Remote Sensing in Ecology and Conservation
Open Access
ZSL LET'S WORK FOR WILDLIFE

ORIGINAL RESEARCH

Predicting defoliator abundance and defoliation measurements using Landsat-based condition scores

Valerie J. Pasquarella[1], James G. Mickley[2,3], Audrey Barker Plotkin[4], Richard G. MacLean[5], Riley M. Anderson[6], Leone M. Brown[2,7], David L. Wagner[2], Michael S. Singer[6] & Robert Bagchi[2]

[1]Department of Earth & Environment, Boston University, 685 Commonwealth Avenue, Boston Massachusetts, 02215,
[2]Department of Ecology & Evolutionary Biology, University of Connecticut, Storrs, Connecticut,
[3]Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon,
[4]Harvard Forest, Harvard University, Petersham, Massachusetts,
[5]Massachusetts Department of Conservation and Recreation Division of Water Supply Protection, Belchertown, Massachusetts,
[6]Department of Biology, Wesleyan University, Middletown, Connecticut,
[7]Environmental Studies Program, Tufts University, Medford, Massachusetts, USA

Keywords
Defoliation, forest pests, Google Earth Engine, gypsy moth, Landsat time series, remote sensing

Correspondence
Valerie J. Pasquarella, Department of Earth & Environment, Boston University, 685 Commonwealth Avenue, Boston, MA 02215. Tel: +1 908 405 6557; Fax: +1 617 353 8399; E-mail: valpasq@bu.edu

Editor: Mat Disney
Associate Editor: Doreen Boyd

Received: 29 September 2020; Revised: 21 March 2021; Accepted: 12 April 2021

doi: 10.1002/rse2.211

Remote Sensing in Ecology and Conservation 2021;**7** (4):592–609

Abstract

Remote sensing imagery can provide critical information on the magnitude and extent of damage caused by forest pests and pathogens. However, monitoring short-term changes in deciduous forest condition caused by defoliating insects is challenging and requires approaches that directly account for seasonal vegetation dynamics. We implemented a previously published harmonic modeling approach for forest condition monitoring in Google Earth Engine and systematically assessed the relative ability of condition change products generated using various model parameterizations for predicting pest abundances and defoliation during the 2016–2018 gypsy moth (*Lymantria dispar*) outbreak in southern New England. Our comparisons revealed that most models made reasonable predictions of changes in canopy condition and egg and larval abundances of *L. dispar*, indicating a strong correlation between our harmonic-based estimates of condition change and defoliator activity. The greatest differences in predictive ability were in the spectral domain, with assessments based on Tasseled Cap Greenness, Simple Ratio, and the Enhanced Vegetation Index ranking among the top models, and the commonly used Normalized Difference Vegetation Index consistently exhibiting poorer performance. We also observed notable differences in the magnitude of scores for different baseline periods. Additionally, we found that Landsat-based condition scores better explained larval abundance than egg mass counts, which have historically been used as a proxy for later-season larval abundance, indicating that our remote sensing approach may be more accurate and cost-effective for generating consistent retrospective assessments of *L. dispar* population abundance in addition to estimates of canopy damage. These findings provide important linkages between spectral changes detected using a harmonic modeling approach and biophysical aspects of defoliator activity, with potential to extend monitoring and prediction to regional or even continental scales.

Introduction

Outbreaks of indigenous and non-native forest pests and pathogens can have far-reaching economic and ecological impacts (Dale et al., 2001; Dukes et al., 2009; Logan et al., 2003) and monitoring the extent and severity of pest and pathogen outbreaks is necessary for effective large-scale forest management (e.g. Hargrove et al., 2009). However, the vast geographical scale of forest monitoring poses a massive logistical and economic challenge, and outbreaks of defoliating insects, such as European gypsy moth (*Lymantria dispar*), forest tent caterpillar

(*Malacosoma disstria*), saddled prominent (*Heterocampa guttivitta*), and spruce budworm (*Choristoneura fumiferana*), remain challenging to assess due to the heterogeneous and ephemeral nature of defoliation severity and vegetation response. With improved access to remote sensing datasets and cloud-based processing resources, both near-real-time monitoring and retrospective assessment are becoming increasingly prevalent (e.g. Coops et al., 2020; Norman & Christie, 2020; Senf et al., 2017). Thus, there remains a critical need to assess the utility of multi-temporal satellite imagery for detecting and characterizing defoliator outbreaks.

Effective monitoring of defoliator impacts using satellite imagery relies on observations of both defoliated and non-defoliated states, and numerous methodologies for quantifying deviations from reference or "baseline" conditions have been proposed (e.g. Chávez et al., 2019; Olsson et al., 2016; Rock et al., 1986; Spruce et al., 2011; Townsend et al., 2012). These methods can produce estimates of both the magnitude and extent of pest damage; however, choice of a baseline to monitor against is a key element of change detection processes (Norman & Christie, 2020). Rather than compare imagery from different time periods, harmonic regression models fit to time series of reflectance observations can be used to generate "synthetic" images that represent predicted reflectance values for specific dates (Zhu et al., 2015). Prior work has demonstrated advantages of a harmonic condition monitoring approach for detecting changes in forest condition in response to a recent *L. dispar* outbreak in southern New England including greater spatial and temporal resolution relative to aerial survey methods (Pasquarella et al., 2017, 2018a). Resulting products have been used in studies correlating spore deposition of the *L. dispar* fungal pathogen *E. maimaiga* with defoliation (Elkinton et al., 2019), determining how nitrogen dynamics interact with defoliation (Conrad-Rooney et al., 2020), and estimating impacts of multi-year defoliation events on streamflow and seasonal water yield (Smith-Tripp et al., unpubl. data). However, improved access to cloud-based datasets and computing resources has resulted in new opportunities for developing and testing methods that rely on dense time series of satellite imagery (Gorelick et al., 2017).

In this study, we used a new google earth engine (GEE) implementation of the condition monitoring approach described in Pasquarella et al. (2017) to evaluate the predictive ability of Landsat-based condition scores for estimating *L. dispar* abundance and damage. Specifically, we address two questions:

1 How well do spectral estimates of vegetation condition change correlate with the abundance of defoliating pests, or with actual defoliation?
2 How do different harmonic baseline parameterizations perform in estimating pest abundance or defoliation?

Our results confirm the utility of Landsat-based condition change scores as a proxy for changes in defoliator populations and canopy condition during a regional-scale outbreak event and offer new insights into the sensitivity of harmonic baselines for forest condition monitoring.

Materials and Methods

Study system

We used field datasets collected during the 2016–2018 gypsy moth outbreak in southern New England for our analyses. Introduced to Massachusetts in 1869, the European gypsy moth is well-established in the northeastern US (Elkinton & Liebhold, 1990). As a generalist defoliator, gypsy moths are known to feed on the foliage of numerous tree species, with *Quercus*, *Populus*, and *Larix* species being among the most common hosts (Liebhold et al., 2000). While periodic outbreaks have occurred for over a century, the 2016–2018 outbreak was the first widespread irruption in this region since the successful establishment of the fungal pathogen *Entomophaga maimaiga* in 1989 (Andreadis & Weseloh, 1990; Hajek et al., 1995). We focused on field data and Landsat-based condition assessments from 2017, as this year represented peak outbreak conditions as well as maximum overlap among our reference datasets.

Landsat-based forest condition monitoring workflow

The GEE platform provides both the Landsat datasets and image processing tools needed to scale harmonic-based condition monitoring analysis and enables more operational product generation. As a key step in this analysis, we adapted the Landsat-based condition monitoring approach presented in Pasquarella et al. (2017) to run from the GEE Code Editor (Fig. 1). Details on methods are provided in Appendix S1.

Baseline model experiments

We used this GEE condition monitoring workflow to develop a series of 32 experiments that consider a variety of parameter choices for establishing multi-year "baseline" models representing relatively stable forest conditions for each Landsat pixel in our study area. We tested a full set of factors including (a) spectral vegetation index used for model fitting, (b) baseline-modeling period, (c) frequencies of harmonic regression terms, and (d) differences in Landsat time series input imagery (Table 1).

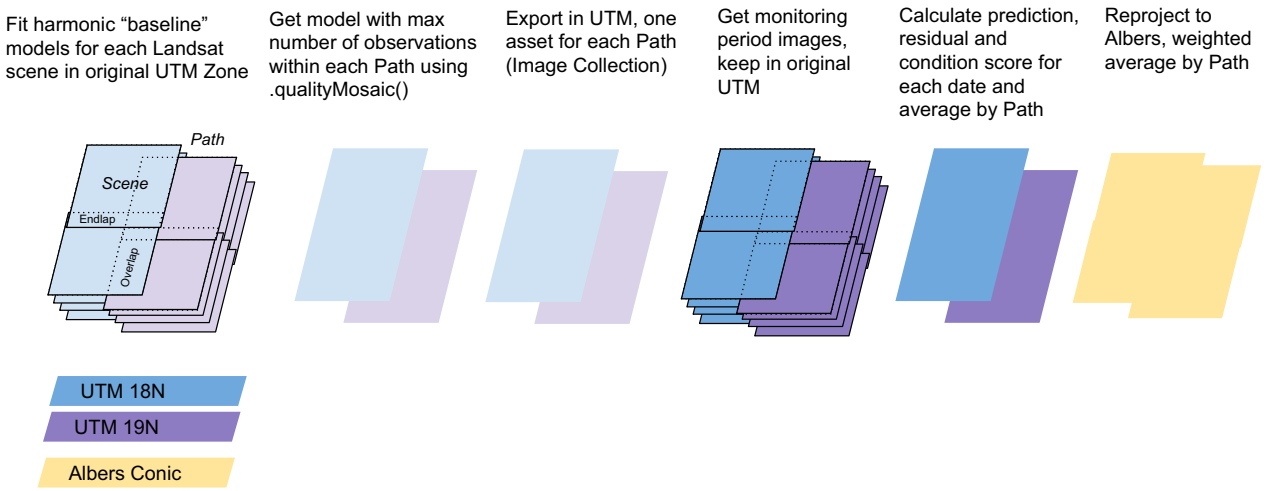

Figure 1. Remote sensing workflow in google earth engine (GEE). Harmonic baseline models are fit for stacks of images within individual Landsat scene footprints. Scene-based results are then merged by orbital Path to remove duplicate models for the same image acquisition date. Baseline model images, including harmonic regression coefficients, number of observations used for fitting and model RMSE, are exported and saved in native UTM projection. Baselines are then used to generate predictions for images acquired during a user-specified monitoring period (May 1–September 30, 2017 in this study), and residuals (observed minus predicted values) as well as condition scores (residuals normalized by baseline RMSE are calculated for each Path acquisition date. Scores are averaged over the monitoring period to generate a mean condition score for each Path. Finally, mean scores for individual Paths are reprojected to a common coordinate system (Albers Equal Area Conic) and combined using a weighted average based on the number of observations used for monitoring. The result is a seamless season-integrated condition assessment product with an average condition score (RMSE-normalized spectral deviation from baseline conditions) for each 30-m pixel.

Table 1. Experimental parameters used for the baseline model comparison.

Parameter	Description	Options tested
Spectral transforms	Vegetation index used for both monitoring and prediction	Tasseled Cap Greenness (TCG) Normalized Difference Vegetation Index (NDVI) Simple Ratio (SR) Enhanced Vegetation Index (EVI)
Baseline period	Years included in baseline model fitting	2000–2010 2005–2015
Harmonic frequencies	Set of harmonic frequencies used for baseline model fitting	1/365.25, 2/365.25 (h12) 1/365.25, 3/365.25 (h13)
Time series image inputs	Frequency of Landsat observations, choice to include or exclude Landsat 7 due to Scan Line Corrector artifacts	All available observations (full) Single-sensor (16d)

Spectral transforms

Vegetation indices that combine the red, near infrared (NIR), and shortwave infrared (SWIR) bands are typically used for detecting changes in canopy cover and condition (Rullan-Silva et al., 2013), and we focused on a subset of indices that are well-characterized by harmonic functions, i.e. indices that rely primarily on the visible and NIR bands. Previous analyses (Pasquarella et al., 2017, 2018a) were based on time series of tasseled cap greenness (TCG), which is a linear combination of six Landsat optical bands that contrasts absorption in the visible and second short-wave infrared bands with reflectance in the near-infrared as distinctive properties of green vegetation (Crist, 1985; Crist & Kauth, 1986). However, due to differences in data structures and sensor calibration, Tasseled Cap coefficients are sensor-dependent, making it difficult to generalize to different instruments. We therefore tested other common vegetation indices, including the normalized difference vegetation index (NDVI; Huete et al., 2002; Jackson & Huete, 1991; Kriegler et al., 1969; Tucker, 1978), the enhanced vegetation index (EVI; Huete et al., 2002), and the simple ratio (SR; Jackson & Huete, 1991).

Baseline period

We considered two different baseline periods, 2000–2010 and 2005–2015, for direct comparison with previous results. Both periods represent pre-outbreak conditions in Connecticut and Massachusetts. If undisturbed forest conditions are assumed to be relatively stable in terms of long-term phenology, any recent subset of years should theoretically result in similar baseline model estimates. However, the 2000–2010 period includes only imagery from Landsat 5 (L5) and Landsat (L7), while the 2005–2015 period also incorporates Landsat 8 (L8) acquisitions, and year-to-year differences in the timing and number of clear observations across the Landsat record as well as assumptions about stable reference periods remain an important source of testable uncertainty in the quality and utility of baseline regression results.

Harmonic frequencies

Baseline modeling was performed using a linear least squares regression with observed spectral values as the response and six predictor terms: intercept, slope, and two harmonic sine and cosine pairs (Pasquarella et al., 2017). For each of the four spectral indices, we tested two different harmonic model specifications that varied the frequencies of the second sine and cosine terms. The first option paired an annual frequency with a bi-annual frequency (i.e. periodicity of 1 and 2). The second paired an annual frequency with a tri-annual frequency (i.e. periodicity of 1 and 3). The choice of harmonic frequencies ("h12" or "h13") used for the baseline regression may result in differences in model fit, particularly when there is strong asymmetry in seasonal reflectance profiles that is not well-characterized by a single annual harmonic. These differences are expected to manifest as variability in baseline model root mean squared error (RMSE), which will in turn influence assessment results, since baseline RMSE is used to normalize differences between observed and predicted values in condition score calculation (Pasquarella et al., 2017).

Time series image inputs

The L7 scan line corrector (SLC) failure causes notable spatial artifacts in condition assessment products, presumably due to differences in number and timing of observations across scan line gaps. Therefore, we experimented with fitting baseline models to time series of all available observations ("full") versus using L7 only when L5 and L8 images are not available ("16d"). This allowed us to test trade-offs between utilizing the maximum temporal density of time series inputs and minimizing impact of SLC artifacts.

Existing "reanalysis" product

We included an existing assessment product, the "reanalysis" result (Pasquarella, 2018b), and to compare the original workflow implementation with new results generated using GEE. The reanalysis product used a 2000–2010 baseline fit to full time series of TCG using annual and tri-annual harmonic frequencies. Rather than assume this existing product represented a target output, we considered it relative to the GEE results in our baseline assessment analysis.

Monitoring period

To test the effect of varying baseline parameterizations on ability to predict metrics of canopy change and defoliator abundance using condition scores, we applied baseline models to a fixed monitoring period that included all Landsat acquisitions from May 1 to September 30, 2017. We generated predicted values for each acquisition date during the monitoring period, calculated condition scores (observed minus predicted spectral value, divided by baseline RMSE), then averaged scores over the monitoring period to produce a final "season-integrated" assessment following methods described in Pasquarella et al. (2017).

The May 1 to September 30 monitoring period was selected to be consistent with previous analyses and captures the phenology of defoliation and recovery in southern New England. *Lymantria dispar* completes one life cycle per season: it hatches in late spring, develops through 5–6 in stars with peak feeding (defoliation) in late June, and pupates by early July (Doane & McManus, 1981). Defoliated trees may produce a new flush of leaves (e.g. de Beurs & Townsend, 2008); however, the defoliation signal is expected to persist even as refoliation occurs. Thus, averaging condition scores for all observations during the monitoring period should produce a relatively robust estimate of overall change in forest condition that is suitable for comparison with field-based measurements collected during the same period.

Field datasets

The 32 baseline model experiments were evaluated using measurements from two networks of field sites (Table 2). The Quabbin Watershed defoliation survey (the "defoliation" study) rapidly examined late-season *L. dispar* damage in 2017. Six ~350 hectare "hotspot representing a range of forest types and Landsat condition scores were selected across the Quabbin Reservoir Watershed (Fig. 2). Within each hotspot, 100 random points were identified and sampled on the ground using horizontal point sampling (2.296 BAF; Bitterlich, 1947) between 21 September and 6

Table 2. Key attributes of the field datasets.

Field plots	Defoliation	*Lymantria dispar* abundance
Location	Central Massachusetts, US	Eastern Connecticut, US
Timing	2017	2017, 2018
	Surveys conducted September 21–October 6	May (egg mass counts), June (larval surveys)
Sample design	Random points stratified by forest type and damage in 2017	Blocked forest fragments of contrasting area selected from CLEAR dataset (Vogt et al., 2007)
Plot design and survey approach	Variable-radius point sampling, rapid assessment (N = 486 sites)	Sample of four trees at vertices of 125-m triangle (>10-cm dbh) (N = 96 sites, 384 trees)
Key measurements	Tree defoliation status, species, and diameter at breast height (dbh)	Egg mass counts, larval abundance

October, 2017, before canopy senescence (e.g. Fig. 2B). For each tree at a point, species; defoliation class (1 = 75%–100% foliage remaining, 2 = 50%–75% foliage, 3 = 25%–50% foliage, 4 = 0%–25% foliage); diameter at breast height (by 5-cm size classes); and canopy exposure (1 = full sun, 2 = partially shaded, 3 = mostly shaded) were recorded. The defoliation estimate for each tree represented net canopy damage after defoliation, including any recovery of new foliage after cessation of gypsy moth larval feeding in early July. Because of time constraints, sampling was completed for only 486 of the 600 points. Of these, three points had no trees because of recent timber harvest, and were therefore excluded from the analysis.

The Eastern Connecticut study (the "*Lymantria*" study) sampled forest fragments across a 3500 km^2 area. Sampled forest fragments ranged from 3 to 1013 ha and were arranged in 13 blocks, each including 2–3 neighboring fragments of different areas. Surveys of *L. dispar* abundance were conducted in 2017 and 2018 at 32 forest sites, though we focused only on 2017 measurements to facilitate comparison with the defoliation study. Each of the 32 fragments contained three sampling points arranged in a triangle, spaced 125 m apart (Fig. 2C). At each point, the closest tree >10-cm diameter at breast height from the center point in each cardinal direction was selected and identified to species (four trees per sampling point for a total of 12 trees per site). In May of each year, we wrapped 20-cm wide burlap at chest height around the tree bole to serve as a day-time refuge and pupation site for late instar larvae (Wagner, 2005). We counted *L. dispar* egg masses on the bottom two meters of tree trunks in May, and *L. dispar* larvae on the bottom two meters of the trunks and under the burlap in early June and again in late June.

Baseline model evaluation

Condition scores were extracted for each field plot location such that plot centers were associated with the nearest Landsat pixel (30-m scale), and sampled values were exported from GEE as CSV data tables. We conducted all model evaluation analyses in R (R Development Core Team, 2019), using the tidyverse package (Wickham et al., 2019) for data wrangling, glmmTMB (Brooks et al., 2017) for running generalized linear mixed models (GLMM), bbmle (Bolker & R Development Core Team, 2020) for model AIC comparisons, and the easystats ecosystem and sjPlot (Lüdecke, 2019; Lüdecke et al., 2019) to calculate R^2 values and generate figures.

To test whether Landsat condition scores predicted observed changes in forest canopy condition at the Massachusetts field sites, we constructed models with mean defoliation score of the trees at each sample point as the response variable. We calculated the mean defoliation score using the midpoint of the defoliation class for each tree such that the mean defoliation score for each point was a continuous proportion bounded by 0.125 and 0.875. We fit GLMMs assuming a beta error distribution (logit link; Douma & Weedon, 2019), and included hotspot as a random intercept term to control for landscape-level variation.

To test whether Landsat condition scores predicted *L. dispar* abundance from the Connecticut field sites, we constructed GLMMs with larval abundance or egg masses as the response variable. We pooled counts of egg masses or larval abundance across the four trees at each survey point, and included a random intercept term for site to control for landscape-level variation. The residuals from initial models that assumed a Poisson distribution of errors were overdispersed, so we assumed negative binomial distributions for the errors in the models presented here (Richards, 2008). Though data were collected in additional years, we fit models using only 2017 data for more direct comparison with the defoliation study. Since *L. dispar* egg mass counts have been used to estimate larval population abundance (Liebhold et al., 1994), we also fit models using egg mass counts as a predictor of larval abundance to

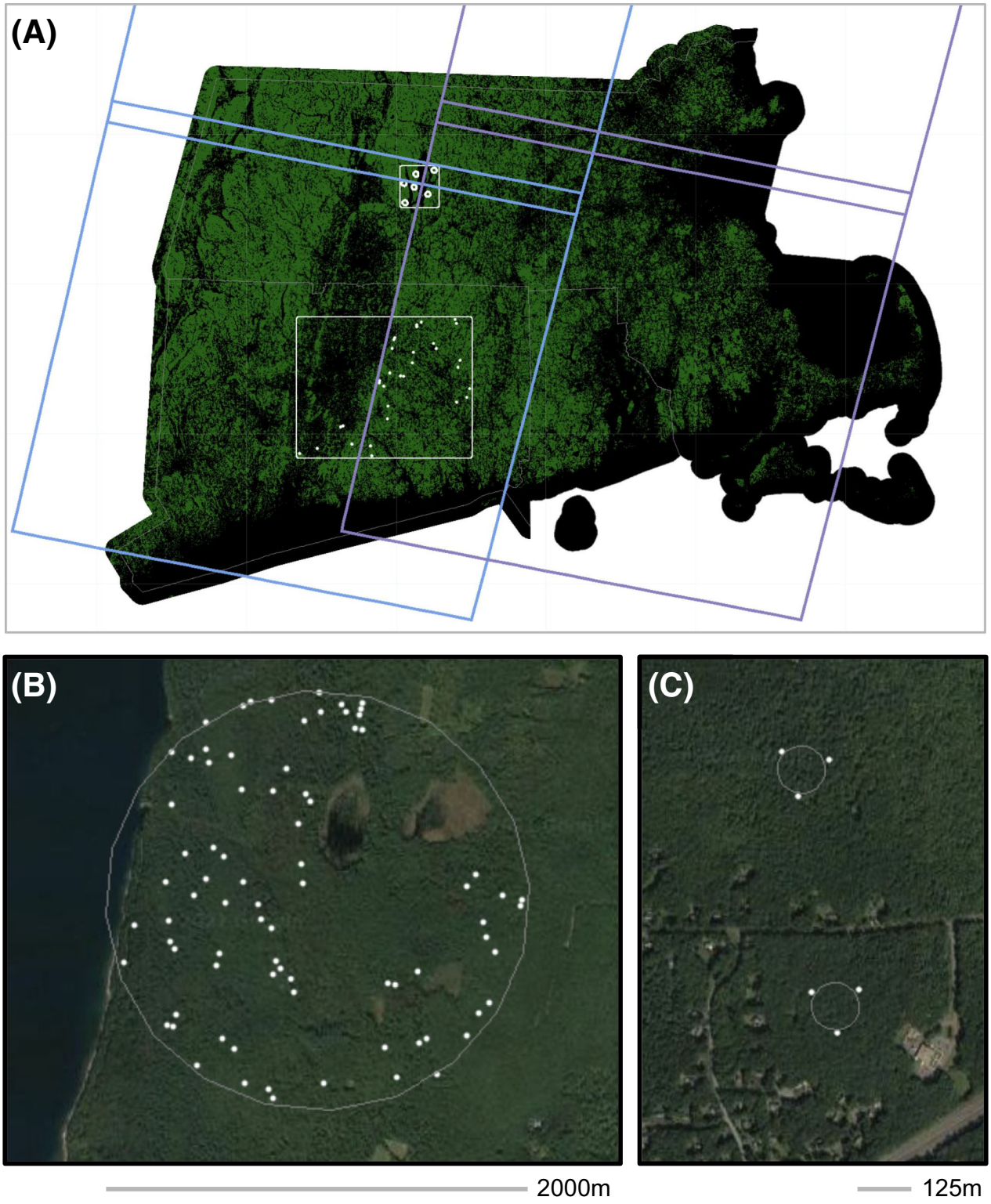

Figure 2. Field sites: (A) overview, showing the Connecticut and Massachusetts plot locations and Landsat scene boundaries, including overlap and endlap) areas; green shading indicates pixels with greater than 75% forest cover (source: USDA USFS, 2016 TCC Product Suite); (B) an example hotspot from the Massachusetts defoliation study; white points indicate defoliation sample sites; (C) an example of two survey sites from *Lymantria dispar* study in Connecticut; white points indicate the vertices of a 125 m triangle within each *Lymantria dispar* study site, egg masses and caterpillars on the bark of burlap-wrapped trees were counted on four trees spaced five meters in each cardinal direction from each vertex.

compare the predictive ability of Landsat change in condition versus egg mass counts.

We used AIC scores to evaluate the performance of different baseline models. We considered all models with a ΔAIC <6 as feasible alternatives within a 95% model set (Richards, 2005). To quantify the ability of each model to explain observed variation in the response variable, we calculated marginal and conditional R^2 values (Nakagawa & Schielzeth, 2013), where marginal R^2 represents the proportion of deviance explained by fixed factors (Landsat condition score), and conditional R^2 represents the deviance explained by both fixed factors and random factors (CT site or MA hotspot). The combination of these metrics provides insights into the relative performance of the 32 GEE baseline experiments, as well as the existing reanalysis product, for predicting field-based measures of *L. dispar* abundance and impacts.

Results

Defoliation dataset (Massachusetts)

Comparison of AIC scores for the defoliation survey data indicate that the 2005–2015 baseline fit to time series of all available TCG observations using annual and tri-annual harmonic frequencies was the best predictor of net changes in canopy condition (Fig. 3), and explained 60% of the variance in ground-observed defoliation (marginal $R^2 = 0.60$, conditional $R^2 = 0.63$). All other models had substantially higher AIC scores ($\Delta AIC > 6$) suggesting that this baseline significantly outperformed other parameterizations. However, marginal and conditional R^2 values were similar to the reanalysis results, indicating that the existing products explained a comparable amount of variability in canopy change estimates. Mapped results for the best-performing baseline for the defoliation survey data showed widespread low- to moderate-magnitude changes in condition, with pockets of higher-magnitude change (Fig. 4A and B), and there was a significant negative relationship between condition assessment scores and net defoliation (Fig. 4C; $\beta = -0.52 \pm 0.039$, $z = -13.34$, $P < 0.001$).

Larval abundance dataset (Connecticut)

Comparison of AIC scores for the *L. dispar* larval dataset indicated that the 2000–2010 baseline fit to a time series of SR observations with a 16-day repeat (limited use of L7 imagery) using annual and tri-annual harmonic frequencies was the best predictor of larval abundance (Fig. 5), and explained 46% of the variance in abundance estimates (marginal $R^2 = 0.46$, conditional $R^2 = 0.73$). However, SR baselines for the same 2000–2010 period fit to the full time series and using annual and bi-annual harmonics showed closely comparable results, and 19 other baselines were within the 95% model set ($\Delta AIC < 6.00$), with similar marginal R^2 values to the top model. While the marginal R^2 for the larval models tended to be lower than those for the defoliation dataset, the conditional R^2 were higher, indicating that site/landscape factors were more important in Connecticut and accounted for a larger portion of variation in larval abundance. The reanalysis product did not perform well for this dataset. Mapped results for the top-ranked baseline for the *L. dispar* larval dataset (Fig. 6A and B) showed similar distribution but lower magnitude condition scores compared to the top model for the defoliation survey (Fig. 4), and a significant negative relationship between larval abundance and defoliation scores (Fig. 6C; $\beta = -0.60 \pm 0.094$, $z = -6.57$, $P < 0.001$).

Egg mass abundance dataset (Connecticut)

Comparison of AIC scores for the *L. dispar* egg masses dataset indicated that the 2000–2010 baseline fit to a time series of all available TCG observations using annual and bi-annual harmonic frequencies was the best predictor of egg mass counts (Fig. 7), and explained 43% of the variance in count data (marginal $R^2 = 0.43$, conditional $R^2 = 0.87$). The comparable model using annual and tri-annual frequencies, as well as the reanalysis product, performed very similarly, and 13 models were within the 95% model set ($\Delta AIC < 6.00$). The difference between marginal R^2 values and conditional R^2 values were even greater for the egg mass dataset compared to larval dataset, indicating more variability of egg mass abundance among sites than larvae or canopy defoliation. Mapped results for the top-ranked baseline for the *L. dispar* egg masses data were very similar to the top-ranked defoliation model, tending to show widespread changes in condition and high magnitude scores (Fig. 8A and B). There was a significant negative relationship between defoliation scores and egg mass abundance (Fig. 8C; $\beta = -1.92 \pm 0.327$, $z = -5.86$, $P < 0.001$).

Comparing the relative ability of egg masses and remote-sensed defoliation to predict larval abundance, we found that egg mass counts were poorer predictors of larval caterpillar abundance (marginal $R^2 = 0.08$, conditional $R^2 = 0.72$, Fig. 9) than Landsat-based condition scores (Fig. 5).

Comparison of condition scores across baselines

The distributions of condition scores from the various baseline experiments and the existing reanalysis product for the *L. dispar* and defoliation plots varied in magnitude

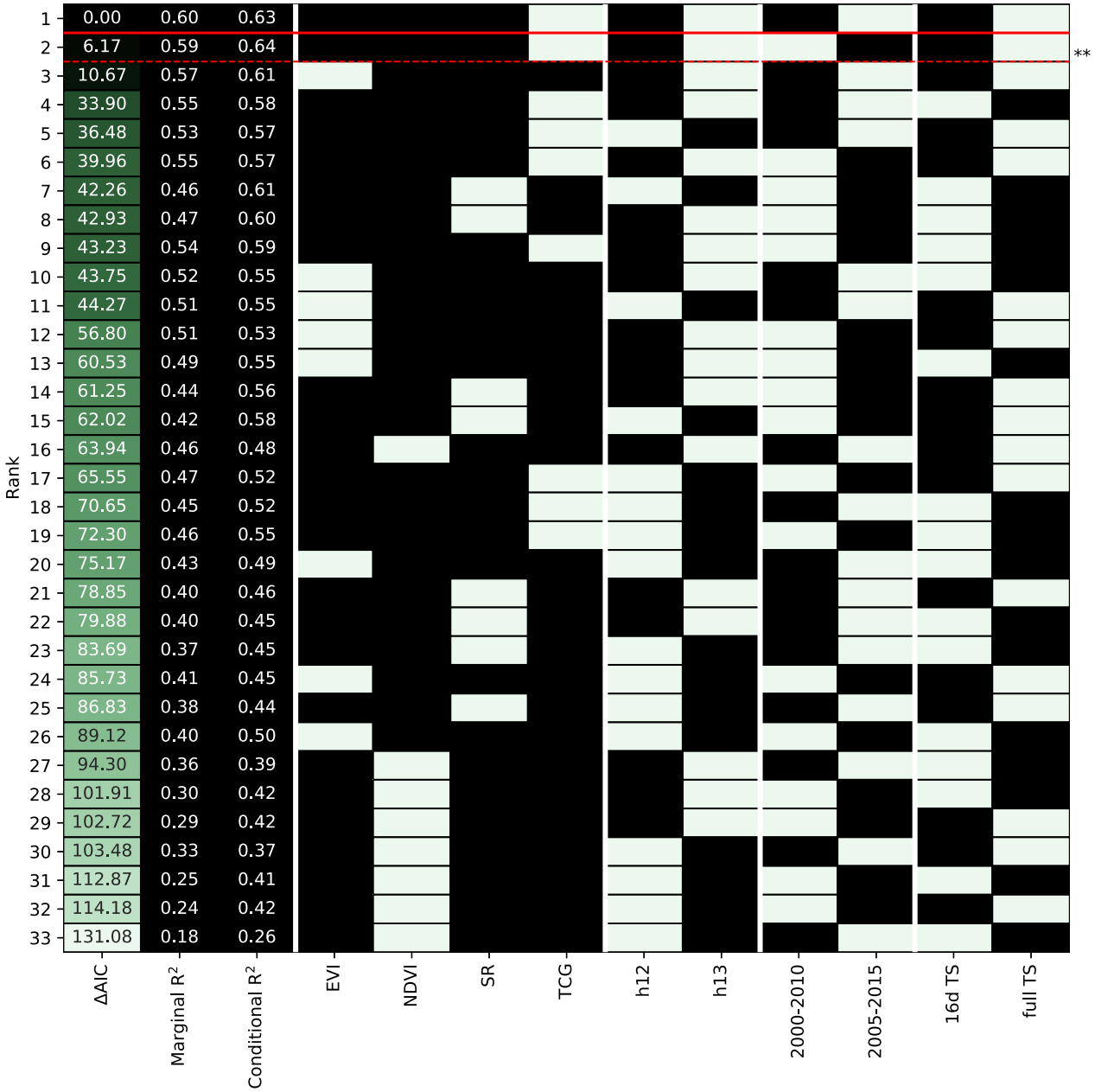

Figure 3. Baseline model comparison for 2017 defoliation results (Massachusetts plot network). Models are ranked according to ΔAIC score (first column, green), and marginal and conditional R^2 values are also included for each model. Red line indicates the approximate 95% model set (models with a ΔAIC < 6). Parameterization options are shown as a binary heatmap, with light cells indicating which option (spectral, harmonic, baseline period, and temporal inputs) was used for a given baseline experiment in order to visualize patterns relative to baseline rankings. ** indicates reanalysis results.

and range (Fig. 10). All condition scores were calculated based on the same 2017 Landsat observations and were normalized by their corresponding baseline RMSE. This accounted for differences in scaling across different spectral vegetation indices such that differences among baselines result from differences in predicted values as well as the quality of baseline fit. Of the four spectral indices considered, NDVI produced scores that generally fell within the range of expected baseline noise (between −1 and 1 times the baseline RMSE) during this known outbreak year, which partially explains its poor performance relative to other indices. The 2005–2015 baselines tended

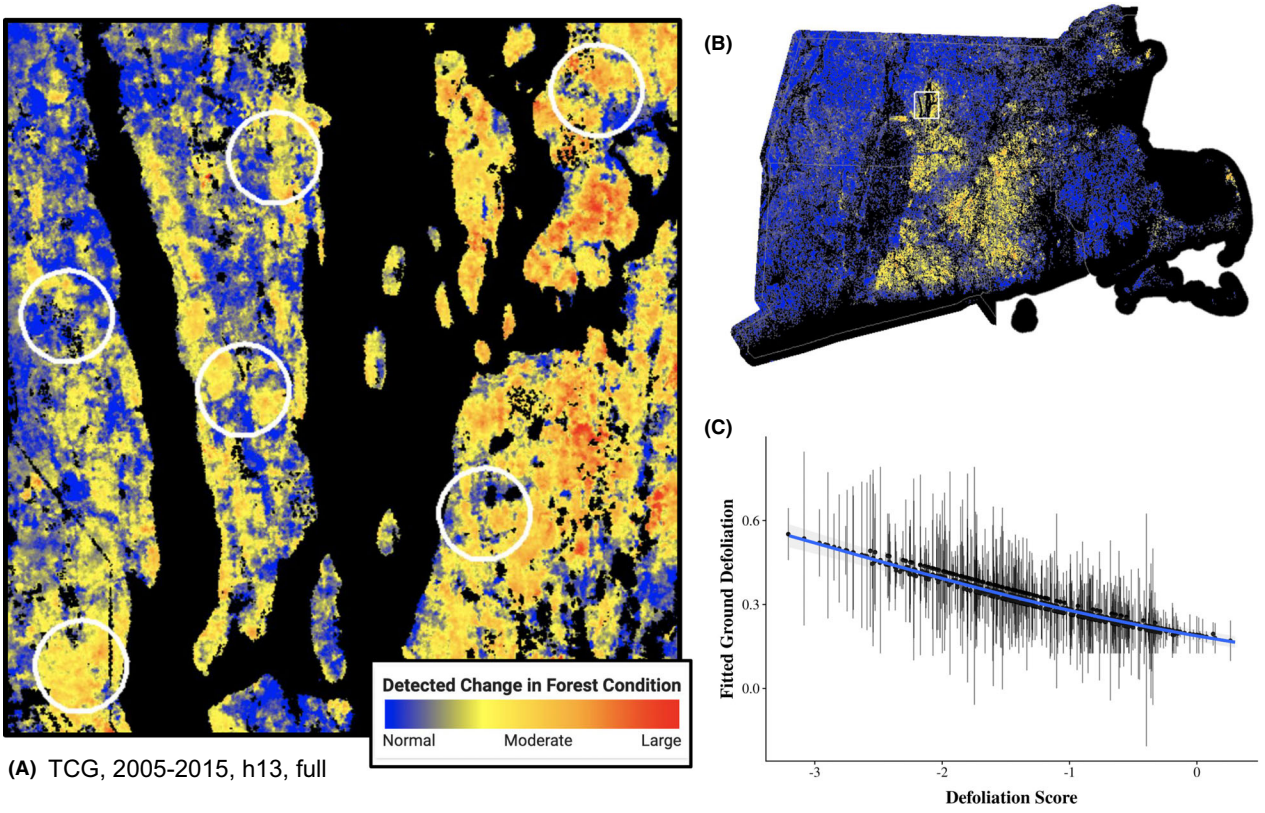

Figure 4. Mapped results and regression analysis for the top model for the defoliation dataset (TCG, 2005–2015, h13, full). (A) Zoomed view of Massachusetts plot sites (n = 486), which are clustered in six "hotspots" indicated by white circles. (B) Overview for the full southern New England study area. (C) Regression model predictions for fitted ground defoliation as a function of Landsat-based defoliation score, black lines indicate model residuals, and the 95% confidence interval on predictions is indicated by gray shading.

to produce more negative scores than comparable 2000–2010 baselines, and these differences are perhaps most apparent in the NDVI and SR experiments.

Discussion

Using only Landsat-based condition scores as a predictor, our models explained up to 60% of the variation in observed canopy change and up to 45% of the variation in *L. dispar* egg mass and larval abundance during the 2017 gypsy moth outbreak across field sites in Connecticut and Massachusetts after accounting for landscape-level variation (Figs. 3, 5, and 7). Historically, egg mass counts have been used as a proxy for later-season larval abundance, especially with regard to predicting where control measures might be needed (Doane & McManus, 1981; Liebhold et al., 1994). When compared to condition scores, however, egg masses only explained 8% of the variation in larval abundance. Thus, while egg mass counts remain useful for predicting larval abundance and damage in the immediate future, our remote sensing approach may be more accurate and cost-effective for generating consistent retrospective assessments of *L. dispar* abundance. Additionally, while it is difficult to rigorously quantify changes in canopy biomass and leaf-area once an outbreak is underway, we found that condition scores were a strong predictor of observed changes in canopy condition and can characterize relative differences in defoliation severity. Therefore, while model rankings and the magnitude of condition scores were somewhat sensitive to Landsat image inputs and baseline model parameterization, many models were able to characterize observed patterns in outbreak populations of *L. dispar* as well as defoliation impacts relatively well, and our results suggest harmonic-based condition scores are generally a suitable proxy for both changes in host condition and pest abundance over broad spatial extents.

Of the parameters tested, the spectral vegetation index used for harmonic model fitting and condition monitoring was perhaps the most important determinant of model quality. TCG and SR were consistently among the top-ranked models, while NDVI-based models exhibited relatively poor explanatory ability across all field-based datasets. Vegetation change studies using MODIS imagery

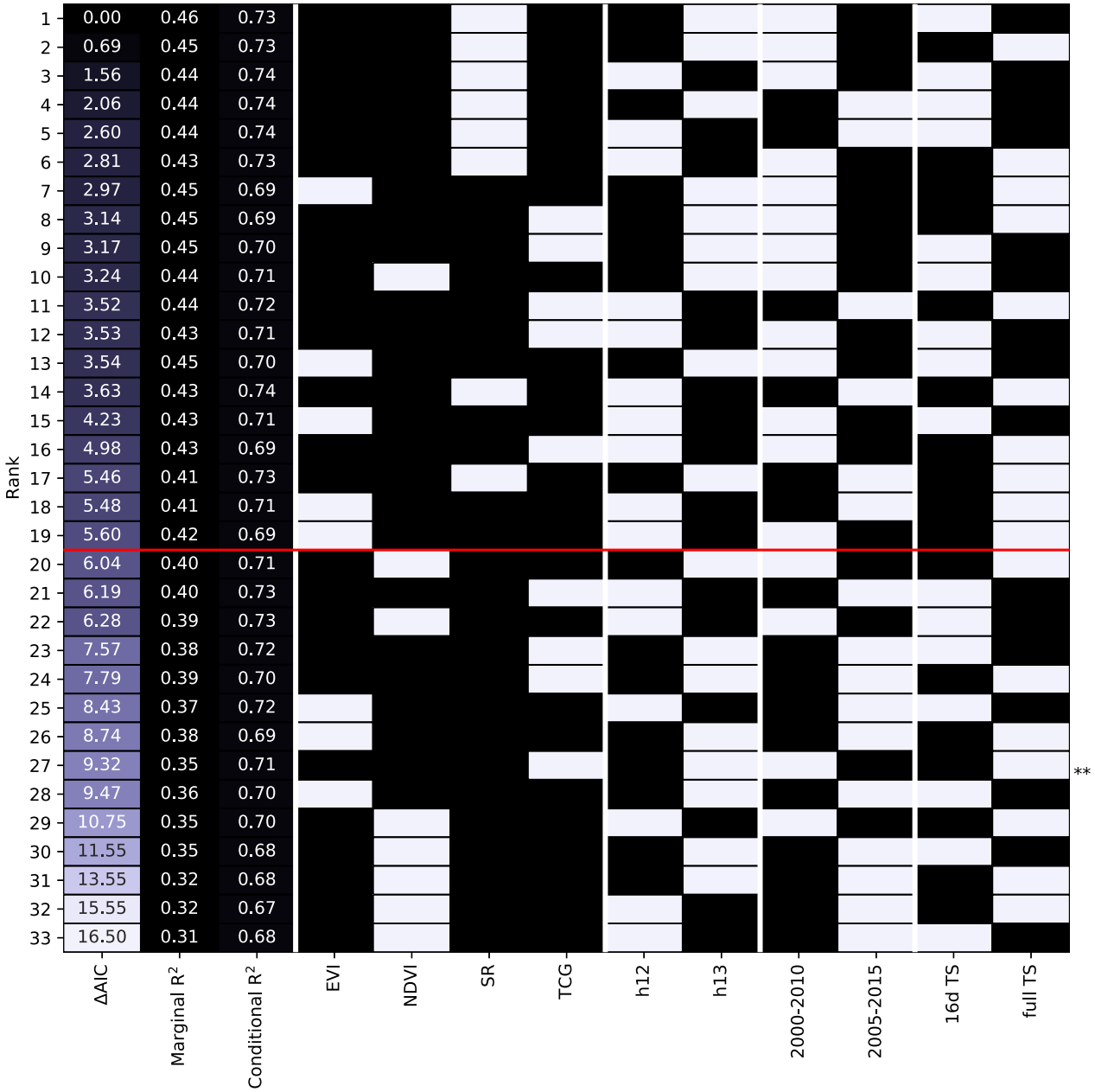

Figure 5. Baseline model comparison for 2017 larval survey/caterpillar results (Connecticut plot network). Models are ranked according to ΔAIC score (first column, purple), and marginal and conditional R^2 values are also included for each model. Red line indicates the approximate 95% model set (models with a ΔAIC < 6). Parameterization options are shown as a binary heatmap, with light cells indicating which option (spectral, harmonic, baseline period, and temporal inputs) was used for a given baseline experiment in order to visualize patterns relative to baseline rankings. ** indicates reanalysis results.

have often relied on NDVI due to higher spatial resolution of the MODIS visible and NIR bands (e.g. Chávez et al., 2019; Jepsen et al., 2009; Spruce et al., 2011). However, NDVI signals are known to saturate in regions of high vegetation biomass (e.g. Huete et al., 1997) and visualizing the distribution of spectral values by month for our field sites confirms low variability in NDVI values during growing season months (June–September) relative to TCG, SR, and EVI (Appendix S2). It is worth noting that comparable analysis using *L. dispar* data for 2018 (a non-outbreak year) did include NDVI among top-ranked models (Appendix S3), suggesting that there may be stronger correlation when population and defoliation signals are both low. However, we suspect that the observed

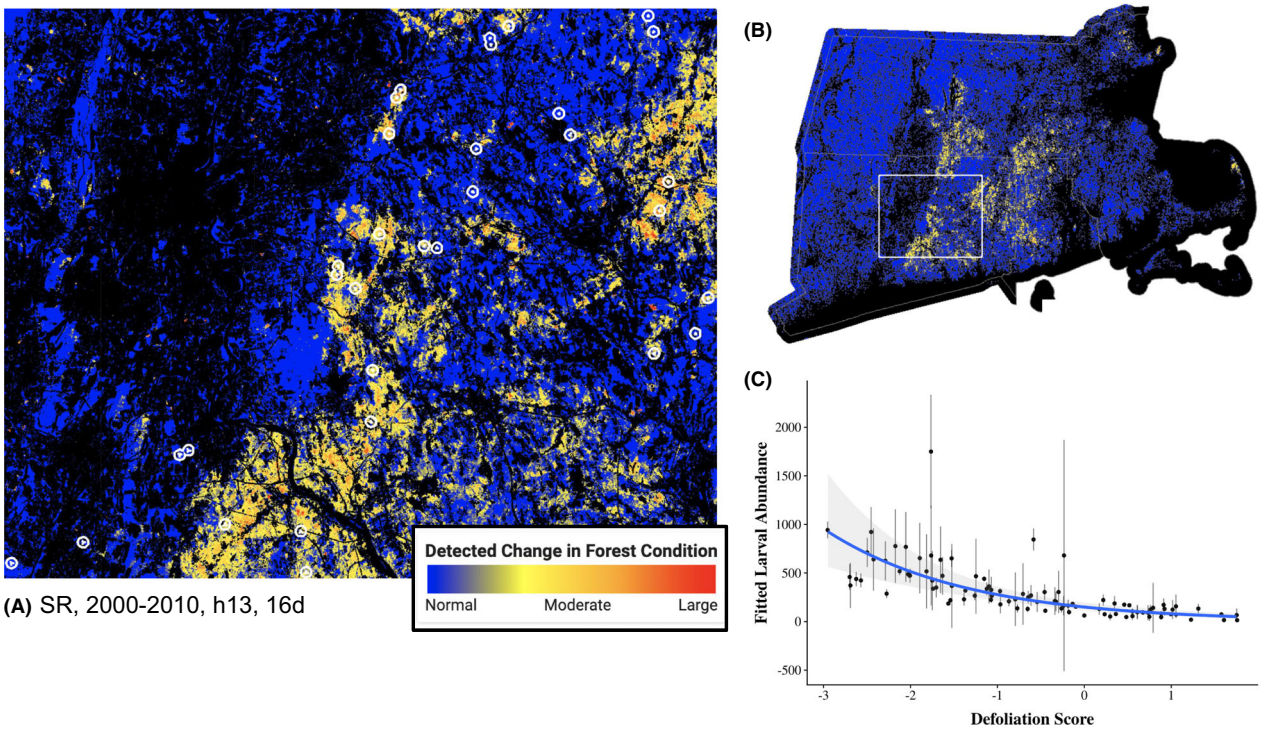

Figure 6. Mapped results and regression analysis for the top model for the larva/caterpillar dataset (SR, 2000–2010, h13, 16d). (A) Zoomed view of Connecticut plot sites (n = 96), which are distributed to sample forest fragments across the broader landscape; plot locations are indicated by circled white points. (B) Overview for the full southern New England study area. (C) Regression model prediction for fitted larval abundance as a function of Landsat-based defoliation score, black lines indicate model residuals, and the 95% confidence interval on predictions is indicated by gray shading.

low dynamic range and saturation effects make NDVI less effective for characterizing more variable conditions during an outbreak year, and thus support our conclusion that NDVI is not a reliable indicator of forest health and condition.

Though we found TCG often outperformed other vegetation indices, potentially due to combined weighting of visible, NIR, and SWIR information, sensor-specific differences in available spectral bands and corresponding wavelengths may limit the utility of the TCG transform for future work. For example, integrating observations from the Sentinel-2 series would require a different set of coefficients and harmonization process (e.g. Shi & Xu, 2019). Commercial high-resolution instruments like PlanetScope sensors also offer opportunities to improve spatial and temporal solution; however, imaging is currently limited to the visible and NIR bands. In this case, our findings suggest that SR is a better choice than NDVI for monitoring changes in forest conditions, at least in temperate deciduous forests.

The somewhat higher rankings of "h13" relative to "h12" harmonic frequencies may be explained by the higher frequency of the second term in the "h13" pairing, which is likely better suited for characterizing asymmetries in the seasonal profiles of vegetation indices. The tests of "full" compared with "16d" time series inputs were inconclusive. It is generally assumed that harmonic regression is robust to variability in input frequency and density of Landsat observations, however, spatial artifacts from Landsat 7 remain evident in mapped results when scan lines are present in baseline and/or monitoring period imagery. These artifacts are most prevalent near scene boundaries, and would be expected to have a greater impact on CT sites, however, differences in scores do not indicate a clear preference among time series options and the distribution of sites may fail to adequately test the effects of this parameterization choice.

Results concerning the choice of years to include in the baseline modeling period were also inconclusive, with the 2005–2015 baseline generally ranking higher in terms of predictive ability for defoliation sites, while the 2000–2010 baseline exhibited stronger performance for the *L. dispar* abundance sites. The mapped results (4, 6, and 8) and the distribution of condition scores for our plot sites (Fig. 10) suggest that the baseline period is a source of considerable variability, and differences in magnitude of

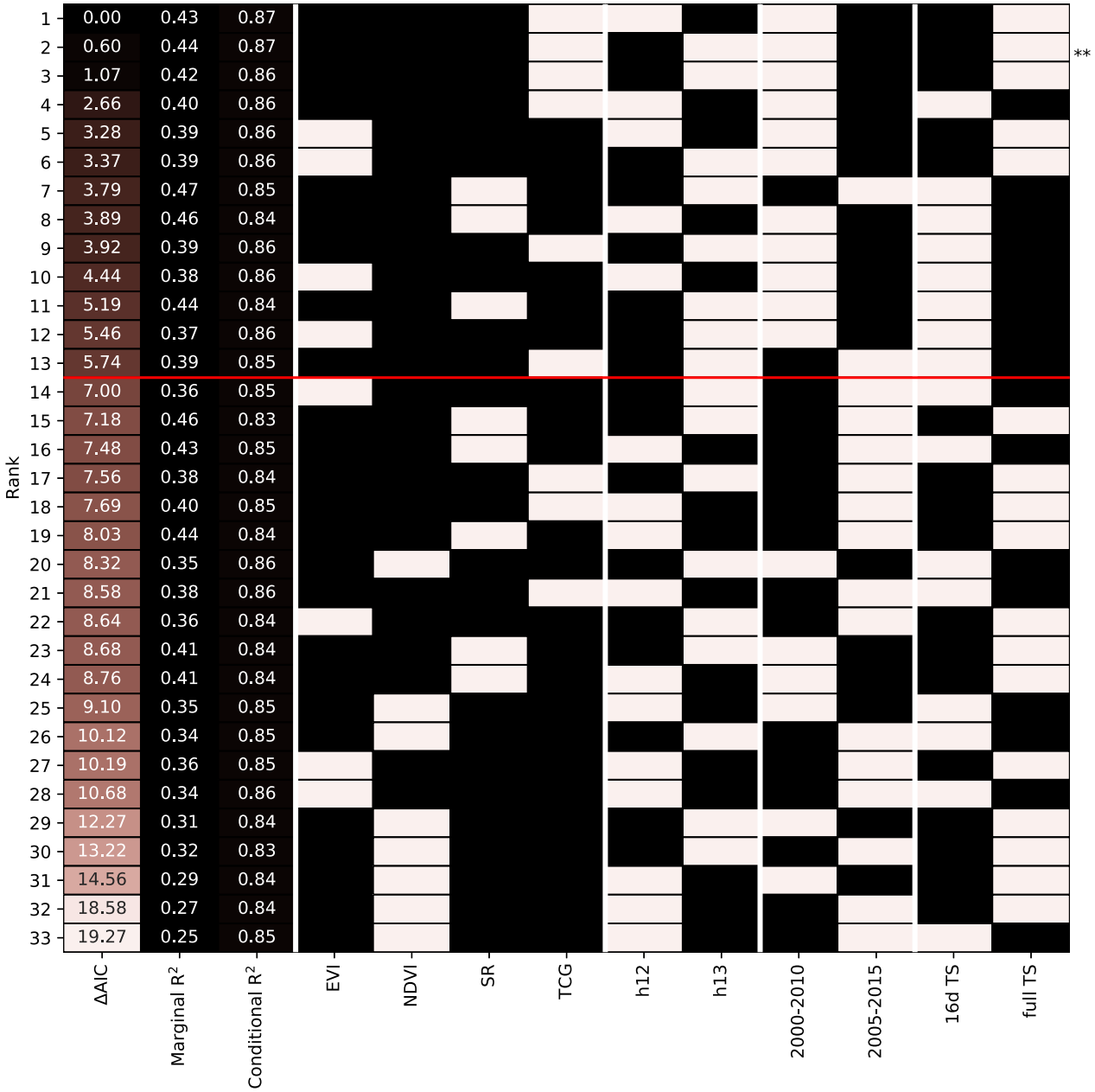

Figure 7. Baseline model comparison for 2017 egg mass count results (Connecticut plot network). Models are ranked according to ΔAIC score (first column, brown), and marginal and conditional R^2 values are also included for each model. Red line indicates the approximate 95% model set (models with a ΔAIC < 6). Parameterization options are shown as a binary heatmap, with light cells indicating which option (spectral, harmonic, baseline period, and temporal inputs) was used for a given baseline experiment in order to visualize patterns relative to baseline rankings. ** indicates reanalysis results.

scores and model rankings warrant further exploration and testing over a larger set of baseline periods. We assumed that longer decadal-scale reference periods were preferable to shorter baselines and that a single, fixed-reference period (e.g. 2000–2010) could be used across a broad spatial extent in our study. However, forest pest outbreaks exhibit complex spatial dynamics (Foster et al., 2013). For gypsy moth outbreaks, the Northeast will differ from more southern and western regions of the US in timing and additional biotic factors may also come into play. For instance, *E. maimaiga* is believed to be responsible for the dampening outbreak population cycles in the Northeast for several decades prior to the 2016–2018 event (i.e. Hajek & van Nouhuys, 2016). Thus, testing

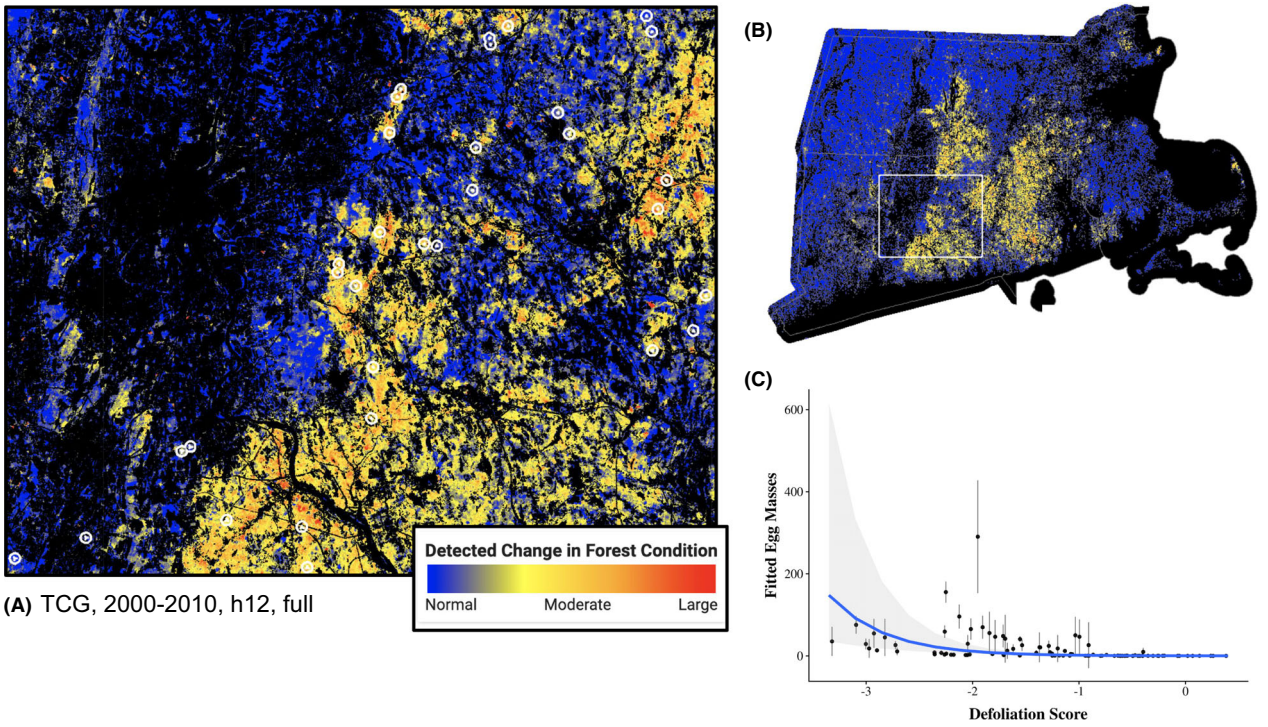

Figure 8. Mapped results and regression analysis for the top model for the egg mass dataset (TCG, 2000–2010, h12, full). (A) Zoomed view of Connecticut plot sites (n = 96), which are distributed to sample forest fragments across the broader landscape; plot locations are indicated by circled white points. (B) Overview for the full southern New England study area. (C) Regression model predictions for fitted egg mass counts as a function of Landsat-based defoliation score, black lines indicate model residuals, and the 95% confidence interval on predictions is indicated by gray shading.

additional baseline options and improving criteria for baseline selection will be important next steps for generalizing the approach used in this study to regions that have experienced more persistent yet spatially and temporally variable outbreak cycles.

We also recommend further exploration of criteria for setting the monitoring period. Defoliator emergence and thus timing of impacts are strongly linked with local climate, particularly spring temperatures and precipitation, as well as host phenology and predator populations (Rullan-Silva et al., 2013). Therefore, flexible monitoring periods will be necessary to account for variability in these factors over larger spatial extents and target optimal "biological windows" for detecting pest-specific damage. Fortunately, gypsy moth population dynamics are well-studied, and there may be opportunities to link condition monitoring efforts with insect life stage models such as BioSIM (Régnière, 1996) and use long-term averages in pupation phenology to adjust monitoring period in a spatially explicit manner.

Differences in scaling between field-based and remotely sensed observations, as well as heterogeneity in the forests and forest conditions sampled may explain differences in the relative importance of site factors across models. While the defoliation field survey was specifically designed to integrate with Landsat-based observations and used preliminary results from 2017 to help stratify plot locations, the *L. dispar* sites were previously established and data were collected for a more limited sample of trees. The field efforts also varied in their spatial scope, with the Massachusetts efforts concentrated within the Quabbin Watershed Management Area while the Connecticut study sampled a much broader landscape (Fig. 2).

Additional site variability likely resulted from localized differences in the severity and timing of defoliation, with areas of Connecticut experiencing more severe defoliation in 2016 than sites in Massachusetts, which were only patchily impacted until the peak of the defoliation event in 2017. We also expect the strength of correlation between larval abundance and Landsat-based condition change estimates to increase over the course of the season, as caterpillars need to be large enough to cause sufficient damage to be detected by remote sensing technology. Yet while later-season images may better characterize maximum damage, the lag between the occurrence of damage and acquisition of clear Landsat observations that can be used for condition assessment, as well as potential for recovery of foliage and noise in spectral measurements including atmospheric

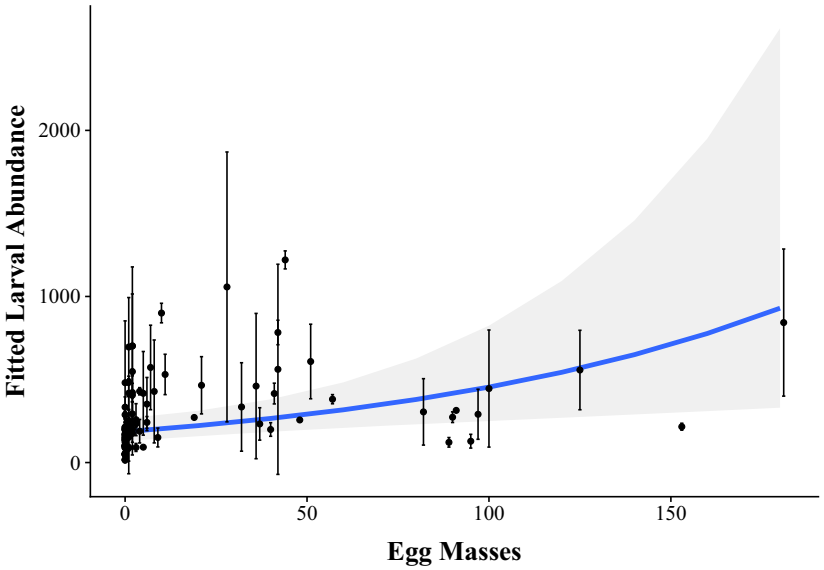

Figure 9. Abundance of *L. dispar* from burlap surveys predicted by egg mass counts at the same sites in 2017. *Lymantria dispar* egg masses predict larval abundance in 2017, but with considerable variation across points sampled.

contamination and cloud shadows, are unavoidable sources of uncertainty. Assessments based on single-date images can provide important perspectives on near-real-time changes in condition; however, averaging across multiple individual assessments to produce condition score products like those used in this study remains key in automating the generation of spatially complete and temporally stable assessments that do not require substantial human interpretation and can be more readily combined with field-based measurements.

It is interesting to note that the "reanalysis" product was ranked second for both the canopy change and egg mass datasets, but much lower for the larva dataset. The existing product also did not consistently outperform the GEE baseline with comparable parameterization, with condition scores from the reanalysis product tending to be more negative with a wider interquartile range. These differences may be attributed to differences in the harmonic regression approach as well as other pre- and post-processing steps, including cloud masking and reprojection, and we recommend the GEE implementation for future use based on both performance across other parameterizations and greater reproducibility.

Conclusions

The results of this study provide evidence of strong correlations between Landsat-based condition scores and various field datasets and suggest it is feasible to predict changes in canopy cover and *L. dispar* populations at different life stages given modeled and observed spectral properties. Both existing products and those generated using a new GEE implementation of our condition monitoring approach serve as suitable proxies for defoliator activity, with models based on time series of TCG and SR tending to have the greatest predictive abilities. Looking to the future, the spatial and temporal resolution of current Landsat-based products could be further enhanced by integrating imagery from other optical sensors such as Sentinel-2 and PlanetScope. However, our present ability to retroactively quantify both the magnitude and extent of outbreak events over large areas using Landsat time series has powerful potential for improving forest pest monitoring and management at regional to continental scales.

Acknowledgments

VJP and ABP were supported by the Harvard Forest LTER (NSF DEB #1832210). VJP was also supported in part by USDA Forest Service Forest Health Monitoring Evaluation Monitoring Award #18-CA-11420004-293 and a USGS/Landsat Science Team award to Dr. Curtis Woodcock (#140G118C0006). RB, DLW, MSS, RMA, JGM, and LMB and the larger Connecticut survey effort were supported by NSF DEB #1557086 and the Doris Duke Conservation Scholars Programs at UC-Santa Cruz and Northern Arizona University. We gratefully acknowledge Cas Carroll, Riley Hawkins-Hecock, Hanna Holcomb, Stephanie Melara, Amanda Minicucci, Hooman Musavi, Adrianna Perez, Nikki Pirtel, Jessica Romero, and Christopher White for field work assistance. The Massachusetts survey was conducted in collaboration with the

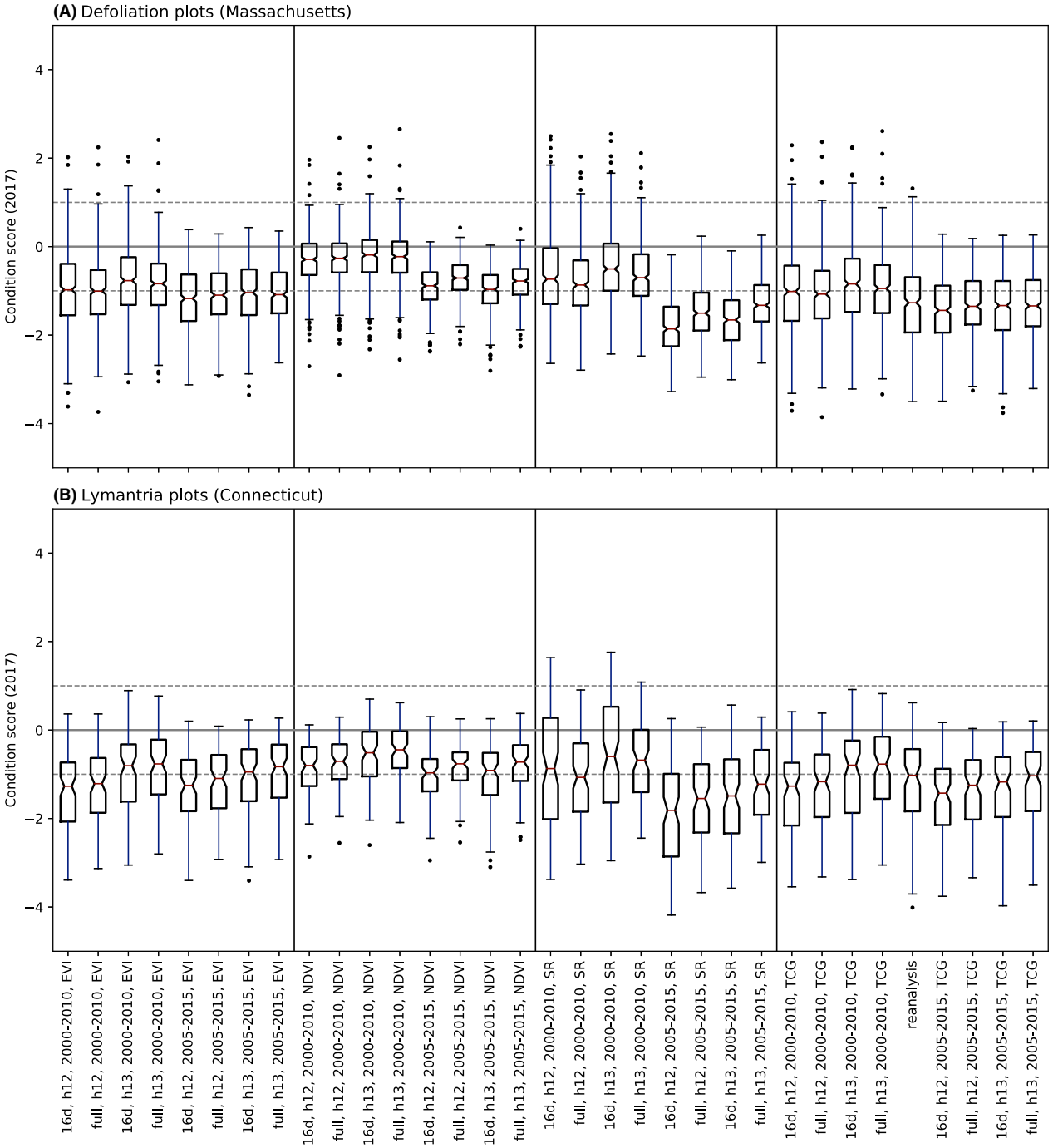

Figure 10. Distribution of 2017 condition assessment scores for the 32 baseline experiments as well as existing "reanalysis" product (which used the tcg_2000-2010_h13_full parameterization) for (A) Massachusetts defoliation and (B) connecticut *Lymantria dispar* plots. Experiments are grouped by spectral index. Dotted lines at −1 and 1 indicate the scores considered to be within the expected range of spectral variability relative to baseline RMSE. More negative scores correspond to loss of vegetation cover and/or decreases in vegetation quality/condition.

Massachusetts Department of Conservation and Recreation Division of Water Supply Protection Office of Watershed Management, and thank Derek Beard, Kenneth Canfield, Virginia Dautreuil, Herm Eck, Helen Johnson, Brian Keevan, and Steven Wood for field work and hotspot data collection. We also thank Vincent Chan for organizing and digitizing datasheets and field notes and two anonymous reviewers for their constructive feedback.

Data Availability Statement

R code and tabular datasets used for the harmonic baseline comparisons are available at github.com/BagchiLab-Uconn/Forest-Condition-Assessment. This repository also includes scripts to reproduce the Earth Engine condition monitoring workflow described in Appendix S1 as well as a set of map visualizations. Baseline experiment map products are available as Earth Engine assets and are archived at doi.org/10.5281/zenodo.4567382.

References

Andreadis, T.G. & Weseloh, R.M. (1990) Discovery of *Entomophaga maimaiga* in North American gypsy moth, , *Lymantria dispar*. *Proceedings of the National Academy of Sciences*, **87**(7), 2461–2465.

Bitterlich, W. (1947) Measurement of basal area per hectare by means of angle measurement. *Allg. Forest. Holzwirtsch, Ztg*, **58**, 94–96.

Bolker, B. & R Development Core Team. (2020). *bbmle: Tools for general maximum likelihood estimation*. R package version 1.0.23. Available at: https://CRAN.R-project.org/package=bbmle.

Brooks, M., Kristensen, K., Benthem, K., Magnusson, A., Berg, C., Nielsen, A. et al. (2017) glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, **9**(2), 378–400.

Chávez, R.O., Rocco, R., Gutiérrez, Á.G., Dörner, M. & Estay, S.A. (2019) A self-calibrated non-parametric time series analysis approach for assessing insect defoliation of broad-leaved deciduous *Nothofagus pumilio* forests. *Remote Sensing*, **11**(2), 204.

Conrad-Rooney, E., Barker Plotkin, A., Pasquarella, V.J., Elkinton, J., Chandler, J.L. & Matthes, J.H. (2020) Defoliation severity is positively related to soil solution nitrogen availability and negatively related to soil nitrogen concentrations following a multi-year invasive insect irruption. *AoB Plants*, **12**(6), plaa059.

Coops, N.C., Shang, C., Wulder, M.A., White, J.C. & Hermosilla, T. (2020) Change in forest condition: Characterizing non-stand replacing disturbances using time series satellite imagery. *Forest Ecol. Management*, **474**, 118370.

Crist, E.P. (1985) A TM tasseled cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, **17**(3), 301–306.

Crist, E.P. & Kauth, R.J. (1986) The tasseled cap de-mystified. *Photogrammetric Engineering Remote Sensing*, **52**, 81–86.

Dale, V.H., Joyce, L.A., Mcnulty, S., Neilson, R.P., Ayres, M.P., Flannigan, M.D. et al. (2001) Climate change and forest disturbances: climate change can affect forests by altering the frequency, intensity, duration, and timing of fire, drought, introduced species, insect and pathogen outbreaks, hurricanes, windstorms, ice storms, or landslides. *BioScience*, **51**(9), 723–734.

De Beurs, K.M. & Townsend, P.A. (2008) Estimating the effect of gypsy moth defoliation using MODIS. *Remote Sensing of Environment*, **112**(10), 3983–3990.

Doane, C.C. & McManus, M.L. (1981) The gypsy moth: research toward integrated pest management. *USDA Forest Service Technical Bulletin*, **1584**, 757.

Douma, J.C. & Weedon, J.T. (2019) Analysing continuous proportions in ecology and evolution: A practical introduction to beta and Dirichlet regression. *Methods in Ecology and Evolution*, **10**(9), 1412–1430.

Dukes, J.S., Pontius, J., Orwig, D., Garnas, J.R., Rodgers, V.L., Brazee, N. et al. (2009) Responses of insect pests, pathogens, and invasive plant species to climate change in the forests of northeastern North America: What can we predict? *Canadian Journal of Forest Research*, **39**(2), 231–248.

Elkinton, J.S., Bittner, T.D., Pasquarella, V.J., Boettner, G.H., Liebhold, A.M., Gould, J.R. et al. (2019) Relating aerial deposition of *Entomophaga maimaiga* conidia (Zoopagomycota: Entomophthorales) to mortality of gypsy moth (Lepidoptera: Erebidae) larvae and nearby defoliation. *Environmental Entomology*, **48**(5), 1214–1222.

Elkinton, J.S. & Liebhold, A.M. (1990) Population dynamics of gypsy moth in North America. *Annual Review of Entomology*, **35**(1), 571–596.

Foster, J.R., Townsend, P.A. & Mladenoff, D.J. (2013) Spatial dynamics of a gypsy moth defoliation outbreak and dependence on habitat characteristics. *Landscape Ecology*, **28**(7), 1307–1320.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. & Moore, R. (2017) Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, **202**, 18–27.

Hajek, A.E., Humber, R.A. & Elkinton, J.S. (1995) Mysterious origin of *Entomophaga maimaiga* in North America. *American Entomologist*, **41**(1), 31–43.

Hajek, A.E. & van Nouhuys, S. (2016) Fatal diseases and parasitoids: From competition to facilitation in a shared host. *Proceedings of the Royal Society B: Biological Sciences*, **283**(1828), 20160154.

Hargrove, W.W., Spruce, J.P., Gasser, G.E. & Hoffman, F.M. (2009) Toward a national early warning system for forest disturbances using remotely sensed canopy phenology. *Photogrammetric Engineering & Remote Sensing*, **75**, 1150–1156.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X. & Ferreira, L.G. (2002) Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, **83**(1–2), 195–213.

Huete, A.R., Liu, H. & van Leeuwen, W.J. (1997) The use of vegetation indices in forested regions: issues of linearity and saturation. In *IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing - A Scientific Vision for Sustainable Development*. IEEE, Vol. **4**, pp. 1966–1968. https://doi.org/10.1109/IGARSS.1997.609169

Jackson, R.D. & Huete, A.R. (1991) Interpreting vegetation indices. *Preventive Veterinary Medicine*, **11**(3–4), 185–200.

Jepsen, J.U., Hagen, S.B., Høgda, K.A., Ims, R.A., Karlsen, S.R., Tømmervik, H. et al. (2009) Monitoring the spatio-temporal dynamics of geometrid moth outbreaks in birch forest using MODIS-NDVI data. *Remote Sensing of Environment*, **113**(9), 1939.

Kriegler, F.J., Malila, W.A., Nalepka, R.F. & Richardson, W. (1969) *Preprocessing transformations and their effects on multispectral recognition*. Proceedings of the Sixth International Symposium on Remote Sensing of Environment, University of Michigan, Ann Arbor, MI, 13–16 October 1969 (Ann Arbor, MI: ERIM), pp. 97–131.

Liebhold, A., Elkinton, J., Williams, D. & Muzika, R.M. (2000) What causes outbreaks of the gypsy moth in North America? *Population Ecology*, **42**(3), 257–266.

Liebhold, A., Thorpe, K., Ghent, J. & Lyons, D.B. (1994) *Gypsy moth egg mass sampling for decision-making: a users' guide*. Radnor, PA: United States Department of Agriculture, Forest Service. NA-TP-94-04.

Logan, J.A., Régnière, J. & Powell, J.A. (2003) Assessing the impacts of global warming on forest pest dynamics. *Frontiers in Ecology and the Environment*, **1**(3), 130–137.

Lüdecke, D. (2019) *sjPlot: Data Visualization for Statistics in Social Science. R package version 2.8.1*. Available online at: https://CRAN.R-project.org/package=sjPlot [Accessed 22nd September 2020].

Lüdecke, D., Waggoner, P.D. & Makowski, D. (2019) Insight: A unified interface to access information from model objects in R. *Journal of Open Source Software*, **4**(38), 1412.

Nakagawa, S. & Schielzeth, H. (2013) A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**(2), 133–142.

Norman, S.P. & Christie, W.M. (2020). *Satellite-based evidence of forest stress and decline across the conterminous United States for 2016, 2017, and 2018. Gen. Tech. Rep. SRS-250*. Asheville, NC: US Department of Agriculture. Forest Service, Southern Research Station, pp. 151–166.

Olsson, P.O., Lindström, J. & Eklundh, L. (2016) Near real-time monitoring of insect induced defoliation in subalpine birch forests with MODIS derived NDVI. *Remote Sensing of Environment*, **181**, 42–53.

Pasquarella, V.J. (2018b). *Landsat-based gypsy moth defoliation assessment (Southern New England) (Version 3.0)*. Zenodo. https://doi.org/10.5281/zenodo.

Pasquarella, V.J., Bradley, B.A. & Woodcock, C.E. (2017) Near-real-time monitoring of insect defoliation using Landsat time series. *Forests*, **8**(8), 275.

Pasquarella, V.J., Elkinton, J.S. & Bradley, B.A. (2018a) Extensive gypsy moth defoliation in Southern New England characterized using Landsat satellite observations. *Biological Invasions*, **20**(11), 3047–3053.

R Development Core Team. (2008) *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Régnière, J. (1996) Generalized approach to landscape-wide seasonal forecasting with temperature-driven simulation models. *Environmental Entomology*, **25**(5), 869–881.

Richards, S.A. (2005) Testing ecological theory using the information-theoretic approach: examples and cautionary results. *Ecology*, **86**(10), 2805–2814.

Richards, S.A. (2008) Dealing with overdispersed count data in applied ecology. *Journal of Applied Ecology*, **45**(1), 218–227.

Rock, B.N., Vogelmann, J.E., Williams, D.L., Vogelmann, A.F. & Hoshizaki, T. (1986) Remote detection of forest damage: Plant responses to stress may have spectral "signatures" that could be used to map, monitor, and measure forest damage. *BioScience*, **36**(7), 439–445.

Rullan-Silva, C.R., Olthoff, A.E., de la Mata, J.A.D. & Alonso, A.P. (2013) Remote monitoring of forest insect defoliation. A review. *Forest Systems*, **22**(3), 377–391.

Senf, C., Seidl, R. & Hostert, P. (2017) Remote sensing of forest insect disturbances: Current state and future directions. *International Journal of Applied Earth Observation and Geoinformation*, **60**, 49–60.

Shi, T. & Xu, H. (2019) Derivation of tasseled cap transformation coefficients for Sentinel-2 MSI at-sensor reflectance data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **12**(10), 4038–4048.

Smith-Tripp, S., Griffith, A., Pasquarella, V. & Hatala Mattes, J. (unpubl. data). *Impacts of a regional multi-year insect defoliation event on seasonal water yield and instantaneous streamflow characteristics*.

Spruce, J.P., Sader, S., Ryan, R.E., Smoot, J., Kuper, P., Ross, K. et al. (2011) Assessment of MODIS NDVI time series data products for detecting forest defoliation by gypsy moth outbreaks. *Remote Sensing of Environment*, **115**(2), 427–437.

Townsend, P.A., Singh, A., Foster, J.R., Rehberg, N.J., Kingdon, C.C., Eshleman, K.N. et al. (2012) A general Landsat model to predict canopy defoliation in broadleaf deciduous forests. *Remote Sensing of Environment*, **119**, 255–265.

Tucker, C.J. (1978) Red and photographic infrared linear combinations for monitoring vegetation. *NASA Technical Memorandum*, **79620**, 37.

USDA USFS. (2016) *TCC product suite*. Available at: https://data.fs.usda.gov/geodata/rastergateway/treecanopycover/ [Accessed 29 Jul 2020].

Vogt, P., Riitters, K., Estrenguil, C., Kozak, J., Wade, T. & Wickham, J. (2007) Mapping spatial patterns with morphological image processing. *Landscape Ecology*, **22**, 171–177.

Wagner, D.L. (2005) *Caterpillars of Eastern North America: A guide to identification and natural history*. Princeton, New Jersey: Princeton University Press, p. 512.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R. et al. (2019) Welcome to the Tidyverse. *Journal of Open Source Software*, **4**(43), 1686.

Zhu, Z., Woodcock, C.E., Holden, C. & Yang, Z. (2015) Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time. *Remote Sensing of Environment*, **162**, 67–83.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Forest condition monitoring using Google Earth Engine.

Appendix S2. Spectral-temporal profiles.

Appendix S3. Assessment using 2018 *L. dispar* survey data.

