Merging Landsat time-series and FIA data to develop vulnerability maps for spruce budworm defoliation decision support

Principal Investigator:

Steven A. Sader, Professor and Director of the Maine Image Analysis Laboratory (MIAL), School of Forest Resources, 260 Nutting Hall, University of Maine, Orono, ME 04469-5755, sasader@maine.edu

Co-Principal Investigators:

Kasey Legaard, University of Maine, kasey.legaard@maine.edu Andrew Lister, U.S. Forest Service Northern Research Station, alister@fs.fed.us David MacLean, University of New Brunswick, macleand@unb.ca Erin Simons-Legaard, University of Maine, erin.simons@maine.edu Jeremy Wilson, Harris Center for Conservation Education, wilson@harriscenter.org

Collaborators:

Dave Struble, Maine Forest Service, dave.struble@maine.gov

Completion Date: July 31, 2012

• Methods were developed to predict and map the vulnerability of northern forest stands to spruce budworm defoliation using Landsat satellite imagery and forest attribute data provided by USFS Forest Inventory and Analysis (FIA) field plots

• Vulnerability classes were mapped across a ~10 million acre northern Maine study area, and satellite-derived data were successfully incorporated into an existing spruce budworm decision support system to demonstrate their potential utility for broad-scale spatial forest planning

Funding support for this project was provided by the Northeastern States Research Cooperative (NSRC), a partnership of Northern Forest states (New Hampshire, Vermont, Maine, and New York), in coordination with the USDA Forest Service. http://www.nsrcforest.org

Project Summary:

Periodic infestations of the eastern spruce budworm cause widespread defoliation, growth reduction, and mortality of balsam fir and spruce trees throughout the northeast United States and eastern Canada. Proactive planning for budworm outbreaks requires knowledge of host species distributions and forest vulnerability so that the consequences of infestation can be better understood and effective policies and response strategies put in place. With the next anticipated outbreak approaching, suitable spatial forest vulnerability data are required for planning and decision support across large, multi-owner landscapes.

We have developed methods to map budworm host species distributions and forest vulnerability classes using Landsat satellite imagery and forest attribute data provided by USFS Forest Inventory and Analysis (FIA) field plots. Budworm vulnerability classes were defined based on known relationships with host and non-host species relative abundance and forest age. Host species relative abundance estimates (expressed as a proportion of total live aboveground biomass) were produced using a predictive modeling framework based on advanced machine learning algorithms. Our approach overcomes multiple deficiencies of more established methods, most notably the tendency to over-predict abundance at low values and under-predict abundance at high values. We have utilized the same approach to predict and map forest disturbance from the nearly 40-year archive of Landsat satellite imagery. Forest disturbance data were combined with host species distributions to map budworm vulnerability. Additionally, satellite-derived maps were integrated with aspatial FIA data to develop a stand map schema that can be applied uniformly across multiple ownerships to model the outcomes of alternative management strategies and budworm outbreak scenarios. Satellite-derived proxy stand maps were integrated with forest estate planning software and the Spruce Budworm Decision Support System (MacLean et al. 2001; Hennigar et al. 2011) to evaluate the outcomes of alternative outbreak and response scenarios across a 300,000 acre portion of our study area.

Project outcomes provide the basis for a large-scale, spatially explicit assessment of forest vulnerability ahead of the next budworm outbreak, and demonstrate the utility of low-cost satellite-derived data for spatial forest planning and decision support. Methods developed over the course of this project are widely applicable to other research and management questions, and transferable to other regions.

Background and Justification:

The eastern spruce budworm (*Choristoneura fumiferana* (Clem.)), a native pest to the northeast United States and eastern Canada, has historically infested these regions every 30 to 50 years causing widespread defoliation, growth reduction, and mortality of balsam fir (*Abies balsamea*) and spruce (*Picea spp.*) trees. Mortality and consequent salvage logging during past outbreaks have dramatically affected forest structure, age class distributions, productivity, wood supply, and wildlife habitat distributions. Without long-term forest planning, budworm outbreaks create considerable uncertainty regarding future economic and ecological conditions.

The most recent outbreak occurred during the 1970s-80s and, according to the prevailing theory of budworm population dynamics, the next outbreak is due. Forest conditions have changed considerably since the last outbreak, limiting our ability to infer future impact from past experience. Large tracts of Maine's spruce/fir forest are now younger, more vigorous, and less vulnerable. Yet greater than 2 million acres consist of vulnerable poletimber and sawtimber. Logging and natural disturbance have generally favored recruitment of fir, the most vulnerable host species. On the other hand, inadequate softwood regeneration following the salvage harvest operations of the 1970s-80s resulted in extensive recruitment of non-host hardwood. Furthermore, following the selective removal of spruce and other softwood, many formerly mixedwood stands are now nearly pure hardwood.

Spruce budworm planning in a number of Canadian jurisdictions is facilitated through the use of the Spruce Budworm Decision Support System (SBWDSS), originally developed by the Canadian Forest Service (MacLean et al. 2001). The SBWDSS evaluates marginal timber supply benefits of protecting vulnerable spruce/fir stands under various outbreak scenarios. Users can evaluate the effects of alternative management strategies on forest structure, timber supply, and non-timber objectives. Coupled with forest estate planning software, the SBWDSS facilitates the development of optimal harvest scheduling, salvage, and protection operations to assist long-term management objectives. Use of the SBWDSS requires a classification of stand types by vulnerability. Yet GIS forest stand maps are generally not compatible or available across multiple forest ownerships. A new approach is needed to obtain spatial vulnerability data that are consistent in time, scale, and information content across large areas.

Background and Justification:

The primary goal of this project was to develop and demonstrate a method of mapping vulnerability to budworm defoliation across large, multi-owner landscapes. Stand susceptibility, or probability of budworm defoliation, is strongly related to species composition (balsam fir > white spruce > red spruce > black spruce). Stand vulnerability, or growth reduction and mortality due to defoliation, is strongly related to both species composition and age (mature fir > immature fir > mature spruce > immature spruce, where stands are considered mature if older than approx. 40 years). Susceptibility and vulnerability are reduced when stands contain a significant non-host (hardwood) component. Although many interrelated factors further influence stand susceptibility and vulnerability, evaluation of budworm impact using the SBWDSS principally relies on stand type information regarding host species abundance and age (Table 1).

Landsat satellite imagery, acquired since the early 1970s, has been widely used to map forest cover and monitor forest disturbance in northern forest types. Several studies have successfully mapped tree species or species groups in northern hardwood and mixed forests using multiple satellite images acquired at different times of the year, exploiting differences between species in the timing of leaf-flush or senescence. Extensive reference datasets, typically field measurements, are required to develop reliable predictive models. Leaf-on Landsat imagery collected in different years can also be used to produce highly accurate maps of forest disturbance. The nearly 40-year depth of the Landsat image archive provides a unique opportunity to map the age of immature forest classes as time since standreplacing disturbance.

Table 1: Criteria used to differentiate vulnerability
classes (following Hennigar et al. 2011).

Total Host Species Abundance	Fir/White Spruce Abundance within Host Component	Vulnerability Class*
75 – 100%	75 – 100%	FW
	50 – 75%	FW-RB
	<50%	RB-FW
10 – 75%	50 – 100%	FW-NH
	<50%	RB-NH
<10%		NH

* Vulnerability classes are further differentiated by age, or time since stand-replacing disturbance

FW: balsam fir/white spruce; RB: red spruce/black spruce; NH: non-host

Background and Justification:

The project study area was originally defined by the availability of existing satellite-derived forest disturbance maps, and encompassed ~4 million acres of largely private forestland in northwest Maine (Figure 1). Early success in modeling and mapping host species distributions contributed to several successful grant applications supporting related research and an expansion of mapping objectives over approx. 10 million acres. This expanded study area encompasses the majority of Maine's commercial forestlands, the majority of the northeastern region's spruce/fir resource, and the greater portion of the historical extent of budworm-induced mortality in the Northern Forest.

Mapped forest attributes include the following:

- Primary host species relative abundance (balsam fir + white spruce)
- Total host species relative abundance (balsam fir + white spruce + red spruce + black spruce)
- Stand-replacing forest disturbance, ca. 1973-2010 (and partial canopy disturbance, ca. 1985-2010)
- Budworm vulnerability classes (Table 1)



Figure 1: The proposed 4 million acre study area and the expanded 10 million acre study area, over which species distributions, forest age, and spruce budworm vulnerability have been mapped. Satellitederived maps were further integrated with forest inventory data to derive proxy stand maps used to evaluate budworm outbreak impacts and mitigation efforts using the Spruce Budworm Decision Support System for a 300,000 acre (14 township) area.

Methods:

Predictor data used to map host species distributions:

- 24 Landsat Thematic Mapper images acquired throughout the growing season (late April through mid-October)
- 17 climatological variables (provided the USFS Rocky Mountain Research Station, Moscow Forestry Sciences Laboratory)
- 23 terrain attributes derived from the National Elevation Dataset and National Hydrography Dataset
- Depth to water table (Murphy et al. 2011; provided by the University of Maine Cooperative Forestry Research Unit)

Reference data for model calibration and validation were provided by USFS Forest Inventory and Analysis (FIA) field plots. Approx. 1500 plot locations provided the statistically rigorous and representative sample of forest conditions required for the calibration of reliable forest attribute models. Confidential plot locations were made available through a collaborative agreement with the USFS Northern Research Station FIA Program.

Forest disturbance time series (ca. 1973-2010) were produced using 59 Landsat Multispectral Scanner and Thematic Mapper images, and reference data obtained by the visual interpretation of satellite imagery augmented with forest age and basal area change estimates obtained at FIA plot locations.







Methods:

Regression models predicting host species relative abundance were initially produced using methods previously applied in other regions. Preliminary analyses highlighted multiple difficulties or deficiencies, including an inability to model highly nonlinear relationships, sensitivity to extreme data values, and a strong tendency to produce biased predictions, where abundance was over-estimated at low values and under-estimated at high values. This form of model bias is a common result of mismatches in scale and location between satellite image pixels and field plots.

We have developed an alternative modeling framework based on advanced machine learning algorithms known as support vector machines (SVMs). SVMs provide the following general capabilities:

- Modeling highly nonlinear relationships between predictor and response variables
- Modeling both continuous and categorical data
- Effective use of a very large number of potentially correlated and noisy predictor variables
- Excellent generalization despite a limited and/or potentially error-prone reference dataset
- Relative insensitivity to extreme data values

Use of SVMs, however, requires the specification of multiple parameters whose values strongly affect model performance. Parameter values can range over multiple orders of magnitude and optimal values are difficult to identify. Additionally, attainment of suitable results may require selection of an optimal subset of predictors. Further improvement may result from the selection an optimal subset of reference data samples for model calibration. Selection of SVM parameter values, predictor variables, and calibration samples provides tremendous flexibility in fitting models to meet various performance criteria. We exploit this flexibility to produce models with minimal total prediction error *and* minimal bias. To do so, we use a multi-objective genetic algorithm (GA) for model calibration.

Methods:

Genetic algorithms mimic evolution by natural selection to efficiently explore a very large range of potential solutions to an optimization problem. Our GA implementation begins by encoding the model calibration problem into a binary string analogous to a chromosome:



An initial set of chromosomes is randomly generated, resulting in an initial population of solutions. Each chromosome is associated with a predictive model calibrated using a particular set of parameter values, predictor variables, and reference samples. The population of models is allowed to evolve over many generations, subject to binary operations that mimic genetic recombination and mutation:



Model performance characteristics determine the relative "fitness" of individuals within the population, and only the fittest individuals are allowed to reproduce at each generation. Fitness metrics include a measure of overall prediction error and a measure of prediction bias. The population gradually converges to a set of models that simultaneously minimize total prediction error and bias.



Figure 2: Predictions of primary budworm host species relative abundance (balsam fir + white spruce; crossvalidated $R^2 = 69.8\%$, mean absolute error = 8.5%). The 1:1 line is shown for reference. Due to differences in Landsat image availability, the project study area was split into two overlapping sections prior to model calibration. Results are presented for the larger of the two sections; very similar results were obtained for the second.

Maps derived from calibrated models were thereafter stitched together to form a seamless mosaic.





Figure 2: Predictions of primary budworm host species relative abundance (balsam fir, white spruce; crossvalidated $R^2 = 69.8\%$, mean absolute error = 8.5%). The 1:1 line is shown for reference. Due to differences in Landsat image availability, the project study area was split into two overlapping sections prior to model calibration. Results are presented for the larger of the two sections; similar results were obtained for the second.



Figure 3: Predictions of total host species relative abundance (balsam fir, white, red, black spruce; crossvalidated $R^2 = 74.8\%$, mean absolute error = 11.9%). The 1:1 line is shown for reference.

To further differentiate vulnerability classes by stand age or maturity (Table 1), we modeled and mapped the occurrence of stand-replacing disturbance using the nearly 40-year Landsat image archive. Whereas Landsat imagery acquired from ca. 1985 to present by the Thematic Mapper instrument is capable of resolving even very light canopy disturbance, earlier imagery acquired by the Multispectral Scanner is generally capable of resolving only heavy disturbance. Using imagery acquired every 3-4 years from 1973 to ca. 1985, we predicted the occurrence of stand-replacing disturbance using the same framework developed for modeling host species abundance. Reference data were obtained through a combination of age estimates provided by FIA plots and visual interpretation of Landsat imagery over FIA plot locations. Fitness metrics for model calibration included measures of prediction accuracy and bias.

Cross-validated accuracy of stand-replacing disturbance maps, 1973-1985:

User's accuracy:	85.6 - 91.1%	(class commission error 8.9 - 14.4%)
Producer's accuracy:	78.0 - 83.5%	(class omission error 16.5 - 22.0%)
Overall map accuracy:	99.4%	

Note that estimated omission error rates for the disturbance class exceed commission error rates. Our modeling framework results in the simultaneous calibration of a set of models with differing error distributions. Some show a nearly perfect balance between omission and commission error. However, models with a small bias toward under-estimation of disturbance extent generally produced maps with superior spatial patterns of mapped disturbance. The end effect may be an over-estimation of mature vulnerability classes relative to immature classes.

Imagery acquired by the Landsat Thematic Mapper instrument ca. 1985 to present is capable of resolving partial canopy disturbance. Using imagery acquired every 2-3 years, we predicted the occurrence of canopy disturbance ca. 1985-2010. Reference data were obtained through visual interpretation of TM imagery and contemporary aerial photography.

Cross-validated accuracy of disturbance maps, 1985-2010:

 User's accuracy:
 90.5 - 91.8%

 Producer's accuracy:
 85.0 - 87.0%

 Overall map accuracy:
 98.6%

(commission error 8.2 - 9.5%) (omission error 13.0 - 15.0%)

Disturbance intensity was subsequently modeled and mapped as a continuous attribute (% of basal area removed) over areas predicted as disturbed. Reference data were obtained from FIA plots identified as disturbed between measurement cycles (1999-2008). Disturbance intensity models were calibrated using the



same approach developed for species relative abundance, resulting in a nearly unbiased predictions, although rapid regrowth following disturbance introduces considerable imprecision. A basal area removal threshold of 70% was used to differentiate stand-replacing from partial canopy disturbance. This threshold value was selected to maintain consistency with related research activities.

Figure 4: Predictions of disturbance intensity (% of basal area removed) at FIA plot locations identified as disturbed between measurement cycles (cross-validated $R^2 = 57.5\%$, mean absolute error = 13.8%). The 1:1 line is shown for reference.



Figure 5: A small sample area extracted from the full 10 million acre map of budworm vulnerability classes. The vulnerability map was produced by intersecting maps of primary host species relative abundance, total host species relative abundance, and standreplacing disturbance according to the classification presented in Table 1. (FW: balsam fir/white spruce; RB: red spruce/black spruce; NH: non-host)

Figure 6: A comparison of the overall extent of mapped budworm vulnerability classes (% of mapped forestland) and the proportion of FIA plots within the project study area that meet the same vulnerability class criteria according to estimates of host species abundance and stand age obtained at plot locations. Results compare reasonably well, and differences may be attributable to uncertainties in either the models used to map vulnerability classes or in FIA estimates, particularly of stand age. Note that FIA data are aspatial, whereas mapped data provide a spatially explicit depiction of forest conditions with a pixel resolution of ~1/4 ac and minimum patch size of ~2 ac.



We have successfully integrated satellite-derived data products with aspatial FIA data to produce a stand map schema that can be applied uniformly across multiple ownerships to model the outcomes of alternative management and outbreak scenarios. Spruce budworm outbreak and response scenarios have been implemented on a 14 township subset of the study area using a version of the SBWDSS specifically designed to work in conjunction with the Remsoft Spatial Planning System (FORUS 2010). Scenarios include a base run without an outbreak, a moderate outbreak with and without spray protection, and a severe outbreak with and without protection.



Figure 7: Impact of alternative outbreak scenarios on spruce-fir standing volume within a 14 township (~300,000 ac) subset of the project study area. Scenarios executed over a period of 100 years.



Period 1 Marten habitat Non-habitat Habitat

Severe outbreak scenario, period 10



Figure 8: Future habitat distributions for American marten (Martes *americana*) under base and severe budworm outbreak scenarios (no spray protection), implemented over a 14 township area.

Implications and applications in the Northern Forest region:

Decision support systems will be needed to plan for and manage the impact of the next spruce budworm outbreak, yet the necessary spatially explicit vulnerability data have been lacking for large, multi-owner landscapes. The principal outcomes of this project are 1) methods for producing nearly unbiased models and maps of forest attributes over large areas, 2) maps of budworm host species relative abundance and budworm vulnerability classes across the majority of the commercial forestlands of northern Maine, and 3) development and demonstration of a methodology by which budworm vulnerability and impact assessments can be applied across ownerships using low-cost geospatial data.

Forest management context has changed considerably since the last outbreak. Factors such as improved road networks and developing markets will presumably provide a greater range of management options than were previously available, but clearcut regulations, social pressures, and ecological realities may have the opposite effect. For example, past forest protection measures were rarely directed at non-timber objectives, but that may change due to the growing recognition of the ecological importance of mature softwood stands. Moreover, Maine's formerly industrial forest is now owned by a diverse set of entities including investment firms, logging contractors, developers, high net-worth individuals, and conservation groups. How these landowners collectively respond to the next budworm outbreak will in large part determine the future condition of the spruce/fir forest. Project outcomes provide a basis for a broad-scale, spatially-explicit assessment of vulnerability ahead of the next outbreak, and could be used to better anticipate management conflicts, tradeoffs, and opportunities.

Our methods are generally transferable to other northern forest regions and applicable to future studies exploring budworm impact mitigation over large landscapes. The linkage of forest inventory data and satellite imagery for the purpose of mapping forest attributes is widely applicable to other research and management questions. Our modeling framework addresses some of the primary difficulties in effectively utilizing existing forest inventory data for this purpose. The integration of satellite-derived maps with existing forest planning and decision support software holds significant potential for large landscapes where traditional stand maps are incomplete, incompatible, or unavailable.

Future Directions:

The sustainable management of forest resources requires a clear understanding of the cumulative effects of natural and anthropogenic disturbance over broad spatial scales and long time horizons. There remains considerable uncertainty regarding future conditions within the Northern Forest due in large part to recent changes in forest management practices, harvest levels, vulnerability to natural disturbance, and climate. Landscape simulation has become an essential tool for understanding the long-term effects of ecological processes and land-use activities across large areas. Spatially explicit simulations of disturbance and forest succession provide information about future forest conditions critical to evaluating interactions between resource management and ecosystem process.

Under funding obtained through NSRC and the NASA/USDA National Institute of Food and Agriculture (NIFA) Carbon Cycle Science program, we are currently applying the LANDIS-II (LANDscape Disturbance and Succession) simulation model over our entire 10 million acre northern Maine study area. LANDIS-II simulates disturbance and succession of cohorts principally defined by age and species. In order to produce suitably detailed spatial data, we have modeled and mapped the distributions of 13 economically and ecologically important tree species and integrated these results with forest disturbance products. Alternative forest management, budworm outbreak, and climate change scenarios will be executed to investigate impacts on wood supply, carbon stocks, and wildlife habitat. Our overarching goal is to provide a better understanding of the interacting effects of dominant drivers of forest landscape change on the long-term sustainability of important forest resources.

- FORUS 2010. Protection Planning System Extension for RSPS. Developed in partnership with the University of New Brunswick and Forest Protection Limited.
- Hennigar, C.R., Wilson, J.S., MacLean, D.A., and Wagner, R.G. 2011. Applying a spruce budworm decision support system to Maine: Projecting spruce-fir volume impacts under alternative management and outbreak scenarios. Journal of Forestry. 109: 332-342.
- MacLean, D.A., Erdle, T.A., MacKinnon, W.E., Porter, K.B., Beaton, K.P., Cormier, G., Morehouse, S., and Budd, M. 2001. The Spruce Budworm Decision Support System: forest protection planning to sustain long-term wood supply. Canadian Journal of Forest Research. 31: 1742-1757.
- Murphy, P.N.C., Ogilvie, J., Meng, F.-R., White, B., Bhatti, J.S., Arp, P.A. 2011. Modelling and mapping topographic variations in forest soils at high resolution: A case study. Ecological Modelling. 222: 2314-2332.

Products:

Manuscripts in preparation on the following topics:

- Forest disturbance mapping by multi-objective optimization of support vector machines (anticipated submission 12/2013)
- Modeling and mapping tree species relative abundance by multi-objective optimization of support vector machines (anticipated submission 4/2014)
- Additional planned manuscripts on the following topics:
 - A spatial assessment of vulnerability to spruce budworm defoliation across the commercial forestlands of northern Maine
 - Semi-automated cloud/shadow masking for Landsat Thematic Mapper and Multispectral Scanner imagery

Presentations:

- Legaard, K.R., S. Sader, J. Wilson, E. Simons-Legaard, and A. Weiskittel. 2013. A spatial assessment of vulnerability to defoliation by spruce budworm across the commercial forestland of northern Maine. Presented at the New England Society of American Foresters Spring Meeting, Bethel, Maine.
- Legaard, K.R., S. Sader, J. Wilson, E. Simons-Legaard, and A. Weiskittel. 2012. Mapping vulnerability to defoliation by spruce budworm using Landsat satellite imagery and FIA field plots. Poster presented at the Eastern CANUSA Forest Science Conference, Durham, New Hampshire.
- Simons-Legaard, E., K. Legaard, J. Wilson, A. Weiskittel, and S. Sader. 2012. Long-term outcomes and tradeoffs of forest policy and management practices on the borad-scale sustainability of forest resources: wood supply, carbon, and wildlife habitat. Poster presented at the Eastern CANUSA Forest Science Conference, Durham, New Hampshire.
- Wilson, J. S., E. M. Simons, K. R. Legaard, S. A. Sader, and J. Leahy. Spatial forest planning for meeting multiple natural resource goals. 2010. Poster presented at Maine EPSCoR State Conference, University of Maine, Orono, Maine.
- Wilson, J. S., E. M. Simons, K. R. Legaard, S. A. Sader, and J. Leahy. 2010. Spatial forest planning to meet multiple natural resource goals. Presented at the Eastern CANUSA Forest Science Conference, Edmundston, New Brunswick.

Products:

Workshops:

• Wilson, J. S. Spruce budworm decision support: Application and pitfalls. Cooperative Forestry Research Unit Fall Field Tour and Forum. Caribou, Maine. October 29, 2009

Partial support for one M.S. and one Ph.D. student:

- Looze, B.E. Forest fragmentation patterns in Maine watersheds and prediction of visible crown diameter in undisturbed forest. M.S. thesis. University of Maine, Orono, ME. (completed 2012)
- Legaard, K.R. A multidecadal analysis of forest disturbance and landscape change in the commercial forestlands of northern Maine. Ph.D. dissertation. University of Maine, Orono, ME. (anticipated completion 12/2013)

Grants resulting in part from the success of this project:

- USDA Forest Service Northeastern States Research Cooperative, Theme 3, 2010: Evaluating the interacting effects of forest management practices and periodic spruce budworm infestation on broad-scale, long-term forest productivity. Lead PI K. Legaard.
- USDA Forest Service Northeastern States Research Cooperative, Theme 1, 2010: Long-term outcomes and tradeoffs of forest policy and management practices on the broad-scale sustainability of forest resources: wood supply, carbon, and wildlife habitat. Lead PI E. Simons-Legaard.
- NASA/USDA National Institute of Food and Agriculture (NIFA) Carbon Cycle Science, 2010: Carbon dynamics and forest management: A retrospective analysis and projection of the potential effects of land use, climate change, and natural disturbance in northeastern forests. Lead PI A. Weiskittel.
- University of Maine's Sustainability Solutions Initiative, 2009: Spatial forest planning to meet multiple natural resource goals: Developing geospatial tools to forecast management outcomes across a diverse landscape of ownership types and stakeholder interests. Lead PI J. Wilson.