identify illegal wetland modification (Clayton, 2004).

We close with several thoughts on directions for future research. First, the likelihood-based expansion technique relies heavily on the quality of the segmentation model. Developing better CAFO-specific segmentation models would thus help improve the results. Second, with a larger labeled set of images, it would be worthwhile comparing the performance of end-to-end deep learning approaches to changepoint detection. Third, it would be worth investigating generalizations of the likelihood approach. Eq. (1) can be modified and researchers may wish to incorporate other information about CAFOs, such as proximity of the added building, conformity with direction of existing sheds, and closures of facilities. Fourth, it could prove beneficial to combine the Dynamic Detection Model (DDM) with the segmenter, as opposed to running it on the image bands. Doing so might reduce the number of false positives arising from abrupt changes in the surrounding landscape (see Appendix B).

Last, as emphasized in the introduction, nothing about this approach is specific to CAFOs besides the initial segmentation model. By adapting this model, the likelihood based approach could be used to detect any structural changes occurring over time. It would be fruitful to apply this method to the range of other environmental challenges, such as zoning violations, habitat modification, and deforestation.

We hope to have provided a useful approach that leverages rapid increases in the availability of satellite imagery to enable remote sensing to provide actionable insights for real-time environmental enforcement. Doing so could dramatically improve the allocation of scarce regulatory resources to where most needed.

CRediT authorship contribution statement

Ben Chugg: Conceptualization; Formal analysis; Methodology; Software; Validation; Visualization; Writing – original draft; Writing – review & editing. **Brandon Anderson:** Conceptualization; Data curation; Formal analysis; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Writing – original draft; Writing – review & editing. **Seiji Eicher:** Conceptualization; Data curation; Methodology; Software; Writing – original draft. **Sandy Lee:** Conceptualization; Data curation; Methodology; Software; Writing – original draft. **Daniel E. Ho:** Conceptualization; Data curation; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Visualization; Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Sensitivity to segmentation

One of the contributions of our approach is a lightweight changepoint detection method that does not rely on the specific implementation of a segmentation algorithm. We believe this to be a strength of this contribution, as researchers can choose any preferred segmentation approach. We nonetheless examine the effect of segmentation accuracy on the results of the MLE approach.

We artificially degrade the quality (pixel accuracy) of the segmenter by 1%, 2% and 5%, and study effects on the MLE approach based on such degraded segmentation. These segmentation differences are representative of differences between state-of-the-art approaches (Ahmed et al., 2020; Cai et al., 2021; Jiwani et al., 2021).

Our findings from this ablation study are twofold. First, as expected, reduction in segmentation accuracy is associated with performance degradation in the MLE approach. Second, this performance degradation is relatively small. For instance, MLE accuracy with the U-Net segmented images yields 78.6% accuracy, which drops to 78.3% with 1% noise, 78.2% with 2% and 78.1% with 5%. We conclude that the MLE approach is likely to remain valuable for change detection of buildings across a range of current segmentation approaches.

Appendix B. Illustration of MLE and DDM

In this Appendix, we illustrate how MLE and DDM perform in our setting. Fig. B.7 displays a setting in which both methods perform well at detecting an added barn. The top panel plots raw images before and after an expansion event. The middle panels plot the pre- and post-expansion models (i.e., $\mathbf{F}^0$ and $\mathbf{F}^+$) generated by our MLE approach, showing that the MLE approach focuses on the added building.

The bottom panel illustrates the DDM model: the left image plots the predicted band immediately prior to an expansion event, based on a weighted sum of the previous 10 weekly images; the middle panel plots the actual band, including the expansion event; and the right panel plots the pixel-wise z-score between the true and predicted images. This DDM result also shows that the approach reasonably focuses on expanded barn location, confirming that this is a reasonable implementation of the DDM baseline.

Fig. B.8 illustrates a non-expansion event in which the MLE model is relatively stable, but the DDM approach yields a false positive due to abrupt changes in the surrounding landscape. This shows a drawback of DDM: drastic changes of any kind can be interpreted to be examples of building expansion. As noted in the Discussion, this kind of anomaly detection may underperform relative to the MLE approach because (a) it does not take advantage of the full time series before and after an expansion event and/or (b) does not focus the change point detection on buildings using segmentation. A potential direction for future work may be to combine DDM with the segmentation approach.

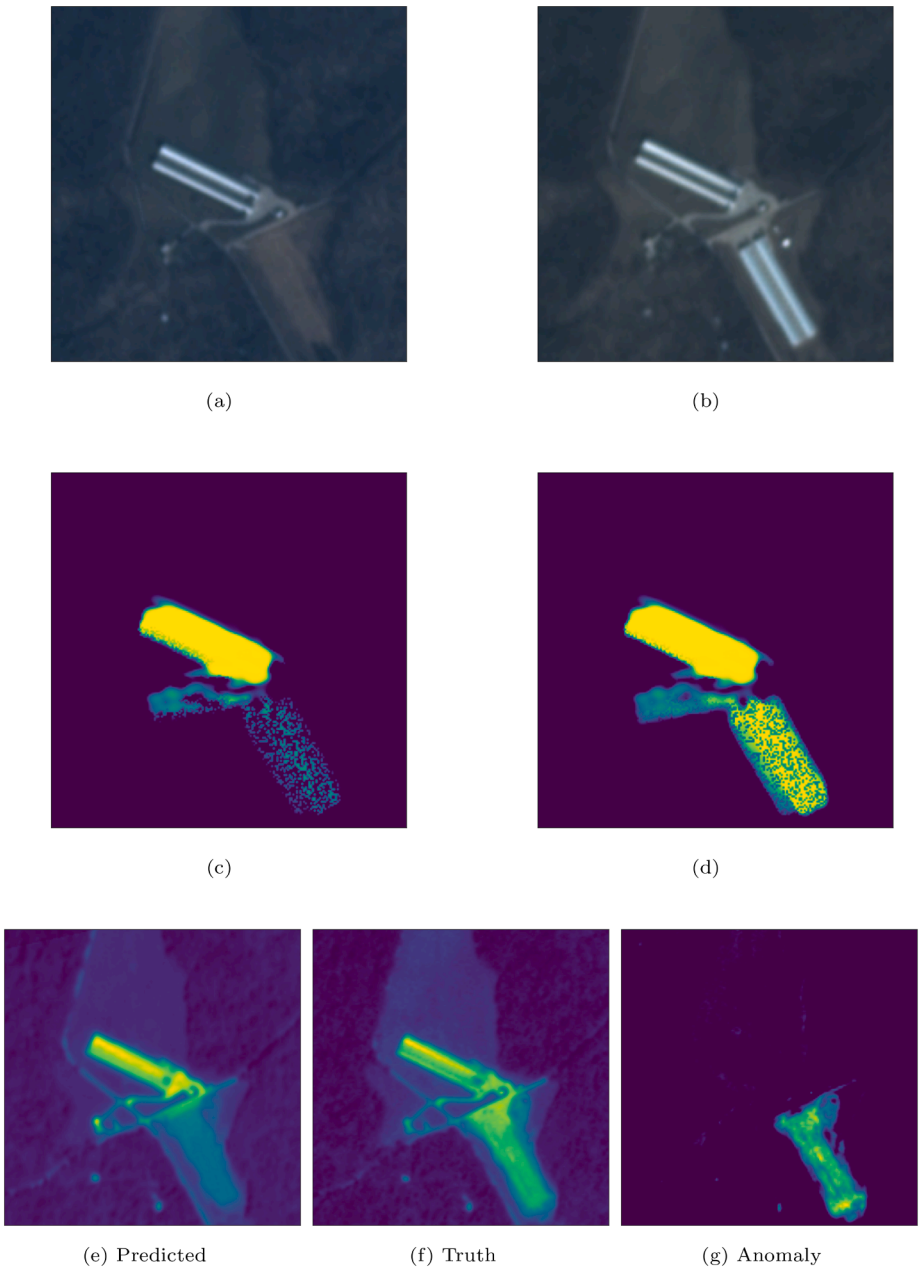
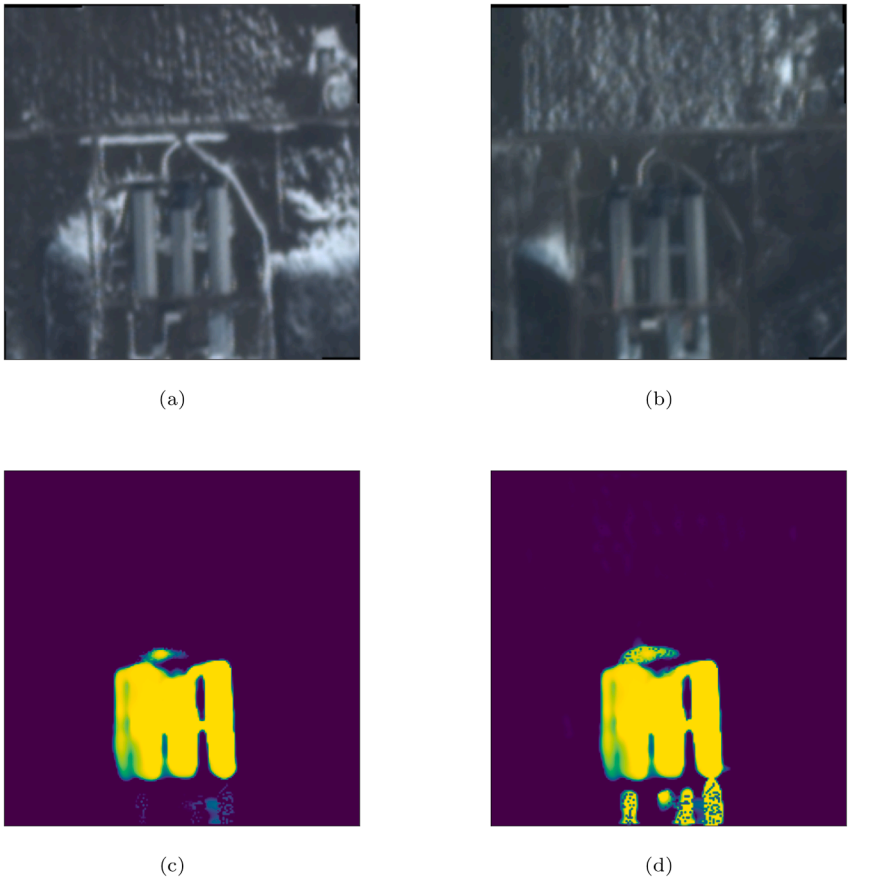


Fig. B.7. (a) and (b) Images of an expanded CAFO at the beginning and end of the year. The expansion occurred during week 28. (c) and (d) The pre and post expansion model generated by MLE. (e), (f) and (g) The predicted green (G) image band at week 28 (based on a weighted average of the previous ten weekly images), the true band at week 28, and the resulting anomaly score $Z_{ij,G}^{28}$ between the two.



(a) (b) (c) (d) (e) Predicted (f) Truth (g) Anomaly

Fig. B.8. (a) and (b) Images of a CAFO at the beginning and end of the year. Note that there was no expansion. (c) and (d) The pre and post expansion model generated by MLE. Because there was no expansion, the pre and post models are very similar, the differences being due to segmentation noise. (e), (f) and (g) The predicted green (G) image band at week 40 (based on a weighted average of the previous ten weekly images), the true band at week 40, and the resulting anomaly score $Z^{40}_{ij,G}$ between the two. Due to abrupt changes in the surrounding fields, there are many pixels with significant anomaly scores, thus resulting in a false positive for DDM.

References

Ahmed, I., Ahmad, M., Khan, F.A., Asif, M., 2020. Comparison of deep-learning-based segmentation models: Using top view person images. IEEE Access 8, 136361–136373.

Anthis, J.R., 2019. US Factory Farming EstimatesPublisher: Sentience Institute. URL https://sentienceinstitute.org/us-factory-farming-estimates.

Bianchi, F.M., Espeseth, M.M., Borch, N., 2020. Large-scale detection and categorization of oil spills from sar images with deep learning. Remote Sens. 12 (14), 2260.

Browning, D.M., Maynard, J.J., Karl, J.W., Peters, D.C., 2017. Breaks in modis time series portend vegetation change: verification using long-term data in an arid grassland ecosystem. Ecol. Appl. 27 (5), 1677–1693.

Burkholder, J., Libra, B., Weyer, P., Heathcote, S., Kolpin, D., Thorne, P.S., Wichman, M., 2007. Impacts of waste from concentrated animal feeding operations on water quality. Environ. Health Perspect. 115 (2), 308–312.

Burkholder, J., Libra, B., Weyer, P., Heathcote, S., Kolpin, D., Thorne, P.S., Wichman, M., 2007. Impacts of waste from concentrated animal feeding operations on water quality. Environ. Health Perspect. 115 (2), 308–312.

Cai, Y., He, H., Yang, K., Fatholahi, S.N., Ma, L., Xu, L., Li, J., 2021. A comparative study of deep learning approaches to rooftop detection in aerial images. Can. J. Remote Sens. 1–19.

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European conference on computer vision (ECCV), pp. 801–818.

Clayton, M., 2004. Wetlands get a new spy cam. Christian Science Monitor. URL https://www.csmonitor.com/2004/0122/p11s02-sten.html.

Copeland, C., 2010. Animal waste and water quality: Epa's response to the waterkeeper alliance court decision on regulation of cafos.

de Bem, P.P., de Carvalho Junior, O.A., Fontes Guimarães, R., Trancoso Gomes, R.A., 2020. Change detection of deforestation in the brazilian amazon using landsat data and convolutional neural networks. Remote Sens. 12(6), 901.

Evans, M.J., Malcom, J.W., 2019. Automated habitat change detection methods using satellite data to improve conservation law implementation.

Fytsilis, A.L., Prokos, A., Koutroumbas, K.D., Michail, D., Kontoes, C.C., 2016. A methodology for near real-time change detection between unmanned aerial vehicle and wide area satellite images. ISPRS J. Photogramm. Remote Sens. 119, 165–186.

GAO, 2008. Concentrated animal feeding operations: EPA needs more information and a clearly defined strategy to protect air and water quality from pollutants of concern. US Government Accountability Office. URL https://www.gao.gov/products/gao-08-944.

Gauthier, M.-F., Weir, L., Ou, Z., Arkett, M., De Abreu, R., 2007. Integrated satellite tracking of pollution: A new operational program. In: 2007 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 967–970.

Glicksman, R.L., Markell, D.L., Monteleoni, C., 2017. Technological innovation, data analytics, and environmental enforcement. Ecol. LQ 44, 41.
Gray, W.B., Shimshack, J.P., 2011. The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence. Rev. Environ. Econ. Policy 5 (1), 3–24.
Guay, T., 2012. Inspector's log. Environ. Compl. Alert 20 (449).
Gurian-Sherman, D., 2008. Cafos uncovered: The untold costs of confined animal feeding operations. URL https://www.ucsusa.org/sites/default/files/2019-10/cafos-uncovered-executive-summary.pdf.
Handan-Nader, C., Ho, D.E., 2019. Deep learning to map concentrated animal feeding operations. Nat. Sustain.
Handan-Nader, C., Ho, D.E., Liu, L.Y., 2020. Deep learning with satellite imagery to enhance environmental enforcement. Data-Driven Insights and Decisions: A Sustainability Perspective. Elsevier.
Hino, M., Benami, E., Brooks, N., 2018. Machine learning for environmental monitoring. Nat. Sustain. 1 (10), 583–588.
Howard, O., 2019. Illinois has open doors for CAFOs. Commun. Word. http://thecommunityword.com/online/blog/2019/07/31/illinois-has-open-door-for-cafos/.
Hribar, C., 2010. Understanding concentrated animal feeding operations and their impact on communities.
Huang, C., Goward, S.N., Masek, J.G., Thomas, N., Zhu, Z., Vogelmann, J.E., 2010. An automated approach for reconstructing recent forest disturbance history using dense landsat time series stacks. Remote Sens. Environ. 114 (1), 183–198.
Jiwani, A., Ganguly, S., Ding, C., Zhou, N., Chan, D.M., 2021. A semantic segmentation network for urban-scale building footprint extraction using rgb satellite imagery. arXiv preprint arXiv: 2104.01263.
Kennedy, R.E., Yang, Z., Cohen, W.B., 2010. Detecting trends in forest disturbance and recovery using yearly landsat time series: 1. LandTrendr–Temporal segmentation algorithms. Remote Sens. Environ. 114 (12), 2897–2910.
Koltunov, A., et al., 2020. eDaRT: The ecosystem disturbance and recovery tracker system for monitoring landscape disturbances and their cumulative effects. Remote Sens. Environ. 238 (1), 111482.
Koltunov, A., Ben-Dor, E., Ustin, S., 2009. Image construction using multitemporal observations and dynamic detection models. Int. J. Remote Sens. 30 (1), 57–83.
Krestenitis, M., Orfanidis, G., Ioannidis, K., Avgerinakis, K., Vrochidis, S., Kompatsiaris, I., 2019. Early identification of oil spills in satellite images using deep cnns. In: International Conference on Multimedia Modeling. Springer, pp. 424–435.
Laradji, I., Rodriguez, P., Kalaitzis, F., Vazquez, D., Young, R., Davey, E., Lacoste, A., 2020. Counting cows: Tracking illegal cattle ranching from high-resolution satellite imagery. arXiv preprint arXiv: 2011.07369.
MacDonald, J.M., Hoppe, R.A., Newton, D., 2018. Three decades of consolidation in us agriculture. Economic information bulletin. USDA.
MacDonald, J.M., McBride, W.D., 2009. The transformation of us livestock agriculture: Scale, efficiency, and risks. Economic information bulletin. USDA.
Maretto, R.V., Fonseca, L.M., Jacobs, N., Körting, T.S., Bendini, H.N., Parente, L.L., 2020. Spatio-temporal deep learning approach to map deforestation in amazon rainforest. IEEE Geosci. Remote Sens. Lett.
Marshall, J.L., 2015. In re: Appeal of groundwater pollution control permit no. ns2014502 smith valley dairy. https://sec.nv.gov/uploads/Appeal-Smith-Valley-Dairy/SOS_Reply_Brief.pdf.
Merced County, 2012. Draft environmental impact report of antonio azevedo dairy expansion CUP09-011. URL https://web2.co.merced.ca.us/pdfs/env_docs/eir/Azevedo_Dairy_DEIR.pdf.
Microsoft, 2018. Microsoft building footprints. https://github.com/microsoft/USBuildingFootprints.
Namoano, B., Starr, A., Emmanouilidis, C., Cristobal, R.C., 2019. Online change detection techniques in time series: An overview. In: 2019 IEEE International Conference on Prognostics and Health Management (ICPHM). IEEE, pp. 1–10.
Nicole, W., 2013. Cafos and environmental justice: The case of north carolina. Environ. Health Perspect. 121 (6), A182 (Online).
Nieto-Hidalgo, M., Gallego, A.-J., Gil, P., Pertusa, A., 2018. Two-stage convolutional neural network for ship and spill detection using slar images. IEEE Trans. Geosci. Remote Sens. 56 (9), 5217–5230.
OIG, 2005. EPA performance measures do not effectively track compliance outcomes. Office of Inspector General, Environmental Protection Agency 2006-P-00006. https://www.epa.gov/sites/production/files/2015-11/documents/20051215-2006-p-00006.pdf.
Peng, D., Zhang, Y., Guan, H., 2019. End-to-end change detection for high resolution satellite images using improved unet++. Remote Sens. 11 (11), 1382.
Purdy, R., 2010. Using earth observation technologies for better regulatory compliance and enforcement of environmental laws. J. Environ. Law 22 (1), 59–87.
Robinson, C., Ortiz, A., Ferres, J.M.L., Anderson, B., aHo, A.E., 2021. Temporal cluster matching for change detection of structures from satellite imagery. In: ACM SIGCAS Conference on Computing and Sustainable Societies.
Ronneberger, O., Fischer, P., Brox, T., 2015a. U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597 [cs]ArXiv: 1505.04597. http://arxiv.org/abs/1505.04597.
Ronneberger, O., Fischer, P., Brox, T., 2015b. U-net: Convolutional networks for biomedical image segmentation. CoRR abs/1505.04597. http://arxiv.org/abs/1505.04597.
Sefrin, O., Riese, F.M., Keller, S., 2021. Deep learning for land cover change detection. Remote Sens. 13 (1), 78.
Setiawan, Y., Kustiyo, K., Darmawan, A., 2016. A simple method for developing near real-time nationwide forest monitoring for indonesia using modis near-and shortwave infrared bands. Remote Sens. Lett. 7 (4), 318–327.
Son, J.-Y., Muenich, R.L., Schaffer-Smith, D., Miranda, M.L., Bell, M.L., 2021. Distribution of environmental justice metrics for exposure to cafos in North Carolina, Usa. Environ. Res. 195, 110862.
Soroka, A., Duren, Z., 2020. Poultry feeding operations on the delaware, maryland, and virginia peninsula from 2016 to 2017: U.s. geological survey data release.
Tang, X., Bullock, E.L., Olofsson, P., Woodcock, C.E., 2020. Can VIIRS continue the legacy of modis for near real-time monitoring of tropical forest disturbance? Remote Sens. Environ. 249, 112024.
USDA, 2017. United states department of agriculture: National agricultural statistics service. URL https://quickstats.nass.usda.gov/results/7812D0C6-7E41-37D8-B977-F6EE524257CA.
van den Burg, G.J., Williams, C.K., 2020. An evaluation of change point detection algorithms arXiv preprint arXiv:2003.06222.
Varghese, A., Gubbi, J., Ramaswamy, A., Balamuralidhar, P., 2018. Changenet: A deep learning architecture for visual change detection. In: Proceedings of the European Conference on Computer Vision (ECCV) Workshops.
Verbesselt, J., Hyndman, R., Newnham, G., Culvenor, D., 2010. Detecting trend and seasonal changes in satellite image time series. Remote Sens. Environ. 114 (1), 106–115.
Verbesselt, J., Hyndman, R., Zeileis, A., Culvenor, D., 2010. Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. Remote Sens. Environ. 114 (12), 2970–2980.
Wan, H., Shao, Y., Campbell, J.B., Deng, X., 2019. Mapping annual urban change using time series Landsat and NLCD. Photogramm. Eng. Remote Sens. 85 (10), 715–724.
Weinstein, B.G., 2018. A computer vision for animal ecology. J. Anim. Ecol. 87 (3), 533–545.
Xue, Y., Wang, T., Skidmore, A.K., 2017. Automatic counting of large mammals from very high resolution panchromatic satellite imagery. Remote Sens. 9 (9), 878.
Ye, S., Rogan, J., Zhu, Z., Eastman, J.R., 2021. A near-real-time approach for monitoring forest disturbance using landsat time series: Stochastic continuous change detection. Remote Sens. Environ. 252, 112167.
Zhu, Z., 2017. Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications. ISPRS J. Photogramm. Remote Sens. 130, 370–384.
Zhu, Z., Woodcock, C.E., 2014. Continuous change detection and classification of land cover using all available landsat data. Remote Sens. Environ. 144, 152–171.